VSV Vectors as Vaccines for Emerging Viruses and as Probes for Entry Pathways

Permanent link

Terms of Use

Share Your Story

Accessibility

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Division of Medical Sciences

Committee on Virology

have examined a dissertation entitled

VSV vectors as vaccines for emerging viruses and as probes for entry pathways

presented by Joseph Timpona

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature: _____

Typed Name: Dr. Max Nibert

Signature: _____
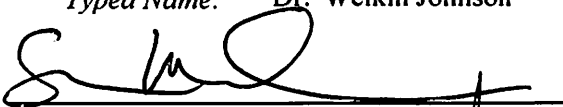
Typed Name: Dr. Lee Gehrke

Signature: _____
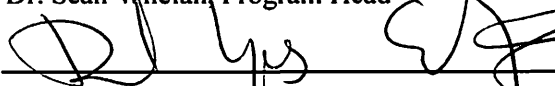
Typed Name: Dr. Todd Allen

Signature: _____

Typed Name: Dr. Welkin Johnson

Dr. Sean Whelan, Program Head

Date: April 06, 2018

Dr. David Lopes Cardozo, Director of Graduate Studies

VSV vectors as vaccines for emerging viruses and as probes for entry pathways

A dissertation presented

by

Joseph Timpona

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Virology

Harvard University

Cambridge, Massachusetts

April 2018

VSV vectors as vaccines for emerging viruses and as probes for entry pathways

Abstract

Viral fusion proteins (VFPs) are membrane-anchored machines that coat the surfaces of enveloped viruses and induce the fusion of virus and cell membranes. While enveloped viruses are evolutionarily diverse, their fusion proteins are organized into three classes based on structural and functional properties. Our work describes the evolutionary and functional characterization of a class-I VFP, EnvPb1, which is derived from an endogenous retrovirus and conserved in primate genomes. Additionally, we investigated vesicular stomatitis virus (VSV) vectors as a means to develop vaccines targeting the class-II VFP of the newly emergent Zika virus (ZIKV).

Since VFPs are surface exposed and mediate viral entry, antibodies that bind and neutralize VFPs are potent inhibitors of infection. Following the Zika virus outbreak in South America, we sought to generate vaccine candidates that would induce a neutralizing antibody response against the ZIKV envelope protein, E. We generated and characterized a panel of recombinant (r)VSV–based vaccines against ZIKV. We show incorporation of functional flavivirus envelope protein into VSV particles. While ZIKV-E incorporation into VSV did not yield independently propagating virus, we discovered that rVSV-ZIKV infected cells readily produce ZIKV virus like particles (VLPs) containing ZIKV-E. We immunized mice with the VLP-producing rVSV-ZIKV or purified VLPs alone. We determined that both the rVSV and the VLPs induce a neutralizing antibody response against ZIKV.

Not only are VFPs targeted by host defenses, but also, in rare cases, they have been captured and repurposed to provide a beneficial function to that host. We developed assays and tools to assess the evolutionary conservation and the function of the endogenous retrovirus envelope protein EnvPb1. Our work revealed that EnvPb1 was captured in a common ancestor to all simians and that evolutionarily distant primate EnvPb1 sequences retain the ability to fuse cells.   Additionally, we detect EnvPb1 protein in placenta tissue – indicating that it may serve a physiological role in this tissue. Moreover, we show that EnvPb1 is released from cells in culture, suggesting it can be packaged into extracellular vesicles. Our work expands upon what is known about EnvPb1 and provides insights and tools to serve future studies.

Table of Contents

List of Figures and Tables

For my family and friends.

Your caring and support helped me get through the Ph.D. with my sanity mostly intact.

For those engaged in scientific pursuits: may the force be with you.

Chapter 1: Introduction

Vesicular stomatitis virus

      Vesicular stomatitis virus (VSV) is the archetypal member of the virus family *Rhabdoviridae* and encodes a single-stranded, negative-sense RNA genome. Its roughly 11kb genome encodes 5 protein products – nucleocapsid (N), phosphoprotein (P), matrix (M), glycoprotein (G), and the large RNA-dependent RNA polymerase, L – bound by terminal leader (Le) and trailer (Tr) sequences (Figure 1.1A). The nucleocapsid protein binds viral RNA forming a nucleocapsid RNA (N-RNA) complex. N-RNA serves as the active template for transcription and replication [1]. The phosphoprotein serves as a cofactor for the VSV polymerase and serves to link N and L to form the ribonucleoprotein (RNP) complex for transcription and replication [2, 3]. The viral polymerase, L, uses N-RNA as a template to transcribe and replicate the VSV genome and additionally serves to cap and polyadenylate viral message [4-6]. The matrix protein is important for assembly and budding of particles as well as for shutoff of host translation by blocking nuclear export of mRNA [7-10]. The primary function of the glycoprotein is to mediate attachment to target cells and initiate fusion of virus and cellular membranes [11]. VSV virions take on a bullet shape, with glycoprotein spikes on the surface (Figure 1.1B) [12]. The inner side of the cell-derived lipid bilayer is coated by matrix protein, which surrounds the ribonucleoprotein (RNP) complex consisting of N, P, L, and viral genome.
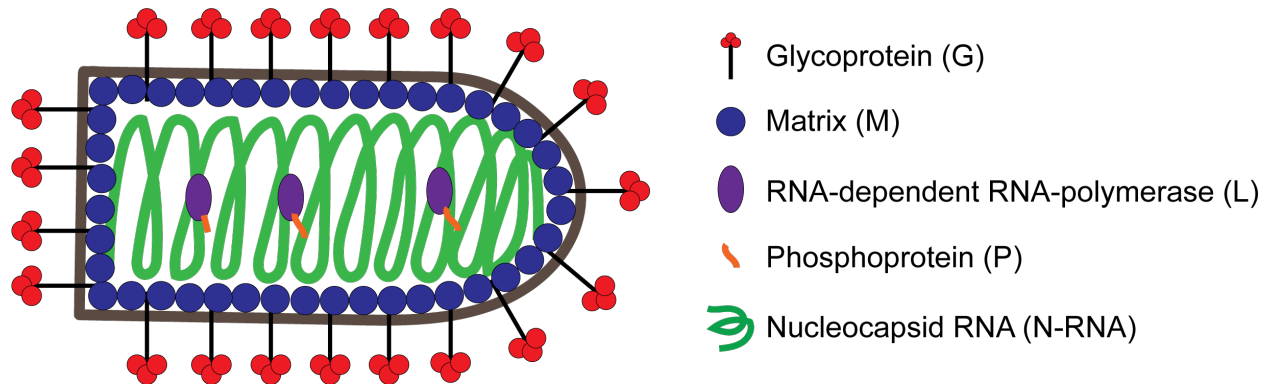
Figure 1.1: VSV genome and particle composition. (A) VSV genome drawn to scale. The VSV genome is a single-stranded negative-sense RNA. VSV genes are shown in color and the 3' leader (Le) and 5' trailer (Tr) are shown as white boxes. (B) VSV particle composition. VSV particles are encased in a lipid bilayer (shown in brown). RNP complexes (N-RNA, P, and L) are condensed by matrix protein to form a bullet shape. Glycoprotein trimers coat the surface of virions and mediate attachment and fusion of viral particles.

VSV replication

The replication cycle of VSV begins when a virion attaches to a target cell and is engulfed by clathrin-mediated endocytosis (Figure 1.2A) [13]. Endosomes acidify as they mature, which leads to low-pH triggering of VSV-G-mediated membrane fusion [14, 15]. This is followed by uncoating (dissociation of M and RNP) and release of the viral RNP into the cytoplasm (Figure 1.2B-D) [16]. Next, L begins to synthesize subgenomic mRNAs corresponding to the 5 viral genes. Due to stuttering and inefficient re-initiation of L between genes, viral genes are transcribed in a cascade in which N mRNA is most abundant and L mRNA is least abundant [17]. As viral genes are translated, they are transported to virus assembly sites at the plasma membrane through distinct pathways [18]. While L initially serves to synthesize viral mRNA, later in infection it switches to replicate the viral genome. N encapsidates nascent genomes, and RNP complexes shuttle to G-containing microdomains at the plasma membrane [19]. Following this, M is recruited to these assembly sites where it condenses RNPs into the characteristic bullet shape, forming a nucleocapsid-matrix protein (NCM) complex [20]. NCMs bud through G microdomains and fission releases virions from the plasma membrane.
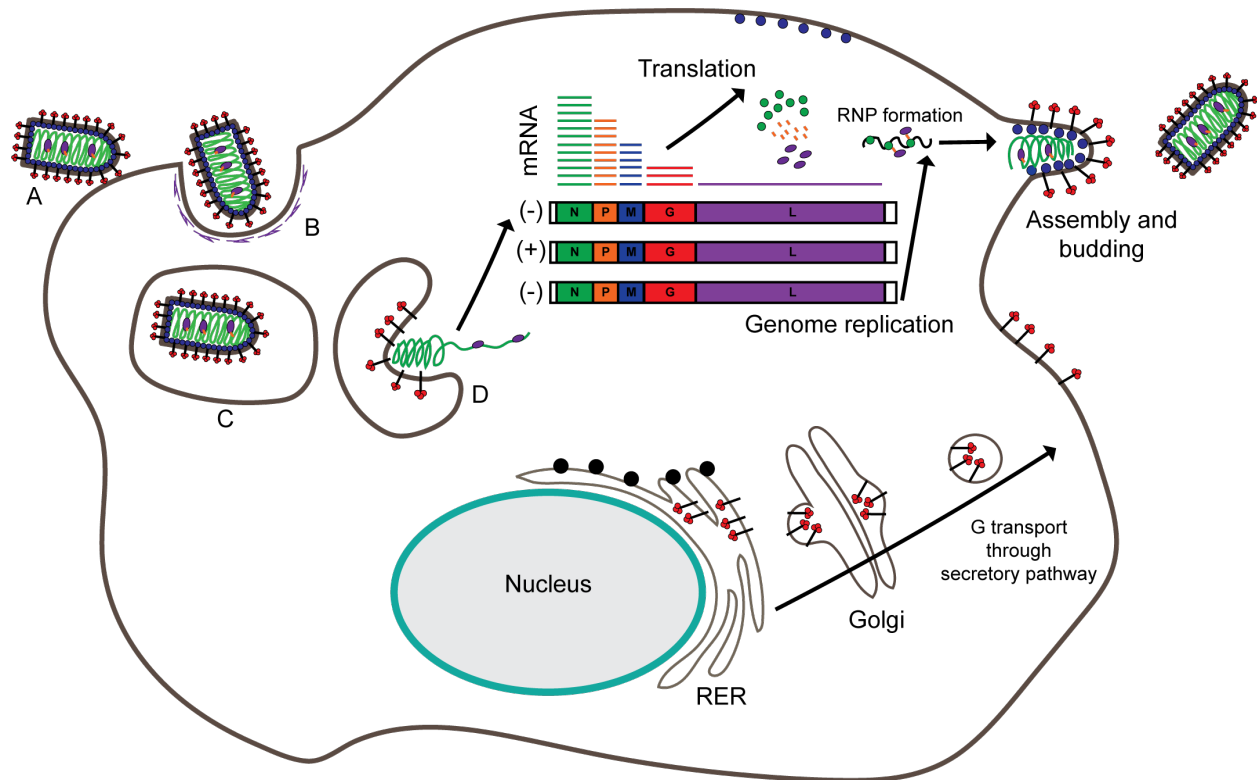
Figure 1.2: Schematic of VSV replication. VSV particles attach to the plasma membrane through G-mediated interactions (A). Clathrin-mediated endocytosis engulfs particles (B) and they traffic to late endosomes (C) where the acidic environment triggers VSV-G fusion of viral and cellular membranes. Uncoating of virus (D) allows for the start of gene transcription by L. Viral proteins are translated, with VSV-G translation occurring in the rough endoplasmic reticulum (RER). VSV-G travels through the secretory pathway en route to the plasma membrane. VSV-M also localizes to the plasma membrane following translation. As L replicates nascent genomes, they are bound by VSV-N, and associate with P and L to form RNPs. RNPs are condensed by M, and budding occurs through VSV-G containing lipid domains. VSV genes and proteins are color-coded as in Figure 1.1. Ribosomes are depicted as black circles on RER and clathrin is depicted as purple shapes around endosomes in B.

Vesicular stomatitis virus as a vector

VSV has long been the archetype mononegavirus, and studies of VSV have provided seminal contributions to the field of mononegavirales including studies of the viral polymerase structure, viral entry, and viral RNA synthesis [21-23]. The development of VSV reverse genetic tools provided a platform upon which researchers could replace VSV genes with their counterparts from other viruses [24, 25]. By generating recombinant VSV (rVSV) in which VSV-G is replaced by a heterologous viral fusion protein, researchers could now study the entry of high biosafety level viruses such as Ebola at the BSL2 level.

rVSVs have yielded significant contributions to virology. For example, rVSV expressing the Ebola virus glycoprotein (GP) was utilized in a haploid genetic screen to identify NPC1 as the intracellular receptor for Ebola virus [26]. This method has yielded similar insights into the entry of other viruses including Rabies virus, Andes virus, Lujo virus, and even an endogenous retrovirus, HERV-K [27-31]. Given the widespread utility of rVSV, rVSVs expressing glycoproteins from a wide range of viruses such as influenza virus and chikungunya virus have been tested as potential vaccine candidates [32, 33]. VSV vectors are promising vaccine candidates due to the ease of making rVSV, the ability of VSV to replicate in essentially all mammalian cells, and the fact that VSV infection is only mildly symptomatic in humans [34-36].

VSV-based vaccines

The most successful rVSV vaccine is rVSV-ZEBOV. Originally developed in 2004 by replacing the coding region of VSV-G with Zaire Ebola virus (ZEBOV) GP, a

single immunization with rVSV-ZEBOV was shown to protect cynomolgus macaques from a lethal challenge with ZEBOV [37, 38]. rVSV-ZEBOV was further characterized in non-human primates; however, it did not progress into humans until 2014 when the unprecedented Ebola virus outbreak in West Africa necessitated rVSV-ZEBOV human trials in efforts to stop the outbreak. Phase I clinical trials in healthy U.S. adults showed that a single dose of 20 million plaque forming units (PFU) of rVSV-ZEBOV induces a ZEBOV-specific antibody response [39]. All recipients experienced viremia at the first follow-up visit, but only 20% still had viremia at day 7. Most importantly, no significant adverse effects were reported, and viremia only correlated with fever. A second dose of rVSV-ZEBOV led to increased short term but not long term antibody titers. Another phase I trial found similar results although it used a higher dose. 11 of 51 patients in Geneva developed short-term arthritis in one of their joints, which led to suspension of that trial [40].

Next, rVSV-ZEBOV was tested in Guinea and Sierra Leone during the Ebola outbreak. A ring-vaccination strategy was used by which contacts and contacts of contacts of a person with confirmed Ebola virus disease were traced and vaccinated with 20 million PFU of rVSV-ZEBOV. In total 5837 individuals received the vaccine and none of them developed Ebola virus disease 10 days or more after vaccination [41]. Despite being administered to almost 6000 individuals, only two serious adverse events related to vaccination were reported, and most adverse events were mild and included headache, injection site pain, and fatigue. Furthermore, the speed with which rVSV-ZEBOV was tested in humans and the ease of generating rVSV for other viruses exemplifies rVSV as an effective tool for generating vaccines against emerging

pathogens. Thus, in chapter two, we aimed to apply this platform to meet the urgent need for Zika virus vaccines.

Viral Fusion Proteins

Enveloped viruses are encapsulated by a lipid bilayer that is acquired during assembly and release of a virus from its host cell. To gain entry into a target cell, enveloped viruses must fuse the viral and target cell membranes. Viral fusion proteins (VFPs) catalyze the fusion of virus and cell membranes to overcome the high kinetic energy barrier to fusion of lipid bilayers [42, 43]. On a virion, VFPs exist in a pre-fusion state (Figure 1.3). Upon localization of the virus to a cell, triggering of conformational changes (often induced by receptor binding or protonation) leads to the formation of an extended intermediate in which a hydrophobic region of the VFP interacts with the target cell membrane. Collapse of the extended intermediate by a fold-back mechanism leads to a trimer of hairpins, which brings the two membranes together. Despite evolutionary differences between enveloped viruses, all VFPs studied follow this same basic mechanism to achieve fusion of virus and cell membranes.
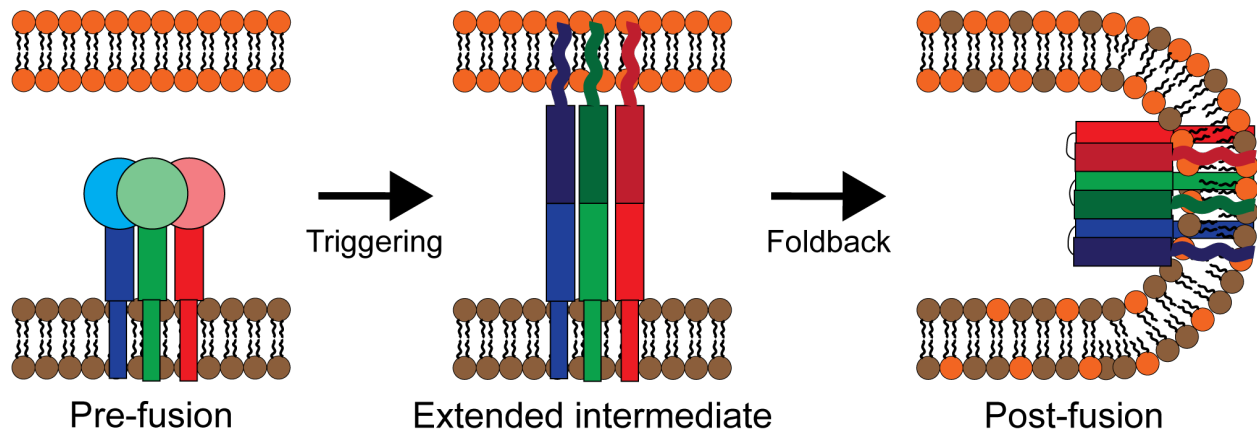
Figure 1.3: Simplified schematic of viral membrane fusion. VFPs reside on the outer surface of enveloped viruses in a pre-fusion state. Upon triggering, VFPs undergo conformational changes upon which a hydrophobic stretch interacts with the target membrane through an extended intermediate. Refolding brings the inserted segment adjacent to the transmembrane anchor, forming a hairpin structure that facilitates mixing of virus and target membrane lipids and fusion pore formation. Model is based off of class I VFP structures.

Class I VFPs

VFPs are classified into three known classes (I-III) based on structural similarities. Class I VFPs are synthesized as a polyprotein precursor. En route to the cell surface, class I VFPs undergo proteolytic cleavage by a host protease in the secretory pathway. Proteolytic processing yields an N-terminal subunit, often containing a receptor-binding domain, and a C-terminal membrane-anchored subunit that harbors the fusion machinery. Class I VFPs form trimeric oligomers in both their pre-fusion and post-fusion states.

Influenza virus HA is the canonical class I VFP. Influenza HA is the precursor to the N-terminal HA1 and the C-terminal HA2. HA1 forms a globular structure containing a binding pocket for attachment of Influenza virions to sialic acid on the surface of cells (Figure 1.4A) [44, 45]. Following attachment, virions are endocytosed, where a proton-induced conformational rearrangement ejects the fusion peptide at the N-terminus of HA2 into the target membrane (Figure 1.4B) [46, 47]. After insertion of the fusion peptide, the HA trimer folds back to form a trimer of hairpins, bringing the fusion peptide, the transmembrane-spanning region of HA2, and the attached lipid bilayers together [48]. The trimer of hairpins is inherently stable and permits the formation of the fusion pore, thus releasing the virus contents into the cell [49]. Approximately three HA trimers are needed to form a fusion pore [50].
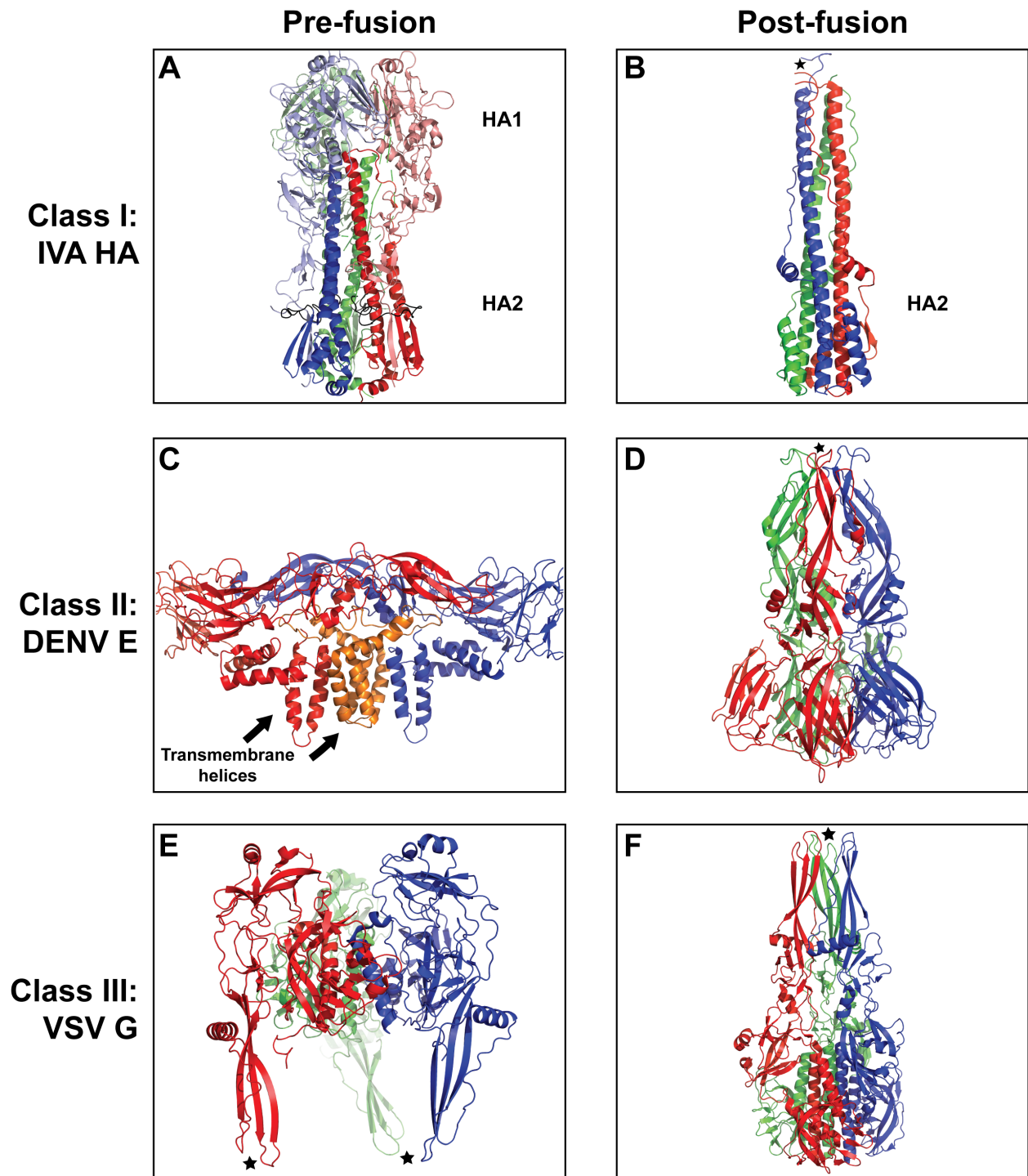
Figure 1.4: Pre- and post-fusion structures of VFPs. Individual subunits are colored in red, blue, and green. Fusion peptides or loops are denoted with a black star. Protein structures were made in PyMol 2.1 [51] (A) Pre-fusion structure of influenza A virus (IVA) HA trimer (protein databank (PDB) entry 2YPG).

Figure 1.4 (continued):

The globular HA1 is lightly colored compared to HA2. (B) Post-fusion structure of IVA

HA2 trimer (PDB: 1QU1). (C) Pre-fusion structure of dengue virus (DENV) envelope (E)

dimer (PDB: 3J27). E monomers are shown in red and blue while membrane (M) is

shown in orange. Arrows denote membrane-spanning helices. (D) Post-fusion

structure of DENV E trimer (PDB: 1OK8). (E) Pre-fusion structure of VSV G trimer

(PDB: 5I2S). Fusion loops are located at the bottom of the structure. (F) Post-fusion

structure of VSV G trimer (PDB: 2CMZ).

While class I VFPs share structural similarities, they differ in the triggers that lead

to the extended intermediate. In the case of HIV, the N-terminal GP120 binds to its

receptor CD4, which induces structural changes in GP120 that expose a co-receptor

binding site [52]. Following binding of GP120 to the co-receptor (CCR5 or CXCR4),

GP120 dissociates and GP41 inserts into the target membrane catalyzing bilayer fusion

[53]. Moreover, the avian sarcoma leucosis virus (ASLV) envelope protein induces

fusion through a two-step mechanism by which receptor binding induces conformational

changes that then permit low pH to drive extension of its transmembrane subunit into a

target membrane [54, 55].

Class II VFPs

Flavivirus envelope (E) proteins best represent class II VFPs. Instead of

proteolytic processing providing the priming step for fusion, class II VFPs have an

associated chaperone protein (premembrane (prM) for flaviviruses), which prevents

triggering of fusion as the virus exits the cell. When prM and E are associated, prM

covers the fusion loop of E [56]. As flavivirus particles traffic through the cell, prM

undergoes proteolytic cleavage by furin in the Golgi to yield pr+M [57]. When the

particle emerges from the cell, pr dissociates from M, which primes E for fusion.

While class II VFPs initiate virus and cell membrane fusion by a similar

mechanism to class I VFPs, they are structurally different. In contrast to the "spike"-like

trimers that are visible projecting out of the surface of class I VFP-containing viruses,

class II VFPs form homodimers that run antiparallel to each other and form a smooth,

ordered surface around a virion (Figure 1.4C) [58]. Upon receptor binding, class II VFP

dimers dissociate and reorganize as trimers (Figure 1.4D) [59]. These trimers catalyze

the fold-back mechanism reaction similar to class I VFPs.

Class III VFPs

Class III VFPs are found on viruses as distinct as rhabdoviruses, herpesviruses,

and insect baculoviruses. Class III VFPs are unique from the other two classes in that

they do not require proteolytic cleavage of the fusion protein or an associated

chaperone to prime fusion. In the pre-fusion structure of VSV-G, the hydrophobic loops

are hidden near the viral membrane (Figure 1.4E) [60]. Following exposure to low pH,

VSV-G rotates 180º around its core, thereby moving the fusion loops to interact with the

target membrane to catalyze membrane fusion through a predicted extended

intermediate (Figure 1.4F) [61]. In the absence of a priming step, the transition between

pre- and post-fusion forms is reversible [62].

Recombinant VSV bearing VFPs

Recombinant VSV (rVSV) incorporate VFPs from all three VFP classes despite the structural differences and priming mechanisms that differentiate VFPs of the three classes [33, 63, 64]. These viruses are capable of propagating in cells while solely relying on the foreign glycoprotein for assembly, budding, and subsequent infection. As such, VSVs bearing foreign glycoproteins are beneficial tools to study entry pathways and to develop vaccines. In chapter two, we utilize VSV vectors to express the class II Zika virus envelope protein and evaluate them as vaccine candidates. In Chapter 3, we establish VSV vectors to investigate the function of a class I retroviral envelope protein, EnvPb1, that has been maintained in simian genomes.

Zika virus emergence and pathogenesis

The first known cases of Zika virus infection of a human occurred in Uganda and Tanzania in 1952 [65]. For the rest of the 20th century, infections occurred sporadically and were limited to Africa and Asia - in total, there were only 14 documented cases of Zika virus disease [66-68]. The first major outbreak of Zika virus occurred in 2007 on Yap Island in The Federated States of Micronesia. Approximately three quarters of the 7,391 Yap residents were thought to be infected with Zika virus and more than 900 sought medical attention with symptoms attributable to Zika virus infection [69]. However, Zika virus infection led to no deaths or hospitalizations during this outbreak.

The next Zika virus outbreak grew in size with a suspected 29,000 cases in French Polynesia between 2013 and 2014 [70]. Importantly, this outbreak provided the first evidence for Zika virus-associated Guillain-Barré syndrome [71]. Nevertheless, it

was not until 2015 when Zika virus emerged in the Western Hemisphere that it was realized to be a significant human pathogen. As of January 4, 2018, 583,451 suspected cases of Zika virus were reported in the Americas, which is likely a major underestimate [72]. Consequently, it was during this outbreak that Zika virus was discovered to be a causative agent of microcephaly [73]. Furthermore, while Guillain-Barré syndrome was first suspected to be associated with Zika virus disease during the outbreak in French Polynesia, the outbreak in the Americas substantiated the association between Guillain-Barré syndrome and Zika virus disease [74-76]. Given the magnitude of the outbreak and the newly discovered neurological pathologies associated with Zika virus disease, the World Health Organization (WHO) proclaimed that a Zika virus vaccine is urgently needed.

Zika virus composition and assembly

Zika virus (ZIKV) belongs to the virus family *Flaviviridae,* which comprises other relevant human pathogens such as Dengue virus, Yellow fever virus, and West Nile virus. As a flavivirus, ZIKV encodes a single-stranded positive sense RNA genome that codes for three structural and seven non-structural proteins. Upon release of the viral genome into a cell following envelope (E) protein-mediated fusion of the virus and cell membranes, the viral genome is recognized as mRNA and directly translated to yield a single polyprotein [77]. After protein synthesis, host and viral proteases cleave the polyprotein to yield the individual viral proteins [78]. Following replication of the viral genome, particles bud into the endoplasmic reticulum (ER). These virions contain 180 total copies of both the premembrane (prM) and envelope (E) protein anchored to the

15

surface lipid membrane through their transmembrane regions. While in the ER, prM

and E are arranged as 60 trimers of prM:E heterodimers [79]. Due to the conformation

of E on the surface, these particles appear spiky, and are referred to as immature

particles because the association of prM with the fusion peptide of E prohibits fusion.

Immature particles traffic through the trans-Golgi network where the mildly acidic

conditions lead to structural rearrangements causing E to lay flat as antiparallel dimers

[58].   Following this, the host protease furin cleaves prM into a pr peptide and the C-

terminal membrane-bound M. Following these conformational changes, the now

"mature" particles consist of 90 dimers of M:E heterodimers [80, 81]. Mature particles

are released from a cell through exocytosis where they can then bind another cell to

restart the process of infection. After release from a cell and a return to neutral pH, pr

dissociates from E, exposing the fusion loop of E and priming the virus for fusion.

Flavivirus virus-like particles

During infection, not all viral particles produced are infectious. Some particles,

know as non-infectious subviral particles, or virus-like particles (VLPs), are produced

and released from cells. VLPs contain the structural proteins M and E but are absent of

the non-structural proteins, capsid, and viral genomes. These particles are highly

ordered and retain the physical and chemical properties of infectious virus [82-85].

Since they are non-infectious but structurally resemble infectious virus, VLPs are a

prime vaccine candidate and much information has been gleaned on how to produce

them.

Early studies with recombinant vaccinia virus expressing Japanese encephalitis virus (JEV) proteins indicated that expression of both prM and E was necessary for extracellular release of E, presumably in the form of VLPs [86]. It was then shown that these recombinant vaccinia viruses were able to induce a protective neutralizing antibody response against JEV E protein [87]. Following this, VLPs were purified from the supernatant of cells infected with the recombinant vaccinia virus and injected into mice. Mice that received the VLPs were protected from a lethal challenge with JEV and had similar neutralizing antibody titers against JEV E as mice who had received the recombinant vaccinia virus [88]. These results indicate that expression of prM and E from viral vectors induces the formation of VLPs, and these VLPs stimulate an adaptive immune response upon immunization.

These findings have been substantiated for other flaviviruses such as dengue virus and yellow fever virus, and it has been shown that expression of prM and E from a plasmid is also sufficient to induce the formation of VLPs [89-91]. Further, the transmembrane domains in the c-terminus of E are necessary for proper VLP formation and stability [92]. prM and E alone are sufficient to form VLPs, and co-expression of capsid has been shown to have in inhibitory effect on VLP formation [90].

Flavivirus vaccines

Vaccines exist for flaviviruses such as yellow fever virus, Japanese encephalitis virus, and dengue virus and range from highly successful (yellow fever) to moderately successful (dengue). The oldest of the three is the Yellow Fever virus vaccine, 17D. 17D was developed in 1937 by cultivating yellow fever virus for three years in minced

whole chick embryos from which the central nervous system was removed [93]. This serial passaging made 17D significantly less neurotropic and viscerotropic than the parent Asibi strain. The first human immunizations occurred shortly after the identification of 17D, and since that time, approximately 500 million people have received 17D with only 12 known cases of yellow fever following vaccination [94-96]. Furthermore, a single dose of 17D confers life-long protective immunity against yellow fever [97]. 17D is considered the gold standard of flavivirus vaccines, and given its efficacy against yellow fever virus; it has served as a backbone for other flavivirus vaccines such as the Sanofi Pasteur CYD tetravalent dengue vaccine (TDV).

Since there are four dengue virus serotypes, CYD TDV is composed of four recombinant viruses in which the prM and E genes of YFV have been replaced with prM and E from each of the four dengue virus serotypes. CYD TDV was the first dengue virus vaccine candidate tested in phase III clinical trials. Its efficacy during the first 25 months was 60.3% for all participants but was markedly lower for those under 9 years of age (44.6%) [98]. Given its low protection, further efforts aimed to determine the correlates of protection against dengue virus and develop more effective vaccines. Recent studies using a tetravalent dengue vaccine in the DENV-2 backbone showed that immunization leads to a CD8+ T-cell response against DENV NS1, NS3, and NS5 [99]. These results indicate that a strong cellular immune response in addition to a neutralizing antibody response may be an optimal correlate of protection against Dengue virus [100]. Thus, the best Dengue virus vaccines may need to induce a cell-mediated immune response against the non-structural proteins in addition to a neutralizing antibody response against all four serotypes.

Vaccines for Zika virus

Given the sudden need for a Zika virus vaccine, many research groups are developing candidates. The first published ZIKV vaccine candidates presented a DNA expressing ZIKV prM and E from the Brazil BeH815744 strain as well as an inactivated virus vaccine from the Puerto Rico PRVABC59 strain [101]. First, the researchers tested the immunogenicity of full-length prM and E as well as a number of truncation variants. Mice immunized with prM-E DNA were completely protected against Zika virus challenge. Furthermore, passive transfer of IgG from vaccinated mice to unvaccinated mice was sufficient to protect against ZIKV challenge. Depletion of CD4 and CD8 T cells in vaccinated mice had no effect on protection, indicating that an E-specific antibody response is the main correlate of protection. The inactivated virus vaccine also developed by this group showed similar results [98]. Therefore, it is possible that neutralizing antibody response against ZIKV-E is sufficient to induce protection against ZIKV challenge.

The DNA vaccine, the inactivated virus vaccine, and an adenovirus (Ad)-based vaccine were subsequently tested in rhesus monkeys [102]. All three provided complete protection against ZIKV challenge although there were subtle differences in the immune response between each platform tested. For example, while all vaccine candidates induced neutralizing antibodies, the DNA vaccine did not elicit a significant neutralizing antibody response following the initial priming immunization while both the inactivated virus and adenovirus vaccines yielded neutralizing antibodies after the initial prime. In all three cases, a boost immunization increased the level of neutralizing antibodies. Further, the inactivated virus vaccine induced low levels of cell-mediated

immunity, mostly against E and capsid. Transfer of IgG from immunized rhesus monkeys to mice or naïve rhesus monkey also afforded protection against ZIKV challenge, again indicating that neutralizing antibodies are sufficient to protect against ZIKV challenge. However, while all three candidates tested elicited neutralizing antibodies, both the inactivated virus and the adenovirus-based vaccines outperformed the DNA vaccine. This could indicate that while neutralizing antibodies are sufficient for protection, viral-based vaccine components might increase the magnitude of the immune response.

Since these studies only tested the immune response immediately after vaccination, the same group assessed the immune response in vaccinated rhesus monkeys one year following immunization. They again assessed the inactivated virus vaccine, the adenovirus (Ad)-based vaccine, and a DNA vaccine. At one year post immunization, 6 of 8 rhesus monkeys immunized with the inactivated virus vaccine, 2 of 7 animals immunized with the DNA vaccine, and 4 of 4 animals immunized with the Ad vaccine were protected against challenge with Zika virus [103]. The authors also defined neutralizing antibodies as the main correlate of protection, with neutralizing antibody titers elicited from vaccination with the Ad vaccine being best maintained at one year following the initial immunization. Seemingly, these results suggest expression of ZIKV-E from a replication competent viral vector may be best suited to generate lasting immunity from a single vaccination.

Other vaccine candidates for ZIKV include RNA nanoparticle vaccines [104], modified RNA vaccines [105], virus-like particle (VLP) vaccines [106], as well as vaccines based on vesicular stomatitis virus [107]. All of these studies bolster the idea

that a neutralizing antibody response is the principal correlate of protection against ZIKV infection. Further, while each ZIKV vaccine study measures the immune response generated against that specific vaccine, little is known mechanistically about how these vaccines induce an immune response – particularly for recombinant vaccines that express ZIKV-E. For example, while the DNA vaccine leads to expression of ZIKV-E in cells, it is not shown whether E is released into cell supernatants and in what form that E is [101]. The same caveat applies to the above published VSV-based vaccine candidate. While these findings were published while our research into VSV-based vaccines was still in progress, our work provides fundamental mechanistic insights into how recombinant vaccines may be inducing a neutralizing antibody response against ZIKV. As a result, our work will aid future generation of vaccine candidates for ZIKV as well as for other emerging viruses.

Endogenous retroviruses

Endogenous retroviruses (ERVs) are the remnants of past retroviral infections that exist throughout the genomes of all vertebrates. When a retrovirus infects a cell, a DNA copy of its RNA genome is integrated into the host's DNA due to the retroviral proteins reverse transcriptase and integrase [108, 109]. The integrated DNA copy is known as a provirus. Infected cells produce virus until the provirus is silenced or the cell dies. ERVs are thought to arise when a retrovirus infects and integrates its genome into a germline cell. If this germline cell goes on to produce progeny, the provirus can be vertically transmitted as a heritable genetic element following traditional Mendelian inheritance and can reach fixation in a population over time (Figure 1.5A).

21
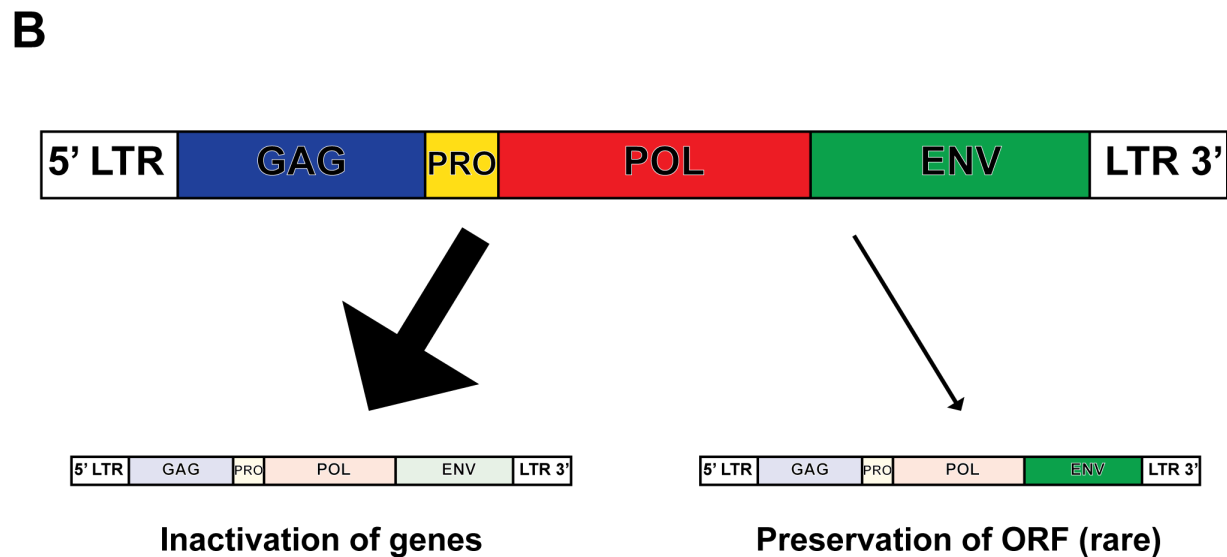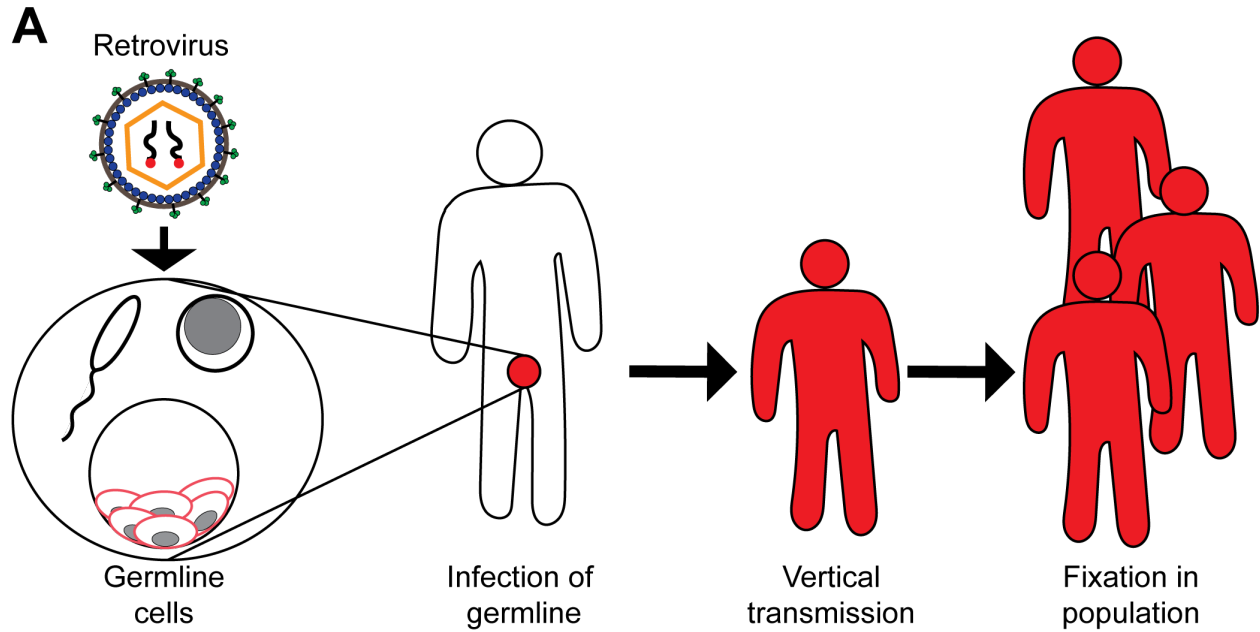
Figure 1.5: Process of endogenization and fates of proviruses. (A) Retroviral infection of germline cells leads to an integrated provirus in that cell. If that germ cell forms offspring, the offspring will have the provirus copy in every cell of its body. Over time, this can be vertically transmitted to others until it eventually reaches fixation in a population. (B) Fates of integrated proviruses.

Figure 1.5 (continued):

An intact, integrated proviral copy contains, at the minimum, 5' and 3' long terminal repeats (LTRs) and the four basic retroviral genes (*gag, pro, pol, and env*). Over time, these genes acquire mutations that render them inactive, but in rare cases, an ORF can be preserved for a beneficial host function.

While most ERVs derive from ancient invasions of mammalian genomes and the mechanism of endogenization is thus inferred, germline colonization is currently occurring in Koala genomes and provides real-time support for the theory that retroviral infection of germ cells leads to endogenization [110]. Further evidence for germline colonization was previously demonstrated in early stage mouse embryos and fertilized chicken eggs [111, 112]. The contribution of human (H) ERV-derived elements to the genome is approximately 8%, and for other mammals it ranges from 5-10% [113, 114]. While HERVs comprise a significant portion of the human genome, they are not all derived from unique viral insertion events. Rather, most HERVs come from a limited number of lineages, the number of which ranges from 31 to more than 300 depending on the method of classification [115, 116]. Therefore, many of HERV elements present in the human genome arose from post-colonization proliferation events such as reinfection, retrotransposition, or complementation *in trans* [116, 117].

