Electronic Thesis and Dissertation Repository

3-24-2023 1:00 PM

Molecular Characterization of the Antiviral Properties of the Small HERC Family of Proteins

Ermela Paparisto, *The University of Western Ontario*

Supervisor: Barr, Stephen D, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Microbiology and Immunology
© Ermela Paparisto 2023

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Immunology of Infectious Disease Commons, and the Virology Commons

Abstract

Although viruses are obligate intracellular parasites, they have their own evolutionary trajectory, their genomes are in a constant battle to overcome the defenses of the host. This thesis investigates the role of the small HERC family of proteins in the battle against two deadly viruses: Human Immunodeficiency Virus -1 (HIV) and Zaire ebolavirus (EBOV). Although their discovery occurred decades ago, little knowledge is available about the small HERC family, their functions, and modes of interactions with other cellular proteins. In the first chapter, the structural evolution of the small HERC family and related functional changes that have occurred over time are explored. We investigate the induction of the small HERC protein by Interferon β and the antiviral activity of HERC5 and HERC6, the most recently emerged members of the small HERC family. We discovered that an ancient form of the HERC5 protein present in the fish species coelacanth has the ability to inhibit the simian immunodeficiency virus (SIV) but not HIV in a similar manner to the human HERC6 protein, whereas the human HERC5 protein inhibits both viruses potently. Focusing more on the interferon induced HERC5 and HERC6 and their contribution to innate immunity, the second chapter focuses on HERC6 and single nucleotide polymorphism which enhances its ability to restrict HIV-1. This SNP confers antiviral activity against HIV-1 in vitro and correlates with disease progression to AIDS in an infected cohort in Uganda. The third chapter illustrates the ability of HERC5 to restrict Ebola virus structural protein VP40 by degrading its RNA, solidifying it as an important weapon in the arsenal of innate immunity. Unfortunately, the glycoprotein of EBOV is able to counteract the restrictive ability of HERC5. The small HERC family and especially HERC5 and HERC6 are emerging as potent antiviral molecules which can combat diverse families of highly pathogenic viruses such as HIV and Ebolavirus. This work advances the knowledge of the small HERC family and more generally the multifaceted ways the innate immune system responds to viral infections in humans.

Keywords

Innate immunity, Restriction factors, Interferon stimulated genes (ISG), Human Immunodeficiency Virus -1, Simian Immunodeficiency Virus, Ebola virus, Small HERC family, single nucleotide polymorphisms, HERC3, HERC4, HERC5, HERC6, ISG15.

Summary for Lay Audience

Innate immunity is the body's ancient and immediate form of defense against harmful organisms. The subtle workings of this protection have evolved over millennia to combat diverse pathogens such as viruses, bacteria, and parasites. Activation of the interferon response is one such defense which produces various proteins to combat these deadly pathogens. Unfortunately, pathogens have developed their own methods of overcoming the immune system, to allow the establishment of the infection. In this thesis, I will discuss the various functions of the small HERC family of proteins, a subset of which are activated by the interferon response to fight a variety of viruses. HERC5 and HERC6, two of the most evolutionarily recent members of the small HERC family are interferon induced and have potent antiviral properties. HERC5 has the ability to inhibit the replication of multiple viruses including Human Immunodeficiency virus (HIV-1) and Ebola virus (EBOV) through various mechanisms of actions. First, through its ability to directly modify viral proteins it can inhibit assembly of viral particles. Second, it can inhibit the exit of viral RNA from the nucleus prohibiting the virus from making its structural proteins and forming new progeny. Lastly, to combat EBOV, HERC5 inhibits the production of its structural protein VP40 through the degradation of its RNA. Interestingly, EBOV, via its envelope protein, is able to inhibit the antiviral activities of HERC5. Another antiviral member of the small HERC family, HERC6, is not able to inhibit HIV-1 replication, however, a single amino acid change can restore this potency. The study of the immune system allows for increased understanding of how the body functions to combat viruses and how viruses have evolved elaborate defenses to continue infecting humans. This knowledge would contribute to the development of therapies and treatments to enhance the immune defenses against viruses or conversely, discover novel ways to inhibit viral replication through inhibitions of their functions.

Co-Authorship Statement

In Chapter 1 the section 1.3.5 ISG15 and ISGylation was partially published in: Mathieu NA, Paparisto E, Barr SD, Spratt DE. HERC5 and the ISGylation Pathway: Critical Modulators of the Antiviral Immune Response. Viruses. 2021 Jun 9;13(6):1102. doi: 10.3390/v13061102. (Appendix 1)

Chapter 2 of this thesis was published in: J Virol. 2018 Jul 1; 92(13): e00528-18. (Appendix 2). SD. Barr and MW Woods were responsible for generating Figure 2.1, 2.2 and 2.6. MW. Woods generated Figure 2.3F and 2.4D. MD. Coleman generated Figure 2.3C and HP Kohio generated Figure 2.4A and B.

In Chapter 3, Figure 3.1 C, D, and E was completed by S. Barr. Figure 3.3 was completed by S. Barr in collaboration with Seyed A. Moghadasi, and Divjyot S. Kochar.

Chapter 4 of this thesis was published in: Cells 2021, 10(9), 2399; https://doi.org/10.3390/cells10092399. (Appendix 1). Figure 4.1 B was adapted from (Watt et al., 2014), copyright © American Society for Microbiology, J. Virol. 88, 2014, 10,511–10,524, doi:10.1128/JVI.01272-14. Figure 4.2 F and G was generated by NR. Hunt, Figure 4.2 H was generated by EJ. Di Gravio. Figure 4.3 and Figure 4.5 A and B were generated by NR. Hunt in collaboration with EJ. Di Gravio. Figure 4.4 was generated by DS. Labach in collaboration with E. Paparisto.

Acknowledgments

Table of Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

1 Introduction

We share our world with a multitude of other organisms: some beneficial to us, most neutral, and a small subset which are pathogenic. Although viruses are obligate intracellular parasites, they have their own evolutionary trajectory, their genome in a constant battle to overcome the defenses of the host. Humans and pathogenic viruses have co-evolved, developing defense mechanisms to invade and resist each other. This chapter will provide a brief introduction of two harmful viruses to humans: Zaire ebolavirus (EBOV) and Human Immunodeficiency Virus Type 1 (HIV) and the way that the innate immune system has simultaneously evolved to detect and eliminate these pathogens. Special attention will be given to the small HERC family of which two members HERC5 and HERC6 are interferon-induced and defend against diverse families of viruses. A portion of section 1.3.5 was previously published in Viruses (Mathieu NA, Paparisto E, Barr SD, Spratt DE. HERC5 and the ISGylation Pathway: Critical Modulators of the Antiviral Immune Response. Viruses. 2021 Jun 9;13(6):1102. doi: 10.3390/v13061102.).

1.1 Global Impact of HIV and Ebola virus

After experiencing the global COVID-19 pandemic, it is difficult to deny the power of viruses to affect our lives. Viral outbreaks impact not only the lives of those infected but the lives of the community and the country where they occur. The far-reaching socio-economic disruptions can last for years after the viral outbreak has been contained[1,2].

The largest recorded *Ebolavirus* outbreak is the West African epidemic between 2014 and 2016 which infected 28,610 people and caused 11,308 deaths in the region[3,4]. This was caused by the Zaire strain of the *Ebolavirus* (also referred to as Ebola virus (EBOV)) rapidly spreading from a rural region in Guinea to the capital city of Conakry, which facilitated the spread to the neighboring countries of Liberia and Sierra Leone[3,5–7]. Until

this outbreak, *Ebolaviruses* were considered an exotic pathogen of little consequence to public health[3].

Ebola virus disease (EVD) caused by the Ebola virus devastated the West African economy and caused a monetary loss of approximately 53.19 billion dollars[1,2,9]. The burden placed on the healthcare system likely resulted in the loss of an additional ~10,000 lives, mostly of people who had other serious health conditions such as HIV, tuberculosis, or malaria and could not access care [10,11]. During this epidemic, healthcare workers were amongst the most affected, around 8 % of which were lost to the Ebola virus infections[6,12]. Due to the vulnerability of the region, recovery was slow and the impact of the outbreak long-lasting. In addition to the death and economic devastation caused by EVD, survivors of the disease are left with life-long disabilities such as autoimmune disorders, ocular problems, encephalitis, neurological disorders, and musculoskeletal pains. This constellation of symptoms is collectively referred to as "Post Ebola Syndrome"[7,13,14] . *Ebolavirus* outbreaks continue to occur in Africa despite the availability of a vaccine. Currently, there is an outbreak in the Democratic Republic of Congo, making it the sixth outbreak in 5 years[15]. In addition to the treatment challenges, harassment of healthcare workers by armed groups, lack of public trust in the health system, and a volatile political climate compound the difficulty in containing this disease[6,15]. The increase in outbreaks over the years, the severity of the disease, and the high socioeconomic cost warrants additional research focused on discovering new, affordable, and broad acting antiviral therapies.

While *Ebolavirus* epidemics occur mostly in the central and western African regions, the HIV pandemic has spread globally, leaving no country untouched[16–18]. Since the first AIDS cases were discovered in the 80's HIV has claimed more than 40 million lives[16–21]. As of 2021, more than 38.4 million people were living with HIV, 1.7 million of which are children under the age of 15[22,23]. Although mortality has dropped almost 40 % since 2010, in 2021 650000 people died from AIDS related disease worldwide[22,24]. In Canada, approximately 62,790 people were living with HIV in 2020, with 1,639 of those being

new cases[25,26]. In addition to the devastating health impact this virus has on people living with HIV, it has a significant burden on the economy. This includes the labor force, ongoing government aid to organizations for providing antiretroviral therapy to patients and testing and prevention programs[26]. The UNAIDS organization previously estimated that in 2020 approximately 26.2 billion dollars will need to be invested in the HIV/ AIDS response and prevention[24]. However, since then funding has been diverted to other global crisis and the progress towards eradicating HIV is slowing rather than progressing[24].

Global efforts have focused on achieving the 95-95-95 goal set by UNAIDS by 2025[24,27]. This strategy proposes to diagnose 95 percent of people infected with HIV, of those, 95 percent should be on antiretroviral treatment, and of those, 95 percent should achieve viral suppression. This goal was greatly helped by the initiation of treatment as soon as a positive diagnosis was confirmed, compared to the initiation of treatment only when the CD4 T cell count dropped below 200 cell per ml. In 2021, 85 percent of people infected with HIV knew their status, 75 percent of HIV positive people were accessing antiretroviral therapy, and of those, 92 percent achieved viral suppression globally[24]. In Canada, the numbers are estimated to be more in line with the goals set out by the UN where 90 percent of HIV positive people are diagnosed, 87% were estimated to be on treatment and 95% of persons on treatment had a suppressed viral load[28]. However, based on the stagnation of progress observed recently, as well as the detrimental impact of COVID-19 and continuous reduction in funding of HIV programs by various governments, most countries have been unable to meet these targets[24].

Both Ebolavirus and HIV disproportionately affect disadvantaged countries where treatment is harder to access. In addition to community outreach and continued education about prevention, new, easy to access treatments are needed to aid those who have already been infected with these viruses.

1.2 Viral Replication Cycle

Since viruses are obligate intracellular parasites, they require access and entry into the host cell to initiate and complete their replication cycle. There are certain fundamental steps that most viruses must go through to replicate: 1) it must attach to and enter the cell; 2), it must uncoat and shed the external layers to expose the genome to the replication machinery; 3) it must usurp the cellular machinery to make its own protein, and lastly 4) it must assemble into a new virion and exit the cell. Viruses have evolved to infect almost all forms of life from Archaea which live in thermal vents to bacterial cells to human organisms to plant cells[29–32].

Due to the lack of a common ancestor for all viruses, it is extremely difficult to classify them taxonomically. Currently, viruses are separated into families based on phenotypic characteristics such as structural proteins, susceptible host organisms, pathogenicity, and interactions with the immune system[33,34]. Viral families are classified into seven different categories as delineated by the Baltimore system: I) double stranded DNA viruses, II) single stranded DNA viruses, III) double stranded RNA viruses, IV) positive sense single stranded RNA viruses, V) negative sense single stranded RNA viruses, VI) single stranded RNA viruses with reverse transcription abilities, and VII) DNA viruses with reverse transcription abilities[33,34].

More recently a megataxonomy proposal was put forth to the International Committee on Taxonomy of Viruses (ICTV). This new proposal classifies viruses into four realms of RNA Viruses and Reverse-Transcribing Viruses (*Riboviria*), ssDNA viruses (*Monodnaviria*), DNA viruses encoding vertical jelly-roll type major capsid proteins (*Varidnaviria*), and dsDNA viruses encoding HK97-type major capsid proteins (*Duplodnaviria*)[35].

The two viruses most discussed in this thesis are HIV and EBOV. HIV belongs to the realm *Riboviria*, Kingdom *Pararnavirae* Phylum *Artverviricota*, Class *Revtraviricetes*, Order *Ort*ervirales, Family *Retroviridae*, Subfamily *Orthoretrovirinae*, Genus *Lentivirus*,

and species *Human immunodeficiency virus 1*[36]. While *Ebolaviruses*, the causative agent of Ebola virus disease (EVD), belong to the realm *Riboviria*, kingdom *Orthornavirae*, phylum *Negarnaviricota*, subphylum *Haploviricotina*, class *Monjiviricetes*, order *Mononegavirales*, family *Filoviridae*, genus *Ebolavirus*. It has six species *Bombali ebolavirus*, *Bundibugy ebolavirus*, *Reston ebolavirus*, *Sudan ebolavirus*, *Tai Forest ebolavirus* and *Zaire ebolavirus*[3,37]. Due to its high pathogenicity in humans, this thesis will focus on the *Zaire ebolavirus* (also referred to as Ebola virus or EBOV) although the replication steps are similar for the other members of the *Ebolavirus* genus[3,4].

While these viruses belong to separate families this thesis provides evidence that members of the small HERC family of proteins are capable of restricting their viral activity in distinct ways.

1.2.1 Human immunodeficiency virus replication cycle

The HIV virus is spherical in shape and contains a capsid core which encapsulates two copies of the viral genome consisting of 9-kB positive sense, unspliced, viral RNA (Figure 1.1 A)[38,39]. This encodes for the *gag* gene, *gag-pol*, *env*, *vif*, *vpr*, *vpu*, *tat*, *rev*, and *nef* (Figure 1.1 A). *Env* produces the gp160 transmembrane protein which is cleaved into gp120 and gp41 in the Golgi complex[40]. The *gag* gene is cleaved into the matrix protein (MA, p17), the capsid protein (CA, p24), the nucleocapsid (NC, p7), p6, and two spacer peptides SP1 and SP2 when the virion matures[41–44]. The *gag-pol* gene created by a ribosome slippage encodes for the above-mentioned *gag* proteins but instead of SP2, protease, reverse transcriptase and integrase are also produced[43,44]. Vif, Vpr, Vpu and Nef are generally considered accessory proteins since they are not strictly essential to produce a virion. They do however contribute greatly to the success of the nascent viral particle[45,46]. The viral genome and replication cycle for HIV are summarized in Figure 1.1 A and B respectively.

Person to person transmission of HIV occurs through the exchange of bodily fluids with infected individuals. This can include the exchange of blood, breast milk, semen, and

vaginal secretions, and typically requires the overcoming of a physical barrier such as the epithelial tissue[47–53]. During sexual transmission, the mucosal epithelium provides a substantial barrier to the virus and microtears are required for the virus to enter the host[47,48,54,55]. The initial virion able to overcome this physical barrier is termed the transmission founder (TF) virus[56–59]. The first cells to encounter HIV in both the female genital tract and the inner foreskin are site specific dendritic cells, epidermal Langerhans cells (eLCs)[60]. These cells then carry the infection to nearby CD4+ T cells either in the submucosa or nearby lymph nodes where the infection can be disseminated[56]. Viral spread to the gut associated lymphoid tissue (GALT) is essential for the establishment of infection. The lymphoid cells in this region of the body are more abundant than anywhere else and are generally in a low state of activation, making them ideal candidates for infection [69]. Although the infection is usually introduced into the body from elsewhere, such as the vaginal mucosa or the peripheral blood, it only establishes a secure foothold once it reaches the GALT and has access to this rich proinflammatory environment with IL2, IL6, and TNF-α which are considered stimulatory cytokines [70,71]. Dissemination of HIV to immunologically protected sites such as the brain form long lasting viral reservoirs [67,68].

Expression of the receptor CD4 and coreceptor CCR5 and CXCR4 are usually required although cells not expressing these receptors such as kidney epithelial cells and astrocytes can be infected via endocytosis or cell-to-cell transfer bypassing the membrane fusion stage of entry[61–63, 66]. The main targets of HIV, T cells, monocytes, immature dendritic cells, and macrophages all express the CD4 receptor, albeit at varying levels[61,62,64,65]. The HIV infection cycle is divided into two stages: the early stages which include attachment, fusion and entry, uncoating, nuclear entry, and integration; and the late stages which include viral transcription, nuclear export, translation, assembly, budding and maturation (Figure 1.1 B).

Attachment and fusion of the virion occurs at the cell membrane; however, these initiating events can vary by cell type as well as the stage of infection. HIV attachment

can occur via interaction with integrins, pattern recognition receptors such as dendritic cell–specific intercellular adhesion, molecular 3-grabbing non-integrin (DC-SIGN) or nonspecific interactions with negatively charged cell-surface heparan sulfate proteoglycans[72,73]. Langerhans cells express a C-type lectin called langerin which is known to bind HIV[68] whereas lymphocytes express integrin α4β7[71]. Although these attachment factors are helpful and likely aid in stabilizing the virion on the plasma membrane to interact with CD4 and the coreceptor they are not strictly essential for infection to occur[71,73].

In cell-free infection, viral particles are released from the cell and proceed to infect non-adjacent cells, usually through fusion with the plasma membrane[74,75]. The envelope protein which studs the surface of the virion is made up of a heterodimer of gp120 and the transmembrane gp41. These heterodimers form a trimer where the gp120 proteins are facing outwards like an umbrella and shields the transmembrane gp41 proteins[76–79]. The gp120 trimer interacts with the CD4 receptor and with the aid of a coreceptor further conformational changes take place which allow for the insertion of the gp41 section of the envelope into the cell membrane and brings them in close proximity to mix with the plasma membrane of the target cell. The transmitter founder virus usually utilizes the CCR5 receptor and viruses utilizing the CXCR4 receptor usually emerge later during the progression of the infection[70]. There are other coreceptors which have been shown to interact with the envelope protein in vitro and these alternate coreceptors could account for the ability of HIV to infect almost all cell types[70,73,76,80,81]. The binding of multiple CD4 receptors to the envelope can speed up the process of opening the gp120 trimer and exposing the gp41 peptides for fusion. A more efficient form of viral dissemination is cell-to-cell transmission whereby a virological synapse, similar to an immunological synapse, is formed to connect two susceptive cells and allow the transmission of virions[75]. Cell-to-cell contact is facilitated by LFA-1 and ICAM-1 receptors on the cell surface maintaining the doner and receiver cell in close contact during the transmission[73,82,83]. This attachment causes additional signal transduction and biochemical changes within both the doner and the receiver cells which influence viral

spread and pathogenesis[74,75]. Both cell-free and cell-to-cell transfer result in the shedding of the viral membrane and release of the viral core into the cytoplasmic milieu.

Following the release of the capsid into the cell, the reverse transcription and uncoating steps are initiated in the cytoplasm. Uncoating is thought to aid in the transport of the HIV genetic material to the nuclear pore where it can enter the nucleus and integrate into the genome. The capsid core is transported along microtubules towards the nucleus[87,90–93]. This movement is aided by the adaptor protein FEZ1 which accommodates the attachment of the capsid molecules to the motor protein Kinesin-1[94]. Although the mechanism of motor movement is incompletely understood it has been demonstrated by several groups that binding to both the dynein and kinesin motors is required for the movement of the core toward the nucleus[94]. Depleting the cell of either motor protein results in accumulation of particles at the plasma membrane and diffusely throughout the cytoplasm[95]. The timing of viral core uncoating is essential for the infectivity of the virus, whereby early uncoating of viral particles can lead to the degradation of the pre-integration complex by the proteasome as well as detection of viral components by the surveillance molecules such as cGAS[96,97]. Imaging studies have suggested that uncoating at or near the nuclear pore is important for the import of the pre-integration complex into the nucleus. Viral docking occurs through association with nucleoporin 358 (NUP358)[91,96,97]. This, in concert with binding of the capsid to NUP153 and CPSF6, are important interaction for the import of the pre-integration complex into the nucleus[97].

The reverse transcription complex which is thought to form shortly after entry into the cytoplasm consist of matrix proteins, capsid proteins, reverse transcriptase, integrase, nucleocapsid, Vif, Tat, Nef, Vpr and protease[84]. Although it is unclear if these are coincidental associations or whether they have a positive effect on the reverse transcription process[85]. The nucleocapsid protein is believed to act as a chaperone and help in the reverse transcription of the RNA. It arranges nucleic acids in the most thermodynamically stable position and aids in the annealing of the tRNA which serves as the primer in the DNA synthesis[86,87]. The reverse transcription process can even occur in

vitro with just the reverse transcriptase (RT), dNTPs, and a template[84,88]. RT has two enzymatic activities: 1) the DNA polymerase function which creates DNA using either RNA or DNA as the template and 2) the RNase H activity which degrades the RNA template only when it is complexed with a DNA strand[88]. The steps of reverse transcription convert a positive sense single-stranded RNA genome into a double-stranded DNA genome. A tRNA primer base binds to the primer binding site near the 5' end of the viral RNA. This initiates reverse transcription to generate the minus strand of viral DNA (vDNA). The RNase H active site of the RT degrades the viral (vRNA) releasing the newly synthesized vDNA strand. This strand folds back and binds to the 3' R region. HIV-1 has packaged two copies of the viral RNA and this DNA fragment can bind to the 3' region of either vRNA strand allowing for the synthesis of the rest of the vRNA genome[84]. As the minus-strand of DNA continues to be synthesized the RNase H continues to degrade the viral RNA with the exception of two small regions called the polypurine tract (ppt) which are located near the 3' region and near the middle of the genome. These ppts act as primers for the initiation of the transcription of the positive-sense DNA strand[84,89]. The plus-strand synthesis can commence from either ppt and continues until the tRNA primer set. This is then removed by the RNase H activity which frees the plus-stranded DNA for transfer to the 5' end of the minus strand, where extension of the plus-strand leads to the completion of the double-stranded viral DNA. The reverse transcribed genome in complex with integrase, some capsid proteins and cellular factors is named the pre-integration complex (PIC)[90].

There has been continuous debate about where and how reverse transcription occurs and how this process is linked to the uncoating of the capsid to expose the pre-integration complex of viral DNA for binding into the host DNA. The model of reverse transcription, uncoating and integration which is beginning to emerge consists of the initiation of reverse transcription in the cytoplasm as the core is moving towards the nuclear membrane[96]. Some groups propose that the uncoating of the virus occurs at the nuclear pore where most of the capsid is shed, however a small amount of capsid travels with the viral replication complex (VRC) into the nucleus[96]. In the nucleus, in association with

CPSF6 the VRC is guided towards nuclear speckles where it integrates within nuclear speckle associated genomic domains (SPADs)[98]. A separate theory proposes, that shedding of the capsid does not occur in the cytoplasm or nuclear pore but that a nearly intact capsid enters the nucleus where it completes the reverse transcription steps and proceeds to uncoat near the integration sites approximately 1.5 µm away from the nuclear envelope[99]. Both studies did find that the host cofactor CPSF6 is important for the facilitation of nuclear entry and post entry events[98,99]. It is clear that there is a complex interplay between the entry of capsid into the nucleus and the direction of PIC into the target site for integration which requires further clarification. The import of the capsid and the integration of the viral genome involves the aid of several cell host proteins such as CypA, RANBP2 (also known as NUP358), CPSF6, TNPO3, NUP153. CPSF6 and NUP153 are involved in both the import of the PIC into the nucleus as well as the selection of integration sites on the host genome[98,99]. Altering the interaction of the capsid with these factors influences not only the entry route of HIV into the nucleus but also the site of integration into lamina-associated DNA (LADs)[90,99,100].

Upon entry into the nucleus, integrase processes the 3' end of the HIV genome by removing two nucleotides from each end and turning the blunt ends of the genome into reactive ends[100]. The viral DNA docks onto the target site on the host genome and forms a target capture complex where the exposed 3' ends of the viral DNA are used as nucleophiles to attack the target DNA at the major grove. Complimentary strands then join with the 5' end of the host genome and host DNA repair machinery join the gaps creating a 5 bp extension onto the integrated provirus[100]. The selection of integration site can determine how actively the viral genome will be transcribed or whether latency will be established[101–103]. Determinants of HIV latency are a highly active area of research and evidence suggests that it may be influenced by both the site of integration and cellular factors. Generally, HIV has been shown to integrate into highly transcribed genes. This preference may be influenced by the cellular factor LEDGF as its depletion leads to a shift away from active genes[104]. Additionally, integration seems to favor the nuclear periphery and open chromatin, and disfavors LADs, heterochromatin condensed regions

and centrally located transcriptionally active regions. Non-canonical B-form DNA has also been shown to influence integration at the nucleotide level[101,105,106].

Once the provirus is integrated into the host DNA, transcription can commence. The initial activation of the provirus largely depends on the state of the host microenvironment including the availability of the transcriptional cofactors and the state of activation of the cell[105]. The availability and abundance of host transcription factors affect the ability of the virus to enter the host phase and subsequently the highly active viral phase that follows[107–110]. This availability also contributes to the heterogenicity in the transcriptional activation and latency observed both in cell culture and in vivo[111] The RNA produced in the latent state is at low levels and the transcripts are short and immature[110]. Activation of the CD4+ T cell and NF-κB leads to a state where the early HIV genes, *tat*, *nef,* and *rev* are expressed[105]. The proviral promoter has a binding site for both NF-κB and NFAT either of which could initiate active transcription of these viral early genes[109,110,112]. The transcriptional circuit of HIV is dependent on the Trans activator of transcription (Tat) protein creating a positive feed-back loop and only jumps into high gear once the Tat protein is transcribed and translated[112]. The LTR contains the core promoter elements which include three tandem SP1 binding sites, a TATA box, and a highly active initiator sequence[109,110,112].

Tat expression leads to a 100-fold increase in transcription of HIV genes and production of full-length RNA transcripts [109,110]. Tat binds to the Trans Activating Response (TAR) element, an RNA stem-loop structure located in the first 59 nucleotides at the 5' LTR. This leads to the recruitment of P-TEFb (positive transcription elongation factor B), the phosphorylation of RNA polymerase II, and increased transcription of viral mRNA[108]. In the absence of Tat, the low-level boost of transcription provided by cellular activation is short-lived and leads to the production of a small amount of viral transcripts. During the host phase of the transcriptional activation the host protein KAP1/TRIM28 seems to be important for the initiation of transcription. However, upon Tat expression KAP1 seems to be dispensable for the recruitment of CDK9 and transcriptional activation [110]. Since the

proviral DNA contains only one transcriptional start site alternate splicing is required to generate the large spectrum of viral mRNAs produced.

Typical conversion of pre-RNA into mRNA involves the removal of introns, addition of a 5' cap and polyadenylation at the 3' end[114,115]. The same holds true of HIV mRNA. HIV RNA has 5 exons which splice together to form the various viral mRNAs. As a rule, the first and last exons are always used[113,115]. Three forms of RNA are translated from the viral DNA: 1) a full-length, unspliced, ~ 9 kB mRNA from which Gag and Gag-Pol are translated; 2) a singly or partially spliced, ~4 kB mRNA from which gp160 (Env), Vif, Vpu and Vpr are translated; and 3) a fully spliced mRNA from which Nef, Tat, and Rev are translated[113,115,116]. Two copies of the unspliced, full-length RNA are also incorporated into the new viral particles[117]. The tendency of the cell is to splice pre-RNA and the default state for HIV RNA is for full splicing to occur. Thus, the 2 Kb multi-spliced transcript of *rev*, *tat* and *nef* are the first produced and they are exported into the cytoplasm through the NXF1 nuclear export pathway similarly to the majority of the host mRNAs[107,113,118]. There, *rev* and *tat* are translated and shuttled back into the nucleus via the nuclear localization signal[107,113]. As mentioned previously, Tat production and nuclear import is essential for increasing the rate of transcription whereas Rev is essential for the export of unspliced and partially spliced viral mRNA from the nucleus. Partial and unspliced mRNAs required virus specific suppression of splicing by overriding the cellular rules to retain introns[115]. Normally, the export of unspliced and intron retained RNA is blocked at the nuclear pore thus the unspliced and partially spliced viral mRNA requires the aid of newly synthesized Rev to mediate export through the CRM1 pathway[107,119]. Intron containing RNA is unstable and quickly degraded in the nucleus, Rev multimerizes around the viral mRNA to protect it from cellular degradation[113]. Expression of Rev leads to decreased levels of fully spliced viral mRNA whereas an increase in fully spliced mRNA is observed when Rev expression is decreased[115,120]. Upon nuclear import, Rev binds to the rev response element (RRE), a short loop on the 5' end of the transcript. This leads to a conformational change and recruitment of other Rev proteins for multimerization[107]. Rev contains a leucine rich nuclear export signal at the

disordered C-terminal domain which forms a complex with RanGTP and CRM1 (also known as exportin1 or XPO1) to shuttle through the nuclear pore complex into the cytoplasm where it disassembles when RanGTP is hydrolyzed into RanGDP[107,121].

In the cytoplasm, the cellular translation machinery is used to generate viral proteins from both the spliced and unspliced viral mRNAs. The 5' end of the unspliced mRNA from which Gag and Gag-Pol are produced contains several structural elements such as the TAR element, the primer binding site, the dimer initiation site, and packaging signal, which, while essential to replication hinder the translation initiation[108,113,122,123]. Recruitment of cellular proteins such as the La protein, TAR RNA-binding protein (TRBP) and Staufen1 aid in overcoming the translation barriers created by these secondary RNA structures[124–127]. While other cellular factors such as cap-binding complex subunit CBP80 and DEAD-box RNA helicases such as eIF4AI and DDX3 cooperate with Rev to recruit the ribosomal subunits for the initiation of the translation of the unspliced and partially spliced mRNAs[107,128,129]. Unspliced mRNA translation produces the Gag proteins while a -1 ribosome frameshift governed by a 6U slippery site, and a downstream stem-loop structure produces the Gag-Pol protein[130,131]. Maintenance of a ratio between the 55 kD Gag (Pr55) and the 160 kD Gag-Pol production at 2% to 10% Gag-Pol is important for the assembly of the virion and may be modulated by the availability of Leu tRNA[131].

Gag is the master orchestrator of the viral particle assembly and Gag containing virus-like particles (VLP) are formed in the absence of any other viral proteins[42,79,132]. Gag and Gag-Pol are translated in the cytoplasm on free ribosomes[41]. The Gag protein consists of a matrix (MA) domain at the N-terminus, followed by the capsid domain (CA), nucleocapsid domain flanked by two spacer region SP1 at the n-terminus and SP2 at the c-terminus, and lastly the p6 domain at the c-terminus of Gag[41,93,133–135]. Although there is some uncertainty about which step comes first, multimerization, RNA binding, or membrane targeting, the end result is that same, assembly of Gag/Gag-Pol, and viral RNA at the plasma membrane where the incorporation of Env and budding occur. The

myristalated matrix domain targets the Gag particles to the inter leaflet of the plasma membrane, the capsid domain mediates Gag-Gag interactions and multimerization into hexamers, the nucleocapsid domain binds two copies of the unspliced viral RNA to be incorporated into the nascent viral particle and the p6 domain recruits the cellular ESCRT machinery to complete membrane cession and release from the cell[42,93].

A myristate moiety is post-translationally attached to the N-terminal of the matrix domain, a modification which together with a basic patch of amino acids targets Gag molecules for transport to the plasma membrane[41,93,134,136]. Elimination of this myristylation changes the destination of the Gag molecules from the plasma membrane to a more diffuse cytoplasmic localization[136,137]. Capsid-capsid interactions are important for the multimerization of the Gag molecules into hexameric sheets as well as the recruitment of essential cell factor cyclophilin A (CypA). CypA incorporation into the viral particle can drastically alter infectivity of the viral particle[91,138]. It is important for protection from cellular restriction factor TRIM5α as well as for successful and timely uncoating upon entry in the cell[138,139]. The 5' end of the unspliced HIV genome contains a highly structured stretch of 150 to 200 bp which encode the packaging signal (psi or ψ). This region contains a nucleocapsid (NC) binding site on stem loop three as well as a dimer initiation signal (DIS) on stem loop one. Packaging of RNA requires an intact psi packaging element as well as parts of the RRE and gag sequence for efficient packaging although in the absence of RNA containing these elements, spliced HIV RNAs will be packaged and, in their absence, cellular RNAs can get packaged into the particles[140]. Interestingly, although the pool of unspliced RNA is both translated into the Gag/ Gag-Pol proteins and incorporated into nascent viral particles, stable RNA-Gag complexes near the plasma membrane are only seen with non-translating viral RNA. The encapsidation of two copies of viral RNA is important for viral recombination as the RT switches between the two templates during the reverse transcription process[41,88,89,141].

Binding of the nucleocapsid domain of Gag to the ψ element is thought to begin the assembly of the immature particle providing a scaffold for the multimerization of the Gag

lattice at the plasma membrane[93,142]. Through actively remodeling the viral assembly HIV acquires a unique membrane protein composition. As Gag multimerizes at the inner leaflet of the plasma membrane phosphatidylinositol 4,5-bisphosphate (PIP2) molecules, through interactions with the highly basic region on MA are incorporated into the assembly site[143]. This kicks off the process of remodeling through trans bilayer coupling and selection of GPI anchored proteins into the virion. As the membrane acquires high curvature dual membrane anchors are recruited to the viral assembly site[143]. It was long thought that Gag multimerization alone provided the required force to bend the plasma membrane for particle formation, however, more recent studies have discovered that BAR domain-containing proteins such as IRSp53 and angiomotin (AMOT) may be required to aid bending the membrane[144–147].

Budding and release of the virions is largely dependent on the cellular endosomal sorting complex required for transport (ESCRT) complex[93,148]. Made up of more than 30 proteins the human ESCRT machinery performs a complex dance regulated by ubiquitin and other regulatory machinery to coordinate the remodeling of the membrane during intraluminal vesicle. This is choreographed by the p6 domain of Gag which contains the PTAP and the YPXL late domain[41,149,150]. TSG101 subunit of the ESCRT-I complex binds to the PTAP domain of Gag to recruit the ESCRT-III proteins CHMP2 and CHMP4 which in conjunction with VSP4 mediate membrane fission through polymerization in a dome structure which pulls together the two sides of the membrane neck[43,147,151]. Similarly, the YPXL domain is the binding site for the ALIX protein which in turn recruits CHMP2 and CHMP4 to complete membrane fission. In addition to aiding in membrane scission VSP4 releases the ESCRT complex back into the cytoplasm[43]. Ubiquitination of Gag by the ubiquitin ligase NEDD4L can rescue the budding deficiencies of Gag which lacs TSG101 and ALIX binding site[41,151]. In addition to its contribution in budding the p6 domain is essential for the recruitment of Vpr into the nascent viral particles[41,152]. The functional consequence of Vpr recruitment into viral particles is uncertain but recent studies have suggested that Vpr is dissociated quickly upon cell entry and any residual molecules are associated with the capsid suggesting a role in early infection[152].

The Env glycoprotein precursor (gp160) contains an ER signaling sequence targeting for translation on the rough endoplasmic reticulum[79]. A hydrophobic membrane anchor at the transmembrane domain of gp160 prevents it from being released into the ER lumen[79,153]. Gp160 undergoes co-translational glycosylation with both *N*-linked and *O*-linked oligosaccharide sidechains. In the ER, gp160 forms trimers which are thought to facilitate the trafficking to the Golgi complex where it acquires extensive sidechain modifications. There, it is proteolytically cleaved by furin or furin-like proteases to produce the surface glycoprotein (SU) gp120 and the transmembrane glycoprotein (TM) gp41, which remain noncovalently associated. The proteolytic cleavage of gp160 is a process required to activate the fusogenic abilities of Env which are essential for entry into the cell and infectivity. The trimeric complexes of gp120 and gp41 is trafficked to the plasma membrane where it is quickly recycled back into the cell via clathrin mediated endocytosis[79]. The weak noncovalent association between the gp120 and the tgp41 subunits often results in sheading of gp120 which together with the endocytosis maintains a relatively low concentration of Env at the plasma membrane and in incorporated particles. There are four, not mutually exclusive, models for how Env is incorporated into viral particles: 1) passive incorporation whereby Env is incorporated into the virion simply due to it being on the cell membrane 2) direct Gag-Env interaction model whereby a direct interaction of the matrix domain and the cytoplasmic tail of Env incorporated it into the virion 3) The Gag-Env co-targeting model where Env incorporation is due to targeting of both Env and Gag to specific plasma membrane microdomains 4) indirect Gag-Env interaction model where a host protein acts as an adaptor molecule linking the cytoplasmic tale of Env with the matrix domain in Gag[79,154]. While some evidence exists for all these models, recent studies utilizing matrix trimerization mutants found that Env incorporation was dependent on the formation of matrix trimers and an intact cytoplasmic tail[155].

The last step in the infection cycle of HIV is maturation of the released particles. In this critical step the viral protease cleaves Gag and Gag-Pol at multiple positions leading to structural rearrangements which prepare the virus for entry into the target cell. The viral

protease is incorporated into the virion as part of the Gag-Pol precursor and during or slightly after budding initiates a highly ordered sequential cleaving cascade[44,130,134]. Protease processing of Gag and Gag-Pol results in production of the matrix, capsid, nucleocapsid, p6, protease, reverse transcriptase, and integrase. The matrix proteins remain bound to the viral envelope while the capsid proteins arrange into a cone shape which surrounds the condensed ribonucleoprotein complex (RNP) formed by the nucleocapsid and RNA[156]. The first cleavage occurs between SP1 and NC domain, followed by MA-CA and SP2-p6 junction, NC-SP2 site, and finally between CA and SP1. The last cleavage liberates the capsid protein to assemble into the conical viral core[43,135]. The cleavage of the CA-SP1 site is particularly important as uncleaved capsid molecules exert a dominant negative effect and severely reduce viral infectivity[157,158]. Consistent with its importance in HIV infectivity, virion maturation has been the successful target of a class of antiretroviral drugs called protease inhibitors.

Studying the replication mechanisms of HIV is essential to the development of antiviral drugs, especially since the development of a vaccine has thus far been unsuccessful. Antiretroviral drugs, on the other hand, are currently the only available treatment for HIV infection. The development of new drugs is required to combat the drug resistant strains which have evolved and keep the infection at bay.
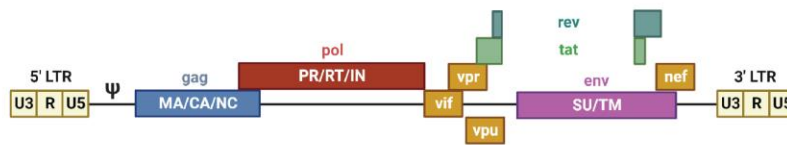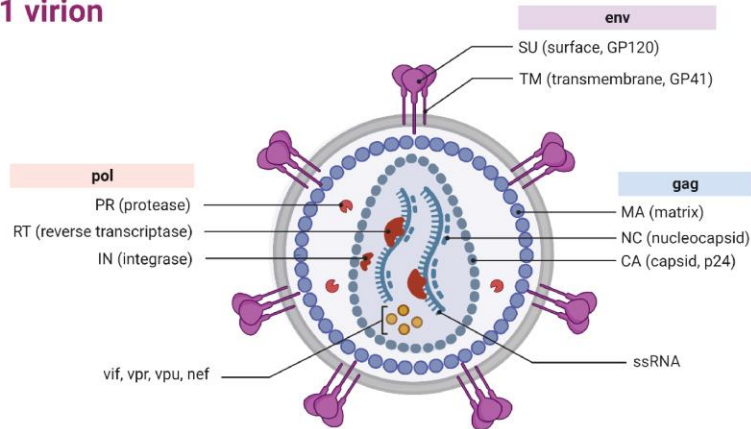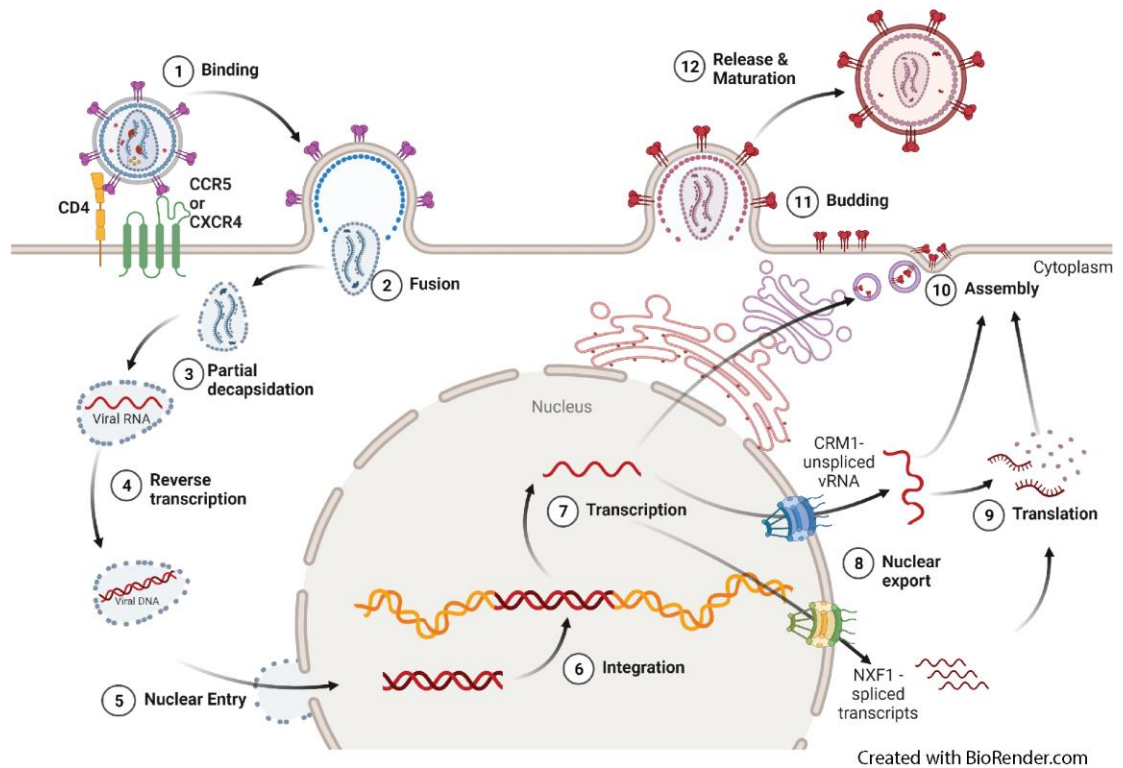
Figure 1.1: HIV-1 genome and replication cycle.

A) Top panel, HIV genome. The HIV genes are flanked by a 5' and a 3' long terminal repeat region. The 5' LTR is followed by the psi (ψ) packaging signal, *gag* and *gag-pol*, *vif, vpr, vpu, env* and *nef*. The *rev* and *tat* genes are produced by alternate splicing. Three forms of RNA are translated from the viral DNA: 1) a full-length, unspliced, ~ 9 kB mRNA from which Gag and Gag-Pol are translated; 2) a singly or partially spliced, ~4 kB mRNA from which gp160 (Env), Vif, Vpu and Vpr are translated; and 3) a fully spliced mRNA from which Nef, Tat, and Rev are translated. Bottom panel, cartoon representation of the mature HIV particle. **B)** HIV replication cycle. 1) The gp120 trimer interacts with the CD4 receptor and with the aid of a coreceptor (CXCR4 or CCR5) further conformational changes take place which allow for fusion of the virion with the plasma membrane. 2) The viral membrane is shed, and the viral core is released into the cytoplasmic milieu. 3) Uncoating aids in the transport of the HIV genetic material to the nuclear pores where it can enter the nucleus and integrate into the genome. 4) The steps of reverse transcription convert a positive sense single-stranded RNA genome into a double-stranded DNA genome. The reverse transcribed genome in complex with integrase, some capsid proteins and cellular factors form the pre-integration complex (PIC). 5) Viral docking at the nuclear membrane occurs through association with nucleoporin 358 (NUP358). This, in concert with binding of the capsid to NUP153 and CPSF6 import the pre-integration complex into the nucleus. 6) In the nucleus, in association with CPSF6 the viral replication complex is guided towards nuclear speckles where it integrates within nuclear speckle associated genomic domains (SPADs). The selection of integration site can determine how actively the viral genome will be transcribed or whether latency will be established. 7) Activation of cells initiates the transcription of the early HIV genes *tat, rev*, and *nef* which kicks into gear the transcription of other HIV genes and production of full-length RNA transcripts. Alternate splicing generates the large spectrum of viral mRNAs produced. 8) Unspliced viral RNAs exit the nucleus via the CRM1 pathway with the aid of Rev while fully spliced viral RNAs exit the nucleus via the NXF1 pathway. 9) both spliced and unspliced viral RNAs are translated into viral proteins in the cytoplasm using cellular translation machinery. 10) Gag is the master orchestrator of the viral particle assembly coordinating

multimerization, RNA binding, and membrane targeting. 11) Virion budding is dependent on the cellular endosomal sorting complex required for transport (ESCRT) complex. The p6 domain of Gag recruits various members of the ESCRT complex to budding and scission of the membrane. 12) In the maturation step the viral protease cleaves Gag and Gag-Pol at multiple positions leading to structural rearrangements which prepare the virus for entry into the target cell.

1.2.2 Ebolavirus Replication Cycle

Predominantly filamentous shaped, Ebolaviruses are membrane bound and contain a 19kB non segmented, negative sense, linear RNA[3,93]. This codes for the structural matrix protein VP40 which can assemble into filamentous particles in the absence of any other viral proteins[159–161]. The EBOV genome also encodes for *GP* the glycoprotein $GP_{1,2}$ as well as the secreted glycoproteins, *NP* the nucleocapsid protein, *VP35* the polymerase cofactor, *VP30* the transcriptional activator, *VP24* the RNA complex-associated protein, and *L* the large protein RNA-directed RNA polymerase (Figure 1.2 A)[3,161].

Most EBOV outbreaks begin with a single spillover event into the human population from an unknown reservoir by unknown means, although there is some evidence for bats being the reservoir host and non-human primates acting as an intermediary species[3,162–165]. Human-to-human transmission can occur through sexual contact, although direct contact with infected biological materials or exposure to contaminated non biological materials such as through an open wound or contact with mucosal membranes is considered the most frequent mode of transmission[3,4,166].

EBOV cell tropism is determined by the glycoprotein $GP_{1,2}$ and the cellular attachment factor NPC intracellular cholesterol transporter 1 (NPC1, also known as Niemann–Pick C1 protein) receptor. The primary targets of infection are mononuclear phagocytes and dendritic cells which upon infection disseminate the virus e regional lymph nodes and to the liver, spleen, and kidney[167–170]. EBOV infection of these organs results in focal necroses and inflammation which leads to multi-organ dysfunction and death[4,166,170,171].

Although EBOV impairs the function of immune cells preventing them from clearing the infection, they are able to secrete cytokines and chemokines resulting in enhanced immune activation resembling septic shock[4,172,173]. Ebola GP activates monocyte-derived macrophages through TLR4 and induces NF-κB as well as IRF3 activation-based production of cytokines and chemokines. This results in increased expression of interleukins IL-1β, IL-6 and IL-8, as well as tumor necrosis factor (TNF)[3,174]. It also strongly induces production of IFN beta and the interferon response[172]. This immune activation results in the recruitment of susceptible immune cells and the failure of the endothelial barrier. Additionally, natural killer cells (NK cells) and T cells undergo apoptosis, despite remaining uninfected, further impairing the ability of the immune system to overcome the disease[3,169].

At the molecular level, EBOV replication cycle begins with attachment to the host cell membrane (Figure 1.2 B). Attachment can be mediated by several host surface proteins including C-type lectin, DC-SIGN and L-SIGN, and integrins such as $\alpha_5\beta_1$ integrin, as well as TIM-1, TIM-4 and the TAM kinases TYRO3, AXL, and MERTK[175–177]. There are somewhat conflicting reports about the strict requirement of any of these receptors for entry since not all are present in the various cell types that EBOV infects. The most likely scenario is that these receptors as well as some yet to be identified act as attachment factors which mediate macropinocytosis of the virion through pathways similar to those used for apoptotic cells[178–185]. Conflicting reports suggest that macropinocytosis occurs in either a clathrin-dependent or a caveolin-dependent manner, however, most agree that remodeling of the actin cytoskeleton is essential for entry[186–188]. Particularly PI3K – AKT – RAC1 axes is important for the maturation of the endosomal compartments containing EBOV, knocking down any of these components resulted in the intracellular accumulation of EBOV in these compartments and inhibited entry[188, 189].

Once the virion is macropinocytosed the vesicle proceeds to mature into an early endosome and then into Rab7-positive late endosome/ lysosome compartments[188]. With the increase in acidity the cathepsin molecules are activated. The GP1 subunit gets

proteolytically processed in the late endosome by cysteine proteases Cathepsins B and L removing the mucin like domain and the glycan cap[177,190,194]. This cleavage exposes the receptor binding site of the molecule and primes GP (now GPCL) for the interaction with NPC1. It is well documented that NPC1 is an essential mediator of entry into the cell. There is unequivocal agreement that release of the capsid into the cytoplasm occurs following the interaction of the GP1 subunit with NPC1[177,190–193]. Binding of GP to the NPC1 receptor trigger conformational changes which are poorly understood to create a fusion pore and facilitating the release of the nucleocapsid into the cytoplasm where the replication of the genome can occur[176,178,183,193,195,196].

The fate of VP40 following release into the cytoplasm is unknown however dissociation of VP24 from the incoming nucleocapsid facilitates the relaxation of the rigid structure allowing for transcription and genomic replication to occur[4]. Primary transcription commences about two to four hours post infection and utilizes the original viral genome and proteins. Transcription occurs within the ribonucleoprotein complex which contains NP, VP30, VP35, and L. EBOV mRNAs are sequentially transcribed from the 3' end to the 5' end with conventional start and stop codons preceding and trailing each gene[4,197]. Between the end of one gene and the start of the next there is a short intragenic region of about 5 nucleotides except for between *vp30* and *vp24* where there is an intragenic region of about 143 nucleotides. The transcription process is not perfect and at times, the polymerase falls off after finishing the transcription of a gene. When this occurs, it cannot pick up where it left off and needs to start the transcription process from the beginning, creating a gradient of transcribed genes and proteins with NP being the most abundant followed by VP35, VP40, GP, VP30, VP24, and L being the least abundant[3,7,8,198]. L acts as both a polymerase as well as a having guanyltransferase and methyltransferase activities to cap the 5'of the nascent viral mRNAs to create translatable viral mRNAs[7,198]. Following transcription, the nascent mRNAs, except for GP, are translated by the free ribosomal machinery in the cytoplasm[198]. When a certain level of NP protein is reached, it localizes near the ER forming the inclusion bodies which are a hallmark of EBOV infection[199,200] . Usually, they appear in cells about 10 hours post infection and are the

site of nascent RNA synthesis. Inclusion bodies contain NP, VP35, VP30, VP24, VP40 and L although only NP is required for their formation[200,201]. How inclusion bodies are formed by EBOV remains unknown, but they are thought to originate through lipid phase separation. Once formed, inclusion bodies become the site of secondary translation, genome replication and nucleocapsid assembly. As more proteins are translated smaller inclusion bodies merge to form larger ones[200–202].

The only EBOV mRNA subjected to transcriptional editing is *GP*. Three different forms of glycoprotein transcripts are generated by L due to transcriptional stuttering at a polyuridine region. *pre-sGP, pre-GP*, and *pre-ssGP* are translated into pre-sGP pre-GP$_0$ and pre-ssGP, respectively in the ER[161,173,203]. About 70% of the transcriptional GP products form a dimer at a unique C-terminus and are secreted (sGP); approximately 25% are comprised of the GP$_0$ gene product which following translation gets cleaved to form the GP$_{1,2}$ protein, and lastly, 5% of the transcripts form a small, secreted form of the glycoprotein (ssGP)[173,204]. sGP is the main transcriptional product and it is proteolytically cleaved and glycosylated as it moves through the classic secretory pathway to be discharged in the extracellular milieu[161,173,205,206]. Retention of sGP in the cell can be highly toxic and has been implicated in acting as a decoy to evade antibody immunity, impairing the chemotaxis of macrophages, and inhibiting the production of proinflammatory cytokines[173,207,208]. The proteolytic cleavage of sGP results in the formation of a small 40 aa protein called the delta (Δ) peptide[209]. This also follows the secretory pathway but is retained in the cell for longer than sGP. Possible functions of this residual protein are inhibition of filovirus to prevent superinfection and acting as a membrane damaging viroporin[173,206,210]. The GP$_0$ gene product is also translated in the ER where it assembles into trimers and follows the classical secretory pathway to the plasma membrane[161,207]. Each GP$_0$ subunit is then post translationally cleaved by the Golgi endoprotease furin to yield disulfide linked GP1 (~55 kDa) and GP2 (~20 kDa) subunits. The final GP assembly, which is an ~450 kDa trimer of GP1,2 heterodimers, is then displayed on the surface of mature EBOV virions and is essential for viral entry[161,173,204,211]. GP1 contains the receptor-binding site and regulates the triggering of

the membrane fusion machinery in the GP2 subunit[173,194]. Expression of GP is highly regulated and large amounts of $GP_{1,2}$ are cleaved from the cell surface by the cellular metalloprotease tumor necrosis factor alpha-converting enzyme (TACE) and shed from the cells[212]. Like sGP, these shed trimers act as decoy from the immune system and sequester antibodies to prevent their binding to the viral particles[208,213]. Additionally, it reduces the cellular toxicity imposed by GP and causes immune activation of monocytes[214]. The last transcription product ssGP is created by a + 2-shift reading frame of the GP gene[204]. Similar to sGP it is heavily glycosylated and forms a homodimer however its function is thus far unknown[161,173].

The inclusion bodies are the site of secondary translation for all the other EBOV proteins as well as genome replication and nucleocapsid assembly[215]. It is unclear how the transcription and genome replication stage of the viral replication cycle are separated in space and time however they are distinguished by the presence of the transcription initiation factor VP30[215]. Transcription of proteins requires the presence of non-phosphorylated VP30 which helps initiate the transcription process[216]. Upon phosphorylation, due to interrupted interaction to VP35, VP30 dissociates from the polymerase complex and its association with NP is strengthened. Enhanced binding to NP allows for its incorporation into the nascent viral particle, a step which is critical for the initiation of primary transcription upon entry into a new host cell[216–218].

During the translation stage of the replication cycle VP35 is bound to NP to prevent oligomerization. When the switch from translation to genomic replication occurs and viral RNA is available NP molecules bind to it displacing the VP35-NP interaction and allowing for the multimerization of NP around the nascent viral RNA[200,219–223]. The EBOV genome contains a bipartite motif within the 3' leader region and the first gene (NP). Access to this promoter instructs the viral polymerase to synthesize a full-length antisense genome[216,217]. This RNA template is uncapped and non-polyadenylated and is associated with NP[223,224]. The 5'end of the genome is composed of a short non transcribed region which is important for replication, transcription initiation, and

encapsidation. The antisense genome (antigenome) serves as a template from which the future genomes are synthesized[4,197]. The replication promoter of the antigenome resides in the complementary promoter region and directs the viral polymerase to begin synthesis of genomic RNA.VP24 and VP35 then come in to complete the condensation of the nucleocapsid structure[224]. Encapsidation by multimerization of NP and VP35 of the viral RNA is cooperative with the extension of the genome[224,225]. Thus, the genome gets encapsidated as soon as it is replicated allowing for a method of protection from degradation[226].

The loose encapsidation of the genome following replication is the first step in the particle formation. This is accomplished by NP, VP35, and L which form a left-handed helical structure[215,216,224,227,228]. Then VP30 binds to the nucleocapsid and determines if the complex will function as a replicase or transcriptase depending on whether VP30 is phosphorylated or not. VP24 binds to the nucleocapsid complex locking it in place and preparing it for transport to the budding site[224]. The presence of VP24 is important for the stabilization and condensation of the nascent nucleocapsid however timing and amount is essential. If there is too much VP24 expressed this can lead to premature capsid condensation and inhibit the movement of L polymerase[197].

Since the synthesis and assembly of the nucleocapsid occurs in the perinuclear region of the cytoplasm in the inclusion bodies this nucleocapsid complex then needs to be transported to the plasma membrane where it can interact with VP40 and GP to form the infectious virion. VP40 dimers are transported in an actin dependent manner, likely, with the aid of GTPase Rab14 and the host COPII vesicular transport system[229–231] to lipid rafts at the plasma membrane were it oligomerizes to forms linear hexamers and octamer rings[232]. Lipid raft composition is important for the VP40 plasma membrane interactions, especially phosphatidylserine concentration[233,234]. VP35, and VP24 are the essential proteins for the transport of the nucleocapsid to the plasma membrane while VP40 is dispensable for this process. Several studies have found that active transport of the nucleocapsid complex is dependent on the actin cytoskeleton and the nucleation

factor ARP2/3. The RAC1/WAVE1/ARP2/3 pathway was recently identified as important for the propulsion of the particles via actin comet tails[227]. It appears that the polymerization of F-actin propels the nucleocapsid forward towards the plasma membrane where it interacts with VP40 to drive the egress of the viral particle.

VP40, the most abundantly expressed EBOV proteins, is critical for particle and VLP production. Assembly and transport of the nucleocapsid can occur in the absence of VP40, however budding and incorporation of GP into the nascent virion cannot. Budding occurs at filopodia in an actin dependent manner likely driven by the interaction of the nucleocapsid with VP40[225,227,235,236].Similar to HIV Gag, EBOV VP40 contains late domains (L-domains) in the N-terminus which facilitate viral egress[237–239]. The EBOV L-domain is unique in that it contains two overlapping late domains, PTAP and PPxY, the deletion of which is sufficient to impaired budding[160,239,240]. Like HIV, the ESCRT complex facilitates the budding of the virion from the plasma membrane and mediates the scission and release. Unlike HIV, the EBOV virion does not need to mature and is infectious as soon as it is released from the cell.

The immune system possesses many tools to combat these viruses. Among which is the small HERC family of proteins.

27
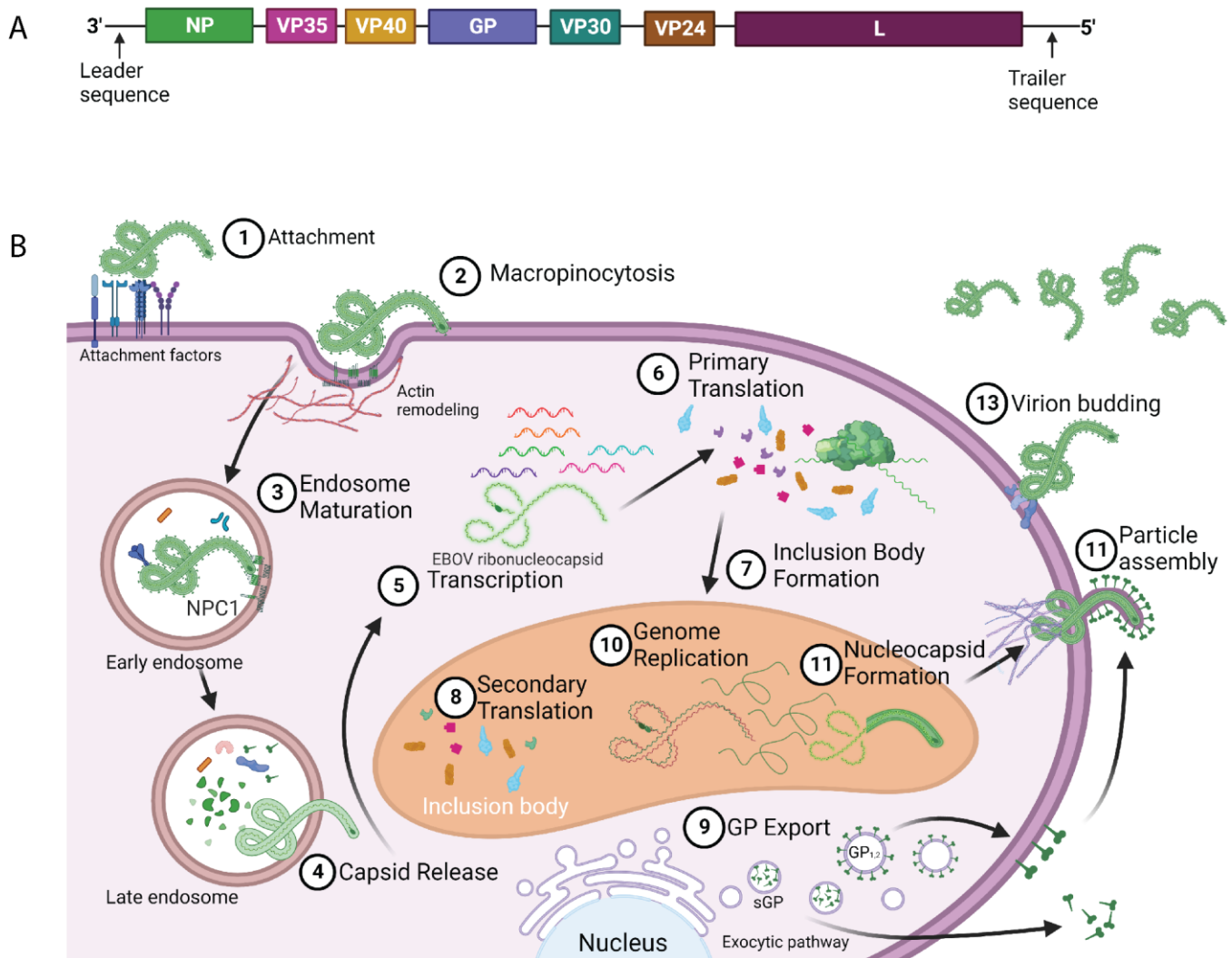
Figure 1.2. EBOV genome and replication cycle

A) The EBOV genome encodes for *GP* the glycoprotein GP$_{1,2}$ as well as the secreted glycoproteins, *NP* the nucleocapsid protein, *VP35* the polymerase cofactor, *VP30* the transcriptional activator, *VP24* the RNA complex-associated protein, and *L* the large protein RNA-directed RNA polymerase. **B)** 1) EBOV replication cycle begins with attachment to the host cell membrane. Attachment can be mediated by several host surface proteins including C-type lectin, DC-SIGN and L-SIGN, and integrins such

as α5β1 integrin, as well as TIM-1, TIM-4, and the TAM kinases TYRO3, AXL, and MERTK. 2) Macropinocytosis of the virion occurs through pathways similar to those used for apoptotic cells. Remodeling of the actin cytoskeleton is essential for this step. 3) The vesicle proceeds to mature into an early endosome and then into Rab7-positive late endosome/ lysosome compartments. The increase in acidity activates the cathepsin molecules. The GP1 subunit gets proteolytically processed in the late endosome by cysteine proteases Cathepsins B and L removing the mucin like domain and the glycan cap. 4) Binding of GP to the NPC1 receptor trigger conformational changes which are poorly understood to create a fusion pore and facilitating the release of the nucleocapsid into the cytoplasm where the replication of the genome can occur. 5) Transcription occurs within the ribonucleoprotein complex which contains NP, VP30, VP30, and L utilising the original viral genome and proteins. The mRNAs are sequentially transcribed from the 3' end to the 5' end with conventional start and stop codons preceding and trailing each gene. 6) Following transcription, all the nascent mRNAs, except for GP, are translated by the free ribosomal machinery in the cytoplasm. 7) When a certain level of NP protein is reached, it localizes near the ER forming the inclusion bodies which are a hallmark of EBOV infection. They appear in cells about 10 hours post infection, contain NP, VP35, VP30, VP24, VP40 and L. 8) Once the inclusion bodies are formed, they become the site of secondary translation. 9) GP is exported through the classical exocytic pathway with sGP being secreted into the extracellular milieu and $GP_{1,2}$ remaining anchored to the plasma membrane awaiting virion assembly. 10) Access to a bipartite motif within the 3' leader region instructs the viral polymerase to synthesize a full-length antisense genome (red). The antisense genome (antigenome) serves as a template from which the future genomes are synthesized. 11) When the switch from translation to genomic replication occurs and viral RNA is available NP molecules bind to it displacing the VP35-NP interaction and allowing for the multimerization of NP around the nascent viral RNA. Encapsidation by multimerization of NP and VP35 of the viral RNA is cooperative with the extension of the genome and signifies the first step of particle formation. 12) The nucleocapsid is transported to the plasma membrane in an actin

dependent manner where it interacts with VP40 and GP to form the viral particle. 13) Virion budding occurs at filopodia in an actin dependent manner, this is likely driven by the interaction of the nucleocapsid with VP40. The ESCRT complex facilitates the budding of the virion from the plasma membrane and mediates the scission and release.

1.3   Small HERC Family of Proteins

Humans possess six HERC proteins and were originally classified as belonging to one HERC family based on their possession of one Regulator of Chromosome Condensation 1-like domain (RLD) and at least one homologous to E6AP C-terminus (HECT) domain. However, evolutionary analysis of animal HECT ubiquitin ligases by Marin in 2010 revealed that they are evolutionarily distant and belong to two subfamilies referred to as the "small" HERCs and the "large" HERCs[241]. Additionally, his analysis proposed that the structural similarity, i.e., the presence of both an RLD and a HECT domain is due to convergence and that the RLD domains of the large HERCs and the small HERCs were attained independently. Evolutionarily, *herc4* appears to be the most ancient of the family from which *herc3* emerged through a gene duplication and chromosomal rearrangement event. This was followed by duplication of *herc3* into *herc6* and a duplication of *herc6* into *herc5*[242,243].

All the small HERC proteins (HERC3-6) contain a single RCC1 like domain at the N-terminus and a single HECT domain at the C-terminus linked by a spacer region (Figure 1.3 A). The spacer region is structurally disorganized and as of now has no known function. Although a recent study found that the RLD and HECT domains of HERC3 were dispensable for its binding to the NF-κB subunit RelA suggesting that the spacer region may be involved in this interaction[244]. Additionally, in mice, the spacer region of HERC6 is essential for facilitating ISGylation. There, the HECT domain alone was not able to ISGylate proteins but when the spacer and the HECT domain were left intact ISGylation was able to take place[245].

The RLD is found in the N-terminus and theoretically has the capacity to act as a guanine nucleotide exchange factor for Ran GTPases, however, its activity has thus far only been detected in HERC5[246,247]. Early studies suggested that the RLD and spacer region is involved in the recognition of protein targets for ISGylation, however global ISGylation is not affected by the removal of the RLD domain in HERC5[248–251].

The HECT domain comprises the C-terminal portion of the small HERCs. It has been shown to possess E3 ligase activity in all the small HERCs and can conjugate ubiquitin (Ub) or ubiquitin-like molecules (Ubl) onto substrates. Unlike other E3 ligases the HERC proteins possess the ability to both bind and transfer Ub/Ubl molecules onto substrates[252]. This process involves a cascade of reactions beginning with the ATP-dependent activation of the Ub/Ubl molecule by adenylation and the formation of a thioester bond with the Ub/Ubl at the active site cystine [252–254]. The Ub/Ubl molecule transferred to the active site cysteine of the E2 conjugating enzyme which then transfers the Ub/Ubl molecule to the E3 ligating enzyme. The HECT domain of these proteins contains a cystine residue (Figure 1.3 A, blue line in HECT domain) which forms a thioester bond with the Ub/Ubl and transfers it onto the substrate. The HECT domains of the small HERC proteins have catalytic activity similar to the E6-AP protein, for which they, are named and do not simply act as docking proteins for the E2 enzymes[252,255,256].

The small HERCs are differentially expressed in different tissues of the body and at varying levels throughout development. There are somewhat conflicting reports about the colocalization of the small HERCs with both the early endosomal compartments (colocalization with Rab5) as well as late endosomal and lysosomal compartments (colocalization with LBPA and CD68)[242,247]. The discrepancies between these reports could be due to the use of different cell lines, different antibodies, different expression methods (endogenously expressed HERC proteins or transiently over-expressed tagged proteins), or lack of imaging resolution. Interestingly, HERC3 and HERC4 were found to coprecipitate with HERC6 in an over-expression assay; however, the functional relevance of this interaction is unknown [242].

Evidence from mRNA expression levels suggests that the function of HERC proteins varies based on their intracellular location (cytoplasmic versus nuclear), the presence of other co-expressed proteins (such as UBE1L, UbcH8, ISG15 and USP18 during IFN stimulation), or the tissue they are expressed in (testis versus epithelial cells). For example, in humans, the majority of ISGylation is performed by the E3 ligase HERC5 whereas in mice this function is performed by HERC6. Based on their temporal and special expression as well as tight regulation of their levels the HERC family of proteins play a key role in development and regulation of various cellular processes. Knowledge about the functions of the small HERCs as well as findings involving the functional consequence of ISGylation will be discussed in the following sections.

A

Domain	RLD (RCC1-like domain)	HECT (Homologous to E6AP COOH terminus)
Structure	Seven-bladed β-propeller	Bilobed

B

HERC3
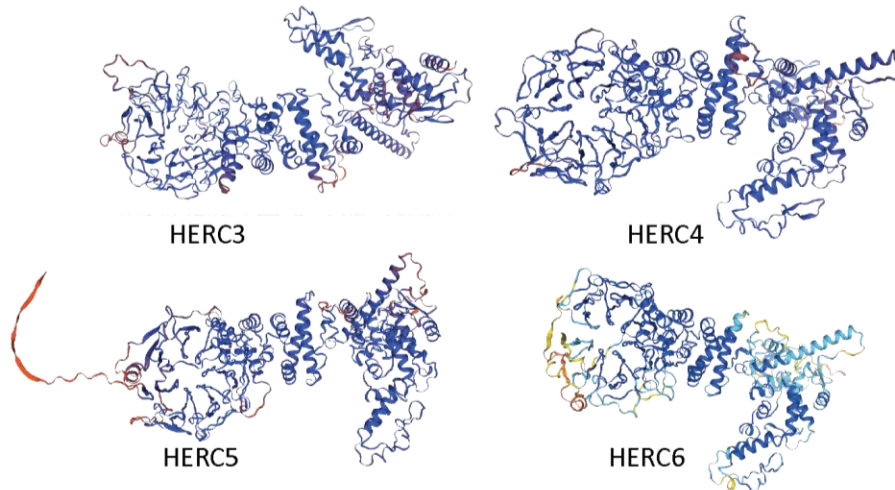
HERC4

HERC5

HERC6

C
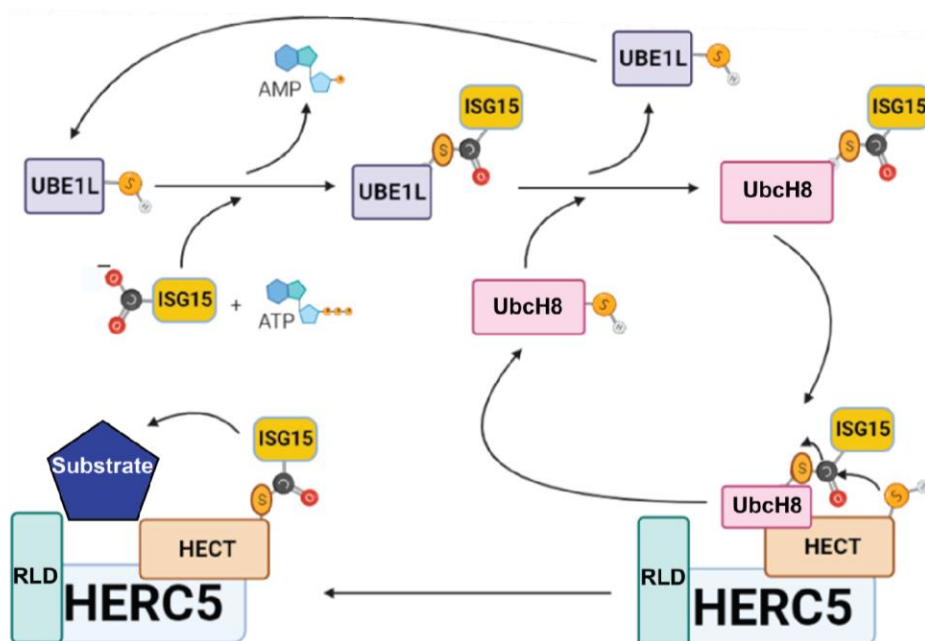
Figure 1.3 Structure of the small HERCs and ISGylation cascade

A) The small HERC proteins (HERC3, HERC4, HERC5, and HERC6) contain a single RCC1 like domain at the N-terminus and a single HECT domain at the C-terminus linked by a spacer region. The RLD domain resembles a seven bladed propeller, and the HECT domain is bilobed. Figure adapted from Sanchez-Tena, Susana & Cubillos-Rojas, Mónica & Schneider, Taiane & Rosa, Jose Luis. (2016). Functional and pathological relevance of HERC family proteins: a decade later. Cellular and Molecular Life Sciences. 73. 10.1007/s00018-016-2139-8. **B)** Predicted structures for human HERC3 (Q15034), HERC4 (Q5GLZ8), HERC5 (Q9UII4), and HERC6 (Q8IVU3). Image from the RCSB PDB (RCSB.org) (RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning (2023) Nucleic Acids Research 51: D488–D508 doi: 10.1093/nar/gkac1077). **C)** Cartoon depiction of the ISGylation pathway. ISG15 is activated in an ATP dependent manner and attached to the E1 activating enzyme UBE1L. It is then transferred to the active site of the E2 conjugating enzyme UbcH8 which then transfers ISG15 to the E3 ligase HERC5 for attachment to the substrate. Image modified from Mathieu NA, Paparisto E, Barr SD, Spratt DE. HERC5 and the ISGylation Pathway: Critical Modulators of the Antiviral Immune Response. Viruses. 2021 Jun 9;13(6):1102. doi: 10.3390/v13061102.

1.3.1 HERC3

The *herc3* gene was first identified in a screen of previously unidentified genes larger than 2 Kb [257]. It is located on chromosome 4q22 and encodes a 117 kDa protein which produces different transcripts based on three different poly(A) sites [243,247,258,259]. HERC3 is ubiquitously expressed and found at low levels in all cell types with significantly higher mRNA levels found in the fetal and adult brain, more specifically neuronal cells[243,260]. Localization studies revealed a cytosolic distribution with a small amount found in membrane fractions. Further fluorescence microscopy studies showed that HERC3 colocalizes with the endosomal markers Rab5, ARF, RhoB and EEA1, and the

late endosomal and lysosomal markers LBPA and CD63 [242,247]. Colocalization with the Golgi network, endoplasmic reticulum, mitochondria, or peroxisomes was not observed [242]. By virtue of its HECT domain, HERC3 is a ubiquitin binding E3 ligase and is itself a target of ubiquitination and proteasomal degradation, suggesting a self-regulatory mechanism [247].

More recently, it was shown that treating T98G cells with the nuclear export inhibitor Leptomycin B resulted in pronounced nuclear localization of HERC3 [261], however, previous studies did not demonstrate such an effect on Bovine aortic endothelial cells [244], suggesting that HERC3 transits the nucleus in a cell type-specific manner.

Further studies have identified HERC3 as a regulator of cell senescence and proliferation as well as a mediator of autophagy through its confirmed interacting partners hPLIC-1 (also known as Ubiquilin1 (UBQLN1)) and hPLIC-2, the RelA component of NF-κB, SMAD7, and MM1 protein [242–244,261–263]. The ubiquitinated version of HERC3 was also shown to interact with human proteins linking integrin-associated proteins and the cytoskeleton (hPLIC-1 and hPLIC-2). Ubiquitination of HERC3 did not result in its degradation but instead was shown to stabilize the protein. It is thought that binding of hPLIC-2 to ubiquitinated HERC3 inhibits its interaction with the proteasome and subsequent degradation[242].

The active site cystine at position 1018 of HERC3 contributes to higher levels of overall [244] ubiquitination but does not appear to be essential for this function. For example, the active site cysteine is not necessary for HERC3 mediated regulation of NF-κB. HERC3 inhibits the nuclear translocation of NF-κB subunit RelA through its ubiquitination at lysine 195 and 315 after disassociation from the IκBα subunit. Together with ubiquilin 1 (UBQLN1), HERC3 targets the ubiquitinated protein to the 26S proteasome. Interaction with UBQLN1 was important for this regulation but the catalytic HECT domain was not necessary[244]. Interestingly, the active site cysteine is indispensable for the ubiquitination of the Myc modulator 1 protein (MM1) by HERC3 [263]. This is part of a regulatory loop of cell senescence whereby ΔNp63α, an isoform of the p63 transcription factor, induces

transcription of HERC3 leading to the ubiquitination and degradation of the MM1 protein by the proteosome. A decrease in cellular MM1 protein leads to aberrant expression of c-Myc and its downstream target cyclin dependent kinase (CDK) 4, an important subunit in the protein complex responsible for the progressing of the cell cycle from G1 to S phase. This results in a loss of cell senescence and an increase in cell proliferation. Accordingly, HERC3, c-Myc and p53 expression is higher in invasive breast carcinoma [263]. Dysregulation of this axis is also evident in UVB-induced photoaging. Exposure to UVB radiation resulted in repression of ΔNp63α and a decrease in *herc3* transcription leading to more MM1 in the cells and subsequent proliferative senescence of keratinocytes [264]. Its involvement in cancer regulation has also been implicated by single nucleotide polymorphism studies whereby SNPs in the *herc3* gene are related to the metastatic phenotype of the tumor[265]. Along the same lines, HERC3 has been implicated in the promotion of epithelial to mesenchymal cells transformation (EMT) in glioblastomas[261]. A common chemotherapy for glioblastoma multiforme (GBM) is Temozolomide which, unfortunately, induces autophagy and EMT[266]. In a recent study, Hong Li, and colleagues[261] found that HERC3 is involved in this process. Temozolomide induces HERC3 expression leading to the ubiquitination/autosomal degradation of SMAD7 and activation of TGFβ signaling pathway. This group also showed that GMB subtypes with low HERC3 expression survived longer and that the chemoresistance of GMB was related to HERC3 levels[261].

In summary, HERC3 was identified as ubiquitin binding E3 ligase and further studies have defined it as a regulator of cell senescence and proliferation as well as a mediator of autophagy in glioblastoma cells. Different methods of investigation and different technologies have provided glimpses into the localization, function, and importance of HERC3. Continuing the research into this important molecule will provide a clearer picture into functional importance in the whole organism.

1.3.2 HERC4

The HERC4 protein was first discovered during a screen for HERC3 and HERC5 proteins in different tissue types. It is comprised of 1057 amino acids encoded by the *herc4* gene on chromosome 10q22[243]. Its localization is predominantly cytoplasmic, although, an altered nuclear localization is reported when transcription factor c-Maf is over-expressed[243]. In terms of tissue, HERC4 is highly expressed in the fetal brain with moderate expression in the adult brain and testis[243].

The most widely studied function of HERC4 is its involvement in various cancers where both pro- and anti-cancer properties have been reported. In solid tumors, upregulation of HERC4 was found to lead to metastasis and invasion of the cancer. Studies using breast cancer cell lines as well as intraductal carcinoma and invasive ductal carcinoma biopsies found that HERC4 was significantly unregulated in these tissues compared to adjacent normal breast tissue at both the mRNA and protein levels[267–269]. The increased expression of HERC4 corelated with the clinical stage and histological grade of invasive ductal carcinoma suggesting that over-expression of HERC4 leads to metastasis of cancer cells[268,270]. In the breast cancer cells lines MCF-7 and MDA-MB-231, HERC4 was upregulated compared to the normal breast cell line MCF-10A and its knockdown reduced cell survival, proliferation, and migration. HERC4 expression in breast cancer cell lines is regulated by miRNA-1285-3p and miRNA-136-5p whereby expression of these miRNAs leads to reduced levels. The mechanism by which HERC4 over-expression induces metastatic properties in these cells appears to involve HERC4-mediated ubiquitination and degradation of the tumor suppressor gene LATS1, which is a component of the Hippo signaling pathway[267]. Similar results were observed in hepatocellular carcinoma and lung cancer cell line A594, where increased expression of HERC4 led to an increase in cell proliferation and migration[271-273,267]. Reduction of HERC4 expression in HPV-infected cervical cancer cell lines via the chemical compound Andrographolide led to an increase in expression of the anti-tumor protein p53 and an increase in cell death[274].

On the other hand, there are several studies which attribute an anti-cancer effect to HERC4 expression. In non-small cell lung cancer (NSCLC) HERC4 protein levels were found to be reduced compared to normal lung tissue[275]. Downregulation and not over-expression of HERC4 lead to cell proliferation in these studies. This was corelated with HERC4 mediated ubiquitylation and degradation of Smoothered (Smo), a highly conserved protein in the hedgehog signaling pathway which plays a crucial role in embryonic development[275]. A similar trend was observed in liquid tumors such as multiple myeloma (MM), where HERC4 downregulation was corelated with cell proliferation and cancer progression[276] . In these experiments, HERC4 was identified as the E3 ligase responsible for the ubiquitination and subsequent degradation of c-Maf, a transcription factor which regulates the expression of several key drivers of cell cycle progression. Interestingly, during c-Maf over-expression HERC4 localization was altered from being cytoplasmic to being nuclear[276]. Altogether, these findings suggest that downregulation of nuclear HERC4 can lead to the formation of cancers while upregulation of HERC4 in certain cell types can lead to the development of metastasis.

Due to its high expression in the testes, some of the first functional experiments investigated the effect of HERC4 on fertility. In male mice, HERC4 deficiency caused a 50 percent reduction in fertility, while their female counterparts did not demonstrate any deficiencies in fertility, physiology, or behavior[277]. The fertility defect was linked to a reduction in sperm motility due to an abnormality in the flagellum. HERC4 deficient mice had a higher percentage of sperm with angulated tails due to the retention of a cytoplasmic droplet[277]. Similarly, dysregulation of HERC4 via administration of *Tripterygium wilfordii,* a plant used in Chinese medicine, or the *parasite Neospora caninum* also resulted in a decrease in fertility in mice[278,279].

HERC4 is an evolutionarily ancient protein believed to have emerged approximately 595 million years ago. In Drosophila it has been implicated in tissue growth and remodeling pathways involving Hippo protein kinases. The HERC4 orthologue was found to target the scaffolding protein Sav to degradation to maintain tight control over its levels[280].

However, in drosophila there is only one large HERC family member (HERC2) and one small HERC family member (HERC4) so whether this effect is translatable to mammals or humans remains to be investigated. Drosophila studies with HERC4 also found that HERC4 ubiquitylates and targets for degradation the Hedgehog pathway protein Smoothened (Smo), which is a G protein-coupled receptor involved in embryogenesis and tissue homeostasis[275,281].

In fish, HERC4 was found to have several antiviral functions. In Senegalese sole, HERC4 levels were significantly higher when fish were infected with a wildtype strain of nervous necrosis virus (NNV) compared to a less lethal mutant strain, a trend also observed for HERC5 and ISG15[282]. The higher expression of HERC4 and other innate immune response elements is believed to be due to the higher viral load observed in the tissues post infection despite the same inoculation provided with both the wildtype and mutant virus. Vaccination of Senegalese sole with a recombinant NNV resulted in upregulation of HERC4 both in response to the vaccine and in response to a viral challenge 30 days post-vaccination[283]. While in Atlantic cod HERC4 expression was slightly lower in the brain of symptomatic fish infected with the Nodavirus Atlantic cod nervous necrosis virus (ACNNV) compared to the asymptomatic Atlantic cod[284]. HERC4 was also identified as one of the most highly upregulated genes in Atlantic cod macrophages challenged with a dsRNA homologue iPC[285]. This coincided with a high expression of ISG15. These findings support the theory that HERC4 is highly induced by viral infection and as such, acts as the main E3 ligase for ISG15 in Atlantic cod. Replacement of HERC5 as the main E3 ligase has been observed in mice and pigs where HERC6 functions as the main cellular E3 ligase for ISG15[286–288]. In human cell lines, HERC4 is not strongly induced by interferon β, but its homologues HERC5 and HERC6 are interferon-induced and act as key mediators of the Type I interferon response.

1.3.3 HERC5

Originally designated as CEB1 (Cyclin E binding protein 1), HERC5 was identified in a yeast two-hybrid screen of cyclin-E and p21 binding partners [289]. The functional results

of this interaction were not investigated but HERC5 expression was increased in cells where the tumor suppressor proteins RB and p53 were absent or non-functional. HERC5 mRNA expression was shown to be induced by interferon, LPS, TNFα, and IL1-β in endothelial cells, suggesting that HERC5 plays an important role in innate immunity [290]. HERC5 proteins and mRNA are moderately expressed in the heart and placenta, and very highly expressed in the testis, spermatogonia and spermatocytes [314] . Among immune cells, its highest expression is within regulatory T-cells and neutrophils. However, it is not considered to be highly expressed under normal conditions but induced by infection and interferon stimulation. HERC5 protein expression does not always correlate with *herc5* mRNA levels since HERC5 protein levels were shown to remain low despite significantly increased mRNA levels, suggesting that HERC5 protein expression is tightly regulated [290].

One of the main functions of HERC5 is its contribution to the conjugation of the ubiquitin-like protein ISG15 to target substrates. HERC5 is the main E3 ligase for ISG15 although TRIM25 and HAARI have also been known to attach ISG15 to specific target substrates [291–293] . Type I interferon treatment induces the expression of ISG15 and all three members of the ISGylation system: the E1 activating enzyme UBE1L, the E2 conjugating enzyme UbcH8, and the E3 ligase HERC5. Interestingly, all three proteins were found to be substrates for ISGylation, suggesting a negative feedback loop could be regulating their expression. There is crosstalk between the ISGylation and the ubiquitination systems at the E2 level whereby both the E1 for ubiquitin and the E1 for ISG15 can bind to and transfer the Ubl molecule to it through a thioester bond. HERC5 in turn can also accept both ISG15 from UbcH8 and Ub form UbcH5 [291–293]. Most of the biochemical studies of HERC5 have involved the conjugation of ISG15 to substrate molecules. These will be discussed in the ISG15 and ISGylation section which follows.

Two ISGylation independent functions of HERC5 have been identified thus far, both involving the RLD domain. First, the RLD domain of HERC5 was found to coprecipitate with Ran and effect the export of REV dependent (unspliced) HIV RNA from the nucleus

to the cytoplasm by reducing intracellular levels of RanGTP and/or inhibiting the association of RanGTP with RanBP1[246]. With the ability of HERC5 to interact with Ran, it is possible that HERC5 binds and sequesters Ran in the cytoplasm, interfering with RanGDP shuttling into the nucleus. Another possibility is that HERC5 stimulates guanine nucleotide release from Ran in the cytoplasm similar to HERC1 which stimulates GDP release from the small GTPase proteins ARF1, ARF6 and Rab [246,294,295]. Other than the disruption of nuclear export of Rev dependent HIV-1 RNA, little is known about the functional effects this relationship has on the cell. The second ISGylation independent function of HERC5 is the ability to reduce intracellular viral RNA of Ebola VP40 in a manner specific to the RLD domain. This is discussed in detail in Chapter 4 of the thesis.

Due to its interferon induced nature, some of the most obvious functions of HERC5 are in pathogen eradication. To that end, several high throughput genomic studies have found a significant relationship between HERC5 expression and pathogenic infections. For example, Chikungunya virus infection increases the amount of HERC5 in cells 250-fold[296] while malaria infection byproducts significantly increased expression in PBMCs n[297]. At 24 hours post ingestion the expression of HERC5 mRNA returned to baseline levels which is consistent with previous findings suggesting that HERC5 expression is tightly regulated and is highest at 21 hours post infection then drops off [293]. Another study found that HERC5, ISG15 and HERC6 expression is related to the severity of Hand Foot and Mouth disease virus. DNA methylation of the region preceding the gene *DDX58*, encoding the dsDNA sensor RIG-1, reduced activation of IRF3 by HFMD EV1 and the expression of HERC5, HERC6 and ISG15, resulting in more severe disease[298].

The technological advancements of RNA sequencing have brought on additional information about HERC5 expression in healthy and diseased states. HERC5 expression has been linked to various cancers, however, HERC5 expression, classification (tumor suppressor or oncoprotein), and binding partners identified differ based on the type of cancer in question. Copy number loss of the chromosome 4q region which contains HERC5 and reduced HERC5 expression has been linked to poor prognosis or advanced

disease stages in pancreatic cancer, colorectal cancer, and non-small cell lung cancer (NSCLC).[299–301]. HERC5 expression was found to be lower in colorectal cancer[300] and HCC[300] samples compared to normal adjacent tissues. It contributes to the ubiquitination and likely degradation of transcriptional coreceptor C-terminal protein 1 (CtBP1) which regulates the suppression of multiple proapoptotic genes[300]. In the colorectal cancer model overexpression of HERC5 or inhibition of CtBP1 resulted in increased sensitivity to chemotherapeutic drugs[300]. These results are similar to the role HERC5 plays in non-small cell lung cancer, where hypermethylation of HERC5 promoter (located at 4q22.1), and thus under-expression of the gene correlated with positive disseminated tumor cells in the bone marrow, brain metastasis, and poor survival in both stage I adenocarcinoma and metastatic lung cancer patients. Normal lung tissue does not have HERC5 methylation but HERC5 methylation was found to be common in brain cancer samples resulting from this metastasis[302,303].

Conversely, HERC5 overexpression increased cell survival in the hepatocellular carcinoma cell lines HepG2 and SMMC-7721 and HERC5 knockdown increased apoptosis, p53 and BAX levels[304]. A chemical compound HZ-6d which downregulates HERC5 and ISG15 expression led to the activation of the p53 pathway[304]. In breast cancer studies HERC5 expression is associated with tumor grade, tumor stage and lymph node metastasis[305]. Here, qRT-PCR results showed that HERC5 was significantly higher in cancer tissues compared to adjacent healthy tissues and knocking down HERC5 inhibited tumor cell proliferation. Interestingly, USP18 was inversely associated with prognosis and its overexpression inhibited tumor cell proliferation, suggesting that the effect HERC5 has on cell proliferation is linked to its role as an ISG15 E3 ligase[305]. Similarly, in epithelial ovarian cancer chemotherapy resistant cell lines express 8 times more HERC5 than chemotherapy sensitive cell lines. The researchers believe this is related to HERC5 ISGylation of and degradation of p53 leading to aberrant cell proliferation[306]. Furthermore, recurrent tumor tissue has increased levels of CCL20 resulting from decreased expression of HERC5. CCL20 is the only chemokine overexpressed in tumor tissue where it can recruit Tregs, creating an immunosuppressive

tumor microenvironment which leads to poor prognosis[307]. Taken together these conflicting results provide a blurry picture about the role of HERC5 in cancer tissues.

In terms of other disease, HERC5 has been identified as a significant differentially expressed gene or influential factor in gestational diabetes mellitus (GDM)[308]. HERC5 mRNA expression is higher in pregnant women experiencing gestational diabetes and through the RLD domain, HERC5 is predicted to interact with the long noncoding RNA GAS5 and the neuropeptide precursor TAC1, both which are known influential factor in GDM [308]. lncGAS5 is reduced in the peripheral blood and renal tissue of patients suffering from GDM, diabetes mellitus, and neuropathy, likely through a mechanism involving HERC5, similar to the mechanism employed to reduces EBOV viral RNA (Chapter 4)[309]. The details of this interaction are theoretical (through predictive interaction models) and further studies are needed to elucidate the mechanism and specificity by which HERC5 can reduce intracellular levels of RNA. Along the same lines HERC5 was identified as a hub gene in protein interaction network for Lupus nephritis (LN)[310–312]. HERC5 expression was increased in samples LN compared to matching controls. Together, this evidence supports a pathological role for aberrant over-expression of HERC5 in the kidneys, where under normal conditions HERC5 has very low protein and RNA expression[313,314].

These studies provide evidence that under normal conditions (immune activation is not occurring), HERC5 regulation is essential to maintain natural function of the cell. Dysregulation of HERC5 expression can lead to a disease state whether it be cancer or autoimmune disorders.

1.3.4 HERC6

Herc6 was first identified in a screen for *herc3* and *herc5* genes in various tissues[243]. Although the initial discovery of HERC6 reported low expression among the cell types tested (< 5-fold), it was later shown that HERC6 is induced by changes in the cellular environment. Like HERC5, HERC6 is interferon-induced and in many animals it is

considered a potent antiviral molecule[286,295,315]. A recent evolutionary study which examined HERC5 and HERC6 proteins over mammalian evolution found evidence that there was a HERC5 duplication and reorganization with HERC6 which formed a chimera containing the HERC5 RLD and spacer regions but the HERC6 HECT domain [316]. The signature of deletions, insertions and multiple amino acid replacements suggests that HERC6 has undergone pathogen driven positive selection. Further, the spacer region of HERC6 was identified as the likely pathogen interaction interface due to the number of positive selection sites it contained across different species, a finding which is in contradiction to the current belief that its E3 activity is the main function of HERC6[316].

The majority of the HERC6 research has focused on mice where HERC6 is the main E3 ligase for ISG15, in similar fashion which HERC5 mediated ISGylation occurs in humans[286]. Recent reports suggest that HERC6, not HERC5, is the main E3 ligase in pigs and the human ISGylation machinery is not able to conjugate porcine ISG15 [404]. This was identified through a mass spectrometry screen of porcine ISGylated proteins and further biochemical analysis is required to confirm these findings. Similar to the other small HERCs, HERC6 is localized predominantly in the cytoplasm. Interestingly, HERC6 was found to interact with both HERC3 and HERC4 in a manner independent of its active site cysteine, suggesting that the small HERCs might form heterodimers or influence each other activity through direct interactions[242].

To study the functions HERC6 might perform independent of its role as the main E3 ligase for ISG15 a system of knockout mouse model was developed. As expected, knockout mice which lack HERC6 also lack ISGylation, cementing its role as the main E3 ligase for ISG15 [286,317]. However, knockout mice for UBE1L, the E1 protein required for the activation of ISG15 also lack ISGylation, thus, physiological characteristics which are shared between the two mice can be considered ISGylation specific and physiological characteristics only found in HERC6 knockout mice can be attributed to HERC6 alone[317]. One such effect is seen in sperm sack morphology of knockout mice[317]. These mice display epithelial hyperplasia in the seminal vesicles, a defect not observed in UBE1L

knockout mice despite their lack of ISGylation. The exact role of HERC6 in male fertility has yet to be deciphered but a study of DNA methylation profiles in human sperm of smokers and non-smokers found that HERC6 DNA methylation is reduced in smokers. Smokers in this study also had reduced semen volume, reduced sperm count, and reduced spermatozoa motility[318].

The recent era of computational biology and genomics has highlighted the expression profile of HERC6 in different situations, mainly organismic stress induced by viral infections and cancer. In terms of its antiviral activities, it has been studied in relation to Influenza B virus, vesicular stomatitis virus (VSV), Newcastle disease virus (NDV), respiratory syncytial virus, human rhinovirus infections, Hepatitis C, West Nile virus, and HIV as well as in relation to parasitic resistance and Mycobacterium avium resistance; however, this has mainly been linked to its ability to act as the main E3 ligase for ISG15 in mice[319–322].

Interestingly, studies looking for gene signatures to distinguish between the different immune responses elicited by viruses and bacteria found that HERC6 expression is part of an 11 gene signature which was able to distinguish between bacterial and viral respiratory infections in a clinical setting[323]. This gene signature which contained 10 other interferon induced genes, including HERC5, could also distinguish influenza virus infection from other viral infections such as respiratory syncytial virus and human rhinovirus infections. Expression of these genes also increased in participants who produced an immune response in reaction to the influenza virus vaccine and can potentially be used to determine if vaccination was successful [323]. This was corroborated in a more recent study examining the gene signatures predictive of SARS-CoV-2 or bacterial respiratory infection. Here, a three gene signature of HERC6, IGF1R and NAGK had a sensitivity of 97.3% (85.8−99.9), and specificity of 100% (63.1−100) when distinguishing between viral and bacterial infection and sensitivity of 88·6%, and specificity of 94.1% when identifying COVID-19 infections, a significant improvement

to the current measurements of C-reactive protein and leukocyte counts currently used for these predictions[324].

HERC6 was also implicated as an important early expression gene in the battel against Hepatitis C. The least expensive treatment for Hepatitis C is a 48-week course of pegIFN/ribavirin, which has a high probability of failure and causes severe side effects to the patient[325]. Lu et al. (2016) found that the early expression of 8 genes including HERC5 and HERC6 were able to accurately predict which patients would respond well (sustain a virological response) and which would benefit from a different course of treatment. Compared to those with non-sustained virological response patients with sustained virological response had increased expression of HERC5, HERC6, ISG15 and USP18 at week one post treatment[325].

Surprisingly, HERC6 has also been implicated as an important gene for agricultural breeders. RNAseq studies have linked HERC6 expression to body weight determinants in sheep and cows. In sheep it was identified as an important gene influencing economic traits such as muscle weight in carcass, body weight, and milk protein percentage[327,328]. HERC3, HERC5, and HERC6 SNPs are associated with resistance to parasites and as such were identified as targets for breeding of sheep to develop antiparasitic resistance[329]. Whereas in cows, HERC6 was identified as a protein important to the transference of feed intake into body fat and protein[330]. Additionally, Cattle which developed severe disease and had to be treated multiple times for respiratory disease had increased expression of type-I interferon activity (HERC6, IFI6, ISG15, MX1).

Overall, these studies suggest that the regulated expression of HERC6 is linked to a healthy immune system which is able to efficiently fight off pathogens and maintain normal metabolic function.

1.3.5 ISG15 and ISGylation

Initially identified as just another interferon induced gene, ISG15 was later classified as a ubiquitin like molecule (Ubl) with potent antiviral properties[331]. ISGylation involves the

covalent posttranslational attachment of ISG15 to host and viral target substrates by a specialized group of IFN-α/β induced E1-E2-E3 ubiquitin cascade enzymes. For example, protein–protein interaction studies have shown that ISG15 only coordinates with five of the over 600 identified E1–E2–E3 enzymatic members[332]. The ISGylation pathway begins when UBE1L activates ISG15 through an ATP-dependent mechanism to form a thioester bond between the C-terminal carboxyl of ISG15 and the catalytic cysteine of UBE1L[333,334]. Following activation, ISG15 is transferred by UBE1L to UbcH8 via a trans-thiolation reaction that forms a thioester bond between the C-terminus of ISG15 and the conserved catalytic cysteine residue of UbcH8[335,336]. The UbcH8–ISG15 thioester complex then transfers its ISG15 cargo to the catalytic cysteine (C994) of the ISG15-specific homologous to E6AP C-terminus (HECT) E3 ligase HERC5. Thereafter, HERC5 catalyzes the covalent attachment of ISG15 onto host and viral substrate proteins by forming an isopeptide bond between the C-terminus of ISG15 and the ε-amino group on the lysine of the respective protein target[337,338]. See Figure 1.3 C.

The ubiquitylation pathway enzymes are constitutively expressed and can form a variety of ubiquitin chain linkages with their target substrates[341–344]. On the other hand, the ISG15-specific E1–E2–E3 cascade enzymes are only induced following inflammation and exclusively form monomer isopeptide linkages[345]. This suggests that the ISG15 signaling pathway provides the cell with a highly specialized and tightly regulated antiviral function at the expense of dysregulating vast cellular activities [346]. While ubiquitination of proteins often leads to their degradation by the proteosome, most studies suggest that ISGylation does not lead to degradation of proteins but more often to the protection from ubiquitination and subsequent degradation. Accordingly, the most reported function of ISGylation is altered protein localization and modified protein function. [245,249–251,288,355–359]. Interestingly, a recent publication observed that hybrid ISG15–ubiquitin chains can also be formed by the cell to regulate protein homeostasis suggesting the interplay between ubiquitin and ISG15 is not completely understood [347].

The expression of the ISGylation system has a significant impact on the cellular environment and normal protein function, making its stringent regulation essential for cell survival[349]. Hyper-ISGylation through the deletion of USP18, the main de-ISGylase, often leads to apoptosis in hemopoietic tissue, brain cell injury, and decreased life expectancy in mice[249]. Whereas in humans, ISG15 deficiencies are extremely rare but not fatal. These diseases are associated with mycobacterial hypersensitivity, brain calcification, and skin lesions[353,354]. Therefore, understanding how ISG15 influences cellular function during a viral infection is essential to clarifying the diverse range of biochemical outcomes that can occur as a consequence of host ISGylation activity.

ISG15 falls in the "ISGF3/IRF-3" group of interferon stimulated genes, meaning that both IRF-3 and ISGF3 can induce its expression in cells[348]. Expression of UBE1L, UbcH8 and HERC5 occurs through the IRF-3 pathway following detection of viral pathogens as well as through the ISGF3 pathway following induction of Interferon α/β[349,350]. HERC5 expression can be induced by IL-1β and TNFα although not to the extent that is observed with Interferon β, and more recent studies have suggested that the ISGylation system can be induced by NF-κB[351,352].

A 2003 study conducted by Malakohova et al. was the first to identify that ISG15 positively regulates the host antiviral response[360]. This was a follow-up study to a high-throughput western blotting screen which identified PLCγ1, ERK-1, JAK-1 and Stat1 as being ISGylated[360]. They found that in UBP43 (also known as a USP18) null cells there was a maintenance of strong interferon response and hypersensitivity to interferon inducers such as poly(I:C). Additionally, the JAK/STAT pathway activation was maintained for longer in USP18 null, ISGylation proficient cells. Reconstitution of protein ISGylation lead to increased interferon signaling in these studies. Meanwhile, in LPS induced inflammation of microglia cells, ISGylation of Stat1 increased its stability and the expression of downstream target proteins implicating it as a positive regulator of the antiviral response[361]. Soon after, additional evidence emerged indicating that ISG15 and ISGylation contribute to the regulation of the innate immune response. For example,

ISGylation of Interferon regulatory factor 3 (IRF3) by HERC5 leads to its stabilization[362–364]. Following detection of viral PAMPs by TLR3 and TLR4 or detection of viral RNA by the cGAS/STING pathway IRF3 is phosphorylated, dimerizes, and is translocated into the nucleus where it acts as a transcription factor for Type1 interferons and proinflammatory genes[365]. The potency of this protein requires a finely tuned response to environmental signals and is evident by the multitude of post-translational modifications which control activation and function. HERC5 ISGylates IRF3 at lysines 193, 360, and 366, inhibiting its ubiquitination by Pin-1 and subsequent degradation. Stabilization of IRF3 by ISGylation acts as a positive feedback loop to sustain the induction of interferons, interferon stimulated genes, and prevent the premature termination of the antiviral response[364].

Additionally, early studies found that UbcH8 and ISG15 conjugation are involved in the regulation of RIG-I by inhibiting its activation via ubiquitination[335,366]. RIG-I and melanoma differentiation-associated protein 5 (MDA5) are major RNA sensors in the cytosol and together represent the RIG-I-like receptors (RLR) family of antiviral proteins. Binding of foreign or mislocated RNA to the C-terminal domain and helicase of RIG-I and MDA5 switches these molecules into the active state, initiating an enzymatic cascade ending with the activation of an antiviral state. ISG15 participates in a negative feedback loop whereby ISGylated RIG-I is not able to be ubiquitinated by TRIM25 and activated for downstream signaling but instead can be ubiquitinated by RNF125 and targeted for degradation[367]. RNF125 increases RIG-I, MDA5 and IPS-1 ubiquitination and destabilization leading to decreased IFNβ stimulation[357]. The recent COVID-19 pandemic stimulated research into the sensing of RNA viruses leading to the discovery that MDA5 ISGylation is important for the induction of an antiviral response[357]. MDA5 is a major RNA sensor in the cytosol and ISGylation at position K32 and K43 following PP1 mediated CARD dephosphorylation lead to the stabilization and higher order oligomerization of MDA5[368]. This regulatory role for ISG15 on MDA5 is contrary to its role on the activation and stability of RIG-I raising the possibility that ISG15 acts as a sensor switching modification in this context. As the levels of ISG15 in the cell rise they

promote MDA5 stability, activation, and cytokine induction while concomitantly deactivating RIG-I[368].

Furthermore, over-expression of the ISGylation system has been found to negatively regulate the activation of the NF-κB pathway. Phosphorylation of TAK1 leads to the ubiquitination through ISGylation and inactivation of Ubc13, an important ubiquitin E3 enzyme in the pathway[355]. More recent studies found that ISGylation of mitochondrial proteins is important for the stability of the mitochondria as well as the antiviral response to Vaccinia virus in murine bone derived macrophages. Additionally, this group found that mitophagy and oxidative phosphorylation were impaired in ISG15 null cells[369].

ISGylation and the ISGylation machinery have also been implicated in the regulation of autophagy[370]. ISGylation of BECN-1 following type I interferon induction was found to inhibit autophagy and restrain viral replication[371]. In 2019, Zhang et al[372] identified addition ISGylation targets in *Listeria monocytogenes* infected mouse livers using quantitative label-free proteomics. They confirmed the ISGylation of 87 substrates previously known to be ISGylated and added an additional 347 proteins to that list. They found that in this model, ISGylation resulted in an increase in extracellular vesicle and protein secretion. Additionally, Listeria infection resulted in an increase of ISGylated mitochondrial proteins. In many cases the site of ISGylation overlapped with acylation sites, dimerization domains, and active sites further supporting ISGylation as a regulatory mechanism in the antimicrobial response.

The studies mentioned above demonstrate that ISGylation feeds into self-regulatory feedback loops which control the expression of a multitude of interferon stimulated genes, for which both positive and negative effects have been observed on upstream effectors of the antiviral response. In biologically relevant contexts, the most likely scenario is that the induction of the ISGylation system has a regulatory effect, the sum result of which is dependent on the quantity of viral proteins and RNAs that are sensed, the expression levels of ISGs, and the temporal progression of the host-virus battle. During the early stages of infection, ISGylation of key immune factors likely feeds into a

positive feedback loop increasing the production of antiviral proteins and inhibiting the premature termination of the interferon response. Later, as the immune response controls the infection, it likely feeds into negative feedback loops, preventing the maintenance of unnecessary inflammation and resulting tissue damage.

The antiviral effects of ISGylation have been identified against several viral proteins. For example, ISGylation (i) stalls IAV replication by blocking NS1 protein homodimerization [346], (ii) promote HCV proliferation via improved cyclophilin A recruitment by NS5A proteins[373] and, alternatively, (iii) inhibit an early stage of HIV assembly by attenuating Gag-particle production [132,374]. In the upcoming sections I will present some of the recent discoveries involving the ISGylation of viral protein targets.

1.3.5.1   Role of ISGylation in HCV replication

The antiviral effects of HERC5 and ISG15 on HCV were first observed by Jung Kim et al., who were examining mechanisms of HCV replication cycle inhibition[375]. By conducting site-directed mutagenesis experiments on the HCV protein NS5A, the group was able to demonstrate that HERC5 inhibits HCV replication by ISGylating NS5A at lysine 379 (K379) which subsequently targets it for K48 polyubiquitylation by an unknown E3 ligase[375]. The researchers confirmed this antiviral activity by demonstrating that HCV replication was unrestrained when ISG15 and its corresponding E1–E2–E3 enzyme cascade were expressed in the presence of HCV NS5A K379R variant proteins. Taken together, these findings suggest that NS5A is the primary HCV target substrate of HERC5-dependent ISGylation, and that K379 is the sole NS5A residue that HERC5 targets for ISGylation. More recently, Abe et al. demonstrated that the HCV NS5A protein is prone to ISGylation at five Lys residues (K44, K68, K166, K215 and K308)[373]. In fact, one of the ISG15 attachment points on NS5A, K308, is located within the NS5A-binding region for cyclophilin A (CypA), a virulence factor that is required for efficient HCV cellular propagation. Thus, HERC5 ISGylation of NS5A at K308 enhances HCV recruitment of CypA to provide a pro-virulent replication activity, a direct contradiction of previous findings from Kim et al.[375]. Moreover, a stand-alone study conducted by

Domingues et al. found that ISG15-related forms of HCV inhibition occur independently of HERC5 activity[376]. To date, no follow up research has been conducted on the HERC5–HCV ISGylation system to confirm whether HERC5 ISGylation activity results in a pro- or anti-virulent response to HCV infection, and whether HERC5 is necessary for catalyzing the modes of ISG15 inhibition that have been observed for HCV NS5A proteins.

1.3.5.2 Role of ISGylation in Influenza Virus replication

New research by Tang et al. has shown that HERC5 targets IAV NS1 for ISGylation to prevent IAV capsid formation[346]. Using pulldown assays and immunoprecipitation analysis, the researchers found that HERC5 interacts with the ribosomal-binding (RBD) and C-terminal effector domains (ED) of NS1, and that both interactions were required to form stable HERC5–NS1 complexes[346]. Lysine residue substitutions in IAV NS1 also revealed that HERC5 attaches ISG15 at multiple NS1 lysine residues (K20, K41, K217, K219, K108, K110 and K126), with the strongest inhibitory effect coming from the ISGylation of the K126 and K217 residues in the ED and RBD domains, respectively[346]. The ISGylation of NS1 subsequently abolished the ability of NS1 to interact with protein kinase R (PKR) and blocked NS1 RBD-dependent homodimerization, thereby significantly inhibiting IAV capsid assembly in vivo[346]. Certain avian flu IAV strains, such as H5N1, demonstrated a higher rate of NS1 K217R mutation compared to most seasonal flu strains[346,377]. It was also observed that avian IAV variants were less susceptible to ISGylation at K126, suggesting that, unlike other IAV strains, avian IAV NS1 proteins may adopt a new structure that obstructs the K126 ISGylation site from HERC5. Given that K126 and K217 have been determined as the primary ISGylation sites involved with preventing IAV capsid formation, these discoveries provide a possible explanation for why avian strains of the flu are more infectious and lethal than other IAV strains. Despite HERC5 demonstrating minimal antiviral activity against certain avian IAV strains, these findings from Tang et al. indicate that HERC5 and its ISGylation activity could serve as prime drug targets for the development of treatments that are used to combat most seasonal IAV strains.

1.3.5.3   Role of ISGylation in HIV and retrovirus replication

A pivotal study by Woods et al. in 2011 confirmed that HERC5 prevents an early stage of HIV-1 viral assembly by ISGylating proteins involved with Gag polyprotein (Pr55Gag) particle production[132] . The researchers used confocal immunofluorescence microscopy to reveal that IFN-I-induced HERC5 localizes to the cytoplasm where it forms punctuate bodies in a variety of cell lineages, and that these bodies interact with polyribosomes[132]. These findings were consistent with a previous study conducted by Durfee et al., who observed that HERC5 associates with the 60S ribosomal subunit of the polyribosome using cell fractionation[248,255]. Additionally, IFN-I-induced HERC5 and HIV Gag proteins colocalize to the plasma membrane where HERC5 ISGylates Gag particles to prevent HIV-1 viral budding. Importantly, the same research group has shown that other retroviruses such as murine leukemia virus (MLV) and simian immunodeficiency virus (SIV) are also inhibited by HERC5[379]. Interestingly, although SIV was inhibited by human HERC5, an ancestral version of HERC5 found in coelacanth fish was unable to inhibit SIV replication, suggesting that the HERC5 gene has evolved to combat lentiviruses in primates[379].

1.3.5.4   Role of ISGylation in Ebolavirus replication

Ebola virus infection induces strong expression of both ISG15 and HERC5 starting at 1 day post infection and lasting at least until day 7[380–383]. Early studies reported that ISG15 over-expression inhibited NEDD4 from ubiquitinating Ebolavirus VP40 and decreased VLP production[384,385]. This effect was dependent on the inhibition of ubiquitination of the late domain PTPY on VP40. More recent studies by the Barr lab have found that HERC5 inhibits EBOV replication in a pseudotype system (See Chapter 4). This inhibition is dependent on the RLD domain of HERC5 and is correlated with a reduction in viral RNA. Interestingly, this inhibition of EBOV replication was antagonized by the EBOV glycoprotein[309].

1.3.5.5   Role of ISGylation in HPV replication

In a proteomic screen to identify ISGylation targets, Durfee et al found that HERC5 ISGylates the L1 capsid protein of HPV16. L1 is the capsid protein of HPV and highly translated during active viral infection. Expression of the ISGylation system significantly reduced viral particle production and infectivity. ISGylation of a small fraction of the total L1 protein led to a dominant inhibitory effect, likely through inhibition of capsid assembly. Since then, no other studies have examined the effect of ISGylation on HPV replication[248].

1.3.5.6   Role of ISGylation in Vaccinia virus replication

In mice it was found that Vaccinia virus replication was enhanced in ISG15 knockout cells and was inhibited in cells expressing wildtype levels of ISG15[386]. This led to the investigation of the antiviral mechanism of ISG15 against Vaccinia virus in ISG15 null cells leading to the discovery that ISG15 deficient cells were more resistant to apoptosis and had impaired phagocytic activity when coming in contact with infected cells[387]. The effect of ISG15 and HERC5 mediated ISGylation of Vaccinia virus is antagonized by the Vaccinia virus E3 protein[388]. When E3 was expressed, there was a significant decrease in ISGylated proteins compared to the delta E3 version of the virus. ISGylation of mitochondrial proteins is important for the regulation of oxidative phosphorylation and decreasing reactive oxygen species[369]. The decrease in ISGylated proteins corresponded with a decrease in ISGylated mitochondrial protein and impaired mitophagy in an infection model. Overall, these studies have demonstrated that control of Vaccinia virus is dependent on the presence of ISG15 and lack of ISGylation, either through ISG15 knockout or deISGylation of proteins by Vaccinia virus E3, leads to enhanced infection kinetics.

1.3.5.7   Role of ISGylation in Kaposi's Sarcoma-Associated Herpesvirus (KSHV) replication

ISG15 and HERC5 were identified as interactors of the KSHV viral homologue of interferon regulatory factor 1 (vIRF1). HERC5 inhibits KSHV in an ISGylation

dependent manner and knockdown of either HERC5 or ISG15 prior to viral reactivation resulted in higher titers of infectious viral particles[389]. Expression of vIRF1 resulted in a decrease of total protein ISGylation. Later studies suggested that expression of ISG15 may lead to viral latency since knockdown of ISG15 and ISG20 lead to lytic reactivation the virus[390]. A more recent study implicated CRM1 in the inhibition of KSHV via retention of autophagy adaptor protein p62 (SQSTM1) in the nucleus which led to increased expression of antiviral genes[391]. Interestingly, the RLD domain of HERC5 also inhibits nuclear export in a CRM1 dependent manner[246]. Although it has not been directly studied, it is possible that in addition to the inhibition of KSHV through ISGylation of cellular and viral proteins, HERC5 inhibits the lytic phase of the virus by a second mechanism involving the CRM1 pathway.

1.3.5.8   Role of ISGylation in Cytomegalovirus replication

Although it has been known for some time that ISG15 and the ISGylation system is induced by cytomegalovirus infection, its role and direct interaction with viral protein was only recently described[392]. Kim et al demonstrated that knocking down ISG15 or HERC5 lead to a significant increase in viral titers while over-expression of the protein ISGylation system led to decreased viral particle production. Additionally, expression of ISG15 and HERC5 inhibited the viral replication cycle by reducing the expression of viral genes as well as disabling viral budding at the plasma membrane[392]. Two antagonists of HERC5 and ISG15 were discovered in this study. First, the IE1 protein suppressed ISGylation through a reduction in ISG15 transcription likely through inhibition of STAT2 dependent interferon response activation. Second, UL26 interacts with ISG15 and HERC5 to inhibit protein ISGylation[392,393]. UL26 has been implicated as a major regulator of the innate immune response against HCMV. Tegument delivered UL26 can limit the induction of ISG expression[394]. Additionally, the pUL50 protein binds to and is involved in the downregulation of UBE1L resulting in decreased global ISGylation[395]. Much of the research into cytomegalovirus infection and its inhibition by HERC5 and ISG15 have focused on the antagonism of the antiviral effects by HCMV proteins however little is known about the specific effect of ISGylation on viral proteins

other than it has a dominant inhibitory effect. Further studies are needed to elucidate the precise mechanism of inhibition by HERC5 and ISG15.

1.3.5.9 Role of ISGylation in Zika Virus replication

Upon Zika virus infection HERC5 and ISG15 are induced 481-fold and 28-fold, respectively, over the uninfected control in primary human brain microvascular endothelial cells. As the infection progresses HERC5 levels drop almost 10-fold while ISG15 levels increase another 60-fold compared to the control by day 9 of the infection[396]. The consequence of increased expression of HERC5 and ISG15 is somewhat controversial with both proviral, and antiviral effects having been proposed[359,397–399]. The discrepancy in the results may be due to either the timeframe of the experiments or the cell type used. Wang et al[359] performed the experiments in a 24-hour infection time frame whereas both this study and others showed that induction of ISG15 in increased with time and corelated with a decrease in viral protein at 96-hours post infection. It is clear that additional studies in primary cell lines with longer periods of follow up are needed to tease out the true effect of ISG15 and HERC5 in Zika Virus infection.

Based on the previous work by the Barr lab and the above mentioned finding we sought to further investigate the small HERC family of proteins with a specific focus on their structural and functional contribution to the innate immune system. Our hypothesis is that the small HERC family of proteins have evolved through gene duplication events to become important interferon induced antiviral proteins, which function to inhibit the activity of diverse RNA viruses such as HIV and EBOV. We found that HERC3 and 4 are evolutionarily distant to HERC5 and HERC6 and are not as strongly induced by interferon treatment. Additionally, they did not exhibit strong antiviral activity against HIV-1. Since HERC5 and HERC6 are the small HERC family members exhibiting the strongest antiviral activity, in Chapter 3 we continued to investigate the substitution of an amino acid residue on HERC6 which confers antiviral activity against HIV as strong as that observed with HERC5. We found that HERC6 R10G inhibits the production of infectious particles by a yet unknown mechanism. In Chapter 4, to determine the breadth

of HERC5 as a strongly induced antiviral protein, we investigated the ability of HERC5 to restrict EBOV. We found that HERC5 strongly inhibits EBOV structural protein VP40 in *in vitro* assays through depletion of its RNA. This inhibition was antagonized by EBOV GP. All together these studies advance the knowledge of the small HERC proteins both structurally and functionally and provides additional information about the host protein interactions during viral infection with HIV and EBOV as well as potential targets for drug development.