Reinfection can occur through infection between germline cells or through infection of germline cells with virus originating in somatic cells. Reinfection is thought to be the dominant method of HERV proliferation due to the fact that the *env* gene,

which is not needed for retrotransposition or complementation *in trans,* is under purifying selection in many ERV lineages [117]. ERV proliferation through retrotransposition occurs through copying of the provirus in a germline cell using the viral machinery, and is evidenced by purifying selection of the *pol* gene but not the *env* gene. Complementation *in trans* is the occurs when the replication machinery is supplied by a separate exogenous or endogenous retrovirus and is thought to account for high copy number HERV lineages such as ERV-9, which do not exhibit purifying selection on any of the proviral genes [116].

Upon integration, an intact provirus includes the four basic retroviral genes; *gag, pro, pol,* and *env,* flanked by two long terminal repeats (LTRs) (Figure 1.5B). Gag is a polyprotein that comprises structural components of the retrovirus core. Pro, the viral protease, aids in the maturation of viral particles. The *pol* gene encodes reverse transcriptase (RT) and integrase, and the *env* gene encodes the surface-bound fusion machinery, Env. Env is a polyprotein consisting of a receptor-binding SU subunit and the transmembrane bound fusion subunit (TM).   Most retroviral Envs resemble class I VFPs.

Since ERVs derive from exogenous viruses, they are classified based on their relation to currently circulating exogenous viruses. Class I ERVs share homology to gammaretroviruses or epsilonretroviruses, class II to betaretroviruses, alpharetroviruses, deltaviruses and lentiviruses, and class III to spumaretroviruses. Phylogeny clustering is most commonly based on comparison of ERV RT domains that are found in all retroviruses [118] (Figure 1.6).
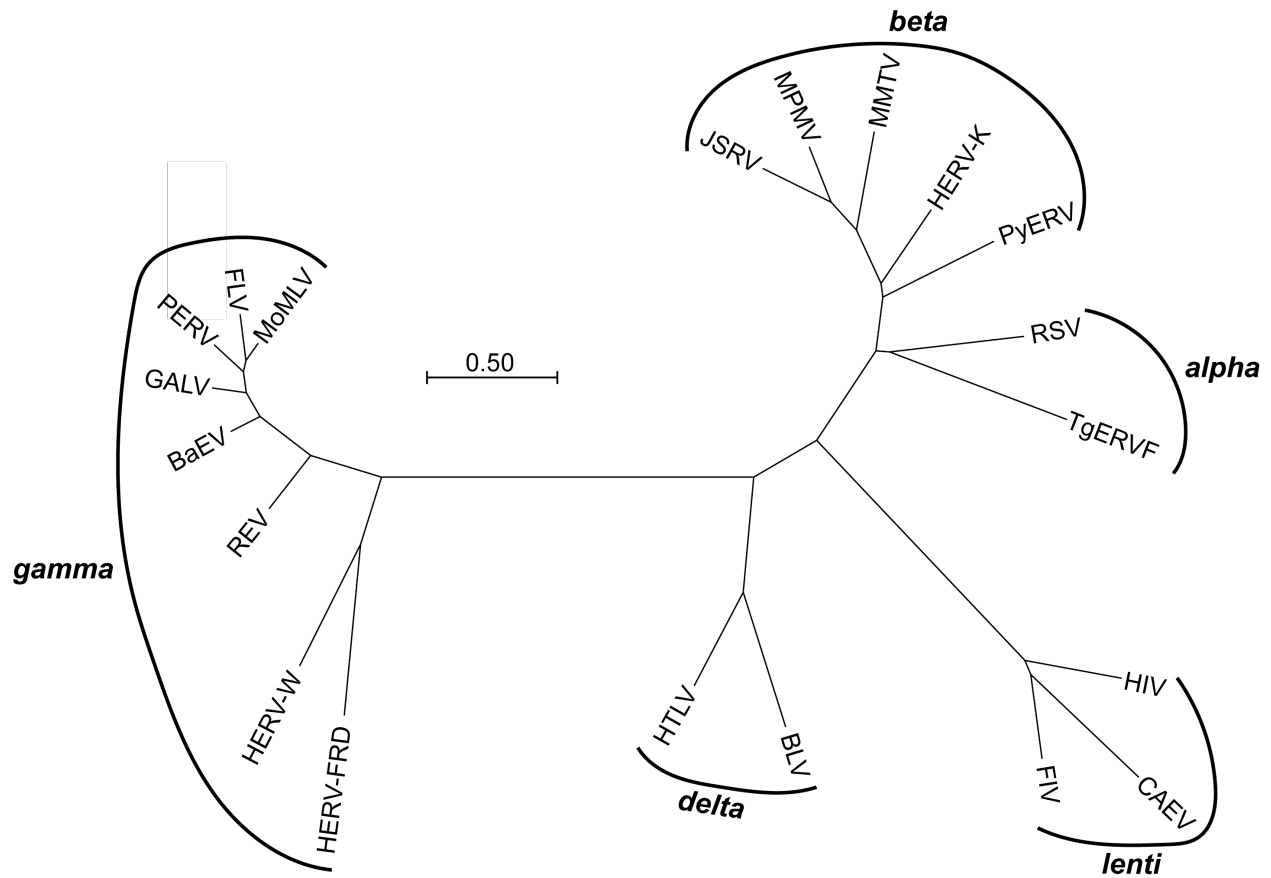
Figure 1.6: Phylogenetic relationship between endogenous and exogenous retroviruses. RT regions of selected retroviruses were aligned, and the evolutionary history was inferred by maximum likelihood based on the General Reverse Transcriptase model [119]. The tree with the highest log likelihood is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Evolutionary analyses were conducted in MEGA7 [120]. Gamma, delta, lenti, alpha, and betaretroviruses are shown. Abbreviations: MoMLV – Moloney murine leukemia virus, FLV – feline leukemia virus, PERV – porcine endogenous retrovirus, GALV – gibbon ape leukemia virus, BaEV – baboon endogenous retrovirus, REV- reticuloendotheliosis virus, HERV-W – human endogenous retrovirus W, HERV-FRD, human endogenous retrovirus FRD, HTLV – human T-lymphotropic virus, BLV – bovine

Figure 1.6 (continued):

leukemia virus, FIV – feline immunodeficiency virus, CAEV – carpine arthritis encephalitis virus, HIV – human immunodeficiency virus, TgERVF – avian endogenous retrovirus F, RSV – Rous sarcoma virus, PyERV – python endogenous retrovirus, HERV-K – human endogenous retrovirus K, MMTV – mouse mammary tumor virus, MPMV – Mason-Pfizer monkey virus, JSRV – jaagsiekte sheep retrovirus. Sequences are as defined in [121].

Additionally, ERV phylogeny can be determined by phylogenetic analysis of conserved regions of the TM domain of Env [122].

Silencing and inactivation of ERVs

To prevent the production of virions, hosts have evolved methods of silencing proviruses. In adult cells, methylation of DNA is enough to silence retroviral transcription, illustrated by the fact that DNA demethylating agents induce ERVs [123, 124]. However, during development when demethylation occurs, other mechanisms are initiated to control ERV activation. One such recently defined mechanism is that TRIM28 binds to primate-specific ERVs in human neural progenitor cells (hNPCs) to recruit H3K9me3 histones to silence ERVs [125]. Knockdown of TRIM28 in hNPCs leads to ERV transcription. While this mechanism and others are able to silence ERV transcription during development and following acquisition, over time, most ERVs acquire debilitating mutations that render them inactive.

Following germline acquisition, ERVs acquire mutations at the rate at which their host acquires mutations during DNA replication. Given this, and because proviral LTRs

are identical at the time of integration, the number of nucleotide differences between

two LTRs of the same provirus can be used to estimate the time of endogenization.

Additionally, estimates based on species divergence times and newly developed

phylogenetic methods can be used to estimate ERV integration times [126]. Through

these methods, it is appreciated that most HERVs arose in a primate ancestor between

25 and 50 million years ago, with the evolutionarily newer HERV-K clade group

members emerging more recently, approximately within the past 2 million years [126-

128].

Since most of the ERVs in the human genome arose 25-50 million years ago,

they have since been inactivated through the acquisition of frameshift mutations, stop

codons, and large deletions that prohibit the production of viral genes [129].

Additionally, the majority of ERV elements have undergone recombinatorial deletion in

which recombination of the proviral LTRs removes the coding region of the ERV [130,

131]. These examples of inactivation apply to almost all HERV genes that have been

identified. However, in rare cases, some ERV genes still contain intact open reading

frames (ORFs) through a process known as co-option or exaptation (Figure 1.5B) [132].

Exaptation of ERV ORFs

The preservation of an ancient ERV ORF when the rest of its genes have

acquired debilitating mutations strongly indicates that that ORF has been repurposed

(or exapted) by the host to provide a beneficial function to the host. The most striking

example of exaptation comes from ERV *env* genes that have been repurposed to play a

positive role in the placenta. These *env* genes have since been termed *syncytins* [133].

Two *syncytins*, *syncytin-1* and *syncytin-2*, exist in the human genome and derive

from the HERV-W and HERV-FRD lineages respectively. *Syncytin-1* was first defined

in 1999 when researchers were searching for multiple sclerosis-associated retroviruses.

Due to its placenta-specific expression, it was hypothesized to play a physiological role

during pregnancy [134]. *In situ* hybridization studies revealed that *syncytin-1* is

expressed specifically in the syncytiotrophoblast cells of the placenta [133]. The

syncytiotrophoblast is a large multinucleated cell layer that constitutes the boundary

between fetus and mother. Further, transfection of DNA plasmids expressing *syncytin-1*

into cultured cells induced cell-cell fusion, and thus, the criteria for a *syncytin* were

established – an evolutionarily conserved, intact endogenous retrovirus envelope

protein that is expressed in the placenta and mediates cell-cell fusion.

Additional bioinformatics searches for intact retroviral envelopes identified a

HERV-FRD envelope (*syncytin-2*) that is fusogenic in human cells and is conserved in

all hominoids, old world monkeys (OWMs), and new world monkeys (NWMs) [135]. In

addition to its placenta-specific expression and fusogenicity, there is strong evidence of

purifying selection ($d_n/d_s < 1$) on the *syncytin-2* gene, indicating that it has been

maintained as a fusogen in the genomes of all simians for at least 40 million years

[136].

Following the discovery of human *syncytins*, it was hypothesized that similar

syncytins must exist in other placental mammals. Since that time, *syncytins* have been

discovered in all placental mammals examined, including mice, ruminants, and even

marsupials [137-139]. Most recently, it was shown that the Mabuya lizard, which

generates an independently evolved placenta similar to the mammalian placenta, also

encodes *syncytins* [121]. The fact that each *syncytin* was independently acquired from a different ERV lineage suggests that acquisition of retroviral *env* genes directly facilitated placenta evolution. While the necessary role for *syncytins* in humans can only be speculated, studies of mouse syncytins provide definitive proof of their indispensability. *Syncytin-A* knockout mouse embryos die between day 11.5 and 13.5 of gestation due to improper placenta morphogenesis [140]. Trophoblast cells fail to fuse and establish the syncytiotrophoblast layer, which leads to imbalanced vascularization and placental transport. Thus, given the similarities between mouse and convergently evolved *syncytins* of other species, it is possible that all *syncytins* are indispensible for embryo survival. As a result, it is possible that an initial *syncytin* was captured in the ancestor to placental mammals and has been replaced with lineage-specific *syncytins* with improved function since that time.

Immunomodulatory function of retroviral Envs

Studies of retroviral Env TM subunits indicated that they exhibit immunosuppressive properties [141, 142]. Further work characterized the immunosuppressive domain to be a 16 amino acid region within the TM subunit that shares significant homology between different retroviruses [143]. Addition of this 16 amino acid region, termed CKS-17, in the form of a synthetic peptide to cell culture media inhibited proliferation of the mouse lymphoblast cell line, CTLL-2, in a dose dependent manner while two partially homologous sequences and two non-homologous sequences did not [143]. The CKS-17 peptide also inhibited proliferation of murine and

human lymphocytes indicating that this region is likely responsible for the immunosuppressive properties exhibited by some retroviral Env TM subunits.

Following the discovery that placenta-expressed ERV Envs also harbored a CKS-17-like domain; it was hypothesized that the immunosuppressive properties could be maintained to play a role in tolerance of the placenta to the maternal immune system. To test this, allogeneic MCA205 tumor cells stably expressing either *syncytin-1* or *syncytin-2* were injected into BALB/c mice. Normally, these tumor cells are rejected, but when they express immunosuppressive retroviral Envs, the tumors can persist and grow [144]. Based on this assay, it was found that syncytin-2, but not syncytin-1, was immunosuppressive [145]. It was determined that one amino acid within the immunosuppressive domain of syncytin-1 was responsible for its lack of *in vivo* immunosuppression but did not affect its fusogenicity. Interestingly, the immunosuppressive domains (ISDs) of syncytin-1 and syncytin-2 are 100% conserved in the primate species in which they are found. Further, the same pattern is found in mice, with one syncytin being immunosuppressive and the other not [145].

Functional HERV Env open reading frames

Given the significance of syncytins to mammalian physiology, other studies have sought to identify ERV ORFs exapted for other physiological roles. One recent study identified a novel HERV *envelope* ORF named HEMO [146]. RNA-sequencing (seq) analysis revealed that HEMO is expressed highly in the placenta and at low levels in other normal human tissues. The ORF has been preserved over 100 million years and contains hallmarks of purifying selection [146]. HEMO lacks a consensus furin cleavage

site, but Heidmann *et al.* showed that a secreted form is released into the supernatant of transfected cells and is shed into a pregnant mother's blood [146]. A function has not yet been ascribed to HEMO, but the existing observations implicate a physiological role that is distinct from cell-cell fusion.

Another potential role for co-opted retroviral Envs is through receptor interference. When expressed in cells, retroviral Envs can traffic to the surface and interact with their receptor, thereby blocking it from interacting with other viral fusion proteins recognizing the same receptor. In some cases, the Env-receptor interaction leads to downregulation of the receptor. This is the case for the HERV-T Env. The HERV-T Env contains an intact open reading frame but lacks a canonical furin cleavage site meaning that it does not function as a fusogen. Paul Bieniasz's group reconstructed an ancestral ERV-T Env harboring an intact furin cleavage site and showed that it was fusogenically functional and able to pseudotype retroviral particles [147]. When cells expressing HERV-T Env were challenged with particles bearing the reconstructed ancestor, the cells were blocked from infection due to HERV-T Env depleting its receptor, MCT-1, from the surface. A similar mechanism was previously demonstrated in mice whereby expression of a proviral env, *Fv-4*, can block infection of ecotropic and polytropic murine leukemia viruses [148].

Given the propensity for retroviral envelope genes to be co-opted for a range of beneficial functions to a host, other studies have searched for intact retroviral envelopes in the human genome. One such study used BLAST searches to detect novel proviral ORFs in the human genome through their similarity to known retroviral sequences [149]. They identified 7,836 total proviral loci, of which 59 contained intact or almost intact

ORFs. Of these 59, 29 are in the *env* gene, and 15 of these were of *gammaretroviral*

origin. One of these *env* sequences was identified due to its sequence similarity to a

Zebrafish endogenous retrovirus (ZFERV). This ZFERV-like *env* was thought to derive

from a gammaretrovirus due to weak similarity in its *gag* and *pol* regions to feline

leukemia virus (FLV). Although no expressed sequence tags (ESTs) were identified for

the ZFERV-like *env*, other studies followed up on it due to its coding potential.

EnvPb1

Theirry Heidmann's group was the first to assess the ZFERV-like Env as a

potential syncytin [150]. A phylogenetic tree based on the TM-region of Env placed this

unique Env in the HERV-I superfamily along with the ZFERV. Using the intact provirus,

they located 3 other copies of this HERV in the human genome. Since all had a primer-

binding site (PBS) in the 5' LTR that most closely resembled Pro-tRNA, this family of

proviruses was named HERV-P(b) to differentiate it from the previously identified

HERV-P group. The fully intact locus was named HERV-P(b1). HERV-P(b1) resides

within an intron of the *RIN3* gene on human chromosome 14 (Figure 1.7). While *env*

encodes an intact ORF, *gag* and *pol* contain deletions and stop codons.
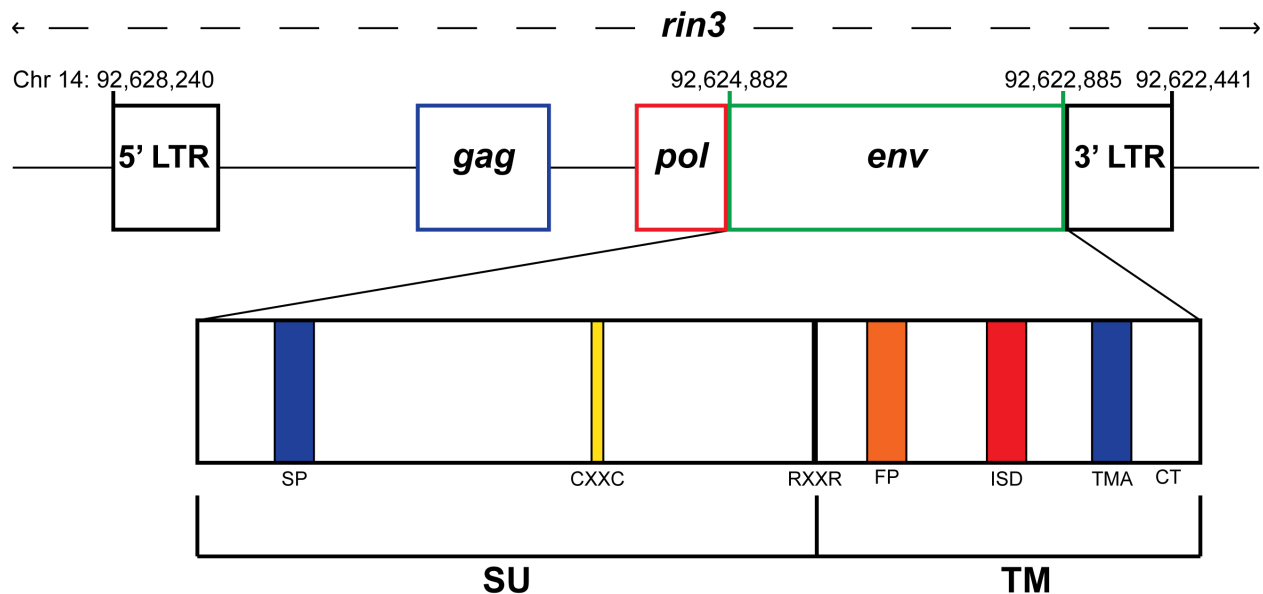
Figure 1.7: Schematic representation of the HERV-Pb1 locus. HERV-Pb1 is

situated within an intron of the *RIN3* gene on chromosome (Chr) 14. The roughly 6-

kilobase (kb) provirus harbors 5' and 3' LTRs, truncated *gag* and *pol* genes, and a fully

intact *env* ORF. The *env* ORF encodes the two retroviral subunits, SU and TM, which

are expected to be cleaved by the RXXR furin cleavage site. It is predicted that SU

contains a signal peptide (SP) and TM harbors a transmembrane anchor (TMA)

(TMHMM Server - http://www.cbs.dtu.dk/services/TMHMM/) [151]. Hyrdophobicity plots

do not predict a canonical fusion peptide (FP) at the N-terminus of TM, but rather there

is a slightly internal hydrophobic stretch that is suggested to serve as a fusion peptide

[152, 153]. The *env* ORF also has a predicted immunosuppressive domain (ISD), and a

CXXC motif [152].

Quantitative reverse transcription polymerase chain reaction (RT-qPCR) of 19 different human tissues revealed that the *envPb1* is expressed at low levels in all tested tissues and exhibits no placenta specificity as the syncytins do [150]. Expression of human *envPb1* in a panel of 9 different cell lines led to fusion in 5 of them, indicating that EnvPb1 may function as a physiological fusogen distinct from the syncytins [150]. This fusion activity was observed despite a canonical fusion peptide at the N-terminus of the TM domain of EnvPb1.

Evolutionary characterization of *envPb1*, showed that the provirus, ERVPb1, is present at the same chromosomal locus in chimpanzees, and fully intact *envPb1* reading frames are present in hominoids and Old World monkeys (OWMs) [152]. Although full-length *envPb1* was not detected in New World monkeys (NWMs), internal regions were amplified indicating the potential presence in NWMs. Molecular clock analysis of the ERVPb1 LTRs indicates integration into an ancestral genome between 27-36 million years ago (mya). Further, no full-length hits with at least 50-70% homology were detected in prosimian, rodent, chicken, or fish genomes [152].

All identified species that harbor an intact EnvPb1 contain a predicted furin cleavage site, a CKS-17-like immunosuppressive domain, a transmembrane region, and a potential internal fusion peptide indicating that other primate EnvPb1 may also function as fusogens.

D_n/D_s analysis of the *envPb1* ORF provided a mean value that did not differ significantly from 1; however, the d_n/d_s varies over the length of the gene with functional regions being highly conserved [152]. This raises the possibility that the most important functional regions are under purifying selection. The *envPb1* gene is the sole intact

gene in the HERV-P(b1) proviral structure, suggesting that this gene has been selectively maintained. Out of 17 single-nucleotide polymorphisms (SNPs) within the HERV-P(b1) locus, only one is located within the *env*-coding region, and that SNP is synonymous [152]. Additionally, there are 14 insertion/deletion events in HERV-P(b1); the three within *envPb1* do not disrupt the reading frame while the others are all reading frame disrupting in their respective genes [152]. It is unlikely for these events to have occurred by chance, and such data are consistent with selective maintenance of the *env* ORF.

Aagaard *et al.* confirmed the ubiquitous expression of *envPb1* in human tissues but found minor differences in relative expression levels compared to Blaise *et al* [150, 152]. Interestingly, *envPb1* is most highly expressed in the thymus. While it is expressed in the placenta, its expression is 100-fold lower than *syncytin-1* [152]. Analysis of RT-PCR products revealed that the transcript was translation competent, and it was hypothesized that the 5' LTR has the capability to drive transcription due to an intact TATA box.

Further work characterized EnvPb1 transcription by 5' and 3' rapid amplification of cDNA ends (RACE) in primary cytotrophoblasts. EnvPb1 transcripts mapped to an initiation site in the HERV-P(b1) 5' LTR near the putative TATA box [154]. Additionally, splice donor and splice acceptor sites were mapped in the 5' LTR, and a poly-A signal, AAUAAA, was mapped in the 3' LTR. This study also generated a plasmid in which the HERV-P(b1) 5' LTR was inserted upstream of a luciferase reporter gene to test transcriptional activity of the LTR. Upon transfection into the trophoblast-derived choriocarcinoma BeWo cell line followed by stimulation with forskolin to induce

syncytium formation, a slight increase in luciferase signal was observed indicating that the HERV-P(b1) 5' LTR is transcriptionally active under conditions thought to stimulate syncytia formation.

Despite being expressed in BeWo cells during steady state and upon stimulation, EnvPb1 does not play a role in BeWo cell fusion events, and it was concluded that it is likely not involved in trophoblast fusion [154]. To test whether EnvPb1 could act in heterotypic cell fusions between cancer cells and epithelial cells, BeWo cells, which express EnvPb1, were added to cultures of HUVEC cells. This led to heterotypic cell fusions [153]. However, following knockdown of EnvPb1 in BeWo cells by shRNA, heterotypic cell fusion was significantly reduced indicating not only that HUVEC cells express the EnvPb1 receptor, but also that EnvPb1 could mediate heterotypic cells fusion in a cancer setting [155].

Immunosuppression is another potential function for ERV Envs. EnvPb1 contains a CKS-17-like ISD domain that was tested *in vivo* for immunosuppressive potential. Allogeneic MCA205 tumor cells expressing EnvPb1 but not an EnvPb1 ISD mutant prevented tumor cell rejection, indicating that EnvPb1 has a functional ISD that may be relevant for its physiological role [145].

Given that EnvPb1 is one of three intact human endogenous retroviral Envs that has retained the ability to fuse cells under physiological conditions and is highly conserved, it is likely that any potential physiological function it plays is related to cell-cell fusion events. However, unlike the placenta-specific expression of the syncytins, EnvPb1 is expressed at low levels in a wide range of human tissues, making it difficult to hypothesize a physiological function based on its expression pattern. Aside from

syncytiotrophoblasts in the placenta, other human cells that undergo cell-cell fusion include sperm-oocyte fusion, monocytes fusing into syncytia osteoclasts, and myoblasts forming multinucleated muscle fibers. While the potential role of retroviral Envs, particularly EnvPb1, has not been extensively examined in these multinucleated cell types, some studies have implicated a potential role for syncytin-1 in generation of osteoclasts as well as myoblasts [156-158]. Additionally, skeletal muscle isolated from human cyclists before and after long-term endurance training revealed upregulated mRNA levels of *syncytin-1*, *envPb1*, and *erv-3 env* while no significant increase in mRNA from 19 other ERV *envs* was exhibited [158]. These results were replicated in cell-culture models using isolated primary myoblasts. Further, myoblast fusion in cell culture was blocked by addition of anti-syncytin-1 antibody, implicating a direct role for syncytin-1 in myoblast fusion. Whether EnvPb1 plays a role in myoblast fusion remains to be determined.

In an effort to better understand if EnvPb1 is physiologically relevant and to assess what role it could be playing, we set out to develop tools to study EnvPb1, test its evolutionary conservation, and understand how it initiates cell-cell fusion. These advances should pave the way for future studies to determine what physiological role it may be playing in primates.

Chapter 2: VSV-based vaccines induce a neutralizing antibody response against

Zika virus

Abstract

Zika virus (ZIKV) was not appreciated to be a significant human pathogen until the Brazilian outbreak of 2015 when it was discovered to be a causal agent of microcephaly. Due to its rapid emergence, no vaccines or treatments were available. To facilitate rapid vaccine development, we utilized recombinant Vesicular stomatitis virus (rVSV) to generate and characterize vaccine candidates for ZIKV. VSV is an ideal vaccine vector because it grows rapidly in cell culture, induces a strong immune response, is not associated with human disease, and has low seroprevalence in the human population. Our findings indicate that rVSV expressing the ZIKV envelope (E) and premembrane (prM) proteins yield significant levels of ZIKV virus-like particles (VLPs). We show that these VLPs alone or the rVSV that best produces them are sufficient to induce a neutralizing antibody response in mice. Additionally, we report VSV incorporation of functional ZIKV-E. Our results have implications for Zika virus and emerging flavivirus vaccine design.

Introduction

Zika virus (ZIKV) is a flavivirus that has infected over 1 million individuals in the Americas since an outbreak began in Brazil in 2015 [159]. Until this outbreak, ZIKV was thought to be a relatively innocuous virus causing only mild, influenza-like symptoms. However, after ZIKV began to spread throughout the Americas, it was realized that ZIKV is a causative agent of microcephaly in unborn fetuses [73, 160]. Given the unprecedented outbreak and the newly appreciated complications of infection, Zika virus was declared a public health emergency of international concern by the World Health Organization (WHO).

Despite the need, no vaccines are currently available for use against ZIKV, although a multitude of candidates are in development. Such candidates include DNA-based, adenovirus vectors, inactivated virus, mRNA, and others [101, 102, 104, 105]. Many of these vaccine candidates employ a vector to express the ZIKV envelope (E). As a result, these vaccines induce a neutralizing antibody response against ZIKV-E. However, while the viral vaccine vectors cited above all express ZIKV-E in cells, further characterization of the ZIKV-E protein product produced by these vaccine vectors is severely lacking. Furthermore, while the boundaries of E and its chaperone, premembrane (prM) are well defined, the various vaccine candidates express different constructs of these proteins. As a result, some vaccine candidates may yield ZIKV-E that is inefficiently processed and released from cells leading to a dampening of the immune response to ZIKV-E. Given the success of a recombinant vesicular stomatitis virus (rVSV)-based Ebola virus vaccine, VSV-ZEBOV, and the fact that VSV vaccines generate a potent immune response and are safe for use in humans, we aimed to

generate VSV recombinants that express the Zika envelope protein as vaccine candidates [161].

Zika virus is a member of the *Flaviviridae* virus family. As such, it contains a positive sense, single-stranded RNA genome. Its open reading frame encodes a polyprotein that is processed into structural and non-structural proteins. The structural proteins consist of capsid (C), precursor membrane (prM), and envelope (E). Similar to other flaviviruses, these structural proteins assemble and bud into the endoplasmic reticulum (ER) to form a non-infectious, immature particle consisting of 60 trimeric spikes of prM:E heterodimers anchored to a lipid membrane [79]. This immature virion traffics through the Golgi where low-pH induces structural rearrangements into 90 dimers of M:E heterodimers. In the trans-Golgi network, prM is cleaved by furin into pr peptide and M protein. The ZIKV virion now has a smooth surface and is primed for membrane fusion following exocytosis from the cell.

This process of assembly and maturation is likely why no VSV recombinant that incorporates functional flavivirus envelope proteins and grows to high titers has been reported despite the fact that viral fusion proteins from all three classes can complement VSV [33, 63, 64]. Flaviviruses assembly and budding occurs in ER-associated membranes whereas VSV assembles and buds at the plasma membrane [19, 77]. Therefore, flavivirus E protein may not be at the proper cellular localization for efficient incorporation into VSV. Additionally, expression of flavivirus E and its chaperone, prM, leads to formation of virus like particles (VLPs) that mature and assemble similarly to infectious virions [82-85, 89]. Therefore, not only is E unlikely to be at the right locale

for incorporation into VSV, but also it is likely already packaged into VLPs and unavailable for VSV incorporation.

We first aimed to determine the ZIKV coding region that would lead to optimal release of E from rVSV-infected cells.   We then investigated incorporation of E in both VLP and VSV fractions. We report incorporation of functional ZIKV-E into VSV virions. Additionally, we defined an optimal coding region of prM and E from which cellular expression generates significant yields of ZIKV virus like particles (VLPs) capable of inducing a neutralizing antibody response against ZIKV-E. Our work demonstrates the utility of using VSV as a means for increasing immunological ZIKV VLP production as well as for a vaccine vector against ZIKV.

Results

Construction of rVSV-ZIKV constructs and recovery of VSV expressing ZIKV-E

We chemically synthesized the region of the ZIKV genome encoding ZIKV prM and E from the Suriname strain, which was isolated in November 2015 [162]. This construct is referred to as prME. Additionally, we made a second construct in which we added the C-terminal transmembrane region of the capsid gene to test if this presumed signal peptide would improve ZIKV-E expression as has been reported for other flaviviruses such as West Nile Virus [163]. The construct containing this signal peptide is referred to as CprME. From these two constructs, we made four recombinant VSV; two in which the ZIKV construct takes the place of VSV-G (rVSV-ΔG-prME and rVSV-ΔG-CprME) and two that have the ZIKV constructs in the first genomic position and retain VSV-G (VSV-prME and VSV-C-prME) (Figure 2.1A).

A

rVSV-eGFP

2.2±0.4 X 10^9 3.1±0.2 X 10^9

rVSV-ΔG-prME

5.0±0.9 X 10^5

rVSV-ΔG-CprME

2.3±2.7 X 10^5

rVSV-prME

0.7±1.3 X 10^7

rVSV-CprME

1.8±0.6 X 10^7

B

L

G

P

N

E

M

C

VSV G

ZIKV E

ZIKV E 10X exp

VSV M

D

rVSV-ΔG-prME   rVSV-ΔG-CprME   rVSV-prME   rVSV-CprME

ZIKV-MR766   rVSV-eGFP
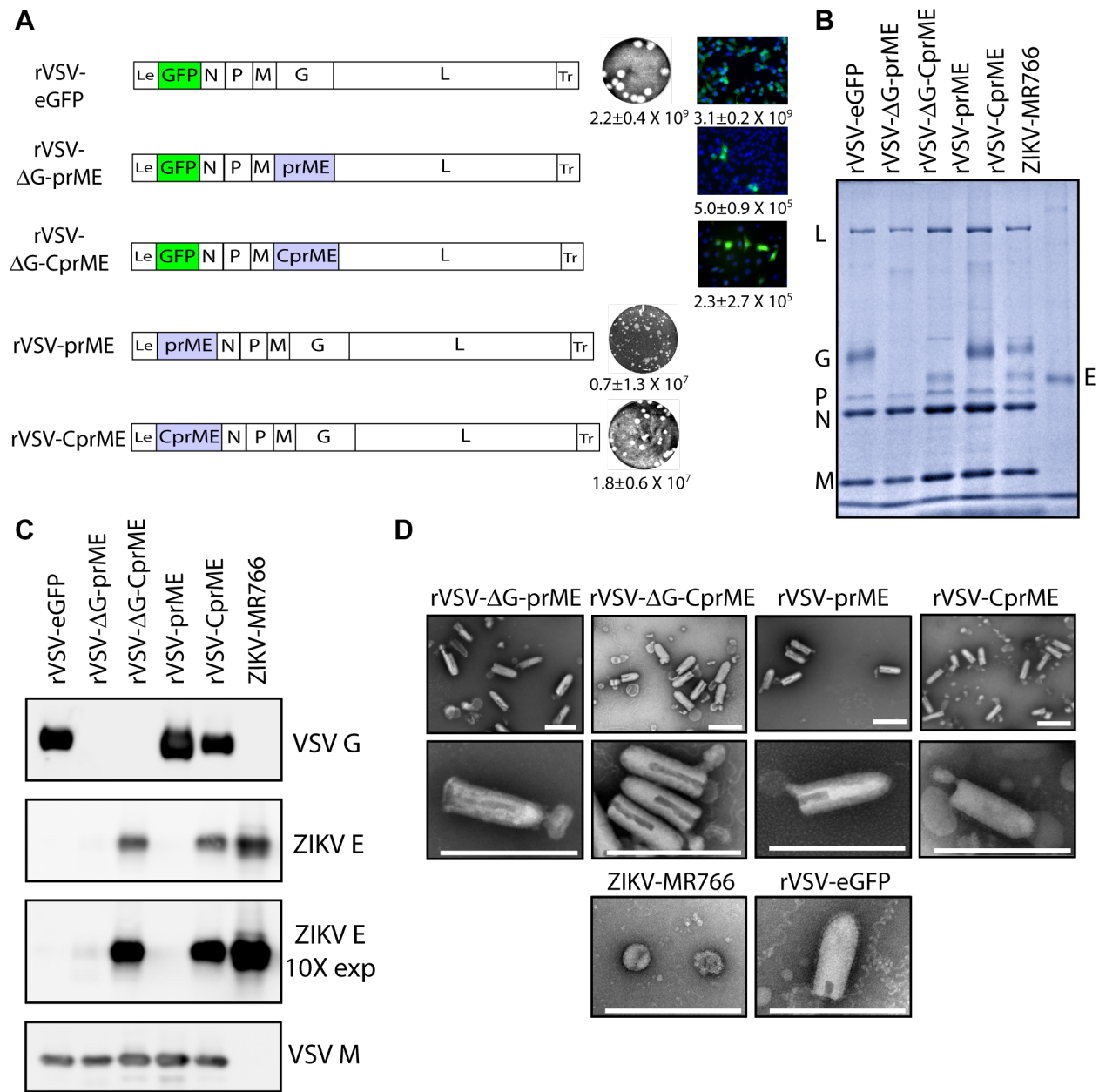
Figure 2.1: Composition and infectivity of rVSV-ZIKV. (A) Schematic diagram of rVSV-ZIKV genomes. The VSV genes encoding N, P, M, G, L, and eGFP are shown flanked by the viral leader (Le) and trailer (Tr) regulatory sequences. The ZIKV precursor membrane envelope (prME) together with the signal peptide portion of the capsid protein (C) is shown together with its position of insertion. Viral titers

43

Figure 2.1 (continued):

determined by plaque assay on Vero cells for viruses encoding VSV G are expressed

as plaque forming units per milliliter (pfu/ml) and representative images shown. For

viruses lacking VSV-G, infectious units were determined by flow cytometry detection of

eGFP positive cells and are expressed as infectious units per milliliter (iu/ml) and

representative images shown at 10X magnification. (B) Protein composition of purified

virions. Equivalent protein amounts of VSV-eGFP, rVSV-ZIKV, and ZIKV (strain:

MR766) were analyzed by SDS-PAGE and Coomassie staining. (C) Glycoprotein

composition of purified virions. VSV-G and ZIKV-E were detected by Western blot using

glycoprotein-specific antibodies together with VSV-M as a VSV control. A 10X exposure

of the 4G2 Western blot is also shown to visualize the relatively low levels of E

incorporated in prME-containing viruses. (D) Structural properties of purified viral

particles. Sucrose gradient purified virions were visualized by transmission electron

microscopy (TEM). VSV virions were stained with 2% phosphotungstic acid (PTA) and

MR766 was stained with 0.5% uranyl formate. Scale bar = 200 nm.

Additionally, rVSV-ΔG- prME and rVSV-ΔG-CprME encode a codon optimized green fluorescent protein (eGFP) for assessing infectivity [164]. Collectively, these viruses will be referred to as rVSV-ZIKV.

The rVSV-ΔG constructs aimed to assess ZIKV-E incorporation into VSV and the first position viruses aimed to produce maximum levels of ZIKV-E in cells since the first VSV gene position in VSV is the most highly transcribed [17]. rVSV-prME and rVSV-CprME grow to titers of $\sim$1x10^7 pfu/ml whereas rVSV-ΔG-prME and rVSV-ΔG-CprME grow to titers of 5x10^5 and 2.3x10^5 infectious units/ml respectively (Figure 2.1A). While rVSV-prME and rVSV-CprME form plaques, rVSV-ΔG-prME and rVSV-ΔG-CprME do not, which indicates that rVSV-ΔG-prME and rVSV-ΔG-CprME do not spread efficiently in cell culture. Indeed, further passaging of rVSV-ΔG-prME and rVSV-ΔG-CprME in various cell lines (Vero, BSRT7, and c6/36) or at different temperatures (28ºC, 30ºC, 34ºC) did not significantly increase the number of infectious particles in the supernatant (data not shown).

Assessing the composition of purified particles

To determine the composition of viral particles, we purified supernatants of cells infected with the rVSV-ZIKV, VSV, or ZIKV strain MR766 by filtration through 0.22µm filters and ultracentrifugation. Viral proteins were analyzed either by Coomassie staining or by western blot using the pan-flavivirus E-protein 4G2 antibody. By Coomassie, we detect approximately equal amounts of the VSV proteins, nucleocapsid (N), phoshoprotein (P), matrix (M), and the RNA-dependent RNA-polymerase (L) between viruses, indicating the release of viral particles (Figure 2.1B). Additionally, we

detect high levels of ZIKV-E in the supernatant of cells infected with rVSV-ΔG-CprME

and rVSV-CprME as evidenced by co-migration with ZIKV-E in the MR766 lane.

However, ZIKV-E is not visible in rVSV-ΔG-prME or rVSV-prME supernatants. Western

blotting with the 4G2 antibody substantiated this result (Figure 2.1C). At longer

exposures, we detect a low-level of ZIKV-E in rVSV-ΔG-prME or rVSV-prME

supernatants; however, it is significantly lower than what we see for rVSV-ΔG-CprME

and rVSV-CprME as well as for MR766. This suggests that the transmembrane region

between capsid and prM is important for either expression or release of ZIKV-E from

cells and could act as a signal peptide as reported for other flaviviruses [165].

Next, we imaged the purified virions by transmission electron microscopy to test

if incorporated ZIKV-E was visible or caused changes in VSV particle shape. Virus was

absorbed onto copper-coated grids and stained with 2% phosphotungstic acid (PTA).

ZIKV-E was not visible on the surface of rVSV-ΔG-prME or rVSV-ΔG-CprME virions

(Figure 2.1D). This could either be because these viruses do not incorporate significant

amounts of ZIKV-E or because flavivirus E proteins in a mature, fusion competent state

are folded close to the virus lipid bilayer and form ordered structures on virion surfaces

that might not be visible on VSV. Conversely, we do see glycoprotein spikes on the

surfaces of rVSV-prME and rVSV-CprME. These spikes resemble those on rVSV-

eGFP, which implies that these viruses preferentially incorporate VSV-G.

ZIKV-E is expressed in rVSV-ZIKV infected cells

Because ZIKV-E was barely detected in the supernatant of cells infected with

rVSV encoding prME, we wanted to assess whether this was due to poor expression in

46

cells. We infected Vero cells at an MOI of 3 with each of our rVSV-ZIKV or rVSV-eGFP

and at an MOI of 10 with MR766. [^{35}S]-methionine was added to infected cells at a time

of optimal viral protein synthesis. Cell lysates were collected and subject to SDS-

PAGE. The gel was exposed overnight to a phosphor screen and the screen was

imaged on a Typhoon phosphorimager. Lanes containing the four rVSV-ZIKV show

similar ratios of VSV protein expression (Figure 2.2A). Furthermore, all rVSV-ZIKV

express roughly equal levels of ZIKV-E in cells, indicating that the lack of ZIKV-E in the

supernatant of cells infected with prME viruses is not due to a deficit in cellular

expression. Importantly, ZIKV-E expressed from VSV runs at the same molecular

weight as that from cells infected with ZIKV-MR766 indicating that it is being cleaved

from prM in cells.
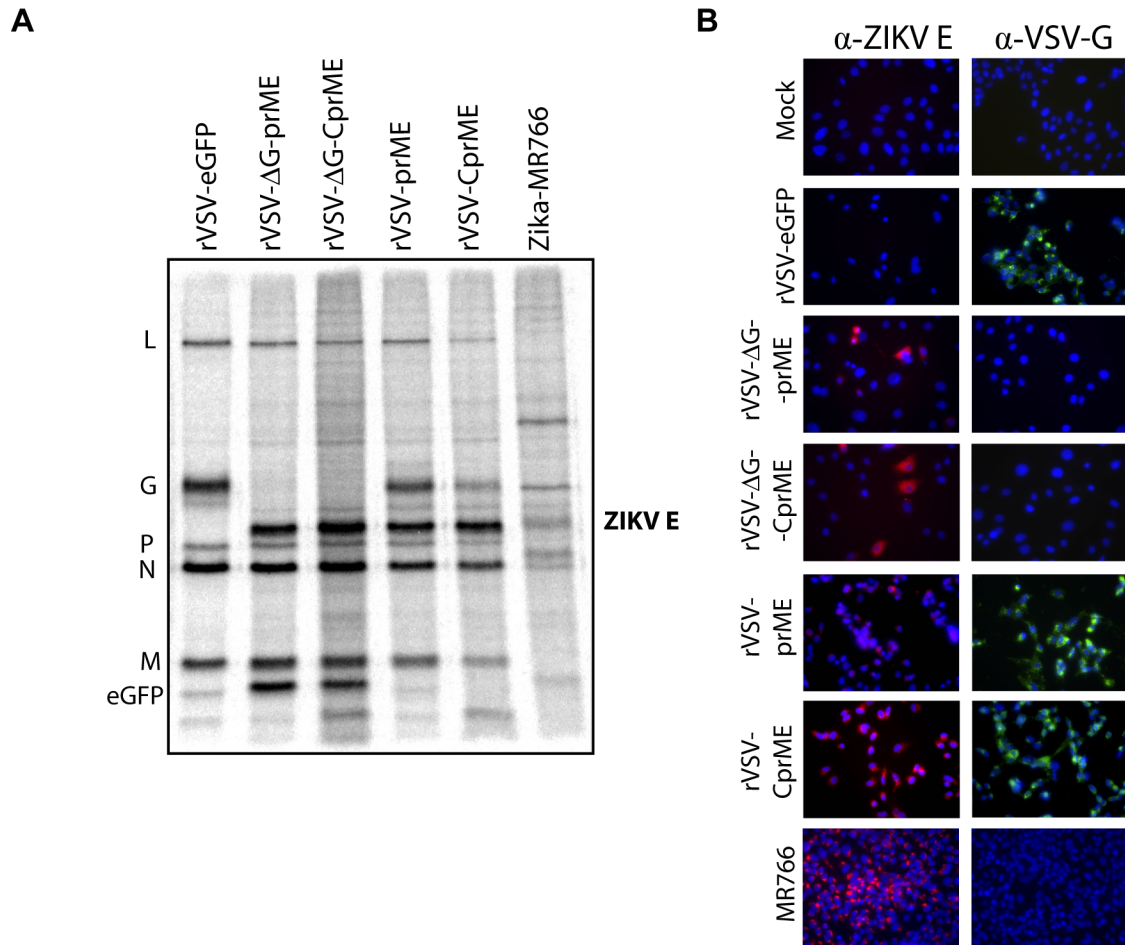
Figure 2.2: rVSV-ZIKV vectors express ZIKV E protein in infected cells. (A) Total

protein synthesis in infected cells. Vero cells were infected (MOI=3) with the indicated

VSV recombinant, and viral protein synthesis was measured by metabolic incorporation

of [^{35}S]-methionine at 5 hours post infection (HPI). Protein synthesis in MR766 infected

cells (MOI=10) was measured by metabolic labeling for 2h at 40 HPI. Total cytoplasmic

extracts were prepared, equivalent amounts analyzed by SDS-PAGE, and labeled

proteins visualized by autoradiography. The VSV proteins, together with the eGFP

reporter are identified on the left side of the gel, and ZIKV-E on the right. (B) Detection

of viral proteins by immunofluorescence microscopy. Vero cells were infected as in

panel A, fixed at 6 HPI (VSV) or 48 HPI (MR766) prior to protein detection using primary

Figure 2.2 (continued):

antibodies 4G2 for ZIKV-E or the VSV-G cytoplasmic tail antibody, V5507. Primary

antibodies were detected using Alexa Fluor 594 goat α-mouse secondary antibody and

images false colored red for ZIKV-E and green for VSV-G. Nuclei (blue) were visualized

by staining with NucBlue™. Images are representative of 3 independent experiments.

      This result suggests that defects in expression or processing are not reasons

why we do not see ZIKV-E in the supernatant of cells infected with the prME viruses.

To further visualize ZIKV-E expression, we performed immunofluorescence on cells

infected with our rVSV-ZIKV (Figure 2.2B). Infection proceeded for 6 hours, after which

cells were fixed, permeabilized, incubated with anti-ZIKV-E or anti-VSV-G antibodies,

and then stained with an Alexa-labeled secondary antibody. We detect ZIKV-E

expression in cells infected with all four rVSV-ZIKV (Figure 2.2B). While rVSV-prME

and rVSV-CprME exhibit similar levels of VSV-G staining, rVSV-prME consistently

exhibited a faint ZIKV-E signal. Furthermore, while ZIKV-E staining in MR766-infected

cells appears in punctae mostly likely confined to the ER, Golgi, and trans-Golgi

network, expression in cells infected with rVSV-ZIKV seems to be more diffuse

throughout the cell. Therefore, expression of ZIKV-E during VSV infection may impact

normal ZIKV-E cellular localization.

Entry of rVSV-∆G-prME and rVSV-∆G-CprME depends on ZIKV-E

To date, only one VSV recombinant has been reported in which VSV-G is replaced by a flavivirus envelope protein [166]. Recombinant VSV with Japanese encephalitis virus (JEV)-E are not highly infectious, but incorporate functional JEV-E [166]. In turn, we wanted to test if our rVSV-ZIKV incorporate mature ZIKV-E. To test if rVSV-ZIKV infection is ZIKV-E-dependent, we utilized a small molecule inhibitor, 3-110-22, that was previously reported to block dengue virus entry into cells by prohibiting the low-pH induced conformational changes that lead to fusion [167]. Unpublished work from the Yang lab confirmed that 3-110-22 also blocks ZIKV entry. To assess E-dependent entry, we incubated our four rVSV-ZIKV, ZIKV MR766, or rVSV-eGFP with 10µM 3-110-22. As a control for VSV-G mediated infection, we separately incubated viruses with 0.1 µg of the VSV-G neutralizing antibody IE2 [168]. The inhibitor/antibody virus mixtures were absorbed onto cells for one hour after which they were removed and fresh media containing the inhibitor or the antibody was added. At 6 HPI for VSV or 48 HPI for MR766, we assessed entry either by observing eGFP expression or by intracellular staining of ZIKV-E with 4G2 antibody (Figure 2.3).
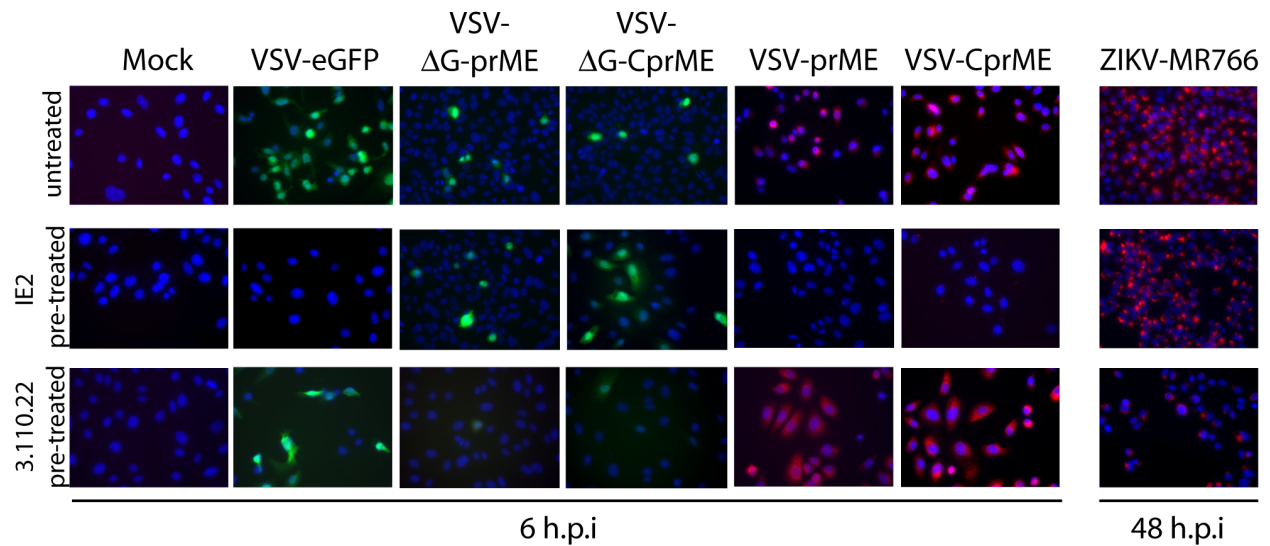
Figure 2.3: ZIKV envelope is functional on VSV particles. Immunofluorescence (red) or eGFP epifluorescence (green) of Vero cells infected with the indicated virus. Where indicated, viruses were pre-incubated for 30 minutes at 34°C with 0.1µg of the VSV G neutralizing monoclonal antibody (IE2) or with 10µM of the dengue virus and Zika virus envelope protein inhibitor 3-110-22. Cells were fixed, permeabilized, and infection was assessed by fluorescence microscopy. Green foci represent infection of cells with viruses expressing eGFP and red foci intracellular detection of ZIKV-E by 4G2 and Alexa Fluor 594. Cells were processed for immunofluorescence microscopy at 6 HPI for VSV and 48 HPI for MR766. Images are representative of 3 independent experiments.

As expected, 3-110-22 significantly reduces MR766 infection. Additionally, incubation of rVSV-ΔG-prME and rVSV-ΔG-CprME with 3-110-22 leads to decreased infectivity, indicating that entry of rVSV-ΔG-prME and rVSV-ΔG-CprME is ZIKV-E-dependent. Furthermore, infection of rVSV-prME and rVSV-CprME is completely abolished by incubation with IE2 antibody but incubation with 3-110-22 does not impact infectivity. This suggests that in cells expressing both VSV-G and ZIKV-E, incorporation of ZIKV-E does not occur readily. Taken together, these results demonstrate that VSV can incorporate ZIKV-E, but it is not an efficient process. It is a surprise to be sure but a welcome one that, despite the differences in ZIKV and VSV assembly, some amount of mature ZIKV-E is incorporated into VSV virions.

The majority of rVSV-ΔG-prME and rVSV-ΔG-CprME particles lack functional E

Given that there are roughly equal levels of VSV proteins in rVSV-ZIKV infected cells (Figure 2.2A) and in cell supernatants (Figure 2.1C), but rVSV-ΔG-prME and rVSV-ΔG-CprME grow to low titers, we hypothesized that the perceived differences between infectious particles and viral proteins in cell supernatants is due to the formation of non-infectious particles. Indeed, VSV is capable of budding and release in the absence of viral glycoproteins [169]. Given that non-infectious particles still contain a viral genome, we used quantitative reverse transcription PCR (RT-qPCR) to estimate the number of viral genomes in infected cell supernatants. We used primers that amplify a region of the VSV-N gene as a means of measuring viral genomes. Our results indicate that there are relatively equal numbers of viral genomes in cell supernatants from each virus infection (Figure 2.4).
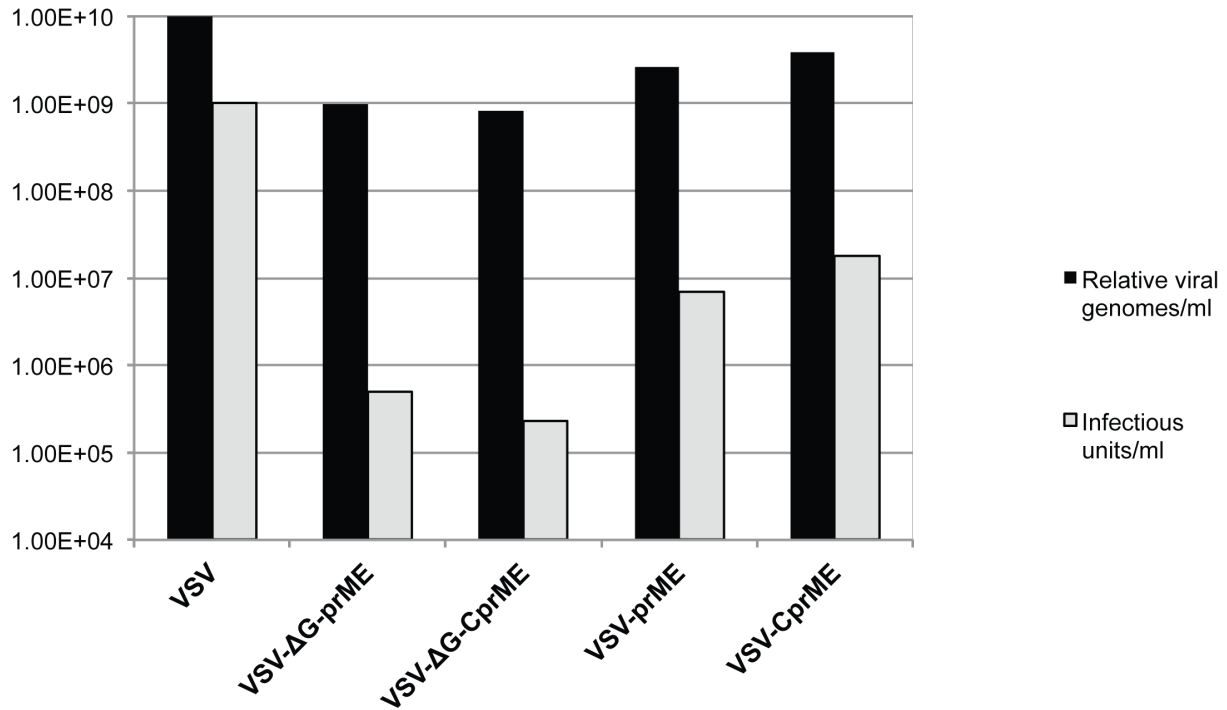
Figure 2.4: Relative infectivity of rVSV-ZIKV. Viral genomes were measured using quantitative RT-PCR. Primers that align to the N gene of VSV were used to compare relative genome levels of the rVSV-ZIKV viruses compared to VSV. Infectious units were measured by plaque assay for VSV-prME and VSV-CprME or by flow cytometry for VSV-ΔG-prME and VSV-ΔG-CprME.

This result is consistent with detection of equal levels of VSV proteins in supernatants of infected cells. However, when we compare the number of genomes to the number of infectious particles, we see significant gaps. For example, VSV particle to infectivity ratios are about 10-50, indicating that there are 10 to 50-fold more viral particles detected in a supernatant compared to infectious particles. For rVSV-ΔG-prME and rVSV-ΔG-CprME, we see genome to infectivity ratios greater than 1000 (1964 and 3590 respectively). The genome to infectivity ratios are decreased for rVSV-prME and rVSV-CprME as would be expected of viruses that retain VSV-G (368 and 211); however, these are still higher than VSV alone indicating that the presence of the ZIKV proteins may have negative consequences for infectious particle generation.

Indeed, rVSV-prME and rVSV-CprME exhibit small plaque morphologies compared to rVSV-eGFP (Figure 2.1A). These results indicate that incorporation of functional ZIKV-E into VSV virions is not favored. This could indicate two possibilities: first, ZIKV-E is incorporated into VSV and is not competent for fusion, while the second possibility is that ZIKV-E is secreted from cells separately than through incorporation into VSV. We set out to test this by assessing viral supernatants for ZIKV virus like particles (VLPs).

ZIKV-E is preferentially exported in VLPs

Since expression of flavivirus prM and E in cells is sufficient to induce the formation of VLPs, we wanted to determine if ZIKV-E detected in viral supernatants incorporated into VSV, VLPs, or is distributed between both. We infected 10 T150 flasks of Vero cells with each virus and pelleted supernatants over a 10% sucrose

cushion. This pelleting step is the stage at which we analyzed supernatants in Figure

2.1. To gain a more nuanced view of the supernatant particle composition, we ran the

supernatant pellets on a 15-45% sucrose gradient. As ZIKV VLPs and VSV have

different protein and lipid compositions, and thus, different buoyant densities, they

migrate differently in the gradient. We then took 500µl fractions from the top (less

dense) to bottom (more dense) of the gradient. These fractions were assessed by

SDS-PAGE followed by Coomassie staining and Western blot with antibodies to ZIKV-E

and VSV-M. For figures 2.5A through 2.5D, Coomassie stained gels are presented on

the top with western blots featured below.

For rVSV-ΔG-prME, ZIKV-E is detected not only in virus fractions, but also in

lighter fractions that do not contain any VSV proteins (Figure 2.5A). These lighter ZIKV-

E-only fractions are indicative of ZIKV VLPs. Comparatively, ZIKV-E from rVSV-ΔG-

CprME supernatants reveals that ZIKV-E is primarily found in VLP fractions, as there is

little ZIKV-E in VSV fractions (Figure 2.5B).
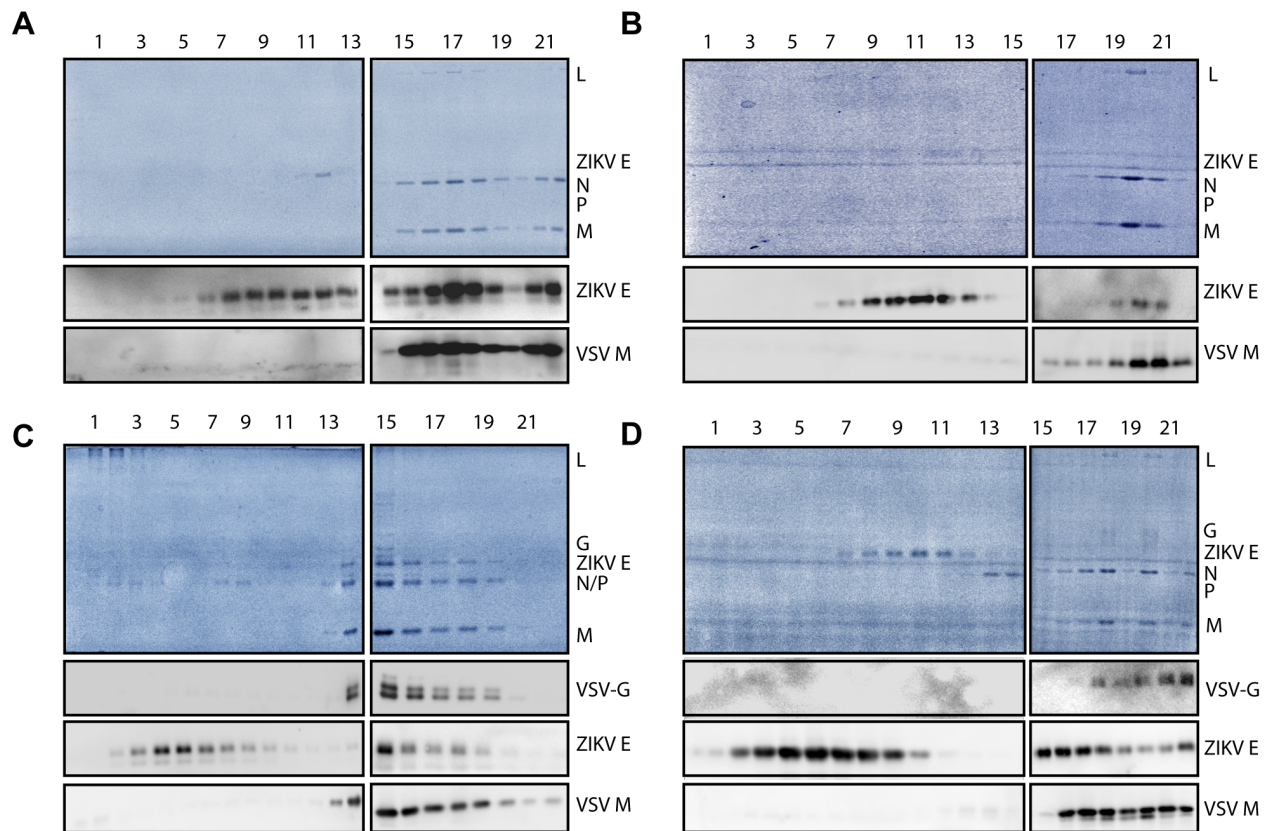
Figure 2.5: ZIKV-E is incorporated into both VSV and VLPs. Cell culture supernatants were collected, and virions and VLPs were recovered by centrifugation, sedimentation through a sucrose cushion, and separation on a 15-45% sucrose gradient. Gradient fractions (0.5 ml) were collected and their protein constituents were analyzed by SDS-PAGE followed by Coomassie staining (blue panels) or western blot (gray scale panels). Fractions are numbered left to right going from least dense (1) to most dense (21). Panel (A) VSV-ΔG-prME, (B) VSV-ΔG-CprME, (C) VSV-prME and (D) VSV-CprME. Blots shown are representative of three independent experiments.

It is surprising that rVSV-ΔG-prME and rVSV-ΔG-CprME exhibit different levels of incorporation of ZIKV-E but grow to similar titers. This could indicate that some ZIKV-E incorporated into rVSV-ΔG-prME as detected by western blot is not competent for membrane fusion. Additionally, as the only difference between prME and CprME is the inclusion of the signal peptide, it seems as though the signal peptide biases ZIKV-E incorporation into VLPs. Without this region, ZIKV-E may not be transported properly through the cell. A similar pattern emerged for rVSV-prME and rVSV-CprME in that ZIKV-E is distributed between VLP and VSV fractions for rVSV-prME and ZIKV-E is primarily detected in VLPs for rVSV-CprME (Figures 2.5C and 2.5D). This further substantiates that inclusion of the signal peptide increases release of ZIKV-E in VLPs. Despite our best attempts, we were unable to generate replication competent rVSV-ΔG expressing ZIKV-E that grew to titers sufficient for vaccination. However, we were intrigued by the observation that these viruses seem to robustly produce ZIKV VLPs. VLPs alone, or the VSV vector that best produces them could serve as vaccine candidates. Therefore, we set out to further characterize VLP production.

rVSV-CprME robustly produces uniform ZIKV VLPs

To directly compare VLP yields between rVSV-prME and rVSV-CprME, we infected the same number of cells at the same MOI with each virus. All subsequent steps through gradient purification and western blotting were performed as in figure 2.5, but VLP fractions from each virus supernatant were run on the same gel. This result shows that infection of cells with rVSV-CprME yields significantly more VLPs than infection with rVSV-prME (Figure 2.6A).
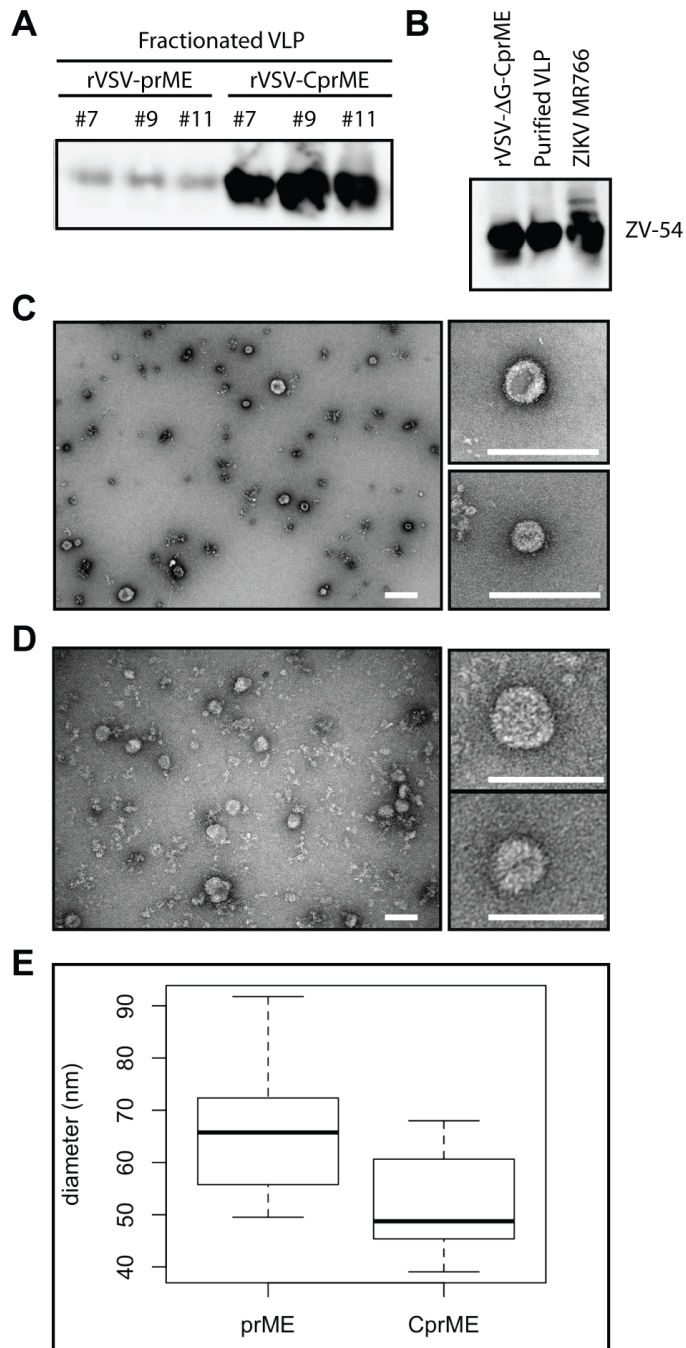
Figure 2.6: Properties of rVSV-produced ZIKV VLPs. (A) VLPs were collected from

cell culture supernatants by precipitation using polyethylene glycol, resuspended and

separated on a 15-45% sucrose gradient and the fractions analyzed by SDS-PAGE and

probed by western blot for ZIKV-E. (B) Detection of the E protein in VSV-ΔG-CprME

virions, VLPs from VSV-CprME and purified ZIKV MR766 virus using the ZIKV

Figure 2.6 (continued):

monoclonal antibody (ZV-54). Transmission electron micrographs of purified VLPs

stained with 0.5% uranyl formate obtained from the VLP fractions of VSV-prME (C) or

VSV-CprME (D). Scale bars represent 100 nm. (E) Box plot of VLP diameters produced

during VSV-prME (prME) and VSV-CprME (CprME) infection. Particle diameters were

measured using ImageJ software.

This result is consistent with the observation that ZIKV-E is primarily detected in

purified supernatant from cells infected with rVSV-CprME but not substantially in rVSV-

prME supernatants (Figure 2.1C). Thus, rVSV-CprME could be used as a viral vector to

express ZIKV VLPs *in vivo* or as a means to produce large quantities of VLPs. Next, we

utilized the ZIKV-specific monoclonal antibody, ZV-54, that recognizes an epitope only

present in mature virions to test whether ZIKV-E in VLPs produced during rVSV-CprME

infection are properly processed [170]. Under non-reducing conditions, we are able to

detect mature ZIKV-E in rVSV-CprME supernatant, purified VLPs, and ZIKV MR766

suggesting that VLPs produced during VSV infection contain properly folded ZIKV-E

(Figure 2.6B).   To confirm that these fractions contained VLPs, and to examine their

morphological properties, we stained purified VLP fractions with 0.5% uranyl formate

and visualized them by transmission electron microscopy. Indeed, VLPs produced from

cells infected with rVSV-prME and rVSV-CprME exhibit the morphological

characteristics of flavivirus VLPs (Figures 2.6C and 2.6D). Namely, they are spherical

in shape and have diameters around the expected size of 50nm. We quantified the

diameters of 10 VLPs from each grid and plotted them using a box plot (Figure 2.6E).

This analysis indicates that VLPs expressed from rVSV-CprME are more uniform in size and cluster closer to the expected 50nm length, while rVSV-prME VLPs exhibit a greater size distribution. As such, rVSV-CprME VLPs represent a better vaccine candidate due to their uniformity and high abundance in supernatants. Therefore, we tested these isolated VLPs as well as rVSV-CprME for their ability to induce a neutralizing antibody response against ZIKV.

rVSV-CprME and VLPs induce a neutralizing antibody response against ZIKV

To assess the immunogenic potential of rVSV-CprME and VLPs produced using this virus, we intramuscularly immunized 5 BALB/c mice per group with either 10^5 pfu rVSV-CprME, 10^5 pfu rVSV-eGFP as a virus control, 20ng of VLPs, or NTE as a vehicle control. All mice appeared healthy following injections. After four weeks, the mice were given a boost with the same material as the initial injection, and one week later, serum was collected (Figure 2.7A).
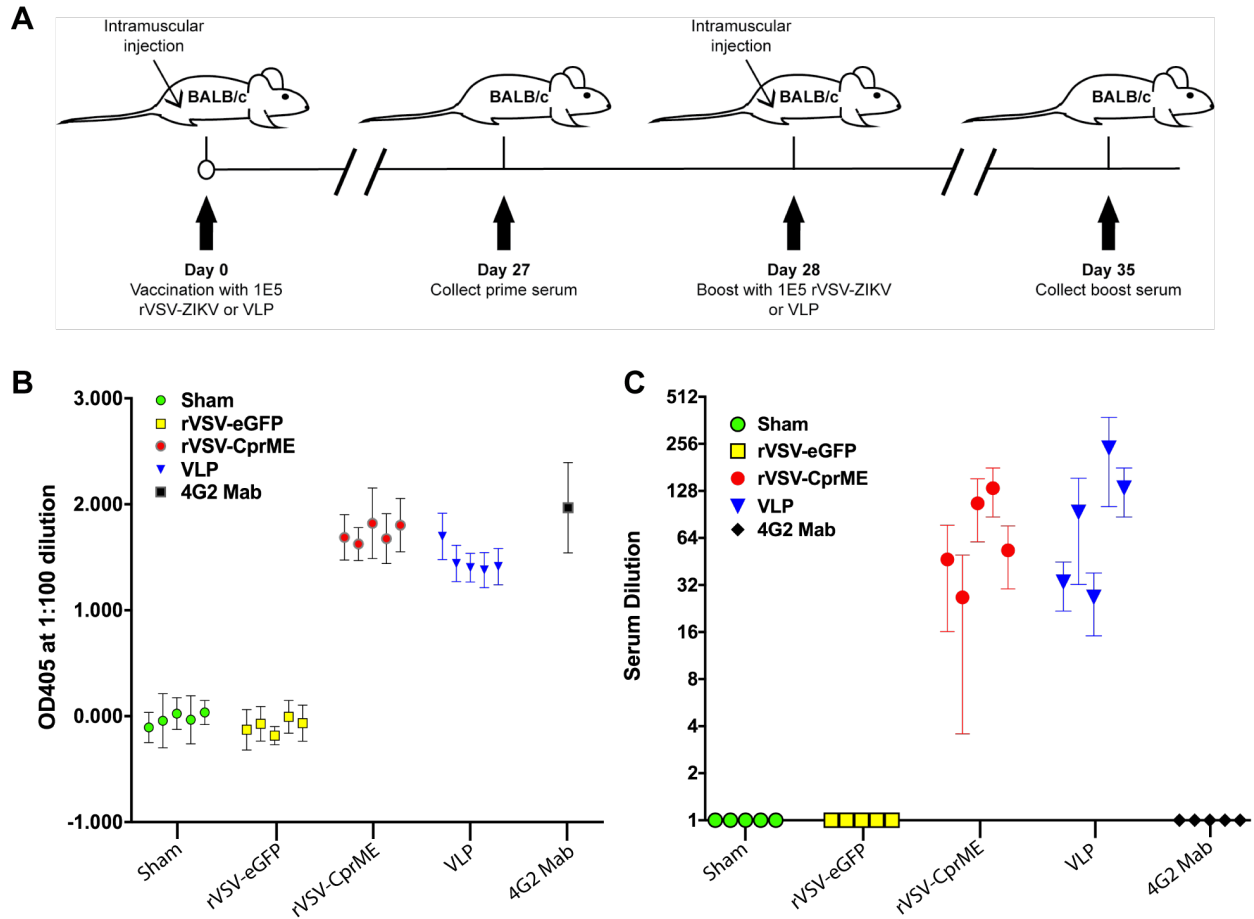
Figure 2.7: VSV-CprME and VLPs are immunogenic in BALB/c mice. (A)

Immunization schedule for BALB/c mice (n=5 /group). (B) Serum ELISA assay for E-

specific antibody present in animals immunized with the indicated immunogens. (C)

Serum microneutralization assay of MR766. Sera from mice inoculated with the

indicated immunogens were assessed for the ability to neutralize the MR766 strain of

Zika (n=3). Sera in 2-fold dilutions were incubated with 100 TCID50 units of MR766 at

37ºC and mixed with cells. Cells were monitored for cytopathic effects and were stained

with crystal violet 8 days following infection. Dilutions of serum that protect Vero cells

from 100 TCID50 units of ZIKV MR766 are shown.

To test sera for antibodies recognizing ZIKV-E, we performed an ELISA using 96-well plates in which purified, recombinant ZIKV-E was bound. Serum samples were diluted 100-fold and added to the ZIKV-E ELISA wells after which the ELISA was performed according to manufacturer specifications. We included 4G2 antibody as a ZIKV-E recognition control. Each animal that received rVSV-CprME or VLPs but not animals receiving rVSV-eGFP or NTE all exhibited high levels or serum antibodies to ZIKV-E (Figure 2.7B). Further, there does not appear to be a significant difference between rVSV-CprME and VLPs for their ability to induce ZIKV-E antibodies.

Following this result, we wanted to test if antibodies generated were capable of neutralizing ZIKV. To do this, we performed a microneutralization assay using ZIKV MR766. MR766 was chosen because the envelope protein is highly similar to that of Suriname strain and because MR766 induces strong cytopathic effects making it a good candidate for visual assays. 100 TCID50 units of MR766 were incubated for two hours at 37ºC with two-fold dilutions of mouse serum. Following this, Vero cells were added to serum-virus mixtures in 96-well tissue culture plates. Cells were monitored for cytopathic effects until visibly cleared wells were observable on day 8. At this time, cells were fixed and stained with crystal violet. The highest serum dilution that provided complete protection against cell killing is plotted in Figure 2.7C. From this, we see that not only do rVSV-CprME and VLPs induce ZIKV-E specific antibodies, but also that they induce antibodies capable of neutralizing ZIKV *in vitro*. Thus, they are likely to be protective *in vivo*. Interestingly, neutralizing antibody titers between rVSV-CprME and VLPs are comparable, indicating that the VSV vector does not lead to an enhanced neutralizing antibody response. Taken together, these results support rVSV-CprME as

a vaccine against ZIKV as well as a means of producing immunogenic VLPs that have potential as ZIKV vaccine candidates.

Discussion

In this chapter, we report the generation and characterization of VSV vectors as vaccines and biological tools for ZIKV. Our vaccine candidates elicit a ZIKV-E neutralizing antibody response, which is reported to be the main correlate of protection against ZIKV [102, 103, 171]. As a result, these vaccine candidates warrant further testing of *in vivo* protection. Ongoing work aims to assess the ability of these vaccine candidates to protect mice against ZIKV challenge, and future studies aim to test the protective effect in non-human primates. While ZIKV VLPs have already been evaluated as vaccine candidates [172], VSV-CprME offers an improved means by which to produce VLPs. At six hours following VSV infection of a cell, over 50% of the translated mRNAs are of viral origin; with the gene in the first position most highly transcribed (William Neidermyer, unpublished data). Thus, VSV infection is likely to yield orders of magnitude greater VLPs produced than standard methods that rely on transfection of DNA plasmids. Furthermore, using VSV instead of DNA to produce ZIKV VLPs would require fewer resources and less time, which would be conducive to rapid vaccine production during outbreaks. To separate VLPs from VSV, supernatants could be run through a 0.1-micron filter, which would exclude VSV and allow VLPs to pass through for further purification.

Our results also indicate that the transmembrane-spanning signal peptide between ZIKV-C and prM proteins is important for ZIKV-E secretion from cells. Indeed,

we detect equal levels of ZIKV-E expression in cells infected with prME-expressing

viruses as we do with CprME expressing viruses, but ZIKV-E levels in supernatants are

significantly higher from CprME virus infected cells. Furthermore, ZIKV-E secreted from

CprME-expressing cells is packaged into more uniform ZIKV VLPs than its prME

counterpart. Given that a signal peptidase cleaves this transmembrane-spanning

region in the ER [173], it is possible that this region leads to retention in the ER or

biases prM-E complexes to bud into the ER, forming VLPs. Without this signal

sequence, it is possible that E is not properly folded and subsequently degraded or it is

mislocalized.

This finding can be extended to the development of other ZIKV vaccines. DNA

and RNA-based vaccines that express prM and E proteins in cells likely also result in

the formation and release of VLPs. To ensure optimal processing and release of ZIKV-

E, other vector-based vaccines for ZIKV should include the C-prM-spanning signal

peptide. Analysis of the sequences used in the initial DNA-based ZIKV vaccine

suggests that this region was omitted, so it stands to reason that that vaccine candidate

could be improved by including the signal peptide [171].

Furthermore, we report generation of recombinant VSV that uses a flavivirus

envelope protein to enter cells. Our results provide understanding for why VSV

incorporation of flavivirus envelope proteins is not favored. Despite high levels of ZIKV-

E expression in rVSV-ΔG-CprME infected cells and supernatants, infectious virus yields

are low and viruses do not propagate. We demonstrate release of bald particles from

rVSV-ΔG-CprME infected cells, suggesting that ZIKV-E is not present at sites of VSV

assembly and budding. This is substantiated by high yields of VLPs in viral

supernatants. Thus, the primary reason why VSV does not readily incorporate flavivirus envelope proteins is likely due to the differences in flavivirus and VSV assembly and budding. Indeed, flavivirus particles assemble in ER-associated membranes and are released through exocytosis. However, VSV assembly and budding occurs at the plasma membrane. Our results suggest that these pathways are mostly separate during rVSV-ZIKV infection. However, we do recover some VSV with functional ZIKV-E on its surface. This may be the result of mislocalization caused by VSV-associated membrane rearrangements or by overexpression of ZIKV-E. Therefore, future work could aim to target flavivirus envelope proteins to the plasma membrane to increase their incorporation into VSV particles.

While recent work by Betancourt *et al.* describes VSV-ZIKV vaccine candidates, there are significant differences between our study and theirs [107]. First, while they show that their rVSV express ZIKV-E in cells, no evidence is provided of ZIKV-E release into cell supernatants. Furthermore, their best vaccine vectors require a mutation in the VSV-M protein that attenuates the virus by abrogating its ability to shut down innate immune protein synthesis. Despite this, ZIKV-E antibody titers induced following immunization with these viruses are low. This suggests that ZIKV-E expressed from their rVSV is not efficiently exported. In contrast, not only are our rVSV highly immunogenic, but also the mechanism by which they induce a neutralizing antibody response as well as the properties of these viruses is well characterized. It would be interesting to directly compare the immunogenicity of the rVSV developed in this chapter to those developed by Betancourt *et al.* as it could shed light into the mechanisms of protection.

Materials and Methods

Cells lines, plasmids and viruses

Syrian golden hamster kidney BSRT7 cells (generously provided by Dr. K. Conzelmann [174] and African green monkey kidney Vero cells (ATCC CCL-81) were maintained at 37°C and 5% CO_2 in Dulbecco's modified Eagle medium (DMEM) (Corning Life Sciences, Tewksbury, MA) supplemented with 10% fetal bovine serum (FBS) (Tissue Culture Biologicals, Tulare, CA).

We commissioned a codon-optimized sequence of ZIKV Suriname strain [162] spanning the *prM* and *E* (prME) coding region from Life Technologies Corporation (Carlsbad, CA). To create CprME, overlap extension PCR was used to introduce the C-terminal transmembrane coding sequence of ZIKV capsid, representing the signal peptide (C) of prME. pVSV-ΔG-prME and pVSV-ΔG-CprME were generated by cloning prME and CprME into pVSV-ΔG-eGFP [175] via MluI and NotI restriction sites as described previously. pVSV-prME and pVSV-CprME were constructed by insertion of prME and CprME into the Xho and MscI sites of pVSV1(+)41 [176].

rVSV-ΔG-prME, rVSV-ΔG-CprME, rVSV-prME, and rVSV-CprME were generated by transfection of the VSV genomic plasmid along with plasmids carrying VSV N, P, L, and, in the case of rVSV-ΔG-prME, rVSV-ΔG-CprME, G as previously described [25]. To increase yield, rVSV-ΔG-prME, rVSV-ΔG-CprME were propagated in BSRT7 cells that were transfected with a plasmid expressing VSV G. After two rounds of trans G complementation, the rVSV-ΔG viruses were passaged on Vero cells to yield virus lacking G on the surface. The rVSV-prME, and rVSV-CprME viruses were

propagated in BSRT7 cells. All virus propagation occurred in DMEM supplemented with

2% FBS and penicillin/streptomycin/kanamycin at 34°C and 5% CO_2. Wild-type ZIKV

(strain: MR766) was kindly provided by the laboratory of Dr. Michael Diamond [177]

and propagated on Vero cells.

Virus titers were determined by plaque assay on Vero cells as previously

described [13]. GFP-expressing fluorescent plaques were imaged using a Typhoon

9400 imager (GE Healthcare, Aurora, OH). For rVSV-ΔG-prME, rVSV-ΔG-CprME, titers

were calculated based on the number of GFP+ cells as determined by flow cytometry

using a BD FACSCalibur flow cytometer (BD Biosciences, San Diego, CA).

Antibodies

Flavivirus E specific Mab 4G2 (hybridoma clone D1-4G2-4-15) was a kind gift

from the laboratory of Dr. Priscilla Yang. Anti ZIKV E ZV-54 Mab [170] was a kind gift

from the laboratory of Dr. Michael Diamond. Anti-VSV-G cytoplasmic tail antibody

(V5507: Sigma Aldrich) and anti-VSV-M antibody 23H12 (a kind gift from Dr. D.S. Lyles

[178]) were used to detect VSV proteins.

Virus growth, purification and fractionation of purified virus

rVSV-ZIKV recombinant viruses were grown in BSRT7 cells as described above.

Viral supernatant collected after 48 HPI was used directly for infectivity experiments or

concentrated by centrifugation and further purified over a 10% sucrose cushion, as

previously described [13]. Sucrose cushion purified virus was used for SDS-PAGE

analysis and Western blots. For further purification and fractionation of the sucrose

cushion purified material, the samples were subject to centrifugation at 25000 rpm for

3h at 4°C in an SW41 rotor over a 15-45% continuous sucrose gradient in NTE buffer.

Following centrifugation, 0.5ml fractions were collected starting from the top of the tube

using a Gilson FC 203B fraction collector (Gilson Inc., Middleton, WI). The fractions

were further analyzed by Western blot and particles were processed for electron

microscopy.

SDS-PAGE analysis and Western blots

Purified virus and virus-like particles were run on a 10% (wt/vol) acrylamide

0.13% (wt/vol) bis-acrylamide gel and transferred to a nitrocellulose membrane or

incubated with Coomassie brilliant blue R (Sigma, St. Louis, MO). Membranes were

blocked with 5% dry milk in 0.1% Tween in PBS (0.1% PBS-Tween) and incubated with

anti-flavivirus E 4G2 antibody (1:500 dilution), anti-VSV-G cytoplasmic tail antibody

(V5507: Sigma Aldrich) (1:5000 dilution), or anti-VSV-M antibody 23H12 (1:5000

dilution) or anti-ZIKV E Mab ZV-54 (1:1000) in 0.1% PBS-Tween buffer, followed by

incubation with goat anti-mouse antibody conjugated HRP (1:5000 dilution).

Chemiluminescence signals were generated using an HRP substrate (Thermo,

Rockford, IL) and detected using an Amersham Imager 600 (GE Healthcare Life

Sciences, Pittsburgh, PA). Gels were incubated in Coomassie reagent overnight, and

after destaining, were imaged using a c200 Gel Imaging System (Azure Biosystems,

Dublin, CA).

Metabolic labeling of proteins in cells

To visualize viral protein synthesis, Vero cells were infected at an MOI of 3 with the VSV-ZIKV recombinants or an MOI of 10 for ZIKV MR766. At 4 HPI for VSV and 40 HPI for ZIKV MR766, cells were starved of methionine and treated with Actinomycin-D (Sigma, St. Louis, MO). At 5 HPI (VSV) and 42 HPI (ZIKV), 50µCi of EasyTag^{TM} L-[^{35}S]-Methionine (PerkinElmer, Waltham, MA) was added to the media. At 6 HPI (VSV) and 44 HPI (MR766), cells were harvested, lysed, and subjected to SDS-PAGE. The gel was fixed for 1h in a gel fix solution (10% acetic acid glacial, 30% methyl alcohol, and 60% H_2O). The fixed gel was dried and exposed overnight on a phosphor screen, which was imaged using a Typhoon FLA 9500 biomolecular imager.

Electron Microscopy (EM)

Purified viruses or VLPs were deposited on carbon coated copper grids and respectively stained with 2% phosphotungstic acid (wt/vol) in H_2O (pH 7.5) or 0.5% uranyl formate (wt/vol) in H_2O. Viruses or VLPs were visualized using a Tecnai G^2 Spirit BioTWIN transmission electron microscope (FEI, Hillsboro, OR).

Purification and fractionation of VLPs and ZIKV

ZIKV MR766 grown in Vero cells and virus-like particles released by rVSV-ZIKV infections were purified by polyethylene glycol (PEG) precipitation. Briefly, ZIKV particles and VLPs were precipitated from cell culture supernatant by PEG (MW = 8000) (0.8 g/10 ml supernatant and NaCl (0.23 g/10 ml supernatant) at 4°C overnight. The precipitate was spun out at 10,000 g for 15min at 4°C in a Sorvall RC 5C Plus centrifuge (Thermofisher, Waltham, MA). The supernatant was aspirated and the pellet

resuspended in 200 µl NTE/50 ml of supernatant. Following this, the samples were centrifuged over a 12.5% sucrose cushion for 100,000 g for 2.5h at 4°C in a TLA 100.2 rotor (Beckmann Coulter, Brea, CA). The pellets were resuspended in 50µl NTE overnight at 4°C. To obtain purified ZIKV or VLPs, samples were centrifuged at 10,000 g for 10min at 4°C and the supernatant collected. Purified ZIKV MR766 and VLPs were subject to sucrose gradient fractionation as described above or used for SDS-PAGE analysis or processed for electron microscopy.

Immunofluorescence and fluorescence microscopy

For immunostaining or fluorescent imaging of eGFP signal in infected Vero cells, cells were infected with rVSV-EGFP, rVSV-ZIKV recombinants or ZIKV MR766 at an MOI of 1. To detect ZIKV-E or VSV-G, cells were fixed with 2% PFA and permeabilized with 0.2% TX-100. ZIKV-E was detected using 4G2 antibody (1:50 dilution), and VSV-G was detected with IE2 antibody (1:50 dilution) at 6 HPI for rVSV-EGFP and rVSV-ZIKV recombinants and at 48 HPI for ZIKV MR766. Secondary detection was performed with goat anti-mouse-Alexa 594-conjugated antibody (1:500 dilution). Cell monolayers were imaged using a 10X objective on a Nikon Eclipse TE 300 microscope (MVI, Avon, MA). Immunostained images showing 4G2 staining and IE2 staining were false colored. Fluorescent imaging of eGFP signal as readout of viral gene expression was imaged directly using the parameters described above.

Antibody neutralization of virus

Purified rVSV-EGFP, rVSV-ZIKV recombinants, and ZIKV MR766 were pre-

incubated with the neutralizing Mab against VSV-G (IE2) for 1h at 34°C at a concentration of 0.1 mg/ml, which inhibits rVSV-EGFP infection of Vero cells. The pre-incubated virus stocks were used to infect Vero cells accordingly, and the cell monolayers were fixed and immunostained for ZIKV-E or directly imaged for eGFP fluorescence as described above.