1.4   References

1.   Huber C, Finelli L, Stevens W. The Economic and Social Burden of the 2014 Ebola Outbreak in West Africa. *J Infect Dis*. 2018;218(suppl_5):S698-S704. doi:10.1093/INFDIS/JIY213

2.   Cost of the Ebola Epidemic | History | Ebola (Ebola Virus Disease) | CDC. Accessed September 14, 2022. https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/cost-of-ebola.html

3.   Jacob ST, Crozier I, Fischer WA, et al. Ebola virus disease. *Nature Reviews Disease Primers 2020 6:1*. 2020;6(1):1-31. doi:10.1038/s41572-020-0147-3

4.   Hoenen T, Groseth A, Feldmann H. Therapeutic strategies to target the Ebola virus life cycle. *Nature Reviews Microbiology 2019 17:10*. 2019;17(10):593-606. doi:10.1038/s41579-019-0233-2

5.   Maganga GD, Kapetshi J, Berthet N, et al. Ebola virus disease in the Democratic Republic of Congo. *New England Journal of Medicine*. 2014;371(22):2083-2091. doi:10.1056/NEJMoa1411099

6.   Bausch DG. West Africa 2013 ebola: From virus outbreak to humanitarian crisis. *Curr Top Microbiol Immunol*. 2017;411:63-92. doi:10.1007/82_2017_69

7.   Feldmann H, Sprecher A, Geisbert TW. Ebola. *New England Journal of Medicine*. 2020;382(19):1832-1842. doi:10.1056/NEJMra1901594

8.   Malvy D, McElroy AK, de Clerck H, Günther S, van Griensven J. Ebola virus disease. *The Lancet*. 2019;393(10174):936-948. doi:https://doi.org/10.1016/S0140-6736(18)33132-5

9.   Centers for Disease Control and Prevention. Cost of the Ebola Epidemic. Published March 8, 2019. Accessed June 17, 2020. https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/cost-of-ebola.html

10.   Hira S, Piot P. The counter effects of the Ebola epidemic on control and treatment of HIV/AIDS, tuberculosis, and malaria in West Africa. *AIDS*. 2016;30(16):2555-2559. doi:10.1097/QAD.0000000000001231

11.   Parpia AS, Ndeffo-Mbah ML, Wenzel NS, Galvani AP. Effects of Response to 2014-2015 Ebola Outbreak on Deaths from Malaria, HIV/AIDS, and Tuberculosis, West Africa. *Emerg Infect Dis*. 2016;22(3):433-441. doi:10.3201/EID2203.150977

12.   Evans DK, Goldstein M, Popova A. Health-care worker mortality and the legacy of the Ebola epidemic. *Lancet Glob Health*. 2015;3(8):e439-e440. doi:10.1016/S2214-109X(15)00065-0

13.   Scott JT, Sesay FR, Massaquoi TA, Idriss BR, Sahr F, Semple MG. Post-Ebola Syndrome, Sierra Leone. *Emerg Infect Dis*. 2016;22(4):641-646. doi:10.3201/EID2204.151302

14.   Jagadesh S, Sevalie S, Fatoma R, et al. Disability among Ebola Survivors and Their Close Contacts in Sierra Leone: A Retrospective Case-Controlled Cohort Study. *Clinical Infectious Diseases*. 2018;66(1):131-133. doi:10.1093/cid/cix705

15.   Makoni M. Ebola outbreak in DR Congo. *Lancet*. 2022;399(10337):1766. doi:10.1016/S0140-6736(22)00819-4

16.   Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med*. 2011;1(1). doi:10.1101/cshperspect.a006841

17.   Faria NR, Rambaut A, Suchard MA, et al. The early spread and epidemic ignition of HIV-1 in human populations. *Science (1979)*. 2014;346(6205):56-61. doi:10.1126/science.1256739

18.   Greene WC. A history of AIDS: Looking back to see ahead. *Eur J Immunol*. 2007;37(SUPPL. 1). doi:10.1002/eji.200737441

19.   Global HIV & AIDS statistics — 2019 fact sheet | UNAIDS. Accessed June 22, 2020. https://www.unaids.org/en/resources/fact-sheet

20.   Simon V, Ho DD, Abdool Karim Q. HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. *Lancet*. 2006;368(9534):489-504. doi:10.1016/S0140-6736(06)69157-5

21.   Eisinger RW, Fauci AS. Ending the HIV/AIDS pandemic. *Emerg Infect Dis*. 2018;24(3):413-416. doi:10.3201/eid2403.171797

22.   Global HIV & AIDS statistics — Fact sheet | UNAIDS. Accessed November 28, 2022. https://www.unaids.org/en/resources/fact-sheet

23.   Global Statistics. Accessed November 28, 2022. https://www.hiv.gov/hiv-basics/overview/data-and-trends/global-statistics

24.   Full report — In Danger: UNAIDS Global AIDS Update 2022 | UNAIDS. Accessed November 28, 2022. https://www.unaids.org/en/resources/documents/2022/in-danger-global-aids-update

25. Canada sees its fourth consecutive year of increasing HIV rates - Canadian AIDS Society. Accessed June 22, 2020. https://www.cdnaids.ca/canada-sees-its-fourth-consecutive-year-of-increasing-hiv-rates/

26. Kingston-riechers J. The Economic Cost of HIV / AIDS in Canada. Published online 2011.

27. Jamieson D, Kellerman SE. The 90 90 90 strategy to end the HIV pandemic by 2030: Can the supply chain handle it? *J Int AIDS Soc*. 2016;19(1):1-4. doi:10.7448/IAS.19.1.20917

28. Estimates of HIV incidence, prevalence, and Canada's progress on meeting the 90-90-90 HIV targets, 2020 - Canada.ca. Accessed November 28, 2022. https://www.canada.ca/en/public-health/services/publications/diseases-conditions/estimates-hiv-incidence-prevalence-canada-meeting-90-90-90-targets-2020.html#a2

29. Koonin E v., Dolja V v., Krupovic M. Origins, and evolution of viruses of eukaryotes: The ultimate modularity. *Virology*. 2015;479-480:2-25. doi:10.1016/J.VIROL.2015.02.039

30. Koonin E v., Makarova KS, Wolf YI, Krupovic M. Evolutionary entanglement of mobile genetic elements and host defense systems: guns for hire. *Nature Reviews Genetics 2019 21:2*. 2019;21(2):119-131. doi:10.1038/s41576-019-0172-9

31. Simmonds P, Aiewsakun P. Virus classification – where do you draw the line? *Arch Virol*. 2018;163(8):2037. doi:10.1007/S00705-018-3938-Z

32. Shi M, Zhang YZ, Holmes EC. Meta-transcriptomics, and the evolutionary biology of RNA viruses. *Virus Res*. 2018;243:83. doi:10.1016/J.VIRUSRES.2017.10.016

33. Koonin E v., Krupovic M, Agol VI. The Baltimore Classification of Viruses 50 Years Later: How Does It Stand in the Light of Virus Evolution? *Microbiol Mol Biol Rev*. 2021;85(3). doi:10.1128/MMBR.00053-21

34. D B. Expression of animal virus genomes. *Bacteriol Rev*. 1971;35(3):235-241. doi:10.1128/BR.35.3.235-241.1971

35. Koonin E v., Dolja V v., Krupovic M, et al. Global Organization and Proposed Megataxonomy of the Virus World. *Microbiol Mol Biol Rev*. 2020;84(2). doi:10.1128/MMBR.00061-19

36. Gorbalenya AE, Krupovic M, Mushegian A, et al. The new scope of virus taxonomy: partitioning the virosphere into 15 hierarchical ranks. *Nat Microbiol*. 2020;5(5):668-674. doi:10.1038/S41564-020-0709-X

37. Kuhn JH, Amarasinghe GK, Basler CF, et al. ICTV Virus Taxonomy Profile: Filoviridae. *J Gen Virol*. 2019;100(6):911-912. doi:10.1099/JGV.0.001252

38. Moore MD, Hu WS. HIV-1 RNA dimerization: It takes two to tango. *AIDS Rev*. 2009;11(2):91-102.

39. Seitz R. Human Immunodeficiency Virus (HIV). *Transfus Med Hemother*. 2016;43(3):203-222. doi:10.1159/000445852

40. Freed EO. HIV-1 replication. *Somat Cell Mol Genet*. 2001;26(1-6):13-33. doi:10.1023/a:1021070512287

41. Lerner G, Weaver N, Anokhin B, Spearman P. Advances in HIV-1 Assembly. *Viruses*. 2022;14(3). doi:10.3390/v14030478

42. Sundquist WI, Kräusslich HG. HIV-1 assembly, budding, and maturation. *Cold Spring Harb Perspect Med*. 2012;2(7):a006924. doi:10.1101/cshperspect.a006924

43. Freed EO. HIV-1 assembly, release, and maturation. *Nat Rev Microbiol*. 2015;13(8):484-496. doi:10.1038/nrmicro3490

44. Bell NM, Lever AML. HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol*. 2013;21(3):136-144. doi:10.1016/j.tim.2012.11.006

45. Strebel K. HIV Accessory Proteins versus Host Restriction Factors. *Curr Opin Virol*. 2013;3(6):692-699. doi:10.1016/J.COVIRO.2013.08.004

46. Frankel AD, Young JAT. HIV-1: Fifteen Proteins and an RNA. *https://doi-org.proxy1.lib.uwo.ca/101146/annurev.biochem6711*. 2003;67:1-25. doi:10.1146/ANNUREV.BIOCHEM.67.1.1

47. Galvin SR, Cohen MS. The role of sexually transmitted diseases in HIV transmission. *Nature Reviews Microbiology 2004 2:1*. 2004;2(1):33-42. doi:10.1038/nrmicro794

48. Burgener A, McGowan I, Klatt NR. HIV and mucosal barrier interactions: consequences for transmission and pathogenesis. *Curr Opin Immunol*. 2015;36:22-30. doi:10.1016/J.COI.2015.06.004

49. Petrova MI, van den Broek M, Balzarini J, Vanderleyden J, Lebeer S. Vaginal microbiota and its role in HIV transmission and infection. *FEMS Microbiol Rev*. 2013;37(5):762-792. doi:10.1111/1574-6976.12029

50. Stone A. Microbicides: a new approach to preventing HIV and other sexually transmitted infections. *Nature Reviews Drug Discovery 2002 1:12*. 2002;1(12):977-985. doi:10.1038/nrd959

51. Devito C, Ellegård R, Falkeborn T, et al. Human IgM monoclonal antibodies block HIV-transmission to immune cells in cervico-vaginal tissues and across polarized epithelial cells in vitro. *Sci Rep*. 2018;8(1). doi:10.1038/S41598-018-28242-Y

62

52. Hurst SA, Appelgren KE, Kourtis AP. Prevention of mother-to-child transmission of HIV type 1: the role of neonatal and infant prophylaxis. *Expert Rev Anti Infect Ther*. 2015;13(2):169-181. doi:10.1586/14787210.2015.999667

53. Kerr T, Small W, Buchner C, et al. Syringe sharing and HIV incidence among injection drug users and increased access to sterile syringes. *Am J Public Health*. 2010;100(8):1449-1453. doi:10.2105/AJPH.2009.178467

54. Saïdi H, Magri G, Nasreddine N, Réquena M, Bélec L. R5- and X4-HIV-1 use differentially the endometrial epithelial cells HEC-1A to ensure their own spread: implication for mechanisms of sexual transmission. *Virology*. 2007;358(1):55-68. doi:10.1016/J.VIROL.2006.07.029

55. Dizzell S, Nazli A, Reid G, Kaushic C. Protective Effect of Probiotic Bacteria and Estrogen in Preventing HIV-1-Mediated Impairment of Epithelial Barrier Integrity in Female Genital Tract. *Cells*. 2019;8(10). doi:10.3390/CELLS8101120

56. Nijmeijer BM, Geijtenbeek TBH. Negative and Positive Selection Pressure During Sexual Transmission of Transmitted Founder HIV-1. *Front Immunol*. 2019;10(JULY). doi:10.3389/FIMMU.2019.01599

57. Oberle CS, Joos B, Rusert P, et al. Tracing HIV-1 transmission: envelope traits of HIV-1 transmitter and recipient pairs. *Retrovirology*. 2016;13(1). doi:10.1186/S12977-016-0299-0

58. Joseph SB, Swanstrom R, Kashuba ADM, Cohen MS. Bottlenecks in HIV-1 transmission: insights from the study of founder viruses. *Nat Rev Microbiol*. 2015;13(7):414-425. doi:10.1038/NRMICRO3471

59. Shaw GM, Hunter E. HIV Transmission. *Cold Spring Harb Perspect Med*. 2012;2(11). doi:10.1101/CSHPERSPECT.A006965

60. Nasr N, Lai J, Botting RA, et al. Inhibition of two temporal phases of HIV-1 transfer from primary Langerhans cells to T cells: the role of langerin. *J Immunol*. 2014;193(5):2554-2564. doi:10.4049/JIMMUNOL.1400630

61. Miller SM, Miles B, Guo K, et al. Follicular Regulatory T Cells Are Highly Permissive to R5-Tropic HIV-1. *J Virol*. 2017;91(17). doi:10.1128/JVI.00430-17

62. Council OD, Joseph SB. Evolution of Host Target Cell Specificity During HIV-1 Infection. *Curr HIV Res*. 2018;16(1):13-20. doi:10.2174/1570162X16666171222105721

63. Joseph SB, Arrildt KT, Swanstrom AE, et al. Quantification of entry phenotypes of macrophage-tropic HIV-1 across a wide range of CD4 densities. *J Virol*. 2014;88(4):1858-1869. doi:10.1128/JVI.02477-13

64. Lee B, Sharron M, Montaner LJ, Weissman D, Doms RW. Quantification of CD4, CCR5, and CXCR4 levels on lymphocyte subsets, dendritic cells, and differentially conditioned monocyte-derived macrophages. *Proc Natl Acad Sci U S A*. 1999;96(9):5215-5220. doi:10.1073/pnas.96.9.5215

65. Xu Y, Ollerton MT, Connick E. Follicular T-cell subsets in HIV infection: recent advances in pathogenesis research. *Curr Opin HIV AIDS*. 2019;14(2):71-76. doi:10.1097/COH.0000000000000525

66. Marras D, Bruggeman LA, Gao F, et al. Replication and compartmentalization of HIV-1 in kidney epithelium of patients with HIV-associated nephropathy. *Nat Med*. 2002;8(5):522-526. doi:10.1038/NM0502-522

67. Valdebenito S, Castellano P, Ajasin D, Eugenin EA. Astrocytes are HIV reservoirs in the brain: A cell type with poor HIV infectivity and replication but efficient cell-to-cell viral transfer. *J Neurochem*. 2021;158(2):429-443. doi:10.1111/jnc.15336

68.   Nasr N, Lai J, Botting RA, et al. Inhibition of two temporal phases of HIV-1 transfer from primary Langerhans cells to T cells: the role of langerin. *J Immunol*. 2014;193(5):2554-2564. doi:10.4049/jimmunol.1400630

69.   Chatzidimitriou D, Tsotridou E, Grigoropoulos P, Skoura L. HIV-1: towards understanding the nature and quantifying the latent reservoir. *Acta Virol*. 2020;64(1):3-9. doi:10.4149/AV_2020_101

70.   Cicala C, Nawaz F, Jelicic K, Fauci JA and AS. HIV-1 gp120: A Target for Therapeutics and Vaccine Design. *Curr Drug Targets*. 2016;17(1):122-135. doi:http://dx.doi.org.proxy1.lib.uwo.ca/10.2174/1389450116666150825120735

71.   Liu Q, Lusso P. Integrin α4β7 in HIV-1 infection: A critical review. *J Leukoc Biol*. 2020;108(2):627-632. doi:10.1002/JLB.4MR0120-208R

72.   Wilen CB, Tilton JC, Doms RW. HIV: cell binding and entry. *Cold Spring Harb Perspect Med*. 2012;2(8). doi:10.1101/cshperspect.a006866

73.   Cicala C, Arthos J, Fauci AS. HIV-1 envelope, integrins and co-receptor use in mucosal transmission of HIV. *J Transl Med*. 2011;9 Suppl 1(Suppl 1). doi:10.1186/1479-5876-9-S1-S2

74.   Li H, Chen BK. Variable infectivity and conserved engagement in cell-to-cell viral transfer by HIV-1 Env from Clade B transmitted founder clones. *Virology*. 2019;526:189-202. doi:10.1016/j.virol.2018.10.016

75.   Pedro KD, Henderson AJ, Agosto LM. Mechanisms of HIV-1 cell-to-cell transmission and the establishment of the latent reservoir. *Virus Res*. 2019;265:115-121. doi:10.1016/j.virusres.2019.03.014

76.   Khasnis MD, Halkidis K, Bhardwaj A, Root MJ. Receptor Activation of HIV-1 Env Leads to Asymmetric Exposure of the gp41 Trimer. *PLoS Pathog*. 2016;12(12):1-29. doi:10.1371/journal.ppat.1006098

77.   Jurado S, Cano-Muñoz M, Morel B, et al. Structural and Thermodynamic Analysis of HIV-1 Fusion Inhibition Using Small gp41 Mimetic Proteins. *J Mol Biol*. 2019;431(17):3091-3106. doi:10.1016/j.jmb.2019.06.022

78.   Lu W, Chen S, Yu J, et al. The Polar Region of the HIV-1 Envelope Protein Determines Viral Fusion and Infectivity by Stabilizing the gp120-gp41 Association. *J Virol*. 2019;93(7). doi:10.1128/JVI.02128-18

79.   Checkley MA, Luttge BG, Freed EO. HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. *J Mol Biol*. 2011;410(4):582-608. doi:10.1016/j.jmb.2011.04.042

80.   Mistry B, D'Orsogna MR, Webb NE, Lee B, Chou T. Quantifying the Sensitivity of HIV-1 Viral Entry to Receptor and Coreceptor Expression. *Journal of Physical Chemistry B*. 2016;120(26):6189-6199. doi:10.1021/acs.jpcb.6b02102

81.   Guzzo C, Ichikawa D, Park C, et al. Virion incorporation of integrin α4β7 facilitates HIV-1 infection and intestinal homing. *Sci Immunol*. 2017;2(11). doi:10.1126/sciimmunol.aam7341

82.   Jolly C, Kashefi K, Hollinshead M, Sattentau QJ. HIV-1 cell to cell transfer across an Env-induced, actin-dependent synapse. *J Exp Med*. 2004;199(2):283-293. doi:10.1084/JEM.20030648

83.   Duncan CJA, Williams JP, Schiffner T, et al. High-multiplicity HIV-1 infection and neutralizing antibody evasion mediated by the macrophage-T cell virological synapse. *J Virol*. 2014;88(4):2025-2034. doi:10.1128/JVI.03245-13

84.   Hu WS, Hughes SH. HIV-1 reverse transcription. *Cold Spring Harb Perspect Med*. 2012;2(10):1-22. doi:10.1101/cshperspect.a006882

85.   Li G, de Clercq E. HIV Genome-Wide Protein Associations: a Review of 30 Years of Research. *Microbiol Mol Biol Rev*. 2016;80(3):679-731. doi:10.1128/MMBR.00065-15

86.   Sleiman D, Goldschmidt V, Barraud P, Marquet R, Paillart JC, Tisné C. Initiation of HIV-1 reverse transcription and functional role of nucleocapsid-mediated tRNA/viral genome interactions. *Virus Res*. 2012;169(2):324-339. doi:10.1016/J.VIRUSRES.2012.06.006

87.   Aiken C, Rousso I. The HIV-1 capsid and reverse transcription. *Retrovirology*. 2021;18(1). doi:10.1186/S12977-021-00566-0

88.   Menéndez-Arias L, Sebastián-Martín A, Álvarez M. Viral reverse transcriptases. *Virus Res*. 2017;234:153-176. doi:10.1016/j.virusres.2016.12.019

89.   Arts EJ, Li Z, Wainberg MA. Analysis of Primer Extension and the First Template Switch during Human Immunodeficiency Virus Reverse Transcription. *J Biomed Sci*. 1995;2(4):314-321. doi:10.1007/BF02255218

90.   Christensen DE, Ganser-Pornillos BK, Johnson JS, Pornillos O, Sundquist WI. Reconstitution and visualization of HIV-1 capsid-dependent replication and integration in vitro. *Science*. 2020;370(6513). doi:10.1126/SCIENCE.ABC8420

91.   Schaller T, Ocwieja KE, Rasaiyaah J, et al. HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS Pathog*. 2011;7(12):e1002439. doi:10.1371/journal.ppat.1002439

92.   Schur FKM, Obr M, Hagen WJH, et al. An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. *Science (1979)*. 2016;353(6298):506 LP - 508. doi:10.1126/science.aaf9620

93.   Novikova M, Zhang Y, Freed EO, Peng K. Multiple Roles of HIV-1 Capsid during the Virus Replication Cycle. *Virol Sin*. 2019;34(2):119-134. doi:10.1007/s12250-019-00095-3

94.   Malikov V, da Silva ES, Jovasevic V, et al. HIV-1 capsids bind and exploit the kinesin-1 adaptor FEZ1 for inward movement to the nucleus. *Nat Commun*. 2015;6. doi:10.1038/ncomms7660

95.   Pawlica P, Berthoux L. Cytoplasmic dynein promotes HIV-1 uncoating. *Viruses*. 2014;6(11):4195-4211. doi:10.3390/v6114195

96.   Dharan A, Bachmann N, Talley S, Zwikelmaier V, Campbell EM. Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. *Nat Microbiol*. Published online 2020. doi:10.1038/s41564-020-0735-8

97.   Ambrose Z, Aiken C. HIV-1 uncoating: Connection to nuclear entry and regulation by host proteins. *Virology*. 2014;454-455(1):371-379. doi:10.1016/j.virol.2014.02.004

98.   Francis AC, Marin M, Singh PK, et al. HIV-1 replication complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains. *Nat Commun*. 2020;11(1). doi:10.1038/s41467-020-17256-8

99.   Burdick RC, Li C, Munshi M, et al. HIV-1 uncoats in the nucleus near sites of integration. *Proceedings of the National Academy of Sciences*. 2020;117(10):5486 LP - 5493. doi:10.1073/pnas.1920631117

100.   Lusic M, Siliciano RF. Nuclear landscape of HIV-1 infection and integration. *Nat Rev Microbiol*. 2017;15(2):69-82. doi:10.1038/nrmicro.2016.162

101.   Ajoge HO, Kohio HP, Paparisto E, et al. G-Quadruplex DNA and Other Non-Canonical B-Form DNA Motifs Influence Productive and Latent HIV-1

Integration and Reactivation Potential. *Viruses*. 2022;14(11):2494. doi:10.3390/v14112494

102. Barr SD, Ciuffi A, Leipzig J, Shinn P, Ecker JR, Bushman FD. HIV Integration Site Selection: Targeting in Macrophages and the Effects of Different Routes of Viral Entry. *Molecular Therapy*. 2006;14(2):218-225. doi:10.1016/j.ymthe.2006.03.012

103. Ciuffi A, Barr SD. Identification of HIV integration sites in infected host genomic DNA. *Methods*. 2011;53(1):39-46. doi:10.1016/j.ymeth.2010.04.004

104. Desimmie BA, Weydert C, Schrijvers R, et al. HIV-1 IN/Pol recruits LEDGF/p75 into viral particles. *Retrovirology*. 2015;12:16. doi:10.1186/s12977-014-0134-4

105. Wu Y. HIV-1 gene expression: lessons from provirus and non-integrated DNA. *Retrovirology*. 2004;1:13. doi:10.1186/1742-4690-1-13

106. McAllister RG, Liu J, Woods MW, Tom SK, Rupar CA, Barr SD. Lentivector integration sites in ependymal cells from a model of metachromatic leukodystrophy: non-B DNA as a new factor influencing integration. *Mol Ther Nucleic Acids*. 2014;3:e187. doi:10.1038/mtna.2014.39

107. Fernandes JD, Booth DS, Frankel AD. A structurally plastic ribonucleoprotein complex mediates post-transcriptional gene regulation in HIV-1. *Wiley Interdiscip Rev RNA*. 2016;7(4):470-486. doi:10.1002/wrna.1342

108. Karn J, Stoltzfus CM. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb Perspect Med*. 2012;2(2):a006916. doi:10.1101/cshperspect.a006916

109. Shukla A, Ramirez NGP, D'Orso I. HIV-1 proviral transcription and latency in the new era. *Viruses*. 2020;12(5):1-41. doi:10.3390/v12050555

110. Morton EL, Forst C v., Zheng Y, et al. Transcriptional Circuit Fragility Influences HIV Proviral Fate. *Cell Rep*. 2019;27(1):154-171.e9. doi:10.1016/j.celrep.2019.03.007

111. Mori L, Valente ST. Key players in HIV-1 transcriptional regulation: Targets for a functional cure. *Viruses*. 2020;12(5):1-35. doi:10.3390/v12050529

112. Karn J. The molecular biology of HIV latency: breaking and restoring the Tat-dependent transcriptional circuit. *Curr Opin HIV AIDS*. 2011;6(1):4-11. doi:10.1097/COH.0b013e328340ffbb

113. de Breyne S, Ohlmann T. Focus on Translation Initiation of the HIV-1 mRNAs. *Int J Mol Sci*. 2018;20(1). doi:10.3390/ijms20010101

114. Ule J, Blencowe BJ. Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution. *Mol Cell*. 2019;76(2):329-345. doi:10.1016/j.molcel.2019.09.017

115. Emery A, Swanstrom R. HIV-1: To Splice or Not to Splice, That Is the Question. *Viruses*. 2021;13(2). doi:10.3390/v13020181

116. Burugu S, Daher A, Meurs EF, Gatignol A. HIV-1 translation and its regulation by cellular factors PKR and PACT. *Virus Res*. 2014;193:65-77. doi:10.1016/j.virusres.2014.07.014

117. Chen J, Liu Y, Wu B, et al. Visualizing the translation and packaging of HIV-1 full-length RNA. *Proc Natl Acad Sci U S A*. 2020;117(11):6145-6155. doi:10.1073/pnas.1917590117

118. Kim SY, Byrn R, Groopman J, Baltimore D. Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. *J Virol*. 1989;63(9):3708-3713. doi:10.1128/JVI.63.9.3708-3713.1989

119. Delaleau M, Borden KLB. Multiple Export Mechanisms for mRNAs. *Cells*. 2015;4(3):452-473. doi:10.3390/cells4030452

120. Malim MH, Hauber J, Fenrick R, Cullen BR. Immunodeficiency virus rev trans-activator modulates the expression of the viral regulatory genes. *Nature*. 1988;335(6186):181-183. doi:10.1038/335181a0

121. Behrens RT, Aligeti M, Pocock GM, Higgins CA, Sherer NM. Nuclear Export Signal Masking Regulates HIV-1 Rev Trafficking and Viral RNA Nuclear Export. *J Virol*. 2017;91(3). doi:10.1128/JVI.02107-16

122. Paillart JC, Dettenhofer M, Yu XF, Ehresmann C, Ehresmann B, Marquet R. First snapshots of the HIV-1 RNA structure in infected cells and in virions. *J Biol Chem*. 2004;279(46):48397-48403. doi:10.1074/jbc.M408294200

123. Seif E, Niu M, Kleiman L. Annealing to sequences within the primer binding site loop promotes an HIV-1 RNA conformation favoring RNA dimerization and packaging. *RNA*. 2013;19(10):1384-1393. doi:10.1261/rna.038497.113

124. Dugré-Brisson S, Elvira G, Boulay K, Chatel-Chaix L, Mouland AJ, DesGroseillers L. Interaction of Staufen1 with the 5' end of mRNA facilitates translation of these RNAs. *Nucleic Acids Res*. 2005;33(15):4797-4812. doi:10.1093/nar/gki794

125. Dorin D, Bonnet MC, Bannwarth S, Gatignol A, Meurs EF, Vaquero C. The TAR RNA-binding protein, TRBP, stimulates the expression of TAR-containing RNAs in vitro and in vivo independently of its ability to inhibit the dsRNA-dependent kinase PKR. *J Biol Chem*. 2003;278(7):4440-4448. doi:10.1074/jbc.M208954200

126. Chang YN, Kenan DJ, Keene JD, Gatignol A, Jeang KT. Direct interactions between autoantigen La and human immunodeficiency virus leader RNA. *J Virol*. 1995;69(1):618-619. doi:10.1128/JVI.69.1.618-619.1995

127. Svitkin Y V, Pause A, Sonenberg N. La autoantigen alleviates translational repression by the 5' leader sequence of the human immunodeficiency virus type 1 mRNA. *J Virol*. 1994;68(11):7001-7007. doi:10.1128/JVI.68.11.7001-7007.1994

128. Toro-Ascuy D, Rojas-Araya B, García-de-Gracia F, et al. A Rev-CBP80-eIF4AI complex drives Gag synthesis from the HIV-1 unspliced mRNA. *Nucleic Acids Res*. 2018;46(21):11539-11552. doi:10.1093/nar/gky851

129. Fröhlich A, Rojas-Araya B, Pereira-Montecinos C, et al. DEAD-box RNA helicase DDX3 connects CRM1-dependent nuclear export and translation of the HIV-1 unspliced mRNA through its N-terminal domain. *Biochim Biophys Acta*. 2016;1859(5):719-730. doi:10.1016/j.bbagrm.2016.03.009

130. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature*. 1988;331(6153):280-283. doi:10.1038/331280a0

131. Korniy N, Goyal A, Hoffmann M, et al. Modulation of HIV-1 Gag/Gag-Pol frameshifting by tRNA abundance. *Nucleic Acids Res*. 2019;47(10):5210-5222. doi:10.1093/nar/gkz202

132. Woods MW, Kelly JN, Hattlmann CJ, et al. Human HERC5 restricts an early stage of HIV-1 assembly by a mechanism correlating with the ISGylation of Gag. *Retrovirology*. 2011;8:95. doi:10.1186/1742-4690-8-95

133. Sumner C, Ono A. Relationship between HIV-1 Gag Multimerization and Membrane Binding. *Viruses*. 2022;14(3). doi:10.3390/v14030622

134. Perlman M, Resh MD. Identification of an intracellular trafficking and assembly pathway for HIV-1 gag. *Traffic*. 2006;7(6):731-745. doi:10.1111/j.1398-9219.2006.00428.x

135. Kleinpeter AB, Freed EO. HIV-1 Maturation: Lessons Learned from Inhibitors. *Viruses*. 2020;12(9). doi:10.3390/v12090940

136. Guo X, Roldan A, Hu J, Wainberg MA, Liang C. Mutation of the SP1 sequence impairs both multimerization and membrane-binding activities of human immunodeficiency virus type 1 Gag. *J Virol*. 2005;79(3):1803-1812. doi:10.1128/JVI.79.3.1803-1812.2005

137. Gallina A, Mantoan G, Rindi G, Milanesi G. Influence of MA internal sequences, but not of the myristylated N-terminus sequence, on the budding site of HIV-1 Gag protein. *Biochem Biophys Res Commun*. 1994;204(3):1031-1038. doi:10.1006/bbrc.1994.2566

138. Selyutina A, Persaud M, Simons LM, et al. Cyclophilin A Prevents HIV-1 Restriction in Lymphocytes by Blocking Human TRIM5α Binding to the Viral Core. *Cell Rep*. 2020;30(11):3766-3777.e6. doi:10.1016/j.celrep.2020.02.100

139. Zhong Z, Ning J, Boggs EA, et al. Cytoplasmic CPSF6 Regulates HIV-1 Capsid Trafficking and Infection in a Cyclophilin A-Dependent Manner. *mBio*. 2021;12(2). doi:10.1128/mBio.03142-20

140. Kharytonchyk S, Brown JD, Stilger K, et al. Influence of gag and RRE Sequences on HIV-1 RNA Packaging Signal Structure and Function. *J Mol Biol*. 2018;430(14):2066-2079. doi:10.1016/j.jmb.2018.05.029

141. Cromer D, Grimm AJ, Schlub TE, Mak J, Davenport MP. Estimating the in-vivo HIV template switching and recombination rate. *AIDS*. 2016;30(2):185-192. doi:10.1097/QAD.0000000000000936

142. Floderer C, Masson JB, Boilley E, et al. Single molecule localisation microscopy reveals how HIV-1 Gag proteins sense membrane virus assembly sites in living host CD4 T cells. *Sci Rep*. 2018;8(1):16283. doi:10.1038/s41598-018-34536-y

143. Sengupta P, Seo AY, Pasolli HA, Song YE, Johnson MC, Lippincott-Schwartz J. A lipid-based partitioning mechanism for selective incorporation of proteins into membranes of HIV particles. *Nat Cell Biol*. 2019;21(4):452-461. doi:10.1038/s41556-019-0300-y

144. Favard C, Chojnacki J, Merida P, et al. HIV-1 Gag specifically restricts PI(4,5)P2 and cholesterol mobility in living cells creating a nanodomain platform for virus assembly. *Sci Adv*. 2019;5(10):eaaw8651. doi:10.1126/sciadv.aaw8651

145. Yandrapalli N, Lubart Q, Tanwar HS, et al. Self assembly of HIV-1 Gag protein on lipid membranes generates PI(4,5)P2/Cholesterol nanoclusters. *Sci Rep*. 2016;6:39332. doi:10.1038/srep39332

146. Inamdar K, Tsai FC, Dibsy R, et al. Full assembly of HIV-1 particles requires assistance of the membrane curvature factor IRSp53. *Elife*. 2021;10. doi:10.7554/eLife.67321

147. Mercenne G, Alam SL, Arii J, Lalonde MS, Sundquist WI. Angiomotin functions in HIV-1 assembly and budding. *Elife*. 2015;4. doi:10.7554/eLife.03778

148. Novikova M, Adams LJ, Fontana J, et al. Identification of a Structural Element in HIV-1 Gag Required for Virus Particle Assembly and Maturation. *mBio*. 2018;9(5). doi:10.1128/mBio.01567-18

149. Göttlinger HG, Dorfman T, Sodroski JG, Haseltine WA. Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. *Proc Natl Acad Sci U S A*. 1991;88(8):3195-3199. doi:10.1073/pnas.88.8.3195

150. Martin-Serrano J, Zang T, Bieniasz PD. HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. *Nat Med*. 2001;7(12):1313-1319. doi:10.1038/nm1201-1313

151. Votteler J, Sundquist WI. Virus budding and the ESCRT pathway. *Cell Host Microbe*. 2013;14(3):232-241. doi:10.1016/j.chom.2013.08.012

152. Wanaguru M, Bishop KN. HIV-1 Gag Recruits Oligomeric Vpr via Two Binding Sites in p6, but Both Mature p6 and Vpr Are Rapidly Lost upon Target Cell Entry. *J Virol*. 2021;95(17):e0055421. doi:10.1128/JVI.00554-21

153. Berman PW, Nunes WM, Haffar OK. Expression of membrane-associated and secreted variants of gp160 of human immunodeficiency virus type 1 in vitro and in continuous cell lines. *J Virol*. 1988;62(9):3135-3142. doi:10.1128/JVI.62.9.3135-3142.1988

154. Burnie J, Guzzo C. The Incorporation of Host Proteins into the External HIV-1 Envelope. *Viruses*. 2019;11(1). doi:10.3390/v11010085

155. Tedbury PR, Novikova M, Alfadhli A, et al. HIV-1 Matrix Trimerization-Impaired Mutants Are Rescued by Matrix Substitutions That Enhance Envelope Glycoprotein Incorporation. *J Virol*. 2019;94(1). doi:10.1128/JVI.01526-19

156. Qu K, Ke Z, Zila V, et al. Maturation of the matrix and viral membrane of HIV-1. *Science*. 2021;373(6555):700-704. doi:10.1126/science.abe6821

157. Checkley MA, Luttge BG, Soheilian F, Nagashima K, Freed EO. The capsid-spacer peptide 1 Gag processing intermediate is a dominant-negative inhibitor of HIV-1 maturation. *Virology*. 2010;400(1):137-144. doi:10.1016/j.virol.2010.01.028

158. Müller B, Anders M, Akiyama H, et al. HIV-1 Gag processing intermediates trans-dominantly interfere with HIV-1 infectivity. *J Biol Chem*. 2009;284(43):29692-29703. doi:10.1074/jbc.M109.027144

159. Johnson RF, Bell P, Harty RN. Effect of Ebola virus proteins GP, NP and VP35 on VP40 VLP morphology. *Virol J*. 2006;3:31. doi:10.1186/1743-422X-3-31

160. Jasenosky LD, Neumann G, Lukashevich I, Kawaoka Y. Ebola virus VP40-induced particle formation and association with the lipid bilayer. *J Virol*. 2001;75(11):5205-5214. doi:10.1128/JVI.75.11.5205-5214.2001

161. Jain S, Martynova E, Rizvanov A, Khaiboullina S, Baranwal M. Structural and Functional Aspects of Ebola Virus Proteins. *Pathogens*. 2021;10(10). doi:10.3390/pathogens10101330

162. Lee-Cruz L, Lenormand M, Cappelle J, et al. Mapping of Ebola virus spillover: Suitability and seasonal variability at the landscape scale. *PLoS Negl Trop Dis*. 2021;15(8):e0009683. doi:10.1371/journal.pntd.0009683

163. Bratcher A, Hoff NA, Doshi RH, et al. Zoonotic risk factors associated with seroprevalence of Ebola virus GP antibodies in the absence of diagnosed Ebola virus disease in the Democratic Republic of Congo. *PLoS Negl Trop Dis*. 2021;15(8):e0009566. doi:10.1371/journal.pntd.0009566

164. Gałas A. The determinants of spread of Ebola virus disease - an evidence from the past outbreak experiences. *Folia Med Cracov*. 2014;54(3):17-25.

165. Goldstein T, Belaganahalli MN, Syaluha EK, et al. Spillover of ebolaviruses into people in eastern Democratic Republic of Congo prior to the 2018 Ebola virus disease outbreak. *One Health Outlook*. 2020;2(1):21. doi:10.1186/s42522-020-00028-1

166. Osterholm MT, Moore KA, Kelley NS, et al. Transmission of Ebola viruses: what we know and what we do not know. *mBio*. 2015;6(2):e00137. doi:10.1128/mBio.00137-15

167. Geisbert TW, Hensley LE, Larsen T, et al. Pathogenesis of Ebola hemorrhagic fever in cynomolgus macaques: evidence that dendritic cells are early and sustained targets of infection. *Am J Pathol*. 2003;163(6):2347-2370. doi:10.1016/S0002-9440(10)63591-2

168. Schnittler HJ, Feldmann H. Marburg and Ebola hemorrhagic fevers: does the primary course of infection depend on the accessibility of organ-specific macrophages? *Clin Infect Dis*. 1998;27(2):404-406. doi:10.1086/517704

169. Bray M, Geisbert TW. Ebola virus: the role of macrophages and dendritic cells in the pathogenesis of Ebola hemorrhagic fever. *Int J Biochem Cell Biol*. 2005;37(8):1560-1566. doi:10.1016/j.biocel.2005.02.018

170. Di Paola N, Sanchez-Lockhart M, Zeng X, Kuhn JH, Palacios G. Viral genomics in Ebola virus research. *Nat Rev Microbiol*. 2020;18(7):365-378. doi:10.1038/s41579-020-0354-7

171. Pinski AN, Messaoudi I. To B or not to B: Mechanisms of protection conferred by RVSV-EBOV-GP and the roles of innate and adaptive immunity. *Microorganisms*. 2020;8(10):1-25. doi:10.3390/microorganisms8101473

172. Olejnik J, Forero A, Deflubé LR, et al. Ebolaviruses Associated with Differential Pathogenicity Induce Distinct Host Responses in Human Macrophages. Lyles DS, ed. *J Virol*. 2017;91(11):e00179-17. doi:10.1128/JVI.00179-17

173. Ning YJ, Deng F, Hu Z, Wang H. The roles of ebolavirus glycoproteins in viral pathogenesis. *Virol Sin*. 2017;32(1):3-15. doi:10.1007/s12250-016-3850-1

174. Wahl-Jensen V, Kurz S, Feldmann F, et al. Ebola virion attachment and entry into human macrophages profoundly effects early cellular gene expression. *PLoS Negl Trop Dis*. 2011;5(10):e1359. doi:10.1371/journal.pntd.0001359

175. Sadewasser A, Dietzel E, Michel S, et al. Anti-Niemann Pick C1 Single-Stranded Oligonucleotides with Locked Nucleic Acids Potently Reduce Ebola Virus Infection In Vitro. *Mol Ther Nucleic Acids*. 2019;16:686-697. doi:https://doi.org/10.1016/j.omtn.2019.04.018

176. Wang H, Shi Y, Song J, et al. Ebola Viral Glycoprotein Bound to Its Endosomal Receptor Niemann-Pick C1. *Cell*. 2016;164(1):258-268. doi:10.1016/j.cell.2015.12.044

177. Jae LT, Brummelkamp TR. Emerging intracellular receptors for hemorrhagic fever viruses. *Trends Microbiol*. 2015;23(7):392-400. doi:10.1016/j.tim.2015.04.006

178. Bo Y, Qiu S, Mulloy RP, Côté M. Filoviruses Use the HOPS Complex and UVRAG To Traffic to Niemann-Pick C1 Compartments during Viral Entry. Dutch RE, ed. *J Virol*. 2020;94(16):e01002-20. doi:10.1128/JVI.01002-20

179. Brindley MA, Hunt CL, Kondratowicz AS, et al. Tyrosine kinase receptor Axl enhances entry of Zaire ebolavirus without direct interactions with the viral glycoprotein. *Virology*. 2011;415(2):83-94. doi:https://doi.org/10.1016/j.virol.2011.04.002

180. Dahlmann F, Biedenkopf N, Babler A, et al. Analysis of Ebola Virus Entry Into Macrophages. *J Infect Dis*. 2015;212(suppl_2):S247-S257. doi:10.1093/infdis/jiv140

181. Jin C, Che B, Guo Z, et al. Single virus tracking of Ebola virus entry through lipid rafts in living host cells. *Biosaf Health*. 2020;2(1):25-31. doi:https://doi.org/10.1016/j.bsheal.2019.12.009

182. Kang YL, Chou YY, Rothlauf PW, et al. Inhibition of PIKfyve kinase prevents infection by Zaire ebolavirus and SARS-CoV-2. *Proc Natl Acad Sci U S A*. 2020;117(34):20803-20813. doi:10.1073/pnas.2007837117

183. Miller EH, Obernosterer G, Raaben M, et al. Ebola virus entry requires the host-programmed recognition of an intracellular receptor. *EMBO J*. 2012;31(8):1947-1960. doi:10.1038/emboj.2012.53

184. Rhein BA, Brouillette RB, Schaack GA, Chiorini JA, Maury W. Characterization of Human and Murine T-Cell Immunoglobulin Mucin Domain 4 (TIM-4) IgV Domain Residues Critical for Ebola Virus Entry. Lyles DS, ed. *J Virol*. 2016;90(13):6097 LP - 6111. doi:10.1128/JVI.00100-16

185. Schornberg KL, Shoemaker CJ, Dube D, et al. Alpha5beta1-integrin controls ebolavirus entry by regulating endosomal cathepsins. *Proc Natl Acad Sci U S A*. 2009;106(19):8003-8008. doi:10.1073/pnas.0807578106

186. Saeed MF, Kolokoltsov AA, Albrecht T, Davey RA. Cellular entry of ebola virus involves uptake by a macropinocytosis-like mechanism and subsequent trafficking through early and late endosomes. *PLoS Pathog*. 2010;6(9):e1001110. doi:10.1371/journal.ppat.1001110

187. Aleksandrowicz P, Marzi A, Biedenkopf N, et al. Ebola Virus Enters Host Cells by Macropinocytosis and Clathrin-Mediated Endocytosis. *J Infect Dis*. 2011;204(suppl_3):S957-S967. doi:10.1093/infdis/jir326

188. Nanbo A, Imai M, Watanabe S, et al. Ebolavirus is internalized into host cells via macropinocytosis in a viral glycoprotein-dependent manner. *PLoS Pathog*. 2010;6(9):e1001121. doi:10.1371/journal.ppat.1001121

189. Saeed MF, Kolokoltsov AA, Freiberg AN, Holbrook MR, Davey RA. Phosphoinositide-3 Kinase-Akt Pathway Controls Cellular Entry of Ebola Virus. *PLoS Pathog*. 2008;4(8):e1000141. https://doi.org/10.1371/journal.ppat.1000141

190. Hunt CL, Lennemann NJ, Maury W. Filovirus entry: a novelty in the viral fusion world. *Viruses*. 2012;4(2):258-275. doi:10.3390/v4020258

191. Côté M, Misasi J, Ren T, et al. Small molecule inhibitors reveal Niemann-Pick C1 is essential for Ebola virus infection. *Nature*. 2011;477(7364):344-348. doi:10.1038/nature10380

192. Herbert AS, Davidson C, Kuehne AI, et al. Niemann-pick C1 is essential for ebolavirus replication and pathogenesis in vivo. *mBio*. 2015;6(3):e00565-15. doi:10.1128/mBio.00565-15

193. Carette JE, Raaben M, Wong AC, et al. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature*. 2011;477(7364):340-343. doi:10.1038/nature10348

194. Bornholdt ZA, Ndungo E, Fusco ML, et al. Host-Primed Ebola Virus GP Exposes a Hydrophobic NPC1 Receptor-Binding Pocket, Revealing a Target for Broadly Neutralizing Antibodies. *mBio*. 2016;7(1):e02154-15. doi:10.1128/mBio.02154-15

195. Lee JE, Saphire EO. Ebolavirus glycoprotein structure and mechanism of entry. *Future Virol*. 2009;4(6):621-635. doi:10.2217/fvl.09.56

196. Qiu S, Leung A, Bo Y, et al. Ebola virus requires phosphatidylinositol (3,5) bisphosphate production for efficient viral entry. *Virology*. 2018;513:17-28. doi:https://doi.org/10.1016/j.virol.2017.09.028

197. Watt A, Moukambi F, Banadyga L, et al. A novel life cycle modeling system for Ebola virus shows a genome length-dependent role of VP24 in virus infectivity. *J Virol*. 2014;88(18):10511-10524. doi:10.1128/JVI.01272-14

198. Cantoni D, Rossman JS. Ebolaviruses: New roles for old proteins. *PLoS Negl Trop Dis*. 2018;12(5):e0006349-e0006349. doi:10.1371/journal.pntd.0006349

199. Geisbert TW, Jahrling PB. Differentiation of filoviruses by electron microscopy. *Virus Res*. 1995;39(2-3):129-150. doi:10.1016/0168-1702(95)00080-1

200. Miyake T, Farley CM, Neubauer BE, Beddow TP, Hoenen T, Engel DA. Ebola Virus Inclusion Body Formation and RNA Synthesis Are Controlled by a Novel Domain of Nucleoprotein Interacting with VP35. Heise MT, ed. *J Virol*. 2020;94(16):e02100-19. doi:10.1128/JVI.02100-19

201. Kolesnikova L, Nanbo A, Becker S, Kawaoka Y. Inside the Cell: Assembly of Filoviruses. *Curr Top Microbiol Immunol*. 2017;411:353-380. doi:10.1007/82_2017_15

202. Noda T, Watanabe S, Sagara H, Kawaoka Y. Mapping of the VP40-binding regions of the nucleoprotein of Ebola virus. *J Virol*. 2007;81(7):3554-3562. doi:10.1128/JVI.02183-06

203. Sanchez A, Trappier SG, Mahy BW, Peters CJ, Nichol ST. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc Natl Acad Sci U S A*. 1996;93(8):3602-3607. doi:10.1073/pnas.93.8.3602

204. Mehedi M, Falzarano D, Seebach J, et al. A new Ebola virus nonstructural glycoprotein expressed through RNA editing. *J Virol*. 2011;85(11):5406-5414. doi:10.1128/JVI.02190-10

205. Escudero-Pérez B, Volchkova VA, Dolnik O, Lawrence P, Volchkov VE. Shed GP of Ebola virus triggers immune activation and increased vascular permeability. *PLoS Pathog*. 2014;10(11):e1004509. doi:10.1371/journal.ppat.1004509

206. Volchkova VA, Feldmann H, Klenk HD, Volchkov VE. The nonstructural small glycoprotein sGP of Ebola virus is secreted as an antiparallel-orientated homodimer. *Virology*. 1998;250(2):408-414. doi:10.1006/viro.1998.9389

207. Volchkov VE, Volchkova VA, Muhlberger E, et al. Recovery of infectious Ebola virus from complementary DNA: RNA editing of the GP gene and viral cytotoxicity. *Science*. 2001;291(5510):1965-1969. doi:10.1126/science.1057269

208. Mohan GS, Li W, Ye L, Compans RW, Yang C. Antigenic subversion: a novel mechanism of host immune evasion by Ebola virus. *PLoS Pathog*. 2012;8(12):e1003065. doi:10.1371/journal.ppat.1003065

209. Volchkova VA, Klenk HD, Volchkov VE. Delta-peptide is the carboxy-terminal cleavage fragment of the nonstructural small glycoprotein sGP of Ebola virus. *Virology*. 1999;265(1):164-171. doi:10.1006/viro.1999.0034

210. Gallaher WR, Garry RF. Modeling of the Ebola virus delta peptide reveals a potential lytic sequence motif. *Viruses*. 2015;7(1):285-305. doi:10.3390/v7010285

211. Bortz RH 3rd, Wong AC, Grodus MG, et al. A virion-based assay for glycoprotein thermostability reveals key determinants of filovirus entry and its inhibition. *J Virol*. Published online July 2020. doi:10.1128/JVI.00336-20

212. Dolnik O, Volchkova V, Garten W, et al. Ectodomain shedding of the glycoprotein GP of Ebola virus. *EMBO J*. 2004;23(10):2175-2184. doi:10.1038/sj.emboj.7600219

213. Cook JD, Lee JE. The secret life of viral entry glycoproteins: moonlighting in immune evasion. *PLoS Pathog*. 2013;9(5):e1003258. doi:10.1371/journal.ppat.1003258

214. Iampietro M, Santos RI, Lubaki NM, Bukreyev A. Ebola Virus Shed Glycoprotein Triggers Differentiation, Infection, and Death of Monocytes Through Toll-Like Receptor 4 Activation. *J Infect Dis*. 2018;218(suppl_5):S327-S334. doi:10.1093/infdis/jiy406

215. Dolnik O, Becker S. Assembly, and transport of filovirus nucleocapsids. *PLoS Pathog*. 2022;18(7):e1010616. doi:10.1371/journal.ppat.1010616

216. Biedenkopf N, Hartlieb B, Hoenen T, Becker S. Phosphorylation of Ebola virus VP30 influences the composition of the viral nucleocapsid complex: impact on viral transcription and replication. *J Biol Chem*. 2013;288(16):11165-11174. doi:10.1074/jbc.M113.461285

217. Biedenkopf N, Schlereth J, Grünweller A, Becker S, Hartmann RK. RNA Binding of Ebola Virus VP30 Is Essential for Activating Viral Transcription. *J Virol*. 2016;90(16):7481-7496. doi:10.1128/JVI.00271-16

218. Biedenkopf N, Lier C, Becker S. Dynamic Phosphorylation of VP30 Is Essential for Ebola Virus Life Cycle. *J Virol*. 2016;90(10):4914-4925. doi:10.1128/JVI.03257-15

219. Basler CF, Mikulasova A, Martinez-Sobrido L, et al. The Ebola Virus VP35 Protein Inhibits Activation of Interferon Regulatory Factor 3. *J Virol*. 2003;77(14):7945-7956. doi:10.1128/jvi.77.14.7945-7956.2003

220. Prins KC, Binning JM, Shabman RS, Leung DW, Amarasinghe GK, Basler CF. Basic residues within the ebolavirus VP35 protein are required for its viral polymerase cofactor function. *J Virol*. 2010;84(20):10581-10591.

221. Ilinykh PA, Lubaki NM, Widen SG, et al. Different Temporal Effects of Ebola Virus VP35 and VP24 Proteins on Global Gene Expression in Human Dendritic Cells. *J Virol*. 2015;89(15):7567-7583. doi:10.1128/JVI.00924-15

222. Leung DW, Borek D, Luthra P, et al. An Intrinsically Disordered Peptide from Ebola Virus VP35 Controls Viral RNA Synthesis by Modulating Nucleoprotein-RNA Interactions. *Cell Rep*. 2015;11(3):376-389. doi:https://doi.org/10.1016/j.celrep.2015.03.034

223. Ivanov A, Ramanathan P, Parry C, et al. Global phosphoproteomic analysis of Ebola virions reveals a novel role for VP35 phosphorylation-dependent regulation of genome transcription. *Cell Mol Life Sci*. 2020;77(13):2579-2603. doi:10.1007/s00018-019-03303-1

224. Banadyga L, Hoenen T, Ambroggio X, Dunham E, Groseth A, Ebihara H. Ebola virus VP24 interacts with NP to facilitate nucleocapsid assembly and genome packaging. *Sci Rep*. 2017;7(1):1-14. doi:10.1038/s41598-017-08167-8

225. Takamatsu Y, Kolesnikova L, Becker S. Ebola virus proteins NP, VP35, and VP24 are essential and sufficient to mediate nucleocapsid transport. *Proc Natl Acad Sci U S A*. 2018;115(5):1075-1080. doi:10.1073/pnas.1712263115

226. Kirchdoerfer RN, Saphire EO, Ward AB. Cryo-EM structure of the Ebola virus nucleoprotein-RNA complex. *Acta Crystallogr F Struct Biol Commun*. 2019;75(Pt 5):340-347. doi:10.1107/S2053230X19004424

227. Grikscheit K, Dolnik O, Takamatsu Y, Pereira AR, Becker S. Ebola Virus Nucleocapsid-Like Structures Utilize Arp2/3 Signaling for Intracellular Long-Distance Transport. *Cells*. 2020;9(7):1728. doi:10.3390/cells9071728

228. Wan W, Kolesnikova L, Clarke M, et al. Structure and assembly of the Ebola virus nucleocapsid. *Nature*. 2017;551(7680):394-397. doi:10.1038/nature24490

229. Yamayoshi S, Noda T, Ebihara H, et al. Ebola virus matrix protein VP40 uses the COPII transport system for its intracellular transport. *Cell Host Microbe*. 2008;3(3):168-177. doi:10.1016/J.CHOM.2008.02.001

230. Fan J, Liu X, Mao F, Yue X, Lee I, Xu Y. Proximity proteomics identifies novel function of Rab14 in trafficking of Ebola virus matrix protein VP40. *Biochem Biophys Res Commun*. 2020;527(2):387-392. doi:10.1016/J.BBRC.2020.04.041

231. Adu-Gyamfi E, Digman MA, Gratton E, Stahelin R v. Single-particle tracking demonstrates that actin coordinates the movement of the Ebola virus matrix protein. *Biophys J*. 2012;103(9). doi:10.1016/J.BPJ.2012.09.026

232. Jasenosky LD, Neumann G, Lukashevich I, Kawaoka Y. Ebola virus VP40-induced particle formation and association with the lipid bilayer. *J Virol*. 2001;75(11):5205-5214. doi:10.1128/JVI.75.11.5205-5214.2001

233. Adu-Gyamfi E, Soni SP, Xue Y, Digman MA, Gratton E, Stahelin R v. The Ebola virus matrix protein penetrates into the plasma membrane: a key step in viral

protein 40 (VP40) oligomerization and viral egress. *J Biol Chem*. 2013;288(8):5779-5789. doi:10.1074/JBC.M112.443960

234. Adu-Gyamfi E, Johnson KA, Fraser ME, et al. Host Cell Plasma Membrane Phosphatidylserine Regulates the Assembly and Budding of Ebola Virus. *J Virol*. 2015;89(18):9440-9453. doi:10.1128/JVI.01087-15

235. Bharat TAM, Noda T, Riches JD, et al. Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. *Proc Natl Acad Sci U S A*. 2012;109(11):4275-4280. Accessed October 2, 2014. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3306676&tool=pmcen trez&rendertype=abstract

236. Schudt G, Dolnik O, Kolesnikova L, Biedenkopf N, Herwig A, Becker S. Transport of Ebolavirus Nucleocapsids Is Dependent on Actin Polymerization: Live-Cell Imaging Analysis of Ebolavirus-Infected Cells. *J Infect Dis*. 2015;212 Suppl:S160-6. doi:10.1093/infdis/jiv083

237. Timmins J, Scianimanico S, Schoehn G, Weissenhorn W. Vesicular release of ebola virus matrix protein VP40. *Virology*. 2001;283(1):1-6. doi:10.1006/VIRO.2001.0860

238. McCarthy SE, Johnson RF, Zhang YA, Sunyer JO, Harty RN. Role for amino acids 212KLR214 of Ebola virus VP40 in assembly and budding. *J Virol*. 2007;81(20):11452-11460. doi:10.1128/JVI.00853-07

239. Licata JM, Simpson-Holley M, Wright NT, Han Z, Paragas J, Harty RN. Overlapping motifs (PTAP and PPEY) within the Ebola virus VP40 protein function independently as late budding domains: involvement of host proteins TSG101 and VPS-4. *J Virol*. 2003;77(3):1812-1819. doi:10.1128/JVI.77.3.1812-1819.2003

240. Neumann G, Ebihara H, Takada A, et al. Ebola virus VP40 late domains are not essential for viral replication in cell culture. *J Virol*. 2005;79(16):10300-10307. doi:10.1128/JVI.79.16.10300-10307.2005

241. Marín I. Animal HECT ubiquitin ligases: evolution and functional implications. *BMC Evol Biol*. 2010;10:56. doi:10.1186/1471-2148-10-56

242. Hochrainer K, Kroismayr R, Baranyi U, Binder BR, Lipp J. Highly homologous HERC proteins localize to endosomes and exhibit specific interactions with hPLIC and Nm23B. *Cell Mol Life Sci*. 2008;65(13):2105-2117. doi:10.1007/s00018-008-8148-5

243. Hochrainer K, Mayer H, Baranyi U, Binder BerndR, Lipp J, Kroismayr R. The human HERC family of ubiquitin ligases: novel members, genomic organization, expression profiling, and evolutionary aspects. *Genomics*. 2005;85(2):153-164. doi:https://doi.org/10.1016/j.ygeno.2004.10.006

244. Hochrainer K, Pejanovic N, Olaseun VA, Zhang S, Iadecola C, Anrather J. The ubiquitin ligase HERC3 attenuates NF-kappaB-dependent transcription independently of its enzymatic activity by delivering the RelA subunit for degradation. *Nucleic Acids Res*. 2015;43(20):9889-9904. doi:10.1093/nar/gkv1064

245. Tang Y, Zhong G, Zhu L, et al. Herc5 attenuates influenza A virus by catalyzing ISGylation of viral NS1 protein. *J Immunol*. 2010;184(10):5777-5790. doi:10.4049/jimmunol.0903588

246. Woods MW, Tong JG, Tom SK, et al. Interferon-induced HERC5 is evolving under positive selection and inhibits HIV-1 particle production by a novel mechanism targeting Rev/RRE-dependent RNA nuclear export. *Retrovirology*. 2014;11:27. doi:10.1186/1742-4690-11-27

247. Cruz C, Ventura F, Bartrons R, Rosa JL. HERC3 binding to and regulation by ubiquitin. *FEBS Lett*. 2001;488(1-2):74-80. doi:10.1016/s0014-5793(00)02371-1

248. Durfee LA, Lyon N, Seo K, Huibregtse JM. The ISG15 conjugation system broadly targets newly synthesized proteins: implications for the antiviral function of ISG15. *Mol Cell*. 2010;38(5):722-732. doi:10.1016/j.molcel.2010.05.002

249. Malakhova OA, Yan M, Malakhov MP, et al. Protein ISGylation modulates the JAK-STAT signaling pathway. *Genes Dev*. 2003;17(4):455-460. doi:10.1101/gad.1056303

250. Takeuchi T, Inoue S, Yokosawa H. Identification and Herc5-mediated ISGylation of novel target proteins. *Biochem Biophys Res Commun*. 2006;348(2):473-477. doi:10.1016/j.bbrc.2006.07.076

251. Zhang D, Zhang DE. Interferon-stimulated gene 15 and the protein ISGylation system. *J Interferon Cytokine Res*. 2011;31(1):119-130. doi:10.1089/jir.2010.0110

252. Huibregtse JM, Scheffner M, Beaudenon S, Howley PM. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proc Natl Acad Sci U S A*. 1995;92(7):2563-2567. doi:10.1073/pnas.92.7.2563

253. Streich FCJ, Lima CD. Structural and functional insights to ubiquitin-like protein conjugation. *Annu Rev Biophys*. 2014;43:357-379. doi:10.1146/annurev-biophys-051013-022958

254. Durfee LA, Kelley ML, Huibregtse JM. The basis for selective E1-E2 interactions in the ISG15 conjugation system. *J Biol Chem*. 2008;283(35):23895-23902. doi:10.1074/jbc.M804069200

255. Durfee LA, Huibregtse JM. Identification and Validation of ISG15 Target Proteins. *Subcell Biochem*. 2010;54:228-237. doi:10.1007/978-1-4419-6676-6_18

256. Scheffner M, Nuber U, Huibregtse JM. Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature*. 1995;373(6509):81-83. doi:10.1038/373081a0

257. Nomura N, Miyajima N, Sazuka T, et al. Prediction of the coding sequences of unidentified human genes. I. The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by analysis of randomly sampled cDNA clones from human immature myeloid cell line KG-1. *DNA Res*. 1994;1(1):27-35. doi:10.1093/dnares/1.1.27

258. Cowley M, Wood AJ, Bohm S, Schulz R, Oakey RJ. Epigenetic control of alternative mRNA processing at the imprinted Herc3/Nap1l5 locus. *Nucleic Acids Res*. 2012;40(18):8917-8926. doi:10.1093/nar/gks654

259. Cruz C, Nadal M, Ventura F, Bartrons R, Estivill X, Rosa JL. The human HERC3 gene maps to chromosome 4q21 by fluorescence in situ hybridization. *Cytogenet Cell Genet*. 1999;87(3-4):263-264. doi:10.1159/000015442

260. Davies W, Smith RJ, Kelsey G, Wilkinson LS. Expression patterns of the novel imprinted genes Nap1l5 and Peg13 and their non-imprinted host genes in the adult mouse brain. *Gene Expr Patterns*. 2004;4(6):741-747. doi:10.1016/j.modgep.2004.03.008

261. Li H, Li J, Chen L, et al. HERC3-Mediated SMAD7 Ubiquitination Degradation Promotes Autophagy-Induced EMT and Chemoresistance in Glioblastoma. *Clin Cancer Res*. 2019;25(12):3602-3616. doi:10.1158/1078-0432.CCR-18-3791

262. Hochrainer K, Racchumi G, Zhang S, Iadecola C, Anrather J. Monoubiquitination of nuclear RelA negatively regulates NF-κB activity independent of proteasomal degradation. *Cell Mol Life Sci*. 2012;69(12):2057-2073. doi:10.1007/s00018-011-0912-2

263. Chen Y, Li Y, Peng Y, et al. DeltaNp63alpha down-regulates c-Myc modulator MM1 via E3 ligase HERC3 in the regulation of cell senescence. *Cell Death Differ*. 2018;25(12):2118-2129. doi:10.1038/s41418-018-0132-5

264. Zheng X, Chen L, Jin S, et al. Ultraviolet B irradiation up-regulates MM1 and induces photoageing of the epidermis. *Photodermatol Photoimmunol Photomed*. 2021;37:395-403. doi:10.1111/phpp.12670

265. Bourneuf E, Estelle J, Blin A, et al. New susceptibility loci for cutaneous melanoma risk and progression revealed using a porcine model. *Oncotarget*. 2018;9(45):27682-27697. doi:10.18632/oncotarget.25455

266. Lu Y, Xiao L, Liu Y, et al. MIR517C inhibits autophagy and the epithelialto-mesenchymal (-like) transition phenotype in human glioblastoma through KPNA2-dependent disruption of TP53 nuclear translocation. *Autophagy*. 2015;11(12):2213-2232. doi:10.1080/15548627.2015.1108507

267. Xu Y, Ji K, Wu M, Hao B, Yao KT, Xu Y. A miRNA-HERC4 pathway promotes breast tumorigenesis by inactivating tumor suppressor LATS1. *Protein Cell*. 2019;10(8):595-605. doi:10.1007/s13238-019-0607-2

268. Zhou H, Shi R, Wei M, Zheng WL, Zhou JY, Ma WL. The expression and clinical significance of HERC4 in breast cancer. *Cancer Cell Int*. 2013;13(1):113. doi:10.1186/1475-2867-13-113

269. Zhou H, Shi R, Chen Y, et al. [Expression of E3 ligase HERC4 in breast cancer and its clinical implications]. *Nan Fang Yi Ke Da Xue Bao*. 2014;34(8):1110-1114.

270. Wang L, Wang Y, Su B, et al. Transcriptome-wide analysis, and modelling of prognostic alternative splicing signatures in invasive breast cancer: a prospective clinical study. *Sci Rep*. 2020;10(1). doi:10.1038/S41598-020-73700-1

271. Zheng Y, Li J, Pan C, et al. HERC4 Is Overexpressed in Hepatocellular Carcinoma and Contributes to the Proliferation and Migration of Hepatocellular Carcinoma Cells. *DNA Cell Biol*. 2017;36(6):490-500. doi:10.1089/dna.2016.3626

272. Zeng WL, Chen YW, Zhou H, Zhou JY, Wei M, Shi R. Expression of HERC4 in lung cancer and its correlation with clinicopathological parameters. *Asian Pac J Cancer Prev*. 2015;16(2):513-517. doi:10.7314/apjcp.2015.16.2.513

273. Wei M, Zhang YL, Chen L, Cai CX, Wang HD. [RNA interference of HERC4 inhibits proliferation, apoptosis, and migration of cervical cancer Hela cells]. *Nan Fang Yi Ke Da Xue Bao*. 2016;37(2):232-237.

274. Udomwan P, Pientong C, Tongchai P, et al. Proteomics analysis of andrographolide-induced apoptosis via the regulation of tumor suppressor p53 proteolysis in cervical cancer-derived human papillomavirus 16-positive cell lines. *Int J Mol Sci*. 2021;22(13). doi:10.3390/IJMS22136806/S1

275. Sun X, Sun B, Cui M, Zhou Z. HERC4 exerts an anti-tumor role through destabilizing the oncoprotein Smo. *Biochem Biophys Res Commun*. 2019;513(4):1013-1018. doi:10.1016/j.bbrc.2019.04.113

276. Zhang Z, Tong J, Tang X, et al. The ubiquitin ligase HERC4 mediates c-Maf ubiquitination and delays the growth of multiple myeloma xenografts in nude mice. *Blood*. 2016;127(13):1676-1686. doi:10.1182/blood-2015-07-658203

277. Rodriguez CI, Stewart CL. Disruption of the ubiquitin ligase HERC4 causes defects in spermatozoon maturation and impaired fertility. *Dev Biol*. 2007;312(2):501-508. doi:10.1016/j.ydbio.2007.09.053

278. Huang D, Li J, He LQ. [Influence of Tripterygium wilfordii on the expression of spermiogenesis related genes Herc4, Ipo11 and Mrto4 in mice]. *Yi Chuan*. 2009;31(9):941-946. doi:10.3724/sp.j.1005.2009.00941

279. Li H, Yang BY, Liu MM, et al. Reproductive injury in male BALB/c mice infected with Neospora caninum. *Parasit Vectors*. 2021;14(1):1-12. doi:10.1186/S13071-021-04664-Y/FIGURES/8

280. Aerne BL, Gailite I, Sims D, Tapon N. Hippo Stabilises Its Adaptor Salvador by Antagonising the HECT Ubiquitin Ligase Herc4. *PLoS One*. 2015;10(6):e0131113. doi:10.1371/journal.pone.0131113

281. Jiang W, Yao X, Shan Z, Li W, Gao Y, Zhang Q. E3 ligase Herc4 regulates Hedgehog signalling through promoting Smoothened degradation. *J Mol Cell Biol*. 2019;11(9):791-803. doi:10.1093/jmcb/mjz024

282. Gémez-Mata J, Labella AM, Bandín I, Borrego JJ, García-Rosado E. Immunogene expression analysis in betanodavirus infected-Senegalese sole using an OpenArray® platform. *Gene*. 2021;774:145430. doi:10.1016/J.GENE.2021.145430

283. Valero Y, Olveira JG, López-Vázquez C, Dopazo CP, Bandín I. BEI Inactivated Vaccine Induces Innate and Adaptive Responses and Elicits Partial Protection upon Reassortant Betanodavirus Infection in Senegalese Sole. *Vaccines (Basel)*. 2021;9(5). doi:10.3390/VACCINES9050458

284. Rise ML, Hall JR, Rise M, et al. Impact of asymptomatic nodavirus carrier state and intraperitoneal viral mimic injection on brain transcript expression in Atlantic cod (Gadus morhua). *Physiol Genomics*. 2010;42(2):266-280. doi:10.1152/physiolgenomics.00168.2009

285. Eslamloo K, Xue X, Booman M, Smith NC, Rise ML. Transcriptome profiling of the antiviral immune response in Atlantic cod macrophages. *Dev Comp Immunol*. 2016;63:187-205. doi:10.1016/j.dci.2016.05.021

286. Oudshoorn D, van Boheemen S, Sánchez-Aparicio MT, Rajsbaum R, García-Sastre A, Versteeg GA. HERC6 is the main E3 ligase for global ISG15 conjugation in mouse cells. *PLoS One*. 2012;7(1). doi:10.1371/JOURNAL.PONE.0029870

287. Li C, Wang Y, Zheng H, et al. Antiviral activity of ISG15 against classical swine fever virus replication in porcine alveolar macrophages via inhibition of autophagy by ISGylating BECN1. *Vet Res*. 2020;51(1):22. doi:10.1186/s13567-020-00753-5

288. Thery F, Eggermont D, Impens F. Proteomics Mapping of the ISGylation Landscape in Innate Immunity. *Front Immunol*. 2021;12:3089. doi:10.3389/FIMMU.2021.720765/BIBTEX

289. Mitsui K, Nakanishi M, Ohtsuka S, et al. A novel human gene encoding HECT domain and RCC1-like repeats interacts with cyclins and is potentially regulated by the tumor suppressor proteins. *Biochem Biophys Res Commun*. 1999;266(1):115-122. doi:10.1006/bbrc.1999.1777

290. Kroismayr R, Baranyi U, Stehlik C, Dorfleutner A, Binder BR, Lipp J. HERC5, a HECT E3 ubiquitin ligase tightly regulated in LPS activated endothelial cells. *J Cell Sci*. 2004;117(Pt 20):4749-4756. doi:10.1242/jcs.01338

291. Zhao C, Beaudenon SL, Kelley ML, et al. The UbcH8 ubiquitin E2 enzyme is also the E2 enzyme for ISG15, an IFN-alpha/beta-induced ubiquitin-like protein. *Proc Natl Acad Sci U S A*. 2004;101(20):7578-7582. doi:10.1073/pnas.0402528101

292. Zhao C, Denison C, Huibregtse JM, Gygi S, Krug RM. Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. *Proc Natl Acad Sci U S A*. 2005;102(29):10200-10205. doi:10.1073/pnas.0504754102

293. Dastur A, Beaudenon S, Kelley M, Krug RM, Huibregtse JM. Herc5, an interferon-induced HECT E3 enzyme, is required for conjugation of ISG15 in human cells. *J Biol Chem*. 2006;281(7):4334-4338. doi:10.1074/jbc.M512830200

294. Garcia-Gonzalo FR, Bartrons R, Ventura F, Rosa JL. Requirement of phosphatidylinositol-4,5-bisphosphate for HERC1-mediated guanine nucleotide

release from ARF proteins. *FEBS Lett*. 2005;579(2):343-348. doi:10.1016/J.FEBSLET.2004.11.095

295. Sánchez-Tena S, Cubillos-Rojas M, Schneider T, Rosa JL. Functional and pathological relevance of HERC family proteins: A decade later. *Cellular and Molecular Life Sciences*. 2016;73(10):1955-1968. doi:10.1007/s00018-016-2139-8

296. Nath P, Chauhan NR, Jena KK, et al. Inhibition of IRGM establishes a robust antiviral immune state to restrict pathogenic viruses. *EMBO Rep*. 2021;22(11). doi:10.15252/embr.202152948

297. Anyona SB, Raballah E, Cheng Q, et al. Differential Gene Expression in Host Ubiquitination Processes in Childhood Malarial Anemia. *Front Genet*. 2021;12. doi:10.3389/FGENE.2021.764759

298. Li YP, Liu CR, Deng HL, et al. DNA methylation and single-nucleotide polymorphisms in DDX58 are associated with hand, foot and mouth disease caused by enterovirus 71. *PLoS Negl Trop Dis*. 2022;16(1). doi:10.1371/JOURNAL.PNTD.0010090

299. Hustinx SR, Cao D, Maitra A, et al. Differentially expressed genes in pancreatic ductal adenocarcinomas identified through serial analysis of gene expression. *Cancer Biol Ther*. 2004;3(12):1254-1261. doi:10.4161/cbt.3.12.1238

300. Zhu L, Wu J, Liu H. Downregulation of HERC5 E3 ligase attenuates the ubiquitination of CtBP1 to inhibit apoptosis in colorectal cancer cells. *Carcinogenesis*. 2021;42(8):1119-1130. doi:10.1093/carcin/bgab053

301. Yoo NJ, Park SW, Lee SH. Frameshift mutations of ubiquitination-related genes HERC2, HERC3, TRIP12, UBE2Q1 and UBE4B in gastric and colorectal carcinomas with microsatellite instability. *Pathology*. 2011;43(7):753-755. doi:10.1097/PAT.0b013e32834c7e78

302. Yoo L, Yoon AR, Yun CO, Chung KC. Covalent ISG15 conjugation to CHIP promotes its ubiquitin E3 ligase activity and inhibits lung cancer cell growth in response to type I interferon. *Cell Death Dis*. 2018;9(2):97. doi:10.1038/s41419-017-0138-9

303. Wrage M, Hagmann W, Kemming D, et al. Identification of HERC5 and its potential role in NSCLC progression. *Int J Cancer*. 2015;136(10):2264-2272. doi:10.1002/ijc.29298

304. Wang Y, Ding Q, Xu T, Li C yao, Zhou D dan, Zhang L. HZ-6d targeted HERC5 to regulate p53 ISGylation in human hepatocellular carcinoma. *Toxicol Appl Pharmacol*. 2017;334:180-191. doi:10.1016/J.TAAP.2017.09.011

305. Tang J, Yang Q, Cui Q, et al. Weighted gene correlation network analysis identifies RSAD2, HERC5, and CCL8 as prognostic candidates for breast cancer. *J Cell Physiol*. 2020;235(1):394-407. doi:10.1002/jcp.28980

306. Sala-Gaston J, Martinez-Martinez A, Pedrazza L, et al. HERC Ubiquitin Ligases in Cancer. *Cancers (Basel)*. 2020;12(6). doi:10.3390/cancers12061653

307. Nowacka M, Ginter-Matuszewska B, Świerczewska M, et al. The significance of HERC5, IFIH1, SAMD4, SEMA3A and MCTP1 genes expression in resistance to cytotoxic drugs in ovarian cancer cell lines . *Medical Journal of Cell Biology*. 2021;9(3):138-147. doi:10.2478/ACB-2021-0019

308. Tang GY, Yu P, Zhang C, Deng HY, Lu MX, Le JH. The Neuropeptide-Related HERC5/TAC1 Interactions May Be Associated with the Dysregulation of lncRNA GAS5 Expression in Gestational Diabetes Mellitus Exosomes. *Dis Markers*. 2022;2022. doi:10.1155/2022/8075285

309. Paparisto E, Hunt NR, Labach DS, et al. Interferon-induced herc5 inhibits ebola virus particle production and is antagonized by ebola glycoprotein. *Cells*. 2021;10(9). doi:10.3390/cells10092399

310. Coit P, Renauer P, Jeffries MA, et al. Renal involvement in lupus is characterized by unique DNA methylation changes in naïve CD4+ T cells. *J Autoimmun*. 2015;61:29-35. doi:10.1016/j.jaut.2015.05.003

311. Dong Z, Dai H, Liu W, et al. Exploring the Differences in Molecular Mechanisms and Key Biomarkers Between Membranous Nephropathy and Lupus Nephritis Using Integrated Bioinformatics Analysis. *Front Genet*. 2022;12:2664. doi:10.3389/FGENE.2021.770902/BIBTEX

312. Shen L, Lan L, Zhu T, et al. Identification and Validation of IFI44 as Key Biomarker in Lupus Nephritis. *Front Med (Lausanne)*. 2021;8:762848. doi:10.3389/fmed.2021.762848

313. Uhlén M, Fagerberg L, Hallström BM, et al. Tissue-based map of the human proteome. *Science (1979)*. 2015;347(6220). doi:10.1126/SCIENCE.1260419

314. The Human Protein Atlas. Accessed December 22, 2022. https://www.proteinatlas.org/

315. Mao X, Sethi G, Zhang Z, Wang Q. The Emerging Roles of the HERC Ubiquitin Ligases in Cancer. *Curr Pharm Des*. 2018;24(15):1676-1681. doi:10.2174/1381612824666180528081024

316. Jacquet S, Pontier D, Etienne L. Rapid Evolution of HERC6 and Duplication of a Chimeric HERC5/6 Gene in Rodents and Bats Suggest an Overlooked Role of HERCs in Mammalian Immunity. *Front Immunol*. 2020;11. doi:10.3389/fimmu.2020.605270

317. Arimoto K ichiro, Hishiki T, Kiyonari H, et al. Murine Herc6 Plays a Critical Role in Protein ISGylation In Vivo and Has an ISGylation-Independent Function in Seminal Vesicles. *J Interferon Cytokine Res*. 2015;35(5):351-358. doi:10.1089/jir.2014.0113

318. Jenkins TG, James ER, Alonso DF, et al. Cigarette smoking significantly alters sperm DNA methylation patterns. *Andrology*. 2017;5(6):1089-1099. doi:10.1111/ANDR.12416

319. Versteeg GA, Hale BG, van Boheemen S, Wolff T, Lenschow DJ, García-Sastre A. Species-specific antagonism of host ISGylation by the influenza B virus NS1 protein. *J Virol*. 2010;84(10):5423-5430. doi:10.1128/JVI.02395-09

320. Zhao C, Hsiang TY, Kuo RL, Krug RM. ISG15 conjugation system targets the viral NS1 protein in influenza A virus-infected cells. *Proc Natl Acad Sci U S A*. 2010;107(5):2253-2258. doi:10.1073/pnas.0909144107

321. PENG QS, LI GP, SUN WC, YANG JB, QUAN GH, LIU N. Analysis of ISG15-Modified Proteins from A549 Cells in Response to Influenza Virus Infection by Liquid Chromatography-Tandem Mass Spectrometry. *Chinese Journal of Analytical Chemistry*. 2016;44(6):850-856. doi:10.1016/S1872-2040(16)60936-2

322. Lai C, Struckhoff JJ, Schneider J, et al. Mice Lacking the ISG15 E1 Enzyme UbE1L Demonstrate Increased Susceptibility to both Mouse-Adapted and Non-Mouse-Adapted Influenza B Virus Infection. *J Virol*. 2009;83(2):1147. doi:10.1128/JVI.00105-08

323. Andres-Terre M, McGuire HM, Pouliot Y, et al. Integrated, Multi-cohort Analysis Identifies Conserved Transcriptional Signatures across Multiple Respiratory Viruses. *Immunity*. 2015;43(6):1199. doi:10.1016/J.IMMUNI.2015.11.003

324. Li HK, Kaforou M, Rodriguez-Manzano J, et al. Discovery and validation of a three-gene signature to distinguish COVID-19 and other viral infections in emergency infectious disease presentations: a case-control and observational cohort study. *Lancet Microbe*. 2021;2(11):e594. doi:10.1016/S2666-5247(21)00145-2

325. Lu MY, Huang CI, Hsieh MY, et al. Dynamics of PBMC gene expression in hepatitis C virus genotype 1-infected patients during combined peginterferon/ribavirin therapy. *Oncotarget*. 2016;7(38):61325-61335. doi:10.18632/oncotarget.11348

326. Chen F, Wang D, Li X, Wang H. Molecular Mechanisms Underlying Intestinal Ischemia/Reperfusion Injury: Bioinformatics Analysis and In Vivo Validation. *Medical Science Monitor*. 2020;26. doi:10.12659/MSM.927476

327. Al-Mamun HA, Kwan P, Clark SA, Ferdosi MH, Tellam R, Gondro C. Genome-wide association study of body weight in Australian Merino sheep reveals an orthologous region on OAR6 to human and bovine genomic regions affecting height and weight. *Genet Sel Evol*. 2015;47(1). doi:10.1186/S12711-015-0142-4

328. Cheng J, Cao X, Hanif Q, et al. Integrating Genome-Wide CNVs Into QTLs and High Confidence GWAScore Regions Identified Positional Candidates for Sheep Economic Traits. *Front Genet*. 2020;11. doi:10.3389/FGENE.2020.00569

329. al Kalaldeh M, Gibson J, Lee SH, Gondro C, van der Werf JHJ. Detection of genomic regions underlying resistance to gastrointestinal parasites in Australian sheep. *Genet Sel Evol*. 2019;51(1). doi:10.1186/S12711-019-0479-1

330. Paradis F, Yue S, Grant JR, Stothard P, Basarab JA, Fitzsimmons C. Transcriptomic analysis by RNA sequencing reveals that hepatic interferon-induced genes may be associated with feed efficiency in beef heifers. *J Anim Sci*. 2015;93(7):3331-3341. doi:10.2527/JAS.2015-8975

331. Reich N, Evans B, Levy D, Fahey D, Knight E, Darnell JE. Interferon-induced transcription of a gene encoding a 15-kDa protein depends on an upstream enhancer element. *Proc Natl Acad Sci U S A*. 1987;84(18):6394-6398. doi:10.1073/PNAS.84.18.6394

332. Kim W, Bennett EJ, Huttlin EL, et al. Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol Cell*. 2011;44(2):325-340. doi:10.1016/J.MOLCEL.2011.08.025

333. Chiu YH, Sun Q, Chen ZJ. E1-L2 Activates Both Ubiquitin and FAT10. *Mol Cell*. 2007;27(6):1014-1023. doi:10.1016/J.MOLCEL.2007.08.020

334. Kim K il, Yan M, Malakhova O, et al. Ube1L and Protein ISGylation Are Not Essential for Alpha/Beta Interferon Signaling. *Mol Cell Biol*. 2006;26(2):472-479. doi:10.1128/MCB.26.2.472-479.2006

335. Zhao C, Beaudenon SL, Kelley ML, et al. The UbcH8 ubiquitin E2 enzyme is also the E2 enzyme for ISG15, an IFN-α/β-induced ubiquitin-like protein. *Proc Natl Acad Sci U S A*. 2004;101(20):7578-7582. doi:10.1073/PNAS.0402528101

336. Kim K il, Giannakopoulos N v., Virgin HW, Zhang DE. Interferon-Inducible Ubiquitin E2, Ubc8, Is a Conjugating Enzyme for Protein ISGylation. *Mol Cell Biol*. 2004;24(21):9592-9600. doi:10.1128/MCB.24.21.9592-9600.2004

337. Wong JJY, Pung YF, Sze NSK, Chin KC. HERC5 is an IFN-induced HECT-type E3 protein ligase that mediates type I IFN-induced ISGylation of protein targets. *Proc Natl Acad Sci U S A*. 2006;103(28):10735-10740. doi:10.1073/PNAS.0600397103

338. Dastur A, Beaudenon S, Kelley M, Krug RM, Huibregtse JM. Herc5, an interferon-induced HECT E3 enzyme, is required for conjugation of ISG15 in human cells. *Journal of Biological Chemistry*. 2006;281(7):4334-4338. doi:10.1074/JBC.M512830200

339. Zou W, Zhang DE. The interferon-inducible ubiquitin-protein isopeptide ligase (E3) EFP also functions as an ISG15 E3 ligase. *Journal of Biological Chemistry*. 2006;281(7):3989-3994. doi:10.1074/JBC.M510787200

340. Okumura F, Zou W, Zhang DE. ISG15 modification of the eIF4E cognate 4EHP enhances cap structure-binding activity of 4EHP. *Genes Dev*. 2007;21(3):255-260. doi:10.1101/GAD.1521607

341. Hill S, Harrison JS, Lewis SM, Kuhlman B, Kleiger G. Mechanism of Lysine 48 Selectivity during Polyubiquitin Chain Formation by the Ube2R1/2 Ubiquitin-Conjugating Enzyme. *Mol Cell Biol*. 2016;36(11):1720-1732. doi:10.1128/MCB.00097-16

342. Xu P, Peng J. Characterization of polyubiquitin chain structure by middle-down mass spectrometry. *Anal Chem*. 2008;80(9):3438-3444. doi:10.1021/AC800016W

343. Fujimuro M, Nishiya T, Nomura Y, Yokosawa H. Involvement of polyubiquitin chains via specific chain linkages in stress response in mammalian cells. *Biol Pharm Bull*. 2005;28(12):2315-2318. doi:10.1248/BPB.28.2315

344. Sowa ME, Harper JW. From loops to chains: unraveling the mysteries of polyubiquitin chain specificity and processivity. *ACS Chem Biol*. 2006;1(1):20-24. doi:10.1021/CB0600020

345. Minakawa M, Sone T, Takeuchi T, Yokosawa H. Regulation of the nuclear factor (NF)-κB pathway by ISGylation. *Biol Pharm Bull*. 2008;31(12):2223-2227. doi:10.1248/BPB.31.2223

346. Tang Y, Zhong G, Zhu L, et al. Herc5 Attenuates Influenza A Virus by Catalyzing ISGylation of Viral NS1 Protein. *The Journal of Immunology*. 2010;184(10):5777-5790. doi:10.4049/JIMMUNOL.0903588

347. Fan JB, Arimoto KL, Motamedchaboki K, Yan M, Wolf DA, Zhang DE. Identification and characterization of a novel ISG15-ubiquitin mixed chain and its role in regulating protein homeostasis. *Sci Rep*. 2015;5. doi:10.1038/SREP12704

348. Nakaya T, Sato M, Hata N, et al. Gene induction pathways mediated by distinct IRFs during viral infection. *Biochem Biophys Res Commun*. 2001;283(5):1150-1156. doi:10.1006/bbrc.2001.4913

349. Malakhova O, Malakhov M, Hetherington C, Zhang DE. Lipopolysaccharide activates the expression of ISG15-specific protease UBP43 via interferon regulatory factor 3. *J Biol Chem*. 2002;277(17):14703-14711. doi:10.1074/jbc.M111527200

350. Au WC, Moore PA, Lowther W, Juang YT, Pitha PM. Identification of a member of the interferon regulatory factor family that binds to the interferon-stimulated response element and activates expression of interferon-induced genes. *Proc Natl Acad Sci U S A*. 1995;92(25):11657-11661. doi:10.1073/pnas.92.25.11657

351. Kroismayr R, Baranyi U, Stehlik C, Dorfleutner A, Binder BR, Lipp J. HERC5, a HECT E3 ubiquitin ligase tightly regulated in LPS activated endothelial cells. *J Cell Sci*. 2004;117(20):4749-4756. doi:10.1242/JCS.01338

352. Hollingsworth J, Lau A, Tone A, et al. BRCA1 Mutation Status and Follicular Fluid Exposure Alters NFκB Signaling and ISGylation in Human Fallopian Tube Epithelial Cells. *Neoplasia*. 2018;20(7):697-709. doi:10.1016/j.neo.2018.05.005

353. Bogunovic D, Byun M, Durfee LA, et al. Mycobacterial disease and impaired IFN-γ immunity in humans with inherited ISG15 deficiency. *Science*. 2012;337(6102):1684-1688. doi:10.1126/science.1224026

354. Speer SD, Li Z, Buta S, et al. ISG15 deficiency and increased viral resistance in humans but not mice. *Nat Commun*. 2016;7:11496. doi:10.1038/ncomms11496

355. Minakawa M, Sone T, Takeuchi T, Yokosawa H. Regulation of the nuclear factor (NF)-kappaB pathway by ISGylation. *Biol Pharm Bull*. 2008;31(12):2223-2227. doi:10.1248/bpb.31.2223

356. Kim K il, Giannakopoulos N v, Virgin HW, Zhang DE. Interferon-inducible ubiquitin E2, Ubc8, is a conjugating enzyme for protein ISGylation. *Mol Cell Biol*. 2004;24(21):9592-9600. doi:10.1128/MCB.24.21.9592-9600.2004

357. Chiang C, Liu G, Gack MU. Viral Evasion of RIG-I-Like Receptor-Mediated Immunity through Dysregulation of Ubiquitination and ISGylation. *Viruses*. 2021;13(2). doi:10.3390/v13020182

358. Zhang M, Li J, Yan H, et al. ISGylation in Innate Antiviral Immunity and Pathogen Defense Responses: A Review. *Front Cell Dev Biol*. 2021;9. doi:10.3389/FCELL.2021.788410

359. Wang Y, Ren K, Li S, Yang C, Chen L. Interferon stimulated gene 15 promotes Zika virus replication through regulating Jak/STAT and ISGylation pathways. *Virus Res*. 2020;287:198087. doi:10.1016/j.virusres.2020.198087

360. Malakhov MP, Kim K il, Malakhova OA, Jacobs BS, Borden EC, Zhang DE. High-throughput immunoblotting. Ubiquitiin-like protein ISG15 modifies key regulators of signal transduction. *J Biol Chem*. 2003;278(19):16608-16613. doi:10.1074/jbc.M208435200

361. Przanowski P, Loska S, Cysewski D, Dabrowski M, Kaminska B. ISG'ylation increases stability of numerous proteins including Stat1, which prevents premature termination of immune response in LPS-stimulated microglia. *Neurochem Int*. 2018;112:227-233. doi:10.1016/j.neuint.2017.07.013

362. Liu S, Cai X, Wu J, et al. Phosphorylation of innate immune adaptor proteins MAVS, STING, and TRIF induces IRF3 activation. *Science (1979)*. 2015;347(6227). doi:10.1126/SCIENCE.AAA2630

363. Seth RB, Sun L, Ea CK, Chen ZJ. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-κB and IRF3. *Cell*. 2005;122(5):669-682. doi:10.1016/J.CELL.2005.08.012

364. Shi HX, Yang K, Liu X, et al. Positive regulation of interferon regulatory factor 3 activation by Herc5 via ISG15 modification. *Mol Cell Biol*. 2010;30(10):2424-2436. doi:10.1128/MCB.01466-09

365. Hiscott J. Triggering the innate antiviral response through IRF-3 activation. *J Biol Chem*. 2007;282(21):15325-15329. doi:10.1074/JBC.R700002200

366. Arimoto KI, Konishi H, Shimotohno K. UbcH8 regulates ubiquitin and ISG15 conjugation to RIG-I. *Mol Immunol*. 2008;45(4):1078-1084. doi:10.1016/J.MOLIMM.2007.07.021

367. Kim MJ, Hwang SY, Imaizumi T, Yoo JY. Negative feedback regulation of RIG-I-mediated antiviral signaling by interferon-induced ISG15 conjugation. *J Virol*. 2008;82(3):1474-1483. doi:10.1128/JVI.01650-07

368. Liu G, Lee JH, Parker ZM, et al. ISG15-dependent Activation of the RNA Sensor MDA5 and its Antagonism by the SARS-CoV-2 papain-like protease. *bioRxiv*. Published online October 27, 2020. doi:10.1101/2020.10.26.356048

369. Baldanta S, Fernández-Escobar M, Acín-Perez R, et al. ISG15 governs mitochondrial function in macrophages following vaccinia virus infection. *PLoS Pathog*. 2017;13(10):e1006651. doi:10.1371/journal.ppat.1006651

370. Bhushan J, Radke JB, Perng YC, et al. ISG15 Connects Autophagy and IFN-γ-Dependent Control of Toxoplasma gondii Infection in Human Cells. *mBio*. 2020;11(5). doi:10.1128/mBio.00852-20

371. Xu D, Zhang T, Xiao J, et al. Modification of BECN1 by ISG15 plays a crucial role in autophagy regulation by type I IFN/interferon. *Autophagy*. 2015;11(4):617-628. doi:10.1080/15548627.2015.1023982

372. Zhang Y, Thery F, Wu NC, et al. The in vivo ISGylome links ISG15 to metabolic pathways and autophagy upon Listeria monocytogenes infection. *Nat Commun*. 2019;10(1):5383. doi:10.1038/s41467-019-13393-x

373. Abe T, Minami N, Bawono RG, et al. ISGylation of Hepatitis C Virus NS5A Protein Promotes Viral RNA Replication via Recruitment of Cyclophilin A. *J Virol*. 2020;94(20). doi:10.1128/JVI.00532-20

374. Okumura A, Lu G, Pitha-Rowe I, Pitha PM. Innate antiviral response targets HIV-1 release by the induction of ubiquitin-like protein ISG15. *Proc Natl Acad Sci U S A*. 2006;103(5):1440-1445. doi:10.1073/PNAS.0510518103/SUPPL_FILE/10518FIG5.PDF

375. Kim MJ, Yoo JY. Inhibition of hepatitis C virus replication by IFN-mediated ISGylation of HCV-NS5A. *J Immunol*. 2010;185(7):4311-4318. doi:10.4049/JIMMUNOL.1000098

376. Domingues P, Bamford CGG, Boutell C, McLauchlan J. Inhibition of hepatitis C virus RNA replication by ISG15 does not require its conjugation to protein substrates by the HERC5 E3 ligase. *J Gen Virol*. 2015;96(11):3236-3242. doi:10.1099/JGV.0.000283

377. Zhao C, Hsiang TY, Kuo RL, Krug RM. ISG15 conjugation system targets the viral NS1 protein in influenza A virus-infected cells. *Proc Natl Acad Sci U S A*. 2010;107(5):2253-2258. doi:10.1073/PNAS.0909144107

378. Sherpa C, Rausch JW, le Grice SFJ, Hammarskjold ML, Rekosh D. The HIV-1 Rev response element (RRE) adopts alternative conformations that promote different rates of virus replication. *Nucleic Acids Res*. 2015;43(9):4676-4686. doi:10.1093/NAR/GKV313

379. Paparisto E, Woods MW, Coleman MD, et al. Evolution-guided structural and functional analyses of the HERC family reveal an ancient marine origin and determinants of antiviral activity. *J Virol*. 2018;92(13). doi:10.1128/JVI.00528-18

380. Hartman AL, Ling L, Nichol ST, Hibberd ML. Whole-genome expression profiling reveals that inhibition of host innate immune response pathways by Ebola virus can be reversed by a single amino acid change in the VP35 protein. *J Virol*. 2008;82(11):5348-5358. doi:10.1128/JVI.00215-08

381. Caballero IS, Honko AN, Gire SK, et al. In vivo Ebola virus infection leads to a strong innate response in circulating immune cells. *BMC Genomics*. 2016;17(1):707. doi:10.1186/s12864-016-3060-0

382. Kotliar D, Lin AE, Logue J, et al. Single-Cell Profiling of Ebola Virus Disease In Vivo Reveals Viral and Host Dynamics. *Cell*. 2020;183(5):1383-1401.e19. doi:10.1016/j.cell.2020.10.002

383. Greenberg A, Huber BR, Liu DX, et al. Quantification of Viral and Host Biomarkers in the Liver of Rhesus Macaques: A Longitudinal Study of Zaire Ebolavirus Strain Kikwit (EBOV/Kik). *Am J Pathol*. 2020;190(7):1449-1460. doi:10.1016/j.ajpath.2020.03.003

384. Malakhova OA, Zhang DE. ISG15 inhibits Nedd4 ubiquitin E3 activity and enhances the innate antiviral response. *J Biol Chem*. 2008;283(14):8783-8787. doi:10.1074/jbc.C800030200

385. Okumura A, Pitha PM, Harty RN. ISG15 inhibits Ebola VP40 VLP budding in an L-domain-dependent manner by blocking Nedd4 ligase activity. *Proc Natl Acad Sci U S A*. 2008;105(10):3974-3979. doi:10.1073/pnas.0710629105

386. Guerra S, Cáceres A, Knobeloch KP, Horak I, Esteban M. Vaccinia virus E3 protein prevents the antiviral action of ISG15. *PLoS Pathog*. 2008;4(7):e1000096. doi:10.1371/journal.ppat.1000096

387. Yánguez E, García-Culebras A, Frau A, et al. ISG15 regulates peritoneal macrophages functionality against viral infection. *PLoS Pathog*. 2013;9(10):e1003632. doi:10.1371/journal.ppat.1003632

388. Eduardo-Correia B, Martínez-Romero C, García-Sastre A, Guerra S. ISG15 is counteracted by vaccinia virus E3 protein and controls the proinflammatory response against viral infection. *J Virol*. 2014;88(4):2312-2318. doi:10.1128/JVI.03293-13

389. Jacobs SR, Stopford CM, West JA, Bennett CL, Giffin L, Damania B. Kaposi's Sarcoma-Associated Herpesvirus Viral Interferon Regulatory Factor 1 Interacts with a Member of the Interferon-Stimulated Gene 15 Pathway. *J Virol*. 2015;89(22):11572-11583. doi:10.1128/JVI.01482-15

390. Dai L, Bai L, Lin Z, et al. Transcriptomic analysis of KSHV-infected primary oral fibroblasts: The role of interferon-induced genes in the latency of oncogenic virus. *Oncotarget*. 2016;7(30):47052-47060. doi:10.18632/oncotarget.9720

391. Meng W, Gao SJ. Targeting XPO1 enhances innate immune response and inhibits KSHV lytic replication during primary infection by nuclear stabilization of the p62 autophagy adaptor protein. *Cell Death Dis*. 2021;12(1):29. doi:10.1038/s41419-020-03303-1

392. Kim YJ, Kim ET, Kim YE, et al. Consecutive Inhibition of ISG15 Expression and ISGylation by Cytomegalovirus Regulators. *PLoS Pathog*. 2016;12(8):e1005850. doi:10.1371/journal.ppat.1005850

393. Bianco C, Mohr I. Restriction of Human Cytomegalovirus Replication by ISG15, a Host Effector Regulated by cGAS-STING Double-Stranded-DNA Sensing. *J Virol*. 2017;91(9). doi:10.1128/JVI.02483-16

394. Goodwin CM, Schafer X, Munger J. UL26 Attenuates IKKβ-Mediated Induction of Interferon-Stimulated Gene (ISG) Expression and Enhanced Protein ISGylation

during Human Cytomegalovirus Infection. *J Virol*. 2019;93(23). doi:10.1128/JVI.01052-19

395. Lee MK, Kim YJ, Kim YE, et al. Transmembrane Protein pUL50 of Human Cytomegalovirus Inhibits ISGylation by Downregulating UBE1L. *J Virol*. 2018;92(15). doi:10.1128/JVI.00462-18

396. Mladinich MC, Schwedes J, Mackow ER. Zika Virus Persistently Infects and Is Basolaterally Released from Primary Human Brain Microvascular Endothelial Cells. *mBio*. 2017;8(4). doi:10.1128/mBio.00952-17

397. Singh PK, Guest JM, Kanwar M, et al. Zika virus infects cells lining the blood-retinal barrier and causes chorioretinal atrophy in mouse eyes. *JCI Insight*. 2017;2(4):e92340. doi:10.1172/jci.insight.92340

398. Singh PK, Singh S, Farr D, Kumar A. Interferon-stimulated gene 15 (ISG15) restricts Zika virus replication in primary human corneal epithelial cells. *Ocul Surf*. 2019;17(3):551-559. doi:10.1016/j.jtos.2019.03.006

399. Hamel R, Dejarnac O, Wichit S, et al. Biology of Zika Virus Infection in Human Skin Cells. *J Virol*. 2015;89(17):8880-8896. doi:10.1128/JVI.00354-15

400. Munoz O, Banga R, Perreau M. Host Molecule Incorporation into HIV Virions, Potential Influences in HIV Pathogenesis. *Viruses*. 2022;14(11):2523. doi:10.3390/V14112523

401. Saphire ACS, Gallay PA, Bark SJ. Proteomic analysis of human immunodeficiency virus using liquid chromatography/tandem mass spectrometry effectively distinguishes specific incorporated host proteins. *J Proteome Res*. 2006;5(3):530-538. doi:10.1021/PR050276B/ASSET/IMAGES/LARGE/PR050276BF00004.JPEG

402. Linde ME, Colquhoun DR, Mohien CU, et al. The Conserved Set of Host Proteins Incorporated into HIV-1 Virions Suggests a Common Egress Pathway in Multiple Cell Types. *J Proteome Res*. 2013;12(5):2045. doi:10.1021/PR300918R

403. Chertova E, Chertov O, Coren L v., et al. Proteomic and Biochemical Analysis of Purified Human Immunodeficiency Virus Type 1 Produced from Infected Monocyte-Derived Macrophages. *J Virol*. 2006;80(18):9039. doi:10.1128/JVI.01013-06

404. Zhu C, Li J, Tian C, Qin M, Wang Z, Shi B, Qu G, Wu C, Nan Y. Proteomic Analysis of ISGylation in Immortalized Porcine Alveolar Macrophage Cell Lines Induced by Type I Interferon. Vaccines (Basel). 2021 Feb 17;9(2):164. doi: 10.3390/vaccines9020164.

Chapter 2

2 Evolution-Guided Structural and Functional Analyses of the HERC Family Reveal an Ancient Marine Origin and Determinants of Antiviral Activity

In humans, homologous to the E6-AP carboxyl terminus (HECT) and regulator of chromosome condensation 1 (RCC1)-like domain-containing protein 5 (HERC5) is an interferon-induced protein that inhibits replication of evolutionarily diverse viruses, including human immunodeficiency virus type 1 (HIV-1). To better understand the origin, evolution, and function of HERC5, we performed phylogenetic, structural, and functional analyses of the entire human small-HERC family, which includes HERC3, HERC4, HERC5, and HERC6. We demonstrated that the HERC family emerged >595 million years ago and has undergone gene duplication and gene loss events throughout its evolution. The structural topology of the RCC1-like domain and HECT domains from all HERC paralogs is highly conserved among evolutionarily diverse vertebrates despite low sequence homology. Functional analyses showed that the human small HERCs exhibit different degrees of antiviral activity toward HIV-1 and that HERC5 provides the strongest inhibition. Notably, coelacanth HERC5 inhibited simian immunodeficiency virus (SIV), but not HIV-1, particle production, suggesting that the antiviral activity of HERC5 emerged over 413 million years ago and exhibits species- and virus-specific restriction. In addition, we showed that both HERC5 and HERC6 are evolving under strong positive selection, particularly blade 1 of the RCC1-like domain, which we showed is a key determinant of antiviral activity. These studies provide insight into the origin, evolution, and biological importance of the human restriction factor HERC5 and the other HERC family members. This chapter was published in Journal of Virology Vol. 92, No. 13 on June 13, 2018 (see Appendix 2 for permissions).

2.1. Introduction

Vertebrates possess multiple defense mechanisms to inhibit the replication of viruses. This defense system is largely composed of specialized hematopoietic cells that react

nonspecifically to pathogens (innate immunity), an antibody-dependent and cell-mediated response (adaptive immunity), and core cellular effector proteins called restriction factors (intrinsic immunity). Restriction factors are considered to be the front line of defense against viral infection, since their activity typically does not require virus-triggered signaling or intercellular communication[1]. The importance of intrinsic immunity in vertebrates is highlighted by the evolutionarily ancient origin and broad antiviral activity of restriction factors, such as bone marrow stromal antigen 2 (BST-2)/tetherin[2]. Other restriction factors, such as apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G (APOBEC3G) and tripartite motif protein 5 alpha (TRIM5α), are unique to placental mammals and appear to play more specialized antiviral roles by targeting a more limited range of viruses, largely retroviruses[3–8].

Interferon-stimulated gene 15 (ISG15) and/or its conjugation to newly translated proteins (referred to as ISGylation) exhibits broad antiviral activity toward evolutionarily diverse viruses, including those belonging to the families Retroviridae, Orthomyxoviridae, Flaviviridae, Togaviridae, Herpesviridae, Poxviridae, Arteriviridae, and Pneumoviridae[9–31]. The main cellular E3 ligase responsible for ISGylation activity is "homologous to the E6-AP carboxyl terminus (HECT) and regulator of chromosome condensation 1 (RCC1)-like domain-containing protein 5" (HERC5), an interferon (IFN)-induced restriction factor that has evolved under strong positive selection in vertebrates[20–23,32–36]. HERC5 belongs to a subfamily of four small HERC proteins, HERC3 to -6. Although referred to as "small," the small HERC proteins are ∼116 kDa in size, each containing a single amino-terminal RCC1-like domain and a carboxyl-terminal HECT domain. The small HERCs are classified as E3 ligases due to the presence of their HECT domains and their ability to conjugate ubiquitin or ubiquitin-like molecules to proteins[32–34]. Although the biological functions of the small-HERC family have not been fully defined, their E3 ligase activities have been implicated in a variety of biological processes, such as protein degradation, cell signaling, spermatogenesis, tumor suppression, and antiviral defense (reviewed in reference[37]).

By virtue of their RCC1-like domains, HERCs also belong to the phylogenetically widespread RCC1 superfamily of proteins[38,39]. The prototypical member of this superfamily is RCC1, characterized by the presence of seven repeats of 51 to 68 amino acids that assume a 7-bladed β-propeller structure. RCC1 is localized in the nuclei of eukaryotic cells and activates the GTPase Ras-related nuclear (Ran) protein[40]. RCC1 maintains a >1,000-fold higher level of RanGTP in the nucleus than in the cytoplasm, which is critical for Crm1-dependent nuclear export of macromolecules[41,42]. We previously showed that human HERC5 inhibits the Crm1-dependent nuclear export of incompletely spliced human immunodeficiency virus type 1 (HIV-1) RNA, resulting in a severe reduction in the level of intracellular HIV-1 Gag protein and production of virus[22]. This mechanism is independent of its E3 ligase activity. Blade 1 of the 7-bladed β-propeller RCC1-like domain structure of HERC5 was critical for this inhibition and contained numerous residues predicted to be evolving under positive selection, identifying the region as a potential key antiviral interface between HERC5 and viruses[22]. Thus far, antiviral activity has been demonstrated only for human HERC5 and its functional homolog in mice, HERC6[20,21,23,32–36]. Here, we investigate the evolutionary origins and antiviral activities of the HERC family members, providing a better understanding of the role these proteins play in intrinsic immunity.

2.2.   Results

2.2.1. The small-HERC gene family has an ancient marine origin more than 595 million years ago.

With the sequencing of many evolutionarily diverse vertebrate and mammalian genomes, we can approximate the emergence and divergence of gene families throughout evolution. We analyzed the most recent genome assemblies (UCSC Genome Browser [https://genome.ucsc.edu]) and NCBI gene and protein sequence databases for the presence of small HERC gene members. The oldest small-HERC member is HERC4, which is present in one of the only surviving lineages of jawless fish, sea lampreys (originating ∼595 million years ago [mya])[43]. To better understand the evolution of the

small-HERC family, we investigated the emergence and divergence of HERC genes in evolutionarily diverse vertebrates. The elephant shark is among the oldest and most slowly evolving jawed vertebrates and has accumulated a small number of chromosomal rearrangements[44]. This allowed us to look for evidence of gene expansion at an early point in vertebrate evolution (~476 mya)[43]. A single copy of HERC4 and multiple copies of HERC3 are present in elephant sharks. Two copies of HERC3 are located immediately adjacent to HERC4, likely representing an early point in vertebrate evolution (~476 to 595 mya), just after the small-HERC family expanded with the duplication and divergence of HERC4 (Figure. 2.1). HERC3 and HERC4 are present in all the jawed vertebrates examined except the platypus, which appears to be the only vertebrate that contains two copies of HERC4. Since the only available assembly for platypus is considered low coverage, future improvements in the assembly are needed to help explain this apparently unique composition of the HERC family in these mammals. Chromosomal rearrangement likely occurred sometime after the divergence of ray-finned fish from cartilaginous fish (~430 mya), giving rise to two different chromosomal HERC loci in most vertebrates, where the HERC3-HERC5-HERC6 locus is flanked by FAM13A and PIGY-PYURF and HERC4 by MYPN and SIRT1 (Figure. 2.1B).
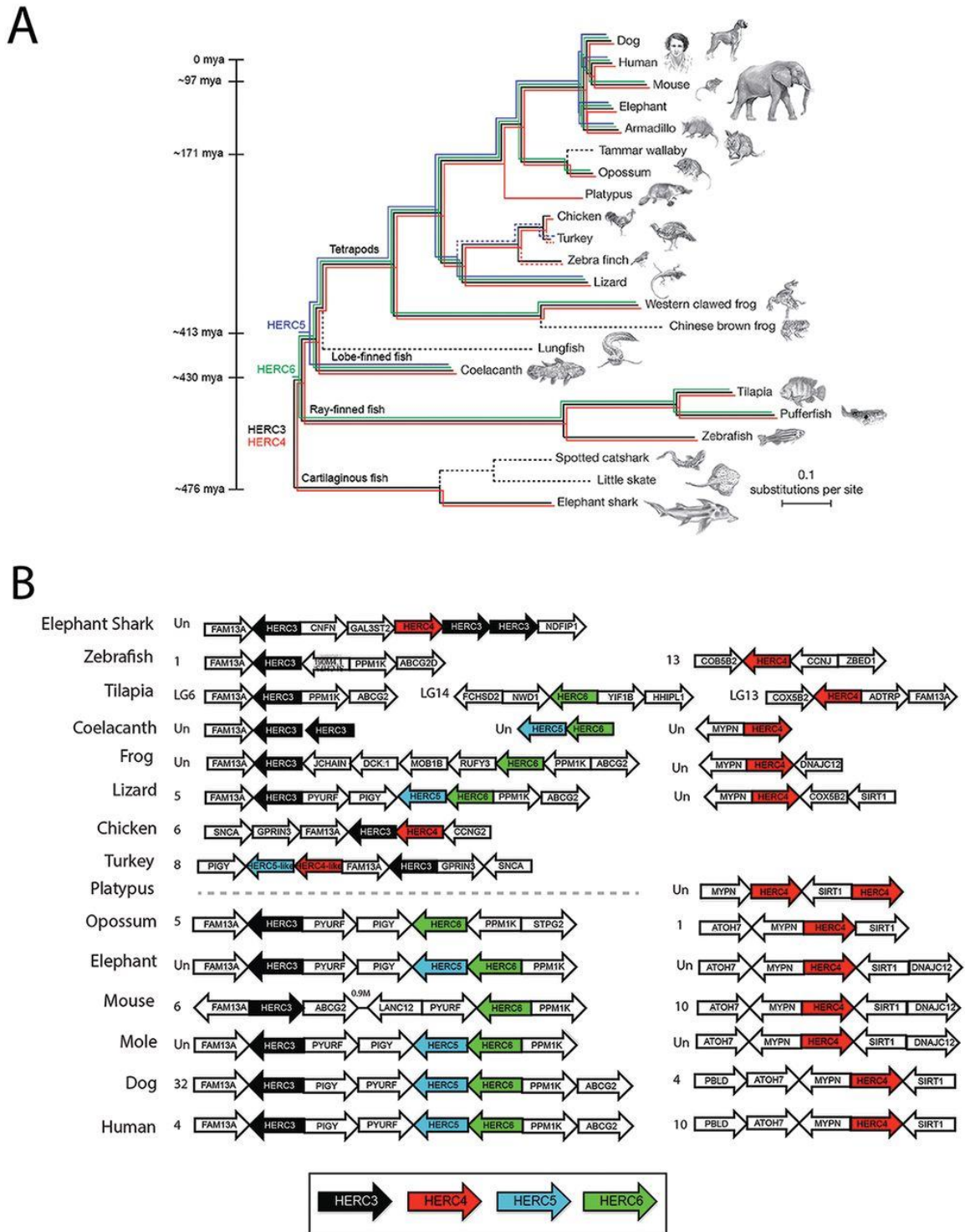
Figure 2.1 Emergence of the small-HERC family

(A) HERC3 (black), HERC4 (red), HERC5 (blue), and HERC6 (green) sequences were searched in genome assemblies using the UCSC Genome Browser (https://genome.ucsc.edu) and NCBI gene and protein sequence databases. The presence of HERC orthologs in species is indicated by colored lines. The approximate dates of divergence among the organisms are indicated by the timeline on the left from the perspective of humans, as previously described by Hedges et al.43. The Bayesian tree was obtained from a multiple-sequence alignment of 251 genes with a 1:1 ratio of orthologs in 22 vertebrates, rooted on cartilaginous fish (support was 100% for all clades but armadillo and elephant, with 45%), as described previously45. The dashed black lines indicate that no HERC orthologs were identified in the species. The dashed red and blue lines indicate the presence of partial HERC4-like and HERC5-like sequences, respectively. (Adapted from reference 45 with permission of the publisher [Springer Nature].) **(B)** Syntenic relationships of the genomic contexts of the small-HERC loci in evolutionarily diverse vertebrates. Chromosome numbers are indicated on the left of each locus. Un, unplaced scaffold.

The HERC family expanded further after the divergence of cartilaginous fish (~430 mya), with the emergence of HERC6, which is present in most jawed vertebrates with an apparent absence in platypus and some fish (e.g., zebrafish) and bird (e.g., chicken, turkey, and zebra finch) species (Figure. 2.1). The last expansion of the HERC family occurred after the divergence of ray-finned fish (~413 mya), with the likely duplication of HERC6, giving rise to HERC5 (Figure. 2.1B). HERC5 is present in the coelacanth, one of the earliest predecessors of tetrapods, and appears to have been lost in some species of frogs, birds (e.g., chicken), metatherian (marsupial) mammals (e.g., opossum), and rodents (e.g., mouse) (Figure. 2.1)[43,45]. A partial HERC5-like gene was identified in turkey, and a partial HERC4-like gene was identified in turkey and finch, possibly indicating erosion of the gene family in birds. In species that appear to have lost HERC orthologs, we cannot rule out the possibility that the orthologs are present but were

missed due to low sequence homology and/or incomplete genome annotation. Together, these findings indicate that the small-HERC family has an ancient marine origin at least 595 mya, before the emergence of jawed vertebrates, and has undergone chromosomal rearrangement, gene duplication, and potential gene loss events during vertebrate evolution.

2.2.2. Evolutionarily distant RCC1-like domains and HECT domains are well conserved.

Phylogenetic analysis of HERC sequences showed segregation of the small HERC genes into four major clusters consisting of HERC3, HERC4, HERC5, and HERC6 (Figure 2.2A). Most HERC orthologs have high sequence homology, ranging from ~70 to 100% amino acid identity, whereas most HERC paralogs have low homology, ranging from ~34 to 57%. Notably, coelacanth and lizard HERC5 sequences clustered on their own, showing more sequence similarity to HERC6 genes than to other HERC5 genes, possibly indicating that these genes are actually HERC6 or a hybrid of HERC5 and HERC6. Similar tree topologies regarding the main branching were predicted using several tree-generating algorithms (maximum likelihood, minimal evolution, unweighted pair group method using average linkages [UPMGA], and neighbor-joining methods). Consistent with the approximate emergence times of the small-HERC family members shown in Figure 2.1, HERC4 is the oldest of the HERC paralogs, followed by HERC3, HERC6, and then HERC5.
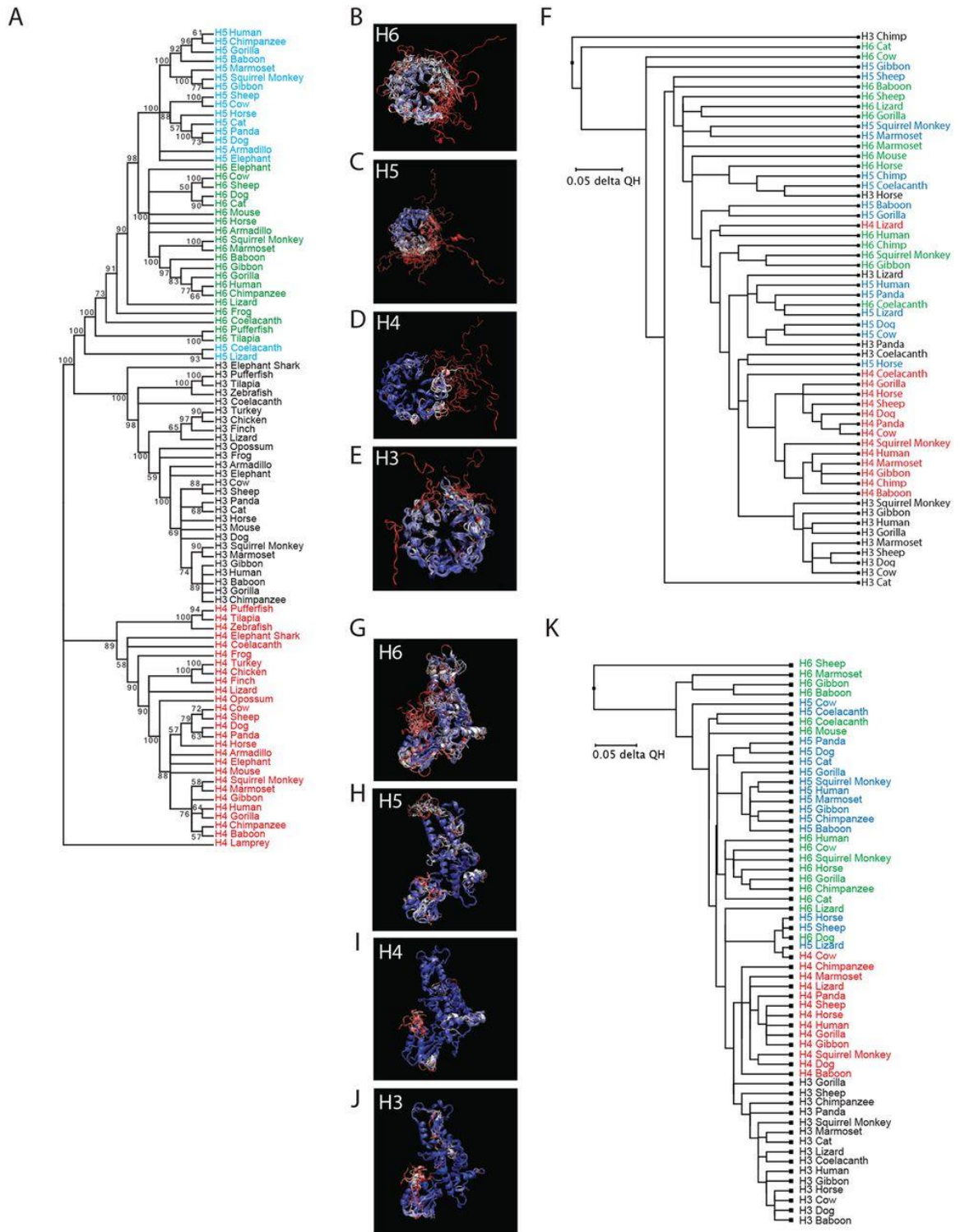
Figure 2.2 Molecular evolution of the small HERC family

A) Molecular phylogenetic analysis using the maximum likelihood method. The evolutionary history was inferred by using the maximum likelihood method based on the JTT matrix-based model46. The bootstrap consensus tree, inferred from 100 replicates, was taken to represent the evolutionary histories of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% of the bootstrap replicates were collapsed. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (100 replicates) are shown next to the branches. Initial trees for the heuristic search were obtained automatically by applying neighbor-joining and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model and then selecting the topology with a superior log likelihood value. The analysis involved 91 amino acid sequences. All positions containing gaps and missing data were removed. There was a total of 433 positions in the final data set. Evolutionary analyses were conducted in MEGA7 (83). H3, HERC3; H4, HERC4; H5, HERC5; H6 HERC6. **(B to F)** Evolutionary conservation of HERC RCC1-like domains. **(B to E)** Predicted structures of the RCC1-like domains were generated using 3D-Jigsaw (v2.0). Multiple structural alignments were generated based on the QH structural measure using the program STAMP, a plug-in in the MultiSeq interface of the Visual Molecular Dynamics (VMD) software (v1.9.2). **(F)** Three-dimensional representation of the structural data colored by structural conservation. Each amino acid is colored according to the degree of conservation within the alignment: blue, highly conserved; white, somewhat conserved; and red, very low or no conservation. The structure-based cladogram was derived from sequence and structural alignments of the predicted tertiary structures of multiple mammalian HERC RCC1-like domains. **(G to K)** Evolutionary conservation of HERC HECT domains. **(G to J)** Three-dimensional representations of the predicted HECT domain structural data colored by structural conservation. **(K)** Structure-based cladogram derived from sequence and structural alignments of the predicted tertiary structures of multiple mammalian HERC HECT domains.

The small HERCs are believed to have arisen from a gene fusion event between an RCC1-like domain and a HECT domain[47]. Although the approximate date of this event is

unknown, the presence of HERC4 in jawless fishes (e.g., lampreys) suggests that the fusion event occurred more than 595 mya. Typically, the primary amino acid sequences of RCC1-like domains have low sequence homology in the RCC1 superfamily; however, their tertiary structures are highly conserved[38,39]. To assess how conserved the predicted tertiary structures are among the different small-HERC members, we generated a phylogenetic tree based on the QH structural measure derived from alignment of the predicted tertiary structures of RCC1-like domains (Figure 2.2B to F). The models were predicted using 3D-Jigsaw v2.0 (https://bmm.crick.ac.uk/~populus/) and showed that each of the RCC1-like domains of HERC3 to -6 adopted the characteristic β-propeller structure of the superfamily, despite their low sequence homology[48–52]. Alignment of the structures was carried out using the program Structural Alignment of Multiple Proteins (STAMP), which is a tool for aligning protein sequences based on their three-dimensional structures[53]. The STAMP algorithm minimizes the Cα distance between aligned residues of each molecule by applying globally optimal rigid-body rotations and translations. STAMP analysis revealed that the RCC1-like domains of the HERC orthologs are well conserved overall, with HERC3 and HERC4 being the most conserved (Figure 2.2B to F). Several paralogs showed greater similarity to each other than to their orthologous counterparts (e.g., lizard HERC4 and human HERC6). Chimpanzee HERC3 differed substantially from the other HERCs in that it lacks blade 3 of the β-propeller. Notably, the amino-terminal blade 1 of each RCC1-like domain is the least conserved region and adopts a more extended conformation than the other blades.

Analysis of the predicted HECT domain structures showed that they all adopted the typical bilobial structure of HECT domains (Figure 2.2G to K). STAMP analysis showed that the different orthologs are well conserved overall, with HERC3 and HERC4 being the most conserved. Some HERC5 and HERC6 paralogs shared more similarity to each other than to their respective orthologs. Notably, coelacanth and mouse HERC6 proteins were more similar to HERC5 paralogs than to other HERC6 orthologs, potentially indicating conservation in structure and function of these HECT domains (Figure 2.2K). This is consistent with the finding that mouse HERC6 is the functional homolog of

human HERC5; these are the main cellular E3 ligases for ISG15 in mice and humans, respectively. Together, these data show that despite low sequence homology, evolutionarily divergent HERC genes share remarkable similarity in the predicted structures of their RCC1-like domains and HECT domains.

2.2.3. Human HERC3 to -6 differentially inhibit HIV-1 particle production.

Given the remarkable similarity in the predicted structures of the HECT and RCC1-like domains, we asked whether human HERC3, HERC4, and HERC6 inhibited HIV-1 particle production like HERC5. To test the effect on HIV-1 replication of knocking down endogenous HERC protein levels, we first screened different HERC short hairpin RNA (shRNA) constructs for the ability to knock down endogenous HERC mRNA and protein levels. As shown in Figure 2.3A, several of the shRNA constructs knocked down HERC mRNA levels by 2- to 5-fold. For each shRNA construct used, no significant differences in mRNA levels were detected for any of the other related small HERCs, demonstrating specificity (Figure 2.3B). Unfortunately, we were unable to readily measure endogenous levels of HERC proteins using several different commercially available antibodies. This could be due to poor recognition of endogenous HERC proteins by the antibodies and/or tightly controlled cytosolic levels of HERC proteins, as previously observed[34]. As such, HERC protein knockdown efficiencies of the shRNA constructs were instead determined using exogenously expressed Flag-tagged HERC constructs (Figure 2.3C). shRNAs that knocked down Flag-tagged HERC3, HERC4, HERC5, and HERC6 protein levels by 10.1-, 7.2-, 5.0-, and 3.6-fold, respectively, were used for subsequent experiments.
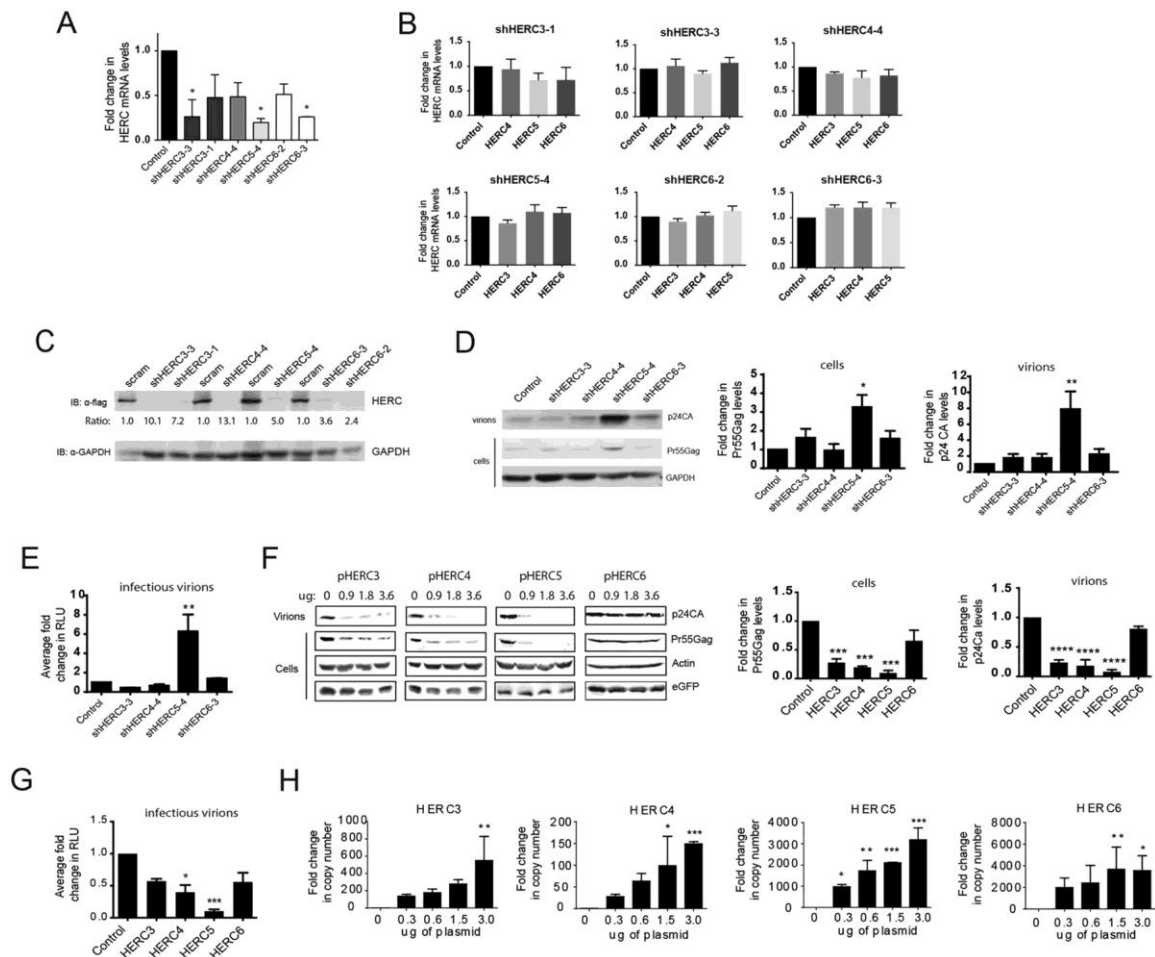
Figure 2.3 HERC3 to -6 differentially restrict HIV-1 particle production

A and B) HOS-CD4-CXCR4 cells were transfected with a plasmid carrying either scrambled shRNA (control) or shRNAs to each of the different HERCs independently. HERC mRNA levels were measured by qPCR using HERC-specific primers. The average fold changes in HERC mRNA levels (with standard errors of the mean [SEM]) from the results of 3 independent experiments are shown. One-way analysis of variance (ANOVA) and Dunnett's multiple-comparison test with the control were performed. **(C)** 293T cells were first transfected with plasmids carrying HERC shRNA and 24 h later with the respective Flag-tagged HERC constructs. HERC protein levels were measured by quantitative Western blotting (immunoblotting [IB]) using anti (α)-Flag or anti-GAPDH (loading control). The numbers below the blots represent the average fold

changes in HERC protein levels compared to the control cells after densitometric quantification of the Flag-tagged HERC bands. scram, scrambled shRNA. **(D)** HOS-CD4-CXCR4 cells were transfected with pR9 and plasmids carrying either scrambled shRNA (control) or shRNAs to each of the different HERCs independently. Seventy-two hours after transfection, cell lysates and HIV-1 virions in the supernatant were harvested. HIV-1 particle production was measured by quantitative Western blotting using anti-p24CA or anti-GAPDH. **(E)** Average (plus SEM) densitometric quantifications of Pr55Gag (cells) and p24CA (virions) bands. Virions produced from the cells were used to infect the luciferase reporter cell line TZM-bl to measure the number of infectious virions. The average fold changes in relative light units (RLU) from the results of at least 3 independent experiments are shown. **(F and G)** 293T cells were cotransfected with pR9 and peGFP (transfection control) and increasing amounts of either empty vector, pHERC3, pHERC4, pHERC5, or pHERC6. Virions in the supernatant and total cell lysates were subjected to quantitative Western blot analysis using anti-p24CA, anti-eGFP, and anti-β-actin as a loading control. The average (plus SEM) densitometric quantification of p24CA and Pr55Gag levels from Western blot images of virions or cell lysate (3.6 μg HERC lane) are shown on the right. The values were normalized to β-actin and eGFP levels. **(G)** Virions produced from cells were used to infect the luciferase reporter cell line TZM-bl to measure the number of infectious virions. **(H)** 293T cells were transfected with either empty vector or increasing amounts of plasmids encoding the different HERCs. Twenty hours post-transfection, total RNA was harvested from the cells and subjected to qPCR to measure HERC mRNA levels. Relative fold changes in HERC mRNA compared to the control cells are shown. Statistical significance was determined using one-way ANOVA and Dunnett's multiple-comparison test with the control cells. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

HOS-CD4-CXCR4 cells were co-transfected with plasmids carrying HERC shRNA and HIV-1 R9 (a full-length replication-competent NL4-3 derivative). After 72 h of

replication (∼2 or 3 rounds), quantitative Western blot analysis of cell lysates or virions produced from cells showed that cells knocked down for HERC5 expression exhibited a significant increase in Gag particle production (∼8-fold), whereas HERC3, -4, and -6 released modestly more virions into the supernatant than the control cells (∼2-fold) (Figure 2.3D). The amount of infectious HIV-1 in the supernatant was also measured by infecting the TZM-bl indicator cell line, which enabled quantitative analysis of HIV-1 using luciferase as a reporter[54–58]. Cells knocked down for endogenous HERC5 expression failed to inhibit production of infectious HIV-1, whereas cells knocked down for HERC3, HERC4, or HERC6 expression released levels of infectious virions similar to those of the control cells (Figure 2.3E).

To test the effect of increased HERC expression on single-round HIV-1 particle production, human 293T cells, which do not support multiround replication, were co-transfected with plasmids carrying HIV-1 (R9) and either empty vector, HERC3, HERC4, HERC5 or HERC6. As expected, HERC5 potently inhibited HIV-1 particle production (Figure. 2.3F). HERC3 or HERC4 also significantly inhibited particle production, but not as potently as HERC5. In contrast, HERC6 modestly inhibited HIV-1 particle production but did not achieve statistical significance. As expected, HERC5 also potently inhibited the production of infectious HIV-1, whereas HERC3, HERC4, and HERC6 modestly inhibited production of infectious HIV-1 (Figure 2.3G). Each transfected HERC construct exhibited robust mRNA expression, although HERC3 and HERC4 levels were less than those of HERC5 and HERC6 (Figure 2.3H). Taken together, these data show that upregulated expression of HERC3 to -6 inhibited HIV-1 particle production and replication to varying degrees, with HERC5 exhibiting the most potent activity and HERC6 the weakest activity. Notably, only endogenous levels of HERC5 significantly inhibited production of infectious HIV-1.