Immunization of mice

For immunogenicity assays, 6-week-old male BALB/c mice (purchased from Taconic Farms) were immunized twice intramuscularly (IM) in the thigh, on days 0 and 28, with 10^5 plaque-forming units (pfu) of rVSV or 20µg ZIKV VLPs in a volume of 50µl in NTE buffer. On day 27, blood samples were collected from the tail vein of the mice. On day 35, the mice were sacrificed and blood samples were collected by cardiac puncture. Serum was separated by centrifugation at 1500g for 10min at room temperature and then stored at -80ºC. End-point Serum samples obtained from immunized mice were heated at 56°C for 30min to inactivate complement.

Enzyme-linked immunosorbant assay (ELISA)

An ELISA for ZIKV E antibodies was performed using a mouse anti-ZIKV virus envelope protein IgG ELISA kit containing purified recombinant ZIKV E (Alpha Diagnostic International, San Antonio, TX). Heat-inactivated serum samples were diluted 1:100 with the provided dilution buffers. 100µl of diluted serum samples, calibrators, and controls, including the 4G2 antibody, were added to wells containing purified recombinant ZIKV E and the assay performed according to the manufacturer's

instructions. Absorbance was read at 405nm using a SpectraMax M3 microplate reader (Molecular Devices, Sunnyvale, CA).

Serum neutralization assay

The heat-inactivated serum samples from rVSV-EGFP, rVSV-CprME, VLP or sham immunized mice were diluted in DMEM in a 2-fold dilution series from 1:20 to 1:640 in duplicate on a 96-well dish. 100 units TCID50 of ZIKV MR766 virus were incubated with each serum dilution in wells for 30min at 34°C. Following incubation of virus and serum, 10^4 Vero cells/well were added and incubated at 34°C for 6 days. Cell monolayers were fixed with 2% PFA and stained with crystal violet and dilutions of completely cleared wells recorded. Neutralization titers for each animal in the four immunization groups were calculated and the neutralization assays were performed thrice.

Animal studies

Animal housing and experiments were conducted according to protocols approved by the Harvard Medical Area Standing Committee on Animals.

Work contributions

Sayantan Bose generated and characterized CprME-containing viruses. Changhong Zhou performed the animal experiments. Melissanne De Wispelaere provided 4G2 antibody and 3-110-22 inhibitor.

Chapter 3: Evolutionary and functional studies of the human endogenous retrovirus envelope protein, EnvPb1

Abstract

The *envelope* gene of the ERV-Pb1 provirus is conserved in hominoids and old world monkeys, and the human copy is fusogenic. We report that ERV-Pb1 integrated in an ancestor of simians and retains coding potential in all species identified. Expression of conserved primate EnvPb1 finds gorilla and marmoset retain fusogenic *envPb1* ORFs. RNA-seq reads in human RNA tissue samples coupled with detection of protein shows *envPb1* expression in human tissues is not as widespread as previously reported. Out of 6 tissues tested for EnvPb1 protein expression, EnvPb1 protein is only detected in the placenta. Transfection of cells with EnvPb1 leads to release into the supernatant in a form that is likely membrane-bound. We developed a recombinant vesicular stomatitis virus bearing EnvPb1 on the surface as a means to understand how EnvPb1 mediates fusion. We show that entry of rVSV-EnvPb1 requires endosomal acidification. Finally, we performed forward genetic screens using rVSV-EnvPb1 in a mutagenized haploid cell line that indicate heparan sulfate is not important for rVSV-EnvPb1 entry.

Introduction

Endogenous retroviruses (ERVs) are the remnants of ancient retroviral infections that persist throughout the genomes of all vertebrates. When a retrovirus infects a cell, a DNA copy of its genome is integrated into the host's DNA. The first step in generating an ERV is when a retrovirus infects and integrates its DNA into a germline cell. If this germline cell goes on to produce progeny, the provirus (the integrated DNA copy) can be vertically transmitted through Mendelian inheritance and can reach fixation in a population over time.

The contribution of human (H) ERVs to the genome is approximately 8% [113]. Due to their ancient nature, most of the proviral genes have been rendered inactive due to mutational decay. However, in rare cases, HERV open reading frames (ORFs) are co-opted by the host and continue to encode transcribed genes. This is the case for two envelope genes known as *syncytin-1* and *syncytin-2* of the HERV-W and HERV-FRD lineages, respectively. Both of these co-opted envelope genes are conserved in numerous primate genera and encode fusogenic proteins that are expressed in the placenta [133, 179]. These proteins have an essential role in placenta formation and embryo survival [180]. Consequently, other studies have aimed to identify novel endogenous retroviral genes that contain intact ORFs [149]. By using BLAST to compare sequences of known retroviral genes to human genome sequences, a novel ERV was identified through similarity to a zebrafish endogenous retrovirus and was predicted to contain an intact *envelope* (*env*) gene. This provirus was subsequently named HERV-Pb1 [149].

The *envelope* gene is the sole intact gene in the HERV-Pb1 proviral structure, which suggests that it has been selectively maintained. Indeed, out of 17 single-nucleotide polymorphisms (SNPs) within the HERV-Pb1 locus, only one is located within the *env* coding region, and that SNP is synonymous [152]. Additionally, 14 insertion/deletion events have occurred in the HERV-Pb1 locus; the three within *envPb1* do not disrupt the reading frame while the others are all reading frame disrupting in their respective genes [152]. It is unlikely for these events to have occurred by chance, and such data are consistent with selective maintenance of the ORF. Not only is the ORF maintained, but also the human copy of EnvPb1 is capable of mediating cell-cell fusion when expressed in a panel of cell lines [150].

These data suggest that EnvPb1 may play a physiological role related to cell-cell fusion. However, unlike the syncytins, it is reported that *envPb1* is broadly expressed in wide panel of human tissues [150, 152]. Given this, it is difficult to hypothesize where EnvPb1 may be physiologically relevant.

In an effort to better understand the physiological role of EnvPb1, we set out to characterize how EnvPb1 mediates cell-cell fusion. We developed a panel of tools including EnvPb1-specific antisera to test for expression in human tissues and a recombinant vesicular stomatitis virus (VSV) bearing EnvPb1 on its surface to assess cellular factors necessary for EnvPb1 function. Further, we characterized the conservation of EnvPb1 using the currently available primate genomes to show that its initial integration is much older than previously reported. Our results substantiate the idea that EnvPb1 may be a physiologically relevant fusogen in simians.

Results

ERV-Pb1 integrated into a simian ancestor and its *envelope* gene is under

purifying selection

Since ERV-Pb1 was initially identified, many more primate genomes have been

sequenced. While ERV-Pb1 was originally shown to be specific to Old World primates

based on LTR divergence and amplification of *envPb1* from OWM genomes, recent

studies indicate that it may have integrated into a primate ancestor around 58 million

years ago (mya) [126, 152]. Thus, we hypothesized that we should be able to identify

envPb1 in NWM genomes. We searched 19 available primate genomes, including 4

NWM and 10 OWM, using BLAST or the UCSC genome browser with human *envPb1*

as the input query [181, 182]. We recovered intact *envPb1* ORFs from all 19 species

searched. ORFs were translated and aligned using the T-Coffee multiple sequence

alignment (MSA) package [183]. The MSA alignment was annotated to highlight

common functional domains (Figure 3.1). All identified sequences show conservation of

the CXXC and CX_6CC motifs involved in thiol-disulfide exchange for assembly and

function of *gammaretroviral* Envs [184, 185]. In addition, all sequences retain a

predicted furin cleavage site (RXXR) suggesting that they retain the ability to be

processed into SU and TM subunits. Furthermore, all sequences retain a CKS-17-like

immunosuppressive domain (ISD) as defined by Mangeney *et al.*, suggesting potential

for immunosuppression *in vivo* [145].

Figure 3.1: MSA of *envPb1* translations. 20 *envPb1* ORFs, Human (hum), chimpanzee (cpz), bonobo (bon), gorilla (gor), orangutan (ora), gibbon (gib), Angola colobus (col), golden snub-nosed monkey (gsn), proboscis monkey (pro), African green monkey (agm), pig-tailed macaque (ptm), rhesus macaque (rhm), crab-eating macaque (cem), sooty mangaby (sty), drill (drl), baboon (bab), marmoset (mar), owl monkey (owl), white-headed capuchin (cap), and squirrel monkey (sqr), were translated and aligned using the T-Coffee multiple sequence alignment package [183]. The sequence alignment was visualized using the ESPript server – http://espript.ibcp.fr [186]. Amino acid positions that are identical between all species are shown in black boxes. Amino acid positions in which greater than 70% of species are identical are shown in grey boxes, with the most common amino acid in bold text. Amino acid positions less than 70% identical are shown in white boxes. Dots represent gaps. Gammaretroviral functional domains are highlighted. CXXC and CX_6CC are shown in yellow. Predicted furin cleavage site (RXXR) is shown in green. CKS-17-like ISD is shown in red, with orange area representing overlap with CX_6CC. Predicted transmembrane-spanning regions (TMHMM Server - http://www.cbs.dtu.dk/services/TMHMM/), signal peptide and transmembrane anchor are shown in blue [151].

Signal peptide

CXXC

Figure 3.1 (continued):

79

```
             350       360       370       380       390       400       410       420       430
hum  GKNLEFLGDLKGCSELKSFQELTNQSALVHARADVWWYCGGHLLDTLPSNWSGTCALIQLAIPFTLAFQQPEKKKPQRRKTREAPQGSFD
cpz  GKDLEFLGDLKGCSELKSFQELTNQSALVHARADVWWYCGGHLLDTLPSNWSGTCALIQLAIPFTLAFQQPEKKKPQRRKTREALQGSFD
bon  GKDLEFLGDLKGCSELKSFQELTNQSALVHARADVWWYCGGHLLDTLPSNWSGTCALIQLAIPFTLAFQQPEKKKPQRRKTREALQGSFD
gor  GKNLEFLGDLKGCSELKSFQELTNQSALVHARADVWWYCGGHLLDTLPSNWSGTCALIQLAIPFTLAFQQPEKKKPQRRKTREAPQGSFD
ora  GENSEFLGDLKGCSELKSFQELTNQSALVHARADVWWYCGGHLLDTLPSNWSGTCALIQLAIPFTLAFQQPEKKKPQRRKTRGAPQGSFD
gib  GKNLQFLGDLQGCSEIKSFQELTNKSALVHARADVWWYCGRHLLDTLPHNWSGTCALIQLAIPFTLAFQQPKKEKLQHRKTREAPEGSFD
col  GKNLEFLGDLKGCSEVKSFQELTNQSPLVHPRADVWWYCGGPLLATLPNNWSGTCALIQLAIPFTLAFQQREKEKPQHRKTRGAPQGSFD
gsn  GKNLEFLGDLKGCSEVKSFQELTNQSPLVHPRADVWWYCGGPLLATLPSNWSGTCALIQLAIPFTLAFQQREKEKSQHRKTRGAPQGSFD
pro  GKNLEFLGDLKGCSEVKSFQKLTNQSPLVHPRADVWWYCGGPLLATLPSNWSGTCALIQLAIPFTLAFQQREKEKPQHRKTRGAPQGSFD
agm  GKNLEFLGDLKGCGEVKSFQELTNQSPLVHPRADVWWYCGGPLLATLPSNWGGTCALIQLAIPFTLAFQR.EKEKPHHRKTRGAPQGSFD
ptm  GKNLEFLGDLKGCSEVKSFQELTNQSPLIHPRADVWWYCGGPLLTTLPSNWSGTCALIQLAIPFTLAFQQ.RKEKPQHRKTRGAPQGSFD
rhm  GKNLEFLGDLKGCSEVKSFQELTNQSPLIHPRADIVWWYCGGPLLITLPSNWGGTCALIQLAIPFTLAFQQREKEKPQHRKTRGAPQGSFD
cem  GKNLEFLGDLKGCSEVKSFQELTNQSPLIHPRADVWWYCGGPLLATLPSNWGGTCALIQLAIPFTLAFQQREKEKPQHRKTRGAPQGSFD
sty  GKNLEFLGDLKGCSEVKSFQELTNQSPLVHPRADVWWYCGGPLLTTLPSNWGGTCALIQLAIPFTLAFQQREKEKPQHRKPRGAPQGSFD
drl  GKNLEFLGDLKGCSEVKSFQELTNQSPLIHPRADVWWYCGGPLLTTLPSNWGGTCALIQLAIPFTLAFQQREKEKPQHRKPRGAPQGSFD
bab  GKNLEFLGDLKGCSEVKSFQELTNQSPLIHPRADVWWYCGGPLLTTLPSNWGGTCALIQLAIPFTLAFQQREKEKPQHRKPRGAPQGSFD
mar  GENLVFFGDLKGCSEVKPFQELTNQSPLVHPRADVWWYCGGPLLDTLPSNWSGTCALVQLAIPFTLAFHETEKEKLQRRKTRAAPLGSFD
owl  GENLAFFGDLKGCSEVKPFQELTNQSALVHPRADVWWYCGGHLLDTLPSNWSGTCALIQLVIPFTLAFHETEKEKLQRRKTREAPLGSFD
cap  GENLAFFGDLKGCSEVKPFQELTNQSALVHPRADVWWYCGGPLLDTLPSNWSGTCALIQLVIPFTLAFHETEKEKLQHRKTRAAPLGSFD
sqr  GENLAFMGDLKGCSEVKPFQELTNQSALVHARADVWWYCGGPLLDTLPSNWGGTCALVQLVIPFTLAFHETEEEKGQHRKRRETPLGSFD
                                                                                     RXXR
```

```
             440       450       460       470       480       490       500       510       520
hum  SHVYVDEIGVPQGVPDRFKARDPVAAGFESLFPMVAINKNVAWINYIYNQQRFINYTRDAIQGIAEQLGPTSQMAWENRMALDMILAEK
cpz  SHVYLDGIGVPQGVPDRFKARDPVAAGFESLFPMVAINKNVAWINYIYYNQQRFINYTRDAIQGIAEQLGPTSQMAWENRMALDMILAEK
bon  SHVYLDGIGVPQGVPDRFKARDPVAAGFESLFPMVAINKNVAWINYIYYNQQRFINYTRDAIQGIAEQLGPTSQMAWENRMALDMILAEK
gor  SHVYLDGIGVPQGVPDRFKARDQIAAGFESLFPIVAINKNVAWINYIYYNQQRFINYTRDAIQGIAEQLGPTSQMAWENRMALDMILAEK
ora  SHVYLDAVGVPQGVPDSFKARDPVAAGFESLLPMVAINKNVGWINYIYYNQQRFINYTRDAIQGIAEQLGPTSQMAWENRMALDMILAEK
gib  SSIYIDAIGVPQGVPDRFKARNPIAAGFESLFPTVAVNKNVGWINYIYYNQQRFINYTRDAIRGIAEQLGPTSQMAWENRMALDMILAEK
col  SQVYIDAIGVPRGVPDRFKARNSIAAGFESLIPIVAINKNVAWINYIYYNQQRFINYTRDAIKGIAEQLGPTSQMAWENRMALDMILAEK
gsn  SQVYIDAIGVPRGVPDRFKARNSIAAGFESLLPVVAINKNVAWINYIYYNQQRFINYTRDAIKGIAEQLGPTSQMAWENRMALDMILAEK
pro  SQVYIDAIGVPRGVPDRFKARNSIAAGFESLLPVVAINKNVAWINYIYYNQQRFINYTRDAIKGIAEQLGPTSQMAWENRMALDMILAEK
agm  SHVYIDAIGVPRGVPDRFKARNQVAAGFESLFPVVAINKNVAWINYIYYNQQRFINYTRDAIKGIAEQLGPTSQMAWENRMVLDMVLAEK
ptm  SHVYIDAIGVPRGVPDRFKARNPVAAGFESLFPVVAINKNVAWINYIYYNQQRFINYTRDAIKGIAEQLGPTSQMAWENRMVLDMVLAEK
rhm  SRVYIDAIGVPRGVPDRFKARNPIAAGFESLLPMVAINKNVAWINYIYYNQQRFINYTRDAIKGIAEQLGPTSQMAWENRMVLDMVLAEK
cem  SRVYIDAIGVPRGVPDRFKARNPIAAGFESLLPVVAINKNVAWINYIYYNQQRFINYTRDAIKGIAEQLGPTSQMAWENRMVLDMVLAEK
sty  SRVYIDAIGVPRGVPDRFKARNPIAAGFESLLPMVAINKNVAWINYIYYNQQRFINYTRDAIKGIAEQLGPTSQMAWENRMVLDMVLAEK
drl  SRVYIDAIGVPRGVPDRFKARNQIAAGFESLLPMVAINKNVAWINYIYYNQQRFINYTRDAIKGIAEQLGPTSQMAWENRMVLDMVLAEK
bab  SRVYIDAIGVPRGVPDRFKARNPIAAGFESLLPIVAINKNVAWINYIYYNQQRFINYTRDAIKGIAEQLGPTSQMAWENRMVLDMVLAEK
mar  SHVYIDAIGVPRGVPDRFKARNPIAAGFESIFPWVAINKNVDWINYIYYNQQRFINYTDAIKGIAEQLGPTSQMAWENRMALDMILAEK
owl  SQVYIDAIGVPRGVPDRFKARNPIAAGFESLFPWVAINKNVDWINYIYYNQQRFINYTDAIKGIAEQLGPTSQMAWENRMALDMILAEK
cap  SHVYIDSIGVPRGVPDRFKARNQIAAGFESLLPWISINKNVDWINYIYYNQQRFINYTDAIKGIAEQLAPTSQMAWENRMALDMILAEK
sqr  SEVYIDAIGVPRGVPDRFKARNQIAAGFESLFPVSINKNVDWINYIYYNQQRFINYTDAIKGIAEQLAPTSQMAWENRMVLDMILAEK
                                                                                     ISD
```

```
             530       540       550       560       570       580       590       600       610
hum  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGITDPFTGWLGQWFGKWKGLMASIVTSLAIAIAVLILVGCCIMPCIR
cpz  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGITDPFTGWLGQWFGRWKGLMASIVTSLAIAIAVLILVGCCIIPCIR
bon  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGITDPFTGWLGQWFGRWKGLMASIVTSLAIAIAVLILVGCCIIPCIR
gor  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGITDPFTGWLGQWFGKWKGLMASIVTSLAIAIAVLILVGCCIIPCIR
ora  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGITDPFTGWLGKWFGKWKGLMASIVTSLAIAIAVLILVGCCIMPCIR
gib  GGVCVMIGTQCCTYIPNNTAPDGTITKALRGLTSLSDELATNSGISDPFTGWLGWFGKWKGLMASIVTSLAIAIAVLILVGCCIMPCIR
col  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSNELATNSGISDPFTRWLGTWFGKWKGLMSIVTSLTITIAVVVVGCCIIPCIR
gsn  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGISDPFTGWLGTWFGQWKGLVASIVTSLAITIAVLMVVGCCIIPCIR
pro  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGISDPFTGWLGTWFGKWKGLMASIVTSVAITIAVLTVVGCCIIPCIR
agm  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGISDPFTGWLGTWFGKWKGLMASIVTSLAIAIAVLLVVGCCIIPCIR
ptm  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGISDPFTGWLGTWFGKWKGLMASIVTSLAIAIAVLLVVGCCIIPCIR
rhm  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGISDPFTGWLGTWFGKWKGLMASIVTSLAIAIAVLMVVGCCIIPCIR
cem  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGISDPFTGWLGTWFGKWKGLMASIVTSLAIAIAVLMVVGCCIIPCIR
sty  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGISDPFTGWLGTWFGKWKGLMASIVTSLAIAIAVLMVVGCCIIPCIR
drl  GGVCVMIGTQCCTYIPNNTAPDGTITKALQGLTSLSDELATNSGISDPFTGWLGTWFGKWKGLMASIVTSLAIAIAVLMVVGCCIIPCIR
bab  GGVCVMIGAQCCTFIPNNTAPNGTITKALQALTSLSNELATNSGIDDPITGWLGKWFGKWKGFFASILSSLAVAMAVLILVGCCIIPCIR
mar  GGVCVMIGAQCCTFIPNNTAPNGTITKALQALTSLSNELATNSGIDDPITGWLGKWFGKWKGFFASILSSLAVAMAVLILVGCCIIPCIR
owl  GGVCVMIGTQCCTFIPNNTAPDGTITKALQALTSLSNELATNSGINDPITGWLGNLFGKWKGFFVSILSSFAAAMVVFILVGCCIIPCIR
cap  GGVCVMIGAQCCTFIPNNTAPNGTITKALHALTSLSNELATNSGINDPFTGWLEKWFGKWKGFFASILSSLAVAMAVLVFVGCCIIPCIR
sqr  GGVCVMIGTQCCTFIPNNTAPDGTITKALQALTALSNELATNSGINDPITGWLEKWFGKWKGFASILSSLAVIAVLILVGCCIIPCIR
     CX6CC                                                           Transmembrane
```

```
             620       630       640       650       660
hum  GLVQRLIETASNKTFPSSSQSYSNKFFPVNEHEI........RIILDRFKAEHV
cpz  GLVQRLIETASNKTFPSSSQSYSNKFFPVNEHEI........RVILDRFKAEHV
bon  GLVQRLIETASNKTFPSSSQSYSNKFFPVNEHEI........RVILDRFKAEHV
gor  GLVQRLIETAVNKTFPSSSQSYSNKFFPVSEHEI........RVILDRFKAEHV
ora  GLVQRLIETAFNKTFPSYSQSYSNKFFPVNEHEI........RVILDRFKAQHV
gib  GLVQRLIETAIHKNFPSSSQSYSNKFFPLNDHQI........QIMLDRFKE.NV
col  GLVQRLIGTAINKTFSSSSKSYSNKSFPLKDQQI........RAMLDKFKEGNV
gsn  GLVQRLIGTAINQTFSSSSKSCSNKLFPLKDQRI........RAMLDKFKEGNV
pro  GLVQRLIGTAINQTFSSSSKSYSNKLFPLKDQQI........RAMLEKFKEGNV
agm  GLVQRLTETAINRTFSSSSKSYSNKLFPLKDQQI........QAMLDKFKEGNV
ptm  GLVQRLTETAINRTFSSSSKSYSNKLFPLKDQQI........QAMLDKFKEGNV
rhm  GLVQRLTETAINRTFSSSSKSYSNKLFPLKDQQI........QAMLDKFKEGNV
cem  GLVQRLTETAINRTFSSSSKSYSNKLFPLKDQQI........QAMLDKFKEGNV
sty  GLVQRLTETAINRTFSSSSKSCSNKLFPLKDQQI........QAILDKFKEGNV
drl  GLVQRLTETAINRTFSSSSKSYGNKLFPLKDQQI........QAMLDKFKEGNV
bab  GLVQRLTETAINRTFSSSSKSYSNKLFPLKDQQI........QAMLDKFKEGNV
mar  GLIQRLVEAAVSKTFPSSSFSYTDKFFHMKELQR........RVMLDKFKEENV
owl  GLIQRLVETAVSKTFPSSSKSYTDKFFRMKELQR........RVMLDKFKEENV
cap  GLIQRLVETAVSKSFPTSSKSYTDKLFVMKELQR........RAMLDKFKEGNV
sqr  GLIQRLVETAVSKSFPSSSKSYTGKSFHMKELQRHMKELQRRVMFDKFKEENV
```

Figure 3.1 (continued):

To test if the *envPb1* ORFs descended from a single integration event in their last common ancestor, we tested if the ORF and the ERV-Pb1 provirus are located in the same genetic locus in all species. We analyzed syntenic loci of hominoid (*Pan troglodytes,* chimp), OWM (*Macaca fascicularis,* macaque), and NWM (*Callithrix jacchus,* marmoset) genomes using the MultiPipMaker alignment tool. Our analysis reveals conservation of the syntenic loci in these species as all *envPb1* ORFs reside within an intron of the *rin3* gene (Figure 3.2A). ERV-Pb was not detected in the *rin3* locus of the tarsier genome (not shown), which suggests that integration occurred after the divergence of tarsiers and simians (NWM, OWM, and hominoids) but before the divergence of NWM and old world apes (OWM and hominoids). This is indicative of ERV-Pb integration occurring between 40 and 60 mya, which is in line with the most recent computational estimate of ~58 mya [126, 187-189].

Next, the nucleotide sequences of all intact ORFs were aligned using the MAFFT alignment algorithm on the GUIDANCE server [190, 191]. From the alignments, we generated a maximum-likelihood (ML) phylogenetic tree using MEGA7 software [120, 192]. As expected, the phylogenetic tree forms 3 distinct clades that respectively include hominoids, OWM, and NWM (Figure 3.2B). The OWM clade forms two distinct branches likely due to Angola colobus, golden snub-nosed monkey, and proboscis monkey containing a 30-nucleotide insertion within the predicted SU domain of EnvPb1. Despite nucleotide differences between the EnvPb1 sequences, none of them contain any internal stop codons. Furthermore, while there are some species-specific insertions or deletions, they all preserve the ORF, which suggests a beneficial function to the host.

A

envPb1 ORF

Chimp
Macaque
Marmoset

1   2k   4k   6k   8k   10k   12k   13998

——— *rin3* intron ———

B

- Drill 83
- Baboon 98
- Sooty Mangaby 61
- Crab eating Mac 98
- Rhesus 74
- Pig tailed Macaque
- AGM 100
- Angola colobus
- Golden Snub Nosed Monkey 87
- Proboscis Monkey 99

(100 / 100) — OWM

- Gibbon 99
- Orangutan 100
- Gorilla 99
- Human 86
- Bonobo 99
- Chimp — Hom.

- Marmoset 100
- Owl Monkey 46
- White headed Capuchin
- Squirrel Monkey 86 — NWM

0.020

C

		OWM Angola colobus	Snub-nosed	Proboscis	Drill	Sooty Mangaby	Crab-eating Mac	Rhesus	Baboon	Pig-tailed Mac	AGM	Hom. Bonobo	Chimp	Human	Gorilla	Orangutan	Gibbon	NWM Capuchin	Owl M.	Marmoset	Squirrel M.
OWM	Angola colobus		0.92	1.13	1.19	1.09	1.18	1.02	1.44	0.94	1.08	1.16	1.07	1.12	1.13	1.15	1.11	0.67	0.73	0.71	0.75
	Snub-nosed	97.20		0.38	0.85	0.74	0.84	0.78	0.98	0.78	0.95	0.92	0.91	0.95	0.95	0.92	0.96	0.60	0.70	0.70	0.74
	Proboscis	97.40	98.45		0.93	0.86	0.90	0.84	1.01	0.85	0.94	1.04	0.96	1.00	1.00	0.99	1.03	0.64	0.72	0.70	0.75
	Drill	96.45	96.35	96.45		0.46	0.90	0.59	1.99	1.04	0.97	0.83	0.78	0.85	0.85	0.84	0.84	0.60	0.70	0.66	0.70
	Sooty Mangaby	96.55	96.55	96.55	99.20		0.96	0.58	1.62	1.10	1.06	0.80	0.75	0.81	0.83	0.81	0.88	0.61	0.67	0.64	0.69
	Crab-eating Mac	96.85	96.75	96.95	99.00	99.20		0.56	2.22	2.84	1.23	0.90	0.84	0.89	0.88	0.88	0.88	0.61	0.69	0.66	0.71
	Rhesus	96.70	96.45	96.60	98.85	99.05	99.65		1.28	1.29	1.03	0.85	0.79	0.84	0.85	0.87	0.83	0.59	0.67	0.64	0.69
	Baboon	96.25	96.15	96.45	99.05	99.00	98.90	98.75		1.92	1.69	0.93	0.87	0.95	0.95	0.94	0.90	0.64	0.70	0.67	0.73
	Pig-tailed Mac	96.65	96.40	96.50	98.35	98.55	99.05	98.80	98.25		1.17	0.89	0.83	0.85	0.85	0.87	0.89	0.58	0.67	0.64	0.69
	AGM	96.45	96.65	96.65	97.95	98.05	98.45	98.10	97.65	98.00		0.89	0.84	0.84	0.86	0.87	0.91	0.59	0.70	0.66	0.72
Hom.	Bonobo	92.44	92.59	92.79	92.54	92.64	92.84	92.59	92.24	92.39	92.49		0.07	0.60	1.00	0.70	1.40	0.74	0.78	0.73	0.82
	Chimp	92.44	92.59	92.79	92.54	92.64	92.84	92.59	92.39	92.39	92.49	99.65		0.52	0.88	0.58	1.40	0.74	0.79	0.73	0.80
	Human	92.54	92.69	92.89	92.54	92.64	92.79	92.54	92.24	92.34	92.59	99.05	99.00		1.18	0.57	1.41	0.73	0.77	0.74	0.80
	Gorilla	92.94	92.99	93.19	92.94	92.94	93.09	92.84	92.64	92.64	92.89	99.00	99.05	99.05		0.73	1.49	0.70	0.75	0.72	0.79
	Orangutan	92.54	92.79	92.89	92.54	92.64	92.79	92.64	92.24	92.24	92.49	97.59	97.54	97.69	97.74		1.59	0.72	0.75	0.73	0.82
	Gibbon	91.84	92.09	92.24	91.94	91.94	92.19	91.94	91.69	91.69	91.79	94.89	94.94	95.24	95.29	95.09		0.75	0.85	0.79	0.86
NWM	Capuchin	87.54	87.39	87.39	87.14	87.19	87.39	87.14	86.74	87.14	87.19	87.74	87.74	87.69	87.99	87.54	86.64		0.61	0.42	0.75
	Owl M.	87.64	87.69	87.84	87.39	87.34	87.54	87.29	87.09	87.24	87.39	88.59	88.64	88.54	88.74	88.34	87.44	95.00		0.66	1.13
	Marmoset	86.59	86.89	86.84	86.39	86.34	86.69	86.44	86.14	86.49	86.54	87.44	87.44	87.34	87.54	87.19	86.29	93.49	94.24		0.78
	Squirrel M.	86.29	86.39	86.14	85.84	85.79	86.04	85.84	85.59	85.79	85.99	86.79	86.94	86.84	87.14	86.44	85.94	93.54	93.39	91.69	

Pairwise % identity (nucleotide) <90 | 90-95 | >95

Pairwise dN/dS <0.9 | 0.9-1.1 | >1.1

Figure 3.2: Conservation and purifying selection of *envPb1*. (A) Syntenic

82

Figure 3.2 (continued):

conservation of the *envPb1* ORF in simians. The genetic locus of *envPb1* was recovered from representative hominoid, OWM, and NWM species using the UCSC genome browser. Conservation of the syntenic locus was assessed using the MultiPipMaker alignment tool [193]. Regions with significant homology to the human sequence as defined by BLASTZ [194] are shown in green, and highly conserved regions (70% identity and no gaps) are shown in red boxes. (B) Maximum-likelihood phylogenetic tree of *envPb1* in all available simian genomes. *EnvPb1* ORFs were recovered from the shown species through the UCSC genome browser or NCBI sequence deposits. Nucleotide sequences were aligned using MAFFT, and the evolutionary history was inferred using the Maximum Likelihood method based on the Tamura-Nei model in MEGA7. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Percentage bootstrap values obtained from 1,000 replicates are shown at the nodes. (C) Purifying selection in simians. Pairwise nucleotide identity is shown in the lower triangle and pairwise d_N/d_S values are shown in the top triangle. Values are color coded based on range.

We next wanted to test whether *envPb1* is under purifying selection, as would be expected for a captured retroviral gene that is playing a beneficial role to the host. This is measured by comparing the number of non-synonymous mutations (d_N) to the number of synonymous mutations (d_S) between species. Genes under purifying selection, such as those conserved for important cellular processes, exhibit a $d_N/d_S < 1$. Conversely, genes under positive selection, such as viral restriction factors that are

adaptively evolving, exhibit a $d_N/d_S > 1$. $d_N/d_S = 1$ indicates neutral evolution. For example, other physiologically relevant captured retroviral envelopes have mean d_N/d_S values that range from 0.37 (*syncytin-2*) to 0.88 (*syncytin-1*) [136, 195]. We measured the mean d_N/d_S for *envPb1* using the CodeML package in the phylogenetic analysis by maximum likelihood (PAML) source code software [196]. The mean d_N/d_S for *envPb1* is 0.838, which is similar to that of *syncytin-1*. This differs from the previous report by Aagaard *et al.* that indicated that the d_N/d_S for *envPb1* did not differ significantly from 1 [152]. This is likely due to the fact that their analysis did not include enough representative sequences, as accurate d_N/d_S values are dependent on sufficient species representation [197]. We next computed pairwise d_N/d_S values between primate species. Most pairwise d_N/d_S values fall near or below the mean value of 0.838 (Figure 3.2C). In some cases we see values above one, but the majority of these occur when comparing species that have 98% or greater nucleotide identity, and thus the number of substitutions may not be high enough to provide accurate estimates of selection. However, when gibbon is compared to other hominoids, there appears to be hallmarks of positive selection ($d_N/d_S > 1.4$), which may indicate unique evolutionary adaptations of EnvPb1 in gibbons. NWMs exhibit the lowest pairwise d_N/d_S values, which provides evidence for purifying selection in NWM and accounts for why previous estimates of *envPb1* d_N/d_S, which lacked NWM sequences did not fall below 1.

EnvPb1-mediated fusion is conserved across multiple species

Since human EnvPb1 is fusogenic *ex vivo,* and other identified *envPb1* sequences have conserved ORFs and exhibit hallmarks of purifying selection, we

wanted to test if fusogenicity is a conserved feature of EnvPb1. Given that the major

functional domains necessary for gammaretroviral Env fusion (i.e. CXXC, CX_6CC, furin

cleavage site, and TM domain) are completely conserved between primate EnvPb1

sequences, we predicted that all EnvPb1 would be fusogenic. To test this, we

chemically synthesized *envPb1* from 3 representative primates: gorilla, baboon, and

marmoset. We cloned these gene fragments into an expression vector and confirmed

their identity by sequencing. EnvPb1-encoding plasmids were transfected into Vero

cells, and syncytia formation was assayed at 24 hours post transfection (HPT) – a time

when large, multinucleated cells are prevalent after expression of human EnvPb1. Our

results indicate that both gorilla and marmoset EnvPb1 are fusogenic in Vero cells but

baboon EnvPb1 is not (Figure 3.3A).

A

Mock

Human

Gorilla

Baboon

Marmoset

B

Mock | Human | Gorilla | Baboon | Marmoset

kDa

135

100

75

58

46

EnvPb1-SU/TM
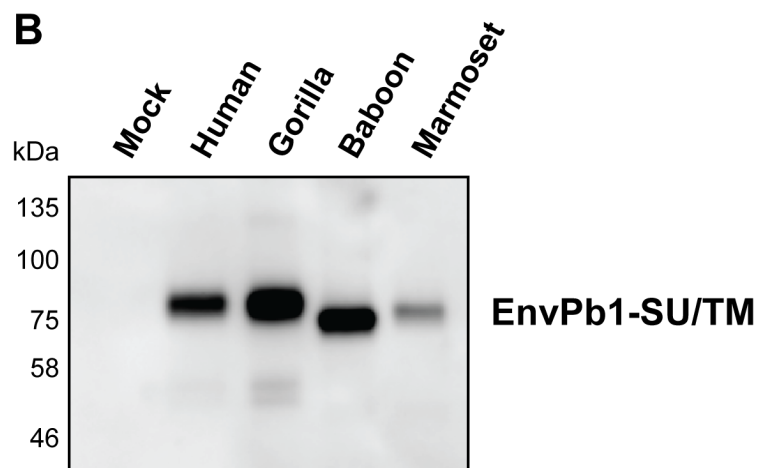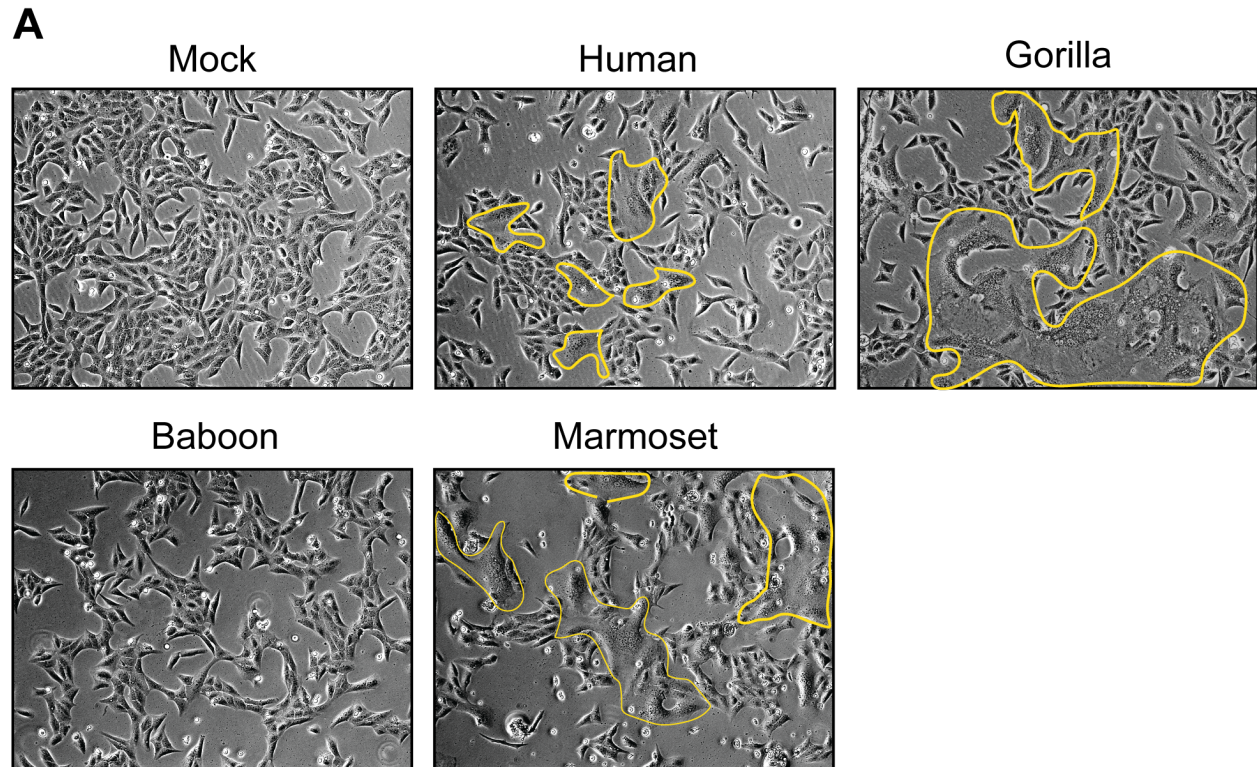
Figure 3.3: Fusogenicity of primate EnvPb1. (A) Select primate EnvPb1 are fusogenic in Vero cells. Vero cells in a 12-well dish were transfected with 1µg of plasmid encoding the indicated primate EnvPb1 sequences. Cells were imaged by phase contrast microscopy 24 HPT to assess syncytia formation. Yellow lines indicate syncytia. Images are representative of three independent experiments.

Figure 3.3 (continued):

(B) Expression of primate EnvPb1 in Vero cells. After assessing syncytia formation, Vero cells were lysed and cellular proteins were collected. SDS-PAGE and western blotting with EnvPb1 antisera at a 1:1000 dilution were used to assess protein expression.

To test for EnvPb1 protein expression in cells, we generated EnvPb1 antisera by inoculating a rabbit with a purified predicted receptor-binding domain (RBD) of human EnvPb1 (amino acids (AA) 51-208). We analyzed cell lysates by SDS-PAGE and western blot with the rabbit anti-EnvPb1 antisera (Figure 3.3B). We detect expression at the predicted molecular weight of the unprocessed precursor (SU/TM) in all samples. Baboon EnvPb1 migrates slightly lower than human and gorilla, which was unexpected given that it contains 666 AA compared to 665 for human and gorilla. This could indicate differences in glycosylation in baboon EnvPb1. Additionally, marmoset EnvPb1 migrates slightly lower than human and gorilla, but this is expected given that it contains 5 fewer AA (660). Given that gorilla and human *envPb1* share 98% amino acid identity, it is not surprising that gorilla EnvPb1 is fusogenic in Vero cells. However, the fact that marmoset EnvPb1 is fusogenic despite sharing only 77.2% amino acid identity with human EnvPb1 suggests that fusion is an evolutionarily selected feature in the capture of EnvPb1 in primate genomes. While baboon EnvPb1 does not fuse Vero cells, it should not be concluded that it is not fusogenic. It is possible that it still retains the ability to fuse cells but its receptor usage is different than that of human EnvPb1.

Another possibility is that its receptor is not expressed or is downregulated in Vero cells. Determining fusogenicity of these EnvPb1 in other cell lines will assess the possibility of differences in receptor usage while assessing more primate EnvPb1 sequences for fusogenicity will determine if conservation of fusogenicity is lineage-specific.

RNA-seq analysis of EnvPb1 expression in human tissues

Previous reports identified *envPb1* transcripts in total RNA samples from human tissues by quantitative reverse transcription polymerase chain reaction (RT-qPCR) [150, 152]. However, given the pitfalls of using RT-qPCR to quantify gene expression in human tissues, we wanted to confirm and expand upon these results by assessing publicly available RNA-sequencing (seq) datasets [198]. We extracted read info from bioproject PRJNA280600 in the NCBI sequence read archive because it contained high-quality total RNA from 20 different human tissues [199]. Then, the reads were aligned to human genome build hg19 that we annotated to contain gene and mRNA annotations for *envPb1*. Following this, aligned reads were mapped to mRNA, and reads per kilobase of transcript per million mapped reads (RPKM) values were calculated for all human genes. We plotted RPKM values for *envPb1* and *GAPDH* (Figure 3.4A). In line with previous reports, our analysis indicates that *envPb1* is widely expressed in human tissues, with the highest levels of expression occurring in the thymus, spleen, and lung [150, 152].
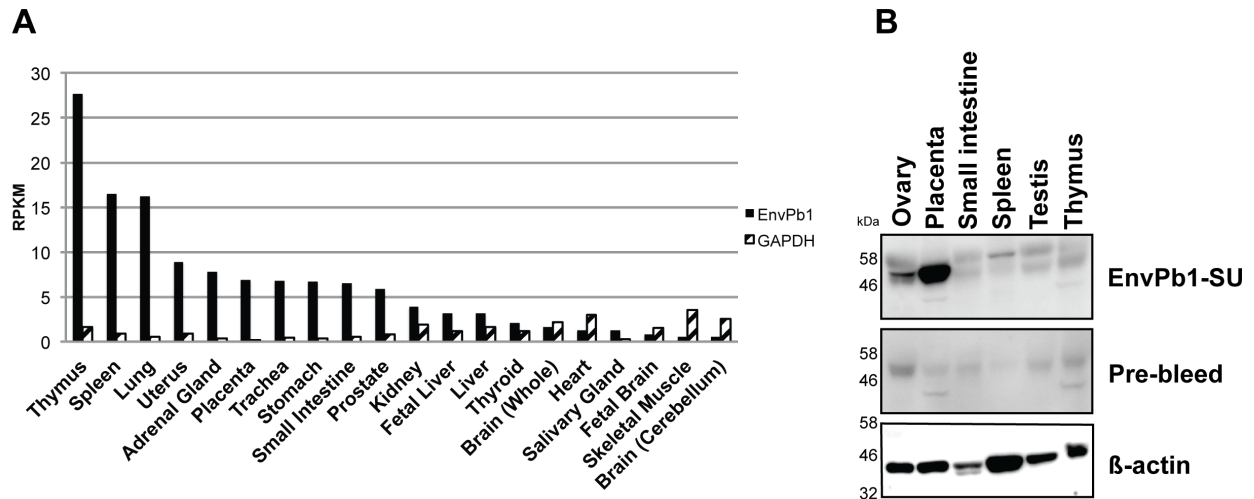
Figure 3.4: EnvPb1 tissue expression. (A) RNA-seq analysis of total RNA from human tissue samples. Reads from a previously published data set [200] were extracted from the NCBI sequence read archive bioproject PRJNA280600. Reads were aligned to human genome build hg19 that was edited to contain an annotation for the *envPb1* ORF. Aligned reads were mapped to mRNA and RPKM was calculated. RPKM is shown for *envPb1 and GAPDH*. (B) EnvPb1 protein expression is placenta-specific. Human protein medleys from ovary, placenta, small intestine, spleen, testis, and thymus tissue were purchased from Clontech. 37.5µg of total protein for each tissue sample was run on a 4-12% SDS-PAGE polyacrylamide gel and assessed for EnvPb1 or ß-actin protein expression by western blot with EnvPb1 antisera or ß-actin antibody. A rabbit pre-bleed serum was included as a non-specific control for EnvPb1 detection.

EnvPb1 protein is detected in human placenta tissue

We next wanted to assess EnvPb1 protein expression in human tissues in which envPb1 transcripts are reported. We chose to analyze thymus, spleen, placenta, ovary, testis, and small intestine based on our RNA-seq analysis as well as the previously published RT-qPCR data [150, 152]. SDS-solubilized protein medleys were purchased from Clontech (Paolo Alto, CA) and boiled with reducing agents. Equal total protein amounts from each tissue were run on a 4-12% SDS polyacrylamide gel, transferred to a nitrocellulose membrane, and assessed for EnvPb1 and ß-actin by western blot. Despite the highest levels of *envPb1* transcripts occurring in the thymus and spleen, the only tissue in which we detect EnvPb1 protein is the placenta, although there also appears to be a faint band in ovary tissue (Figure 3.4B). Indeed, this result does not correlate with the RNA-based tissue expression data. We therefore wanted to test whether the reads identified by our RNA-seq analysis are specific to translation competent EnvPb1 transcripts.

As it has been reported that the ERV-Pb1 5' LTR is capable of driving transcription of the EnvPb1 gene, we wanted to test if we could detect reads corresponding to splice junctions between the ERV-Pb1 5' LTR and *env* ORF in the RNA-seq human tissue profiles [154]. We analyzed reads from the thymus data set using HISAT2 [201], and no splice sites were observed within the EnvPb1 ORF that mapped to the 5'LTR or elsewhere that could serve as a transcription start site (Figure 3.5A).

Figure 3.5: Visualization of mapped reads from thymus total RNA. (A) Splice

junction mapping near the ERV-Pb1 locus. HISAT2 was used to identify reads mapping

to splice sites as previously described [201]. Mapped reads (red vertical lines) and

splice junctions (red parabolic lines) are presented above genes. Blue rectangles

represent exons, and blue lines represent intronic regions. The ERV-Pb1 locus is

annotated such that the left rectangle represents the ORF and the right rectangle

represents the 5' LTR. Other lines and rectangles correspond to *RIN3*. (B)

Visualization of reads near the ERV-Pb1 locus. Reads (vertical lines) are shown above

annotated genes. Blue rectangles represent exons, and blue lines represent intronic

Figure 3.5 (continued):

regions. Arrows indicate directionality. The middle RIN3 box represents the gene. (C) Comparison of reads mapping to *RIN3* to a neighboring gene, *SLC24A4*.

Since EnvPb1 is located within an intron of the *RIN3* gene, it is possible that reads assigned to *envPb1* actually derive from *RIN3* introns. This is further substantiated by the fact that a high density of reads aligns to *RIN3* introns over the length of the gene, indicating that the ones overlapping with *envPb1* are not specific (Figure 3.5B). These are likely to be bona fide reads due to the fact that the neighboring gene, *SLC24A4*, has comparatively low coverage (Figure 3.5C). BLAST searches on polyA-selected thymus mRNA RNA-seq data sets in the SRA (ERR1857349 and ERR1857367) do not return any reads aligning to the *envPb1* ORF, which seems to confirm our observation. Despite this, we do detect EnvPb1 protein in the placenta, thus indicating that EnvPb1 is expressed in the placenta. This possibility is intriguing because EnvPb1 had previously been ruled out as a syncytin mainly because it did not exhibit placenta-specific mRNA expression. Our results challenge this idea and call for further characterization of EnvPb1 in the placenta. Given this, we wanted to better understand how EnvPb1 is processed in cells.

EnvPb1-mediated cell-cell fusion depends on R428

The EnvPb1 amino acid sequence does not contain a prototypic furin-cleavage site (R-X-R/K-R), but rather, it contains a "weak" furin-cleavage site (R^{428}-K^{429}-T^{430}-R^{431}) as predicted by the ProP 1.0 server [202]. Given this, we wanted to test if this site leads

to proteolytic processing of EnvPb1 in cells. We used site-directed mutagenesis to

mutate R^{428} to S^{428}, which was predicted to completely abolish any recognition by furin

at this site (Figure 3.6A). We also generated a mutant in which T^{430} was mutated to

R^{430} to yield a consensus furin cleavage site. These mutants are referred to as SKTR

and RKRR respectively. We expressed these mutants as well as wild-type (WT)

EnvPb1 in Vero cells and assessed protein processing in cell lysates and cell

supernatants by SDS-PAGE followed by western blot with EnvPb1 antisera (Figure

3.6B). In cell lysates, EnvPb1 expressed by all three constructs is primarily in the

unprocessed (SU/TM) form, although there do appear to be lower levels of processed

(SU) EnvPb1 in SKTR transfected cells (Figure 3.6B).
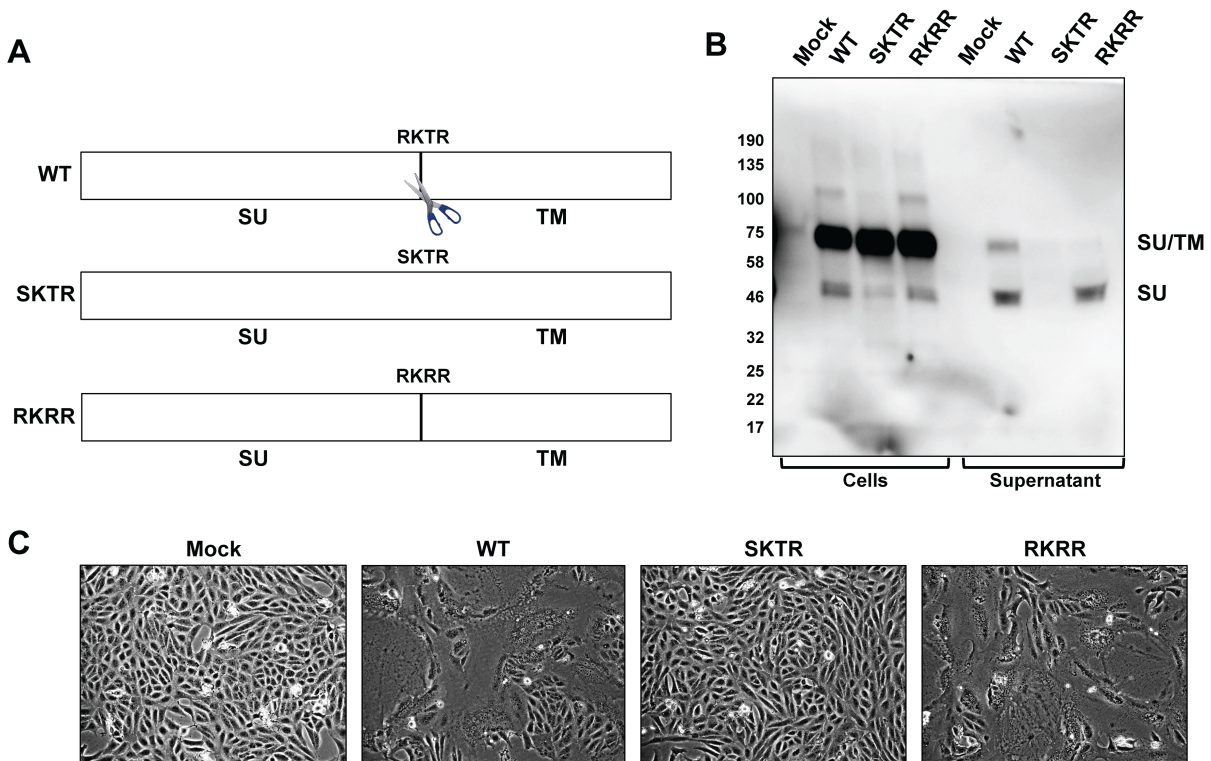
Figure 3.6: Dependence of EnvPb1 processing and fusion on the predicted furin cleavage site. (A) Schematic of furin cleavage mutant constructs. WT refers to the full coding sequence of EnvPb1. RKTR = predicted furin cleavage site. SKTR refers to a construct generated in which the R^{428} was mutated to an S, and RKRR refers to a construct generated to have a consensus furin cleavage site. (B) Western blot of cell lysates or supernatants from cells transfected with furin mutant constructs shows cellular expression and supernatant processing. (C) Fusion assay in Vero cells indicates that the SKTR mutant does not fuse cells.

However, EnvPb1 released into the supernatant does correlate with the nature of its cleavage site. WT is a mix between processed and unprocessed EnvPb1 in the supernatant, which is in line with it having a sub-optimal cleavage site. Further, SKTR does not show any visible processed EnvPb1 in the supernatant, substantiating that this site undergoes proteolytic processing, most likely by furin. Moreover, supernatant RKRR is primarily in the processed form. Since the sub-optimal RKTR furin cleavage site is conserved in all EnvPb1 except squirrel monkey, which has a consensus furin cleavage site (RKRR), it is possible that this sub-optimal furin cleavage site is selectively maintained in primate EnvPb1.

We next tested the requirement of proteolytic processing of EnvPb1 for cell-cell fusion. Syncytia formation is not observed in SKTR-transfected Vero cells, indicating that processing at this site is required for cell-cell fusion (Figure 3.6C). As expected, the RKRR mutant fuses. We were interested that we could detect EnvPb1 SU in the supernatant of transfected cells because it implies that EnvPb1 could be secreted or packaged into extracellular vesicles. However, since EnvPb1 fuses Vero cells, the other interpretation of this result is that SU in the supernatant is just a by-product of the fusion reaction. Therefore, we wanted to test if we could detect EnvPb1 secretion in cells that are not fused by EnvPb1.

EnvPb1 release occurs independent of cell-cell fusion

To test whether EnvPb1 SU release is dependent on EnvPb1 fusion, we transfected 293T cells with plasmids encoding WT EnvPb1, SKTR, or RKRR. It has been previously reported that EnvPb1 does not fuse 293T cells [150], and our results

confirm this (Figure 3.7A). We collected supernatant from transfected cells at 24 and 48

hours post transfection (HPT) and harvested cells lysates at 48 HPT. These samples

were subjected to SDS-PAGE followed by western blot with EnvPb1 antisera. Despite

the fact that EnvPb1 does not fuse 293T cells, we detect EnvPb1 in the supernatant of

transfected 293T cells at both 24 and 48hpt (Figure 3.7B).
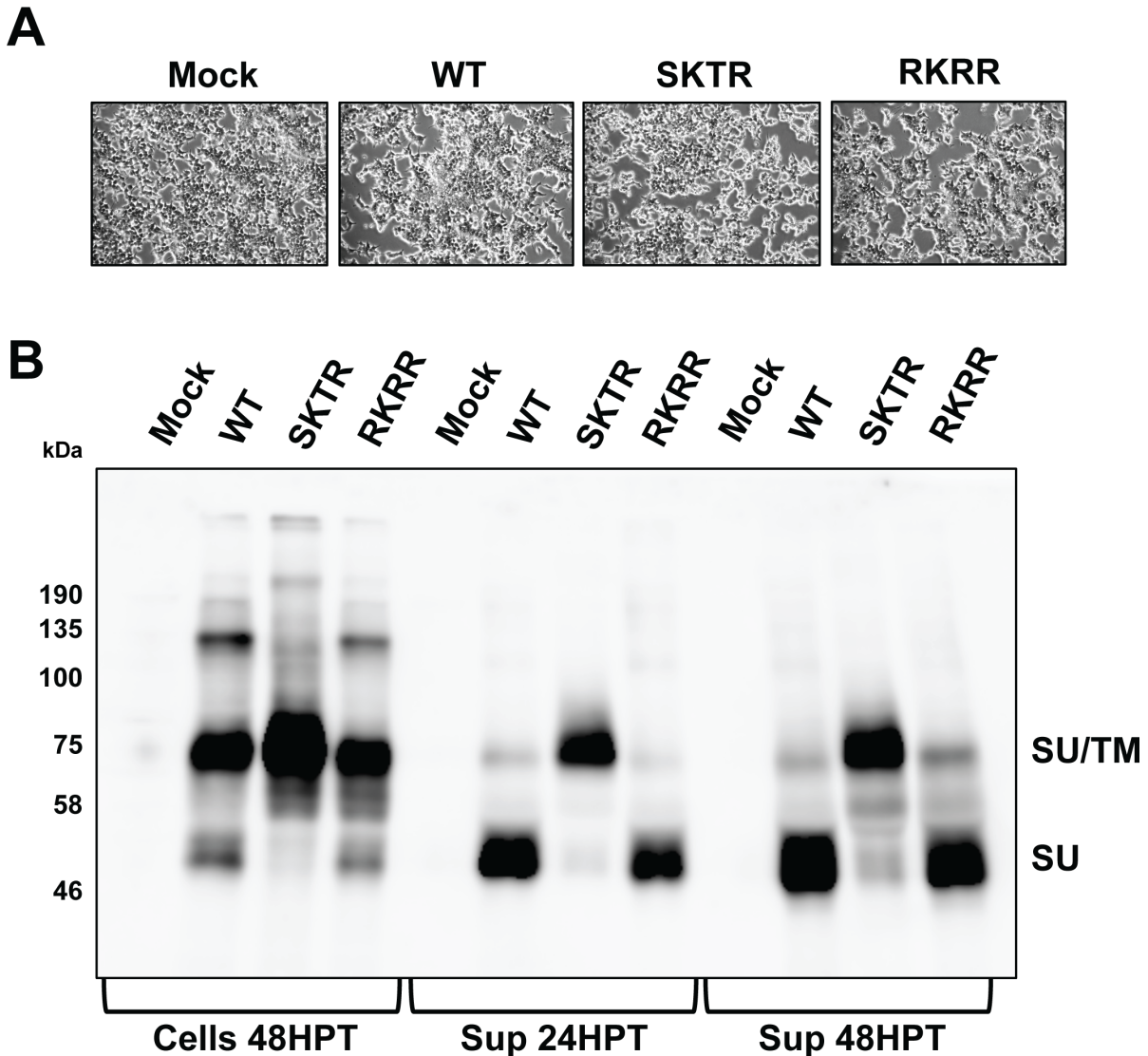
Figure 3.7: EnvPb1 release occurs independent of cell-cell fusion. (A) Fusion assay in 293T cells. 293T cells were transfected with vector expressing WT, furin-cleavage null (SKTR), or consensus furin-cleavage (RKRR) EnvPb1. Cells were assessed for syncytia formation at 24 HPT by phase contrast microscopy. (B) Western blot of EnvPb1 transfected 293T cells and supernatants. Supernatant (sup) from transfected cells was collected, clarified by centrifugation, and filtered through a 0.45μm filter at 24 and 48 HPT. Cells were collected and lysed at 48 HPT. Western blot was performed using rabbit anti-EnvPb1 sera.

In line with our observations in Vero cells, the integrity of the furin cleavage site impacts EnvPb1 processing in cells. In cell supernatants we see the release of EnvPb1, indicating that this process occurs in the absence of cell-cell fusion. Interestingly, although the SKTR mutant is not cleaved, we detect it in cell supernatants. Without proteolytic processing of the furin cleavage site, it is likely that the unprocessed form we detect is membrane bound. Thus, the observation that SKTR is released from cells suggests that EnvPb1 detected in cell supernatants is membrane-bound. These results were confirmed in another non-fused cell line (BSRT7) and another fused cell line (U2OS) (not shown). This raises the possibility that EnvPb1 is incorporated in extracellular vesicles such as exosomes. Exosomes and retroviruses exploit similar biogenesis pathways, so given the conservation of EnvPb1, it is possible that it could be packaged into exosomes [203]. This observation warrants further characterization of extracellular form of EnvPb1 and opens up a potential physiological role for a fusogenic retroviral Env that is distinct from cell-cell fusion events. Given this potential, we next wanted to develop tools to better understand how EnvPb1 mediates fusion.

Recovery of recombinant VSV-envPb1

To better understand the mechanism by which EnvPb1 induces fusion and to search for cellular factors necessary for fusion, we generated recombinant VSV expressing EnvPb1 (Figure 3.8A). We were not able to rescue virus in which the full EnvPb1 ORF replaced VSV-G in the VSV genome. Since the cytoplasmic tail of EnvPb1 is 46 amino acids (compared to 23 for VSV-G), we wondered if the cytoplasmic tail was preventing efficient incorporation into VSV particles. Since cytoplasmic tail

truncations can increase pseudotyping of viral vectors, we constructed an EnvPb1 in which we replaced the EnvPb1 cytoplasmic tail with the 8 C-terminal amino acids comprising the VSV-G cytoplasmic tail [204-206]. This truncated EnvPb1 retained fusogenicity in Vero cells and allowed us to recover infectious virus (Figure 3.8A). Initial recovery of rVSV-EnvPb1 yielded low titers that were not significantly improved by passaging of virus in different cell types or at different temperatures. Therefore, to increase the titer of rVSV-EnvPb1, we passaged the virus on cells that were transfected with a VSV-G expression plasmid. This way, budding virions would incorporate VSV-G to improve particle assembly and release, but VSV-G would not be present in the genome. After three rounds of passaging with VSV-G complementation, we generated a large stock of rVSV-EnvPb1 bearing VSV-G on the surface that grew to 1×10^8 plaque forming units (pfu)/ml. We passaged this virus stock on Vero cells to remove VSV-G bearing virus and yield rVSV-EnvPb1. Through this method, rVSV-EnvPb1 viral titers were increased to 4.65×10^6 pfu/ml.
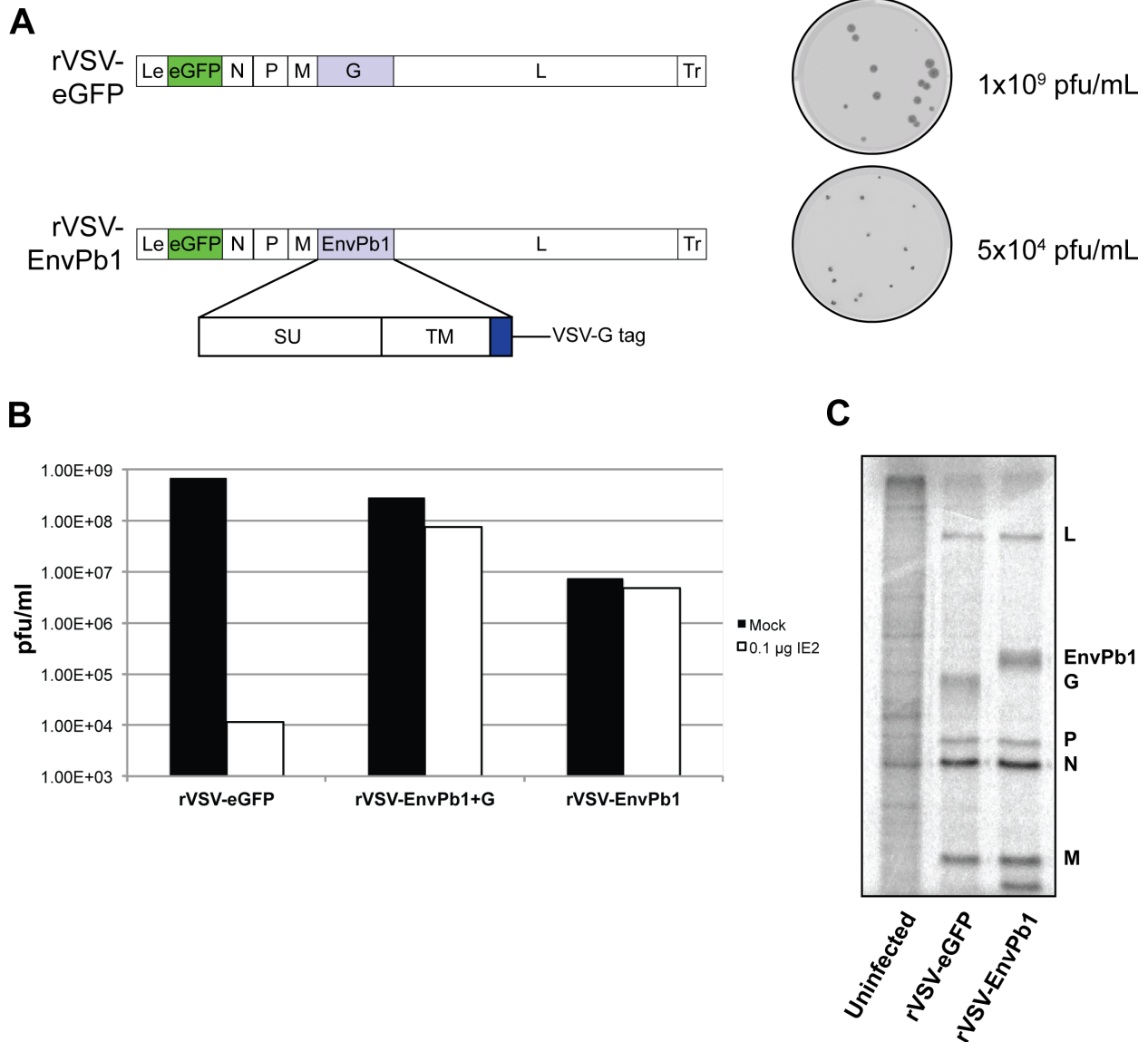
Figure 3.8: **Composition and infectivity of rVSV-EnvPb1.** (A) Schematic diagram of

rVSV-EnvPb1 genome. The VSV genes encoding N, P, M, G, L, and eGFP are shown

flanked by the viral leader (Le) and trailer (Tr) regulatory sequences. Viral titers

determined by plaque assay on Vero are expressed as plaque forming units per milliliter

(pfu/ml) and representative fluorescent scans shown. (B) rVSV-EnvPb1 is resistant to

IE2 neutralization. To increase viral titers, rVSV-EnvPb1 was passaged 3 times in

BSRT7 cells expressing VSV-G. A final round of infection was performed on Vero cells.

Figure 3.8 (continued):

rVSV-eGFP, virus passaged in VSV-G-expressing cells (rVSV-EnvPb1+G), and the final passage of rVSV-EnvPb1 was incubated for 30 minutes at 37ºC with 0.1µg of the VSV-G neutralizing antibody IE2. To determine if VSV-G mediated infection was occurring, viruses incubated with IE2 were plaqued on Vero cells. (C) Total protein synthesis in infected cells. Vero cells were infected (MOI=3) with rVSV-EnvPb1 or rVSV-eGFP, and viral protein synthesis was measured by metabolic incorporation of [^{35}S]-methionine at 5.5 hours post infection (HPI). Total cytoplasmic extracts were prepared, equivalent amounts analyzed by SDS-PAGE and proteins visualized by autoradiography. Viral proteins are labeled on the right.

To ensure that these virions did not retain VSV-G-mediated infectivity, we assessed their sensitivity to neutralization with the VSV-G neutralizing antibody IE2 [168]. VSV-eGFP, rVSV-EnvPb1 passaged in VSV-G expressing cells (rVSV-EnvPb1+G), and rVSV-EnvPb1 were incubated with IE2 and titered by plaque assay on Vero cells (Figure 3.8B). Incubation of VSV with IE2 leads to a greater than 4-log decrease in viral titers, while rVSV-EnvPb1 is largely unaffected indicating that its infectivity is dependent on EnvPb1. Additionally, incubation of rVSV-EnvPb1+G with IE2 results in minimal decreases in titer indicating that particles produced in VSV-G-expressing cells also incorporate EnvPb1.

Next we infected cells with rVSV-eGFP or rVSV-EnvPb1 to assess viral protein synthesis. Infected cells were pulsed with [^{35}S]-methionine and cysteine at 5.5 hours post infection in methionine and cysteine-free media. After one hour of labeling, cells

were collected, lysed, and assessed by SDS-PAGE and autoradiography. We detect approximately equal levels of VSV protein synthesis between rVSV-eGFP and rVSV-EnvPb1 (Figure 3.8C). Furthermore, EnvPb1 seems to be produced to similar levels as VSV-G indicating that it is well expressed. Interestingly, this protein product migrates at a molecular weight of approximately 75 kDa, which is the predicted molecular weight of SU-TM before proteolytic processing. Since this assay labels newly synthesized proteins, it is possible that EnvPb1 detected here has not yet trafficked to the Golgi for cleavage by furin.

rVSV-EnvPb1 is sensitive to lysosomotropic agents

Since EnvPb1 induces cell-cell fusion at neutral pH, similar to many gammaretroviral envelope proteins, it would be expected that receptor binding is a sufficient trigger to catalyze the membrane fusion reaction. Conversely, viruses that require acidic pH to trigger fusion are endocytosed and undergo conformational changes leading to fusion in the low-pH compartments of the endosome or lysosome. To test if EnvPb1-mediated fusion occurs at the plasma membrane, and to begin to understand the cellular requirements for EnvPb1 fusion, we treated cells with a panel of known lysosomotropic agents that block the acidification of intracellular vesicles and infected them with rVSV-EnvPb1. Ammonium chloride (NH_4Cl) is a weak base that is taken up by endosomes and lysosomes where it acts as an alkalinizing agent, and bafilomycin A1 (Baf A1) is an inhibitor of the H+ ATPase pump. As controls, we tested VSV and VSV bearing the parainfluenza virus 5 (PIV5) attachment (HN) and fusion (F) proteins (rVSV-PIV5). VSV membrane fusion occurs at low pH and is sensitive to

lysosomotropic agents [207]. Conversely, PIV5 F-mediated fusion occurs at the plasma membrane at neutral pH and is insensitive to lysosomotropic agents [208].

We infected BSRT7 cells at an MOI of 1 with rVSV-eGFP, rVSV-PIV5, or rVSV-EnvPb1. Cells were pretreated with either 0.1μM Baf A1 or DMSO for 30 minutes prior to infection. Viruses were absorbed for 1 hour in the presence of inhibitor, and infection proceeded in the presence of the inhibitor. At 6HPI, cells were assessed for infection by fluorescence microscopy. As expected, VSV infection is completely blocked by Baf A1, whereas Baf A1 has no effect on rVSV-PIV5 entry (Figure 3.9A)**.**
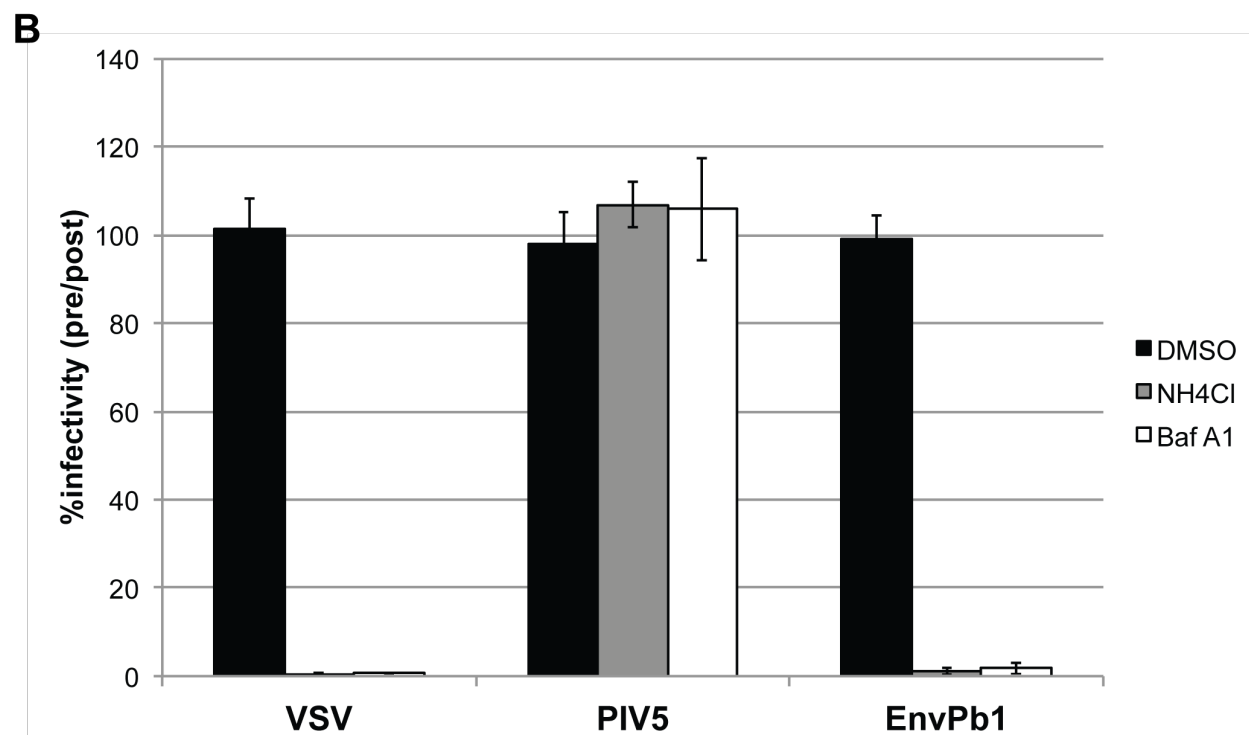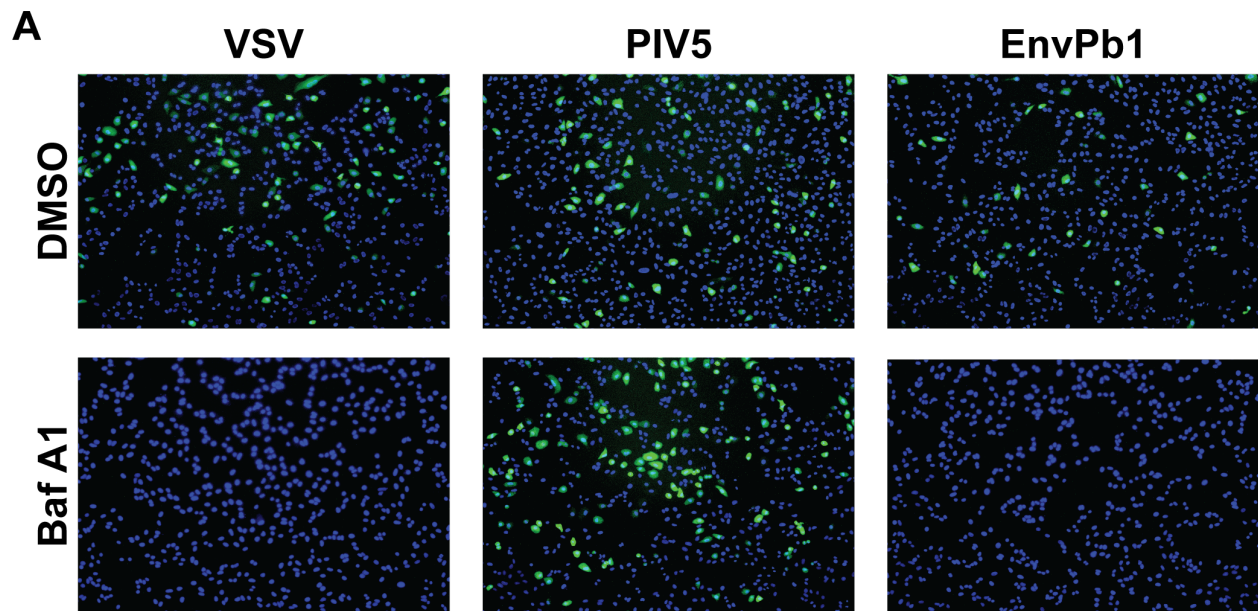
Figure 3.9: Lysosomotropic agents block rVSV-EnvPb1 entry. (A) Bafilomycin A1

blocks rVSV-EnvPb1entry. BSRT7 cells were treated with 10µM Bafilomycin A1 (Baf

A1) 30 minutes prior to infection. Virus was absorbed at an MOI of 1 in the presence of

Baf A1 for 1 hour. Bafilomycin was left on and cells were imaged by fluorescence

Figure 3.9 (continued):

microscopy at 6 HPI. GFP denotes infected cells and blue are DAPI stained nuclei. (B) Ammonium chloride (NH_4Cl) and bafilomycin A1 (Baf A1) block entry of rVSV-EnvPb1 in Vero cells. Vero cells were treated with 10mM NH_4Cl or 10µM Baf A1 for 30 minutes. Virus was absorbed at an MOI of 1 for 1 hour in the presence of inhibitor. Post-treated controls received inhibitor at 2 hours post infection. At 6 hours post infection, cells were assessed for GFP expression by flow cytometry. Values represent average %infectivity in pre-treated samples normalized to %infectivity in post-treated samples from three independent experiments. Error bars signify two standard deviations.

However, rVSV-EnvPb1 infectivity is completely blocked in the presence of Baf A1. To rule out that this was a cell-specific phenomenon, we further substantiated this result in Vero cells that were pre or post-treated with Baf A1 or NH_4Cl. We infected Vero cells at an MOI of 1 with rVSV-eGFP, rVSV-PIV5, or rVSV-EnvPb1. Cells were either pretreated with 10mM NH_4Cl or with 0.1µM Baf A1 for 30 minutes prior to infection or treated at 2 hours following infection to control for any post entry off-target effects of the inhibitors. DMSO-treated cells were included as a vehicle control. At 6 HPI, cells were collected and assessed for GFP expression by flow cytometry. The percentage of infected cells in the inhibitor-pretreated condition is normalized to the post-treatment control. Again, the lysosomotropic agents block VSV entry but not rVSV-PIV5 (Figure 3.9B). Furthermore, we observe that both NH_4Cl and Baf A1 block entry of VSV-EnvPb1. These results implicate a role for endosomal acidification in EnvPb1-

mediated fusion, despite cell-cell fusion occurring at neutral pH. Thus, we wanted to test if the role of low-pH is to induce conformational changes in EnvPb1.

EnvPb1 is resistant to low-pH inactivation

One interpretation of the result in Figure 3.8 is that low pH acts to induce conformational changes in EnvPb1 that result in the triggering of fusion. For viruses that undergo fusion in an acidic intracellular compartment, such as influenza A or Semliki Forest virus (SFV), the exposure of virions to acidic pH is often sufficient to trigger the fusion protein [209].   Pretreatment of virus with low pH induces irreversible conformational changes in the fusion protein, and the virus is unable to infect cells. Viruses that fuse at the plasma membrane are resistant to low-pH inactivation. Since VSV-G undergoes a reversible conformational change, exposure to low-pH induces conformational changes leading to fusion; however, if the pH is neutralized, VSV-G will return to its pre-fusion state [210, 211]. We set out to test if acidic pH could trigger EnvPb1 by pre-incubating viruses at pH 5 before neutralizing the inoculum and infecting cells. VSV and VSV bearing the SFV spike protein were included as controls. Viruses were incubated at 37ºC for 30 minutes in pH 5 or pH 7.4 PBS. Following this, the virus inoculum was neutralized with cell culture media at 37ºC for 15 minutes. The inoculum was absorbed onto cells for 1 hour after which it was removed and fresh media was added. Cells were harvested and assessed for infectivity at 6 HPI by flow cytometry. Infectivity is measured in percent GFP positive cells, and values were normalized to infectivity at pH 7.4. As expected, VSV-SFV exposed to pH 5 is unable to mediate infection due to triggering of its fusion protein (Figure 3.10).
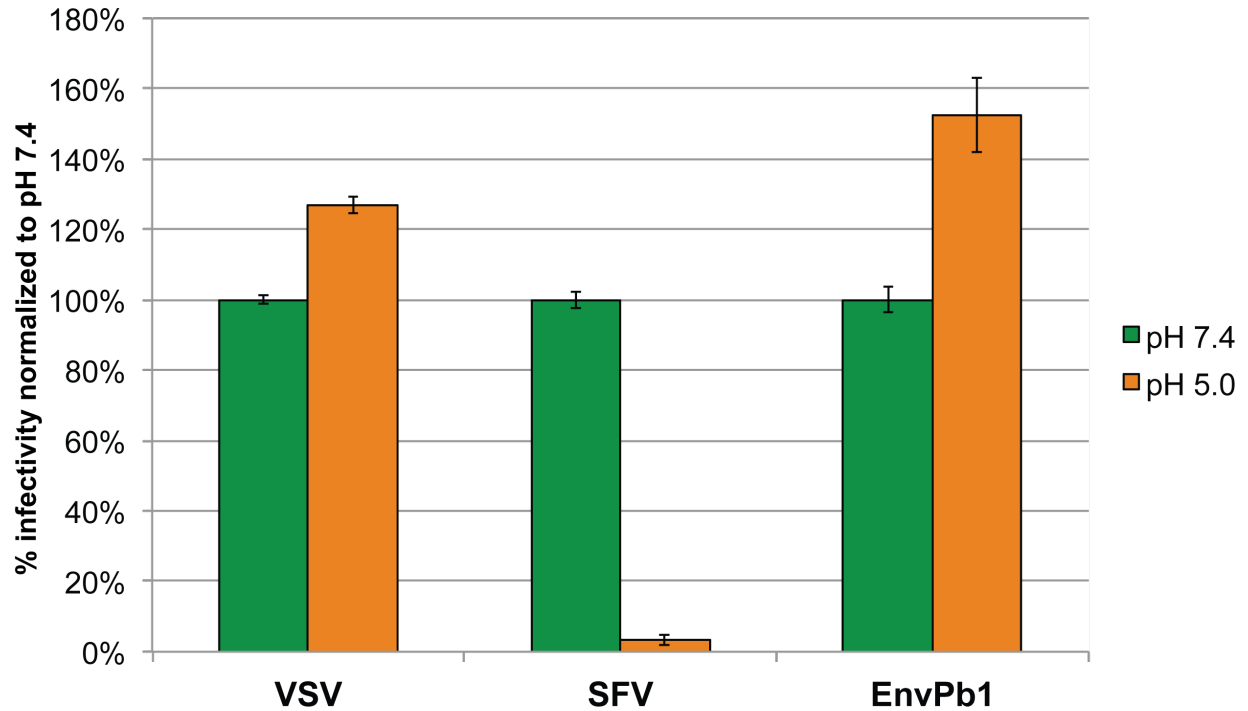
Figure 3.10: rVSV-EnvPb1 is resistant to low pH inactivation. rVSV-EnvPb1, VSV-eGFP, and rVSV-SFV were incubated in pH5 or pH7.4 buffer for 30 minutes at 37°C. Following this, DMEM was added to return viruses to neutral pH. The viruses were then absorbed onto Vero cells for one hour and GFP+ cells were detected by flow cytometry at 6 HPI. Values were normalized to percent infectivity at pH 7.4. Values represent averages of three independent replicates and error bars represent two standard deviations.

VSV treated at pH 5 retains its infectivity, suggesting that glycoprotein triggering was reversed following neutralization. Despite being blocked by lysosomotropic agents, rVSV-EnvPb1 is resistant to low-pH inactivation, indicating that acidic pH alone does not induce conformational changes in EnvPb1. This indicates that the requirement for acidification in EnvPb1-mediated entry likely serves another purpose. This is in line with the fact that EnvPb1 induces cell-cell fusion at neutral pH and suggests that the requirement for endosomal acidification could be for a cellular factor involved in the fusion of EnvPb1. To understand what cellular factors are involved in EnvPb1-mediatied fusion we set out to perform a forward genetic screen to identify cellular factors involved in EnvPb1 fusion.

Discussion

In this chapter, we report the evolutionary and functional characterization of the endogenous retrovirus envelope protein EnvPb1. Our analysis of the currently available primate genomes indicates that the *envPb1* ORF is highly conserved in monkeys and apes. Despite some primate *envPb1* ORFs harboring insertions or deletions when compared to the human copy, all insertions or deletions preserve the reading frame, suggesting that preservation of the ORF provides a beneficial function to the host. In line with this, we find that *envPb1* is under purifying selection in simians with clade-specific differences. Pairwise d_n/d_s values are most indicative of purifying selection in NWM. This could indicate that purifying selection is strongest in NWM or that other sequences do not have enough sequence divergence for accurate estimates of

selective pressure. Despite this, our analysis strongly suggests that EnvPb1 has been exapted to provide a beneficial function to the host.

Furthermore, we show that some primate *envPb1* ORFs still encode fusogenic proteins. Marmoset and gorilla EnvPb1 are both fusogenic in Vero cells whereas baboon EnvPb1 is not. These data suggest that EnvPb1 fusogenicity is an evolutionarily conserved feature. The fusogenicity of more *envPb1* ORFs needs to be tested to demonstrate if there are lineage specific differences in fusogenicity. For example, despite orthologous *envV2* sequences in hominoids, OWM, and NWM being highly conserved and retaining furin cleavage sites, CXXC motifs, fusion peptides, and transmembrane spanning regions, only OWM and marmoset EnvV2 induce cell-cell fusions. This indicates that subtle sequence changes can differentiate fusogenic and non-fusogenic ERV Envs. Given that all EnvPb1 defined have fully conserved furin-cleavage sites, CXXC, CX_6CC motifs, and transmembrane domains, it is likely that they all retain the ability to induce cell-cell fusion and our results merely indicate that there are cell-type specific differences such as receptor usage or receptor interference that lead to baboon EnvPb1 not being fusogenic in Vero cells. Therefore, the fusogenicity of these EnvPb1 should be tested in other cell lines. Given the high degree of amino acid conservation, sequence comparison between EnvPb1 with different receptor or fusogenic properties could help define receptor-binding regions in SU.

While we were able to confirm previously published reports of EnvPb1 transcripts in human tissues by analyzing publicly available RNA-seq datasets, our analysis of EnvPb1 protein levels in six human tissues is in contrast to the transcriptome data [150, 152]. Indeed, we only detect EnvPb1 in placental tissues. The most likely explanation

for this observation is the fact that in all three transcriptome analyses, total RNA was used instead of polyA-selected mRNA. As EnvPb1 resides in an intron of the *RIN3* gene, it is possible that these suspected EnvPb1 transcripts actually derive from *RIN3*RNA. When we analyzed the RNA-seq data from human thymus total RNA for splice junctions between EnvPb1 and its 5'LTR, we did not find any despite an abundance of reads mapping to the *envPb1* ORF. Furthermore, there were no other splice junctions aligning to EnvPb1 transcripts that could represent possible transcription initiation sites. Indeed, BLAST searches of polyA-selected thymus RNA indicate that no transcripts map to the *envPb1* ORF. These data suggest that the presumed *envPb1* transcripts detected by us and others derive from *RIN3* introns. In support of this, it was reported that no expressed sequence tags, which represent RNA message, map to the *envPb1* ORF [149].   Despite this, Blaise *et al.* did detect spliced *envPb1* transcripts in placenta tissue by using RT-qPCR primers residing in the 5'LTR and *envPb1* ORF, and others have detected *envPb1* transcription in cytotrophoblasts [152, 154]. Taken together, our work and these previous observations suggest that *envPb1* may be specifically expressed in the placenta. To test this, we are currently assessing EnvPb1 transcripts and protein in placental cells and tissues as well as evaluating EnvPb1 splice junctions in RNA-seq data from placenta. To confirm our assessments, future studies should enrich for polyA-selected mRNA to test for *envPb1* transcript levels in human tissues.

Given the presumed widespread tissue tropism of EnvPb1, it was never considered to play a positive role in the placenta like the syncytin proteins derived from HERV-FRD and HERV-W. However, since we only detect EnvPb1 in human placenta

tissue, it is possible that EnvPb1 does play a positive role in the placenta. Indeed this physiological role is likely related to fusion, as EnvPb1 is one of only three HERV Env ORFs that is capable of mediating cell-cell fusion at physiological conditions. While other HERV Envs such as Env-T and HEMO have been hypothesized to play physiologically relevant roles in human health, they are not proteolytically processed indicating that their function is distinct from fusion. Thus, conservation of fusogenicity is rare and significant when it occurs. While it is possible that EnvPb1 may play a role in syncytiotrophoblast fusion similar to the syncytins, another possibility is that it is released from cells in extracellular vesicles.

Our data indicate that EnvPb1 is secreted from cells – even those it does not fuse. When cells are transfected with plasmid encoding EnvPb1 with or without a functional furin cleavage site, we detect the presence of EnvPb1 in supernatants. Placenta exosomes are thought to play a role in immunotolerance and development, so it is intriguing to consider that a retroviral Env could be packaged into such a vesicle and permit its attachment and uptake in a target cell [212]. Indeed, evidence exists for the incorporation of Syncytin-1 and Syncytin-2 in placenta-derived exosomes, so it is possible that EnvPb1 could be playing a similar role given our data [213]. To test this, supernatants of EnvPb1-transfected cells could be enriched for exosomes, and these exosomes tested for incorporation of EnvPb1. Future studies could utilize our EnvPb1 antisera to address if placental exosomes contain EnvPb1.

Additionally, we developed a recombinant vesicular stomatitis virus that bears functional EnvPb1 on the surface. We utilized this virus to better understand how EnvPb1 interacts with cells. Our work with rVSV-EnvPb1 indicates that EnvPb1

requires endosomal acidification, and that low pH alone does not induce conformational

changes in EnvPb1. This suggests a role for endosomal uptake that is distinct from low

pH-induced fusion. Therefore, it is possible that EnvPb1 binding at the cell surface

leads to endosomal uptake where it either binds another cellular protein or is acted on

by an endosomal protein that is dependent on the low pH environment of the

endosome. Use of rVSV-EnvPb1 will help to further interrogate how EnvPb1 mediates

fusion and what the cellular requirements for EnvPb1 fusion are.