2.2.4. Human HERC3 to -6 differentially inhibit nuclear export of incompletely spliced RNA.

We previously showed that human HERC5 blocked nuclear export of Rev-dependent HIV-1 RNA[22]. To determine if HERC3, HERC4, and HERC6 also blocked nuclear export of Rev-dependent HIV-1 RNAs, we cotransfected 293T cells with plasmids carrying full-length HIV-1 R9 and either empty vector, HERC3, HERC4, HERC5, or HERC6. A plasmid encoding enhanced green fluorescent protein (eGFP) was also cotransfected to serve as a transfection control. Total RNA was harvested from the total cell extract or the cytoplasmic extract only and subjected to quantitative PCR (qPCR) with primers specific for either Gag (unspliced HIV-1 genomic RNA), Rev (fully spliced RNA), β-actin (loading control), or eGFP. Each of the small HERC proteins exhibited significant reductions in the amount of HIV-1 genomic RNA present in the cytoplasm, with HERC5 exhibiting the strongest activity (Figure 2.4A). In contrast, no significant reductions in the export of fully spliced HIV-1 transcripts were observed (Figure 2.4B).
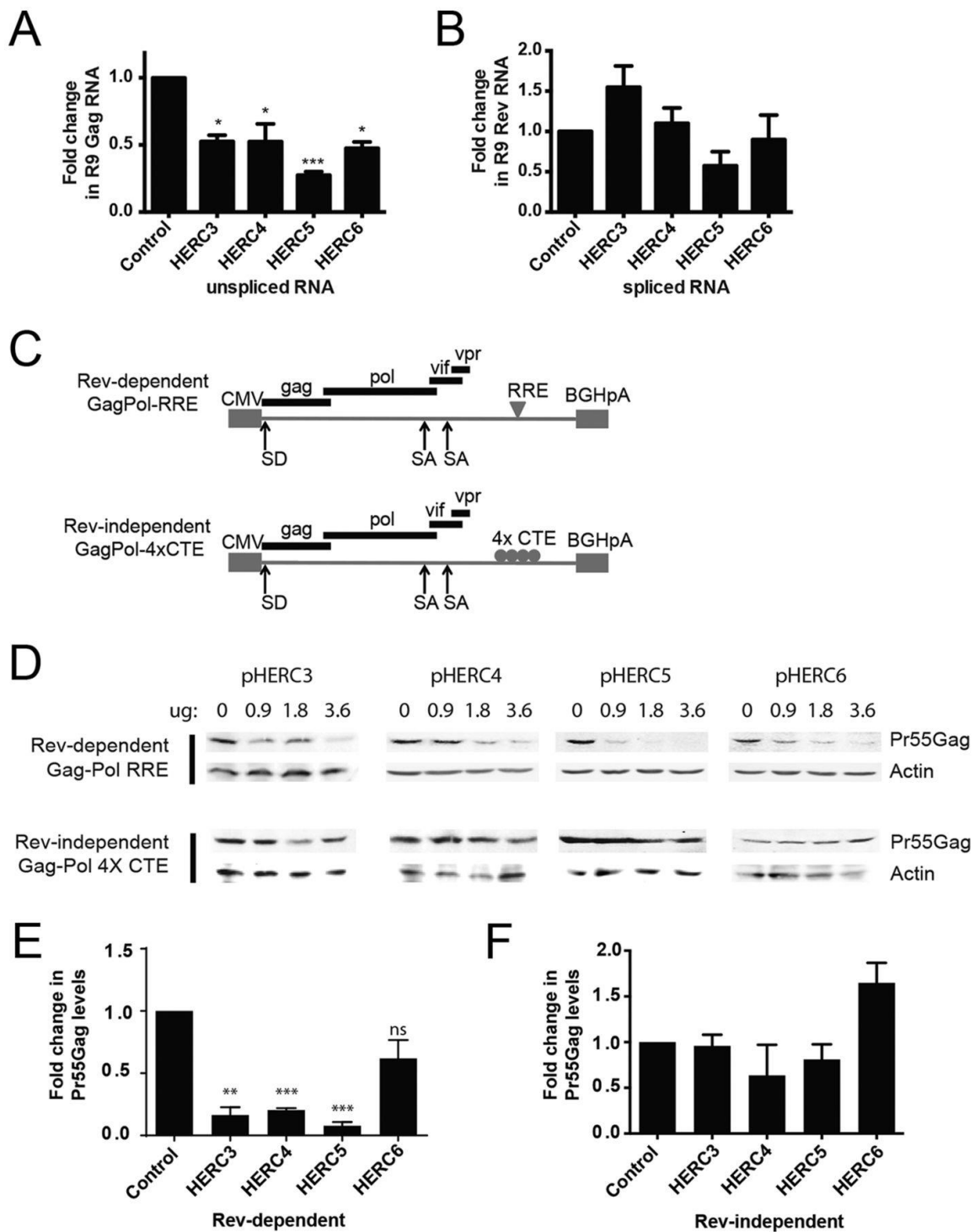
Figure 2.4 HERC3 to -5 inhibit cytoplasmic accumulation of unspliced HIV-1 RNA.

A and B) 293T cells were cotransfected with pR9 and peGFP (transfection control) and either empty vector, pHERC3, pHERC4, or pHERC5. Forty-eight hours after transfection, total RNA was extracted and reverse transcribed into cDNA from whole-cell lysates or from the cytoplasmic fraction only. Quantitative PCR was performed on each fraction using primers specific for unspliced HIV-1 genomic RNA (e.g., Gag), fully spliced RNA (e.g., Rev), β-actin (loading control), or eGFP (transfection control). The proportion of unspliced or fully spliced HIV-1 RNA in the cytoplasmic fraction compared to the total amount of HIV-1 RNA (nuclear plus cytoplasmic) was determined for control cells and cells expressing HERC after normalization to β-actin and eGFP levels. The fold changes in copy numbers relative to control cells are shown. The data shown represent the averages (plus SEM) from the results of four independent experiments. **(C)** Schematic depicting the different Gag-Pol constructs used in the experiment shown in panel D. CMV, cytomegalovirus; SA, splice acceptor; SD, splice donor. **(D)** 293T cells were cotransfected with increasing amounts of plasmids encoding HERC3, HERC4, or HERC5 and either Rev-dependent GagPol-RRE (plus pRev) or Rev-independent GagPol-4×CTE. The total DNA transfected was kept equal with the empty-vector plasmid. Gag levels within the cell lysates were analyzed by quantitative Western blotting using anti-p24CA and anti-β-actin as a loading control. **(E and F)** Densitometric quantification of Pr55Gag bands from the lanes containing the largest amount of each HERC in the Western blots from panel D was performed. Shown are the average fold changes (plus SEM) in Pr55Gag levels relative to the empty-vector control after normalization to β-actin levels. Statistical significance was determined using ANOVA with Dunnett's multiple-comparison test. *, $P < 0.05$; **, $P < 0.01$, ***, $P < 0.001$; ns, not significant.

To further support this finding, we tested the abilities of the small HERCs to inhibit Gag expression from Rev-dependent (e.g., GagPol-RRE) and Rev-independent (e.g., GagPol-4×CTE) constructs, as previously described[22,59]. HIV-1 Rev promotes nuclear export of incompletely spliced HIV-1 mRNAs by binding to a specific cis-acting element called the Rev-response element (RRE) located within an HIV-1 intron (Figure 2.4C). HIV-1

mRNA containing four copies of the Mason-Pfizer monkey virus constitutive export element (4×CTE) in place of the RRE is not dependent on Rev for nuclear export and thus serves as a Rev-independent control[60]. Successful export of incompletely spliced RNA can be assessed by Western blotting for Gag protein expression. 293T cells were cotransfected with a plasmid encoding Rev and increasing concentrations of plasmids encoding HERC, with or without pGagPol-RRE or pGagPol-4×CTE. As shown in Figure. 2.4D and E, each of the small HERCs differentially inhibited nuclear export of Rev-dependent RNA, where HERC5 exhibited the highest level of inhibition and HERC6 the weakest inhibition. In contrast, none of the HERCs significantly inhibited nuclear export of Rev-independent RNA (Figure 2.4D and F). Together, these findings indicate that the small-HERC members differentially inhibit nuclear export of Rev-dependent RNAs.

2.2.5. Antiviral activity of HERC5 evolved more than 413 million years ago.

We next asked whether the antiretroviral activity of human HERC5 has an evolutionarily ancient origin. Since the coelacanth was the oldest vertebrate in which we identified HERC5, we tested the ability of coelacanth HERC5 to inhibit HIV-1 virus production. To assess potential virus-specific effects, we also tested the antiviral activity toward another, related nonhuman retrovirus, simian immunodeficiency virus (SIV) (SIVmac239, a full-length rhesus macaque derivative lacking the 5′ long terminal repeat [LTR]), which is thought to be at least 32,000 years older than HIV-1[61]. For comparison, we included human HERC5 and human HERC6, which exhibited the strongest and weakest anti-HIV-1 activities, respectively. Human 293T cells were cotransfected with plasmids carrying SIVmac239 or HIV-1 R9 and increasing concentrations of either empty vector, coelacanth HERC5, human HERC5, or human HERC6. Forty-eight hours after transfection, virus released into the supernatant was measured by Western blotting. As expected, human HERC5 exhibited strong inhibition, whereas both coelacanth HERC5 and human HERC6 exhibited little inhibition of HIV-1 (Figure 2.5A to C). In contrast, each of the HERCs inhibited SIVmac239 virus production, with human HERC5 being the

most potent (Figure 2.5D to F). Together, these results demonstrate that the antiretroviral activity of HERC5 has an ancient marine origin at least 413 mya and that HERC5 and HERC6 exhibit species- and virus-specific antiviral activity.
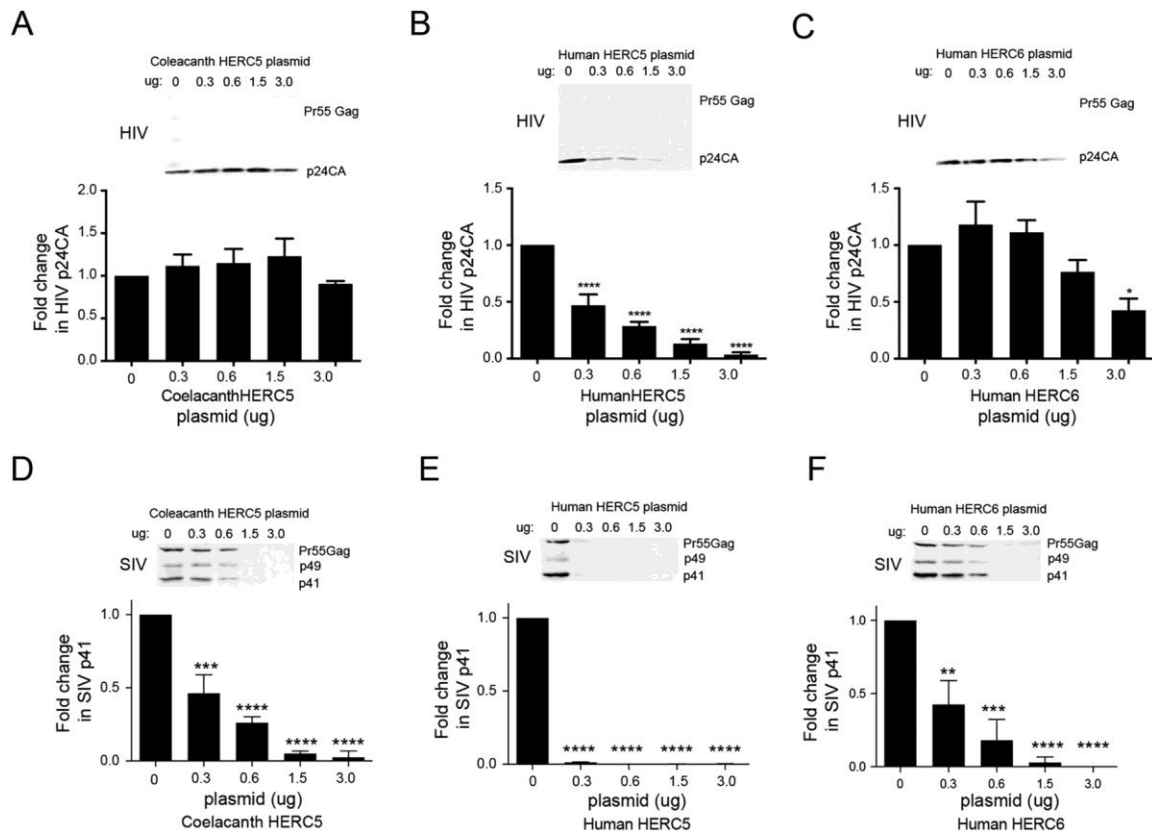
Figure 2.5 Coelacanth HERC5 restricts SIV, but not HIV-1, particle production. HERC6 is evolving under positive selection.

293T cells were co-transfected with a plasmid carrying HIV-1 (pR9) **(A to C)** or SIV (pSIVmac239) **(D to F)** and increasing amounts of plasmids encoding either coelacanth HERC5, human HERC5, or human HERC6. Forty-eight hours post-transfection, virus released into the supernatant was measured by quantitative Western blotting of Gag proteins using monoclonal anti-p24CA (183-H12-5C) or anti-SIVp17 (KK59). The average relative fold changes (plus SEM) in HIV p24CA or SIV p24CA protein levels compared to the control cells after densitometric quantification of 3 independent Western

blot images are shown. Statistical significance was determined using one-way ANOVA with Dunnett's multiple-comparison test with the control cells. **, P < 0.01; ***, P < 0.001; ****, P < 0.0001.

We previously showed that HERC5, especially blade 1 of its RCC1-like domain, has been evolving under positive selection for >100 million years[22]. We performed a similar analysis for each of the small-HERC members to determine if the other members of the small-HERC family have been evolving under positive selection. HERC evolution in mammals was evaluated under several standard models of sequence evolution using the Server for the Identification of Site-Specific Positive Selection and Purifying Selection (Selecton) program[62–65]. This comprised two nested pairs of models (M8a and M8; M7 and M8), in which the first model of each pair is nested in the second model. The M8 model, but not the M8a or M7 model, allows sites to evolve under positive selection. A nonnested-pair (M8a and MEC) model comparison was also performed. The MEC model differs from the other models in that it takes into account the differences between amino acid replacement rates[62]. The nested models were compared using the likelihood ratio test.

Analysis of 12 evolutionarily diverse HERC sequences using Selecton revealed that HERC6, but not HERC3 and HERC4, is evolving under positive selection (Figure 2.6). Allowing sites to evolve under positive selection (M8) gave a significantly better fit to the HERC6 sequence data than the corresponding model without positive selection (M8a and M7) (Figure 2.6B). The MEC model, which allows positive selection, was compared with the M8a null model, which does not allow positive selection. Comparison of the Akaike Information Criterion (AICc) scores (M8a, 25,806; MEC, 25,557) revealed that the MEC model fits the HERC6 data better than the M8a model. The results of the MEC analysis were projected onto the primary sequence of human HERC6 (Figure 2.6C). Notably, ~23% (23 of 102) of the codons cluster within the first 80 amino acids of the amino terminus of the RCC1-like domain, encompassing blade 1 of its predicted β-

propeller structure. Another ~32% (33 of 102) cluster at the carboxyl terminus of the spacer region (amino acids ~630 to 680). These results show that strong positive selection is operating on HERC6, with a large number of codons in blade 1 of the RCC1-like domain and the carboxyl terminus of the spacer region evolving under positive selection.
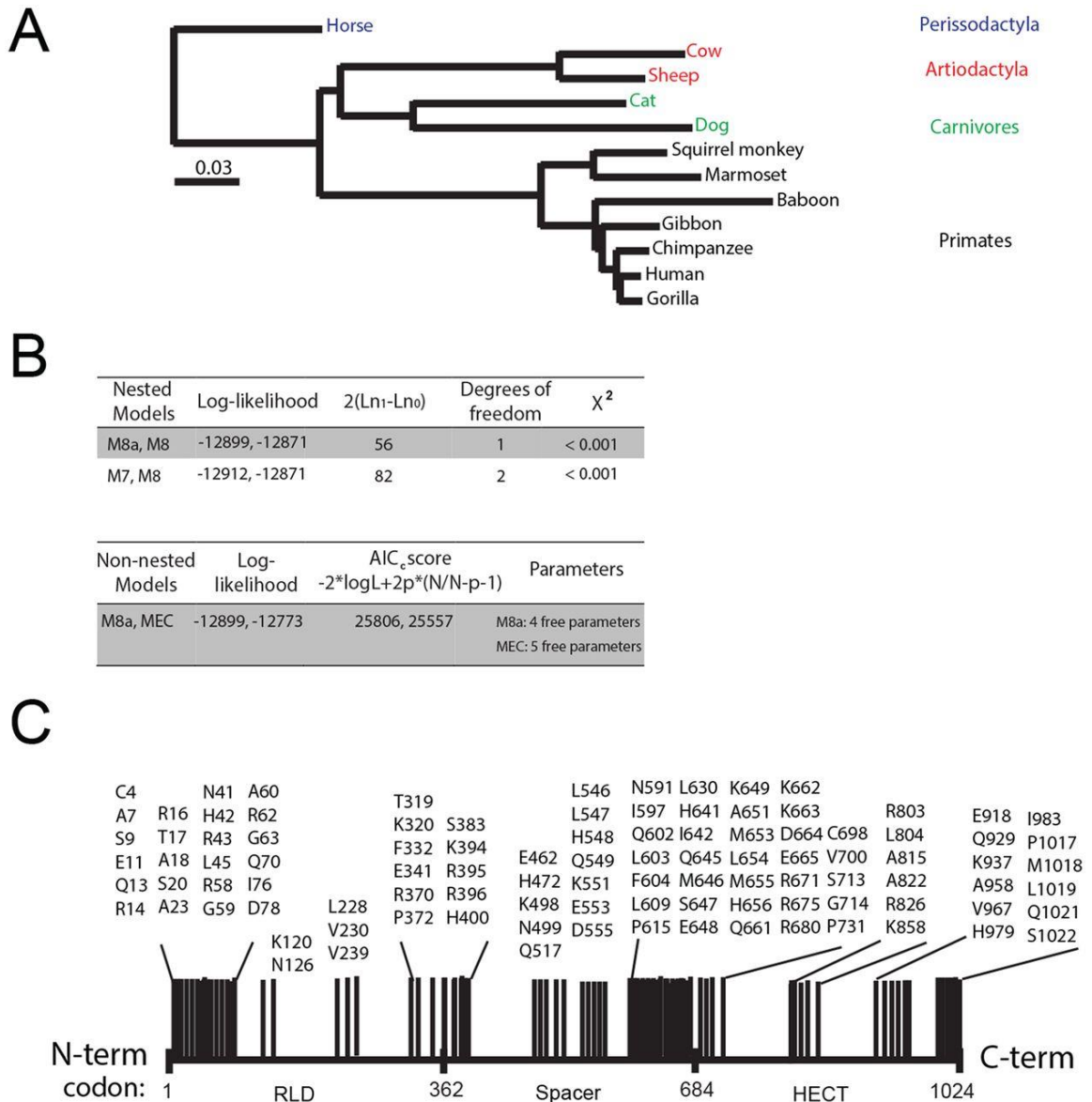
Figure 2.6 HERC6 is evolving under strong positive evolutionary selection.

A) Neighbor-joining phylogenetic tree for progressive alignment of 12 different HERC6 species using COBALT for multiple protein sequences. The branch lengths are proportional to the amount of inferred evolutionary change. **(B)** Analysis for positive selection was performed using HERC6 amino acid sequences from human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla gorilla*), marmoset (*Callithrix jacchus*), baboon (*Papio anubis*), squirrel monkey (*Saimiri boliviensis*), gibbon (*Nomascus leucogenys*), horse (*Equus caballus*), sheep (*Ovis aries*), cow (*Bos taurus*), dog (*Canis lupus familiaris*), and cat (*Felis catus*). Evolutionary analysis for positive selection in HERC6 used various models of evolution, where M8 and MEC allowed sites to evolve under positive selection and M7 and M8a did not. L represents the likelihood of the model given the data, p represents the number of free parameters, and N represents the sequence length. The lower the AICC score, the better the fit of the model to the data, and hence, the model is considered more justified**. (C)** Schematic showing the results of a Bayesian analysis approach identifying positively selected sites with a ratio of the number of nonsynonymous substitutions per nonsynonymous site (Ka) to the number of synonymous substitutions per synonymous site (Ks) (Ka/Ks) with a value of >1.5 and a 95% confidence interval larger than 1 and therefore considered statistically significant. The HERC6 reference sequence accession number is NM_017912.3

2.2.6. Blade 1 of the RCC1-like domain of human HERC6 is an important determinant of anti-HIV-1 activity.

Given the evolutionary similarities between human HERC5 and human HERC6, we asked why they differed in their antiviral activities. Since we previously showed that blade 1 of HERC5 is required for its anti-HIV-1 activity and that blades 1 of both HERC5 and HERC6 contain numerous residues evolving under positive selection, we asked if blade 1 of HERC5 can confer antiviral activity on HERC6. We replaced either the entire RCC1-like domain (H6:H5RLD) or blade 1 (H6:H5blade1) from human HERC5 with the corresponding region in human HERC6. We then measured the abilities of these HERC6

mutants to inhibit HIV-1 particle production. As shown in Figure 2.7A, the H6:H5RLD and H6:H5blade1 mutants potently inhibited HIV-1 particle production similarly to wild-type HERC5. This inhibition occurred despite levels of H6:H5RLD and H6:H5blade1 protein expression slightly lower than that of wild-type HERC5 (Figure 2.7B). This result indicates that blade 1 is an important determinant of antiviral activity.
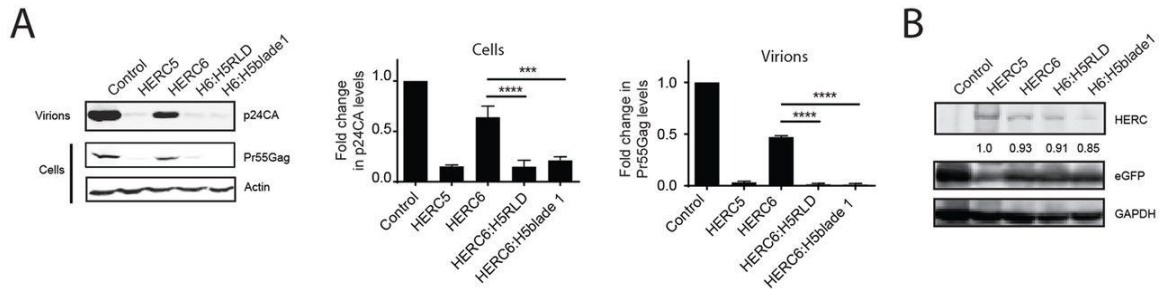
Figure 2.7 Blade 1 of the human HERC5 RCC1-like domain confers potent antiviral activity on human HERC6.

A) 293T cells were cotransfected with pR9, peGFP, and either empty vector, pHERC5, pHERC6, pH 6:H5RLD, or pH6:H5blade1. Forty-eight hours post transfection, virions in the supernatant and total cell lysates were subjected to quantitative Western blot analysis using anti-p24CA or anti-β-actin. Densitometric quantification of p24CA (virions) and Pr55Gag (cells) from the results of three independent experiments is shown beside the blots after normalization to the β-actin loading controls. The error bars indicate SEM. ***, $P < 0.001$; ****, $P < 0.0001$. **(B)** Cell lysates from panel A were subjected to Western blot analysis using anti-Flag, anti-eGFP, or anti-GAPDH. Average densitometric quantification of the HERC bands is shown below the blot after normalization to anti-eGFP or anti-GAPDH from the results of three independent experiments.

2.3.   Discussion

We showed here that the small-HERC family has an ancient marine origin, where HERC4 emerged at least 595 mya and expansion of the family occurred sometime after

the divergence of jawed vertebrates from jawless vertebrates (∼476 to 595 mya). Elephant sharks are among the oldest and most slowly evolving jawed vertebrates and have accumulated a small number of chromosomal rearrangements[44]. Thus, analysis of their genome allows us to gain insight into the early evolution and expansion of gene families. The presence of a single copy of HERC4 and multiple copies of HERC3 in the elephant shark likely represents an early time in vertebrate evolution when the HERC4 ancestral gene duplicated and evolved into HERC3. Although the evolutionary pressures in vertebrates triggering expansion of the small-HERC family are unknown, it is plausible, given their antiretroviral activity, that this expansion involved retroviruses. Endogenous retroviruses (ERVs) comprise a substantial portion of vertebrate genomes and appear to have an ancient marine origin, with evidence of ERV sequences found in the genomes of elephant shark, coelacanth, and possibly lamprey[66–69]. A recent panvertebrate comparative genomic analysis showed that retroviruses have an unprecedented capacity for rampant host switching among distantly related vertebrates, undoubtedly exerting substantial evolutionary pressure on their hosts[69]. Pressures like these can trigger antiviral gene duplication and neofunctionalization events in the hosts, allowing them to evolve more rapidly in order to maintain evolutionary dominance over viruses. Several examples where gene duplication/neofunctionalization has given rise to restriction factor families in primates are MX1, IFITM, TRIM5, and APOBEC3[5,6,70–75]. These genes, including HERC5 and HERC6, exhibit strong signatures of positive selection, which is consistent with repeated exposure to such evolutionary pressures.

An interesting feature of the small-HERC family is the highly conserved topology of the RCC1-like domain, despite limited sequence homology. By allowing numerous amino acid substitutions while maintaining the overall protein configuration and antiviral activity, these HERC proteins may be able to interfere with the binding of diverse viral antagonists. Evidence of such an evolutionary battle lies in the strong signatures of positive selection in both human HERC5 and HERC6, especially blades 1 of their RCC1-like domains, which we have shown to be important determinants of antiviral activity. HERCs are not the only restriction factors likely to have played an important antiviral

role early in vertebrate evolution. BST2/tetherin has also been shown to have an ancient marine origin, emerging >450 mya, before the separation of cartilaginous fish from bony vertebrates[2]. Like the HERC family, the general topology of BST2/tetherin orthologs are also highly conserved despite low sequence homology, and it is the overall protein configuration that is important for its antiviral activity[2,76,77]. This evolutionary strategy may help BST2/tetherin and HERC proteins maintain evolutionary dominance over viruses.

We observed that some small HERC genes have been lost in some species, most notably birds and rodents. For birds, this is not too surprising, given that their genomes have been subjected to lineage-specific erosion of repetitive elements, large segmental deletions, and gene loss (>1,000 genes), resulting in a smaller repertoire of immune genes than in humans[78,79]. Although HERC5 and HERC6 are missing in most bird species, HERC3 and/or HERC4 are present. Given their modest antiretroviral activity in humans, it will be interesting to learn if HERC3 and HERC4 play an antiviral role in birds, perhaps compensating for the loss of HERC5 and HERC6. Other potent antiviral restriction factor genes, such as BST2/tetherin, TRIM22, TRIM5, and APOBEC3G, are also notably absent from birds; however, they do possess other members of the BST, TRIM, and APOBEC families, whose antiviral activities remain largely uncharacterized in birds. Rodents also have HERC3 and HERC4 but also possess at least one of the HERC5 and HERC6 genes. For example, mice, rats, and hamsters possess HERC6 but lack HERC5, ground squirrels possess HERC5 but lack HERC6, and guinea pigs possess both HERC5 and HERC6. Since no rodent lacks both HERC5 and HERC6, it is likely that one of these genes has assumed the role of the main cellular E3 ligase for ISG15 in the absence of the other, potentially adding a new antiviral defense for rodents. For instance, this is the case in mice, which possess only HERC6, which encodes the main cellular E3 ligase for ISG15 and serves as a critical antiviral defense mechanism in mice[80,81]. Our phylogenetic analysis, where we showed that the predicted structures of the HECT domains of mouse HERC6 and human HERC5 (the main cellular E3 ligase for ISG15 in humans) share a high degree of similarity, also supports this finding. One possibility for the loss of

HERC5 or HERC6 in rodents is that retroviruses or other viral pathogens have not provided constant selective pressure to maintain both genes in these species. A similar dynamic history of gene expansion and loss is evident for other restriction factors, such as TRIM22, TRIM5, and BST2/tetherin[2,7,82].

As genes duplicate, neofunctionalize, and diverge in response to evolutionary pressures, reliance on the activities of the ancestral genes may diminish, or they may be replaced altogether by their more advantageous descendants (reviewed in reference 83)[83]. Therefore, it is not surprising that the human small-HERC family exhibits differential antiviral activity. Despite still possessing antiviral activity when overexpressed in vitro, it is unlikely that HERC3 and HERC4 play significant biological roles as antiviral proteins in humans, since they are not IFN induced, nor have they been evolving under positive selection. This role was likely assumed by HERC5 and HERC6 after the divergence of ray-finned fish from cartilaginous fish (~430 mya). However, HERC3 and HERC4 do exhibit differential tissue-specific expression (reviewed in reference 37)[37]. Therefore, it is possible that HERC3 and HERC4 play more dominant antiviral roles in tissues where their basal expression is already much more elevated, such as in the brain, heart, and stomach for HERC3 and the brain, lung, and testis for HERC4. Moreover, species such as elephant shark, coelacanth, and platypus that contain duplicated copies of HERC3 or HERC4 genes may also express higher levels of these HERC proteins due to increased gene copy numbers. Although we did not test the antiviral activities of other evolutionarily diverse HERC3 or HERC4 proteins, our phylogenetic analyses demonstrated that the HECT and RCC1-like domains of these proteins show remarkable structural similarity to their human counterparts, which do exhibit antiviral activity at elevated levels. It will be interesting to determine if the antiviral function of HERC3 and HERC4 is conserved in these ancient vertebrates, before the emergence of HERC5 and HERC6.

Interactions between host antiviral proteins and their viral-protein targets can be a critical requirement for their antiviral activity. When viral proteins mutate to evade such

interactions, the antiviral protein frequently develops rapid amino acid replacements at the protein-protein interface in an attempt to restore those interactions and maintain evolutionary dominance over the virus. HERC5 is known to interact with evolutionarily diverse viral proteins and, like HERC6, is evolving under strong positive selection[20–23,32,35,36]. Therefore, it is highly likely that these proteins contain one or more protein-protein interaction interfaces between viral and host proteins. Our findings that blade 1 of the RCC1-like domains of HERC5 and HERC6 contain numerous positively selected residues and that these residues are important determinants of antiretroviral activity indicate that blade 1 is likely one such interface. Although there is currently no evidence that retroviruses have driven positive selection of blade 1, it is interesting that blade 1 of HERC5 is sufficient to confer antiretroviral activity on HERC6. It is possible that the topology of blade 1 from HERC5 is such that it promotes interaction with a cellular protein required for activity that blade 1 of HERC6 prevents. However, our finding that wild-type HERC6 potently inhibited SIVmac239, but not HIV-1, in the same cell type suggests that virus-specific differences are more likely to account for the observed differential antiviral activity between HERC5 and HERC6. Additional structure-function studies are needed to differentiate between these possibilities and others.

In conclusion, the small-HERC family has an evolutionarily ancient origin more than 595 mya, with the latest expansion of the family occurring more than 413 mya. We showed that the structural topologies of the HECT and RCC1-like domains are highly conserved despite low sequence homology and that the antiretroviral activity of HERC5 has an ancient marine origin. HERC5 and HERC6 are evolving under strong positive selection, and a patch of positively selected residues in blade 1 of the RCC1-like domain is a strong determinant of antiviral activity. Altogether, our study highlights the potential importance of the HERC family in intrinsic immunity.

2.4.   Materials and Methods

2.4.1. Cell lines.

293T cells were obtained from the American Type Culture Collection. HOS-CD4-CXCR4 and TZM-bl cells were from the NIH AIDS Reagent Program. The cells were maintained in standard growth medium (Dulbecco's modified Eagle's medium [DMEM]) supplemented with 10% heat-inactivated fetal bovine serum (FBS), 100 U/ml penicillin, and 100 μg/ml streptomycin at 37°C with 5% CO2.

2.4.2. Analyses of sequences and synteny.

The MEGA 7.0 package was used for phylogenetic analysis[84]. The amino-terminal end of the small HERCs varies in length among the different members and was not included for phylogenetic analysis. The first 30 amino acids were omitted from HERC3 and HERC4 and the first 23 amino acids from HERC5 and HERC6. Accession numbers used are listed in Table 1.

Table 1: Accession numbers used for analyses of sequences and synteny

Protein	Species	Accession Number
HERC3	*Homo sapiens* (human)	NP_055421.1
	Gorilla gorilla gorilla (gorilla)	XP_004039158.1
	Pan troglodytes (chimpanzee)	XP_517337.3
	Nomascus leucogenys (gibbon)	XP_003265945.1
	Papio anubis (baboon)	XP_003898995.1
	Saimiri boliviensis boliviensis (squirrel monkey)	XP_003924058.1
	Callithrix jacchus (marmoset)	XP_002745644.1

	Bos taurus (cow)	NP_001077132.1
	Ovis aries (sheep)	XP_004009758.1
	Felis catus (cat)	XP_003985228.1
	Canis lupus familiaris (dog)	XP_535653.3
	Ailuropoda melanoleuca (giant panda)	XP_002913643.1
	Equus caballus (horse)	XP_001496703.3
	Callorhinchus_milii (elephant shark)	XM_007902532.1
	Danio rerio (zebrafish)	NM_001145624.1
	Takifugu rubripes (pufferfish)	XM_011612933.1
	Oreochromis niloticus (tilapia)	XM_005456948.2
	Latimeria chalumnae (coelacanth)	XM_006005071.2
	Xenopus tropicalis (frog)	XM_002938630.3
	Anolis carolinensis (lizard)	XM_008111037.2
	Geospiza fortis (finch)	XM_005418124.1
	Meleagris gallopavo (turkey)	XM_003205471.2
	Gallus gallus (chicken)	XM_015276202.1
	Monodelphis domestica (opossum)	XM_007495699.2
	Mus musculus (mouse)	NM_028705.3

	Dasypus novemcinctus (armadillo)	XM_004480548.2
HERC4	*Homo sapiens* (human)	NP_071362.1
	Gorilla gorilla gorilla (gorilla)	XP_004049535.1
	Pan troglodytes (chimpanzee)	XP_001167753.1
	Nomascus leucogenys (gibbon)	XP_003258260.1
	Papio anubis (baboon)	XP_003903924.1
	Saimiri boliviensis boliviensis (squirrel monkey)	XP_003928728.1
	Callithrix jacchus (marmoset)	XP_002756356.1
	Bos taurus (cow)	NP_001070362.2
	Ovis aries (sheep)	XP_004021455.1
	Canis lupus familiaris (dog)	XP_849808.1
	Ailuropoda melanoleuca (giant panda)	XP_002913788.1
	Equus caballus (horse)	XP_001503636.1
	Callorhinchus_milii (elephant shark)	XM_007897346.1
	Danio rerio (zebrafish)	XM_005173035.3
	Takifugu rubripes (pufferfish)	XM_011602912.1
	Oreochromis niloticus (tilapia)	XM_005473848.2
	Latimeria chalumnae (coelacanth)	XM_005992449.2

	Xenopus tropicalis (frog)	NM_001128650.1
	Anolis carolinensis (lizard)	XM_008115175.2
	Geospiza fortis (finch)	XM_005428501.2
	Meleagris gallopavo (turkey)	XM_010714265.1
	Gallus gallus (chicken)	XM_015278711.1
	Ornithorhynchus anatinus (platypus)	XM_007667149.1
	Monodelphis domestica (opossum)	XM_016421882.1
	Mus musculus (mouse)	NM_030114.2)
	Dasypus novemcinctus (armadillo)	XM_012529817.1
	Petromyzon marinus (lamprey)	ENSPMAG00000001626
HERC5	*Homo sapiens* (human)	NP_057407.2
	Pan troglodytes (chimpanzee)	XP_003310459.1
	Gorilla gorilla gorilla (gorilla)	XP_004039179.1
	Callithrix jacchus (marmoset)	XP_002745648.1
	Papio anubis (baboon)	XP_003898997.1
	Saimiri boliviensis boliviensis (squirrel monkey)	XP_003924055.1
	Nomascus leucogenys (gibbon)	XP_003265940.1
	Equus caballus (horse)	XP_001915115.2

	Ailuropoda melanoleuca (giant panda)	XP_002913645.1
	Ovis aries (sheep)	XP_004009762.1
	Bos taurus (cow)	NP_001095465.1
	Canis lupus familiaris (dog)	XP_535652.3
	Felis catus (cat)	XP_003985249.1
	Latimeria chalumnae (coelacanth)	XM_014498805.1
	Anolis carolinensis (lizard)	XM_008111035.2
	Dasypus novemcinctus (armadillo)	XM_012520464.1
HERC6	*Homo sapiens* (human)	NP_060382.3
	Gorilla gorilla gorilla (gorilla)	XP_004039178.1
	Pan troglodytes (chimpanzee)	XP_001160851.1
	Nomascus leucogenys (gibbon)	XP_003265938.1
	Papio anubis (baboon)	XP_003899001.1
	Saimiri boliviensis boliviensis (squirrel monkey)	XP_003924053.1
	Callithrix jacchus (marmoset)	XP_002745681.1
	Bos taurus (cow)	NP_001179573.1
	Ovis aries (sheep)	XP_004010096.1
	Felis catus (cat)	XP_003985250.1

	Canis lupus familiaris (dog)	XP_851549.1
	Equus caballus (horse)	XP_001494887.1
	Takifugu rubripes (pufferfish)	XM_011618146.1
	Oreochromis niloticus (tilapia)	XM_005474674.1
	Latimeria chalumnae (coelacanth)	XM_014498807.1
	Xenopus tropicalis (frog)	XM_002938624.2
	Anolis carolinensis (lizard)	XM_008111034.2
	Monodelphis domestica (opossum)	XM_007495704.1
	Mus musculus (mouse)	NM_025992.2
	Dasypus novemcinctus (armadillo)	XM_012520459.1

2.4.3. Synteny.

Synteny maps were derived using the reference assemblies from Table 2.

Table 2: Reference assemblies used for synteny maps.

Species	**Common name**	**Assembly**
Callorhinchus milii	elephant shark	6.1.3 GCF_000165045.1
Danio rerio	zebrafish	GRCz10 GCF_000002035.5
Takifugu rubripes	pufferfish	FUGU5 GCF_000180615.1

Oreochromis niloticus	tilapia	Orenil1.1 GCF_000188235.2
Latimeria chalumnae	coelacanth	LatCha1 GCF_000225785.1
Xenopus tropicalis	frog	Xtropicalis_v7 GCF_000004195.2
Anolis carolinensis	lizard	AnoCar2.0 GCF_0000090745.1
Geospiza fortis	finch	GeoFor_1.0 GCF_000277835.1
Meleagris gallopavo	Turkey	Turkey_5.0 GCF_000146605.2
Gallus gallus	chicken	5.0 GCF_000002315.4
Ornithorhynchus anatinus	platypus	5.0.1 GCF_000002275.2
Monodelphis domestica	opossum	MonDom5 GCF_000002295.2
Dasypus novemcinctus	armadillo	Dasnovv3.0 GCF_000208655.1
Loxodonta africana	elephant	Loxafr3.0 GCF_000001905.1
Mus musculus	mouse	GRCm38.p3 GCF_000001635.23
Rattus norvegivcus	rat	Rnor_6.0 GCF_000001895.5
Elephantulus edwardii	shrew	EleEdw1.0 GCF_000299155.1
Chrysochloris asiatica	cape golden mole	ChrAsi1.0 GCF_000296735.1
Canis lupus familiaris	dog	CanFam3.1 GCF_000002285.3
Homo sapiens	human	GRCh38.p2 GCF_000001405.28

2.4.4. Positive selection.

Positive-selection analysis was performed as previously described[22]. HERC sequences were aligned, and a phylogenetic tree was generated using COBALT (constraint-based alignment tool) (http://www.ncbi.nlm.nih.gov/tools/cobalt/)[85]. The following HERC sequences were obtained from GenBank: for HERC3, *Homo sapiens* (human) (NP_055421.1), *Gorilla gorilla gorilla* (gorilla) (XP_004039158.1), *Pan troglodytes* (chimpanzee) (XP_517337.3), *Nomascus leucogenys* (gibbon) (XP_003265945.1), *Papio anubis* (baboon) (XP_003898995.1), *Saimiri boliviensis boliviensis* (squirrel monkey) (XP_003924058.1), *Callithrix jacchus* (marmoset) (XP_002745644.1), *Bos taurus* (cow) (NP_001077132.1), *Ovis aries* (sheep) (XP_004009758.1), *Felis catus* (cat) (XP_003985228.1), *Canis lupus familiaris* (dog) (XP_535653.3), *Ailuropoda melanoleuca* (giant panda) (XP_002913643.1), and *Equus caballus* (horse) (XP_001496703.3); for HERC4, human (NP_071362.1), gorilla (XP_004049535.1), chimpanzee (XP_001167753.1), gibbon (XP_003258260.1), baboon (XP_003903924.1), squirrel (XP_003928728.1), marmoset (XP_002756356.1), cow (NP_001070362.2), sheep (XP_004021455.1), dog (XP_849808.1), panda (XP_002913788.1), and horse (XP_001503636.1); for HERC5, human (NP_057407.2), chimpanzee (XP_003310459.1), gorilla (XP_004039179.1), marmoset (XP_002745648.1), baboon (XP_003898997.1), squirrel monkey (XP_003924055.1), gibbon (XP_003265940.1), horse (XP_001915115.2), giant panda (XP_002913645.1), sheep (XP_004009762.1), cow (NP_001095465.1), dog (XP_535652.3), and cat (XP_003985249.1); for HERC6, human (NP_060382.3), gorilla (XP_004039178.1), chimpanzee (XP_001160851.1), gibbon (XP_003265938.1), baboon (XP_003899001.1), squirrel (XP_003924053.1), marmoset (XP_002745681.1), cow (NP_001179573.1), sheep (XP_004010096.1), cat (XP_003985250.1), dog (XP_851549.1), and horse (XP_001494887.1). At least 2 independent sequences were available for human, sheep, baboon, marmoset, gibbon, and squirrel monkey. The following sequences were not independently validated: cat, dog, cow, horse, sheep, and giant panda. The identification of site-specific positive selection and purifying selection was generated using the Selecton server

(http://selecton.tau.ac.il/index.html). The HERC5 phylogenetic tree was used in the Selecton analysis. Nested pairs of models (M8a and M8; M7 and M8) and a nonnested pair (M8a and MEC) were compared using the likelihood ratio test implemented in the Selecton program.

Plasmids, transfections, antibodies, and Western blotting.

2.4.4.1. Plasmids.

Plasmids encoding Flag-tagged HERC3, HERC4, HERC5, and HERC6 were created by first PCR amplifying the various HERC coding regions from their respective templates (as described above) using the primers in Table 3 rows 1 to 4. The amplified products were then cloned into p3xFLAG-CMV10 using the following restriction enzymes: for HERC3, EcoRI and KpnI; for HERC4, EcoRI and XbaI; for HERC5, BglII and BamHI; and for HERC6, BglII and EcoRV. The resulting plasmids were named pHERC3, pHERC4, pHERC5, and pHERC6.

To generate pH6:H5RLD, the HERC5 RCC1-like domain was PCR amplified from pHERC5 using the primers in Table 3 row 5. The backbone of pHERC6 was PCR amplified using the primers in Table 3 row 6. The two amplified fragments were cloned using the fast cloning technique[86].

To generate pH6:H5blade1, blade 1 of HERC5 (amino acids 1 to 100) was PCR amplified from pHERC5 using the primers from Table 3 row 7. The backbone of pHERC6 was PCR amplified using the primers from Table 3 row 8. The two amplified fragments were cloned using the Gibson cloning technique per the manufacturer's instructions (New England Biolabs).

pHERC3-ΔRLD, pHERC4-ΔRLD, pHERC6-ΔRLD, pHERC3-C1018A, pHERC4-C1025A, and pHERC6-C985A were generated similarly, using the primers from Table 3 row 9 to 14 respectively. pHERC5-ΔRLD and pHERC5-C994A were generated previously (86).

The promoterless empty-vector plasmid pGL3, p3xFLAG-CMV10, and peGFP were obtained from Promega, Sigma, and Clontech, respectively. pLKO.1/scrambled shRNA and pLKO.1/HERC5 shRNA were previously described[22,23]. The following pLKO.1 shRNA constructs were obtained from Dharmacon: HERC3-#1 (TRCN0000000291), HERC3-#3 (TRCN0000000293), HERC4-#4 (TRCN0000034302), HERC6-#2 (TRCN0000160017), and HERC6-#3 (TRCN0000160044). The coding regions of HERC3, HERC4, HERC5, and HERC6 were obtained from the following sources: HERC3 (NM_014606.2), HERC6 (NM_017912.3)[87], and HERC5 (NP_057407.2)[33]. HERC4 (NM_015601.3) was isolated from HeLa cells by first reverse transcribing total RNA and then PCR amplification of cDNA using the primers in Table 3 row 15. All the constructs were verified by sequencing. Transfections were performed using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions unless otherwise indicated. Cotransfections of HERC plasmids with pR9 were performed at a ratio of 10:1 unless otherwise noted. Standard Western blot analyses were performed as previously described[22]. Densitometric analysis was performed using ImageJ 1.43u software 64-bit version (NIH).

For the construction of pSIVmac239 (pREC_nfl_SIV239), the SIVmac239 Spx vector was obtained from the NIH AIDS Reagent Program. We previously constructed pREC_nfl_HIV and pCMV_cplt vectors for Saccharomyces cerevisiae-based cloning of diverse HIV-1 strains[88]. We developed a similar method for SIV cloning. To generate pREC_nfl_SIV239, the 5′ half of the HIV genome in the pREC_nfl_HIV vector was first replaced with URA3 and then with the 5′ half of the SIV239 genome through the yeast recombination technique described below. Yeast colonies were selected on C-leu plates supplemented with uracil but lacking leucine for selection of pREC_nfl_HIVΔ5′HIV/URA3 and on C-leu supplemented with 5-fluoro-1,2,3,6-tetra-hydro-2,6-dioxo-4-pyrimidine carboxylic acid (5-FOA) for selection of pREC_nfl_5′SIV239/3′HIV. C-leu plates allow growth only when a plasmid containing the leucine gene is transformed into the yeast. The 3′ half of SIV239 was introduced using the same procedure to form the vector pREC_nfl_SIV239; this vector contains

nearly the full-length SIV239 genome and lacks the 5′ repeat (R) and unique (U5) regions. Approximately 95% of the FOA-resistant yeast colonies harbored pREC_nfl_SIV239. A crude yeast lysate was then used to transform bacteria and to amplify these ampicillin-resistant DNA plasmids for purification, as described previously[88]. For yeast recombination, *S. cerevisiae* Hanson (MYA-906) (MATα ade6 can1 his3 leu2 trp1 URA3) was obtained from the American Type Culture Collection (ATCC). The yeast was grown at 30°C in appropriate medium (yeast extract peptone dextrose [YEPD] or complete [C] minimal medium C-LEU-URA3, C-LEU, or C-LEU/5-FOA), depending on the cloning step. Transformations/recombinations were performed using the lithium acetate (LiAc) method. Briefly, the linearized vector DNA (~1 μg) and PCR product (~3 μg) were added to competent cells at a 1:3 ratio, along with 50 μg of single-stranded salmon sperm carrier DNA (BD Biosciences/Clontech, Palo Alto, CA) and sterile polyethylene glycol (50%)-TE (10 mM Tris-Cl, 1 mM EDTA)-LiAc (100 mM). Following agitation for 30 min at 30°C, the yeast was heat shocked at 42°C for 15 min and plated on C-leu agar plates containing the appropriate selection.

2.4.4.2. Antibodies.

The following reagents were obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: HIV-1 p24 monoclonal antibody (183-H12-5C) from Bruce Chesebro and Kathy Wehrly and anti-SIVmac p17 monoclonal antibody (KK59) from Karen Kent and Caroline Powell. Anti-FLAG and anti-hemagglutinin (HA) (clone 3F10) were purchased from Sigma, anti-myc and anti-β-actin were purchased from Rockland, anti-eGFP was purchased from Clontech, and anti-GAPDH (glyceraldehyde-3-phosphate dehydrogenase) (clone 6C5) was purchased from EMD/Millipore. Anti-HERC3 (H00008916-B01P), anti-HERC4 (H00026091-A01), anti-HERC5 (H00051191-A01), and anti-HERC6 (H00055008-A01) were purchased from Abnova.

2.4.5. Quantitative PCR.

Total RNA was extracted using the PureLink RNA mini kit (Ambion, Life Technologies). Three micrograms of RNA were reverse transcribed to cDNA using Moloney murine leukemia virus (MMLV) reverse transcriptase and oligo(dT) primers (Life Technologies). Prior to qPCR, the cDNA samples were diluted 1:5 with water. Each PCR mixture consisted of 10 µl of SYBR green master mix, 2 µl of gene-specific primers (1 µl of 10 µM forward primer and 1 µl of 10 µM reverse primer), 5 µl of diluted cDNA, and water to a total volume of 20 µl. For quantification of incompletely and fully spliced HIV RNAs, qPCR was run on the Rotor-Gene 6000 qPCR machine (Corbett Life Science) under the following cycling conditions: 10 min at 95°C and 40 cycles of 10 s at 95°C, 15 s at 60°C, and 20 s at 72°C. The Rotor-Gene 6000 series software (version 1.7) was used to determine the cycle threshold (CT) for each PCR. The gene-specific forward and reverse primer are listed in Table 3 rows 16 and 17. Quantification of endogenous HERC mRNA was run on the QuantStudio5 qPCR machine (Applied Biosystems) under the following cycling conditions: 2 min at 95°C and 40 cycles of 5 s at 95°C, 10 s at 60°C, and 20 s at 72°C. QuantStudio design and analysis desktop software (version 1.4) was used to determine the CT for each PCR. The primers are listed in Table 3 rows 18 to23. To ensure no carryover of DNA into each total purified RNA sample, 3 µg of the purified RNA was used directly as the template without reverse transcription for qPCR, using the primer sets described above.

Table 3: List of primers

1	pHERC3	5' ACG TGA ATT CCA TGT TAT GTT GGG GAT ATT GG 3'	5' ACG TGG TAC CTC AGG CCA AAC TAA ACC TTT CAT AGT TGT C 3'
2	pHERC4	5'ACG TGA ATT CTA TGT TGT GCT GGG GAA ATG C 3'	5' ACG TTC TAG ATT ATA TTA AAC TGA AGC CTT CAT TGT G 3';

3	pHERC5	5′ AAT CGA GAT CTT ATG GAG CGC CGC AGC 3′	5′ TAT GCG GAT CCT CAG CCA AAT CCT CTG 3′
4	pHERC6	5′ AGA TAA GAT CTT ATG TAC TTC TGT TGG GGC 3′,	5′ TAG GAG ATA TCT TAT GAC TGT GTG AGC ATG 3′
5	pH6:H5RLD - HERC5	5′ GGA TGA CGA TGA CAA GAT GGA GCG CCG CAG CC 3′	5′ TAT GTT CCA GCA AAA ATT ATT AAC TCC TTT TCT GAG GTA TGG CTT TCA AG 3′
6	pH6:H5RLD - pHERC6 backbone	5′ TTT TTG CTG GAA CAT ATG CCA ACT TTG 3′	5′ CTT GTC ATC GTC ATC CTT GTA ATC GAT G 3′
7	pH6:H5 blade1 - HERC5	5′ GGA TGA CGA TGA CAA GAT GGA GCG CCG CAG CCG CCC CAA CAG AAG TAC ATC TTG TCA TCG TCA TCC TTG TAA TCG ATG 3′	5′ GCT CCT TCC CGC AGC TCA CGT GGA TCT TCA TGT TCT TGC CCA GC 3′
8	pH6:H5 blade1 - pHERC6 backbone	GCT GGG CAA GAA CAT GAA GAT CCA CAG CTG CGG GAA GGA GCA C 3′	5′ GGC TGC GGC GCT CCA TCT TGT CAT CGT CAT CCT TGT AAT CGA TG 3′
9	pHERC3-ΔRLD	5′ GAC GAT GAC AAG ATG AGC TCA CCA CCA GAT GTT GAA G 3′	5′ CAT CTG GTG GTG AGC TCA TCT TGT CAT CGT CAT CCT TGT AAT CG 3′

10	pHERC4-ΔRLD	5′ ACG ATG ACA AGA TGA ATT GGT ACC CCT ATA ATG GGC AGT G 3′	5′ TAG GGG TAC CAA TTC ATC TTG TCA TCG TCA TCC TTG TAA TCG 3′
11	pHERC6-ΔRLD	5′ CGA TGA CAA GAT GAT TTT GCT GGG AAC ATA TGC CAA C 3′	5′ GTT CCA GCA AAA TCA TCT TGT CAT CGT CAT CCT TGT T 3′
12	pHERC3-C1018A	5′ CGG TGG CCC ACA CTG CTT ACA ACC TTC TTG 3′	5′ GAG GTC AAG AAG GTT GTA CGC AGT GTG GGC C 3′
13	pHERC4-C1025A	5′ CCC AGT TTC CCA TAC TGC TTT TAA TCT TCT G 3′	5′ GAA GAT CCA GAA GAT TAA AAG CAG TAT GGG AAA C 3′
14	pHERC6-C985A	5′ CCA ACA TCA ATA ACT GCT CAT AAT ATT CTC TCC C 3′	5′ GGG AGG GAG AGA ATA TTA TGA GCA GTT ATT GAT G 3′
15	HERC4 (NM_015601.3)	5′ACG TGA ATT CTA TGT TGT GCT GGG GAA ATG C 3′	5′ ACG TTC TAG ATT ATA TTA AAC TGA GCC TTC ATT GTG 3′
16	Gag	5′ CAT ATA GTA TGG GCA AGC AGG G 3′	5′ CTG TCT GAA GGG ATG GTT GTA G 3′
17	Rev	5′ GAG CTC ATC AGA ACA GTC AGA C 3′	5′ CGA ATG GAT CTG TCT CTG TCT C 3′
18	HERC3_qpcr	5′ CAG TGC CCA GGT TAA TAC AAA AG 3′	5′ GAA CTC CTT CCC TAA GCC AAG 3′

19	HERC4_qpcr	5′ TTC ATG TGG AGA AGC TCA TAC G 3′	5′ CAT CAG AAT CGA GAC CCC AAG 3′
20	HERC5_qpcr	5′ ATG AGC TAA GAC CCT GTT TGG 3′	5′ CCC AAA TCA GAA ACA TAG GCA AG 3′
21	HERC6_qpcr	5′ GCG TCA ATT AAG TCA AGC TGA AGC 3′	5′ GAA ACC ACA TGC AGG AAC CC 3′
22	GAPDH_qpcr	5′ CAT GTT CGT CAT GGG TGT GAA CCA 3′	5′ AGT GAT GGC ATG GAC TGT GGT CAT 3′
23	eGFP_qpcr	5′ GAC AAC CAC TAC CTG AGC AC 3'	5′ CAG GAC CAT GTG ATC GCG 3′

2.4.6. Statistical analyses.

GraphPad Prism v6 was used for all statistical analyses mentioned in the text. The P values and statistical tests used are mentioned where appropriate. P values of less than 0.05 were deemed significant.

2.5.   References

1.   Bieniasz PD. Intrinsic immunity: A front-line defense against viral attack. *Nat Immunol*. 2004;5(11):1109-1115. doi:10.1038/NI1125

2.   Heusinger E, Kluge SF, Kirchhoff F, Sauter D. Early Vertebrate Evolution of the Host Restriction Factor Tetherin. *J Virol*. 2015;89(23):12154-12165. doi:10.1128/JVI.02149-15

3.   Münk C, Willemsen A, Bravo IG. An ancient history of gene duplications, fusions and losses in the evolution of APOBEC3 mutators in mammals. *BMC Evol Biol*.