Materials and methods

Cells lines, plasmids and viruses

Syrian golden hamster kidney BSRT7 cells (generously provided by Dr. K. Conzelmann

[174], human embryonic kidney 293T cells (ATCC CRL-3216), and African green

monkey kidney Vero cells (ATCC CCL-81) were maintained at 37°C and 5% CO_2 in

Dulbecco's modified Eagle medium (DMEM; Corning Life Sciences, Tewksbury, MA)

supplemented with 10% fetal bovine serum (FBS; Tissue Culture Biologicals, Tulare,

CA). The human haploid cell line, HAP1 (derived from male chronic myelogenous

leukemia (CML) cell line KBM-7 [214] and generously provided by Thijn Brummelkamp),

the Chinese hamster ovary, CHO-K1 (ATCC CCl-61), and the Chinese hamster ovary

deficient in xylosyltransferase, pgsA-745 (ATCC CRL-2242) were maintained at 37ºC

and 5% CO_2 in Iscove's Modified Dulbecco's Medium (IMDM) with 10% FBS.

We commissioned a codon-optimized human *envPb1* (NCBI: DQ247958) from

Life Technologies Corporation (Carlsbad, CA). Gorilla, gibbon, baboon, and marmoset

envPb1 sequences were collected from the UCSC genome browser by BLAT searches

with human *envPb1* as query [182]. These sequences were synthesized using the FragmentGENE service from Genewiz (South Plainfield, NJ). All primate *envPb1* were cloned into the pVRC vector (provided by Kevin McCarthy in the laboratory of Steve Harrison [215]) using SalI and NotI cut sites. Vectors expressing the EnvPb1 furin cleavage mutants were generated using site-directed mutagenesis on pVRC-EnvPb1. pVSV-ΔG-EnvPb1 was generated by cloning EnvPb1 into pVSV-ΔG-eGFP [175] through MluI and NotI restriction sites as described previously. Vector inserts were confirmed by sequencing.

rVSV-ΔG-EnvPb1 was generated by transfection of the VSV genomic plasmid along with plasmids carrying VSV-N, -P, -L, and -G as previously described [25]. To increase yield, rVSV-ΔG-EnvPb1 was propagated in BSRT7 cells that were transfected with a plasmid expressing VSV-G. After three rounds of trans G complementation, rVSV-ΔG-EnvPb1 was passaged on Vero cells to yield particles lacking VSV-G. rVSV-ΔG-EBOV, a recombinant VSV expressing the Ebola virus glycoprotein [175], rVSV-ΔG-PIV5, a recombinant VSV expressing the parainfluenza virus 5 hemagglutinin-neuramididase (HN) and fusion glycoprotein (F), and rVSV-ΔG-EnvPb1 were amplified, purified, and maintained as previously described .   All virus propagation occurred in DMEM supplemented with 2% FBS and penicillin/streptomycin/kanamycin at 34°C and 5% CO_2. Virus titers were determined by plaque assay on Vero cells as previously described [13]. GFP-expressing fluorescent plaques were imaged using a Typhoon 9400 imager (GE Healthcare, Aurora, OH).

Database screening and sequence analysis

Primate genomes were searched in the available primate genome builds in the UCSC genome browser (http://genome.ucsc.edu/) for homology to human *envPb1* using the BLAT search tool. Other sequences were acquired through BLAST searches using human *envPb1* as query (https://blast.ncbi.nlm.nih.gov/blast/Blast.cgi). Predicted homologues were searched for transcription start and stop sites and syntenic loci were mapped.

Nucleotide sequences were aligned using the MAFFT alignment algorithm on the GUIDANCE server (http://guidance.tau.ac.il/) [190, 191]. Sequence alignments were imported into MEGA version 7.0 (https://www.megasoftware.net/), and maximum likelihood phylogenetic nucleotide trees were constructed based on the Tamura-Nei model with bootstrap percentages calculated after 1,000 replicates [216]. A discreet Gamma distribution was used to model evolutionary rate differences among sites (43 categories (+G, parameter = 0.5823)). dN/dS ratios were calculated using the codeml package in PAML (http://abacus.gene.ucl.ac.uk/software/paml.html).

ERV-Pb proviral loci were analyzed using RetroTector searches on primate *envPb1* coding regions plus 10,000 nucleotides up and downstream (http://retrotector.neuro.uu.se/).

RNA-seq data mining

Raw RNA-seq reads from 20 human tissues were acquired from the NCBI sequence read archive (SRA) from project PRJNA280600. Reads were imported into the CLC genomics workbench and aligned to human genome build hg19 that was annotated to include the *envPb1* ORF. Unique reads were intersected with mRNA and

only exact matches were analyzed. Read counts were normalized by length of the transcripts. Analysis of splice junctions was performed using HISAT2 as previously described [201].

Human tissue samples

Human tissue protein medleys were purchased from Clontech (Mountain View, CA, and source information is as follows. Normal, whole placentas pooled from 15 females, age 19-33. Normal, whole small intestines pooled from 9 male/females, ages 21-45 who died from trauma and sudden death. Normal, whole spleen from a 55-year-old male who died a sudden death. Normal, whole testes pooled from 18 adult males ages 19-64 who died of trauma. Normal thymus from a 21-year-old male who died a sudden death. Normal, whole ovary from a 30-year-old female who died of trauma.

Development of EnvPb1 antisera

In collaboration with Lindsey Robinson-McCarthy and Kevin McCarthy, we generated a region of EnvPb1 (aa 51-208) that was predicted to contain the receptor-binding region of the SU subunit through sequence homology and predicted structural similarities to MLV and FeLV receptor binding domains (RBDs) [217, 218]. A rabbit was inoculated with EnvPb1-RBD and we received sera at various times following infection. The final bleed sera was tested and confirmed to detect both unprocessed (SU/TM) and processed (SU alone) EnvPb1 in transfected cells. Further, the anti-EnvPb1 serum, but not the pre-bleed serum, is capable of blocking EnvPb1-mediated cell-cell fusion, which implies that it indeed recognizes the receptor-binding domain.

SDS-PAGE analysis and Western blots.

Cell lysates or tissue samples were run on a 4-15% polyacrylamide gel and transferred to a nitrocellulose membrane. Membranes were blocked with 5% dry milk in 0.1% Tween in PBS (0.1% PBS-Tween) and incubated with rabbit anti-EnvPb1 sera (1:1000 dilution), rabbit pre-bleed sera (1:1000 dilution), or mouse anti-ß-actin antibody (A5316: Sigma Aldrich; 1:5000 dilution) in 0.1% PBS-Tween buffer, followed by incubation with goat anti-rabbit or goat anti-mouse secondary antibody conjugated HRP (1:5000 dilution). Chemiluminescence signals were generated using an HRP substrate (Thermo, Rockford, IL) and detected using an Amersham Imager 600 (GE Healthcare Life Sciences, Pittsburgh, PA).

Cell fusion and supernatant release assays

Vero cells at 70-90% confluency in a 12-well dish were transfected with pVRC plasmid encoding different primate *envPb1*. 1µg of plasmid was mixed with 4µl of lipofectamine 2000 (Thermo Fisher, Waltham, MA) in a total volume of 200µl Opti-Mem (Thermo Fisher, Waltham, MA) for 15 minutes. Transfection mixture was added to 500µl Opti-mem for 4 hours, after which DMEM + 2%FBS was added. At 24 hours post transfection, cells were visualized by phase contrast microscopy at 10x magnification. Cells were then trypsinized and cell lysates were collected for analysis by western blot.

To assess supernatant release, 80-90% confluent 293T cells in 60mm dishes were transfected with 2µg of pVRC encoding human *envPb1* with 8µl of lipofectamine. Transfections were performed as in the cell fusion assay. Supernatants were harvested

at 24 or 48 hours post transfection and spun at 2,000 RPM for 10 minutes to remove

cell debris. Clarified supernatants were then filtered (0.45µm Millipore filter) and either

stored at -80ºC or mixed with SDS buffer containing BME and boiled for western blot

analysis.

Metabolic labeling of proteins in cells

To visualize viral protein synthesis, Vero cells were infected at an MOI of 3 with

rVSV-eGFP or rVSV-EnvPb1. At 5 HPI cells were starved of methionine and cysteine.

At 5.5 HPI, 50µCi of EasyTag^{TM} EXPRESS^{35}S Protein Labeling Mix (PerkinElmer,

Waltham, MA) was added to the media. At 6 HPI cells were harvested, lysed, and

subjected to SDS-PAGE gel electrophoresis. The gel was fixed for 1h in a gel fix

solution (10% acetic acid glacial, 30% methyl alcohol, and 60% H_2O) and then washed

2x with methyl alcohol for 1h. The fixed gel was dried and exposed overnight on a

phosphor screen, which was imaged using a Typhoon FLA 9500 biomolecular imager.

Inhibitor assays

Ammonium chloride (NH_4Cl, Sigma Aldrich) and bafilomycin A1 (Baf A1,

Calbiochem, Millipore) were administered at concentrations of 10mM and 10µM

respectively. DMSO (Sigma Aldrich) was included as a control and administered equal

to the total volume of Baf A1 added per condition. Vero or BSRT7 cells were treated

either 30 min prior to infection or at 2 HPI with DMEM plus the indicated inhibitors. VSV

inoculations in DMEM were incubated for 1h at 37ºC with an MOI of 3. Following

inoculation, cells were washed and media was replaced with DMEM. For inhibitor pre-

treated cells, inhibitors were maintained during inoculation, washing, and replacement of media. Cells were processed for cytofluorimetry at 6 HPI. Cells were washed with PBS and harvested with PBS containing 0.5mM EDTA. % GFP positive cells were counted using a modified FACSCalibur (Cytek Development, Fremont, CA) and data were analyzed using FlowJo (Tree Star Industries, Ashland, OR).

Low pH inactivation assay

PBS was buffered to pH5 or pH7.4. 200µl of pH buffer was added to 20µl of virus and incubated at 37ºC for 30min. Buffered viruses were neutralized by adding 200µl of DMEM and incubated at 37ºC for 15min. 300µl of neutralized inoculum were absorbed onto 90% confluent Vero cells in a 6-well dish at 37ºC for 1h. Initial neutralized virus was diluted such that the final inoculum would yield an MOI of 1. After absorption, inoculums were removed, cells were washed, and fresh DMEM was added. Cells were processed for and analyzed by cytofluorimetry at 6 HPI as in the inhibitor assay.

Work contributions

Lindsey Robinson-McCarthy and Kevin McCarthy developed EnvPb1 antiserum. Phillip Tomezsko, a dang hunk, performed splice junction analysis and IGV read visualization.

Chapter 4: General discussion

Studies of vesicular stomatitis virus (VSV) vectors bearing foreign viral fusion proteins (VFPs) have provided seminal insights into virology and vaccinology. For example, VSV bearing the Lassa fever virus (LASV) glycoprotein (GP) led to the discovery that LASV undergoes a pH-induced receptor switch to use LAMP1 as an intracellular receptor [219]. Importantly, studies of virus-host interaction using recombinant VSV have been corroborated with authentic viruses [26, 31, 219]. In addition, VSV vectors expressing foreign VFPs have proven to induce potent immune responses – mostly in small animal models [32, 33]. However, use of rVSV-ZEBOV during the 2014-2016 Ebola virus outbreak indicated that VSV vaccines are safe and effective for use in humans [41].

Our work expands on the use of VSV as both a biological tool and as a vaccine vector. In chapter two, we developed VSV vectors expressing ZIKV proteins. We tested the immunogenicity of one of these VSV vectors, rVSV-CprME, in mice, and discovered that it induces a neutralizing antibody response against ZIKV-E. Additionally, we demonstrated that while VSV can incorporate ZIKV-E, the majority of ZIKV-E is released from cells in the form of virus-like particles (VLPs). By assessing different ZIKV constructs, we showed that inclusion of the transmembrane-spanning region between capsid (C) and premembrane (prM) significantly improves release of E into cell supernatants. This indicates that this region likely acts as a signal peptide to ensure proper insertion of prM and E into membranes, which favors assembly and exocytosis of VLPs. This observation can aid the design of other vector-based vaccines for Zika virus and emerging flaviviruses.

The work of chapter three utilized rVSV-EnvPb1 to understand how this ERV envelope protein that is conserved in primate evolution mediates fusion. Subjecting rVSV-EnvPb1 to lysosomotropic agents revealed that entry requires endosomal acidification, but exposing rVSV-EnvPb1 to low pH suggests that low pH alone does not trigger fusion. Additionally, we show that EnvPb1 does not require glycosaminoglycans for entry. Further work with EnvPb1 revealed that it is evolutionarily conserved in simians and that two of three species tested are fusogenic in Vero cells. Additionally, we show that EnvPb1 seems to be expressed specifically in the placenta.

VSV incorporation of flavivirus VFPs

VSV recombinants incorporate all three known classes of viral fusion proteins [27, 32, 33, 64, 220]. This is most surprising for the class II VFPs given the structure of class II VFPs (Figure 1.4). Electron cryo-microscopy studies of viruses encoding class II VFPs revealed that these proteins form highly ordered, icosahedral morphologies on spherical virions [221-223]. This is in contrast to class I and class III VFPs whose spike-like structures are less conducive to the formation of highly ordered lattices and are amenable to incorporation into virions of irregular shape [224]. Despite this, VSV recombinants expressing class II VFPs from the alphavirus chikungunya virus and the bunyavirus Andes virus grow to high titers [33, 220]. This suggests that while there may be morphological barriers to VSV incorporation of class II viral fusion proteins, these alone are not prohibitive.

In the context of this literature, our studies of the rVSV-ΔG-ZIKV recombinants suggest that glycoprotein trafficking is the major determinant of class II VFP

incorporation into VSV. First, despite expressing high levels of ZIKV-E in cells, these viruses grow to low titers (Figure 2.1A). Secondly, we detect nearly 4-fold more viral genomes than infectious particles in rVSV-ΔG-ZIKV supernatants, which is indicative of the budding of naked particles or of particles that do not contain functional ZIKV-E (Figure 2.4). Further, we show that infection of cells with all rVSV-ZIKV results in the release of ZIKV VLPs (Figure 2.6). These observations are most likely attributable to the fact that flaviviruses assemble in ER-associated membranes and are exocytosed from cells while VSV assembles and buds at the plasma membrane. In support of this, class II VFPs that are readily incorporated into VSV exhibit plasma membrane localization. Alphavirus budding occurs at the plasma membrane, and alphavirus VFPs are retained in the plasma membrane in the absence of alphavirus nucleocapsids [225]. Additionally, while many bunyaviruses assemble in the Golgi and are released through exocytosis, New World bunyaviruses, which includes Andes virus, have been shown to bud from the plasma membrane, and the Andes virus VFPs Gn and Gc localize to the plasma membrane [226, 227]. While we do detect incorporation of functional ZIKV-E into VSV particles, these viruses grow to low titers and do not spread (Figure 2.1 and 2.3). The observations of chapter two suggest a model in which expression of ZIKV-E during VSV infection results in most ZIKV-E assembling and budding into ER-membranes, and some ZIKV-E trafficking to the plasma membrane where it is incorporated into VSV (Figure 4.1).
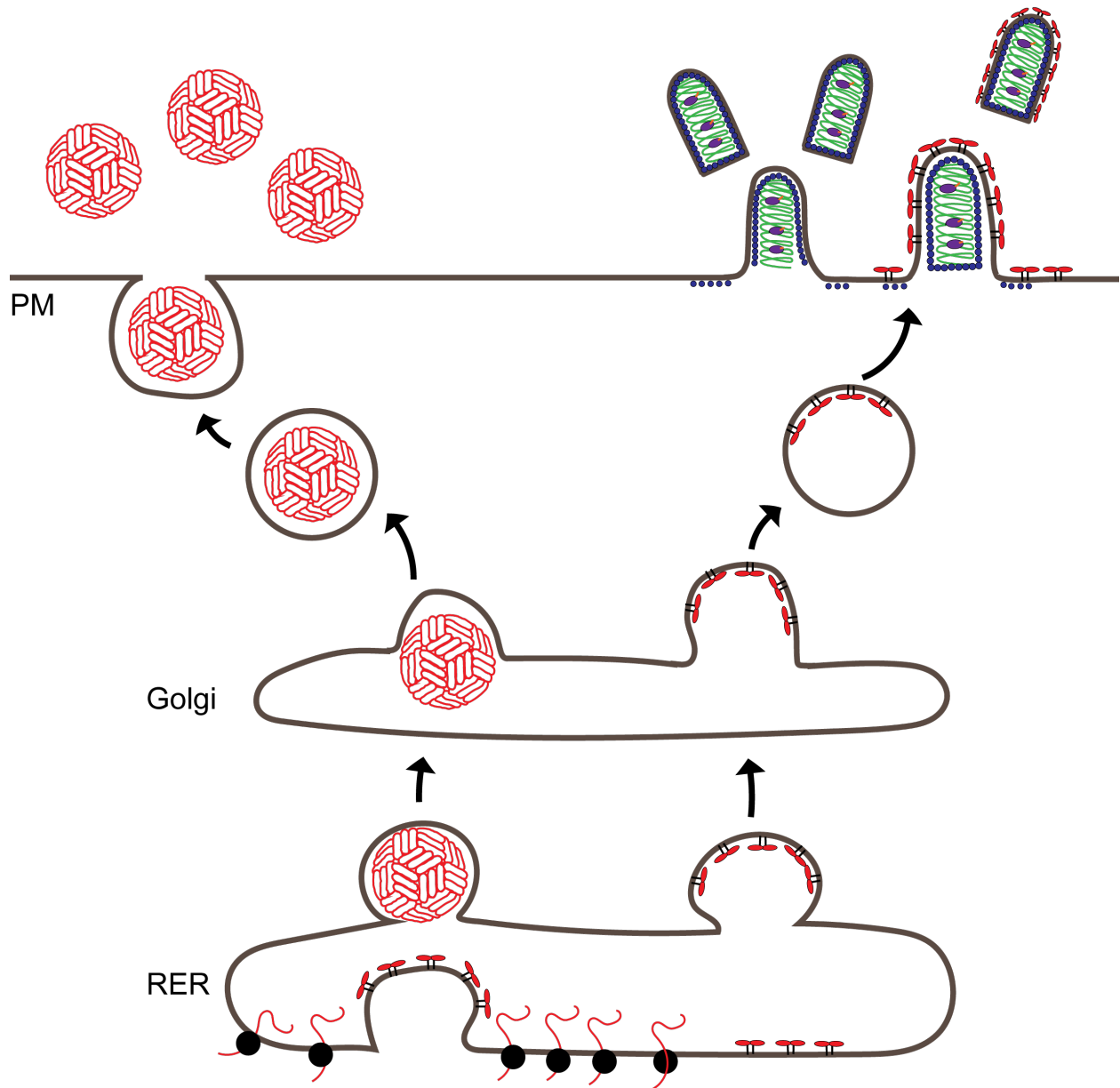
Figure 4.1: Proposed model of ZIKV-E trafficking in rVSV infected cells. During infection with rVSV-ZIKV, ZIKV-E is translated by ribosomes on the rough endoplasmic reticulum (RER) where a majority of it begins to assemble into VLPs. VLPs mature through the Golgi and are eventually exocytosed. During this time, VSV is replicating, and many non-infectious particles are released. However, some particles

Figure 4.1 (continued):

incorporate functional ZIKV-E, which suggests that a subset of ZIKV-E traffics to the

plasma membrane (PM) in VSV infected cells. ZIKV-E is shown as membrane-bound,

red proteins, and ZIKV virions are shown as red spheres. VSV particles are bullet

shaped with VSV-M in blue, nucleocapsid-RNA in green, L in purple, and P in orange.

While it is surprising that VSV incorporates functional ZIKV-E, it is possible that

the high levels of ZIKV-E protein expression during VSV infection result in the

localization of a subset ZIKV-E to the cell surface. Assessing surface-bound ZIKV-E by

flow cytometry in rVSV-ZIKV infected cells could test this. Additionally, while it is widely

demonstrated that flaviviruses assemble in ER-associated membranes, there are

reports of flavivirus budding taking place at the plasma membrane under abnormal

conditions such as during the treatment of infected cells with Brefeldin A (BFA) [228]. It

would be interesting to test if BFA treatment of rVSV-ZIKV infected cells could lead to

the targeting of more ZIKV-E to the plasma membrane. As BFA inhibits VSV

replication, this treatment would probably not significantly increase viral titers; however,

we could test how BFA treatment affects ZIKV-E distribution in VLP and VSV fractions

similarly to our assay in Figure 2.5 [229]. Other studies have detected flavivirus

envelope proteins at the plasma membrane and it was hypothesized that the

glycosylation state of the envelope protein may affect its ability to localize to the plasma

membrane [230, 231]. A panel of rVSV expressing different flavivirus prM and E

proteins could be generated and assessed for plasma membrane localization. If any

flavivirus prM and E were found to better localize to the plasma membrane, sequence differences between that flavivirus E and others could be identified and tested for their effect on cellular localization. This could result in the identification of residues that modulate the localization of flavivirus E proteins, and these could be introduced into ZIKV-E to improve incorporation into VSV. Taken together, our results provide hope that targeting of ZIKV-E to the plasma membrane can increase yields of rVSV-ΔG-ZIKV viruses.

VSV vaccine vectors

While various VSV-vectored vaccines have been tested in animal models, only two, rVSV-ZEBOV and rVSV expressing HIV-1 *gag*, have been tested in humans [40, 232]. While the rVSV-HIV1 vaccine was shown to have an acceptable safety profile in humans [232], there are concerns about adverse effects with rVSV-ZEBOV. A phase I trial in Geneva using the highest total dose of rVSV-ZEBOV had to be pulled due to 22% of participants developing arthritis [233]. Given the cytotoxicity of ZEBOV GP [234], it is possible that these adverse effects are due to ZEBOV GP and not the VSV backbone, but only further testing of VSV-vectored vaccines in humans can assess this. Despite this, the result of phase III rVSV-ZEBOV trials was mostly positive – no individuals vaccinated fell ill to Ebola virus disease [41]. These results are seemingly acceptable given the dire consequences of the Ebola virus outbreak, but costs and benefits need to be weighed for testing other VSV-vectored vaccines in humans. While we show that rVSV-CprME induces a neutralizing antibody response against ZIKV-E, a major issue is that this virus retains VSV-G in the viral genome. VSV-G facilitates virus

entry into most cell types. Additionally, VSV has been associated with severe encephalitis in humans [235]. Given that rVSV-CprME contains all of the VSV genes, it would likely have to be attenuated to reduce concerns about adverse affects to be a viable vaccine candidate in humans. VSV vectors harboring N gene rearrangements and G cytoplasmic tail truncations are attenuated and have reduced neurovirulence in animal models [236]. Therefore, CprME could be introduced into virus backbones containing these or other attenuating mutations. Additionally, CprME could be introduced into a replication competent VSV vector in which VSV-G is replaced with another relevant VFP. For example, VSV harboring chikungunya virus (CHIKV) E1 and E2 in the place of VSV-G grows to high titers [33]. Viruses expressing CHIKV and ZIKV antigens could yield titers high enough for vaccine production and stimulate an immune response against both viruses. Given that these viruses have overlapping areas of infection, this idea should be explored.

Ongoing studies aim to determine if our vaccine candidates protect against ZIKV infection using South American virus strains. As they both induce neutralizing antibodies, we predict that they both would provide immunity to ZIKV infection. As both VLPs alone and rVSV-CprME elicit similar titers of neutralizing antibodies, it would be interesting to see if there are differences in protection. Such a result would suggest differences in immune stimulation. Indeed, a DNA-based ZIKV vaccine is not as protective as an inactivated virus vaccine or an adenovirus-vectored vaccine [103]. Therefore, it could be possible that the VSV-vectored vaccine would protect better than VLPs alone. If this were the case, the costs and benefits would need to be compared.

Given the concerns with side effects in humans, and the multitude of other

vaccine platforms against ZIKV, VSV-CprME, as is, is probably not suited for use in

humans. Our studies do however provide an improved method by which to produce

flavivirus VLPs. In addition, we show the rapid development of vaccine candidates to

meet the need of an ongoing outbreak – an area in which VSV-vectored vaccines can

thrive in the future. Future studies should substantiate the human immune response to

VSV infection to better understand how to identify correlates of protection and potential

side effects of VSV infection.

VSV vectors as VLP producers

In chapter two, we show that both VLPs isolated from rVSV-CprME-infected cells

and infectious rVSV-CprME induce a neutralizing antibody response against ZIKV

MR766 (Figure 2.7C). These vaccine candidates induced similar titers of neutralizing

antibodies against ZIKV-E, which suggests that ZIKV-E expression from a VSV vector

does not lead to significant increases in ZIKV-E neutralizing antibodies compared to

non-replicating VLPs. Given that ZIKV-E neutralizing antibodies are reported to be the

main correlate of protection against ZIKV, it is reasonable that VLPs and rVSV-CprME

would elicit similar levels of protection [103]. However, this remains to be tested.

VLP vaccines have been used extensively to lower morbidity and mortality

against viruses such as hepatitis B virus and human papilloma virus [237]. There are a

number of advantages to VLP vaccines: they are morphologically similar to native

virions and thus preserve the conformation of the immunogenic protein, they do not

have to be inactivated – a process which can alter antigenicity [238], and VLPs can be

readily produced by expressing their components in cells. Our data suggest that VSV

vectors can be utilized as flavivirus VLP producers. Given that large-scale production of

VLPs is a roadblock to developing effecting VLP vaccines, VSV vectors present an

improvement over expression-based methods. VSV replicates in a wide range of cell

types, including insect cells, which are a preferred cell type in which to grow vaccines

due to reduced concerns of mammalian virus contaminants [239-241]. While

expression-based methods of VLPs require the generation of expression vectors and

materials to transfect cells, VSV replication only requires suitable cells in which to grow,

leading to reduced production costs. Therefore, VSV vectors should be considered as a

possible means for producing VLP-based vaccines. Direct comparisons of VLP yields

between VSV-based and DNA expression-based methods will be informative in

determining the benefits to using VSV to produce VLPs.

VSV-vectors to study entry mechanisms

In chapter 3, we generated a VSV recombinant bearing the VFP of an

endogenous retrovirus that is conserved in primate evolution. This allowed us to probe

the mechanism by which EnvPb1 fuses cells. We show that EnvPb1 infection is

blocked by lysosomotropic inhibitors, which implicates are role for endosomal

acidification despite EnvPb1 fusing cells at neutral pH (Figure 3.9). By exposing virions

to acid pH, we show that this presumed low pH step does not lead to triggering of

EnvPb1, thus suggesting that it plays another role in infection (Figure 3.10). These

results are seemingly in contrast with each other as EnvPb1-mediated fusion occurring

at neutral pH and rVSV-EnvPb1 resistance to low pH triggering indicate fusion at the

128

plasma membrane, while rVSV-EnvPb1 infection being blocked by lysosomotropic agents suggests receptor-mediated endocytosis.

Additionally, through our studies of EnvPb1 expression in cell culture, we show that EnvPb1 is detected in cell supernatants, even when the furin-cleavage site is mutated (Figure 3.6). Since the furin-cleavage mutant is transmembrane bound, it is likely associated with cellular-derived membranes in the supernatant. This indicates that in addition to plasma membrane-bound EnvPb1 that mediates cell-cell fusion; some EnvPb1 may be secreted in extracellular vesicles. It is then possible that EnvPb1 may function differently on vesicles than it does when present on the plasma membrane. For example, when EnvPb1 is plasma membrane bound, it might have different densities or localize differently than it does on vesicles or virions. Determining the cellular requirements for fusion will help assess how EnvPb1 functions on membranes.

EnvPb1 expression

Our most unexpected result in chapter three was the observation that EnvPb1 protein is detected in placenta tissue (Figure 3.4). This is because our analysis of *envPb1* transcripts in human tissues as well as those by Aagaard *et al.* and Blaise *et al.* suggested that *envPb1* is widely expressed in human tissues [150, 152]. However, these conclusions were derived from the analysis of total cellular RNA, which is not enriched for transcripts. This can be a confounding factor because *envPb1* lies within an intron of the cellular gene *RIN3.* It therefore raises the possibility that transcripts assigned to *envPb1* are not actually relevant. It has been shown that *envPb1* transcription is driven from its 5' LTR [154], so we looked for splice junctions

corresponding to the *envPb1* ORF and the 5' LTR in RNA-seq data from human

thymus. Despite a high density of reads within the *envPb1* ORF, we were unable to

detect any splice sites for *envPb1* suggesting that these reads do not derive from

translation competent RNA (Figure 3.5A). Alone, this would not be concerning;

however, we detect reads evenly distributed throughout *RIN3* introns when visualizing

mapped reads with the integrated genome viewer (IGV) (Figure 3.5B). These reads

align in both the sense and antisense direction throughout *RIN3* introns indicating that

they are not specific to *envPb1* mRNA. Our analysis suggests that the existing view of

envPb1 tissue expression should be re-examined. This is supported by the initial

observation that no human expressed sequence tags (ESTs), which represent mRNA,

mapped to the *envPb1* ORF [149]. Future analysis of *envPb1* transcripts should be

conducted using mRNA-enriched samples, and bona fide transcripts can be assessed

using primers that span the ERV-Pb1 5'LTR and *env* ORF. Currently, only placenta

tissue has been shown to harbor translation competent *envPb1* mRNA, although it was

significantly lower than *syncytin-1* mRNA [152]. This coincides with our observation by

western blot of human tissue samples with EnvPb1 antisera. Out of six tissue samples

tested, we only detect a band corresponding to EnvPb1 SU in placenta tissue (Figure

3.4B). While tissue samples are heterogeneous, these data suggest that EnvPb1 is not

as significantly expressed in human tissues as was previously hypothesized. Indeed, it

is difficult to rationalize how a functional fusogen could be widely expressed especially

in tissues in which no cell-cell fusions are observed unless it is either not targeted to the

cell surface or it does not localize with cells expressing its yet to be determined

receptor. Thus, our data suggest that EnvPb1 is not as widely expressed as previously reported but appears to be expressed in the placenta.

While it was previously shown that EnvPb1 is not involved of fusion of placenta choriocarcinoma cell line, BeWo, this result should not be interpreted to mean that EnvPb1 cannot be involved in placenta cell fusions [155]. Cell-cell fusions are tightly regulated events and therefore self-fusion is generally not a desirable outcome. In fact, it is most likely that BeWo cells do not express the receptor for EnvPb1 as EnvPb1 expressed in BeWo cells can lead to fusion of BeWo cells with the human umbilical vain cell line, HUVEC [155].

Future studies should expand upon our findings in the placenta. Since our tissue samples were from full-term placentas, placenta tissue at different pregnancy stages should be tested for EnvPb1 transcripts and protein. Future transcriptional profiling should be done with primers that result in the amplification of a product that spans the 5'LTR splice junction. In addition, human trophoblast stem cells were recently derived that can differentiate into three trophoblast lineages such as cytotrophoblast or syncytiotrophoblast cells [242]. The trophoblast stem cells functionally mimic primary cells, and thus could serve as a model to begin to unravel if EnvPb1 is expressed in the placenta. Assessing EnvPb1 transcription in primary placenta cells such as syncytiotrophoblasts and cytotrophoblasts could substantiate any results obtained from trophoblast stem cells. Tissues with positive signals could be profiled for protein expression using our EnvPb1 antisera.

Conservation of EnvPb1 fusion

For ERV Env proteins thought to be involved in a physiological process related to cell-cell fusion, fusion activity is often evolutionarily conserved between members of different families in the same order. Hominoid Syncytin-1 is fusogenic while in OWMs it is inactive [195]. In comparison, syncytin-2 of hominoids, OWM, and NWM is fusogenic while EnvV2, another potential syncytin, of OWM but not NWM and hominoids is fusogenic [135, 243]. This leads to a model in which hominoids retain functional syncytin-1 and syncytin-2, OWMs retain functional syncytin-2 and EnvV2, and NWMs retain functional syncytin-2. Our finding that marmoset encodes a functional EnvPb1 raises the possibility that other NWM may encode fusogenic EnvPb1. Moreover, since fusion activity is conserved in gorilla, it is possible that other hominoids retain fusogenic EnvPb1. Therefore, it is possible that EnvPb1 represents an ancestral syncytin that is specific to different primate lineages, similar to EnvV2. EnvV2 exhibits strong purifying selection over all branches [243] whereas EnvPb1 is more variable except in NWMs where all members exhibit strong evidence of purifying selection. It could be that EnvPb1 is a relevant syncytin in NWMs and that in other lineages it has either lost fusion activity or is lowly expressed. First, other primate *envPb1* ORFs should be thoroughly tested for fusogenicity. Following the identification of fusogenic primate EnvPb1s, *envPb1* tissue expression could be tested in these primates. Indeed, if EnvPb1 were found to be specifically expressed in the placenta of any other primates, *in situ* hybridization studies could assess whether it is confined to the syncytiotrophoblast boundary as would be expected for a syncytin.

Identifying the EnvPb1 receptor

132

In our studies, we developed a recombinant VSV that incorporates EnvPb1 (rVSV-EnvPb1) as a means to understand the cellular requirements for EnvPb1-mediated membrane fusion. This led to the observation that endosomal acidification is necessary for rVSV-EnvPb1 entry despite low-pH not acting as the trigger for EnvPb1 fusion (Figure 3.9 and Figure 3.10). Additionally, we used this virus to carry out forward genetic screens in mutagenized HAP1 cells. Despite the top hits of the screen residing within the heparan sulfate biosynthesis pathway, we found that heparan sulfate was not required for entry of rVSV-EnvPb1 despite the selected cells from the screen being resistant to EnvPb1-mediated entry (Figures A.1 and A.2).

Ongoing studies aim to determine the EnvPb1 receptor using different methods. We generated a recombinant VSV containing the furin-cleavage null EnvPb1, SKTR. rVSV-SKTR does not spread, but can grown in cells expressing VSV-G. Since this virus does not fuse cells, labeled virus can be used to stain cells to which it binds. In a mutagenized pool of cells, it will only stain cells that bear attachment factors, so unstained cells can be sorted to look for knockouts that block binding. Alternatively, transfection of cDNA libraries from cells fused by EnvPb1 (Vero) into non-fused cells (MDCK) could be used to assess genes that permit EnvPb1 fusion. cDNA library screens have a long history of defining retroviral Env receptors [244-246]. Identification of the EnvPb1 receptor will greatly enhance studies of EnvPb1 as a physiological fusogen. For an ERV *env* to function as a fusogen, both it and its receptor need to be localized to the proper cellular interface. Therefore, until a receptor is defined, it can only be speculated what role EnvPb1 plays.

Acknowledgements

I would first like to thank my mentor, Sean Whelan, for his involvement in my scientific training. I first spoke to Sean when he called me to invite me to interview for the program, and I was absolutely ecstatic. Since that time, Sean has been my scientific role model. I thank him for his investment in my training.

I could not have made it this far without the love and support of my parents, Steve and Karen Timpona. They have always instilled in me the idea that one can achieve any goal he sets for himself, and that lesson served me well through the ups and downs of the Ph.D.

I would also like to thank everyone who I collaborated with throughout this work – Sayantan Bose, Melissanne De Wispelaere, Lindsey Robinson-McCarthy, Kevin McCarthy, Changhong Zhou, and all of the Whelan lab members for thoughtful insights over the years.

References

1.   Emerson, S.U. and R.R. Wagner, *Dissociation and reconstitution of the transcriptase and template activities of vesicular stomatitis B and T virions.* J Virol, 1972. **10**(2): p. 297-309.

2.   Emerson, S.U. and Y. Yu, *Both NS and L proteins are required for in vitro RNA synthesis by vesicular stomatitis virus.* J Virol, 1975. **15**(6): p. 1348-56.

3.   Gupta, A.K., D. Shaji, and A.K. Banerjee, *Identification of a novel tripartite complex involved in replication of vesicular stomatitis virus genome RNA.* J Virol, 2003. **77**(1): p. 732-8.

4.   Emerson, S.U. and R.R. Wagner, *L protein requirement for in vitro RNA synthesis by vesicular stomatitis virus.* J Virol, 1973. **12**(6): p. 1325-35.

5.   Hercyk, N., S.M. Horikami, and S.A. Moyer, *The vesicular stomatitis virus L protein possesses the mRNA methyltransferase activities.* Virology, 1988. **163**(1): p. 222-5.

6.   Hunt, D.M., E.F. Smith, and D.W. Buckley, *Aberrant polyadenylation by a vesicular stomatitis virus mutant is due to an altered L protein.* J Virol, 1984. **52**(2): p. 515-21.

7.   Li, Y., L. Luo, M. Schubert, R.R. Wagner, and C.Y. Kang, *Viral liposomes released from insect cells infected with recombinant baculovirus expressing the matrix protein of vesicular stomatitis virus.* J Virol, 1993. **67**(7): p. 4415-20.

8.   Justice, P.A., W. Sun, Y. Li, Z. Ye, P.R. Grigera, and R.R. Wagner, *Membrane vesiculation function and exocytosis of wild-type and mutant matrix proteins of vesicular stomatitis virus.* J Virol, 1995. **69**(5): p. 3156-60.

9.   Black, B.L. and D.S. Lyles, *Vesicular stomatitis virus matrix protein inhibits host cell-directed transcription of target genes in vivo.* J Virol, 1992. **66**(7): p. 4058-64.

10.   Quan, B., H.S. Seo, G. Blobel, and Y. Ren, *Vesiculoviral matrix (M) protein occupies nucleic acid binding site at nucleoporin pair (Rae1 * Nup98).* Proc Natl Acad Sci U S A, 2014. **111**(25): p. 9127-32.

11.   Bishop, D.H., P. Repik, J.F. Obijeski, N.F. Moore, and R.R. Wagner, *Restitution of infectivity to spikeless vesicular stomatitis virus by solubilized viral components.* J Virol, 1975. **16**(1): p. 75-84.

12.   Ge, P., J. Tsao, S. Schein, T.J. Green, M. Luo, and Z.H. Zhou, *Cryo-EM model of the bullet-shaped vesicular stomatitis virus.* Science, 2010. **327**(5966): p. 689-93.

13.   Cureton, D.K., R.H. Massol, S. Saffarian, T.L. Kirchhausen, and S.P.J. Whelan, *Vesicular Stomatitis Virus Enters Cells through Vesicles Incompletely Coated with Clathrin That Depend upon Actin for Internalization.* PLoS Pathogens, 2009. **5**(4): p. e1000394.

14. White, J., K. Matlin, and A. Helenius, *Cell fusion by Semliki Forest, influenza, and vesicular stomatitis viruses.* The Journal of Cell Biology, 1981. **89**(3): p. 674.

15. Matlin, K.S., H. Reggio, A. Helenius, and K. Simons, *Pathway of vesicular stomatitis virus entry leading to infection.* J Mol Biol, 1982. **156**(3): p. 609-31.

16. Mire, C.E., J.M. White, and M.A. Whitt, *A Spatio-Temporal Analysis of Matrix Protein and Nucleocapsid Trafficking during Vesicular Stomatitis Virus Uncoating.* PLOS Pathogens, 2010. **6**(7): p. e1000994.

17. Ball, L.A. and C.N. White, *Order of transcription of genes of vesicular stomatitis virus.* Proceedings of the National Academy of Sciences of the United States of America, 1976. **73**(2): p. 442-446.

18. Knipe, D.M., D. Baltimore, and H.F. Lodish, *Separate pathways of maturation of the major structural proteins of vesicular stomatitis virus.* J Virol, 1977. **21**(3): p. 1128-39.

19. Swinteck, B.D. and D.S. Lyles, *Plasma membrane microdomains containing vesicular stomatitis virus M protein are separate from microdomains containing G protein and nucleocapsids.* J Virol, 2008. **82**(11): p. 5536-47.

20. Newcomb, W.W. and J.C. Brown, *Role of the vesicular stomatitis virus matrix protein in maintaining the viral nucleocapsid in the condensed form found in native virions.* J Virol, 1981. **39**(1): p. 295-9.

21. Liang, B., Z. Li, S. Jenni, A.A. Rahmeh, B.M. Morin, T. Grant, N. Grigorieff, S.C. Harrison, and S.P.J. Whelan, *Structure of the L-protein of vesicular stomatitis virus from electron cryomicroscopy.* Cell, 2015. **162**(2): p. 314-327.

22. Morin, B., A.A. Rahmeh, and S.P.J. Whelan, *Mechanism of RNA synthesis initiation by the vesicular stomatitis virus polymerase.* The EMBO Journal, 2012. **31**(5): p. 1320-1329.

23. Cureton, D.K., R.H. Massol, S.P.J. Whelan, and T. Kirchhausen, *The Length of Vesicular Stomatitis Virus Particles Dictates a Need for Actin Assembly during Clathrin-Dependent Endocytosis.* PLoS Pathogens, 2010. **6**(9): p. e1001127.

24. Lawson, N.D., E.A. Stillman, M.A. Whitt, and J.K. Rose, *Recombinant vesicular stomatitis viruses from DNA.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(10): p. 4477-4481.

25. Whelan, S.P., L.A. Ball, J.N. Barr, and G.T. Wertz, *Efficient recovery of infectious vesicular stomatitis virus entirely from cDNA clones.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(18): p. 8388-8392.

26.   Carette, J.E., M. Raaben, A.C. Wong, A.S. Herbert, G. Obernosterer, N. Mulherkar, A.I. Kuehne, P.J. Kranzusch, A.M. Griffin, G. Ruthel, P. Dal Cin, J.M. Dye, S.P. Whelan, K. Chandran, and T.R. Brummelkamp, *Ebola virus entry requires the cholesterol transporter Niemann-Pick C1.* Nature, 2011. **477**(7364): p. 340-3.

27.   Piccinotti, S., T. Kirchhausen, and S. Whelan, *Uptake of rabies virus into epithelial cells by clathrin-mediated endocytosis depends upon actin.* Journal of virology, 2013. **87**(21): p. 11637-11647.

28.   Jae, L.T., M. Raaben, M. Riemersma, E. van Beusekom, V.A. Blomen, A. Velds, R.M. Kerkhoven, J.E. Carette, H. Topaloglu, P. Meinecke, M.W. Wessels, D.J. Lefeber, S.P. Whelan, H. van Bokhoven, and T.R. Brummelkamp, *Deciphering the glycosylome of dystroglycanopathies using haploid screens for lassa virus entry.* Science, 2013. **340**(6131): p. 479-83.

29.   Raaben, M., L.T. Jae, A.S. Herbert, A.I. Kuehne, S.H. Stubbs, Y.-y. Chou, V.A. Blomen, T. Kirchhausen, J.M. Dye, T.R. Brummelkamp, and S.P. Whelan, *NRP2 and CD63 Are Host Factors for Lujo Virus Cell Entry.* Cell Host & Microbe, 2017. **22**(5): p. 688-696.e5.

30.   Robinson, L.R. and S.P.J. Whelan, *Infectious Entry Pathway Mediated by the Human Endogenous Retrovirus K Envelope Protein.* Journal of Virology, 2016. **90**(7): p. 3640-3649.

31.   Kleinfelter, L.M., R.K. Jangra, L.T. Jae, A.S. Herbert, E. Mittler, K.M. Stiles, A.S. Wirchnianski, M. Kielian, T.R. Brummelkamp, J.M. Dye, and K. Chandran, *Haploid Genetic Screen Reveals a Profound and Direct Dependence on Cholesterol for Hantavirus Membrane Fusion.* MBio, 2015. **6**(4): p. e00801.

32.   Ryder, A.B., L. Buonocore, L. Vogel, R. Nachbagauer, F. Krammer, and J.K. Rose, *A viable recombinant rhabdovirus lacking its glycoprotein gene and expressing influenza virus hemagglutinin and neuraminidase is a potent influenza vaccine.* J Virol, 2015. **89**(5): p. 2820-30.

33.   Chattopadhyay, A., E. Wang, R. Seymour, S.C. Weaver, and J.K. Rose, *A chimeric vesiculo/alphavirus is an effective alphavirus vaccine.* J Virol, 2013. **87**(1): p. 395-402.

34.   Johnson, K.M., J.E. Vogel, and P.H. Peralta, *Clinical and serological response to laboratory-acquired human infection by Indiana type vesicular stomatitis virus (VSV).* Am J Trop Med Hyg, 1966. **15**(2): p. 244-6.

35.   Hanson, R.P., A.F. Rasmussen, Jr., C.A. Brandly, and J.W. Brown, *Human infection with the virus of vesicular stomatitis.* J Lab Clin Med, 1950. **36**(5): p. 754-8.