2012;12(1). doi:10.1186/1471-2148-12-71

4.   Johnson WE, Sawyer SL. Molecular evolution of the antiretroviral TRIM5 gene. *Immunogenetics*. 2009;61(3):163-176. doi:10.1007/S00251-009-0358-Y

5.   Conticello SG, Thomas CJF, Petersen-Mahrt SK, Neuberger MS. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol Biol Evol*. 2005;22(2):367-377. doi:10.1093/MOLBEV/MSI026

6.   Tareen SU, Sawyer SL, Malik HS, Emerman M. An expanded clade of rodent Trim5 genes. *Virology*. 2009;385(2):473-483. doi:10.1016/J.VIROL.2008.12.018

7.   Sawyer SL, Emerman M, Malik HS. Discordant evolution of the adjacent antiretroviral genes TRIM22 and TRIM5 in mammals. *PLoS Pathog*. 2007;3(12):1918-1929. doi:10.1371/JOURNAL.PPAT.0030197

8.   Conticello SG. The AID/APOBEC family of nucleic acid mutators. *Genome Biol*. 2008;9(6). doi:10.1186/GB-2008-9-6-229

9.   Hsiang T-Y, Zhao C, Krug RM. Interferon-Induced ISG15 Conjugation Inhibits Influenza A Virus Gene Expression and Replication in Human Cells. *J Virol*. 2009;83(12):5971-5977. doi:10.1128/JVI.01667-08

10.   Lenschow DJ, Lai C, Frias-Staheli N, et al. IFN-stimulated gene 15 functions as a critical antiviral molecule against influenza, herpes, and Sindbis viruses. *Proc Natl Acad Sci U S A*. 2007;104(4):1371-1376. doi:10.1073/PNAS.0607038104

11.   Giannakopoulos N V., Arutyunova E, Lai C, Lenschow DJ, Haas AL, Virgin HW. ISG15 Arg151 and the ISG15-Conjugating Enzyme UbE1L Are Important for Innate Immune Control of Sindbis Virus. *J Virol*. 2009;83(4):1602-1610. doi:10.1128/JVI.01590-08

12.   Lenschow DJ, Giannakopoulos N V., Gunn LJ, et al. Identification of Interferon-

Stimulated Gene 15 as an Antiviral Molecule during Sindbis Virus Infection In Vivo. *J Virol*. 2005;79(22):13974-13983. doi:10.1128/JVI.79.22.13974-13983.2005

13.   Malakhova OA, Zhang DE. ISG15 inhibits Nedd4 ubiquitin E3 activity and enhances the innate antiviral response. *J Biol Chem*. 2008;283(14):8783-8787. doi:10.1074/JBC.C800030200

14.   Okumura A, Pitha PM, Harty RN. ISG15 inhibits Ebola VP40 VLP budding in an L-domain-dependent manner by blocking Nedd4 ligase activity. *Proc Natl Acad Sci U S A*. 2008;105(10):3974-3979. doi:10.1073/PNAS.0710629105

15.   Guerra S, Cáceres A, Knobeloch KP, Horak I, Esteban M. Vaccinia virus E3 protein prevents the antiviral action of ISG15. *PLoS Pathog*. 2008;4(7). doi:10.1371/JOURNAL.PPAT.1000096

16.   Dai J, Pan W, Wang P. ISG15 facilitates cellular antiviral response to dengue and west nile virus infection in vitro. *Virol J*. 2011;8. doi:10.1186/1743-422X-8-468

17.   Sun Z, Li Y, Ransburgh R, Snijder EJ, Fang Y. Nonstructural Protein 2 of Porcine Reproductive and Respiratory Syndrome Virus Inhibits the Antiviral Function of Interferon-Stimulated Gene 15. *J Virol*. 2012;86(7):3839-3850. doi:10.1128/JVI.06466-11

18.   Jacobs SR, Stopford CM, West JA, Bennett CL, Giffin L, Damania B. Kaposi's Sarcoma-Associated Herpesvirus Viral Interferon Regulatory Factor 1 Interacts with a Member of the Interferon-Stimulated Gene 15 Pathway. *J Virol*. 2015;89(22):11572-11583. doi:10.1128/JVI.01482-15

19.   González-Sanz R, Mata M, Bermejo-Martín J, et al. ISG15 Is Upregulated in Respiratory Syncytial Virus Infection and Reduces Virus Growth through Protein ISGylation. *J Virol*. 2016;90(7):3428-3438. doi:10.1128/JVI.02695-15

20.   Durfee LA, Lyon N, Seo K, Huibregtse JM. The ISG15 Conjugation System
       Broadly Targets Newly Synthesized Proteins: Implications for the Antiviral
       Function of ISG15. *Mol Cell*. 2010;38(5):722-732.
       doi:10.1016/J.MOLCEL.2010.05.002

21.   Tang Y, Zhong G, Zhu L, et al. Herc5 Attenuates Influenza A Virus by Catalyzing
       ISGylation of Viral NS1 Protein. *J Immunol*. 2010;184(10):5777-5790.
       doi:10.4049/JIMMUNOL.0903588

22.   Woods MW, Tong JG, Tom SK, et al. Interferon-induced HERC5 is evolving
       under positive selection and inhibits HIV-1 particle production by a novel
       mechanism targeting Rev/RRE-dependent RNA nuclear export. *Retrovirology*.
       2014;11(1). doi:10.1186/1742-4690-11-27

23.   Woods MW, Kelly JN, Hattlmann CJ, et al. Human HERC5 restricts an early stage
       of HIV-1 assembly by a mechanism correlating with the ISGylation of Gag.
       Retrovirology. 2011;8:95. doi:10.1186/1742-4690-8-95

24.   Kim YJ, Kim ET, Kim YE, et al. Consecutive Inhibition of ISG15 Expression and
       ISGylation by Cytomegalovirus Regulators. *PLoS Pathog*. 2016;12(8).
       doi:10.1371/JOURNAL.PPAT.1005850

25.   Lai C, Struckhoff JJ, Schneider J, et al. Mice Lacking the ISG15 E1 Enzyme
       UbE1L Demonstrate Increased Susceptibility to both Mouse-Adapted and Non-
       Mouse-Adapted Influenza B Virus Infection. *J Virol*. 2009;83(2):1147-1151.
       doi:10.1128/JVI.00105-08

26.   Okumura A, Lu G, Pitha-Rowe I, Pitha PM. Innate antiviral response targets HIV-
       1 release by the induction of ubiquitin-like protein ISG15. *Proc Natl Acad Sci U S
       A*. 2006;103(5):1440-1445. doi:10.1073/PNAS.0510518103

27.   Künzi MS, Pitha PM. Role of interferon-stimulated gene ISG-15 in the interferon-
       ω-mediated inhibition of human immunodeficiency virus replication. *J Interf*

Cytokine Res. 1996;16(11):919-927. doi:10.1089/JIR.1996.16.919

28.   Chua PK, McCown MF, Rajyaguru S, et al. Modulation of alpha interferon anti-hepatitis C virus activity by ISG15. *J Gen Virol*. 2009;90(12):2929-2939. doi:10.1099/VIR.0.013128-0

29.   Kim M-J, Yoo J-Y. Inhibition of Hepatitis C Virus Replication by IFN-Mediated ISGylation of HCV-NS5A. *J Immunol*. 2010;185(7):4311-4318. doi:10.4049/JIMMUNOL.1000098

30.   Domingues P, Bamford CGG, Boutell C, McLauchlan J. Inhibition of hepatitis C virus RNA replication by ISG15 does not require its conjugation to protein substrates by the HERC5 E3 ligase. *J Gen Virol*. 2015;96(11):3236-3242. doi:10.1099/JGV.0.000283

31.   Hsiao NW, Chen JW, Yang TC, et al. ISG15 over-expression inhibits replication of the Japanese encephalitis virus in human medulloblastoma cells. *Antiviral Res*. 2010;85(3):504-511. doi:10.1016/J.ANTIVIRAL.2009.12.007

32.   Dastur A, Beaudenon S, Kelley M, Krug RM, Huibregtse JM. Herc5, an interferon-induced HECT E3 enzyme, is required for conjugation of ISG15 in human cells. *J Biol Chem*. 2006;281(7):4334-4338. doi:10.1074/JBC.M512830200

33.   Wong JJY, Pung YF, Sze NSK, Chin KC. HERC5 is an IFN-induced HECT-type E3 protein ligase that mediates type I IFN-induced ISGylation of protein targets. *Proc Natl Acad Sci U S A*. 2006;103(28):10735-10740. doi:10.1073/PNAS.0600397103

34.   Kroismayr R, Baranyi U, Stehlik C, Dorfleutner A, Binder BR, Lipp J. HERC5, a HECT E3 ubiquitin ligase tightly regulated in LPS activated endothelial cells. *J Cell Sci*. 2004;117(20):4749-4756. doi:10.1242/JCS.01338

35.   Zhao C, Hsiang TY, Kuo RL, Krug RM. ISG15 conjugation system targets the

viral NS1 protein in influenza A virus-infected cells. *Proc Natl Acad Sci U S A*. 2010;107(5):2253-2258. doi:10.1073/PNAS.0909144107

36. Versteeg GA, Hale BG, van Boheemen S, Wolff T, Lenschow DJ, García-Sastre A. Species-Specific Antagonism of Host ISGylation by the Influenza B Virus NS1 Protein. *J Virol*. 2010;84(10):5423-5430. doi:10.1128/JVI.02395-09

37. Sánchez-Tena S, Cubillos-Rojas M, Schneider T, Rosa JL. Functional and pathological relevance of HERC family proteins: A decade later. *Cell Mol Life Sci*. 2016;73(10):1955-1968. doi:10.1007/S00018-016-2139-8

38. Hadjebi O, Casas-Terradellas E, Garcia-Gonzalo FR, Rosa JL. The RCC1 superfamily: From genes, to function, to disease. *Biochim Biophys Acta - Mol Cell Res*. 2008;1783(8):1467-1479. doi:10.1016/J.BBAMCR.2008.03.015

39. Stevens TJ, Paoli M. RCC1-like repeat proteins: A pangenomic, structurally diverse new superfamily of β-propeller domains. *Proteins Struct Funct Genet*. 2008;70(2):378-387. doi:10.1002/PROT.21521

40. Bischoff FR, Ponstingl H. Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature*. 1991;354(6348):80-82. doi:10.1038/354080A0

41. Weis K. Regulating access to the genome: Nucleocytoplasmic transport throughout the cell cycle. *Cell*. 2003;112(4):441-451. doi:10.1016/S0092-8674(03)00082-5

42. Harel A, Forbes DJ. Importin beta: conducting a much larger cellular symphony. *Mol Cell*. 2004;16(3):319-330. doi:10.1016/J.MOLCEL.2004.10.026

43. Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. Tree of life reveals clock-like speciation and diversification. *Mol Biol Evol*. 2015;32(4):835-845. doi:10.1093/MOLBEV/MSV037

44. Venkatesh B, Lee AP, Ravi V, et al. Elephant shark genome provides unique

insights into gnathostome evolution. *Nature*. 2014;505(7482):174-179. doi:10.1038/NATURE12826

45. Amemiya CT, Alfoldi J, Lee AP, et al. The African coelacanth genome provides insights into tetrapod evolution. *Nature*. 2013;496(7445):311-316. doi:10.1038/NATURE12027

46. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*. 1992;8(3):275-282. doi:10.1093/BIOINFORMATICS/8.3.275

47. Hochrainer K, Mayer H, Baranyi U, Binder BR, Lipp J, Kroismayr R. The human HERC family of ubiquitin ligases: Novel members, genomic organization, expression profiling, and evolutionary aspects. *Genomics*. 2005;85(2):153-164. doi:10.1016/J.YGENO.2004.10.006

48. Offman MN, Fitzjohn PW, Bates PA. Developing a move-set for protein model refinement. *Bioinformatics*. 2006;22(15):1838-1845. doi:10.1093/BIOINFORMATICS/BTL192

49. Contreras-Moreira B, Fitzjohn PW, Offman M, Smith GR, Bates PA. Novel Use of a Genetic Algorithm for Protein Structure Prediction: Searching Template and Sequence Alignment Space. *Proteins Struct Funct Genet*. 2003;53(SUPPL. 6):424-429. doi:10.1002/PROT.10549

50. Söding J. Protein homology detection by HMM-HMM comparison. *Bioinformatics*. 2005;21(7):951-960. doi:10.1093/BIOINFORMATICS/BTI125

51. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol*. 1999;292(2):195-202. doi:10.1006/JMBI.1999.3091

52. Canutescu AA, Shelenkov AA, Dunbrack RL. A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci*. 2003;12(9):2001-2014.

doi:10.1110/PS.03154503

53. Russell RB, Barton GJ. Multiple protein sequence alignment from tertiary structure comparison: Assignment of global and residue confidence levels. *Proteins Struct Funct Bioinforma*. 1992;14(2):309-323. doi:10.1002/PROT.340140216

54. Platt EJ, Bilska M, Kozak SL, Kabat D, Montefiori DC. Evidence that Ecotropic Murine Leukemia Virus Contamination in TZM-bl Cells Does Not Affect the Outcome of Neutralizing Antibody Assays with Human Immunodeficiency Virus Type 1. *J Virol*. 2009;83(16):8289-8292. doi:10.1128/JVI.00709-09

55. Takeuchi Y, McClure MO, Pizzato M. Identification of Gammaretroviruses Constitutively Released from Cell Lines Used for Human Immunodeficiency Virus Research. *J Virol*. 2008;82(24):12585-12588. doi:10.1128/JVI.01726-08

56. Wei X, Decker JM, Liu H, et al. Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob Agents Chemother*. 2002;46(6):1896-1905. doi:10.1128/AAC.46.6.1896-1905.2002

57. Derdeyn CA, Decker JM, Sfakianos JN, et al. Sensitivity of Human Immunodeficiency Virus Type 1 to Fusion Inhibitors Targeted to the gp41 First Heptad Repeat Involves Distinct Regions of gp41 and Is Consistently Modulated by gp120 Interactions with the Coreceptor. *J Virol*. 2001;75(18):8605-8614. doi:10.1128/JVI.75.18.8605-8614.2001

58. Platt EJ, Wehrly K, Kuhmann SE, Chesebro B, Kabat D. Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1. *J Virol*. 1998;72(4):2855-2864. doi:10.1128/JVI.72.4.2855-2864.1998

59. Swanson CM, Puffer BA, Ahmad KM, Doms RW, Malim MH. Retroviral mRNA

nuclear export elements regulate protein function and virion assembly. *EMBO J*. 2004;23(13):2632-2640. doi:10.1038/SJ.EMBOJ.7600270

60.   Pasquinelli AE, Ernst RK, Lund E, et al. The constitutive transport element (CTE) of Mason-Pfizer monkey virus (MPMV) accesses a cellular mRNA export pathway. *EMBO J*. 1997;16(24):7500-7510. doi:10.1093/EMBOJ/16.24.7500

61.   Worobey M, Telfer P, Souquière S, et al. Island biogeography reveals the deep history of SIV. *Science (80-)*. 2010;329(5998):1487. doi:10.1126/SCIENCE.1193550

62.   Stern A, Doron-Faigenboim A, Erez E, Martz E, Bacharach E, Pupko T. Selecton 2007: Advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Res*. 2007;35(SUPPL.2). doi:10.1093/NAR/GKM382

63.   Doron-Faigenboim A, Stern A, Mayrose I, Bacharach E, Pupko T. Selection: A server for detecting evolutionary forces at a single amino-acid site. *Bioinformatics*. 2005;21(9):2101-2103. doi:10.1093/BIOINFORMATICS/BTI259

64.   Sawyer SL, Wu LI, Emerman M, Malik HS. Positive selection of primate TRIM5α identifies a critical species-specific retroviral restriction domain. *Proc Natl Acad Sci U S A*. 2005;102(8):2832-2837. doi:10.1073/PNAS.0409853102

65.   Yap MW, Nisole S, Stoye JP. A single amino acid change in the SPRY domain of human Trim5α leads to HIV-1 restriction. *Curr Biol*. 2005;15(1):73-78. doi:10.1016/J.CUB.2004.12.042

66.   Han GZ, Worobey M. An endogenous foamy-like viral element in the coelacanth genome. *PLoS Pathog*. 2012;8(6). doi:10.1371/JOURNAL.PPAT.1002790

67.   Han GZ. Extensive retroviral diversity in shark. *Retrovirology*. 2015;12(1). doi:10.1186/S12977-015-0158-4

68. Herniou E, Martin J, Miller K, Cook J, Wilkinson M, Tristem M. Retroviral diversity and distribution in vertebrates. *J Virol*. 1998;72(7):5955-5966. doi:10.1128/JVI.72.7.5955-5966.1998

69. Hayward A, Cornwallis CK, Jern P. Pan-vertebrate comparative genomics unmasks retrovirus macroevolution. *Proc Natl Acad Sci U S A*. 2015;112(2):464-469. doi:10.1073/PNAS.1414980112

70. OhAinle M, Kerns JA, Malik HS, Emerman M. Adaptive Evolution and Antiviral Activity of the Conserved Mammalian Cytidine Deaminase APOBEC3H . *J Virol*. 2006;80(8):3853-3862. doi:10.1128/JVI.80.8.3853-3862.2006

71. LaRue RS, Jónsson SR, Silverstein KAT, et al. The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. *BMC Mol Biol*. 2008;9. doi:10.1186/1471-2199-9-104

72. Münk C, Beck T, Zielonka J, et al. Functions, structure, and read-through alternative splicing of feline APOBEC3 genes. *Genome Biol*. 2008;9(3). doi:10.1186/GB-2008-9-3-R48

73. Bogerd HP, Cullen BR. Single-stranded RNA facilitates nucleocapsid: APOBEC3G complex formation. *RNA*. 2008;14(6):1228-1236. doi:10.1261/RNA.964708

74. Siegrist F, Ebeling M, Certa U. The small interferon-induced transmembrane genes and proteins. *J Interf Cytokine Res*. 2011;31(1):183-197. doi:10.1089/JIR.2010.0112

75. Staeheli P, Haller O. Interferon-induced Mx protein: a mediator of cellular resistance to influenza virus. *Interferon*. 1987;8:1-23. http://www.ncbi.nlm.nih.gov/pubmed/2445687

76. Perez-Caballero D, Zang T, Ebrahimi A, et al. Tetherin Inhibits HIV-1 Release by Directly Tethering Virions to Cells. *Cell*. 2009;139(3):499-511. doi:10.1016/J.CELL.2009.08.039

77. Sauter D. Counteraction of the multifunctional restriction factor tetherin. *Front Microbiol*. 2014;5(APR). doi:10.3389/FMICB.2014.00163

78. Zhang G, Li C, Li Q, et al. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science (80-)*. 2014;346(6215):1311-1320. doi:10.1126/SCIENCE.1251385

79. Magor KE, Miranzo Navarro D, Barber MRW, et al. Defense genes missing from the flight division. *Dev Comp Immunol*. 2013;41(3):377-388. doi:10.1016/J.DCI.2013.04.010

80. Oudshoorn D, van Boheemen S, Sánchez-Aparicio MT, Rajsbaum R, García-Sastre A, Versteeg GA. HERC6 is the main E3 ligase for global ISG15 conjugation in mouse cells. *PLoS One*. 2012;7(1). doi:10.1371/JOURNAL.PONE.0029870

81. Ketscher L, Basters A, Prinz M, Knobeloch KP. MHERC6 is the essential ISG15 E3 ligase in the murine system. *Biochem Biophys Res Commun*. 2012;417(1):135-140. doi:10.1016/J.BBRC.2011.11.071

82. Sawyer SL, Wu LI, Akey JM, Emerman M, Malik HS. High-frequency persistence of an impaired allele of the retroviral defense gene TRIM5α in humans. *Curr Biol*. 2006;16(1):95-100. doi:10.1016/J.CUB.2005.11.045

83. Hurles M. Gene duplication: The genomic trade in spare parts. *PLoS Biol*. 2004;2(7). doi:10.1371/JOURNAL.PBIO.0020206

84. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*. 2016;33(7):1870-1874.

doi:10.1093/MOLBEV/MSW054

85. Papadopoulos JS, Agarwala R. COBALT: Constraint-based alignment tool for multiple protein sequences. *Bioinformatics*. 2007;23(9):1073-1079. doi:10.1093/BIOINFORMATICS/BTM076

86. Li C, Wen A, Shen B, Lu J, Huang Y, Chang Y. FastCloning: A highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. *BMC Biotechnol*. 2011;11. doi:10.1186/1472-6750-11-92

87. Hochrainer K, Kroismayr R, Baranyi U, Binder BR, Lipp J. Highly homologous HERC proteins localize to endosomes and exhibit specific interactions with hPLIC and Nm23B. *Cell Mol Life Sci*. 2008;65(13):2105-2117. doi:10.1007/S00018-008-8148-5

88. Dudley DM, Gao Y, Nelson KN, et al. A novel yeast-based recombination method to clone and propagate diverse HIV-1 isolates. *Biotechniques*. 2009;46(6):458-467. doi:10.2144/000113119

Chapter 3

3 A single nucleotide polymorphism in the RLD domain of HERC6 impacts its antiviral activity against HIV-1

HERC5 is an interferon induced member of the small HERC family which inhibits replication of evolutionarily diverse viruses including human immunodeficiency virus type 1 (HIV). HERC6, a closely related, member of the same family does not inhibit HIV to the same extent. To better understand the discrepancy between the HIV restriction ability of these two proteins we performed structural analysis and discovered that a single nucleotide polymorphism, where the substitution of an arginine at position 10 for a glycine, can alter the theoretical folding of HERC6. This SNP enhanced its ability to inhibit the infectivity of HIV. While most of the human population possesses the defective form of HERC6, people of certain ethnicities possess the active form, which we postulate provides them with an enhanced antiviral defense. In fact, we discovered that possession of the active form of HERC6 is potentially associated with slower disease progression to AIDS in HIV-1-infected individuals in Uganda. Using western blotting, qRT-PCR, and infectivity assays we discovered that overexpression of HERC6-R10G results in less infectious viral particles. To determine the difference between the content of the supernatant released from cells overexpression HERC6 or HERC6-R10G we performed mass spectrometry. We found that the most down-regulated biological processes are viral transcription, viral gene expression, and protein translation-related processes. In our analysis we also noticed a reduction in CADM1 a surface protein which can be involved in viral attachment, however, neutralization of CADM1 did not result in reduced HIV particle infectivity. This study advances the knowledge of the small HERC family and more specifically HERC6, establishing it as an antiviral molecule.

3.1 Introduction

The steps in HIV replication cycle utilize cellular machinery to convert a single infections virion into a multitude of viral particles. Typically, this begins with the attachment of the infectious virions via the interaction of the envelope protein (gp120)

with the CD4 molecule and CCR5 or CXCR4 coreceptor on the surface of CD4+ T cells[1,2]. Fusion of the viral particle and cell membranes lead to the release of the capsid into the cytoplasm[1,3–5]. Reverse transcription and uncoating steps are initiated in the cytoplasm and aid in the transport of the HIV pre-integration complex to the nuclear pore where it can enter the nucleus and integrate into the genome[4,6–8]. Reverse transcription convert a positive sense single-stranded RNA genome into a double-stranded DNA genome which after entering the nucleus integrates into the host DNA[5,9,10]. Transcription of the HIV genome produces three forms of RNA: 1) a full-length unspliced mRNA that is incorporated into new virions and from which Gag and Gag-pol proteins are translate; 2) a singly spliced mRNA from which gp160, Vif, Vpu and Vpr proteins are translated; and 3) a doubly spliced mRNA from which Nef, Tat, and Rev proteins are translated[11–13]. Upon its translation Rev is imported into the nucleus and binds to the Rev Response Element (RRE) forming a complex with RanGTP and CRM1 (also known as exportin1 or XPO1) to shuttle the unspliced viral RNA through the nuclear pore complex into the cytoplasm where it disassembles and RanGTP is hydrolyzed into RanGDP[14–18]. The newly translated viral proteins are incorporated into new virions at the plasma membrane commencing with the multimerization of Gag[19]. Following multimerization the p6 domain of Gag recruits the endosomal sorting complexes required for transport (ESCRT) proteins which complete membrane cision[20,21]. A maturation step then occurs where the Gag and Gag-Pol polyproteins are cleaved by the HIV protease (PR) protein to produce infectious particles that can go on to infect neighbouring cells[19,22,23].

Sensing of viral infections stimulates an immediate response from the innate immune system triggering the production of interferons and interferon stimulated genes[24]. From the small HERC family of proteins HERC5 and HERC6 are the only two highly expressed interferon stimulated genes (ISGs). Similar to their other family members, HERC5 and HERC6 both consist of a N-Terminal (RCC1- like) RLD domain linked by a disordered spacer region to the HECT domain[25–27]. The RLD domain is a β-propeller structure consisting of 7 blades and the HECT domain is a C-lobe shape and contains an "active site" cystine which is involved in the ubiquitination or ISGylation of proteins[28,29].

By virtue of its HECT domain, HERC5 is the main E3 ligase for the conjugation of ubiquitin like proteins called interferon stimulated gene 15 (ISG15) to candidate substrates[27,28,30–33]. ISGylation can have several consequences for the modified protein although the most common is an alteration of function or localization.

Our laboratory and others have clearly established a role for HERC5, and more recently HERC6, in the inhibition of HIV-1 infection[14,29,34]. HERC5 inhibits HIV-1 replication through two distinct mechanisms. The first involves the RLD domain and inhibits the export of Rev-dependent (unspliced) HIV RNA from the nucleus to the cytoplasm by reducing intracellular levels of RanGTP and/or inhibiting the association of RanGTP with RanBP1[14]. The second mechanism of inhibition of HERC5 is dependent on its ISGylation activity and correlated with the ISGylation of Gag[35,36]. Although ISGylation of Gag results in decreased viral particle production and arrested budding, it is unclear whether this is due to the ability of ISG15 to replace the ubiquitin molecule needed to recruit the ESCRT protein necessary for successful budding, or whether it interferes with Gag multimerization which is essential for virion budding.

Recently, our laboratory discovered that HERC6 has a modest ability to inhibit HIV-1[34]. However, this ability was largely increased when the RLD domain of HERC6 was replaced with that of HERC5, suggesting that differences in the RLD domains between HERC5 and HERC6 impact anti-HIV activity. Moreover, swapping blade 1 of the RLD domain (first 100 amino acids) from HERC6 with that from HERC5 was sufficient to confer potent inhibition of HIV-1 particle production[34]. These experiments left many open questions about the contribution of the RLD domain to the anti-HIV activity of the interferon induced HERC proteins. Here we report the discovery of a single nucleotide polymorphism that significantly enhances the ability of HERC6 to inhibit production of infectious HIV-1 particles by a mechanism that potentially disrupts viral transcription and/or viral gene expression.

3.2 Results

3.2.1 Blade 1 of the HERC6 RLD is a determinant of antiviral activity

We previously showed that HERC6 exhibited modest antiviral activity towards HIV-1[37]. Swapping the HERC6 RLD with the HERC5 RLD, and more specifically the first 100 amino acids (blade 1) of HERC5, conferred potent anti-HIV-1 activity similar to that observed by wild type HERC5. Given the high degree of structural homology between the HERC5 and HERC6 RLDs, we hypothesized that structure-altering amino acid differences in the HERC6 RLD result in diminished antiviral activity.

To further characterize the contribution of the HERC6 RLD to its antiviral activity, we first generated a flag tagged RLD domain deletion mutant of HERC6 (Figure 3.1A). For comparison, we also generated a flag tagged RLD deletion mutant of HERC5. No significant differences in the expression of the mutant proteins were observed (Figure 3.1A). Consistent with our previous observations[14], HERC5 lacking its RLD increased both the amount of intracellular Gag protein and viral particles released from cells after 24 hours ($P<0.05$) (Figure 3.1A and B). Deletion of the HERC6 RLD domain resulted in a significant increase in viral particles released from cells. In contrast with HERC5, expression of the HERC6 RLD domain mutant did not result in an increase in intracellular Gag protein, suggesting that the HERC6 RLD is important for inhibiting a later stage of viral particle production.

To identify potential amino acid determinants of this antiviral activity, we compared predicted three-dimensional structures of the HERC5 and HERC6 RLDs. As shown in Figure 3.1C, the RLDs of both HERC5 and HERC6 assume a β-propeller structure consistent with that of the RCC1 family of proteins[34]. Interestingly, blade 1 of the HERC5 RLD is more disorders and extends away from the β-propeller structure whereas blade 1 of HERC6 folds back to interact with blade 2 before extending away from the β-propeller. A positively charged arginine residue at amino acid position 10 (R10) and a negatively charged glutamic acid residue at amino acid position 67 (E67) are located at

this point of interaction in the HERC6 RLD, likely contributing to this interaction (Figure 3.1D). To assess the predicted structural significance of these specific residues, we generated mutant HERC6 sequences containing uncharged alanine mutations at positions R10 (H6-R10A), E67 (H6-E67A) or both R10 and E67 (H6-R10A/E67A) using site-directed mutagenesis. Molecular modeling of these RLD mutants containing the R10A, E67A or R10A/E67A mutations predicted that blade 1 would assume a more disordered structure (without interaction with blade 2), similar to blade 1 of HERC5 (Figure 3.1C).

We then tested the functional significance of the HERC6 R10A, E67A or R10A/E67A mutants in inhibiting HIV-1 particle production. For comparison, we also included HERC5:HERC6 RLD and blade 1 swaps where the HERC5 RLD (H6:H5RLD) or the first 100 amino acids of the HERC5 RLD (H6:H5 RLD$_{1-100}$) replaced the corresponding region in HERC6. As expected, HERC5 and to a much lesser extent HERC6, inhibited intracellular production of HIV-1 Gag proteins and inhibited production of viral particles (Figure 3.1F and 3.1G). Despite lower levels of intracellular expression, H6:H5RLD, and especially H6:H5RLD$_{1-100}$ exhibited increased inhibition of intracellular Gag protein production and viral particle production. H6-R10A, H6-E67A and H6-R10A/E67A modestly increased the ability of HERC6 to inhibit intracellular Gag protein levels and viral particle production after 24 hours. This data suggests that blade 1 of the HERC6 RLD and, potentially, residues R10 and E67 are determinants of HERC6 anti-HIV-1 activity.
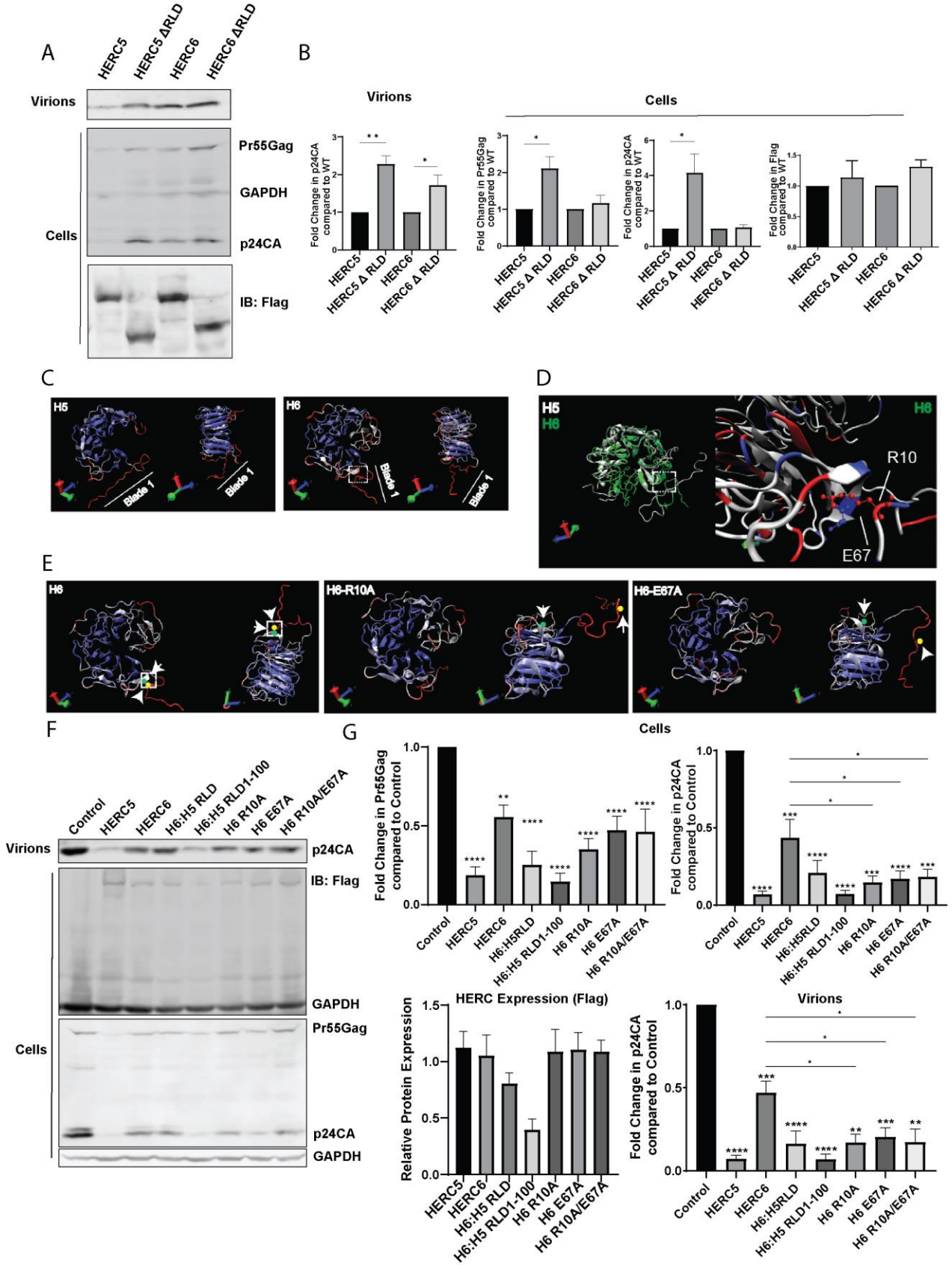
Figure 3.1 The RLD domain of HERC5 and HERC6 is a determinant of antiviral activity

A) 293T cells were co-transfected with pR9 (HIV-1) and either, pHERC5, pHERC6, pHERC5ΔRLD, or pHERC6ΔRLD. Twenty-four hours post-transfection, virions released into the supernatant and total cell lysates were subjected to quantitative Western blot analysis using anti-flag, anti-p24CA, or anti-GAPDH as loading controls. **B)** Average densitometric quantification (+/- S.E.M.) quantification of the p24CA (virions) and Pr55Gag, p24CA, and flag bands (cells) from four independent experiments is shown. Statistical significance was measured using Students t-test and the ΔRLD proteins were compared with wildtype. **C)** Two different views of the three-dimensional representation of the predicted structural data of human HERC5 and HERC6 RLDs. Each amino acid is colored per the degree of evolutionary conservation with other HERC homologs. **D) Left**, structural alignment of the RLDs of HERC5 (white) and HERC6 (green). **D) Right**, close-up image showing the interaction of the positively charged R10 residue (red ball and stick) and negatively charged E67 residue (blue ball and stick) of the human HERC6 RCC1-like domain. **E)** Two different views of the three-dimensional representation of the predicted structural data of human HERC6 RLDs containing either R10A (HERC6-R10A) or E67A (HERC6-E67A) mutations. Yellow and green circles denote the locations of the R10 and E67 mutated residues respectively. **F)** 293T cells were co-transfected with pR9 (HIV-1) and either empty vector, pHERC5, pHERC6, pHERC6:HERC5 RLD, HERC6:HERC5 RLD1-100, pHERC6:R10A, pHERC6:E67A or pHERC6:R10A/E67A. Twenty-four hours post-transfection, viral particles released into the supernatant and total cell lysates were subjected to quantitative Western blot analysis using anti-p24CA, anti-Flag and anti-β actin as a loading control. **G)** Average densitometric quantification (+/- S.E.M.) of p24CA (virions) and Pr55Gag, p24CA, and HERC expression (Flag) bands (cells) from four independent experiments is shown. Statistical significance was measured using

One-way ANOVA, Dunnett's multiple comparisons test with either control (above the bar) or wildtype HERC6 (indicated by a line). $P > 0.05$, not significant (ns); * $P < 0.05$, ** $P < 0.01$ *** $P < 0.001$, **** $P < 0.0001$

3.2.2 SNP rs111670008 (R10G) is associated with slower disease progression

To determine if residues R10 and E67 of HERC6 have potential clinical relevance for HIV-1 infection, we first analyzed the human SNP database to determine if polymorphisms have been detected at these residues in the human population. Indeed, SNPs are present in the human population at residues 10 and 67 of HERC6 (rs111670008 and rs775530962 respectively). Since rs111670008 contained more information pertaining to its prevalence in the human population, we asked if rs111670008 is associated with HIV-1 disease progression.

The rs111670008 SNP contains an alanine (contig reference) to guanine nucleotide polymorphism, resulting in an arginine to glycine amino acid change. Analysis of the global allele and genotype frequencies showed that the minor allele frequency of the G nucleotide allele is 2.5% with 13.6% heterozygosity with the A allele. The G allele is rare in the African, East Asian, and South Asian populations and most prevalent in the American and European populations (Figure 3.2A). In the uninfected African population, 99.2% of individuals are homozygous for the A allele, resulting in arginine at position 10; whereas 0% are homozygous for the G allele, resulting in glycine. 0.8% of the population is heterozygous for the A and G alleles. After sequencing the genomic DNA region surrounding amino acid residue 10, we calculated the frequency of each allele from 29 HIV-1-infected individuals in a Ugandan cohort who were classified as either rapid, normal, or slow progressors (Figure 3.2B). Sequence analysis of the 29 patient samples yielded no significant difference in the homozygous (A/A) or heterozygous (A/G) allele frequencies between the uninfected population and rapid progressors; however, there was a significant association between the heterozygous (A/G) individuals and slower disease progression to AIDS (dominant model, odds ratio = 157.4, 95% CI = 35.9 to 690.1, P <

0.0001) (Figure 3.2B). This data identifies a potential association of the HERC6 SNP rs111670008 with slower disease progression to AIDS in infected individuals in Uganda.

Global populations

| | A |
| | G |

	All	African	American	East Asian	European	South Asian
Allele frequency	A: 97.5 % G: 2.5 %	A: 100 % G: 0 %	A: 95 % G: 5 %	A: 100% G: 0 %	A: 93 % G: 7 %	A: 99 % G: 1 %
Genotype frequency	A|A: 0.951 (2382) A|G: 0.048 (119) G|G: 0.001 (3)	A|A: 0.992 (656) A|G: 0.008 (5)	A|A: 0.905 (314) A|G: 0.089 (31) G|G: 0.006 (2)	A|A: 1.000 (504)	A|A: 0.861 (433) A|G: 0.137 (69) G|G: 0.002 (1)	A|A: 0.971 (475) A|G: 0.029 (14)

African sub-populations

	African Caribbean in Barbados	African Ancestry in Southwest US	Esan in Nigeria	Luhya Webuye Kenya	Mandinka in the Gambia	Mende in Sierra Leone	Yoruba in Ibadan Nigeria
Allele frequency	A: 99 % G: 1 %	A: 98 % G: 2 %	A: 100 % G: 0 %	A: 99 % G: 1 %	A: 100 % G: 0 %	A: 100 % G: 0 %	A: 100 % G: 0 %
Genotype frequency	A|A: 0.990 (95) A|G: 0.010 (1)	A|A: 0.951 (58) A|G: 0.049 (3)	A|A: 1.000 (99)	A|A: 0.990 (98) A|G: 0.010 (1)	A|A: 1.000 (113)	A|A: 1.000 (85)	A|A: 1.000 (108)

Allele frequency

Allele	Healthy African Population	Rapid Progressors	Normal Progressors	Slow Progressors
R/R	99.2% (656)	85.7% (6)	81.8% (9) **	45.5% (5) ****
R/G	0.8% (5)	14.3% (1)	18.2% (2) **	54.5% (6) ****
G/G	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)

dominant model, ** P = 0.0049; odds ratio = 29.2, 95% CI = 5.0-170.7;
**** P < 0.0001; odds ratio = 157.4, 95% CI = 35.9-690.1

Figure 3.2 HERC6 SNP rs111670008 (R10G) is associated with slower disease progression in a Ugandan cohort of HIV-1-infected individuals.

A) Allele and genotype frequencies of the R10G HERC6 SNP rs111670008 in the global population and African sup-populations. Data was obtained from the 1000 Genomes database. **B)** Genomic DNA encompassing the HERC6 SNP rs111670008 was amplified from genomic DNA isolated from 29 HIV-1 infected

individuals categorized as fast, normal, or slow progressors. The amplified DNA was sequenced to identify the percentages of patients homozygous or heterozygous for the A or G allele at amino acid position 10 of HERC6 and compared to that of the normal healthy African population.

3.2.3 Arginine 10 of blade 1 in the HERC6 RLD is a critical determinant of anti-HIV-1 activity

To further understand the contribution of the rs111670008 SNP (R10G) to the antiviral activity of HERC6, we generated a HERC6 construct bearing the R10G polymorphism and compared its ability to inhibit HIV-1 particle production to that of wild type HERC5 and HERC6. 293T cells were co-transfected with plasmids carrying full-length infectious HIV-1 R9 with or without plasmids carrying flag-tagged HERC5, HERC6 or HERC6-R10G. Twenty-four hours after expression, cells and supernatant were collected and subjected to Western blot analysis. As expected, cells expressing HERC5 and to a lesser extent HERC6 exhibited a significant reduction in intracellular Gag protein and release of viral particles from cells (Figure 3.3A and B). Cells expressing HERC6-R10G exhibited a similar reduction in intracellular Gag protein and release of viral particles from cells as observed with HERC6.

Next, we sought to determine if the R10G SNP impacts the infectivity of virions produced from cells. HOS-CD4-CXCR4 cells, which support robust HIV-1 replication, were co-transfected with pR9 (HIV-1) and either empty vector, flag-tagged pHERC5, pHERC6 or pHERC6-R10G. Seventy-two hours post-transfection, supernatant containing released viral particles were used to infect the luciferase reporter cell line HeLa TZM-bl to measure the level of infectious virions. As previously observed, HERC5 significantly reduced the infectivity of released HIV-1 virions (Figure 3.3A). In contrast, no significant reduction in infectivity was observed for HERC6. Intriguingly, cells expressing HERC6-R10G significantly inhibited the release of infectious virions to a similar degree as HERC5 (Figure 3.3C).

To better understand how HERC6-R10G inhibits the infectivity of HIV-1 virions, we asked if viral RNA levels were affected by HERC6-R10G expression. HIV-1 RNA was isolated from the supernatant of cells co-expressing HERC5, HERC6 or HERC6-R10G and measured using quantitative RT-PCR (qRT-PCR) (Figure 3.3D). As expected, HERC5 significantly reduced the amount of detectable viral RNA in the supernatant. Conversely, no significant reduction was observed with HERC6 or HERC6-R10G. We previously showed that blade 1 of the HERC5 RLD plays a critical role in blocking the production of infectious HIV-1 particles by inhibiting REV-dependent nuclear export of unspliced HIV-1 RNA, which are exported via the CRM1-dependent nuclear export pathway[14]. To determine if the R10G substitution in blade 1 of HERC6 conferred a similar ability, we compared the abilities of HERC5, HERC6 and HERC6-R10G to inhibit the export of unspliced HIV-1 RNA from the nucleus using qRT-PCR. As previously observed[14], HERC5 significantly inhibited the export of unspliced RNA from the nucleus (Figure 3.3E). In contrast, no significant inhibition of nuclear export was observed for HERC6 or HERC6-R10G. As a control, none of the HERC constructs inhibited nuclear export of REV-independent RNA, which are exported via the NXF1/TAP-dependent nuclear export pathway. Together, these data suggest that R10G inhibits the production of infectious HIV-1 particles by a mechanism not involving a reduction in virion RNA levels or by inhibiting nuclear export of HIV-1 RNA.
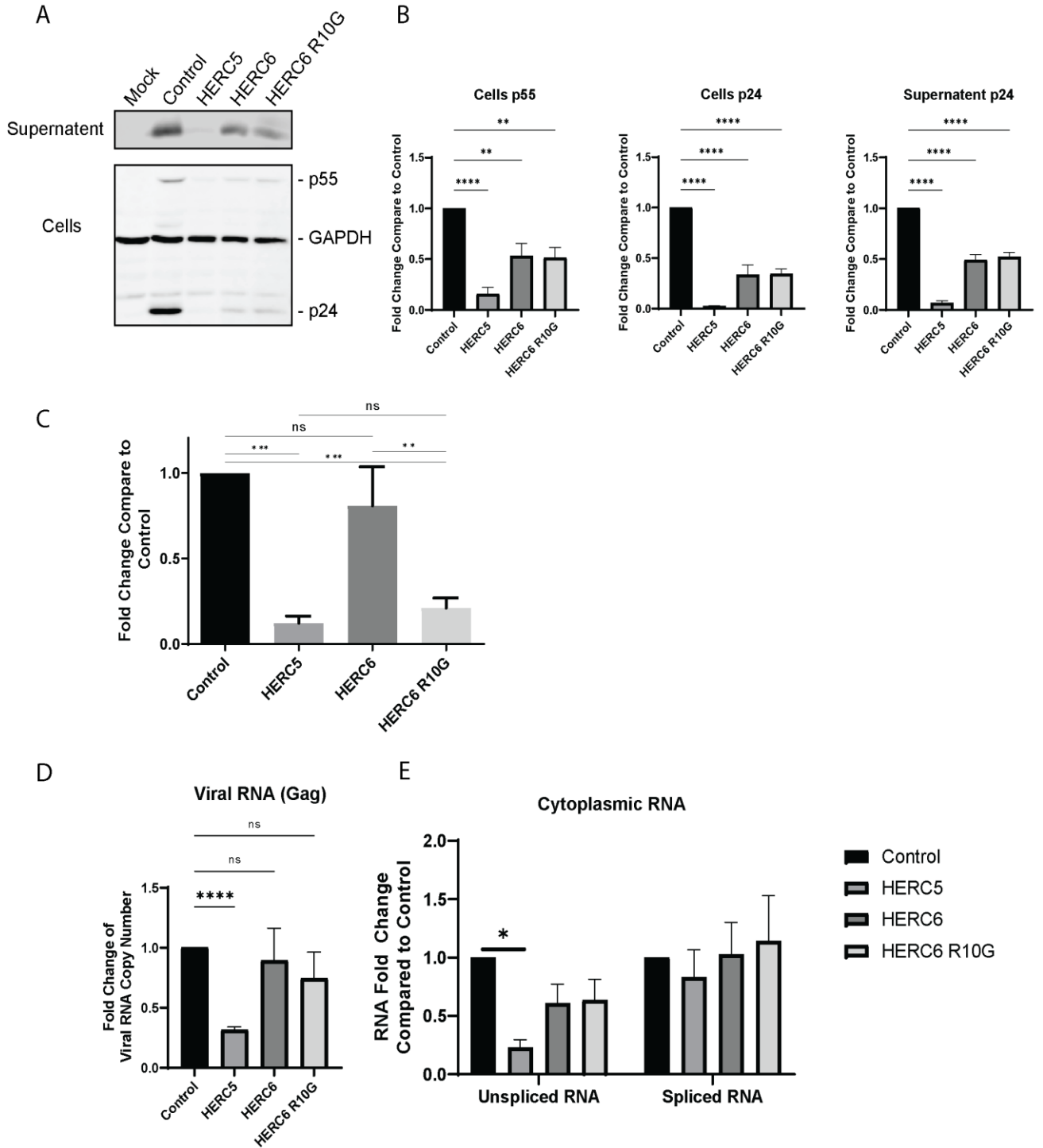
Figure 3.3 HERC6-R10G reduces the infectivity of viral particles.

A) 293T cells were co-transfected with pR9 (HIV-1) and either empty vector, flag-tagged pHERC5, pHERC6 or pHERC6-R10G. Twenty-four hours post-transfection, virions released into the supernatant and total cell lysates were subjected to quantitative Western blot analysis using anti-p24CA or anti-GAPDH as a loading control. **B)** Densitometric quantification of the p24CA (virions in the supernatant) and Pr55Gag and p24CA bands (cells) from four independent experiments is shown. Statistical significance is based on One-way ANOVA, Dunnett's multiple comparisons test with either control or wildtype HERC6. $P > 0.05$, not significant (ns); * $P < 0.05$, ** $P < 0.01$ *** $P < 0.001$, **** $P < 0.0001$. **C)** HOS-CD4-CXCR4 cells, which support robust HIV-1 replication, were co-transfected with pR9 (HIV-1) and either empty vector, flag-tagged pHERC5, pHERC6 or pHERC6-R10G. Seventy-two hours post-transfection, supernatant containing released viral particles were used to infect the luciferase reporter cell line HeLa TZM-bl to measure the level of infectious virions. **D)** 293T cells were co-transfected with pR9 (HIV-1) and either empty vector, flag-tagged pHERC5, pHERC6 or pHERC6-R10G. Twenty-four hours post-transfection, viral RNA was isolated from viral particles released into the supernatant and subjected to qRT-PCR. Statistical significance was measured using one-way ANOVA with Tukey's multiple comparisons test. **E)** 293T cells were co-transfected with full-length replication-competent HIV-1 (pR9) and peGFP (transfection control) and either empty vector (control), pHERC5, pHERC6 or pHERC6-R10G. Forty-eight hours after transfection, total RNA was extracted, and reverse transcribed into cDNA from whole cell lysates or from the cytoplasmic fraction only. Quantitative PCR was performed on each fraction using primers specific to unspliced HIV-1 genomic RNA (Gag), fully spliced RNA (Rev), or GAPDH (loading control). The proportion of unspliced or fully spliced HIV-1 RNA in the cytoplasmic fraction compared to total amount of HIV-1 RNA (whole cell) was determined for control cells and cells expressing HERC. Fold-change in copy number relative to control cells is shown for each fraction. Data shown represents the average (+/− SEM) from four independent experiments.

3.2.4 HERC6-R10G alters the protein content of virus-containing particles released from cells

To better understand the contribution of the R10G polymorphism to the antiviral activity of HERC6, we performed mass spectrometry analysis on virus-containing particles released from control cells or cells expressing HERC6 or HERC6-R10G. The relative expression levels of proteins contained in the virus-containing particles were determined and the most significantly identified proteins were compared to the group mean to obtain relative fold changes in expression levels (Figure 3.4A). We used a significant cut-off of 10, a down-regulation cut-off of 0.8-fold change and an up-regulation cut-off of 1.2-fold change from HERC6. As observed in the heatmap in Figure 3.4A, several cellular proteins are significantly upregulated and downregulated in the particles released into the supernatant of control cells. In comparison, particles released from cells expressing HERC6 exhibited contrasting expression of the majority of these proteins (e.g., red versus blue or blue versus red), while some protein levels remained unchanged (e.g., orange versus orange or blue versus blue).

Interestingly, cells expressing HERC6-R10G exhibited strong contrasting protein expression levels when compared to HERC6. The most down-regulated proteins were Ribosomal protein L6 , L35, and L18 (RPL6, RPL35, RPL18), ABRA C-Terminal-Like Protein (ABRACL), Nuclear Casein Kinase And Cyclin Dependent Kinase Substrate 1 (NUKS1), HEAT Repeat Containing 6 (HEATR6), Heterogeneous Nuclear Ribonucleoprotein D (HNRNPD), Histone-lysine N-trimethyltransferase SMYD5, Alpha 2-HS Glycoprotein (AHSG), H1.2 Linker Histone(H1-2), and Cell Adhesion Molecule 1 (CADM1). Proteins that are over-represented consist of Calmodulin 1, 2 and 3 (CALM1, CALM2, CALM3), Albumin (ALB), ubiquinol-cytochrome c reductase core protein 2 (UQCRC2), Prohibitin 2 (PHB2), Voltage Dependent Anion Channel 1 (VDAC1), Elongator Acetyltransferase Complex Subunit 2 (ELP2), Heterogeneous Nuclear Ribonucleoprotein U (HNRNPU), Stomatin-Like 2(STOML2), Ribophorin I (RPN1), DEAD-Box Helicase 42 (DDX42), Heat Shock Protein Family A (Hsp70) Member 9 (HSPA9), Nuclear Receptor Binding Protein 1 (NRBP1), ATP Synthase Peripheral Stalk

Subunit D (ATP5PD) and Destrin (DSTN). Interestingly, many of the proteins identified are ribosomal proteins or related to ribosomal processes. Proteins such as ABRACL and DSTN are involved in the regulation of actin filamin processes while PHB2, UQCRC2, and ATP5PD are involved in mitochondrial processing.

To better understand the biological processes most effected by HERC6-R10G, we performed gene ontology (GO) analysis of the proteins most under-represented and over-represented compared to HERC6. Calcium processes are among the most up-regulated by HERC6-R10G expression such as detection of calcium ions, detection of chemical stimulus, mitochondrial calcium ion transmembrane transport, negative regulation of calcium ion transmembrane transport, and cytosolic calcium ion transport (Figure 3.4C). The most down-regulated biological processes are viral transcription, viral gene expression, followed by translation-related processes such as co-translational protein targeting to the membrane, protein targeting to the ER, establishment of protein localization to ER, and protein targeting (Figure 3.4D). These results suggest that compared to HERC6, HERC6-R10G disrupts viral gene expression processes.

Given our previous observations that HERC6 expression exhibited a modest effect in inhibiting intracellular Gag levels and production of virus-containing particles compared to the control, we were interested in identifying proteins whose expression levels differed substantially in the HERC6-R10A samples but were relatively similar between the control and HERC6 samples (Figure 3.4A, box 1 and box 2). Figure 3.4A box 1 highlights several protein candidates that are significantly down-regulated in HERC6-R10G compared to the control/HERC6. GO analysis revealed that these proteins are mostly involved in regulation of cytoplasmic translation. Figure 3.4A box2 highlights several protein candidates that are significantly up-regulated in HERC6-R10G compared to the control/HERC6. Notably among this list of proteins is DDX42. DDX42 was recently shown to potently inhibit HIV-1 and other positive-sense RNA viruses by binding viral RNA[38]. Taken together, these data show that HERC6 R10G expression

alters the protein composition of HIV-1-containing particles released from cells, leading to an increase in proteins known to disrupt virus replication.

A Proteins Expression by Sample

B Differential Expression of Proteins in HERC6 R10G vs. HERC6

C GO biological processes upregulated in HERC6 R10G vs. HERC6 samples

D GO biological processes downregulated in HERC6 R10G vs. HERC6 samples
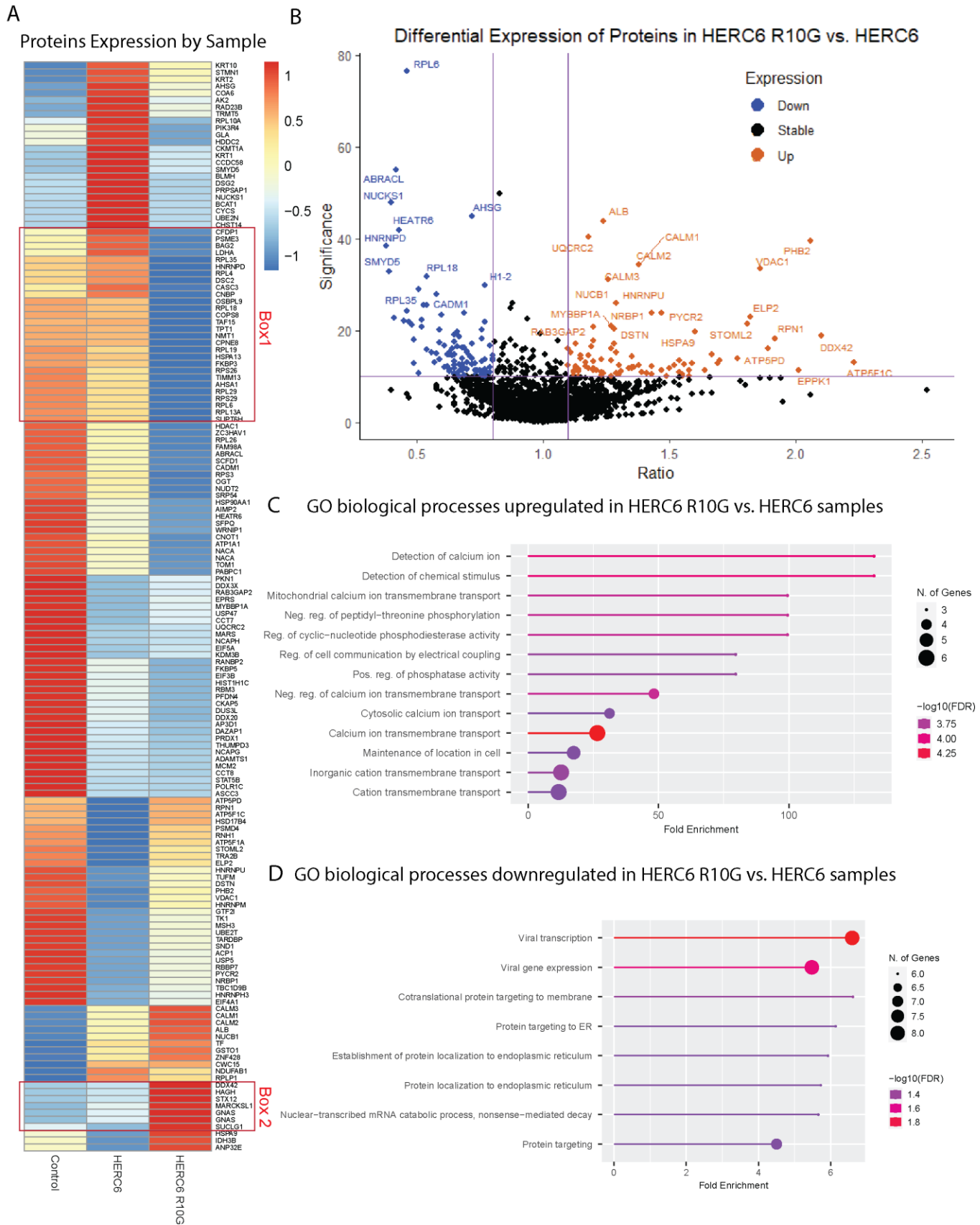
Figure 3.4 Mass Spectrometry analysis of virus containing particles.

293T cells were transfected with pR9 containing plasmid and empty vector (control), pHERC6 or pHERC6-R10G. Forty-eight hours post transfection the samples were processed and analyzed using LC-MS/MS (see methods) **A)** Protein expression profile of the top 150 differentially expressed genes based on z-scores. Box 1 highlights genes which are detected in HERC6 and control samples but not in HERC6-R10G. Box 2 highlights genes which are highly expressed in HERC6-R10G samples but not in HERC6 and control samples. **B)** Volcano plot showing significance and ratio of proteins comparing their abundance in HERC6-R10G to HERC6 samples. Significance cutoff of 13 was used corresponding to a p-value of 0.05. Proteins beyond a ratio of 1.2 and below 0.8 are highlighted in red and blue respectively. **C)** and **D)** the proteins identified in the volcano plot in B were entered into ShinyGo to obtain significant (FDR ≥ 0.05) upregulated **(C)** and downregulated **(D)** biological processes.

3.2.5 CADM1 neutralization does not affect viral particle infectivity

The proteomic data suggests that HERC6-R10G disrupts viral gene expression and transcriptional processes. To determine if HERC6-R10G produces altered viral protein expression we examined the viral protein detected in the mass spec samples mentioned above (Figure 3.5A). There was a large difference in the amount of Pol and Env when comparing HERC6 and HERC6-R10G to control samples and a slight decrease in Gag levels. The amount of Nef remained relatively unchanged in all samples and Rev was only detectable in the control sample. Additionally, we have not observed a difference in the RNA content within the viral particles released from cells transfected with HERC6-R10G or HERC6 (Figure 3.3 D).

To further understand if there is a single cellular protein responsible for the difference observed in the infectivity of particles produced in HERC6-R10G expressing cells vs HERC6 expressing cells we looked to see if attachment molecules were differentially

expressed. Cell adhesion molecules such as ICAM-1 are actively incorporated into the viral membrane to facilitate the interaction between the virion and the target cells[39–41]. We noticed that CADM1 is significantly reduced in the HERC6-R10G sample compared to the HERC6 sample. CADM1 is a cell adhesion molecule mostly known for its role in the attachment of human mast cells to fibroblasts, human airway smooth muscle cells and neurons[42–44]. To determine if this effects the infectivity of the viral particles we performed a neutralization assay with a monoclonal antibody against the ectodomain (9D2) of CADM1 which is known to inhibit its adhesive function[42,45,46]. The HIV-1 neutralizing antibody VRC01 was used as a positive control[47]. We found that neutralization of the CADM1 molecule on the surface of viral particles does not influence the infectivity of the particles. This was the case for both the R9 and the NL43 virus particles. Here, we determined that the effect of the R10G SNP on inhibiting particle infectivity is not due to its modulation of the attachment molecule CADM1.
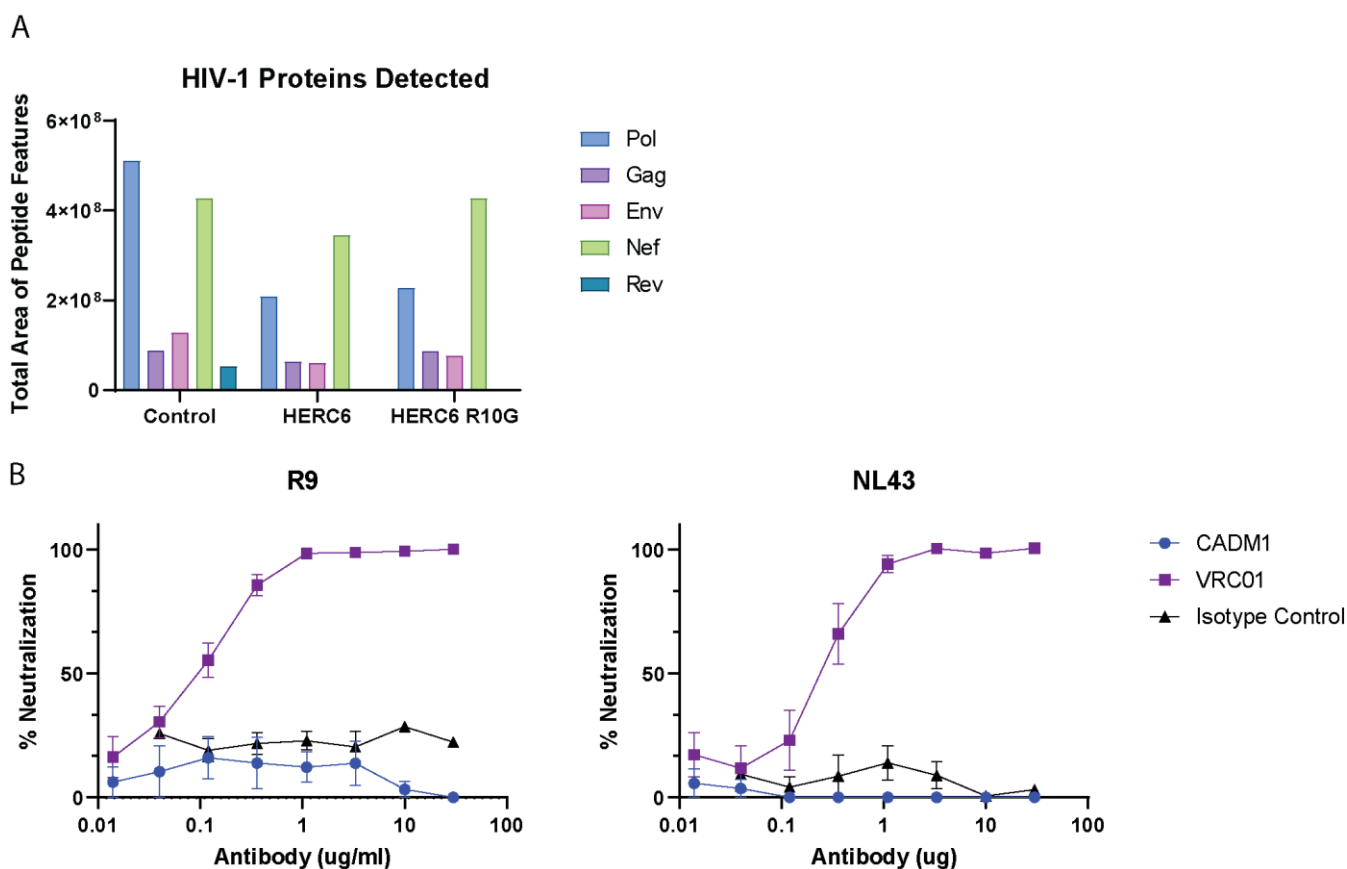
Figure 3.5 Anti-CADM1 antibody is not able to neutralize HIV-1 R9 or NL43 strains

A) HIV-1 proteins from the mass spectrometry results of figure 4 were filtered out and their abundance in each condition examined. The graph shows the sum of the area for each HIV-1 protein detected. **B)** HeLa TZM-bl cell neutralization assay using CADM1 neutralizing antibody (9D2), isotype control, or VRC01 antibody. Left panel is neutralization percent of R9 virus and right panel is neutralization percent of NL43 virus stocks. Results are representative of three independent experiments

3.3 Discussion

HIV-1 is one of the deadliest, most widespread viruses on earth. Although the advancement of medical treatments significantly extends the life span of those infected, a true cure or even a preventative vaccine has yet to be found. Here, we report the discovery of a novel SNP in the interferon induced HERC6 molecule which contributes

to its antiviral activity against HIV-1. The HERC6-R10G heterozygous allele is over-represented in HIV-1 slow progressors in a Uganda cohort. Although this SNP does not have an effect on the particle production or intracellular levels of Gag protein (Figure 3.3 A and B), on the nuclear export of viral RNA (both spiced and unspliced) (Figure 3.3E), or on the incorporation of viral RNA into viral particles (Figure 3.3.D), it does have a significant effect on the infectivity of the virions produced in cells expressing it. Using mass spectrometry, we examined the differentially expressed proteins in particles produced in cells expressing either wildtype HERC6 or HERC6-R10G. We found that several proteins such as ribosomal protein 6, 18 and 35 as well as CADM1 are significantly lower in particles produced in HERC6-R10G expressing cells. More importantly, proteins related to the GO biological processes of viral transcription and viral gene expression are significantly under-represented in these virions compared to those produced in cells expressing wildtype HERC6 (Figure 3.4D). Viral membrane composition is important for HIV-1 infectivity. For example, ICAM-1 incorporation into the viral envelope promotes viral attachment through the interaction with LFA-1 on the membrane of CD4+ T cells[48]. Since the cellular adhesion molecule CADM1 was under-represented in the HERC6-R10G samples, we performed a neutralization assay to determine if antibody neutralization of CADM1 on the surface of virions would reduce particle infectivity. Unfortunately, the CADM1 neutralizing antibody did not perform better than the isotype control in our assay, although previous studies have found that the 9D2 IgY mAb against CADM1 inhibited the binding of mast cells to human airway smooth muscle cells[44] and sensory neurons[49]. Conversely, binding of CADM1 to the extracellular matrix of human airway smooth muscle cells was affected by expression levels of CADM1 and not by antibody binding[42]. It is possible that a similar mechanism might be affecting the binding mediated by CADM1 of the HIV particles to target cells, whereby the reduced expression of CADM1 on virions will inhibit their binding to target cells but antibody neutralization does not hinder it. Future research is needed to determine if CADM1 is a true HIV adhesion molecule which affects infectivity.

A small subset of people are resistant to HIV infection and from the beginning of the HIV-1 epidemic researchers have searched for host genetic factors which contribute to protection from the disease with the hope of capitalizing on these discoveries to deliver a cure. A meta-analysis of 6315 individuals determined that approximately 25% of the variation in viremia setpoint is due to host genetic variation[50,51]. Most of this variation is attributed to the contribution of HLA and CCR5 mutations, leaving approximately 5.5% of variance associated with other regions of the genome[50,51]. Expression of HLA-B*57, B*27, B*58:01, B*51, B*13, and B*81:01 are amongst the most protective correlates while, HLA-B*58:02 and HLA-B*35Px alleles are associated with more rapid disease progression[50]. CCR5, specifically the CCR5Δ32, is the only genetic variant which has been proven to definitively protect from HIV infection[52–55]. So far, three separate patients have been cured of HIV by receiving a bone marrow or stem cell transplant from individuals with the CCR5Δ32 mutation and have remained HIV-free without the aid of antiretroviral therapies (ART) for more than 10 years[52]. In two other cases, the HIV positive recipients of a bone marrow transplant (not containing the CCR5Δ32 mutation) rebounded approximately 32 weeks after the cessation of ART[54]. These results highlight the power of genetic variation on controlling HIV-1 disease progression.

Other immune sensing molecules such as Toll Like Receptors (TLR's) and restriction factor SNPs have also been shown to influence disease progression. For example, a SNP on TLR4 (D299G) is associated with higher HIV-1 infection[56–58] while on TLR3, P554S and L412F are thought to be a protective alleles[59]. TLR7 Q11L, TLR9 G1174A and G1635A are associated with accelerated HIV-1 disease progression[60–62]. APOBEC3G houses a protective SNP (rs35228531) which reduced the risk of HIV infection for carriers of the minor allele T[63]. In elite controllers, variation in TRIM5α (rs10838525) may contribute to viral suppression while rapid progression was found in those harboring a mutation in cyclophilin A (rs6850)[64]. Additionally, 136Q polymorphism of TRIM5α along with the −2GG genotype was associated with faster disease progression (The Effect of Trim5 Polymorphisms on the Clinical Course of HIV-1 Infection). In tetherin (BST-2) several SNPs (rs3217318, rs10415893, rs145303329, rs12609479) have been reported to

interfere with the expression and function of the potent antiviral molecule[65,66]. On the other hand HIV-1 infected individuals who possess the BST-2 rs12609479-GG/rs113189798-AG alleles have higher T-cell counts and this mutation is enriched in HIV-1 elite controllers[67] and the rs9576CA allele is associated with higher CD4 counts and lower viral loads[68].

With the data obtained in this chapter, we speculate that HERC6-R10G SNP affects the proteins packaged into the viral particles and released from cells in the form of extracellular vesicles. Specifically, proteins involved in the biological processes of viral transcription and gene expression are significantly reduced in these virions, and while these processes are not affected in the producer cells, they seem to have a deleterious effect in the establishment of infection in the target cells. Several proteins have been previously identified as influencing the infectivity of viral particles based on their incorporation into the virion[48,69]. Incorporation of host protein into viral particles can occur actively, through direct interactions with viral proteins such as the incorporation of LEDGF through its interaction with the integrase portion of Gag-Pol[70], or Heat shock protein (Hsp) 70 through its interactions with Gag[71]; or passively, due to their proximity to the plasma membrane such as tetraspanins[69]. Omitting these proteins from the viral particle often leads to defective virions and abortive infections[48,69].

Similar to ICAM-1, integrin α4β7 is incorporated into the viral envelope to facilitate viral homing to the gut and preclusion of these attachment molecules reduces the infectivity of viral particles[72]. On the other hand, incorporation of host SERINC3 and SERINC5 reduces the infectivity of viral particles, although they are counteracted by HIV-1 Nef [73,74]. The packaging of viral proteins into extracellular vesicles can also have an effect on the infectivity of the viral particles. For example, packaging of Nef into exosomes suppresses the immune response of neighboring cells and makes them more permissive to infection[75]. Liu et al[76] found that virions produced in cells over-expressing PSGL-1 were significantly less infectious. PSGL-1 is a lymphocyte and myeloid cell specific factor which is induced by IFNγ during HIV-1 infection of primary CD4+ T cells. This study

did not elucidate the exact mechanism by which PSGL-1 reduces particle infectivity, but it illustrates that cellular factors can impact the infectivity of the produced virions in unknown and unexpected ways.

Our study had several limitations. We were only able to repeat the mass spec experiment with one replicate and due to the nature of the isolation technique used we were not able to separate exosomes secreted by the cells from the viral particles. However, our results are consistent with other studies which examined HIV-1 particle composition[77,78]. Consistently, we identified heat shock proteins (e.g. HSP70, HSP90, and HSP27), actin and actin binding proteins (e.g., ACTN1, moesin, and ezrin) and surface proteins such as HLA-A which have been previously described to be actively incorporated into viral particles[69]. These studies utilized different cell lines (e.g., Sup-T1, THP-1, CEM, and H9) in their experiments so some cell type-specific variation can be expected, however, the overall results are similar. While not conclusive, our results provide the basis for future studies to build on and better elucidate the relationship between the HERC6-R10G SNP, the infectivity of viral particle and consequences on disease progression in a larger cohort.

Our finding that the R10G allele is over-represented in individuals categorized as slow progressors in conjunction with decreased infectivity for virions formed in cells expressing HERC6-R10G suggest that this SNP has a significant role in innate resistance to HIV-1 infection. Although further studies are needed to better understand the mechanism by which this restriction occurs, here, we outline the discovery and initial experiments used to determine the contribution of the R10G SNP to the antiviral activity of HERC6. Further understanding of the multitude of ways the HIV virus interacts with the host restriction factors will pave the way for the discovery of novel antiviral therapeutics to aid in the treatment and/or prevention of AIDS.

3.4   Materials and Methods

3.4.1 Cell Lines

293T cells were obtained from American Type Culture Collection. TZM-bl cells were obtained from NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID. Cells were maintained in standard growth medium (Dulbecco's Modified Eagle's Medium (DMEM)), supplemented with 10% heat-inactivated Fetal Bovine Serum (FBS), 100 U/ml Penicillin and 100 µg/ml Streptomycin) at 37°C with 5% CO2.

3.4.2 Plasmids, transfections, antibodies, and Western blotting

Plasmids: The promoterless empty vector plasmid pGL3, p3xFLAG-CMV10 and peGFP were obtained from Promega, Sigma and Clontech respectively. Coding regions of HERC5 and HERC6 were obtained from the following sources: HERC6 (NM_017912.3), and HERC5 (NP_057407.2). Details pertaining to the construction of plasmids encoding the various flag-tagged HERC5, and HERC6 constructs can be found in the methods section of Chapter 2 and Table 3. All constructs were verified by sequencing. Transfections were performed using Lipofectamine 2000 (Invitrogen) per manufacturer's instructions unless otherwise stated. Co-transfections of HERC plasmids with pR9 were performed at a ratio of 10:1 respectively, unless otherwise noted. Standard Western blot analyses were performed as previously described[34]. Densitometric analysis was performed using ImageJ 1.43u 64-bit version software (NIH, USA).

To generate pH6:R10A, pH6:E67A, pH6:R10A/E67A, and pR10G site-directed mutagenesis was performed on pHERC6 using the primers in Table 4 rows 1 to 4.

Antibodies: The following reagents were obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: HIV-1 p24 Monoclonal Antibody (183-H12-5C) from Dr. Bruce Chesebro and Kathy Wehrly. Anti-FLAG was purchased from Sigma, anti-myc and anti-GAPDH was purchased from Rockland, and anti-eGFP was purchased from Clontech. anti-GAPDH (clone 6C5) was purchased from

EMD/Millipore. CADM1 (9D2) was obtained from MBL International and Chicken IgG Isotype Control was obtained from Novus Biologicals.

3.4.3 Quantitative real-time PCR

The total RNA was extracted using the PureLink RNA mini kit (Ambion, Life Technologies). Using the M-MLV reverse transcriptase and Oligo(dT) primers (Eurofins), 500 ng of RNA was reverse transcribed to cDNA. Prior to qPCR, cDNA samples were diluted 1:5with water. Each PCR reaction consisted of 10µL of SYBR Green Master Mix, 1.6µL of gene-specific primers (0.8µL of 10µM forward primer and 0.8µL of 10µM reverse primer),4µL of diluted cDNA, and water to a total volume of 20µL. Quantification of endogenous mRNA was run on the QuantStudio5 qPCR machine (Applied Biosystems) under the following cycling conditions: 2 min at 95∘C and 40 cycles of 5 sec at 95∘C, 10 s at 60∘C,and 20 s at 72∘C. The QuantStudio Design and Analysis Desktop Software (version 1.4) was used to determine the CT for each PCR reaction. Gene-specific forward and reverse primer sets are described in Table 4 rows 5 to 7.

Table 4 List of primers

1	pH6:R10A	5′ TTC TGT TGG GGC GCC GAC TCC GCG GAG CTG CAG CGC CGG AGG 3′	, 5′ CCT CCG GCG CTG CAG CTC CGC GGA GTC GGC GCC CCA ACA GAA 3′
2	pH6:E67A	5′ GCA GCG CGG GGA GCT GCC AGC ACC AAT TCA GGC ATT GGA AAC C 3′	5′ GGT TTC CAA TGC CTG AAT TGG TGC TGG CAG CTC CCC GCG CTG C 3′

3	pH6:R10A/E67A	5′ GCA GCG CGG GGA GCT GCC AGC ACC AAT TCA GGC ATT GGA AAC C 3′	5′ GGT TTC CAA TGC CTG AAT TGG TGC TGG CAG CTC CCC GCG CTG C 3′
4	pR10G	5′ CGC CGA CTC CGG GGA GCT GCA 3′	5′ TGC AGC TCC CCG GAG TCG GCG 3′
5	Gag	5′ CAT ATA GTA TGG GCA AGC AGG G 3′	5′ CTG TCT GAA GGG ATG GTT GTA G 3′
6	Rev	5′ GAG CTC ATC AGA ACA GTC AGA C 3′	5′ CGA ATG GAT CTG TCT CTG TCT C 3′
7	GAPDH_qpcr	5′ CAT GTT CGT CAT GGG TGT GAA CCA 3′	5′ AGT GAT GGC ATG GAC TGT GGT CAT 3′
8	HERC6 SNP rs111670008	5' GAT ATT CAG GAA ATC ATC GCG CAC C 3'	5' CAT GTA CTG GGG TCC TCA CAC C 3'

3.4.4 Isolation of PCR fragments containing HERC6 SNP rs111670008 from patients

Genomic DNA was isolated from frozen PBMCs from HIV-1 positive patients within the Ugandan cohort using the QIAGEN QIAamp DNA Blood Mini Kit per the manufacturers protocol. Genomic regions flanking the HERC6 SNP rs111670008 were PCR amplified using the primers in Table 4 row 8. Sequencing was performed at the London Regional Genomics Centre (Robarts Research Institute, Western University).

3.4.5 Ethics Statement

Details pertaining to the Uganda study population have been reported previously [46–50]. Briefly, women who became HIV infected while participating in the Hormonal Contraception and Risk of HIV Acquisition Study in Uganda were enrolled upon primary infection with HIV-1 into a subsequent study, the Hormonal Contraception and HIV-1 Genital Shedding and Disease Progression among Women with Primary HIV Infection (GS) Study. Ethical approval was obtained from the Institutional Review boards (IRBs) from the Joint Clinical Research Centre and UNST in Uganda, from University of Zimbabwe, from the University Hospitals of Cleveland, and recently, from Western University. All adult subjects provided written informed consent and no child participants were included in the study. Protocol numbers and documentation of these approvals/renewals are available upon request. Blood and cervical samples were collected every month for the first six months, then every three months for the first two years, and then every six months up to 9.5 years. Women who had CD4 lymphocyte counts of 200 cells/ml and/or who developed severe symptoms of HIV infection (WHO clinical stage IV or advanced stage III disease) were offered combination ART (cART) and trimethoprim-sulfamethoxazole (for prophylaxis against bacterial infections and *Pneumocystis jeroveci* pneumonia). HIV-1 (subtype A or D) positive patients were stratified into three groups based on their rate of disease progression. Rapid Progressors were characterized by a CD4+ T-cell count < 250 cells/µL within the first 3 years post infection (1095 days). Normal Progressors were characterized by a CD4+ T-cell count between 250 cells/µL and 499 cells/µL 5 years post infection (1825 days) without antiretroviral treatment. Slow progressors were characterized by a CD4+ T-cell count >500 cells/µL 5 years post infection (1825 days) without antiretroviral treatment.

3.4.6 Analyses of Sequences, Synteny and Positive Selection

The MEGA 7.0 package was used for the phylogenetic analysis[80]. The amino-terminal end of the small HERCs varies in length among the different members and was not

included for phylogenetic analysis. The first 23 amino acids from HERC5 and HERC6.
Modeling of HERC5 and HERC6 RLD was performed with VMD version 1.9.2[81]

3.4.7 Mass Spectrometry Sample preparation and Data analysis

Mock (no virus just empty vector control plasmid), Control (empty vector control and
R9), HERC6 (HERC6 and R9), HERC6-R10G (HERC6-R10G and R9) virus samples
were created in HEK 293T. The virus was concentrated using 100 KDa MWCO
centrifuge tubes (Amicon, USA) and washed several times using PBS. Fresh urea lysis
buffer (8M Urea, 2% SDS, 10mM DTT, 50mM Ammonium Bicarbonate) was used to
lyse the samples. Protein concentration was measured with Pierce 660nm Protein Assay
and 100 ug of starting material was used. The samples were reduced and alkylated using
10mM DTT and 100 mM (final concentration) Iodoacetamide solution. Methanol-
Chloroform precipitation method was used to clean up the protein samples (see [82]).
Proteins were resuspended in 100 ul of 50 mM Ammonium Bicarbonate solution and
digested with Trypsin/LysC (Promega) and LysC (Wako) enzymes at a ratio of 100:1
sample to enzyme. Samples were incubated overnight at 37°C. Samples were filtered in a
Microcon-10kDa centrifugal filter unit with ultracel-10 membrane (Millipore). Samples
were analyzed on a Thermo Scientific Q Exactive Plus Orbitrap LC-MS/MS System.
Raw files were searched against the human Uniprot databased (20, 274 entries) using the
de novo search engine PEAKS® (version 8). Parent and fragment mass error tolerances
were set to 20 ppm and 0.05 Da, respectively. Maximum missed cleavages were set to 3
and 1 non-specific cleavage was allowed. Carbamidomethylation was set as a fixed
modification, and deamidation, oxidation and acetylation (protein N-term) were included
as variable modifications with a maximum of 3 PTMs per peptide allowed. To correct for
the sample loading and technical variability, peak areas for each peptide were normalized
to the total ion current (TIC). Normalized peak areas of 0 were assumed to be missing not
at random and imputed with the lowest ratio detected for the given peptide. Figures were
generated using R version 4.1.2 (2021-11-01) with Bioconductor[83] packages tidyverse[84],
pheatmap[85], EnhancedVolcano[86] and annotated with human genome database

org.Hs.eg.db [87]. Gene ontology analysis was performed with ShinyGo 0.76.3 with the setting of FDR cutoff 0.05[88].

3.4.8 CADM1 Neutralization Assay

Methods were adapted from Sarzotti-Kelsoe et al., 2014 [89]. Briefly, R9 or NL4-3 infectious virus was incubated with either CADM1, VRC01, or isotype control neutralizing antibodies at increasing concentration of 0.014, 0.12, 0.36, 1.1, 3.3, 10, or 30 ug of antibody and 30ug/ml of DEAE dextran for 1 hour at 37°C. 10000 TZM-bl cells were added to the virus-antibody mixture and were incubated for 48 hours at 37 °C. Britelite plus Reporter Gene Assay System (PerkinElmer) was used to detect luciferase activity as per manufacturer's instructions.

3.4.9 Statistical analyses

GraphPad Prism v8 was used for all statistical analyses stated in the text. P values and statistical tests used are stated in the text where appropriate. P values less than 0.05 were deemed significant.

3.5   References

1.   Wang IH, Burckhardt CJ, Yakimovich A, Greber UF. Imaging, tracking and computational analyses of virus entry and egress with the cytoskeleton. *Viruses*. 2018;10(4):1-29. doi:10.3390/v10040166

2.   Shukla A, Ramirez NGP, D'Orso I. HIV-1 proviral transcription and latency in the new era. *Viruses*. 2020;12(5):1-41. doi:10.3390/v12050555

3.   Wilen CB, Tilton JC, Doms RW. HIV: cell binding and entry. *Cold Spring Harb Perspect Med*. 2012;2(8). doi:10.1101/cshperspect.a006866

4.   Novikova M, Zhang Y, Freed EO, Peng K. Multiple Roles of HIV-1 Capsid during the Virus Replication Cycle. *Virol Sin*. 2019;34(2):119-134. doi:10.1007/s12250-019-00095-3

5.   Dharan A, Bachmann N, Talley S, Zwikelmaier V, Campbell EM. Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. *Nat Microbiol*. Published online 2020. doi:10.1038/s41564-020-0735-8

6.   Christensen DE, Ganser-Pornillos BK, Johnson JS, Pornillos O, Sundquist WI. Reconstitution and visualization of HIV-1 capsid-dependent replication and integration in vitro. *Science*. 2020;370(6513). doi:10.1126/SCIENCE.ABC8420

7.   Ambrose Z, Aiken C. HIV-1 uncoating: Connection to nuclear entry and regulation by host proteins. *Virology*. 2014;454-455(1):371-379. doi:10.1016/j.virol.2014.02.004

8.   Francis AC, Marin M, Singh PK, et al. HIV-1 replication complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains. *Nat Commun*. 2020;11(1). doi:10.1038/s41467-020-17256-8

9.   Barrera A, Olguín V, Vera-Otarola J, López-Lastra M. Cap-independent translation initiation of the unspliced RNA of retroviruses. *Biochim Biophys Acta - Gene Regul Mech*. 2020;1863(9):194583. doi:https://doi.org/10.1016/j.bbagrm.2020.194583

10.   Wu Y. HIV-1 gene expression: lessons from provirus and non-integrated DNA. *Retrovirology*. 2004;1:13. doi:10.1186/1742-4690-1-13

11.   de Breyne S, Ohlmann T. Focus on Translation Initiation of the HIV-1 mRNAs. *Int J Mol Sci*. 2018;20(1). doi:10.3390/ijms20010101

12.   Burugu S, Daher A, Meurs EF, Gatignol A. HIV-1 translation and its regulation by cellular factors PKR and PACT. *Virus Res*. 2014;193:65-77. doi:10.1016/j.virusres.2014.07.014

13.   Freed EO. HIV-1 replication. *Somat Cell Mol Genet*. 2001;26(1-6):13-33. doi:10.1023/a:1021070512287

14.   Woods MW, Tong JG, Tom SK, et al. Interferon-induced HERC5 is evolving under positive selection and inhibits HIV-1 particle production by a novel mechanism targeting Rev/RRE-dependent RNA nuclear export. *Retrovirology*. 2014;11:27. doi:10.1186/1742-4690-11-27

15.   Toro-Ascuy D, Rojas-Araya B, García-de-Gracia F, et al. A Rev-CBP80-eIF4AI complex drives Gag synthesis from the HIV-1 unspliced mRNA. *Nucleic Acids Res*. 2018;46(21):11539-11552. doi:10.1093/nar/gky851

16.   Behrens RT, Aligeti M, Pocock GM, Higgins CA, Sherer NM. Nuclear Export Signal Masking Regulates HIV-1 Rev Trafficking and Viral RNA Nuclear Export. *J Virol*. 2017;91(3). doi:10.1128/JVI.02107-16

17.   Fernandes JD, Booth DS, Frankel AD. A structurally plastic ribonucleoprotein complex mediates post-transcriptional gene regulation in HIV-1. *Wiley Interdiscip Rev RNA*. 2016;7(4):470-486. doi:10.1002/wrna.1342

18.   Sherpa C, Rausch JW, Le Grice SFJ, Hammarskjold ML, Rekosh D. The HIV-1 Rev response element (RRE) adopts alternative conformations that promote different rates of virus replication. *Nucleic Acids Res*. 2015;43(9):4676-4686. doi:10.1093/NAR/GKV313

19.   Freed EO. HIV-1 assembly, release and maturation. *Nat Rev Microbiol*. 2015;13(8):484-496. doi:10.1038/nrmicro3490

20.   Tran P-T-H, Chiramel AI, Johansson M, Melik W. Roles of ESCRT Proteins ALIX and CHMP4A and Their Interplay with Interferon-Stimulated Gene 15 during Tick-Borne Flavivirus Infection. *J Virol*. 2022;96(3). doi:10.1128/JVI.01624-21

21.   Votteler J, Sundquist WI. Virus budding and the ESCRT pathway. *Cell Host Microbe*. 2013;14(3):232-241. doi:10.1016/j.chom.2013.08.012

22. Bell NM, Lever AML. HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol*. 2013;21(3):136-144. doi:10.1016/j.tim.2012.11.006

23. Lerner G, Weaver N, Anokhin B, Spearman P. Advances in HIV-1 Assembly. *Viruses*. 2022;14(3). doi:10.3390/v14030478

24. Freitas BT, Scholte FEM, Bergeron É, Pegan SD. How ISG15 combats viral infection. *Virus Res*. 2020;286. doi:10.1016/J.VIRUSRES.2020.198036

25. Hochrainer K, Mayer H, Baranyi U, Binder B, Lipp J, Kroismayr R. The human HERC family of ubiquitin ligases: novel members, genomic organization, expression profiling, and evolutionary aspects. *Genomics*. 2005;85(2):153-164. doi:https://doi.org/10.1016/j.ygeno.2004.10.006

26. Hochrainer K, Kroismayr R, Baranyi U, Binder BR, Lipp J. Highly homologous HERC proteins localize to endosomes and exhibit specific interactions with hPLIC and Nm23B. *Cell Mol Life Sci*. 2008;65(13):2105-2117. doi:10.1007/s00018-008-8148-5

27. Dastur A, Beaudenon S, Kelley M, Krug RM, Huibregtse JM. Herc5, an interferon-induced HECT E3 enzyme, is required for conjugation of ISG15 in human cells. *J Biol Chem*. 2006;281(7):4334-4338. doi:10.1074/jbc.M512830200

28. Durfee LA, Lyon N, Seo K, Huibregtse JM. The ISG15 conjugation system broadly targets newly synthesized proteins: implications for the antiviral function of ISG15. *Mol Cell*. 2010;38(5):722-732. doi:10.1016/j.molcel.2010.05.002

29. Woods MW, Kelly JN, Hattlmann CJ, et al. Human HERC5 restricts an early stage of HIV-1 assembly by a mechanism correlating with the ISGylation of Gag. *Retrovirology*. 2011;8:95. doi:10.1186/1742-4690-8-95

30. Im E, Yoo L, Hyun M, Shin WH, Chung KC. Covalent ISG15 conjugation positively regulates the ubiquitin E3 ligase activity of parkin. *Open Biol*.

2016;6(8). doi:10.1098/rsob.160193

31.   Durfee LA, Huibregtse JM. Identification and Validation of ISG15 Target
      Proteins. *Subcell Biochem*. 2010;54:228-237. doi:10.1007/978-1-4419-6676-6_18

32.   Zhang X, Bogunovic D, Payelle-Brogard B, et al. Human intracellular ISG15
      prevents interferon-α/β over-amplification and auto-inflammation. *Nature*.
      2015;517(7532):89-93. doi:10.1038/NATURE13801

33.   Versteeg GA, Hale BG, van Boheemen S, Wolff T, Lenschow DJ, García-Sastre
      A. Species-specific antagonism of host ISGylation by the influenza B virus NS1
      protein. *J Virol*. 2010;84(10):5423-5430. doi:10.1128/JVI.02395-09

34.   Paparisto E, Woods MW, Coleman MD, et al. Evolution-guided structural and
      functional analyses of the HERC family reveal an ancient marine origin and
      determinants of antiviral activity. *J Virol*. 2018;92(13). doi:10.1128/JVI.00528-18

35.   Woods MW, Kelly JN, Hattlmann CJ, et al. Human HERC5 restricts an early stage
      of HIV-1 assembly by a mechanism correlating with the ISGylation of Gag.
      Retrovirology. 2011;8:95. doi:10.1186/1742-4690-8-95

36.   Okumura A, Lu G, Pitha-Rowe I, Pitha PM. Innate antiviral response targets HIV-
      1 release by the induction of ubiquitin-like protein ISG15. *Proc Natl Acad Sci U S
      A*. 2006;103(5):1440-1445.
      doi:10.1073/PNAS.0510518103/SUPPL_FILE/10518FIG5.PDF

37.   Paparisto E, Woods MW, Coleman MD, et al. Evolution-Guided Structural and
      Functional Analyses of the HERC Family Reveal an Ancient Marine Origin and
      Determinants of Antiviral Activity. *J Virol*. 2018;92(13). doi:10.1128/jvi.00528-18

38.   Bonaventure B, Rebendenne A, Luiza A, et al. The DEAD box RNA helicase
      DDX42 is an intrinsic inhibitor of positive-strand RNA viruses. *EMBO Rep*.
      Published online September 26, 2022:e54061. doi:10.15252/EMBR.202154061

39. Yu X, Shang H, Jiang Y. ICAM-1 in HIV infection and underlying mechanisms. *Cytokine*. 2020;125:154830. doi:10.1016/j.cyto.2019.154830

40. Sanders RW, de Jong EC, Baldwin CE, Schuitemaker JHN, Kapsenberg ML, Berkhout B. Differential transmission of human immunodeficiency virus type 1 by distinct subsets of effector dendritic cells. *J Virol*. 2002;76(15):7812-7821. doi:10.1128/jvi.76.15.7812-7821.2002

41. Wu L, KewalRamani VN. Dendritic-cell interactions with HIV: infection and viral dissemination. *Nat Rev Immunol*. 2006;6(11):859-868. doi:10.1038/nri1960

42. Moiseeva EP, Straatman KR, Leyland ML, Bradding P. CADM1 controls actin cytoskeleton assembly and regulates extracellular matrix adhesion in human mast cells. *PLoS One*. 2014;9(1):e85980. doi:10.1371/journal.pone.0085980

43. Arthur G, Bradding P. New Developments in Mast Cell Biology: Clinical Implications. *Chest*. 2016;150(3):680-693. doi:10.1016/j.chest.2016.06.009

44. Yang W, Kaur D, Okayama Y, et al. Human lung mast cells adhere to human airway smooth muscle, in part, via tumor suppressor in lung cancer-1. *J Immunol*. 2006;176(2):1238-1243. doi:10.4049/jimmunol.176.2.1238

45. Hagiyama M, Mimae T, Wada A, et al. Possible Therapeutic Utility of anti-Cell Adhesion Molecule 1 Antibodies for Malignant Pleural Mesothelioma. *Front cell Dev Biol*. 2022;10. doi:10.3389/FCELL.2022.945007

46. Moiseeva EP, Roach KM, Leyland ML, Bradding P. CADM1 Is a Key Receptor Mediating Human Mast Cell Adhesion to Human Lung Fibroblasts and Airway Smooth Muscle Cells. *PLoS One*. 2013;8(4). doi:10.1371/JOURNAL.PONE.0061579

47. Li Y, O'Dell S, Walker LM, et al. Mechanism of Neutralization by the Broadly Neutralizing HIV-1 Monoclonal Antibody VRC01. *J Virol*. 2011;85(17):8954.

doi:10.1128/JVI.00754-11

48. Burnie J, Guzzo C. The Incorporation of Host Proteins into the External HIV-1 Envelope. *Viruses*. 2019;11(1). doi:10.3390/v11010085

49. Magadmi R, Meszaros J, Damanhouri ZA, Seward EP. Secretion of Mast Cell Inflammatory Mediators Is Enhanced by CADM1-Dependent Adhesion to Sensory Neurons. *Front Cell Neurosci*. 2019;13:262. doi:10.3389/fncel.2019.00262

50. McLaren PJ, Coulonges C, Bartha I, et al. Polymorphisms of large effect explain the majority of the host genetic contribution to variation of HIV-1 virus load. *Proc Natl Acad Sci U S A*. 2015;112(47):14658-14663. doi:10.1073/pnas.1514867112

51. Naranbhai V, Carrington M. Host genetic variation and HIV disease: from mapping to mechanism. *Immunogenetics*. 2017;69(8-9):489-498. doi:10.1007/s00251-017-1000-z

52. Ni J, Wang D, Wang S. The CCR5-Delta32 Genetic Polymorphism and HIV-1 Infection Susceptibility: a Meta-analysis. *Open Med*. 2018;13(1):467. doi:10.1515/MED-2018-0062

53. Zajac V. Evolutionary view of the AIDS process. *J Int Med Res*. 2018;46(10):4032-4038. doi:10.1177/0300060518786919

54. Gupta PK, Saxena A. HIV/AIDS: Current Updates on the Disease, Treatment and Prevention. *Proc Natl Acad Sci India Sect B*. 2021;91(3):495-510. doi:10.1007/S40011-021-01237-Y

55. Hütter G, Bodor J, Ledger S, et al. CCR5 Targeted Cell Therapy for HIV and Prevention of Viral Escape. *Viruses*. 2015;7(8):4186-4203. doi:10.3390/V7082816

56. Vidyant S, Chatterjee A, Dhole TN. A single-nucleotide polymorphism in TLR4 is linked with the risk of HIV-1 infection. *Br J Biomed Sci*. 2019;76(2):59-63.

doi:10.1080/09674845.2018.1559486

57.   Kim Y-C, Jeong B-H. Strong Association of the rs4986790 Single Nucleotide
       Polymorphism (SNP) of the Toll-Like Receptor 4 (TLR4) Gene with Human
       Immunodeficiency Virus (HIV) Infection: A Meta-Analysis. *Genes (Basel)*.
       2020;12(1). doi:10.3390/genes12010036

58.   Pine SO, McElrath MJ, Bochud P-Y. Polymorphisms in toll-like receptor 4 and
       toll-like receptor 9 influence viral load in a seroincident cohort of HIV-1-infected
       individuals. *AIDS*. 2009;23(18):2387-2395. doi:10.1097/QAD.0b013e328330b489

59.   Sironi M, Biasin M, Cagliani R, et al. A common polymorphism in TLR3 confers
       natural resistance to HIV-1 infection. *J Immunol*. 2012;188(2):818-823.
       doi:10.4049/jimmunol.1102179

60.   Oh D-Y, Baumann K, Hamouda O, et al. A frequent functional toll-like receptor 7
       polymorphism is associated with accelerated HIV-1 disease progression. *AIDS*.
       2009;23(3):297-307. doi:10.1097/QAD.0b013e32831fb540

61.   Selvaraj P, Harishankar M, Singh B, Jawahar MS, Banurekha V V. Toll-like
       receptor and TIRAP gene polymorphisms in pulmonary tuberculosis patients of
       South India. *Tuberculosis (Edinb)*. 2010;90(5):306-310.
       doi:10.1016/j.tube.2010.08.001

62.   Bochud P-Y, Hersberger M, Taffé P, et al. Polymorphisms in Toll-like receptor 9
       influence the clinical course of HIV-1 infection. *AIDS*. 2007;21(4):441-446.
       doi:10.1097/QAD.0b013e328012b8ac

63.   Compaore TR, Soubeiga ST, Ouattara AK, et al. APOBEC3G Variants and
       Protection against HIV-1 Infection in Burkina Faso. *PLoS One*.
       2016;11(1):e0146386. doi:10.1371/journal.pone.0146386

64.   Amanya SB, Nyiro B, Waswa F, et al. Variations in Trim5α and Cyclophilin A

genes among HIV-1 elite controllers and non controllers in Uganda: a laboratory-based cross-sectional study. *Retrovirology*. 2020;17(1):19. doi:10.1186/s12977-020-00527-z

65. Laplana M, Caruz A, Pineda JA, Puig T, Fibla J. Association of BST-2 gene variants with HIV disease progression underscores the role of BST-2 in HIV type 1 infection. *J Infect Dis*. 2013;207(3):411-419. doi:10.1093/infdis/jis685

66. Gupte DS, Patil A, Selvaa Kumar C, et al. Risk association of BST2 gene variants with disease progression in HIV-1 infected Indian cohort. *Infect Genet Evol*. 2020;80:104139. doi:10.1016/j.meegid.2019.104139

67. Dias BDC, Paximadis M, Martinson N, Chaisson RE, Ebrahim O, Tiemessen CT. The impact of bone marrow stromal antigen-2 (BST2) gene variants on HIV-1 control in black South African individuals. *Infect Genet Evol*. 2020;80:104216. doi:10.1016/j.meegid.2020.104216

68. Mlimi H, Naidoo KK, Mabuka J, Ndung'u T, Madlala P. Bone marrow stromal antigen 2 (BST-2) genetic variants influence expression levels and disease outcome in HIV-1 chronically infected patients. *Retrovirology*. 2022;19(1):3. doi:10.1186/s12977-022-00588-2

69. Ott DE. Cellular proteins detected in HIV-1. *Rev Med Virol*. 2008;18(3):159-175. doi:10.1002/rmv.570

70. Desimmie BA, Weydert C, Schrijvers R, et al. HIV-1 IN/Pol recruits LEDGF/p75 into viral particles. *Retrovirology*. 2015;12:16. doi:10.1186/s12977-014-0134-4

71. Gurer C, Cimarelli A, Luban J. Specific incorporation of heat shock protein 70 family members into primate lentiviral virions. *J Virol*. 2002;76(9):4666-4670. doi:10.1128/jvi.76.9.4666-4670.2002

72. Guzzo C, Ichikawa D, Park C, et al. Virion incorporation of integrin α4β7

facilitates HIV-1 infection and intestinal homing. *Sci Immunol*. 2017;2(11). doi:10.1126/sciimmunol.aam7341

73. Xu S, Zheng Z, Pathak JL, et al. The Emerging Role of the Serine Incorporator Protein Family in Regulating Viral Infection. *Front cell Dev Biol*. 2022;10:856468. doi:10.3389/fcell.2022.856468

74. Gonzalez-Enriquez GV, Escoto-Delgadillo M, Vazquez-Valls E, Torres-Mendoza BM. SERINC as a Restriction Factor to Inhibit Viral Infectivity and the Interaction with HIV. *J Immunol Res*. 2017;2017:1548905. doi:10.1155/2017/1548905

75. Welch JL, Stapleton JT, Okeoma CM. Vehicles of intercellular communication: exosomes and HIV-1. *J Gen Virol*. 2019;100(3):350. doi:10.1099/JGV.0.001193

76. Liu Y, Fu Y, Wang Q, et al. Proteomic profiling of HIV-1 infection of human CD4+ T cells identifies PSGL-1 as an HIV restriction factor. *Nat Microbiol*. 2019;4(5):813-825. doi:10.1038/s41564-019-0372-2

77. Santos S, Obukhov Y, Nekhai S, Bukrinsky M, Iordanskiy S. Virus-producing cells determine the host protein profiles of HIV-1 virion cores. *Retrovirology*. 2012;9(1):1-28. doi:10.1186/1742-4690-9-65/FIGURES/4

78. Linde ME, Colquhoun DR, Ubaida Mohien C, et al. The conserved set of host proteins incorporated into HIV-1 virions suggests a common egress pathway in multiple cell types. *J Proteome Res*. 2013;12(5):2045-2054. doi:10.1021/pr300918r

79. Li C, Wen A, Shen B, Lu J, Huang Y, Chang Y. FastCloning: A highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. *BMC Biotechnol*. 2011;11. doi:10.1186/1472-6750-11-92

80. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*. 2016;33(7):1870-1874.

doi:10.1093/MOLBEV/MSW054

81.   Stone JE, McGreevy R, Isralewitz B, Schulten K. GPU-accelerated analysis and visualization of large structures solved by molecular dynamics flexible fitting. *Faraday Discuss*. 2014;169:265-283. doi:10.1039/c4fd00005f

82.   Wessel D, Flügge UI. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem*. 1984;138(1):141-143. doi:10.1016/0003-2697(84)90782-6

83.   Huber W, Carey VJ, Gentleman R, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods*. 2015;12(2):115-121. http://www.nature.com/nmeth/journal/v12/n2/full/nmeth.3252.html

84.   Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *J Open Source Softw*. 2019;4(43):1686. doi:10.21105/JOSS.01686

85.   Package "pheatmap." Published online 2022.

86.   Blighe K, Rana S, Lewis M. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. Published online 2021. https://github.com/kevinblighe/EnhancedVolcano

87.   Carlson M. org.Hs.eg.db: Genome wide annotation for Human. Published online 2021.

88.   Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*. 2020;36(8):2628-2629. doi:10.1093/bioinformatics/btz931

89.   Sarzotti-Kelsoe M, Bailer RT, Turk E, Lin CL, Bilska M, Greene KM, Gao H, Todd CA, Ozaki DA, Seaman MS, Mascola JR, Montefiori DC. Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing

antibodies against HIV-1. J Immunol Methods. 2014 Jul;409:131-46. doi: 10.1016/j.jim.2013.11.022.

Chapter 4

4 Interferon-Induced HERC5 Inhibits Ebola Virus Particle Production and Is Antagonized by Ebola Glycoprotein

Survival following Ebola virus (EBOV) infection correlates with the ability to mount an early and robust interferon (IFN) response. The host IFN-induced proteins that contribute to controlling EBOV replication are not fully known. Among the top genes with the strongest early increases in expression after infection *in vivo* is IFN-induced HERC5. Using a transcription- and replication-competent VLP system, we showed that HERC5 inhibits EBOV virus-like particle (VLP)replication by depleting EBOV mRNAs. The HERC5 RCC1-like domain was necessary and sufficient for this inhibition and did not require zinc finger antiviral protein (ZAP). Moreover, we showed that EBOV (Zaire) glycoprotein (GP) but not Marburg virus GP antagonized HERC5 early during infection. Our data identify a novel 'protagonist–antagonistic' relationship between HERC5 and GP in the early stages of EBOV infection that could be exploited for the development of novel antiviral therapeutics. This chapter was published in Cells on September 13, 2021 (see Appendix 1 for publication permissions).

4.1 Introduction

Ebola virus (EBOV) is a member of the Filoviridae family of single-stranded negative-sense RNA viruses with a filamentous morphology. EBOV infection results in severe hemorrhagic fever and can lead to death 6-16 days after the onset of symptoms in up to 90% of cases, making EBOV one of the most virulent pathogens to infect humans[1] .Studies involving primate models, and human studies carried out during the 2013–2016 outbreak, showed that EBOV exposure results in an early and robust immune response, largely characterized by the up-regulation of IFN-stimulated genes[2–12]. A contributing factor to the pathophysiology of EBOV infection is the ability of the virus to evade the host IFN response[7,13–16] . Using *in vitro* models of infection, it was shown that EBOV is able to evade the innate immune response through various IFN antagonisms, notably involvingVP24 and VP35 proteins[17–19]. The key mediators of this early cellular IFN

response to EBOV and how EBOV withstands this early response are not fully characterized. Restriction factors are key intrinsic mediators of the early IFN response and potently inhibit different steps in the life cycle of evolutionarily diverse viruses in the absence of viral antagonists[20]. Bone marrow stromal cell antigen 2 (BST-2)/tetherin is one such factor that potently inhibits the release of EBOV from cells by tethering virions to the surface of cells[21,22]. This inhibition is counteracted by EBOV GP[23–26]. IFN-inducible trans-membrane proteins 1–3 (IFITM1–3) comprise another family of factors that restrict the cellular entry of EBOV, although an EBOV antagonist to these proteins has yet to be identified[27,28]. HECT and RCC1-like containing domain 5 (HERC5) are some of the genes with the strongest early increases in expression in multiple tissues after EBOV infection[3,5,6,29]. HERC5 is an evolutionarily ancient restriction factor that inhibits the replication of diverse viruses[30–36]. By virtue of its C-terminal HECT domain, HERC5 is the main cellular E3 ligase for conjugating ISG15 to substrates and localizes to polyribosomes to modify newly translated viral proteins, thereby disrupting key aspects of viral particle production[30,37–39]. E3 ligase-independent antiviral activity has also been demonstrated towards HIV-1, where it inhibits the nuclear export of incompletely spliced viral RNAs by a mechanism requiring its N-terminal RCC1-like domain (RLD)[30] . Here, we examined the antiviral activity of HERC5 towards EBOV VLP production and replication. We identified a novel E3 ligase-independent mechanism by which HERC5inhibits viral particle production involving the depletion of EBOV mRNAs. In addition, we demonstrated that EBOV GP antagonizes HERC5 activity and rescues EBOV production and replication.

4.2   Results

4.2.1 HERC5 Inhibits EBOV trVLP Replication

Previous studies have identified HERC5 as a potent inhibitor of diverse viruses [30,31,32,33,34,35,36]. To determine if HERC5 restricts EBOV particle production and replication, we used an EBOV (Zaire) transcription- and replication-competent VLP (trVLP) system. This system utilizes a tetracistronic minigenome ('4cis') carrying a

luciferase reporter gene together with VP40, VP24, and GP (Figure 4.1A)[40,41]. The advantage of this system over conventional VLP assays is that the viral proteins VP40, GP and VP24 are encoded by the minigenome and expressed from the EBOV promoter in a more natively regulated fashion[40]. The co-expression of this minigenome with NP, VP35, VP30, and L drive genome replication and transcription, synthesis of the minigenome-encoded proteins, and formation of infectious trVLPs. These trVLPs incorporate minigenomes and are capable of undergoing multiple rounds of replication and infection in target cells that express NP, VP35, VP30, L and Tim-1 (Figure 4.1B). The replication of these trVLPs was quantified over multiple passages (every three days) by measuring the luciferase reporter activity within cells. As a negative control, the plasmid carrying the Ebola L gene was omitted from the transfections, which abrogated the trVLP formation. Compared to the control cells transfected with an empty vector plasmid, cells expressing HERC5 exhibited a significant reduction in trVLP replication over four passages (Figure 4.1C). The reduction in luciferase reporter activity also correlated with a reduction in GP and VP40 mRNA levels (Figure 4.1D).

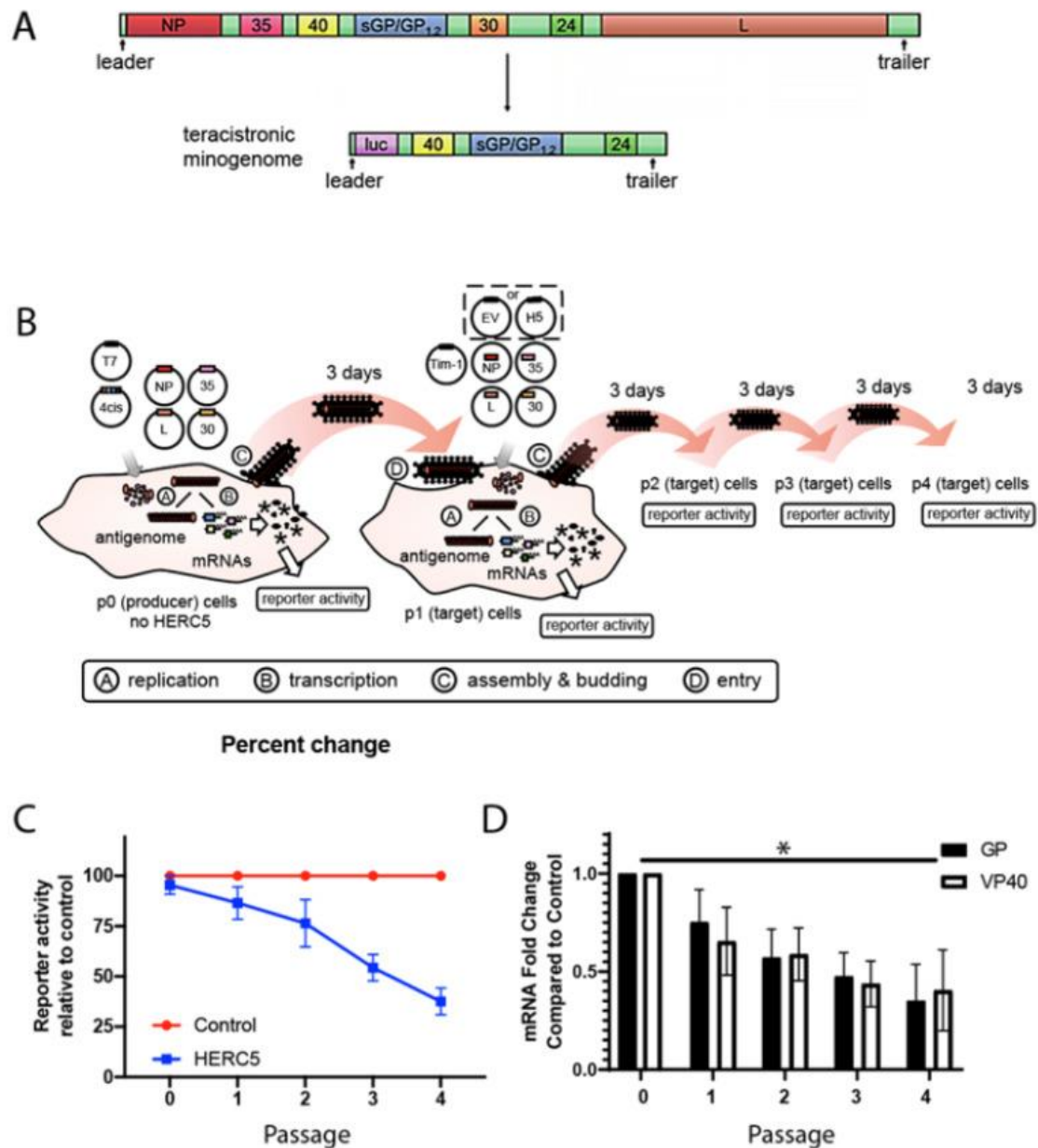**Figure 4.1 HERC5 inhibits EBOV trVLP replication. HERC5 Inhibits EBOV VP40 Particle Production**

A) Schematic depicting EBOV full-length genome and the derived tetracistronic minigenome. **(B)** The trVLP propagation assay. A tetracistronic EBOV minigenome (4cis) is expressed in cells together with the viral ribonucleoprotein complex (RNP) proteins (NP, VP35, VP30 and L). After the initial transcription by a co-expressed T7

polymerase, the minigenome is replicated and transcribed by the RNP proteins. Expression of VP40, GP and VP24 from the minigenome leads to the formation of infectious trVLPs containing minigenomes, which can infect target cells. Multiple infectious cycles can be modeled in cells expressing NP, VP35, L, VP30 and Tim-1 without the need for additional transfections of plasmids carrying VP40, GP and VP24. The figure was adapted from40, copyright © American Society for Microbiology, J. Virol. 88, 2014, 10,511–10,524, doi:10.1128/JVI.01272-14. **(C)** Quantification of trVLP propagation in the presence and absence of HERC5. The trVLP propagation assay was performed using tetracistronic minigenomes carrying a luciferase reporter, EBOV VP40, VP24 and EBOV GP over four passages (spanning 12 days). All EBOV minigenomes and plasmids carrying the EBOV proteins are based on EBOV H. sapiens-tc/COD/1976/Yambuku-Mayinga. Luciferase reporter activity relative to the control (trVLPs propagated in the absence of HERC5) is shown. The data shown represent the average (+/− S.E.M.) of four independent experiments. Linear regression analysis, F = 39.14. DFn = 1, DFd = 36; p < 0.0001. **(D)** The mRNA of GP and VP40 was measured using qRT-PCR at each passage. The data shown represent the average (+/− S.E.M.) of the four independent experiments represented in part C. * p < 0.05; One-way ANOVA with Dunnet's multiple comparisons test compared to the control.

Previous studies showed that HERC5 interferes with the function of key viral structural proteins[30,31,36,37]. The EBOV structural protein VP40 is necessary and sufficient for the assembly and budding of virus particles. When expressed in the absence of any other viral protein, VP40 can form VLPs that bud and are released from cells similar to wild-type EBOV[42–44]. To determine if HERC5 targets VP40, we co-transfected 293T cells with a plasmid carrying VP40 and increasing concentrations of plasmids carrying either empty vector control or FLAG-tagged HERC5. VP40 protein levels within cells and in VLPs were measured using quantitative Western blotting. HERC5 transfection did not alter cell viability (Appendix 3). As shown in Figure 4.2A and Appendix 3B, C, HERC5 inhibited

the production of VP40 VLPs in a dose-dependent manner when VP40 is tagged with either GFP or with FLAG but had no effect on intracellular GFP levels. As a control, transfection with HERC4, a closely related member of the small HERC family, did not significantly alter cell viability, VP40 or GFP levels (Appendix 3D–F). In contrast, when HERC5 mRNA levels were reduced using RNA interference, an increase in intracellular VP40 protein levels and an increase in the production of VP40 VLPs were observed compared to the control cells (Figure 4.2B, C).

We also assessed the impact of HERC5 expression on VLPs using confocal microscopy and transmission electron microscopy (TEM). As expected, cells expressing VP40 with enhanced green fluorescent protein fused at its amino-terminus (VP40-EGFP) exhibited punctate fluorescence at the cell surface (Figure 4.2D). In contrast, cells co-expressing VP40-EGFP and HERC5 exhibited substantially less punctate fluorescence at the cell surface compared to the control cells. The presence of VP40 protein at the cell surface was also confirmed using TEM and immunogold TEM (Figure 4.2E, F). In cells expressing VP40-EGFP alone, an accumulation of immunogold particles was observed in budding structures at the cell surface, which was significantly different from a random distribution (Appendix 5 and Appendix 6). Cells expressing HERC5 exhibited markedly fewer VP40-EGFP-containing structures at the cell surface compared to the control cells. In addition, cells expressing HERC5 exhibited on average eight-fold fewer immunogold particles per cell compared to the control cells (Figure 4.2G). Notably, the few VP40-EGFP-containing structures that were observed in cells expressing HERC5 were located predominantly in a region under the plasma membrane.

We then asked whether the reduced VP40 protein levels correlated with reduced intracellular VP40 mRNA levels. The quantitative PCR showed that 293T cells co-expressing HERC5 and FLAG-tagged VP40 exhibited reduced intracellular levels of VP40 mRNA (nine-fold) compared to the control cells not expressing HERC5 (Figure 4.2H). Similar results were obtained when HERC5 was co-expressed with a VP40-EGFP fusion protein (Appendix 4). As a control, HERC5 expression had no significant effect on

EGFP mRNA levels when EGFP was expressed alone (Figure 4.2H and Appendix 4). To determine if the effect of HERC5 is specific for VP40 mRNA, we assessed the impact of HERC5 expression on the level of other EBOV mRNAs. Cells co-expressing HERC5 and either VP30, VP35, L or NP exhibited a two- to five-fold reduction in mRNA levels compared to the control cells (Figure 4.2H). Together, these data show that HERC5 inhibits EBOV VP40 particle production by a mechanism involving the depletion of EBOV mRNAs.

A. VLPs / Cells blots showing VP40, HERC5, VP40, GAPDH with pHERC5 (ug): 0, 0.4, 0.8, 1.5. Graphs: Cells — Fold change in VP40 protein vs pHERC5 (ug): 0, 0.4, 0.8*, 1.5**. VLPs — Fold change in VP40 protein vs pHERC5 (ug): 0, 0.4***, 0.8****, 1.5****.

B. Fold change in HERC5 mRNA: Scram vs HERC5 shRNA**.

C. VLPs/Cells blots (Scram, H5 shRNA): VP40, HERC5, VP40, GAPDH. Graphs: Fold change in HERC5 protein (Scram vs shHERC5****); Cells Fold change in VP40 protein (Scram vs shHERC5*); VLPs Fold change in VP40 protein (Scram vs shHERC5*).

D. Immunofluorescence: Hoechst, VP40-eGFP, anti-FLAG, Merge for Control and FLAG-HERC5.

E. TEM: Control, VP40-eGFP, VP40-eGFP + HERC5 (Major phenotype, Minor phenotype).

F. Immunogold TEM: same conditions.

G. Average # of gold particles per cell: Mock, Control, HERC5**.

H. Fold change in mRNA copy number: Control, VP40****, VP30**, VP35**, L****, NP***, GP**, GFP.
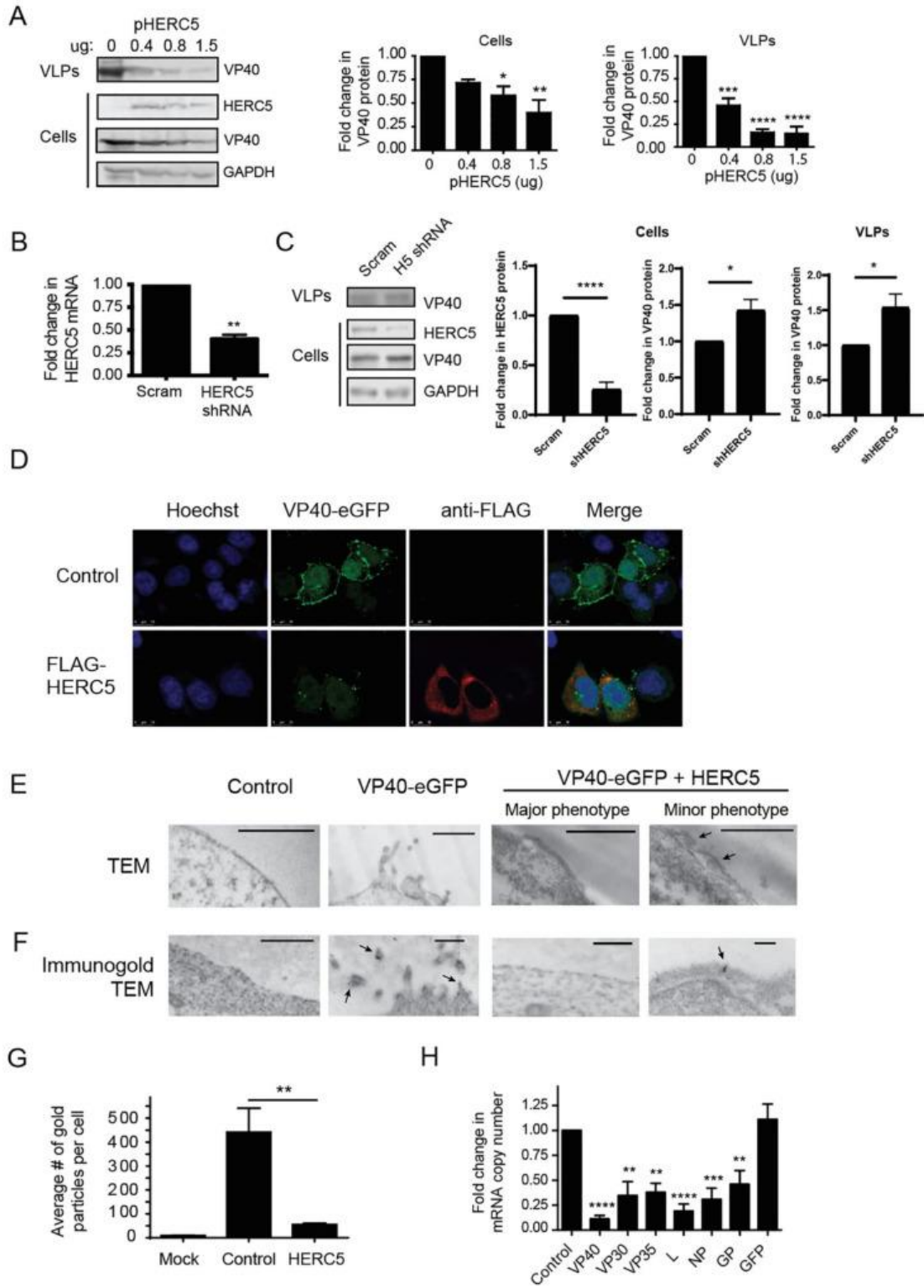
Figure 4.2 HERC5 inhibits EBOV VP40 particle production.

A) 293T cells were co-transfected with plasmids carrying FLAG-tagged VP40 (pFLAG-VP40) and increasing concentrations of FLAG-tagged HERC5 (pFLAG-HERC5). Empty vector plasmid was transfected in the condition with no HERC5 and used to ensure equal amounts of DNA were transfected in each condition. Forty-eight hours post-transfection, purified VLPs released into the cell supernatant and intracellular protein were subjected to quantitative Western blot analysis using anti-FLAG, anti-VP40 and anti-GAPDH. The average densitometric quantification of VP40 protein bands is shown to the right after normalization to GAPDH levels (+/− S.E.M.). A representative Western blot of four independent experiments is shown. **(B)** 293T cells were co-transfected with pFLAG-VP40 and either scrambled short-hairpin RNA (shRNA) (scram) or HERC5shRNA (shHERC5). Forty-eight hours after transfection, intracellular levels of HERC5 mRNA were quantified via qPCR. Data shown is the average (+/− S.E.M.) of three independent experiments. **(C)** 293T cells were transfected with either scrambled short-hairpin RNA (shRNA) (scram) or HERC5shRNA (shHERC5) for 24 h and then with pFLAG-HERC5 and pFLAG-VP40 for forty-eight hours. Purified VLPs released into the cell supernatant and intracellular protein were subjected to quantitative Western blot analysis using anti-FLAG and anti-GAPDH. The average densitometric quantification of VP40 protein bands is shown to the right after normalization to GAPDH levels (+/− S.E.M.). A representative Western blot of four independent experiments is shown. **(D)** HeLa cells were co-transfected with pVP40-EGFP and either empty vector (control) or pFLAG-HERC5 and visualized using confocal microscopy 48 h post-transfection. **(E)** 293T cells were "mock" transfected (control), transfected with empty vector and pVP40-EGFP, or transfected with pFLAG-HERC5 and pVP40-EGFP and analyzed via transmission electron microscopy (TEM) after 48 h. Virus particles beneath the plasma membrane are indicated with arrows. **(F)** Representative immunogold TEM images of 293T cells transfected as in **(E)** and labelled with 5 (+/− 2) nm anti-GFP immunogold particles. Immunogold-labelled VLPs are indicated with arrows. Scale bars = 500 nm. **(G)** The number of gold particles per positive cell was counted and presented as the average

number of particles per cell (+/− S.E.M). **(H)** 293T cells were co-transfected with plasmids carrying FLAG-HERC5 (or empty vector) and either EBOV VP40, VP30, VP35, L, NP, GP or GFP at a ratio of 10:1 (HERC5: EBOV plasmids). Forty-eight hours post-transfection viral mRNA was measured using qPCR after normalization to GAPDH mRNA levels. Data shown are representative of three independent experiments (+/− S.E.M.). **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$; One-way ANOVA with Dunnet's multiple comparisons test compared to the control (A, G); Student's paired t-test (B, C, H).

4.2.2 HERC5 RLD Is Necessary and Sufficient for Inhibition of VP40 Particle Production

To determine if the RLD or HECT domains of HERC5 are required for inhibition, we tested the ability of several HERC5 mutants to inhibit VP40 particle production. 293T cells were co-transfected with plasmids carrying VP40 and either empty vector (control), wild type HERC5 or HERC5 mutants lacking the RCC1-like domain (HERC5-ΔRLD), spacer region (HERC5-Δspacer) or HECT domain (HERC5-ΔHECT). We also tested the HERC5 RLD alone (HERC5-RLDonly) or HERC5 containing a cysteine to an alanine point mutation of residue 994 (HERC5-C994A), which specifically inactivates its E3 ligase activity (4.3A). Each of the FLAG-tagged mutant proteins was expressed at similar levels in 293T cells (Figure 4.3B).

As shown in Figure 4.3C, cells expressing wild type HERC5, HERC5-ΔHECT or HERC5-C994A reduced VP40 protein levels, which also correlated with reduced VP40 VLP production. In contrast, cells expressing HERC5-ΔRLD, and to a lesser extent HERC5-Δspacer, exhibited a diminished capacity to reduce VP40 protein levels and VP40 VLP production. Notably, expression of the HERC5 RLD alone (HERC5-RLDonly) reduced VP40 protein levels and VP40 VLP production similar to wild-type HERC5 (Figure 4.3D). We also examined the ability of the different HERC5 mutants to reduce VP40 mRNA levels. All HERC5 mutants except for HERC5-ΔRLD significantly reduced VP40 mRNA levels (4.3E). Taken together, these data show that the HERC5

RLD is necessary and sufficient to reduce VP40 mRNA levels and VP40 particle production.
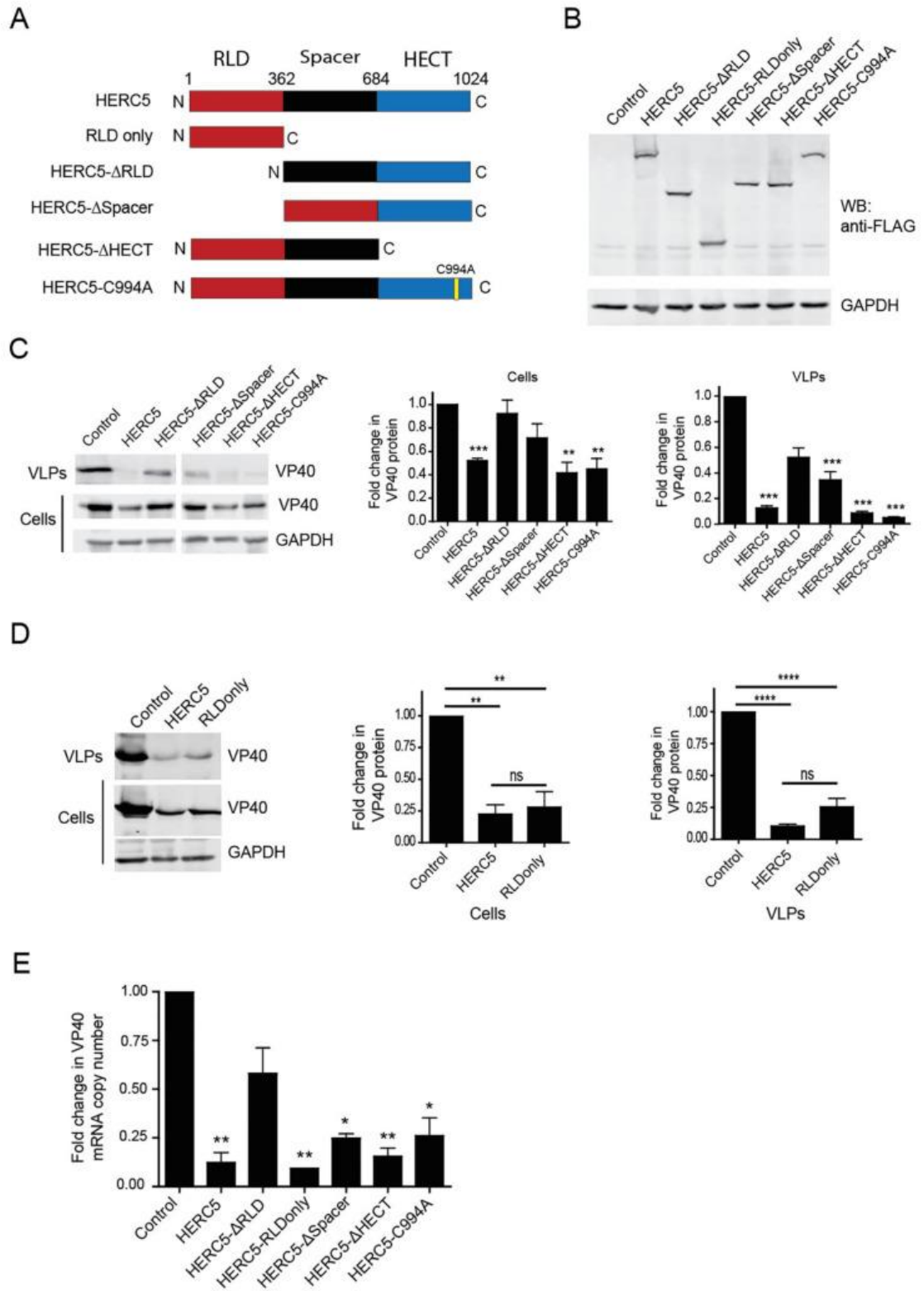
Figure 4.3 The RLD is necessary and sufficient for HERC5-mediated restriction.

A) Schematic of the different HERC5 mutant constructs. **(B)** Representative Western blot showing consistent expression of wild-type HERC5 and mutant forms of HERC5. 293T cells were transfected with either empty vector or plasmids carrying FLAG-tagged HERC5, HERC5-ΔRLD, HERC5-RLDonly, HERC5-ΔSpacer, HERC5-ΔHECT or HERC5-C994A. Forty-eight hours after transfection, cell lysate was subjected to Western blot analysis using anti-FLAG and anti-GAPDH. **(C)** 293T cells were co-transfected with plasmids carrying FLAG-tagged VP40 and either empty vector, wild-type HERC5 or one of the HERC5 mutants listed in (A). Forty-eight hours post-transfection, purified VLPs released into the supernatant and intracellular protein were examined by Western blotting using anti-FLAG and anti-GAPDH. VP40 protein levels were quantified densitometrically after normalization to GAPDH levels (graphs on the right). **(D)** 293T cells were co-transfected with plasmids carrying VP40-EGFP and either empty vector, HERC5 or HERC5-RLDonly. Cell lysates and VLPs were analyzed via Western blotting using anti-GFP and anti-GAPDH. VP40-EGFP protein levels were quantified densitometrically (graphs on the right). **(E)** 293T cells were co-transfected with plasmids carrying FLAG-tagged VP40 and either empty vector, HERC5, HERC5-ΔRLD, HERC5-RLDonly, HERC5-ΔSpacer, HERC5-ΔHECT or HERC5-C994A. Forty-eight hours post-transfection, mRNA was isolated and used to measure intracellular VP40 mRNA levels using qPCR. All data shown are representative of three independent experiments (+/− S.E.M.). **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * $p \leq 0.05$, ns (not significant) $p > 0.05$; One-way ANOVA with Dunnet's multiple comparisons test compared to the control.

4.2.3 HERC5 Depletes VP40 mRNA Independently of ZAP

ZAP (also called Zinc finger CCCH-type, antiviral 1, ZC3HAV1, and Poly (ADP-ribose) polymerase 13, PARP13) is an antiviral protein that causes significant loss of viral mRNAs from evolutionarily diverse RNA viruses, including Filoviridae, Retroviridae, Togaviridae and Hepadnaviridae[45–51]. We, therefore, asked if ZAP was required for

HERC5-mediated depletion of EBOV mRNA. We co-expressed VP40 and HERC5 in 293T cells that were knocked out for all ZAP isoforms and measured VP40 mRNA and protein levels using qPCR and Western blotting[52,53]. Cells expressing HERC5 in the absence of ZAP significantly reduced VP40 mRNA levels (Figure 4.4A). Exogenous expression of ZAP (short isoform) in the ZAP knockout cells reduced VP40 mRNA levels as previously shown[48,52]. Co-expression of HERC5 and ZAP together resulted in an enhanced loss of VP40 mRNA (Figure 4.4A). In support of this observation, cells expressing HERC5 in the absence of ZAP significantly reduced intracellular VP40 protein and VP40 VLPs the cell supernatant (Figure 4.4B, C). Together, these data show that ZAP is not required for HERC5-mediated reduction of VP40 mRNA.
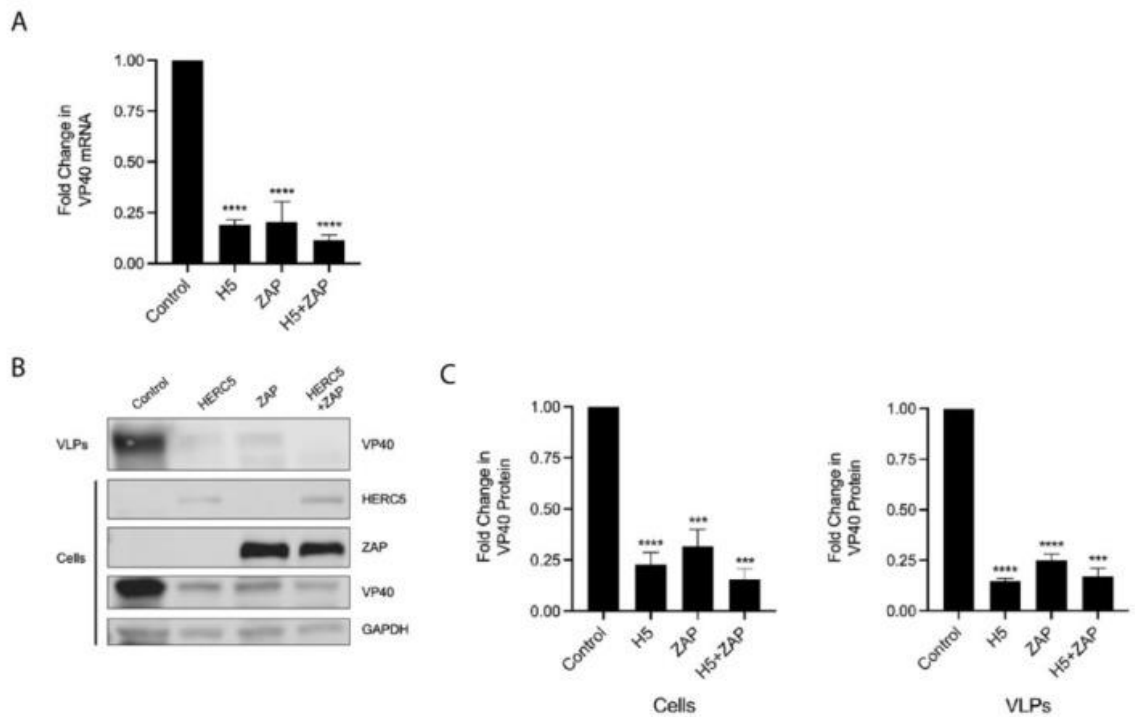

Figure 4.4 HERC5 restricts VP40 independently of ZAP.