36.   Patterson, W.C., L.O. Mott, and E.W. Jenney, *A study of vesicular stomatitis in man.* J Am Vet Med Assoc, 1958. **133**(1): p. 57-62.

37. Garbutt, M., R. Liebscher, V. Wahl-Jensen, S. Jones, P. Möller, R. Wagner, V. Volchkov, H.-D. Klenk, H. Feldmann, and U. Ströher, *Properties of Replication-Competent Vesicular Stomatitis Virus Vectors Expressing Glycoproteins of Filoviruses and Arenaviruses.* Journal of Virology, 2004. **78**(10): p. 5458-5465.

38. Jones, S.M., H. Feldmann, U. Ströher, J.B. Geisbert, L. Fernando, A. Grolla, H.-D. Klenk, N.J. Sullivan, V.E. Volchkov, E.A. Fritz, K.M. Daddario, L.E. Hensley, P.B. Jahrling, and T.W. Geisbert, *Live attenuated recombinant vaccine protects nonhuman primates against Ebola and Marburg viruses.* Nature Medicine, 2005. **11**: p. 786.

39. Regules, J.A., J.H. Beigel, K.M. Paolino, J. Voell, A.R. Castellano, Z. Hu, P. Muñoz, J.E. Moon, R.C. Ruck, J.W. Bennett, P.S. Twomey, R.L. Gutiérrez, S.A. Remich, H.R. Hack, M.L. Wisniewski, M.D. Josleyn, S.A. Kwilas, N. Van Deusen, O.T. Mbaya, Y. Zhou, D.A. Stanley, W. Jing, K.S. Smith, M. Shi, J.E. Ledgerwood, B.S. Graham, N.J. Sullivan, L.L. Jagodzinski, S.A. Peel, J.B. Alimonti, J.W. Hooper, P.M. Silvera, B.K. Martin, T.P. Monath, W.J. Ramsey, C.J. Link, H.C. Lane, N.L. Michael, R.T. Davey, and S.J. Thomas, *A Recombinant Vesicular Stomatitis Virus Ebola Vaccine.* New England Journal of Medicine, 2015. **376**(4): p. 330-341.

40. Agnandji, S.T., A. Huttner, M.E. Zinser, P. Njuguna, C. Dahlke, J.F. Fernandes, S. Yerly, J.-A. Dayer, V. Kraehling, R. Kasonta, A.A. Adegnika, M. Altfeld, F. Auderset, E.B. Bache, N. Biedenkopf, S. Borregaard, J.S. Brosnahan, R. Burrow, C. Combescure, J. Desmeules, M. Eickmann, S.K. Fehling, A. Finckh, A.R. Goncalves, M.P. Grobusch, J. Hooper, A. Jambrecina, A.L. Kabwende, G. Kaya, D. Kimani, B. Lell, B. Lemaître, A.W. Lohse, M. Massinga-Loembe, A. Matthey, B. Mordmüller, A. Nolting, C. Ogwang, M. Ramharter, J. Schmidt-Chanasit, S. Schmiedel, P. Silvera, F.R. Stahl, H.M. Staines, T. Strecker, H.C. Stubbe, B. Tsofa, S. Zaki, P. Fast, V. Moorthy, L. Kaiser, S. Krishna, S. Becker, M.-P. Kieny, P. Bejon, P.G. Kremsner, M.M. Addo, and C.-A. Siegrist, *Phase 1 Trials of rVSV Ebola Vaccine in Africa and Europe.* New England Journal of Medicine, 2015. **374**(17): p. 1647-1660.

41. Henao-Restrepo, A.M., A. Camacho, I.M. Longini, C.H. Watson, W.J. Edmunds, M. Egger, M.W. Carroll, N.E. Dean, I. Diatta, M. Doumbia, B. Draguez, S. Duraffour, G. Enwere, R. Grais, S. Gunther, P.-S. Gsell, S. Hossmann, S.V. Watle, M.K. Kondé, S. Kéïta, S. Kone, E. Kuisma, M.M. Levine, S. Mandal, T. Mauget, G. Norheim, X. Riveros, A. Soumah, S. Trelle, A.S. Vicari, J.-A. Røttingen, and M.-P. Kieny, *Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ça Suffit!).* The Lancet, 2017. **389**(10068): p. 505-518.

42. Siegel, D.P., *Energetics of intermediates in membrane fusion: comparison of stalk and inverted micellar intermediate mechanisms.* Biophysical Journal, 1993. **65**(5): p. 2124-2140.

43.   Kuzmin, P.I., J. Zimmerberg, Y.A. Chizmadzhev, and F.S. Cohen, *A quantitative model for membrane fusion based on low-energy intermediates.* Proc Natl Acad Sci U S A, 2001. **98**(13): p. 7235-40.

44.   Wilson, I.A., J.J. Skehel, and D.C. Wiley, *Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution.* Nature, 1981. **289**(5796): p. 366-73.

45.   Watowich, S.J., J.J. Skehel, and D.C. Wiley, *Crystal structures of influenza virus hemagglutinin in complex with high-affinity receptor.* Structure, 1994. **2**(8): p. 719-731.

46.   Yoshimura, A., K. Kuroda, K. Kawasaki, S. Yamashina, T. Maeda, and S. Ohnishi, *Infectious cell entry mechanism of influenza virus.* J Virol, 1982. **43**(1): p. 284-93.

47.   Bullough, P.A., F.M. Hughson, J.J. Skehel, and D.C. Wiley, *Structure of influenza haemagglutinin at the pH of membrane fusion.* Nature, 1994. **371**(6492): p. 37-43.

48.   Chen, J., J.J. Skehel, and D.C. Wiley, *N- and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA<sub>2</sub> subunit to form an N cap that terminates the triple-stranded coiled coil.* Proceedings of the National Academy of Sciences, 1999. **96**(16): p. 8967.

49.   Borrego-Diaz, E., M.E. Peeples, R.M. Markosyan, G.B. Melikyan, and F.S. Cohen, *Completion of trimeric hairpin formation of influenza virus hemagglutinin promotes fusion pore opening and enlargement.* Virology, 2003. **316**(2): p. 234-244.

50.   Danieli, T., S.L. Pelletier, Y.I. Henis, and J.M. White, *Membrane fusion mediated by the influenza virus hemagglutinin requires the concerted action of at least three hemagglutinin trimers.* J Cell Biol, 1996. **133**(3): p. 559-69.

51.   Schrodinger, LLC, *The PyMOL Molecular Graphics System, Version 1.8*. 2015.

52.   Myszka, D.G., R.W. Sweet, P. Hensley, M. Brigham-Burke, P.D. Kwong, W.A. Hendrickson, R. Wyatt, J. Sodroski, and M.L. Doyle, *Energetics of the HIV gp120-CD4 binding reaction.* Proceedings of the National Academy of Sciences, 2000. **97**(16): p. 9026.

53.   Buzon, V., G. Natrajan, D. Schibli, F. Campelo, M.M. Kozlov, and W. Weissenhorn, *Crystal Structure of HIV-1 gp41 Including Both Fusion Peptide and Membrane Proximal External Regions.* PLOS Pathogens, 2010. **6**(5): p. e1000880.

54.   Cardone, G., M. Brecher, J. Fontana, D.C. Winkler, C. Butan, J.M. White, and A.C. Steven, *Visualization of the two-step fusion process of the retrovirus avian sarcoma/leukosis virus by cryo-electron tomography.* J Virol, 2012. **86**(22): p. 12129-37.

55.   Mothes, W., A.L. Boerger, S. Narayan, J.M. Cunningham, and J.A.T. Young, *Retroviral Entry Mediated by Receptor Priming and Low pH Triggering of an Envelope Glycoprotein.* Cell, 2000. **103**(4): p. 679-689.

56.   Guirakhoo, F., F.X. Heinz, C.W. Mandl, H. Holzmann, and C. Kunz, *Fusion activity of flaviviruses: comparison of mature and immature (prM-containing) tick-borne encephalitis virions.* Journal of General Virology, 1991. **72**(6): p. 1323-1329.

57.   Stadler, K., S.L. Allison, J. Schalich, and F.X. Heinz, *Proteolytic activation of tick-borne encephalitis virus by furin.* Journal of Virology, 1997. **71**(11): p. 8475-8481.

58.   Zhang, X., P. Ge, X. Yu, J.M. Brannan, G. Bi, Q. Zhang, S. Schein, and Z.H. Zhou, *Cryo-EM structure of the mature dengue virus at 3.5-A resolution.* Nat Struct Mol Biol, 2013. **20**(1): p. 105-10.

59.   Modis, Y., S. Ogata, D. Clements, and S.C. Harrison, *Structure of the dengue virus envelope protein after membrane fusion.* Nature, 2004. **427**(6972): p. 313-9.

60.   Roche, S., F.A. Rey, Y. Gaudin, and S. Bressanelli, *Structure of the prefusion form of the vesicular stomatitis virus glycoprotein G.* Science, 2007. **315**(5813): p. 843-8.

61.   Roche, S., S. Bressanelli, F.A. Rey, and Y. Gaudin, *Crystal Structure of the Low-pH Form of the Vesicular Stomatitis Virus Glycoprotein G.* Science, 2006. **313**(5784): p. 187.

62.   Roche, S. and Y. Gaudin, *Characterization of the equilibrium between the native and fusion-inactive conformation of rabies virus glycoprotein indicates that the fusion complex is made of several trimers.* Virology, 2002. **297**(1): p. 128-35.

63.   Zimmer, G., S. Locher, M. Berger Rentsch, and S.J. Halbherr, *Pseudotyping of vesicular stomatitis virus with the envelope glycoproteins of highly pathogenic avian influenza viruses.* J Gen Virol, 2014. **95**(Pt 8): p. 1634-9.

64.   Rogalin, H.B. and E.E. Heldwein, *Characterization of Vesicular Stomatitis Virus Pseudotypes Bearing Essential Entry Glycoproteins gB, gD, gH, and gL of Herpes Simplex Virus 1.* J Virol, 2016. **90**(22): p. 10321-10328.

65.   Dick, G.W., S.F. Kitchen, and A.J. Haddow, *Zika virus. I. Isolations and serological specificity.* Trans R Soc Trop Med Hyg, 1952. **46**(5): p. 509-20.

66.   Olson, J.G., T.G. Ksiazek, Suhandiman, and Triwibowo, *Zika virus, a cause of fever in Central Java, Indonesia.* Trans R Soc Trop Med Hyg, 1981. **75**(3): p. 389-93.

67.   Fagbami, A.H., *Zika virus infections in Nigeria: virological and seroepidemiological investigations in Oyo State.* The Journal of Hygiene, 1979. **83**(2): p. 213-219.

68.   Simpson, D.I., *Zika Virus Infection in Man.* Trans R Soc Trop Med Hyg, 1964. **58**: p. 335-8.

69. Duffy, M.R., T.-H. Chen, W.T. Hancock, A.M. Powers, J.L. Kool, R.S. Lanciotti, M. Pretrick, M. Marfel, S. Holzbauer, C. Dubray, L. Guillaumot, A. Griggs, M. Bel, A.J. Lambert, J. Laven, O. Kosoy, A. Panella, B.J. Biggerstaff, M. Fischer, and E.B. Hayes, *Zika Virus Outbreak on Yap Island, Federated States of Micronesia.* New England Journal of Medicine, 2009. **360**(24): p. 2536-2543.

70. Roth, A., A. Mercier, C. Lepers, D. Hoy, S. Duituturaga, E. Benyon, L. Guillaumot, and Y. Souarès, *Concurrent outbreaks of dengue, chikungunya and Zika virus infections – an unprecedented epidemic wave of mosquito-borne viruses in the Pacific 2012–2014.* Eurosurveillance, 2014. **19**(41): p. 20929.

71. Cao-Lormeau, V.M., A. Blake, S. Mons, S. Lastere, C. Roche, J. Vanhomwegen, T. Dub, L. Baudouin, A. Teissier, P. Larre, A.L. Vial, C. Decam, V. Choumet, S.K. Halstead, H.J. Willison, L. Musset, J.C. Manuguerra, P. Despres, E. Fournier, H.P. Mallet, D. Musso, A. Fontanet, J. Neil, and F. Ghawché, *Guillain-Barré Syndrome outbreak caused by ZIKA virus infection in French Polynesia.* Lancet (London, England), 2016. **387**(10027): p. 1531-1539.

72. Organization, P.A.H.O.W.H., *Zika suspected and confirmed cases reported by countries and territories in the Americas Cumulative cases, 2015-2017. Updated as of 04 January 2018.* 2018.

73. de Araújo, T.V.B., R.A.d.A. Ximenes, D.d.B. Miranda-Filho, W.V. Souza, U.R. Montarroyos, A.P.L. de Melo, S. Valongueiro, M.d.F.P.M. de Albuquerque, C. Braga, S.P.B. Filho, M.T. Cordeiro, E. Vazquez, D.d.C.S. Cruz, C.M.P. Henriques, L.C.A. Bezerra, P.M.d.S. Castanha, R. Dhalia, E.T.A. Marques-Júnior, C.M.T. Martelli, and L.C. Rodrigues, *Association between microcephaly, Zika virus infection, and other risk factors in Brazil: final report of a case-control study.* The Lancet Infectious Diseases, 2017.

74. Salinas, J.L., D.M. Walteros, A. Styczynski, F. Garzón, H. Quijada, E. Bravo, P. Chaparro, J. Madero, J. Acosta-Reyes, J. Ledermann, Z. Arteta, E. Borland, P. Burns, M. Gonzalez, A.M. Powers, M. Mercado, A. Solano, J.J. Sejvar, and M.L. Ospina, *Zika virus disease-associated Guillain-Barré syndrome—Barranquilla, Colombia 2015–2016.* Journal of the Neurological Sciences, 2017. **381**: p. 272-277.

75. Styczynski, A.R., J.M.A.S. Malta, E.R. Krow-Lucal, J. Percio, M.E. Nóbrega, A. Vargas, T.M. Lanzieri, P.L. Leite, J.E. Staples, M.X. Fischer, A.M. Powers, G.-J.J. Chang, P.L. Burns, E.M. Borland, J.P. Ledermann, E.C. Mossel, L.B. Schonberger, E.B. Belay, J.L. Salinas, R.D. Badaro, J.J. Sejvar, and G.E. Coelho, *Increased rates of Guillain-Barré syndrome associated with Zika virus outbreak in the Salvador metropolitan area, Brazil.* PLOS Neglected Tropical Diseases, 2017. **11**(8): p. e0005869.

76. da Silva, I., J.A. Frontera, A. Bispo de Filippis, O. Nascimento, and R.I.O.G.B.S.Z.R.G. for the, *Neurologic complications associated with the zika virus in brazilian adults.* JAMA Neurology, 2017. **74**(10): p. 1190-1198.

77.   Apte-Sengupta, S., D. Sirohi, and R.J. Kuhn, *Coupling of Replication and Assembly in Flaviviruses.* Current opinion in virology, 2014. **0**: p. 134-142.

78.   Bollati, M., K. Alvarez, R. Assenberg, C. Baronti, B. Canard, S. Cook, B. Coutard, E. Decroly, X. de Lamballerie, E.A. Gould, G. Grard, J.M. Grimes, R. Hilgenfeld, A.M. Jansson, H. Malet, E.J. Mancini, E. Mastrangelo, A. Mattevi, M. Milani, G. Moureau, J. Neyts, R.J. Owens, J. Ren, B. Selisko, S. Speroni, H. Steuber, D.I. Stuart, T. Unge, and M. Bolognesi, *Structure and functionality in flavivirus NS-proteins: Perspectives for drug design.* Antiviral Research, 2010. **87**(2): p. 125-148.

79.   Zhang, Y., J. Corver, P.R. Chipman, W. Zhang, S.V. Pletnev, D. Sedlak, T.S. Baker, J.H. Strauss, R.J. Kuhn, and M.G. Rossmann, *Structures of immature flavivirus particles.* The EMBO Journal, 2003. **22**(11): p. 2604.

80.   Kuhn, R.J., W. Zhang, M.G. Rossmann, S.V. Pletnev, J. Corver, E. Lenches, C.T. Jones, S. Mukhopadhyay, P.R. Chipman, E.G. Strauss, T.S. Baker, and J.H. Strauss, *Structure of Dengue Virus.* Cell, 2002. **108**(5): p. 717-725.

81.   Zhang, X., P. Ge, X. Yu, J.M. Brannan, G. Bi, Q. Zhang, S. Schein, and Z.H. Zhou, *Cryo-EM structure of the mature dengue virus at 3.5-Å resolution.* Nature Structural &Amp; Molecular Biology, 2012. **20**: p. 105.

82.   Lorenz, I.C., J. Kartenbeck, A. Mezzacasa, S.L. Allison, F.X. Heinz, and A. Helenius, *Intracellular assembly and secretion of recombinant subviral particles from tick-borne encephalitis virus.* J Virol, 2003. **77**(7): p. 4370-82.

83.   Schalich, J., S.L. Allison, K. Stiasny, C.W. Mandl, C. Kunz, and F.X. Heinz, *Recombinant subviral particles from tick-borne encephalitis virus are fusogenic and provide a model system for studying flavivirus envelope glycoprotein functions.* J Virol, 1996. **70**(7): p. 4549-57.

84.   Corver, J., A. Ortiz, S.L. Allison, J. Schalich, F.X. Heinz, and J. Wilschut, *Membrane Fusion Activity of Tick-Borne Encephalitis Virus and Recombinant Subviral Particles in a Liposomal Model System.* Virology, 2000. **269**(1): p. 37-46.

85.   Ferlenghi, I., M. Clarke, T. Ruttan, S.L. Allison, J. Schalich, F.X. Heinz, S.C. Harrison, F.A. Rey, and S.D. Fuller, *Molecular Organization of a Recombinant Subviral Particle from Tick-Borne Encephalitis Virus.* Molecular Cell, 2001. **7**(3): p. 593-602.

86.   Mason, P.W., S. Pincus, M.J. Fournier, T.L. Mason, R.E. Shope, and E. Paoletti, *Japanese encephalitis virus-vaccinia recombinants produce particulate forms of the structural membrane proteins and induce high levels of protection against lethal JEV infection.* Virology, 1991. **180**(1): p. 294-305.

87.   Konishi, E., S. Pincus, B.A.L. Fonseca, R.E. Shope, E. Paoletti, and P.W. Mason, *Comparison of protective immunity elicited by recombinant vaccinia viruses that*

synthesize E or NS1 of Japanese encephalitis virus. Virology, 1991. **185**(1): p. 401-410.

88.   Konishi, E., S. Pincus, E. Paoletti, R.E. Shope, T. Burrage, and P.W. Mason, *Mice immunized with a subviral particle containing the japanese encephalitis virus prM/M and E proteins are protected from lethal JEV infection.* Virology, 1992. **188**(2): p. 714-720.

89.   Allison, S.L., K. Stadler, C.W. Mandl, C. Kunz, and F.X. Heinz, *Synthesis and secretion of recombinant tick-borne encephalitis virus protein E in soluble and particulate form.* Journal of Virology, 1995. **69**(9): p. 5816-5820.

90.   Fonseca, B.A.L., S. Pincus, R.E. Shope, E. Paoletti, and P.W. Mason, *Recombinant vaccinia viruses co-expressing dengue-1 glycoproteins prM and E induce neutralizing antibodies in mice.* Vaccine, 1994. **12**(3): p. 279-285.

91.   Pincus, S., P.W. Mason, E. Konishi, B.A.L. Fonseca, R.E. Shope, C.M. Rice, and E. Paoletti, *Recombinant vaccinia virus producing the prM and E proteins of yellow fever virus protects mice from lethal yellow fever encephalitis.* Virology, 1992. **187**(1): p. 290-297.

92.   Purdy, D.E. and G.-J.J. Chang, *Secretion of noninfectious dengue virus-like particles and identification of amino acids in the stem region involved in intracellular retention of envelope protein.* Virology, 2005. **333**(2): p. 239-250.

93.   Theiler, M. and H.H. Smith, *THE EFFECT OF PROLONGED CULTIVATION IN VITRO UPON THE PATHOGENICITY OF YELLOW FEVER VIRUS.* The Journal of Experimental Medicine, 1937. **65**(6): p. 767.

94.   Theiler, M. and H.H. Smith, *THE USE OF YELLOW FEVER VIRUS MODIFIED BY IN VITRO CULTIVATION FOR HUMAN IMMUNIZATION.* The Journal of Experimental Medicine, 1937. **65**(6): p. 787.

95.   Barrett, A.D.T., *Yellow fever live attenuated vaccine: A very successful live attenuated vaccine but still we have problems controlling the disease.* Vaccine, 2017. **35**(44): p. 5951-5955.

96.   World Health, O., *WHO position on the use of fractional doses – June 2017, addendum to vaccines and vaccination against yellow fever WHO: Position paper – June 2013.* Vaccine, 2017. **35**(43): p. 5751-5752.

97.   Gotuzzo, E., S. Yactayo, and E. Córdova, *Efficacy and Duration of Immunity after Yellow Fever Vaccination: Systematic Review on the Need for a Booster Every 10 Years.* The American Journal of Tropical Medicine and Hygiene, 2013. **89**(3): p. 434-444.

98.   Hadinegoro, S.R., J.L. Arredondo-García, M.R. Capeding, C. Deseda, T. Chotpitayasunondh, R. Dietze, H.I. Hj Muhammad Ismail, H. Reynales, K. Limkittikul,

D.M. Rivera-Medina, H.N. Tran, A. Bouckenooghe, D. Chansinghakul, M. Cortés, K. Fanouillere, R. Forrat, C. Frago, S. Gailhardou, N. Jackson, F. Noriega, E. Plennevaux, T.A. Wartel, B. Zambrano, and M. Saville, *Efficacy and Long-Term Safety of a Dengue Vaccine in Regions of Endemic Disease.* New England Journal of Medicine, 2015. **373**(13): p. 1195-1206.

99. Chu, H., S.L. George, D.T. Stinchcomb, J.E. Osorio, and C.D. Partidos, *CD8+ T-cell Responses in Flavivirus-Naive Individuals Following Immunization with a Live-Attenuated Tetravalent Dengue Vaccine Candidate.* J Infect Dis, 2015. **212**(10): p. 1618-28.

100. Chu, H., S.L. George, D.T. Stinchcomb, J.E. Osorio, and C.D. Partidos, *CD8+ T-cell Responses in Flavivirus-Naive Individuals Following Immunization with a Live-Attenuated Tetravalent Dengue Vaccine Candidate.* The Journal of Infectious Diseases, 2015. **212**(10): p. 1618-1628.

101. Larocca, R.A., P. Abbink, J.P.S. Peron, P.M. de A. Zanotto, M.J. Iampietro, A. Badamchi-Zadeh, M. Boyd, D. Ng'ang'a, M. Kirilova, R. Nityanandam, N.B. Mercado, Z. Li, E.T. Moseley, C.A. Bricault, E.N. Borducchi, P.B. Giglio, D. Jetton, G. Neubauer, J.P. Nkolola, L.F. Maxfield, R.A. De La Barrera, R.G. Jarman, K.H. Eckels, N.L. Michael, S.J. Thomas, and D.H. Barouch, *Vaccine protection against Zika virus from Brazil.* Nature, 2016. **536**: p. 474.

102. Abbink, P., R.A. Larocca, R.A. De La Barrera, C.A. Bricault, E.T. Moseley, M. Boyd, M. Kirilova, Z. Li, D. Ng'ang'a, O. Nanayakkara, R. Nityanandam, N.B. Mercado, E.N. Borducchi, A. Agarwal, A.L. Brinkman, C. Cabral, A. Chandrashekar, P.B. Giglio, D. Jetton, J. Jimenez, B.C. Lee, S. Mojta, K. Molloy, M. Shetty, G.H. Neubauer, K.E. Stephenson, J.P.S. Peron, P.M.d.A. Zanotto, J. Misamore, B. Finneyfrock, M.G. Lewis, G. Alter, K. Modjarrad, R.G. Jarman, K.H. Eckels, N.L. Michael, S.J. Thomas, and D.H. Barouch, *Protective efficacy of multiple vaccine platforms against Zika virus challenge in rhesus monkeys.* Science, 2016. **353**(6304): p. 1129.

103. Abbink, P., R.A. Larocca, K. Visitsunthorn, M. Boyd, R.A. De La Barrera, G.D. Gromowski, M. Kirilova, R. Peterson, Z. Li, O. Nanayakkara, R. Nityanandam, N.B. Mercado, E.N. Borducchi, A. Chandrashekar, D. Jetton, S. Mojta, P. Gandhi, J. LeSuer, S. Khatiwada, M.G. Lewis, K. Modjarrad, R.G. Jarman, K.H. Eckels, S.J. Thomas, N.L. Michael, and D.H. Barouch, *Durability and correlates of vaccine protection against Zika virus in rhesus monkeys.* Science Translational Medicine, 2017. **9**(420).

104. Chahal, J.S., T. Fang, A.W. Woodham, O.F. Khan, J. Ling, D.G. Anderson, and H.L. Ploegh, *An RNA nanoparticle vaccine against Zika virus elicits antibody and CD8+ T cell responses in a mouse model.* Scientific Reports, 2017. **7**(1): p. 252.

105. Richner, J.M., S. Himansu, K.A. Dowd, S.L. Butler, V. Salazar, J.M. Fox, J.G. Julander, W.W. Tang, S. Shresta, T.C. Pierson, G. Ciaramella, and M.S. Diamond, *Modified mRNA Vaccines Protect against Zika Virus Infection.* Cell, 2017. **169**(1): p. 176.

106. Espinosa, D., J. Mendy, D. Manayani, L. Vang, C. Wang, T. Richard, B. Guenther, J. Aruri, J. Avanzini, F. Garduno, P. Farness, M. Gurwith, J. Smith, E. Harris, and J. Alexander, *Passive Transfer of Immune Sera Induced by a Zika Virus-Like Particle Vaccine Protects AG129 Mice Against Lethal Zika Virus Challenge.* EBioMedicine, 2018. **27**: p. 61-70.

107. Betancourt, D., N.M.G.P. de Queiroz, T. Xia, J. Ahn, and G.N. Barber, *Cutting Edge: Innate Immune Augmenting Vesicular Stomatitis Virus Expressing Zika Virus Proteins Confers Protective Immunity.* The Journal of Immunology, 2017. **198**(8): p. 3023.

108. Baltimore, D., *Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of RNA Tumour Viruses.* Nature, 1970. **226**: p. 1209.

109. Katz, R.A., G. Merkel, J. Kulkosky, J. Leis, and A.M. Skalka, *The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro.* Cell, 1990. **63**(1): p. 87-95.

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

111. Jaenisch, R., *Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus.* Proceedings of the National Academy of Sciences of the United States of America, 1976. **73**(4): p. 1260-1264.

112. Salter, D.W., E.J. Smith, S.H. Hughes, S.E. Wright, and L.B. Crittenden, *Transgenic chickens: Insertion of retroviral genes into the chicken germ line.* Virology, 1987. **157**(1): p. 236-240.

113. Lander, E.S., L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J.P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J.C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R.H. Waterston, R.K. Wilson, L.W. Hillier, J.D. McPherson, M.A. Marra, E.R. Mardis, L.A. Fulton, A.T. Chinwalla, K.H. Pepin, W.R. Gish, S.L. Chissoe, M.C. Wendl, K.D. Delehaunty, T.L. Miner, A. Delehaunty, J.B. Kramer, L.L. Cook, R.S. Fulton, D.L. Johnson, P.J. Minx, S.W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R.A. Gibbs, D.M. Muzny, S.E. Scherer, J.B. Bouck, E.J. Sodergren, K.C. Worley, C.M. Rives, J.H. Gorrell, M.L. Metzker, S.L. Naylor, R.S. Kucherlapati, D.L. Nelson, G.M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P.

Wincker, D.R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H.M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R.W. Davis, N.A. Federspiel, A.P. Abola, M.J. Proctor, R.M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D.R. Cox, M.V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G.A. Evans, M. Athanasiou, R. Schultz, B.A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W.R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J.A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D.G. Brown, C.B. Burge, L. Cerutti, H.C. Chen, D. Church, M. Clamp, R.R. Copley, T. Doerks, S.R. Eddy, E.E. Eichler, T.S. Furey, J. Galagan, J.G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L.S. Johnson, T.A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W.J. Kent, P. Kitts, E.V. Koonin, I. Korf, D. Kulp, D. Lancet, T.M. Lowe, A. McLysaght, T. Mikkelsen, J.V. Moran, N. Mulder, V.J. Pollara, C.P. Ponting, G. Schuler, J. Schultz, G. Slater, A.F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y.I. Wolf, K.H. Wolfe, S.P. Yang, R.F. Yeh, F. Collins, M.S. Guyer, J. Peterson, A. Felsenfeld, K.A. Wetterstrand, A. Patrinos, M.J. Morgan, P. de Jong, J.J. Catanese, K. Osoegawa, H. Shizuya, S. Choi and Y.J. Chen, *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.

114. Mouse Genome Sequencing, C., *Initial sequencing and comparative analysis of the mouse genome.* Nature, 2002. **420**: p. 520.

115. Pačes, J., A. Pavlíček, and V. Pačes, *HERVd: database of human endogenous retroviruses.* Nucleic Acids Research, 2002. **30**(1): p. 205-206.

116. Belshaw, R., A. Katzourakis, J. Pačes, A. Burt, and M. Tristem, *High Copy Number in Human Endogenous Retrovirus Families is Associated with Copying Mechanisms in Addition to Reinfection.* Molecular Biology and Evolution, 2005. **22**(4): p. 814-817.

117. Belshaw, R., V. Pereira, A. Katzourakis, G. Talbot, J. Pačes, A. Burt, and M. Tristem, *Long-term reinfection of the human genome by endogenous retroviruses.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(14): p. 4894-4899.

118. Xiong, Y. and T.H. Eickbush, *Origin and evolution of retroelements based upon their reverse transcriptase sequences.* Embo j, 1990. **9**(10): p. 3353-62.

119. Dimmic, M.W., J.S. Rest, D.P. Mindell, and R.A. Goldstein, *rtREV: an amino acid substitution matrix for inference of retrovirus and reverse transcriptase phylogeny.* J Mol Evol, 2002. **55**(1): p. 65-73.

120. Kumar, S., G. Stecher, and K. Tamura, *MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.* Mol Biol Evol, 2016. **33**(7): p. 1870-4.

121. Cornelis, G., M. Funk, C. Vernochet, F. Leal, O.A. Tarazona, G. Meurice, O. Heidmann, A. Dupressoir, A. Miralles, M.P. Ramirez-Pinilla, and T. Heidmann, *An endogenous*

retroviral envelope syncytin and its cognate receptor identified in the viviparous placental Mabuya lizard. Proceedings of the National Academy of Sciences, 2017. **114**(51): p. E10991.

122. Bénit, L., P. Dessen, and T. Heidmann, *Identification, Phylogeny, and Evolution of Retroviral Elements Based on Their Envelope Genes.* Journal of Virology, 2001. **75**(23): p. 11709-11719.

123. Maksakova, I.A., D.L. Mager, and D. Reiss, *Endogenous retroviruses.* Cellular and Molecular Life Sciences, 2008. **65**(21): p. 3329-3347.

124. Groudine, M., R. Eisenman, and H. Weintraub, *Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation.* Nature, 1981. **292**(5821): p. 311-7.

125. Brattås, P.L., M.E. Jönsson, L. Fasching, J. Nelander Wahlestedt, M. Shahsavani, R. Falk, A. Falk, P. Jern, M. Parmar, and J. Jakobsson, *TRIM28 Controls a Gene Regulatory Network Based on Endogenous Retroviruses in Human Neural Progenitor Cells.* Cell Reports, 2017. **18**(1): p. 1-11.

126. Martins, H. and P. Villesen, *Improved Integration Time Estimation of Endogenous Retroviruses with Phylogenetic Data.* PLOS ONE, 2011. **6**(3): p. e14745.

127. 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, 2011. **8**(1): p. 90.

128. Barbulescu, M., G. Turner, M.I. Seaman, A.S. Deinard, K.K. Kidd, and J. Lenz, *Many human endogenous retrovirus K (HERV-K) proviruses are unique to humans.* Curr Biol, 1999. **9**(16): p. 861-8.

129. Tristem, M., *Identification and characterization of novel human endogenous retrovirus families by phylogenetic screening of the human genome mapping project database.* J Virol, 2000. **74**(8): p. 3715-30.

130. Belshaw, R., J. Watson, A. Katzourakis, A. Howe, J. Woolven-Allen, A. Burt, and M. Tristem, *Rate of recombinational deletion among human endogenous retroviruses.* J Virol, 2007. **81**(17): p. 9437-42.

131. Stankiewicz, P. and J.R. Lupski, *Genome architecture, rearrangements and genomic disorders.* Trends in Genetics, 2002. **18**(2): p. 74-82.

132. Frank, J.A. and C. Feschotte, *Co-option of endogenous viral sequences for host cell function.* Curr Opin Virol, 2017. **25**: p. 81-89.

133. Mi, S., X. Lee, X.-p. Li, G.M. Veldman, H. Finnerty, L. Racie, E. LaVallie, X.-Y. Tang, P. Edouard, S. Howes, J.C. Keith Jr, and J.M. McCoy, *Syncytin is a captive retroviral*

envelope protein involved in human placental morphogenesis. Nature, 2000. **403**: p. 785.

134.   Blond, J.-L., F. Besème, L. Duret, O. Bouton, F. Bedin, H. Perron, B. Mandrand, and F. Mallet, *Molecular Characterization and Placental Expression of HERV-W, a New Human Endogenous Retrovirus Family.* Journal of Virology, 1999. **73**(2): p. 1175-1185.

135.   Blaise, S., N. de Parseval, L. Bénit, and T. Heidmann, *Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution.* Proceedings of the National Academy of Sciences, 2003. **100**(22): p. 13013.

136.   de Parseval, N., G. Diop, S. Blaise, F. Helle, A. Vasilescu, F. Matsuda, and T. Heidmann, *Comprehensive search for intra- and inter-specific sequence polymorphisms among coding envelope genes of retroviral origin found in the human genome: genes and pseudogenes.* BMC Genomics, 2005. **6**(1): p. 117.

137.   Dupressoir, A., C. Vernochet, F. Harper, J. Guegan, P. Dessen, G. Pierron, and T. Heidmann, *A pair of co-opted retroviral envelope syncytin genes is required for formation of the two-layered murine placental syncytiotrophoblast.* Proc Natl Acad Sci U S A, 2011. **108**(46): p. E1164-73.

138.   Cornelis, G., O. Heidmann, S.A. Degrelle, C. Vernochet, C. Lavialle, C. Letzelter, S. Bernard-Stoecklin, A. Hassanin, B. Mulot, M. Guillomot, I. Hue, T. Heidmann, and A. Dupressoir, *Captured retroviral envelope syncytin gene associated with the unique placental structure of higher ruminants.* Proc Natl Acad Sci U S A, 2013. **110**(9): p. E828-37.

139.   Cornelis, G., C. Vernochet, Q. Carradec, S. Souquere, B. Mulot, F. Catzeflis, M.A. Nilsson, B.R. Menzies, M.B. Renfree, G. Pierron, U. Zeller, O. Heidmann, A. Dupressoir, and T. Heidmann, *Retroviral envelope gene captures and syncytin exaptation for placentation in marsupials.* Proc Natl Acad Sci U S A, 2015. **112**(5): p. E487-96.

140.   Dupressoir, A., C. Vernochet, O. Bawa, F. Harper, G. Pierron, P. Opolon, and T. Heidmann, *Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene.* Proceedings of the National Academy of Sciences, 2009. **106**(29): p. 12127.

141.   Mathes, L.E., R.G. Olsen, L.C. Hebebrand, E.A. Hoover, J.P. Schaller, P.W. Adams, and W.S. Nichols, *Immunosuppressive Properties of a Virion Polypeptide, a 15,000-Dalton Protein, from Feline Leukemia Virus.* Cancer Research, 1979. **39**(3): p. 950.

142.   Hebebrand, L.C., R.G. Olsen, L.E. Mathes, and W.S. Nichols, *Inhibition of Human Lymphocyte Mitogen and Antigen Response by a 15,000-Dalton Protein from Feline Leukemia Virus.* Cancer Research, 1979. **39**(2 Part 1): p. 443.

143. Cianciolo, G.J., T.D. Copeland, S. Oroszlan, and R. Snyderman, *Inhibition of lymphocyte proliferation by a synthetic peptide homologous to retroviral envelope proteins.* Science, 1985. **230**(4724): p. 453.

144. Blaise, S., M. Mangeney, and T. Heidmann, *The envelope of Mason–Pfizer monkey virus has immunosuppressive properties.* Journal of General Virology, 2001. **82**(7): p. 1597-1600.

145. Mangeney, M., M. Renard, G. Schlecht-Louf, I. Bouallaga, O. Heidmann, C. Letzelter, A. Richaud, B. Ducos, and T. Heidmann, *Placental syncytins: Genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins.* Proceedings of the National Academy of Sciences, 2007. **104**(51): p. 20534.

146. Heidmann, O., A. Béguin, J. Paternina, R. Berthier, M. Deloger, O. Bawa, and T. Heidmann, *HEMO, an ancestral endogenous retroviral envelope protein shed in the blood of pregnant women and expressed in pluripotent stem cells and tumors.* Proceedings of the National Academy of Sciences, 2017. **114**(32): p. E6642.

147. Blanco-Melo, D., R.J. Gifford, and P.D. Bieniasz, *Co-option of an endogenous retrovirus envelope for host defense in hominid ancestors.* eLife, 2017. **6**: p. e22519.

148. Taylor, G.M., Y. Gao, and D.A. Sanders, *Fv-4: identification of the defect in Env and the mechanism of resistance to ecotropic murine leukemia virus.* J Virol, 2001. **75**(22): p. 11244-8.

149. Villesen, P., L. Aagaard, C. Wiuf, and F.S. Pedersen, *Identification of endogenous retroviral reading frames in the human genome.* Retrovirology, 2004. **1**: p. 32.

150. Blaise, S., N. de Parseval, and T. Heidmann, *Functional characterization of two newly identified Human Endogenous Retrovirus coding envelope genes.* Retrovirology, 2005. **2**: p. 19.

151. Moller, S., M.D. Croning, and R. Apweiler, *Evaluation of methods for the prediction of membrane spanning regions.* Bioinformatics, 2001. **17**(7): p. 646-53.

152. Aagaard, L., P. Villesen, A.L. Kjeldbjerg, and F.S. Pedersen, *The approximately 30-million-year-old ERVPb1 envelope gene is evolutionarily conserved among hominoids and Old World monkeys.* Genomics, 2005. **86**(6): p. 685-91.

153. Wilkins, M.R., E. Gasteiger, A. Bairoch, J.C. Sanchez, K.L. Williams, R.D. Appel, and D.F. Hochstrasser, *Protein identification and analysis tools in the ExPASy server.* Methods Mol Biol, 1999. **112**: p. 531-52.

154. Vargas, A., M. Thiery, J. Lafond, and B. Barbeau, *Transcriptional and functional studies of Human Endogenous Retrovirus envelope EnvP(b) and EnvV genes in human trophoblasts.* Virology, 2012. **425**(1): p. 1-10.

155. Aagaard, L., B. Bjerregaard, A.L. Kjeldbjerg, F.S. Pedersen, L.-I. Larsson, and J.J. Rossi, *Silencing of endogenous envelope genes in human choriocarcinoma cells shows that envPb1 is involved in heterotypic cell fusions.* The Journal of General Virology, 2012. **93**(Pt 8): p. 1696-1699.

156. Søe, K., T.L. Andersen, A.-S. Hobolt-Pedersen, B. Bjerregaard, L.-I. Larsson, and J.-M. Delaissé, *Involvement of human endogenous retroviral syncytin-1 in human osteoclast fusion.* Bone, 2011. **48**(4): p. 837-846.

157. Møller, A.M.J., J.-M. Delaissé, and K. Søe, *Osteoclast Fusion: Time-Lapse Reveals Involvement of CD47 and Syncytin-1 at Different Stages of Nuclearity.* Journal of Cellular Physiology, 2017. **232**(6): p. 1396-1403.

158. Frese, S., M. Ruebner, F. Suhr, T.M. Konou, K.A. Tappe, M. Toigo, H.H. Jung, C. Henke, R. Steigleder, P.L. Strissel, H. Huebner, M.W. Beckmann, P. van der Keylen, B. Schoser, T. Schiffer, L. Frese, W. Bloch, and R. Strick, *Long-Term Endurance Exercise in Humans Stimulates Cell Fusion of Myoblasts along with Fusogenic Endogenous Retroviral Genes In Vivo.* PLOS ONE, 2015. **10**(7): p. e0132099.

159. Organization, P.A.H.O.W.H., *Zika suspected and confirmed cases reported by countries and territories in the Americas Cumulative cases, 2015-2017. Updated as of 04 January 2018.* 2018, PAHO/WHO: Washington, D.C.

160. Rasmussen, S.A., D.J. Jamieson, M.A. Honein, and L.R. Petersen, *Zika Virus and Birth Defects — Reviewing the Evidence for Causality.* New England Journal of Medicine, 2016. **374**(20): p. 1981-1987.

161. Henao-Restrepo, A.M., A. Camacho, I.M. Longini, C.H. Watson, W.J. Edmunds, M. Egger, M.W. Carroll, N.E. Dean, I. Diatta, M. Doumbia, B. Draguez, S. Duraffour, G. Enwere, R. Grais, S. Gunther, P.S. Gsell, S. Hossmann, S.V. Watle, M.K. Konde, S. Keita, S. Kone, E. Kuisma, M.M. Levine, S. Mandal, T. Mauget, G. Norheim, X. Riveros, A. Soumah, S. Trelle, A.S. Vicari, J.A. Rottingen, and M.P. Kieny, *Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ca Suffit!).* Lancet, 2017. **389**(10068): p. 505-518.

162. Enfissi, A., J. Codrington, J. Roosblad, M. Kazanji, and D. Rousset, *Zika virus genome from the Americas.* The Lancet. **387**(10015): p. 227-228.

163. Taylor, T.J., F. Diaz, R.C. Colgrove, K.A. Bernard, N.A. DeLuca, S.P.J. Whelan, and D.M. Knipe, *Production of immunogenic West Nile virus-like particles using a herpes simplex virus 1 recombinant vector.* Virology, 2016. **496**: p. 186-193.

164. Zhang, G., V. Gurtu, and S.R. Kain, *An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells.* Biochem Biophys Res Commun, 1996. **227**(3): p. 707-11.

165.   Lobigs, M., E. Lee, M.L. Ng, M. Pavy, and P. Lobigs, *A flavivirus signal peptide balances the catalytic activity of two proteases and thereby facilitates virus morphogenesis.* Virology, 2010. **401**(1): p. 80-89.

166.   Tani, H., M. Shiokawa, Y. Kaname, H. Kambara, Y. Mori, T. Abe, K. Moriishi, and Y. Matsuura, *Involvement of ceramide in the propagation of Japanese encephalitis virus.* J Virol, 2010. **84**(6): p. 2798-807.

167.   Schmidt, A.G., K. Lee, P.L. Yang, and S.C. Harrison, *Small-Molecule Inhibitors of Dengue-Virus Entry.* PLoS Pathogens, 2012. **8**(4): p. e1002627.

168.   Lefrancois, L. and D.S. Lyles, *The interaction of antibody with the major surface glycoprotein of vesicular stomatitis virus I. Analysis of neutralizing epitopes with monoclonal antibodies.* Virology, 1982. **121**(1): p. 157-167.

169.   Schnitzer, T.J., C. Dickson, and R.A. Weiss, *Morphological and biochemical characterization of viral particles produced by the tsO45 mutant of vesicular stomatitis virus at restrictive temperature.* Journal of Virology, 1979. **29**(1): p. 185-195.

170.   Zhao, H., E. Fernandez, K.A. Dowd, S.D. Speer, D.J. Platt, M.J. Gorman, J. Govero, C.A. Nelson, T.C. Pierson, M.S. Diamond, and D.H. Fremont, *Structural basis of Zika virus specific antibody protection.* Cell, 2016. **166**(4): p. 1016-1027.

171.   Larocca, R.A., P. Abbink, J.P.S. Peron, P.M. de A. Zanotto, M.J. Iampietro, A. Badamchi-Zadeh, M. Boyd, D. Ng'ang'a, M. Kirilova, R. Nityanandam, N.B. Mercado, Z. Li, E.T. Moseley, C.A. Bricault, E.N. Borducchi, P.B. Giglio, D. Jetton, G. Neubauer, J.P. Nkolola, L.F. Maxfield, R.A. De La Barrera, R.G. Jarman, K.H. Eckels, N.L. Michael, S.J. Thomas, and D.H. Barouch, *Vaccine protection against Zika virus from Brazil.* Nature, 2016. **536**(7617): p. 474-478.

172.   Boigard, H., A. Alimova, G.R. Martin, A. Katz, P. Gottlieb, and J.M. Galarza, *Zika virus-like particle (VLP) based vaccine.* PLOS Neglected Tropical Diseases, 2017. **11**(5): p. e0005608.

173.   Zhang, R., J.J. Miner, M.J. Gorman, K. Rausch, H. Ramage, J.P. White, A. Zuiani, P. Zhang, E. Fernandez, Q. Zhang, K.A. Dowd, T.C. Pierson, S. Cherry, and M.S. Diamond, *A CRISPR screen defines a signal peptide processing pathway required by flaviviruses.* Nature, 2016. **535**: p. 164.

174.   Buchholz, U.J., S. Finke, and K.-K. Conzelmann, *Generation of Bovine Respiratory Syncytial Virus (BRSV) from cDNA: BRSV NS2 Is Not Essential for Virus Replication in Tissue Culture, and the Human RSV Leader Region Acts as a Functional BRSV Genome Promoter.* Journal of Virology, 1999. **73**(1): p. 251-259.

175.   Wong, A.C., R.G. Sandesara, N. Mulherkar, S.P. Whelan, and K. Chandran, *A Forward Genetic Strategy Reveals Destabilizing Mutations in the Ebolavirus Glycoprotein That*

Alter Its Protease Dependence during Cell Entry. Journal of Virology, 2010. **84**(1): p. 163-175.

176. Whelan, S.P.J., J.N. Barr, and G.W. Wertz, *Identification of a Minimal Size Requirement for Termination of Vesicular Stomatitis Virus mRNA: Implications for the Mechanism of Transcription.* Journal of Virology, 2000. **74**(18): p. 8268-8276.

177. Lazear, H.M., J. Govero, A.M. Smith, D.J. Platt, E. Fernandez, J.J. Miner, and M.S. Diamond, *A Mouse Model of Zika Virus Pathogenesis.* Cell host & microbe, 2016. **19**(5): p. 720-730.

178. Lyles, D.S., L. Puddington, and B.J. McCreedy, *Vesicular stomatitis virus M protein in the nuclei of infected cells.* Journal of Virology, 1988. **62**(11): p. 4387-4392.

179. Blaise, S., N. de Parseval, L. Bénit, and T. Heidmann, *Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(22): p. 13013-13018.

180. Lavialle, C., G. Cornelis, A. Dupressoir, C. Esnault, O. Heidmann, C. Vernochet, and T. Heidmann, *Paleovirology of 'syncytins', retroviral env genes exapted for a role in placentation.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**(1626): p. 20120507.

181. Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman, *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.

182. Kent, W.J., C.W. Sugnet, T.S. Furey, K.M. Roskin, T.H. Pringle, A.M. Zahler, and D. Haussler, *The human genome browser at UCSC.* Genome Res, 2002. **12**(6): p. 996-1006.

183. Notredame, C., D.G. Higgins, and J. Heringa, *T-Coffee: A novel method for fast and accurate multiple sequence alignment.* J Mol Biol, 2000. **302**(1): p. 205-17.

184. Pinter, A., R. Kopelman, Z. Li, S.C. Kayman, and D.A. Sanders, *Localization of the labile disulfide bond between SU and TM of the murine leukemia virus envelope protein complex to a highly conserved CWLC motif in SU that resembles the active-site sequence of thiol-disulfide exchange enzymes.* J Virol, 1997. **71**(10): p. 8073-7.

185. Opstelten, D.J., M. Wallin, and H. Garoff, *Moloney murine leukemia virus envelope protein subunits, gp70 and Pr15E, form a stable disulfide-linked complex.* J Virol, 1998. **72**(8): p. 6537-45.

186. Robert, X. and P. Gouet, *Deciphering key features in protein structures with the new ENDscript server.* Nucleic Acids Res, 2014. **42**(Web Server issue): p. W320-4.

187. Schrago, C.G. and C.A.M. Russo, *Timing the Origin of New World Monkeys.* Molecular Biology and Evolution, 2003. **20**(10): p. 1620-1625.

188. Goodman, M., C.A. Porter, J. Czelusniak, S.L. Page, H. Schneider, J. Shoshani, G. Gunnell, and C.P. Groves, *Toward a phylogenetic classification of primates based on DNA evidence complemented by fossil evidence.* Molecular phylogenetics and evolution, 1998. **9**(3): p. 585-598.

189. Janečka, J.E., W. Miller, T.H. Pringle, F. Wiens, A. Zitzmann, K.M. Helgen, M.S. Springer, and W.J. Murphy, *Molecular and Genomic Data Identify the Closest Living Relative of Primates.* Science, 2007. **318**(5851): p. 792.

190. Sela, I., H. Ashkenazy, K. Katoh, and T. Pupko, *GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters.* Nucleic Acids Res, 2015. **43**(W1): p. W7-14.

191. Katoh, K. and D.M. Standley, *MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability.* Molecular Biology and Evolution, 2013. **30**(4): p. 772-780.

192. Huelsenbeck, J.P. and K.A. Crandall, *Phylogeny Estimation and Hypothesis Testing Using Maximum Likelihood.* Annual Review of Ecology and Systematics, 1997. **28**(1): p. 437-466.

193. Elnitski, L., R. Burhans, C. Riemer, R. Hardison, and W. Miller, *MultiPipMaker: a comparative alignment server for multiple DNA sequences.* Curr Protoc Bioinformatics, 2010. **Chapter 10**: p. Unit10.4.

194. Schwartz, S., W.J. Kent, A. Smit, Z. Zhang, R. Baertsch, R.C. Hardison, D. Haussler, and W. Miller, *Human-mouse alignments with BLASTZ.* Genome Res, 2003. **13**(1): p. 103-7.

195. Cáceres, M. and J.W. Thomas, *The Gene of Retroviral Origin Syncytin 1 is Specific to Hominoids and is Inactive in Old World Monkeys.* Journal of Heredity, 2006. **97**(2): p. 100-106.

196. Yang, Z., *PAML: a program package for phylogenetic analysis by maximum likelihood.* (0266-7061 (Print)).

197. McBee, R.M., S.A. Rozmiarek, N.R. Meyerson, P.A. Rowley, and S.L. Sawyer, *The effect of species representation on the detection of positive selection in primate gene data sets.* Mol Biol Evol, 2015. **32**(4): p. 1091-6.

198. Bustin, S.A. and T. Nolan, *Pitfalls of Quantitative Real-Time Reverse-Transcription Polymerase Chain Reaction.* Journal of Biomolecular Techniques : JBT, 2004. **15**(3): p. 155-166.

199.   Duff, M.O., S. Olson, X. Wei, S.C. Garrett, A. Osman, M. Bolisetty, A. Plocik, S.E. Celniker, and B.R. Graveley, *Genome-wide identification of zero nucleotide recursive splicing in Drosophila.* Nature, 2015. **521**(7552): p. 376-9.

200.   Duff, M.O., S. Olson, X. Wei, S.C. Garrett, A. Osman, M. Bolisetty, A. Plocik, S.E. Celniker, and B.R. Graveley, *Genome-wide identification of zero nucleotide recursive splicing in Drosophila.* Nature, 2015. **521**: p. 376.

201.   Kim, D., B. Langmead, and S.L. Salzberg, *HISAT: a fast spliced aligner with low memory requirements.* Nature methods, 2015. **12**(4): p. 357-360.

202.   Duckert, P., S. Brunak, and N. Blom, *Prediction of proprotein convertase cleavage sites.* Protein Eng Des Sel, 2004. **17**(1): p. 107-12.

203.   Prada, I. and J. Meldolesi, *Binding and Fusion of Extracellular Vesicles to the Plasma Membrane of Their Cell Targets.* International Journal of Molecular Sciences, 2016. **17**(8): p. 1296.

204.   Christodoulopoulos, I. and P.M. Cannon, *Sequences in the cytoplasmic tail of the gibbon ape leukemia virus envelope protein that prevent its incorporation into lentivirus vectors.* J Virol, 2001. **75**(9): p. 4129-38.

205.   Johnson, J.E., W. Rodgers, and J.K. Rose, *A plasma membrane localization signal in the HIV-1 envelope cytoplasmic domain prevents localization at sites of vesicular stomatitis virus budding and incorporation into VSV virions.* Virology, 1998. **251**(2): p. 244-52.

206.   Johnson, J.E., M.J. Schnell, L. Buonocore, and J.K. Rose, *Specific targeting to CD4+ cells of recombinant vesicular stomatitis viruses encoding human immunodeficiency virus envelope proteins.* J Virol, 1997. **71**(7): p. 5060-8.

207.   Carneiro, F.A., A.S. Ferradosa, and A.T. Da Poian, *Low pH-induced conformational changes in vesicular stomatitis virus glycoprotein involve dramatic structure reorganization.* J Biol Chem, 2001. **276**(1): p. 62-7.

208.   Bissonnette, M.L., S.A. Connolly, D.F. Young, R.E. Randall, R.G. Paterson, and R.A. Lamb, *Analysis of the pH requirement for membrane fusion of different isolates of the paramyxovirus parainfluenza virus 5.* J Virol, 2006. **80**(6): p. 3071-7.

209.   Mueller, Daniela S., T. Kampmann, R. Yennamalli, Paul R. Young, B. Kobe, and Alan E. Mark, *Histidine protonation and the activation of viral fusion proteins.* Biochemical Society Transactions, 2008. **36**(1): p. 43.

210.   Gaudin, Y., *Reversibility in Fusion Protein Conformational Changes The Intriguing Case of Rhabdovirus-Induced Membrane Fusion*, in *Fusion of Biological Membranes and Related Problems*, H. Hilderson and S. Fuller, Editors. 2002, Springer US: Boston, MA. p. 379-408.

211. Gaudin, Y., R.W. Ruigrok, M. Knossow, and A. Flamand, *Low-pH conformational changes of rabies virus glycoprotein and their role in membrane fusion.* J Virol, 1993. **67**(3): p. 1365-72.

212. Record, M., *Intercellular communication by exosomes in placenta: A possible role in cell fusion?* Placenta, 2014. **35**(5): p. 297-302.

213. Vargas, A., S. Zhou, M. Éthier-Chiasson, D. Flipo, J. Lafond, C. Gilbert, and B. Barbeau, *Syncytin proteins incorporated in placenta exosomes are important for cell uptake and show variation in abundance in serum exosomes from patients with preeclampsia.* The FASEB Journal, 2014. **28**(8): p. 3703-3719.

214. Andersson, B.S., V.P. Collins, R. Kurzrock, D.W. Larkin, C. Childs, A. Ost, A. Cork, J.M. Trujillo, E.J. Freireich, M.J. Siciliano, and et al., *KBM-7, a human myeloid leukemia cell line with double Philadelphia chromosomes lacking normal c-ABL and BCR transcripts.* Leukemia, 1995. **9**(12): p. 2100-8.

215. Chao, L.H., D.E. Klein, A.G. Schmidt, J.M. Pena, and S.C. Harrison, *Sequential conformational rearrangements in flavivirus membrane fusion.* Elife, 2014. **3**: p. e04389.

216. Tamura, K. and M. Nei, *Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees.* Mol Biol Evol, 1993. **10**(3): p. 512-26.

217. Fass, D., R.A. Davey, C.A. Hamson, P.S. Kim, J.M. Cunningham, and J.M. Berger, *Structure of a murine leukemia virus receptor-binding glycoprotein at 2.0 angstrom resolution.* Science, 1997. **277**(5332): p. 1662-6.

218. Barnett, A.L., D.L. Wensel, W. Li, D. Fass, and J.M. Cunningham, *Structure and mechanism of a coreceptor for infection by a pathogenic feline retrovirus.* J Virol, 2003. **77**(4): p. 2717-29.

219. Jae, L.T., M. Raaben, A.S. Herbert, A.I. Kuehne, A.S. Wirchnianski, T.K. Soh, S.H. Stubbs, H. Janssen, M. Damme, P. Saftig, S.P. Whelan, J.M. Dye, and T.R. Brummelkamp, *Virus entry. Lassa virus entry requires a trigger-induced receptor switch.* Science, 2014. **344**(6191): p. 1506-10.

220. Brown, K.S., D. Safronetz, A. Marzi, H. Ebihara, and H. Feldmann, *Vesicular stomatitis virus-based vaccine protects hamsters against lethal challenge with Andes virus.* J Virol, 2011. **85**(23): p. 12781-91.

221. Zhang, R., C.F. Hryc, Y. Cong, X. Liu, J. Jakana, R. Gorchakov, M.L. Baker, S.C. Weaver, and W. Chiu, *4.4 A cryo-EM structure of an enveloped alphavirus Venezuelan equine encephalitis virus.* Embo j, 2011. **30**(18): p. 3854-63.

222. Freiberg, A.N., M.B. Sherman, M.C. Morais, M.R. Holbrook, and S.J. Watowich, *Three-dimensional organization of Rift Valley fever virus revealed by cryoelectron tomography.* J Virol, 2008. **82**(21): p. 10341-8.

223. Hasan, S.S., M. Sevvana, R.J. Kuhn, and M.G. Rossmann, *Structural biology of Zika virus and other flaviviruses.* Nature Structural & Molecular Biology, 2018. **25**(1): p. 13-20.

224. Beniac, D.R. and T.F. Booth, *Structure of the Ebola virus glycoprotein spike within the virion envelope at 11 Å resolution.* Scientific Reports, 2017. **7**: p. 46374.

225. Suomalainen, M., P. Liljestrom, and H. Garoff, *Spike protein-nucleocapsid interactions drive the budding of alphaviruses.* J Virol, 1992. **66**(8): p. 4737-47.

226. Goldsmith, C.S., L.H. Elliott, C.J. Peters, and S.R. Zaki, *Ultrastructural characteristics of Sin Nombre virus, causative agent of hantavirus pulmonary syndrome.* Arch Virol, 1995. **140**(12): p. 2107-22.

227. Cifuentes-Munoz, N., G.P. Barriga, P.D. Valenzuela, and N.D. Tischler, *Aromatic and polar residues spanning the candidate fusion peptide of the Andes virus Gc protein are essential for membrane fusion and infection.* J Gen Virol, 2011. **92**(Pt 3): p. 552-63.

228. Ng, M.L., V. Howe J Fau - Sreenivasan, J.J. Sreenivasan V Fau - Mulders, and J.J. Mulders, *Flavivirus West Nile (Sarafend) egress at the plasma membrane.* (0304-8608 (Print)).

229. Irurzun, A., L. Pérez, and L. Carrasco, *Brefeldin A blocks protein glycosylation and RNA replication of vesicular stomatitis virus.* FEBS Letters, 2001. **336**(3): p. 496-500.

230. Wright, P.J., *Envelope protein of the flavivirus Kunjin is apparently not glycosylated.* J Gen Virol, 1982. **59**(Pt 1): p. 29-38.

231. Hase, T., P.L. Summers, K.H. Eckels, and W.B. Baze, *An electron and immunoelectron microscopic study of dengue-2 virus infection of cultured mosquito cells: maturation events.* Arch Virol, 1987. **92**(3-4): p. 273-91.

232. Fuchs, J.D., I. Frank, M.L. Elizaga, M. Allen, N. Frahm, N. Kochar, S. Li, S. Edupuganti, S.A. Kalams, G.D. Tomaras, R. Sheets, M. Pensiero, M.A. Tremblay, T.J. Higgins, T. Latham, M.A. Egan, D.K. Clarke, J.H. Eldridge, M. Mulligan, N. Rouphael, S. Estep, K. Rybczyk, D. Dunbar, S. Buchbinder, T. Wagner, R. Isbell, V. Chinnell, J. Bae, G. Escamilla, J. Tseng, R. Fair, S. Ramirez, G. Broder, L. Briesemeister, and A. Ferrara, *First-in-Human Evaluation of the Safety and Immunogenicity of a Recombinant Vesicular Stomatitis Virus Human Immunodeficiency Virus-1 gag Vaccine (HVTN 090).* Open Forum Infectious Diseases, 2015. **2**(3): p. ofv082-ofv082.

233. Agnandji, S.T., A. Huttner, M.E. Zinser, P. Njuguna, C. Dahlke, J.F. Fernandes, S. Yerly, J.A. Dayer, V. Kraehling, R. Kasonta, A.A. Adegnika, M. Altfeld, F. Auderset, E.B. Bache,

N. Biedenkopf, S. Borregaard, J.S. Brosnahan, R. Burrow, C. Combescure, J. Desmeules, M. Eickmann, S.K. Fehling, A. Finckh, A.R. Goncalves, M.P. Grobusch, J. Hooper, A. Jambrecina, A.L. Kabwende, G. Kaya, D. Kimani, B. Lell, B. Lemaitre, A.W. Lohse, M. Massinga-Loembe, A. Matthey, B. Mordmuller, A. Nolting, C. Ogwang, M. Ramharter, J. Schmidt-Chanasit, S. Schmiedel, P. Silvera, F.R. Stahl, H.M. Staines, T. Strecker, H.C. Stubbe, B. Tsofa, S. Zaki, P. Fast, V. Moorthy, L. Kaiser, S. Krishna, S. Becker, M.P. Kieny, P. Bejon, P.G. Kremsner, M.M. Addo, and C.A. Siegrist, *Phase 1 Trials of rVSV Ebola Vaccine in Africa and Europe.* N Engl J Med, 2016. **374**(17): p. 1647-60.

234. Yang, Z.Y., H.J. Duckers, N.J. Sullivan, A. Sanchez, E.G. Nabel, and G.J. Nabel, *Identification of the Ebola virus glycoprotein as the main viral determinant of vascular cell cytotoxicity and injury.* Nat Med, 2000. **6**(8): p. 886-9.

235. Quiroz, E., N. Moreno, P.H. Peralta, and R.B. Tesh, *A human case of encephalitis associated with vesicular stomatitis virus (Indiana serotype) infection.* Am J Trop Med Hyg, 1988. **39**(3): p. 312-4.

236. Clarke, D.K., F. Nasar, M. Lee, J.E. Johnson, K. Wright, P. Calderon, M. Guo, R. Natuk, D. Cooper, R.M. Hendry, and S.A. Udem, *Synergistic attenuation of vesicular stomatitis virus by combination of specific G gene truncations and N gene translocations.* J Virol, 2007. **81**(4): p. 2056-64.

237. Wang, J.W. and R.B.S. Roden, *Virus-like particles for the prevention of human papillomavirus-associated malignancies.* Expert review of vaccines, 2013. **12**(2): p. 10.1586/erv.12.151.

238. Fan, Y.-C., H.-C. Chiu, L.-K. Chen, G.-J.J. Chang, and S.-S. Chiou, *Formalin Inactivation of Japanese Encephalitis Virus Vaccine Alters the Antigenicity and Immunogenicity of a Neutralization Epitope in Envelope Protein Domain III.* PLOS Neglected Tropical Diseases, 2015. **9**(10): p. e0004167.

239. Wyers, F., C. Richard-Molard, D. Blondel, and S. Dezelee, *Vesicular stomatitis virus growth in Drosophila melanogaster cells: G protein deficiency.* J Virol, 1980. **33**(1): p. 411-22.

240. Zárate, S. and I.S. Novella, *Vesicular Stomatitis Virus Evolution during Alternation between Persistent Infection in Insect Cells and Acute Infection in Mammalian Cells Is Dominated by the Persistence Phase.* Journal of Virology, 2004. **78**(22): p. 12236-12242.

241. Dalemans, W., *Insect cells as a new substrate for vaccine production.* Dev Biol (Basel), 2006. **123**: p. 235-41; discussion 265-6.

242. Okae, H., H. Toh, T. Sato, H. Hiura, S. Takahashi, K. Shirane, Y. Kabayama, M. Suyama, H. Sasaki, and T. Arima, *Derivation of Human Trophoblast Stem Cells.* Cell Stem Cell, 2018. **22**(1): p. 50-63.e6.

243. Esnault, C., G. Cornelis, O. Heidmann, and T. Heidmann, *Differential Evolutionary Fate of an Ancestral Primate Endogenous Retrovirus Envelope Gene, the EnvV Syncytin, Captured for a Function in Placentation.* PLOS Genetics, 2013. **9**(3): p. e1003400.

244. Bacquin, A., C. Bireau, M. Tanguy, C. Romanet, C. Vernochet, A. Dupressoir, and T. Heidmann, *A Cell Fusion-Based Screening Method Identifies Glycosylphosphatidylinositol-Anchored Protein Ly6e as the Receptor for Mouse Endogenous Retroviral Envelope Syncytin-A.* J Virol, 2017. **91**(18).

245. Tailor, C.S., A. Nouri, Y. Zhao, Y. Takeuchi, and D. Kabat, *A sodium-dependent neutral-amino-acid transporter mediates infections of feline and baboon endogenous retroviruses and simian type D retroviruses.* J Virol, 1999. **73**(5): p. 4470-4.

246. Quigley, J.G., C.C. Burns, M.M. Anderson, E.D. Lynch, K.M. Sabo, J. Overbaugh, and J.L. Abkowitz, *Cloning of the cellular receptor for feline leukemia virus subgroup C (FeLV-C), a retrovirus that induces red cell aplasia.* Blood, 2000. **95**(3): p. 1093-9.

247. Raaben, M., L.T. Jae, A.S. Herbert, A.I. Kuehne, S.H. Stubbs, Y.-y. Chou, V.A. Blomen, T. Kirchhausen, J.M. Dye, T.R. Brummelkamp, and S.P. Whelan, *NRP2 and CD63 Are Host Factors for Lujo Virus Cell Entry.* Cell Host & Microbe. **22**(5): p. 688-696.e5.

248. Carette, J.E., C.P. Guimaraes, M. Varadarajan, A.S. Park, I. Wuethrich, A. Godarova, M. Kotecki, B.H. Cochran, E. Spooner, H.L. Ploegh, and T.R. Brummelkamp, *Haploid genetic screens in human cells identify host factors used by pathogens.* Science, 2009. **326**(5957): p. 1231-5.

249. Sarrazin, S., W.C. Lamanna, and J.D. Esko, *Heparan Sulfate Proteoglycans.* Cold Spring Harbor Perspectives in Biology, 2011. **3**(7): p. a004952.

250. Busse, M., A. Feta, J. Presto, M. Wilen, M. Gronning, L. Kjellen, and M. Kusche-Gullberg, *Contribution of EXT1, EXT2, and EXTL3 to heparan sulfate chain elongation.* J Biol Chem, 2007. **282**(45): p. 32802-10.

251. O'Hearn, A., M. Wang, H. Cheng, C.M. Lear-Rooney, K. Koning, E. Rumschlag-Booms, E. Varhegyi, G. Olinger, and L. Rong, *Role of EXT1 and Glycosaminoglycans in the Early Stage of Filovirus Entry.* J Virol, 2015. **89**(10): p. 5441-9.

252. Bai, X., G. Wei, A. Sinha, and J.D. Esko, *Chinese hamster ovary cell mutants defective in glycosaminoglycan assembly and glucuronosyltransferase I.* J Biol Chem, 1999. **274**(19): p. 13017-24.

253. Carette, J.E., C.P. Guimaraes, I. Wuethrich, V.A. Blomen, M. Varadarajan, C. Sun, G. Bell, B. Yuan, M.K. Muellner, S.M. Nijman, H.L. Ploegh, and T.R. Brummelkamp, *Global gene disruption in human cells to assign genes to phenotypes by deep sequencing.* Nature Biotechnology, 2011. **29**: p. 542.

254.   Langmead, B., C. Trapnell, M. Pop, and S.L. Salzberg, *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biol, 2009. **10**(3): p. R25.

Appendix: Haploid Screens to Identify EnvPb1 Entry Factors

Introduction

The EnvPb1 receptor is currently unknown. Determining the EnvPb1 receptor well help define the tropism and mechanism of fusion of EnvPb1, which will aid in understanding its putative physiological role. Additionally, other host proteins may be necessary for the function of this envelope protein. Therefore, we set out to define the cellular requirements for EnvPb1 entry by performing a forward genetic screen using recombinant VSV expressing EnvPb1.

Our lab has previously applied haploid genetic screens to identify cellular entry factors for Ebola virus, Lassa fever virus, and Lujo virus [26, 28, 219, 247]. The human cell line, HAP1, contains one copy of each chromosome; as such, insertional mutagenesis throughout the genome can lead to knockout of non-essential genes [248]. Infecting HAP1 cells with a retroviral gene trap vector creates a mutagenized library of knockout cells. These cells can then be subject to a selective pressure such as infection with a virus. Since VSV lyses infected cells, it will kill susceptible HAP1 cells. However, if a cell harbors a mutation in a gene that is necessary for infection by the virus, that cell will be resistant to infection and thus persist. Sequencing of the gene trap insertion site in surviving cells reveals genes that result in resistance to infection with the recombinant VSV. Since the envelope protein dictates the route of internalization, these genes are likely to play a role in entry.

Results

To yield a significant number of hits, approximately 50 million cells need to be infected at an MOI that would kill nearly all unmutagenized cells – roughly an MOI of 3.

162

Since titers of rVSV-envPb1 did not rise above 10^6 pfu/ml even after attempts at concentrating the virus, we had to devise another method of infecting the HAP1 cells at a high MOI with virus bearing EnvPb1 on the surface. To do so, we employed rVSV-EnvPb1 that was passaged on VSV-G expressing cells and then treated with the VSV-G-neutralizing antibody IE2 [168]. As shown in Figure 3.7B, IE2 significantly reduces VSV infectivity but IE2 has little effect rVSV-EnvPb1+G suggesting that it blocks VSV-G-mediated infection of this virus and permits entry through EnvPb1.

From this, we tested different dilutions of IE2-neutralized rVSV-EnvPb1+G to find an optimal amount that would lead to >99% killing of unmutagenized HAP1 cells. We calculated this MOI to be approximately 4 based on the titer of rVSV-EnvPb1+G in the presence of IE2. We then infected 50 million mutagenized HAP1 cells at this MOI. The virus IE2 mixture was allowed to bind and enter cells for 1 hour, after which the inoculum was removed, cells were washed, and fresh cell culture media was added. At 24 hours post infection, it was observed that most cells were dead, with a maximum of one surviving cell per field of view at 10X magnification. Dead cells were removed, and fresh cell culture media was added. Surviving cells grew until colonies were visible at 8 days post infection. At this point, cells were pooled and re-infected with IE2-neutralized rVSV-envPb1+G at an MOI of 3. Again, there was a significant amount of cell killing. However, there were more resistant cells visible per field of view during this round of infection. These surviving cells were pooled at 24 hours post infection and infected with rVSV-EnvPb1. We were able to infect these surviving cells with rVSV-EnvPb1 at an MOI of 3 to ensure that the final selection was done in the absence of both VSV-G and IE2S since there were fewer surviving cells at this stage than the 50 million initially

infected. Many of the cells survived this round of infection, indicating that they harbor

mutations necessary for EnvPb1-mediated entry into cells. These cells were expanded,

and genomic DNA was extracted for sequencing of the gene trap insertion sites. A

schematic of this screening method is shown in Figure A.1A.
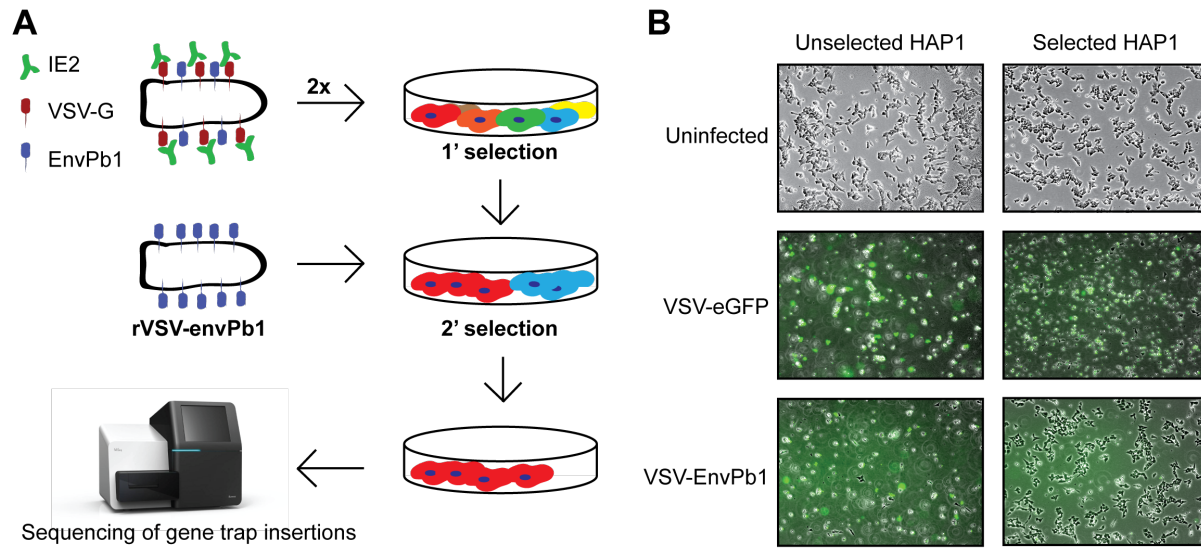
Figure A.1: Haploid genetic screens for rVSV-EnvPb1. (A) Schematic of screening approach for rVSV-EnvPb1. (B) Final selected cells are specifically resistant to rVSV-EnvPb1. Unselected or selected cells were plated at a density of 1 million cells and infected with virus at an MOI of 3. Images shown are at 24 HPI. GFP+ cells represent infected cells. (C) Top hits in haploid screen with rVSV-EnvPb1. The size of each bubble corresponds to the number of insertions in the gene. The significance

165

score was calculated by taking the $-\log_{10}$ of the p-value generated from a one-sided

Fisher exact test. Hits with a significance score greater than 2 are plotted.

To ensure that these cells were specifically resistant to rVSV-EnvPb1-mediated

entry, the pooled cells from the final selection were subject to infection by either rVSV-

eGFP or rVSV-EnvPb1 at an MOI of 3. Unselected, mutagenized HAP1 cells were

included as a control. Infection was allowed to proceed for 24 hours to allow for virus

propagation and cell killing. The selected cells were completely resistant to infection

with rVSV-EnvPb1 over the 24-hour infection period while the majority of unselected

cells were killed by the same amount of virus (Figure A.1B). Furthermore, rVSV-eGFP

kills the majority of the selected cells indicating that most of these cells are specifically

resistant to EnvPb1-mediated entry. Interestingly, not all of the selected cells were

killed by VSV-eGFP, suggesting that a small subset of cells may harbor mutations in

genes that are involved with post-entry steps of VSV infection.

Haploid genetic screen implicates heparan sulfate biosynthesis as a major

component of rVSV-EnvPb1 entry

Reads returned from the sequencing of rVSV-EnvPb1 selected cells were

trimmed, collapsed, and aligned to hg18. Aligned reads that mapped to genes were

extracted. The number of unique reads per gene was counted, and the enrichment for

a gene in the selected set was calculated by comparing the frequency of insertions in

the selected data set to the number of insertions in an unselected data set. A p-value

166

was calculated for each gene by using a one-sided Fisher exact test. These data were

then intersected with RNA-seq data from unselected, mutagenized HAP1 cells, and only

the genes that had an RPKM > 1 (expressed genes) were retained. These data were

graphed using a bubble plot where the size of the circle corresponds to the number of

unique insertions and the y-axis represents significance (negative log of p-value)

(Figure A.1C, Table A.1).   Eight of the top hits are known genes in the heparan sulfate

biosynthesis pathway, and another, TM9SF2, is likely involved in that pathway.

Heparan sulfate (HS) is a glycosaminoglycan that is attached to proteins to form

heparan sulfate proteoglycans. Heparan sulfate proteoglycans reside at the cell surface

and extracellular matrix and have been implicated as attachment factors for a variety of

viruses [249]. Other hits include genes encoding proteins involved in Golgi transport

(COGs). Since the most significant hits are involved in heparan sulfate biosynthesis, we

decided to probe the necessity of HS in rVSV-EnvPb1 entry by using HS knockout cell

lines.

Table A.1: List of top hits in rVSV-EnvPb1 haploid screen. Top hits were defined as having a significance score greater than 3 and RPKM values greater than 1. Gene names are listed as well as the number of inactivating insertions in the selected and unselected data sets.

Gene	Selected	Unselected	p-value	Significance score	RPKM
SLC35B2	211	8	0	Inf	19.65408391
EXTL3	164	63	9.67E-227	226.0146923	6.403240487
EXT1	159	113	1.11E-196	195.9538485	2.759342339
NDST1	116	15	1.79E-181	180.7473412	12.33811068
TM9SF2	96	20	3.25E-144	143.4878065	31.77455642
COG7	29	15	1.67E-39	38.77816851	3.012904875
EXT2	28	26	5.90E-34	33.22945772	11.13775144
C3orf58	27	43	2.22E-28	27.6534973	9.906785206
COG5	25	75	4.65E-21	20.33301426	4.848367108
B4GALT7	12	2	1.75E-19	18.75761094	2.818949406
PTAR1	20	47	7.84E-19	18.10587344	12.10783844
XYLT2	8	10	5.84E-10	9.233696933	4.496599013
UBE2V2	8	24	1.11E-07	6.955092091	13.93279467
ATP6V0A2	7	24	1.48E-06	5.829197457	3.738649273
UNC50	6	20	7.41E-06	5.130048638	9.779555957
RPP40	3	1	2.63E-05	4.579939691	3.923081451
VAMP3	3	2	6.48E-05	4.18816999	26.20231787
COG6	4	9	7.84E-05	4.105893489	2.71515634
SYNJ2	6	34	9.86E-05	4.00606408	3.05433795
CYP2S1	3	3	0.000127849	3.893304336	4.52314638
LDLR	6	38	0.000170071	3.769369895	4.671750563
TMEM187	8	77	0.000209515	3.678785897	5.351866084
UBE2O	6	41	0.000246542	3.60810967	14.62562589
EDEM1	6	46	0.000431472	3.365046915	8.564972664
INSR	8	87	0.000448457	3.348278828	4.821572584
IFT52	4	16	0.000477876	3.320684677	8.962774975
EHMT1	4	17	0.000581518	3.235436465	4.872304139
RRP12	3	7	0.000724852	3.139750422	11.10883302
NOP2	4	20	0.000987028	3.005670454	22.6735861
ZNF326	4	20	0.000987028	3.005670454	8.393479909

rVSV-EnvPb1 infects heparan sulfate knockout cells

We first tested the ability of rVSV-EnvPb1 to infect *ext1*^{KO} HAP1 cells. Ext1 is involved in the chain elongation step of heparan sulfate biosynthesis and was a top hit in our screen [250]. rVSV-EnvPb1 and control viruses were absorbed onto *ext1*^{KO} HAP1 cells for one hour at 37ºC and cells were assayed for virus entry at 6 HPI by flow cytometry. As expected, rVSV-EBOV, which uses heparan sulfate to attach to cells, minimally infected these cells, and rVSV-eGFP was unaffected (Figure A.2A) [251].
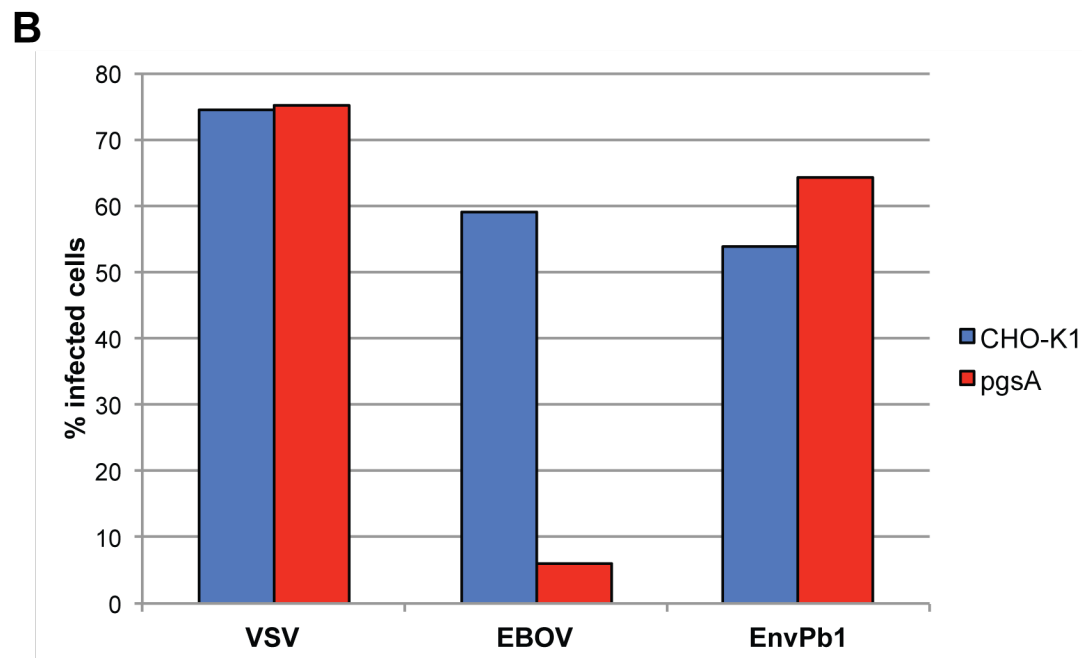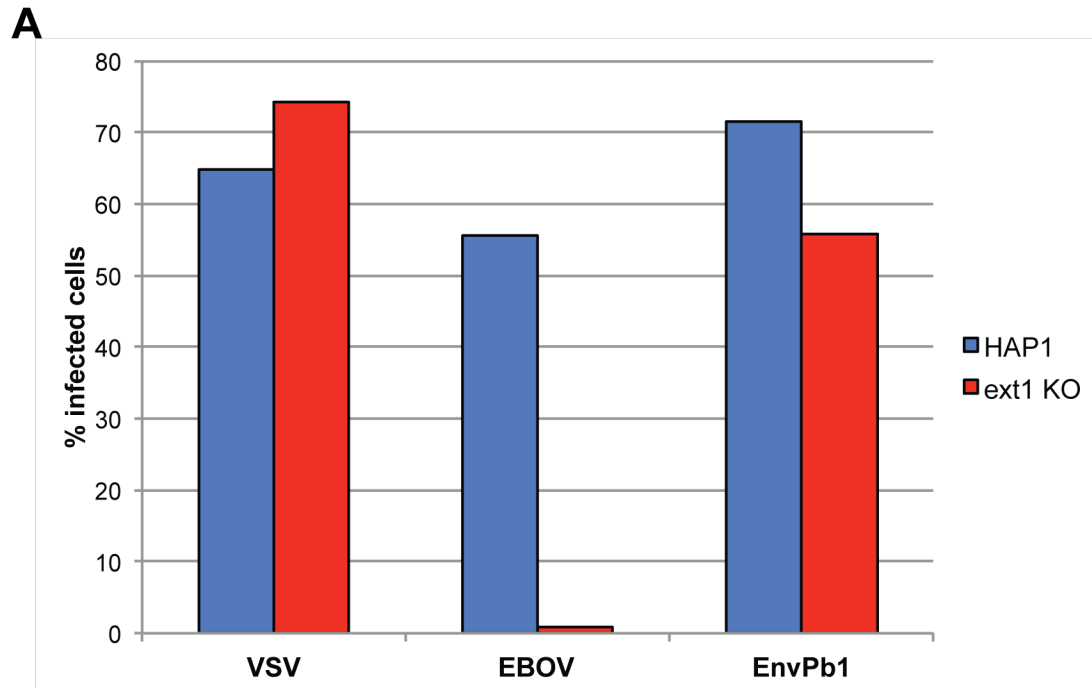
Figure A.2: HS is not required for rVSV-EnvPb1 infection. (A) rVSV-EnvPb1 infects heparan sulfate knockout Hap1 cells. VSV-EBOV serves as a positive control for heparan sulfate-dependent entry. Virus at an MOI=1 was absorbed onto wild-type HAP1 or *ext1*[KO] HAP1 cells for 1 hour at 37ºC. GFP expression was confirmed by

Figure 3.12 (continued):

fluorescence microscopy at 6 HPI and cells were fixed and analyzed for GFP

expression by flow cytometry. %infection represents the percent of cells that are GFP+.

(B) rVSV-EnvPb1 infects CHO cells lacking glycosaminoglycans. Virus at an MOI=1

was absorbed onto CHO-K1 or glycosaminoglycan lacking CHO pgsA cells for 1 hour at

37ºC. GFP expression was confirmed by fluorescence microscopy at 6 HPI and cells

were fixed and analyzed for GFP expression by flow cytometry. %infection represents

the percent of cells that are GFP+.

However, there was only a minimal decrease in infectivity with rVSV-EnvPb1.

This result is in contrast to the screen data. Given the result, we wanted to test if this

was specific to Ext1 or if knockout of other aspects of heparan sulfate biosynthesis

would validate our screen results. Thus, we tested infectivity of rVSV-EnvPb1 in CHO

pgsA-745 cells, which are deficient in xylosyltransferase and thus have no

glycosaminoglycans [252]. Infectivity of rVSV-EnvPb1 in these cells was comparable to

infectivity in WT CHO cells at the same MOI indicating that EnvPb1-mediated entry can

occur in the absence of heparan sulfate (Figure A.2B). Whereas the screen analysis

strongly implicates heparan sulfate, these experiments suggest that it is not important

for rVSV-EnvPb1 entry. We cannot disregard the possibility that HS plays a role distinct

from attachment in EnvPb1-mediated entry; however, we did not test this given the aims

of our screen. Therefore, the EnvPb1 receptor is still unknown, and ongoing work is

being directed to its identification.

Discussion

We aimed to test what cellular proteins are involved in EnvPb1-mediated cell fusion events. Haploid genetic screens with rVSV-EnvPb1 identified the heparan sulfate biosynthesis pathway as being important for rVSV-EnvPb1 entry. However, we show that rVSV-EnvPb1 infects $ext1^{KO}$ cells as well as cells lacking glycosaminoglycans. This suggests that heparan sulfate is not a major entry factor for rVSV-EnvPb1. It is possible that heparan sulfate does play a role in EnvPb1 entry, but in the absence of heparan sulfate, EnvPb1 can bind and enter cells through a different mechanism. Another possibility is that the rVSV-EnvPb1-resistant cells in the haploid screen contained multiple insertions that skewed the interpretation of our results. This is supported by the observation that only 22,903 of 1,861,242 collapsed reads map to genes in the rVSV-EnvPb1-selected data set (1.23%). Further, in the unselected data set, only 637,817 of 16,494,868 collapsed reads map to genes (3.87%), with the rest presumably residing outside of genes. Conversely, in the previously published VSV-LASV screen, 17.67% and 37.58% of reads in the selected and unselected data sets respectively mapped to genes [28]. These data suggest that the screened HAP1 cells were not thoroughly mutagenized, had multiple insertions per cell, or a combination of both. Given that some of the lower significance hits had few insertions in the selected data set, it is likely that some aspect of the screen was not saturating and thus decreased the likelihood of identifying relevant genes. Therefore, these results should be viewed as preliminary, and other methods should be explored to identify the EnvPb1 receptor.

Methods

Haploid screens

Gene trap virus was produced in 293T cells and concentrated as previously described [26]. HAP1 cells were infected with the gene trap virus by spin infection, and infected cells were expanded to yield sufficient cells counts for multiple screens. Genomic DNA was isolated from a portion of these mutagenized cells to create a control dataset. For the screen, 50 million mutagenized cells were exposed to IE2-neutralized rVSV-EnvPb1+G at an MOI ~4. The resistant colonies were expanded to ~10 million cells and reinfected at an MOI of 3. These resistant cells were expanded to ~ 5 million cells and infected with rVSV-EnvPb1. Surviving cells were expanded to ~50 million cells and used for genomic DNA isolation.

Sequencing and analysis of insertion sites

Genomic DNA was isolated using the QIamp DNA mini kit (Qiagen) and processed as previously described [253]. PCR-amplified gene trap insertion sites were sequenced using Illumina HiSeq 2000 at the Whitehead Institute (Cambridge, MA). Reads were trimmed and sorted by quality score using the Galaxy server (usegalaxy.org) and aligned to human genome build hg18 using Bowtie alignment software [254]. Aligned reads were intersected with genes, and insertions that occurred in the positive sense of an intron or in either direction of an exon were taken as inactivating insertions. Total inactivating insertions per gene were counted. Inactivating insertions per gene from the rVSV-EnvPb1 selected dataset were compared to inactivating insertions per gene in the unselected data set, and a p-value was calculated using a one-sided Fischer exact test

in R (www.r-project.org). Significance scores were calculated by taking the negative log

of the p-value. The significance score and number of insertions per gene were used to

generate bubble plots in Microsoft Excel.