293T ZAP knockout cells were co-transfected with plasmids carrying FLAG-tagged VP40 and either empty vector control, HERC5, ZAP (short isoform), or HERC5 and ZAP (short isoform). Twenty-four hours post-transfection, cell lysates and VLP-containing supernatants were harvested. **(A)** Intracellular VP40 mRNA levels were measured using

qPCR and normalized to GAPDH. The data shown is representative of four independent experiments. **(B)** Purified VLPs released into the cell supernatant and intracellular proteins were subjected to Western blot analysis using anti-FLAG and anti-GAPDH. Representative Western blot of three independent experiments is shown. **(C)** The average densitometric quantification of VP40 protein bands from B is shown after normalization to GAPDH levels. Results are presented as mean ($\pm$ SEM) fold changes in VP40 protein or mRNA. **** $p < 0.001$, *** $p < 0.001$, One-way ANOVA with Tukey's multiple comparisons test.

4.2.4 EBOV GP and L Proteins Antagonize HERC5

Despite an early and robust IFN-signaling response to EBOV infection, EBOV proteins ultimately suppress this response leading to pathogenesis[2–12]. Given the potent antiviral activity of HERC5 towards EBOV mRNAs, we asked if any of the EBOV proteins could antagonize this activity. VP40 mRNA levels in cells co-expressing HERC5 and various EBOV proteins were measured by qPCR. As shown in Figure 4.5A, VP40 mRNA levels were rescued in cells co-expressing GP or L protein, but not VP30, VP35, NP or the non-EBOV protein vesicular stomatitis virus-G (VSV-G) protein. Western blot analysis of cell lysates correlated with the qPCR data where only L and GP proteins rescued intracellular VP40 protein levels (Figure 4.5B). Western blot analysis of VP40 VLPs in the supernatant revealed that GP but not L protein rescued VLP production, indicating that only GP was able to fully rescue VLP production.

To determine if the ability of EBOV GP to antagonize HERC5 is specific to the Ebolavirus genus, we tested the ability of Marburg virus (MARV) GP, which belongs to the Marburgvirus genus, to antagonize HERC5. In contrast with EBOV GP, co-expression of MARV GP failed to rescue VP40 VLP production (Figure 4.5C). Together these data show that EBOV GP antagonizes HERC5 activity, and that this antagonism does not appear to be conserved between filovirus genera.
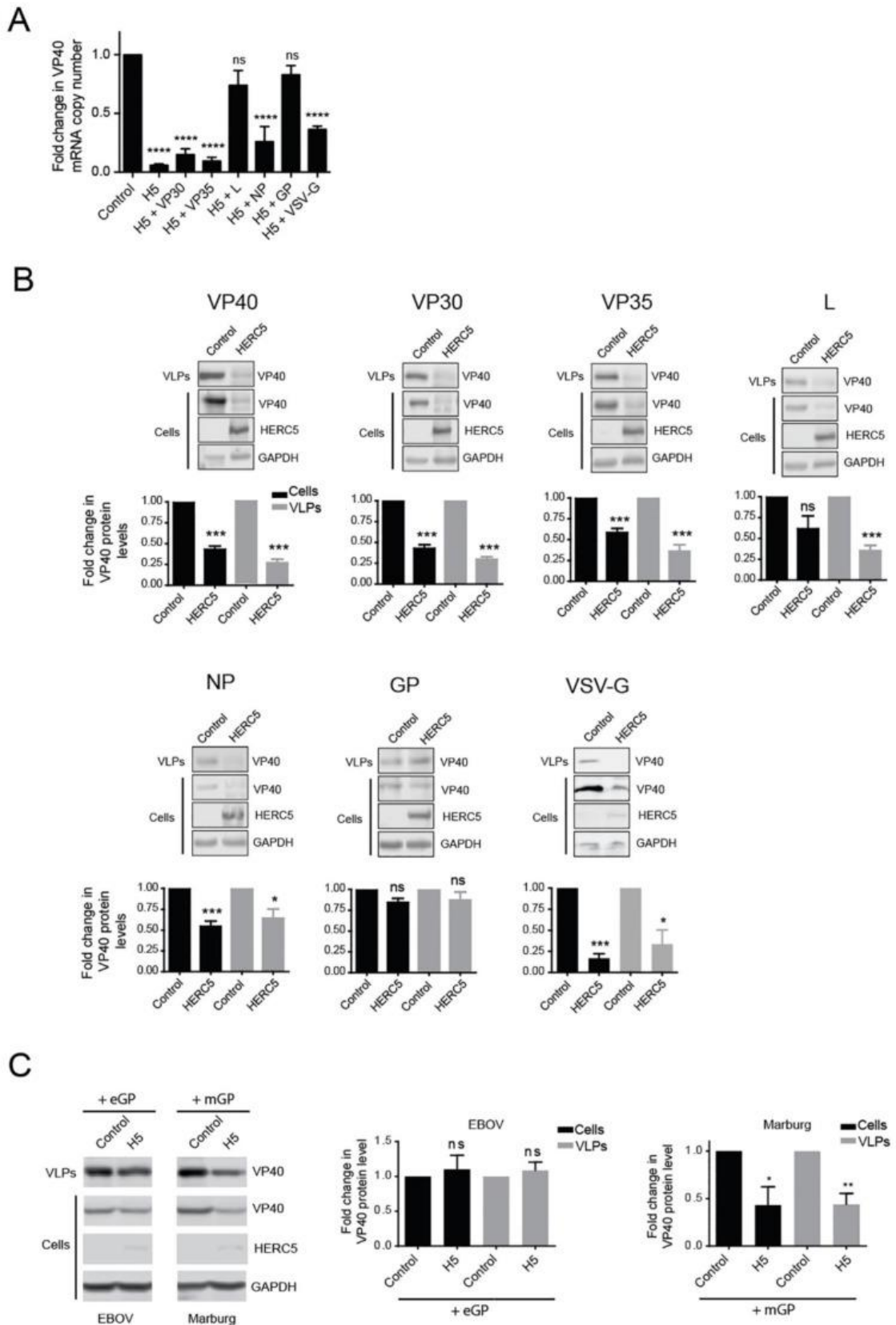
Figure 4.5 EBOV GP and L antagonize HERC5.

293T cells were co-transfected with plasmids carrying FLAG-tagged VP40 and either empty vector or HERC5 and one plasmid carrying either EBOV VP30, VP35, L NP, GP, or VSV-G. Forty-eight hours post-transfection, VP40 mRNA was measured using qPCR **(A)** and VP40 protein levels in cell lysates and VLPs released into supernatant were analyzed by quantitative Western blotting and quantified densitometrically after normalization to GAPDH levels **(B). (C)** 293T cells were co-transfected with plasmids carrying FLAG-tagged VP40 and either empty vector or HERC5, and one of EBOV GP (eGP) or MARV GP (mGP). Forty-eight hours post-transfection, VP40 protein levels in cell lysates and VLPs released into the supernatant were analyzed via Western blotting using anti-FLAG and anti-GAPDH. The data shown represent the average (+/− S.E.M.) of three independent experiments. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, ns (not significant) $p > 0.05$; One-way ANOVA with Dunnett's multiple comparisons test compared to the control (A); Student's paired t-test (B, C).

4.2.5 EBOV and MARV GP Differentially Antagonize HERC5 Inhibition of EBOV trVLP Replication

We utilized the EBOV trVLP system described in Figure 4.1 to determine if genus-specific GP (EBOV or MARV) could antagonize the ability of HERC5 to inhibit trVLP replication. To test the effect of different GPs on trVLP replication, two different sets of trVLP particles were generated at P0. One set contained EBOV GP (trVLP EBOV GP) and was generated as described in Figure 4.1A. The second set was generated in an identical way except that the EBOV GP gene in the '4cis' plasmid minigenome was substituted with the MARV GP gene (trVLP MARV GP). This allowed us to test the impact of different GPs in the VLPs while maintaining the same background of EBOV proteins. As a negative control, the plasmid carrying the Ebola L gene was omitted from the transfections, which abrogates trVLP formation. Compared to the control cells not expressing HERC5, cells expressing HERC5 exhibited significantly reduced levels of trVLP MARV GP and trVLP EBOV GP replication over four passages (spanning 12 days) (Figure 4.6). Notably, HERC5 inhibited trVLP MARV GP replication significantly

more than trVLP EBOV GP replication over two passages ($p < 0.01$, Two-way ANOVA). Together, these data show that EBOV GP and MARV GP differentially antagonize HERC5 inhibition of EBOV trVLP replication.
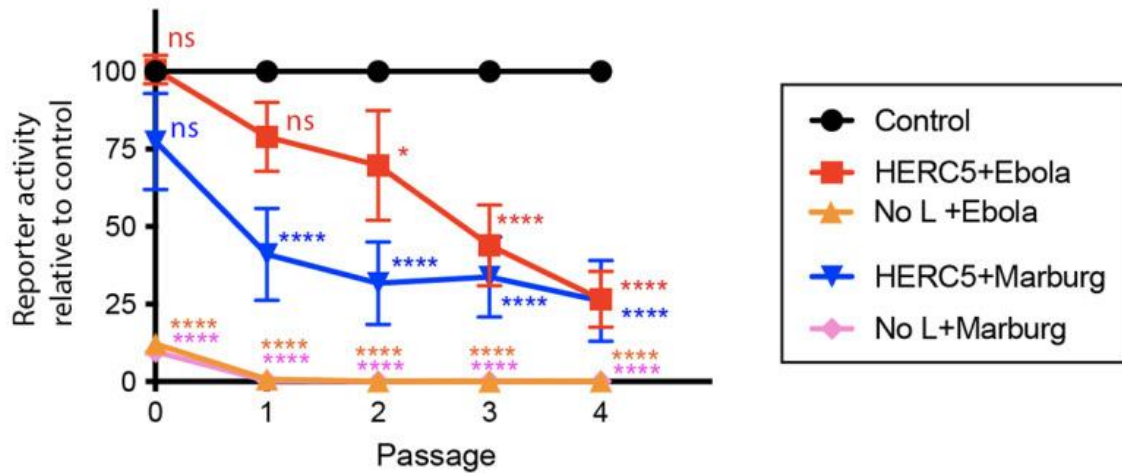

Figure 4.6 EBOV GP and MARV GP differentially antagonize HERC5.

Quantification of trVLP propagation in the presence and absence of HERC5. The trVLP propagation assay was performed using tetracistronic minigenomes carrying a luciferase reporter, EBOV VP40, VP24 and either EBOV GP or MARV GP over four passages (spanning 12 days). All EBOV minigenomes and plasmids carrying the EBOV proteins are based on EBOV H. sapiens-tc/COD/1976/Yambuku-Mayinga. As a negative control ('No L'), the plasmid carrying the Ebola L gene was omitted from the transfections. Luciferase reporter activity relative to the control (trVLPs propagated in the absence of HERC5) is shown. The data shown represent the average ($+/-$ S.E.M.) of at least six independent experiments. **** $p < 0.0001$, * $p \leq 0.05$, ns (not significant) $p > 0.05$; Two-way ANOVA with Dunnett's multiple comparisons test compared to the no HERC5 control.

4.3 Discussion

Hundreds of IFN-induced proteins are part of the early and robust immune response to EBOV infection in primates[2–13]. Characterization of the key effector proteins of this

defense and how EBOV overcomes them will provide a better understanding of the virus–host interactions that occur early in infection. HERC5 is one of the most up-regulated antiviral proteins in the early response to EBOV infection in vivo; however, its role in EBOV replication was previously unknown[3,6,8,28].

In this study, we showed that HERC5 inhibits EBOV VLP replication via a novel E3 ligase-independent mechanism. This mechanism involves the depletion of viral mRNAs and requires the RLD domain of HERC5. We previously showed that HERC5 inhibits the nuclear export of HIV-1 RNA genomes by a different E3 ligase-independent mechanism, one that also requires the RLD domain of HERC5[30]. These E3 ligase-independent antiviral activities, together with its well-documented E3 ligase-dependent antiviral activities[54], identifies HERC5 as a multifunctional antiviral protein. It is perhaps not surprising that HERC5 has evolved multiple mechanisms of restriction of viruses. The ancestral HERC gene is believed to have arisen from a gene fusion event between an RCC1-like gene and a HECT gene[55,56]. This fusion event gave rise to a family of small HERC proteins containing an amino-terminal RLD and a carboxyl-terminal HECT domain that is highly conserved among vertebrates spanning >595 million years of evolution[55–57]. Moreover, HERC5 has been evolving under strong positive selection, which is characteristic of many host restriction factors involved in an evolutionary struggle with viruses[30,56,58]. The ability of HERC5 to inhibit viruses via both E3 ligase-dependent and -independent mechanisms would confer a strong evolutionary advantage to its host, making it more difficult for viruses to evolve countermeasures to HERC5.

Like HERC5, ZAP is present in evolutionarily diverse vertebrates and has evolved under strong positive selection[30,57,59]. ZAP targets diverse viruses such as HIV-1, MoLV and XMRV (Retroviridae), Ebola and Marburg viruses (Filoviridae), alphavirus, Sindbis, Semliki Forest and Ross River viruses (Togaviridae), hepatitis B virus (Hepadnaviridae) and double-stranded DNA murine gamma herpesvirus (Herpesviridae)[45,47,49,51,53,60–64]. ZAP is known to inhibit a wide range of antiviral activities, including recruiting the exosome complex to target viral RNAs for degradation[45,47,49,51,53,60–64]. ZAP also exhibits

virus specificity since it has no antiviral effect on vesicular stomatitis, poliovirus, yellow fever, and herpes simplex I viruses[45]. We showed here that HERC5 depletes EBOV mRNAs in a ZAP-independent manner. Our finding that the HERC5 RLD is necessary and sufficient for EBOV mRNA depletion further supports an E3 ligase-independent mechanism of restriction. It was previously shown that the RLD is required for the association of HERC5 with polyribosomes[37]. It is possible that HERC5 exploits this interaction to recruit other RNA degradation machinery to EBOV mRNAs.

Although we showed that the RLD alone was necessary and sufficient to inhibit particle production, HERC5 lacking the RLD failed to completely inhibit VP40 VLP particle production. Since the RLD is important but not essential for its E3 ligase activity, it is possible that the E3 ligase activity of HERC5 also confers some antiviral activity towards VLP production via ISGylation of viral and/or host proteins involved in particle production[39,65,66]. It was previously shown that over-expression of ISG15 alone inhibited budding of EBOV VP40 VLPs by disrupting Nedd4 function and subsequent ubiquitination of VP40, which is necessary for viral egress[67]. It is unknown whether HERC5 was involved in this activity since it was not investigated. Although our data show that the predominant mechanism by which HERC5 inhibits EBOV VLP production involves the depletion of EBOV mRNAs, visual inspection of cells co-expressing EBOV VP40 and HERC5 by TEM and confocal microscopy revealed an accumulation of the VP40 protein at the localized regions in the plasma membrane in some cells, consistent with the idea of a second mechanism of inhibition acting later in particle production. HERC5-induced trapping of virus particles at the plasma membrane has also been observed with HIV-1[65]. However, it is also possible that these accumulations represent particles in the process of budding that have escaped HERC5 restriction. HERC5 reduced intracellular mRNA levels of viral protein expressed both from a plasmid system (Figure 4.2H) and of viral mRNA expressed from a tetracistronic minigenome. It is unknown how HERC5 can target viral RNAs but not non-viral RNA such as GFP. Perhaps virus-specific RNA sequences recruit HERC5 and/or RNA depletion machinery similar to how

ZAP selectively recognizes high CpG-containing viral RNA. Further studies are needed to decipher this novel antiviral function of HERC5.

Animal model studies have suggested that the Type I IFN response plays an important role in restricting EBOV replication and that the ability of EBOV to overcome this response may be a requirement for lethal infection[68,69]. Although EBOV VP24 and VP35 can act broadly to dampen the IFN response, several IFN-induced antiviral proteins, including HERC5, are also highly upregulated early in response to other stimuli associated with infection, such as pro-inflammatory cytokines[70–72]. As such, it is likely that EBOV evolved additional antagonists of such antiviral proteins. Indeed, EBOV GP can directly antagonize the restriction factor BST-2/tetherin without altering BST-2/tetherin expression levels or cellular localization[73–80]. As shown herein, EBOV GP also antagonizes HERC5 without altering HERC5 expression levels. Although controversial, GP sequence diversity has been shown to affect EBOV transmission and virulence, as demonstrated in the 2013-2016 EBOV epidemic[81,82]. We showed here that variations in GP sequence, such as those found between different filovirus genera (e.g., EBOV and MARV), also influence the potency of antagonism of HERC5 during the early stages of EBOV trVLP replication. It is unclear how GP, which is predominantly localized to the plasma membrane, can rescue EBOV mRNA levels. GP expression is known to alter the expression and trafficking of select cellular proteins; therefore, it is possible that proteins involved in viral RNA stability are affected by GP expression[83–85]. Important next steps will be to characterize the mechanism of GP antagonism and to test the importance of this HERC5-GP axis early in infection using animal models.

It is interesting that EBOV L protein was also able to rescue HERC5-induced VP40 mRNA depletion but unable to antagonize the release of VP40 VLPs into the cell supernatant. The mechanism underlying this antagonism is not fully understood; however, it was previously shown that L protein antagonizes ZAP[48]. It is possible that L protein also specifically antagonizes HERC5-induced depletion of mRNAs. However, we speculate that the E3 ligase activity of HERC5 remains functional, leading to the

ISGylation of viral and/or host proteins and subsequent arrest of later steps in viral particle production.

In conclusion, we showed that HERC5 inhibits EBOV virus particle production by a mechanism involving the depletion of EBOV mRNAs. Our data also identifies a novel 'protagonist–antagonistic' relationship between HERC5 and GP early in EBOV infection. With the ability to inhibit HERC5 and other restriction factors, GP is an attractive target for the development of small molecule compounds that interfere with this antagonism.

4.4   Materials and Methods

4.4.1 Cell Lines

293T and HeLa cells were obtained from American Type Culture Collection. 293T ZC3HAV1 (ZAP) knockout cells were obtained from Dr. Takaoka (Hokkaido University, Japan) via Dr. Li (University of California, Los Angeles, CA, USA) and Dr. MacDonald (The Rockefeller University, New York, NY, USA). Cells were maintained in standard growth medium (Dulbecco's Modified Eagle's Medium (DMEM)), supplemented with 10% heat-inactivated Fetal Bovine Serum (FBS), 100 U/mL Penicillin and 100 μg/mL Streptomycin) at 37 °C with 5% CO_2.

4.4.2 Plasmids, Transfections, Antibodies and Quantitative Western Blotting

Expression plasmids carrying FLAG-tagged HERC5, HERC5-ΔRLD, HERC5-ΔHECT and HERC5-C994A, and HERC4 have been described previously[36]. The plasmid carrying FLAG-tagged RLD only (pFLAG-RLDonly) was generated by standard restriction enzyme cloning of the HERC5 RLD (containing a 3′ stop codon) into p3xFLAG-CMV-10 (Sigma). The promoterless empty vector plasmid pGL3, pEGFP-C1 (pEGFP) and pZAP (short isoform) were obtained from Promega, Clontech and Dharmacon, respectively. pLKO.1/scrambled shRNA and pLKO.1/HERC5 shRNA were previously described[30,65]. VP40 and GP were cloned into p3xFLAG-CMV-10 (Sigma) to generate pFLAG-VP40, pFLAG GP and pEGFP-C1 (containing a CMV promoter)

(Clontech) to generate pVP40-EGFP using standard restriction enzyme cloning. EBOV expression plasmids: pCAGGS plasmids (containing a CMV enhancer, chicken beta-actin promoter and beta-actin intron sequence) carrying only EBOV (Zaire) VP40, VP30, VP35, L, NP, or GP were obtained from Dr. Kawaoka (University of Wisconsin)[86]. Plasmids for the trVLP assay were provided by Dr. Hoenen (Friedrich-Loeffler-Institute, Germany): Plasmids carrying NP, VP35, VP30, L, Tim-1, T7 and the tetracistronic minigenomes (p4cisvRNA-hrLuc, p4cis-vRNA-EGFP) have been previously described [40,87]. All EBOV gene sequences in the minigenomes and plasmids carrying NP, VP35, VP30, and L originated from the Zaire EBOV isolate *H. sapiens*-tc/COD/1976/Yambuku-Mayinga. The EBOV GP and MARV GP expression plasmids were kind gifts of Dr. Cunningham (Brigham and Women's Hospital) [88,89]. Transfections were performed using Lipofectamine 2000 (Invitrogen) per manufacturer's instructions unless otherwise stated. Co-transfections of HERC5 plasmids with pVP40 were performed at a ratio of 10:1, respectively, unless otherwise noted. VP40 VLPs were purified from cell supernatants by centrifugation over a 20% sucrose cushion at $21,000\times g$ for 2 h. Cell lysates and VP40 VLP pellets were subjected to quantitative Western blot analyses using LI-COR, as previously described[65]. Densitometric analysis was performed using ImageJ 1.53e 64-bit version software. Antibodies: Anti-FLAG was purchased from Sigma, anti-ZAP from AbCam (Cat. #ab154680), anti-VP40 from GeneTex (Cat. #GTX134034), anti-MARV GP from Alpha Diagnostic International (Cat. #MVGP12-A), anti-EBOV GP from Bio-Techne (Cat. #MAB9016), anti-β-actin from Rockland, anti-EGFP from Clontech and anti-GAPDH (clone 6C5) from EMD/Millipore.

4.4.3 Confocal Immunofluorescence Microscopy

HeLa cells were cultured in 12-well plates on 18 mm coverslips and co-transfected with either pFLAG-HERC5 and pVP40-EGFP (10:1 ratio) or pGL3 and pVP40-EGFP (10:1 ratio). Twenty-four hours after transfection, the coverslips containing the cells were washed twice with PF buffer (1× PBS + 1% FBS), fixed for 10 min in 1× PBS containing 4% formaldehyde and 2% sucrose, permeabilized in 1× PBS containing 0.1% Triton X 100 (Sigma) and then washed twice more with PF buffer. Coverslips were incubated with

primary antibody rabbit anti-FLAG (1:500 dilution) for 1 h, washed 3× with PF buffer and incubated with either secondary antibody anti-rabbit 594 (1:1000) for 1 h. Coverslips were washed 3×, incubated in Hoechst 33342 (1:10,000 dilution) (Life Technologies) for 5 min and washed 6× with PF buffer. Coverslips were then mounted on glass slides with 10 µL Vectashield mounting media (Vector Laboratories Inc., Burlingame, CA, USA) and sealed with nail polish. Confocal micrographs were obtained using a Leica TCS SP8 (Leica Microsystems) microscope, and Leica Application Software X was used for image acquisition.

4.4.4 Transmission Electron Microscopy

Cells were co-transfected with empty vector or pFLAG-HERC5 and pVP40-EGFP at a 10:1 ratio. After 48 h, cells were resuspended in media, fixed in 2.5% glutaraldehyde in 0.1 M sodium cacodylate (pH 7.4) for 2 h, and washed 3× in 0.1 M sodium cacodylate. Cells were pelleted and fixed with 2% osmium tetroxide in sodium cacodylate. After ~1 h in the dark, cells were washed 3× in ddH$_2$O. Water was discarded, and samples were left at 4 °C overnight. Samples were dehydrated by adding 1 mL 20% acetone in ddH$_2$O, mixed and incubated for 10 min at room temperature. Cells were pelleted, acetone removed, and the procedure was repeated with 50%, 70%, 90%, 100%, 100% and 100% acetone. Cells were embedded in resin by adding 1 mL of a 2:1 mix of acetone: resin (Epon) and incubated for ~4 h at room temperature in a rotating tube shaker. Cells were pelleted, acetone: resin mix was discarded and repeated with a 1:1 mix overnight, 1:2 mix overnight, and finally, resin only overnight. Samples were cut in 70 nm slices using a Sorval Ultracut ultramicrotome and placed onto 400 mesh nickel grids (Embra). Grids were placed on drops of 2% uranyl acetate in ddH$_2$O to stain for 20 min in the dark and washed 5–6× in ddH$_2$O for 1 min. Samples were then stained in drops of Sato's lead citrate (5 mM calcined lead citrate, 11 mM lead nitrate, 11 mM lead acetate, 95 mM sodium citrate) for 1 min and washed using ddH$_2$O. Samples were imaged using a Phillips CM10 Transmission Electron Microscope. The AMT Advantage digital imaging system was used for image acquisition.

4.4.5 Quantitative PCR

The total RNA was extracted using the PureLink RNA mini kit (Ambion, Life Technologies). Using the M-MLV reverse transcriptase and Oligo(dT) primers (Eurofins), 500 ng of RNA was reverse transcribed to cDNA. Prior to qPCR, cDNA samples were diluted 1:5 with water. Each PCR reaction consisted of 10 µL of SYBR Green Master Mix, 1.6 µL of gene-specific primers (0.8 µL of 10 µM forward primer and 0.8 µL of 10 µM reverse primer), 4 µL of diluted cDNA, and water to a total volume of 20 µL. Quantification of endogenous mRNA was run on the QuantStudio5 qPCR machine (Applied Biosystems) under the following cycling conditions: 2 min at 95 °C and 40 cycles of 5 sec at 95 °C, 10 s at 60 °C, and 20 s at 72 °C. The QuantStudio Design and Analysis Desktop Software (version 1.4) was used to determine the C_T for each PCR reaction. Primer pairs are shown in Table 5. To ensure no carry-over of DNA into each total purified RNA sample, 100 ng of RNA was used directly as a template without reverse transcription for qPCR using the primer sets described above.

Table 5 List of primers

HERC5_qpcr	5′ ATG AGC TAA GAC CCT GTT TGG 3′	5′ CCC AAA TCA GAA ACA TAG GCA AG 3′
ZAP_qpcr	5′ CGC TTA ATG GTA GCT GCA GC 3′	5′ CTA CAG AAC AGA GGT GGA TTC C 3′
GAPDH_qpcr	5′ CAT GTT CGT CAT GGG TGT GAA CCA 3′	5′ AGT GAT GGC ATG GAC TGT GGT CAT 3′
eGFP_qpcr	5′ GAC AAC CAC TAC CTG AGC AC 3'	5′ CAG GAC CAT GTG ATC GCG 3′
EBOV VP40	5′ GCT TCC TCT AGG TGT CGC TG 3′	5′ GGT TGC CTT GCC GAA ATG G 3'

EBOV GP	5′ GTG AAT GGG CTG AAA ACT GC 3′	5′ CCG TTC CTG ATA CTT TGT GC 3′
EBOV VP30	5′ CCA GAC AGC ATT CAA GGG 3′	5′ GCT GGA GGA ACT GTT AAT GG 3′
EBOV VP35	5′ CGA CTC AAA ACG ACA GAA TGC 3′	5′ GGT TTG GCT TCG TTT GTT GC 3′
EBOV NP	5′ GCC AAC TTA TCA TAC AGG CC 3′	5′ CCA AAT ACT TGA CTG CGC C 3′
EBOV L	5′ CCT AGT CAC TAG AGC TTG CG 3′	5′ GGC TCA ACA GGA CAG AAT CC 3′

4.4.6 trVLP Assay

Expression plasmids carrying tim-1, T7, NP, VP35, VP30, L, and the tetracistronic minigenome (p4cis-vRNA-hrLuc) carrying luciferase, VP40, GP and VP24 have been previously described[40,90]. trVLP assays were performed as previously described, with the following changes[40,41,88]. Passage zero (p0) cells were seeded in 12-well plates and transfected at 50% confluency using Transit LT-1 (Mirus Bio LLC, Madison, WI, USA) with expression plasmids carrying T7-polymerase (125 ng; all amounts per well), the viral proteins NP (62.5 ng), VP35 (62.5 ng), VP30 (37.5 ng), L (500 ng), a tetracistronic minigenome (125 ng), and Firefly luciferase (100 ng) following the manufacturer's instructions. Twenty-four hours prior to infection of p1, p2, p3 and p4 cells, target cells were pre-transfected with expression plasmids carrying NP (62.5 ng), VP35 (62.5 ng), VP30 (37.5 ng), L (500 ng), Tim-1 (125 ng) and either HERC5 (125 ng) or empty vector (125 ng).

4.4.7 Cell Viability Assay

293T cells were co-transfected with pFLAG-VP40, GFP-VP40 or GFP alone, as well as increasing concentrations of pFLAG-HERC5 or empty vector control plasmid. Forty-eight hours post-transfection Cell Counting Kit-8 (CCK-8) (GLPBIO) was used to measure cell viability as per the manufacturer's instructions.

4.4.8 Statistical Analyses

GraphPad Prism v9 was used for all statistical analyses stated in the text. p values and statistical tests used are stated in the text where appropriate. p values less than 0.05 were deemed significant. Quantification of immunogold labelling for statistical analysis was performed as described[31].

4.5 References

1. Feldmann H, Geisbert TW. Ebola hemorrhagic fever. *The Lancet*. 2011;377(9768):849-862. doi:10.1016/S0140-6736(10)60667-8

2. Rubins KHKH, Hensley LE, Wahl-Jensen V, et al. The temporal program of peripheral blood gene expression in the response of nonhuman primates to Ebola hemorrhagic fever. *Genome Biol*. 2007;8(8). doi:10.1186/GB-2007-8-8-R174

3. Caballero IS, Honko AN, Gire SK, et al. In vivo Ebola virus infection leads to a strong innate response in circulating immune cells. *BMC Genomics*. 2016;17(1). doi:10.1186/S12864-016-3060-0

4. Kash JC, Walters KA, Kindrachuk J, et al. Longitudinal peripheral blood transcriptional analysis of a patient with severe Ebola virus disease. *Sci Transl Med*. 2017;9(385). doi:10.1126/SCITRANSLMED.AAI9321

5. Speranza E, Bixler SL, Altamura LA, et al. A conserved transcriptional response to intranasal Ebola virus exposure in nonhuman primates prior to onset of fever. *Sci Transl Med*. 2018;10(434). doi:10.1126/SCITRANSLMED.AAQ1016

6.   Versteeg K, Menicucci AR, Woolsey C, et al. Infection with the Makona variant results in a delayed and distinct host immune response compared to previous Ebola virus variants. *Sci Rep*. 2017;7(1). doi:10.1038/S41598-017-09963-Y

7.   Cilloniz C, Ebihara H, Ni C, et al. Functional Genomics Reveals the Induction of Inflammatory Response and Metalloproteinase Gene Expression during Lethal Ebola Virus Infection. *J Virol*. 2011;85(17):9060-9068. doi:10.1128/JVI.00659-11

8.   Garamszegi S, Yen JY, Honko AN, et al. Transcriptional Correlates of Disease Outcome in Anticoagulant-Treated Non-Human Primates Infected with Ebolavirus. *PLoS Negl Trop Dis*. 2014;8(7). doi:10.1371/JOURNAL.PNTD.0003061

9.   Speranza E, Altamura LA, Kulcsar K, et al. Comparison of Transcriptomic Platforms for Analysis of Whole Blood from Ebola-Infected Cynomolgus Macaques. *Sci Rep*. 2017;7(1). doi:10.1038/S41598-017-15145-7

10.   Eisfeld AJ, Halfmann PJ, Wendler JP, et al. Multi-platform 'Omics Analysis of Human Ebola Virus Disease Pathogenesis. *Cell Host Microbe*. 2017;22(6):817-829.e8. doi:10.1016/J.CHOM.2017.10.011

11.   Yen JY, Garamszegi S, Geisbert JB, et al. Therapeutics of Ebola hemorrhagic fever: Whole-genome transcriptional analysis of successful disease mitigation. *Journal of Infectious Diseases*. 2011;204(SUPPL. 3). doi:10.1093/INFDIS/JIR345

12.   Liu X, Speranza E, Muñoz-Fontela C, et al. Transcriptomic signatures differentiate survival from fatal outcomes in humans infected with Ebola virus. *Genome Biol*. 2017;18(1). doi:10.1186/S13059-016-1137-3

13.   Kash JC, Mühlberger E, Carter V, et al. Global Suppression of the Host Antiviral Response by Ebola- and Marburgviruses: Increased Antagonism of the Type I Interferon Response Is Associated with Enhanced Virulence. *J Virol*. 2006;80(6):3009-3020. doi:10.1128/JVI.80.6.3009-3020.2006

14.   Hartman AL, Ling L, Nichol ST, Hibberd ML. Whole-Genome Expression Profiling Reveals That Inhibition of Host Innate Immune Response Pathways by Ebola Virus Can Be Reversed by a Single Amino Acid Change in the VP35 Protein. *J Virol*. 2008;82(11):5348-5358. doi:10.1128/JVI.00215-08

15.   Hartman AL, Bird BH, Towner JS, Antoniadou ZA, Zaki SR, Nichol ST. Inhibition of IRF-3 Activation by VP35 Is Critical for the High Level of Virulence of Ebola Virus. *J Virol*. 2008;82(6):2699-2704. doi:10.1128/JVI.02344-07

16.   Prins KC, Delpeut S, Leung DW, et al. Mutations Abrogating VP35 Interaction with Double-Stranded RNA Render Ebola Virus Avirulent in Guinea Pigs. *J Virol*. 2010;84(6):3004-3015. doi:10.1128/JVI.02459-09

17.   Basler CF, Amarasinghe GK. Evasion of interferon responses by ebola and marburg viruses. *Journal of Interferon and Cytokine Research*. 2009;29(9):511-520. doi:10.1089/JIR.2009.0076

18.   Messaoudi I, Amarasinghe GK, Basler CF. Filovirus pathogenesis and immune evasion: Insights from Ebola virus and Marburg virus. *Nat Rev Microbiol*. 2015;13(11):663-676. doi:10.1038/NRMICRO3524

19.   Luthra P, Ramanan P, Mire CE, et al. Mutual antagonism between the ebola virus VP35 protein and the RIG-I activator PACT determines infection outcome. *Cell Host Microbe*. 2013;14(1):74-84. doi:10.1016/J.CHOM.2013.06.010

20.   Bieniasz PD. Intrinsic immunity: A front-line defense against viral attack. *Nat Immunol*. 2004;5(11):1109-1115. doi:10.1038/NI1125

21.   Neil SJD, Zang T, Bieniasz PD. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature*. 2008;451(7177):425-430. doi:10.1038/NATURE06553

22.   van Damme N, Goff D, Katsura C, et al. The Interferon-Induced Protein BST-2 Restricts HIV-1 Release and Is Downregulated from the Cell Surface by the Viral Vpu Protein. *Cell Host Microbe*. 2008;3(4):245-252. doi:10.1016/J.CHOM.2008.03.001

23.   Jouvenet N, Neil SJD, Zhadina M, et al. Broad-Spectrum Inhibition of Retroviral and Filoviral Particle Release by Tetherin. *J Virol*. 2009;83(4):1837-1844. doi:10.1128/JVI.02211-08

24.   Bates P, Kaletsky RL, Francica JR, Agrawal-Gamse C. Tetherin-mediated restriction of filovirus budding is antagonized by the Ebola glycoprotein. *Proc Natl Acad Sci U S A*. 2009;106(8):2886-2891. doi:10.1073/PNAS.0811014106

25.   Sakuma T, Noda T, Urata S, Kawaoka Y, Yasuda J. Inhibition of Lassa, and Marburg Virus Production by Tetherin. *J Virol*. 2009;83(5):2382-2385. doi:10.1128/JVI.01607-08

26.   Radoshitzky SR, Dong L, Chi X, et al. Infectious Lassa Virus, but Not Filoviruses, Is Restricted by BST-2/Tetherin. *J Virol*. 2010;84(20):10569-10580. doi:10.1128/JVI.00103-10

27.   Huang IC, Bailey CC, Weyer JL, et al. Distinct patterns of IFITM-mediated restriction of filoviruses, SARS coronavirus, and influenza A virus. *PLoS Pathog*. 2011;7(1). doi:10.1371/JOURNAL.PPAT.1001258

28.   Wrensch F, Karsten CB, Gnirß K, et al. Interferon-Induced Transmembrane Protein-Mediated Inhibition of Host Cell Entry of Ebolaviruses. *Journal of Infectious Diseases*. 2015;212:S210-S218. doi:10.1093/INFDIS/JIV255

29.   Menicucci AR, Versteeg K, Woolsey C, et al. Transcriptome analysis of circulating immune cell subsets highlight the role of monocytes in Zaire Ebola Virus Makona pathogenesis. *Front Immunol*. 2017;8(OCT). doi:10.3389/FIMMU.2017.01372/PDF

30.   Woods MW, Tong JG, Tom SK, et al. Interferon-induced HERC5 is evolving under positive selection and inhibits HIV-1 particle production by a novel mechanism targeting

Rev/RRE-dependent RNA nuclear export. *Retrovirology*. 2014;11(1). doi:10.1186/1742-4690-11-27

31.   Woods MW, Kelly JN, Hattlmann CJ, et al. Human HERC5 restricts an early stage of HIV-1 assembly by a mechanism correlating with the ISGylation of Gag. *Retrovirology*. 2011;8:95. doi:10.1186/1742-4690-8-95

32.   Tang Y, Zhong G, Zhu L, et al. Herc5 Attenuates Influenza A Virus by Catalyzing ISGylation of Viral NS1 Protein. *The Journal of Immunology*. 2010;184(10):5777-5790. doi:10.4049/JIMMUNOL.0903588

33.   Versteeg GA, Hale BG, van Boheemen S, Wolff T, Lenschow DJ, García-Sastre A. Species-specific antagonism of host ISGylation by the influenza B virus NS1 protein. *J Virol*. 2010;84(10):5423-5430. doi:10.1128/JVI.02395-09

34.   Zhao C, Hsiang TY, Kuo RL, Krug RM. ISG15 conjugation system targets the viral NS1 protein in influenza A virus-infected cells. *Proc Natl Acad Sci U S A*. 2010;107(5):2253-2258. doi:10.1073/PNAS.0909144107

35.   Durfee LA, Huibregtse JM. Identification and Validation of ISG15 Target Proteins. *Subcell Biochem*. 2010;54:228-237. doi:10.1007/978-1-4419-6676-6_18

36.   Paparisto E, Woods MW, Coleman MD, et al. Evolution-guided structural and functional analyses of the HERC family reveal an ancient marine origin and determinants of antiviral activity. *J Virol*. 2018;92(13). doi:10.1128/JVI.00528-18

37.   Durfee LA, Lyon N, Seo K, Huibregtse JM. The ISG15 conjugation system broadly targets newly synthesized proteins: implications for the antiviral function of ISG15. *Mol Cell*. 2010;38(5):722-732. doi:10.1016/j.molcel.2010.05.002

38.   Wong JJY, Pung YF, Sze NSK, Chin KC. HERC5 is an IFN-induced HECT-type E3 protein ligase that mediates type I IFN-induced ISGylation of protein targets. *Proc Natl Acad Sci U S A*. 2006;103(28):10735-10740. doi:10.1073/PNAS.0600397103

39. Dastur A, Beaudenon S, Kelley M, Krug RM, Huibregtse JM. Herc5, an interferon-induced HECT E3 enzyme, is required for conjugation of ISG15 in human cells. *Journal of Biological Chemistry*. 2006;281(7):4334-4338. doi:10.1074/JBC.M512830200

40. Watt A, Moukambi F, Banadyga L, et al. A Novel Life Cycle Modeling System for Ebola Virus Shows a Genome Length-Dependent Role of VP24 in Virus Infectivity. *J Virol*. 2014;88(18):10511-10524. doi:10.1128/JVI.01272-14

41. Hoenen T, Watt A, Mora A, Feldmann H. Modeling the lifecycle of ebola virus under biosafety level 2 conditions with virus-like particles containing tetracistronic minigenomes. *Journal of Visualized Experiments*. 2014;(91). doi:10.3791/52381

42. Frick C, Ollmann-Saphire E, Stahelin R. Live-Cell Imaging of Ebola Virus Matrix Protein VP40. *The FASEB Journal*. 2015;29(S1). doi:10.1096/FASEBJ.29.1_SUPPLEMENT.886.4

43. Noda T, Sagara H, Suzuki E, Takada A, Kida H, Kawaoka Y. Ebola Virus VP40 Drives the Formation of Virus-Like Filamentous Particles Along with GP. *J Virol*. 2002;76(10):4855-4865. doi:10.1128/JVI.76.10.4855-4865.2002

44. Johnson KA, Taghon GJF, Scott JL, Stahelin R v. The Ebola Virus matrix protein, VP40, requires phosphatidylinositol 4,5-bisphosphate (PI(4,5)P 2) for extensive oligomerization at the plasma membrane and viral egress. *Sci Rep*. 2016;6. doi:10.1038/SREP19125

45. Bick MJ, Carroll JWN, Gao G, Goff SP, Rice CM, MacDonald MR. Expression of the Zinc-Finger Antiviral Protein Inhibits Alphavirus Replication. *J Virol*. 2003;77(21):11555-11562. doi:10.1128/JVI.77.21.11555-11562.2003

46. Kerns JA, Emerman M, Malik HS. Positive selection and increased antiviral activity associated with the PARP-containing isoform of human zinc-finger antiviral protein. *PLoS Genet*. 2008;4(1):0150-0158. doi:10.1371/JOURNAL.PGEN.0040021

47. Mao R, Nie H, Cai D, et al. Inhibition of Hepatitis B Virus Replication by the Host Zinc Finger Antiviral Protein. *PLoS Pathog*. 2013;9(7). doi:10.1371/JOURNAL.PPAT.1003494

48. Müller S, Möller P, Bick MJ, et al. Inhibition of Filovirus Replication by the Zinc Finger Antiviral Protein. *J Virol*. 2007;81(5):2391-2400. doi:10.1128/JVI.01601-06

49. Wang X, Tu F, Zhu Y, Gao G. Zinc-finger antiviral protein inhibits XMRV infection. *PLoS One*. 2012;7(6). doi:10.1371/JOURNAL.PONE.0039159

50. Zhang Y, Burke CW, Ryman KD, Klimstra WB. Identification and Characterization of Interferon-Induced Proteins That Inhibit Alphavirus Replication. *J Virol*. 2007;81(20):11246-11255. doi:10.1128/JVI.01282-07

51. Zhu Y, Chen G, Lv F, et al. Zinc-finger antiviral protein inhibits HIV-1 infection by selectively targeting multiply spliced viral mRNAs for degradation. *Proc Natl Acad Sci U S A*. 2011;108(38):15834-15839. doi:10.1073/PNAS.1101676108

52. Li MMH, Aguilar EG, Michailidis E, et al. Characterization of Novel Splice Variants of Zinc Finger Antiviral Protein (ZAP). *J Virol*. 2019;93(18). doi:10.1128/JVI.00715-19

53. Hayakawa S, Shiratori S, Yamato H, et al. ZAPS is a potent stimulator of signaling mediated by the RNA helicase RIG-I during antiviral responses. *Nat Immunol*. 2011;12(1):37-44. doi:10.1038/NI.1963

54. Sánchez-Tena S, Cubillos-Rojas M, Schneider T, Rosa JL. Functional and pathological relevance of HERC family proteins: A decade later. *Cellular and Molecular Life Sciences*. 2016;73(10):1955-1968. doi:10.1007/S00018-016-2139-8

55. Hochrainer K, Mayer H, Baranyi U, Binder BR, Lipp J, Kroismayr R. The human HERC family of ubiquitin ligases: Novel members, genomic organization, expression profiling, and evolutionary aspects. *Genomics*. 2005;85(2):153-164. doi:10.1016/J.YGENO.2004.10.006

56.   Jacquet S, Pontier D, Etienne L. Rapid Evolution of HERC6 and Duplication of a Chimeric HERC5/6 Gene in Rodents and Bats Suggest an Overlooked Role of HERCs in Mammalian Immunity. *Front Immunol*. 2020;11. doi:10.3389/FIMMU.2020.605270

57.   Paparisto E, Woods MW, Coleman MD, et al. Evolution-Guided Structural and Functional Analyses of the HERC Family Reveal an Ancient Marine Origin and Determinants of Antiviral Activity. *J Virol*. 2018;92(13). doi:10.1128/JVI.00528-18

58.   Duggal NK, Emerman M. Evolutionary conflicts between viruses and restriction factors shape immunity. *Nat Rev Immunol*. 2012;12(10):687-695. doi:10.1038/NRI3295

59.   Daugherty MD, Young JM, Kerns JA, Malik HS. Rapid Evolution of PARP Genes Suggests a Broad Role for ADP-Ribosylation in Host-Virus Conflicts. *PLoS Genet*. 2014;10(5). doi:10.1371/JOURNAL.PGEN.1004403

60.   Guo X, Carroll JWN, MacDonald MR, Goff SP, Gao G. The Zinc Finger Antiviral Protein Directly Binds to Specific Viral mRNAs through the CCCH Zinc Finger Motifs. *J Virol*. 2004;78(23):12781-12787. doi:10.1128/JVI.78.23.12781-12787.2004

61.   Zhu Y, Wang X, Goff SP, Gao G. Translational repression precedes and is required for ZAP-mediated mRNA decay. *EMBO Journal*. 2012;31(21):4236-4246. doi:10.1038/EMBOJ.2012.271

62.   Karki S, Li MMH, Schoggins JW, Tian S, Rice CM, MacDonald MR. Multiple interferon stimulated genes synergize with the zinc finger antiviral protein to mediate anti-alphavirus activity. *PLoS One*. 2012;7(5). doi:10.1371/JOURNAL.PONE.0037398

63.   Chen G, Guo X, Lv F, Xu Y, Gao G. p72 DEAD box RNA helicase is required for optimal function of the zinc-finger antiviral protein. *Proc Natl Acad Sci U S A*. 2008;105(11):4352-4357. doi:10.1073/PNAS.0712276105

64. Guo X, Ma J, Sun J, Gao G. The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. *Proc Natl Acad Sci U S A*. 2007;104(1):151-156. doi:10.1073/PNAS.0607063104

65. Woods MW, Kelly JN, Hattlmann CJ, et al. Human HERC5 restricts an early stage of HIV-1 assembly by a mechanism correlating with the ISGylation of Gag. *Retrovirology*. 2011;8:95. doi:10.1186/1742-4690-8-95

66. Durfee LA, Lyon N, Seo K, Huibregtse JM. The ISG15 Conjugation System Broadly Targets Newly Synthesized Proteins: Implications for the Antiviral Function of ISG15. *Mol Cell*. 2010;38(5):722-732. doi:10.1016/J.MOLCEL.2010.05.002

67. Okumura A, Pitha PM, Harty RN. ISG15 inhibits Ebola VP40 VLP budding in an L-domain-dependent manner by blocking Nedd4 ligase activity. *Proc Natl Acad Sci U S A*. 2008;105(10):3974-3979. doi:10.1073/PNAS.0710629105

68. Bray M. The role of the type I interferon response in the resistance of mice of filovirus infection. *Journal of General Virology*. 2001;82(6):1365-1373. doi:10.1099/0022-1317-82-6-1365

69. Ebihara H, Takada A, Kobasa D, et al. Molecular determinants of Ebola virus virulence in mice. *PLoS Pathog*. 2006;2(7):0705-0711. doi:10.1371/JOURNAL.PPAT.0020073

70. Bailey CC, Huang IC, Kam C, Farzan M. Ifitm3 Limits the Severity of Acute Influenza in Mice. *PLoS Pathog*. 2012;8(9). doi:10.1371/JOURNAL.PPAT.1002909

71. Guzzo C, Jung M, Graveline A, Banfield BW, Gee K. IL-27 increases BST-2 expression in human monocytes and T cells independently of type i IFN. *Sci Rep*. 2012;2. doi:10.1038/SREP00974

72. Kroismayr R, Baranyi U, Stehlik C, Dorfleutner A, Binder BR, Lipp J. HERC5, a HECT E3 ubiquitin ligase tightly regulated in LPS activated endothelial cells. *J Cell Sci*. 2004;117(20):4749-4756. doi:10.1242/JCS.01338

73.   Kühl A, Banning C, Marzi A, et al. The Ebola virus glycoprotein and HIV-1 VPU employ different strategies to counteract the antiviral factor tetherin. *Journal of Infectious Diseases*. 2011;204(SUPPL. 3). doi:10.1093/INFDIS/JIR378

74.   Lopez LA, Yang SJ, Hauser H, et al. Ebola Virus Glycoprotein Counteracts BST-2/Tetherin Restriction in a Sequence-Independent Manner That Does Not Require Tetherin Surface Removal. *J Virol*. 2010;84(14):7243-7255. doi:10.1128/JVI.02636-09

75.   Lopez LA, Yang SJ, Exline CM, Rengarajan S, Haworth KG, Cannon PM. Anti-Tetherin Activities of HIV-1 Vpu and Ebola Virus Glycoprotein Do Not Involve Removal of Tetherin from Lipid Rafts. *J Virol*. 2012;86(10):5467-5480. doi:10.1128/JVI.06280-11

76.   González-Hernández M, Hoffmann M, Brinkmann C, et al. A GXXXA Motif in the Transmembrane Domain of the Ebola Virus Glycoprotein Is Required for Tetherin Antagonism. *J Virol*. 2018;92(13). doi:10.1128/JVI.00403-18

77.   vande Burgt NH, Kaletsky RL, Bates P. Requirements within the Ebola viral glycoprotein for tetherin antagonism. *Viruses*. 2015;7(10):5587-5602. doi:10.3390/V7102888

78.   Gustin JK, Bai Y, Moses A v., Douglas JL. Ebola Virus Glycoprotein Promotes Enhanced Viral Egress by Preventing Ebola VP40 from Associating with the Host Restriction Factor BST2/Tetherin. *Journal of Infectious Diseases*. 2015;212:S181-S190. doi:10.1093/INFDIS/JIV125

79.   Brinkmann C, Nehlmeier I, Walendy-Gnirß K, et al. The Tetherin Antagonism of the Ebola Virus Glycoprotein Requires an Intact Receptor-Binding Domain and Can Be Blocked by GP1-Specific Antibodies. *J Virol*. 2016;90(24):11075-11086. doi:10.1128/JVI.01563-16

80.   Gnirß K, Fiedler M, Krämer-Kühl A, et al. Analysis of determinants in filovirus glycoproteins required for tetherin antagonism. *Viruses*. 2014;6(4):1654-1671. doi:10.3390/V6041654

81.   Wang MK, Lim SY, Lee SM, Cunningham JM. Biochemical Basis for Increased Activity of Ebola Glycoprotein in the 2013–16 Epidemic. *Cell Host Microbe*. 2017;21(3):367-375. doi:10.1016/J.CHOM.2017.02.002

82.   Marzi A, Chadinah S, Haddock E, et al. Recently Identified Mutations in the Ebola Virus-Makona Genome Do Not Alter Pathogenicity in Animal Models. *Cell Rep*. 2018;23(6):1806-1816. doi:10.1016/J.CELREP.2018.04.027

83.   Sullivan NJ, Peterson M, Yang Z yong, et al. Ebola Virus Glycoprotein Toxicity Is Mediated by a Dynamin-Dependent Protein-Trafficking Pathway. *J Virol*. 2005;79(1):547-553. doi:10.1128/JVI.79.1.547-553.2005

84.   Iampietro M, Younan P, Nishida A, et al. Ebola virus glycoprotein directly triggers T lymphocyte death despite of the lack of infection. *PLoS Pathog*. 2017;13(5). doi:10.1371/JOURNAL.PPAT.1006397

85.   Stewart CM, Phan A, Bo Y, et al. Ebola virus triggers receptor tyrosine kinase-dependent signaling to promote the delivery of viral particles to entry-conducive intracellular compartments. *PLoS Pathog*. 2021;17(1). doi:10.1371/JOURNAL.PPAT.1009275

86.   Watanabe S, Watanabe T, Noda T, et al. Production of Novel Ebola Virus-Like Particles from cDNAs: an Alternative to Ebola Virus Generation by Reverse Genetics. *J Virol*. 2004;78(2):999-1005. doi:10.1128/JVI.78.2.999-1005.2004

87.   Wendt L, Kamper L, Schmidt ML, Mettenleiter TC, Hoenen T. Analysis of a Putative Late Domain Using an Ebola Virus Transcription and Replication-Competent Virus-Like Particle System. *Journal of Infectious Diseases*. 2018;218:S355-S359. doi:10.1093/INFDIS/JIY247

88.   Côté M, Misasi J, Ren T, et al. Small molecule inhibitors reveal Niemann-Pick C1 is essential for Ebola virus infection. *Nature*. 2011;477(7364):344-348. doi:10.1038/NATURE10380

89.   Chandran K, Sullivan NJ, Felbor U, Whelan SP, Cunningham JM. Endosomal Proteolysis of the Ebola Virus Glycoprotein Is Necessary for Infection. *Science (1979)*. 2005;308(5728):1643 LP - 1645. doi:10.1126/science.1110656

90.   Schmidt ML, Hoenen T. Characterization of the catalytic center of the Ebola virus L polymerase. *PLoS Negl Trop Dis*. 2017;11(10). doi:10.1371/JOURNAL.PNTD.0005996

5   Discussion

This dissertation highlights the potency of the small HERC family of proteins in combating evolutionarily diverse families of viruses. This ongoing battle between restriction factors and viruses has shaped this group of proteins and their function. Although the small HERC family of proteins has been implicated in regulation of cell senescence[1,2], cell proliferation[3], autophagy[4], tissue growth and remodeling[5], regulation of nuclear export[6,7], spermatogenesis[8], as well as several pro and anti-cancer properties[4,9–15], their antiviral properties, especially those of HERC3 and HERC4, are less well characterized in humans. Here, we build upon the current knowledge of the antiviral characteristics of the small HERC family of protein with special emphasis on the interferon induced members HERC5 and HERC6. This investigation led to the discovery of a single nucleotide polymorphism in HERC6 which inhibits the production of infections HIV particles, similar to HERC5, the most potent antiviral member of the small HERCs. We found that HERC5 reduces intracellular levels of EBOV VP40 RNA, a mechanism which is antagonized by EBOV GP. This work showcases the small HERCs as a functionally important family of proteins which contribute to the antiviral cellular state. Specifically, it advances the knowledge of HERC5 and HERC6 and their role in antiviral immunity against HIV and EBOV.

5.1   Summary of Findings

Believed to have originated from a gene fusion event which linked an RCC1-like domain (RLD) with a HECT domain, the small HERC family of proteins began with HERC4 over 595 mya. In Chapter 2 we investigated the evolutionary origin and antiviral activities of these proteins to define the protection they provide against pathogens. Based on the available genome assemblies of vertebrates we estimate the emergence of HERC4 was followed by HERC3 approximately 476 mya, HERC6 emerged approximately 430 mya and HERC5 approximately 413 mya (Figure 2.1). Our analysis suggests that the small HERC family has undergone gene duplication events, chromosomal rearrangements, and potential gene loss events over the course of vertebrate evolution.

Closer examination of the amino acid homology of the RCC1-like and HECT domains of the small HERC family revealed that although the amino acid homology is rather low, their predicted structural homology is highly conserved (Figure 2.2). Next, we tested the ability of the small HERC proteins to inhibit HIV. Although we found that HERC3 and HERC4 are not strongly induced by IFN β (Appendix 7) they maintain some inhibitory function at higher levels of expression (Figure 2.3 F and G). As expected, based on the previous findings of the Barr lab[6,16], HERC5 potently inhibits intracellular Gag and released viral particles (Figure 2.3 F). Knocking down HERC5, but not HERC3 and HERC4 allowed for enhanced HIV particle production (Figure 2.3 D and E). Interestingly, HERC6, HERC5s closest relative, did not significantly inhibit HIV particle production or infectivity (Figure 2.3 F and G). This result was further investigated in Chapter 3 of the thesis. Continuing with the exploration of the small HERC family we found that all four members inhibit Rev dependent nuclear export of Gag RNA to varying extents with HERC5 being the most potent (Figure 2.4). To determine whether the antiviral activity of HERC5 has an ancient origin we used a coelacanth HERC5 construct, the oldest vertebrate in which HERC5 was identified. We found that while coelacanth HERC5 and HERC6 cannot inhibit HIV, they retain some antiviral activity against SIV, exhibiting differentia species- and specific- antiviral activity (Figure 2.5). Further investigation of HERC6 revealed that several codons are evolving under strong positive selection (Figure 2.6 C) and that blade 1 of HERC6 is an important determinant of anti-HIV activity (Figure 2.7).

The discrepancy between the antiviral activity of HERC5 and HERC6 while being the most closely related members of the small HERC family was intriguing and led us to investigate the structural differences between these proteins. By comparing the structure of the RLD domain of HERC5 and HERC6 we discovered that blade 1 of HERC6 is folded toward blade 2 while blade 1 of HERC5 is folded away from the rest of the protein structure, a structural change alleviated when the arginine at position 10 and glutamic acid at position 64 were mutated to alanine (Figure 3.1). Interestingly, a small section of the human population possesses a SNP at position 10 of HERC6 (rs111670008) which

substitutes the wildtype arginine with a glycine (R10G). This polymorphism was associated with slower disease progression in a cohort of HIV infected individuals in Uganda (Figure 3.2). We found that overexpression of HERC6 containing the R10G SNP resulted in less infectious HIV particles similar to the effect observed with overexpression of HERC5 (Figure 3.3 C). No difference was observed between HERC6 and HERC6 R10G in Gag content, both intracellular and released in viral particles; viral RNA in the released particles; and nuclear export of unspliced RNA. To better understand the mechanism by which HERC6 R10G overexpression restricts production of infectious HIV particles we performed mass spectrometry analysis on particles produces in cells expressing either HERC6 or HERC6 R10G. GO analysis revealed that in the particles released from the cells expressing HERC6 R10G contained less proteins related to viral transcription, viral gene expression and cotranslational protein targeting to the membrane than those expressing wildtype HERC6 (Figure 3.4 D). Our next step was to determine if attachment of viral particles produced in HERC6 R10G cells was affected compared to those produced in HERC6 expressing cells. Since CADM1, a cell attachment molecule was significantly reduced in the particles isolated from HERC6 R10G cell (Figure 3.4 A) we wondered if neutralization of CADM1 would replicate the loss of infectivity we observed. Unfortunately, treatment with the CADM1 neutralizing antibody did not perform better than the isotype control, suggesting that CADM1 proteins are not important for the infectivity of HIV particles. While this result is not conclusive, it is consistent with other research which has not identified it as an important host protein incorporated in HIV particles and affecting HIV infectivity[17–24]. Further research is needed to understand the influence of HERC6 R10G on particle infectivity.

Focusing on HERC5 and defining its antiviral properties we asked if HERC5 can inhibit the matrix protein of Ebola virus similar to its ability to inhibit HIV Gag protein[16]. While previous studies found that ISGylation of Ebola matrix protein VP40 lead to a decrease in the release of virus-like particles from the cells[25,26] we aimed to determine the role of HERC5, if any, in the inhibition of Ebola virus. To that end we used a trVLP assay to investigate the effect of HERC5 overexpression on EBOV particle production. Compared

to control cells, HERC5 expression reduced the infectivity of the produced virions over time, an effect correlated with a reduction in GP and VP40 (Figure 4.1). We found that increased amounts of HERC5 reduced both intracellular VP40 protein and virus like particles produced (Figure 4.2 A, D, E and F). Noticing that intracellular VP40 protein levels were reduced, a phenotype not expected if ISGylation of VP40 is inhibiting its ubiquitination and budding from the plasma membrane, we asked if RNA levels of VP40 are unchanged. Quantitative PCR showed that VP40 levels were significantly reduced in HERC5 overexpressing cells compared to control, an effect also observed on other EBOV proteins to a lesser extent (Figure 4.2 H). Using HERC5 domain truncations we determined that the RLD domain of HERC5 is necessary and sufficient for HERC5 mediated restriction of VP40 at both the RNA and protein levels (Figure 4.3). Since the level of VP40 in our trVLP assay did not coincide with the drastic reduction in VP40 protein observed in our RT-qPCR and western blot assays we asked if EBOV proteins may be antagonizing HERC5. We found that both GP and to a lesser extent L protein antagonized HERC5 mediated restriction of VP40 (Figure 4.5). This antagonistic effect was genus specific, as the MARV GP did not inhibit HERC5 mediated restriction of VP40 (Figure 4.4 C and Figure 4.5). These results highlight the multifaceted way HERC5 inhibits EBOV protein expression and the antagonism which does not allow this restriction to combat Ebola virus infections *in vivo*.

5.2   Future directions

The goal of scientific research is to systematically find answers. However, more often than not, research leads to new questions rather than answers. While in this thesis I have illuminated certain aspects of the small HERC family and their workings, several new questions require attention. The world of the small HERC family and their diverse functions remains a largely unexplored one. Several of their functions are inferred from RNA sequencing during various disease states and their contribution to said state and mechanism of action remains a mystery[27–39]. While these studies are important to focus the research on the proteins most affected during the disease states it is important for the development of cures to have a thorough understanding of the workings of these proteins.

Several questions remain unanswered about HERC3 and HERC4 and their ability to inhibit viral pathogens. Although they are not interferon induced in humans it is likely their expression modulated by the state of the tissue where they are expressed. For example, Senegalese sole and Atlantic cod, HERC4 expression is induced by viral infection and dsRNA [40–42]. Direct testing of the effect these proteins have against diverse viruses will reveal the protection they provide.

Along the same lines HERC6 research has largely been focused on mice where it acts as the main E3 ligase for ISG15. The function of HERC6 independent of ISGylation is likely involved in male fertility in mice but further studies are needed to determine the mechanism of action or whether this translates to human biology [43]. Additionally, the antiviral properties of HERC6 in humans are understudied. Recently, HERC6 has been identified as an upregulated protein in several transcriptomic studies of viral infection in humans and other mammals [44–48]. Several questions remain unknown about whether this is linked to the enzymatic activities of the HECT domain of HERC6 or whether, as we found in Chapter 3, this is linked to the RLD domain and its antiviral functions. Further studies are needed to explore the effect of HERC6 R10G and the mechanism of inhibition. Based on our finding in Chapter 3 we suspect that HIV virions produced in HERC6 R10G expressing cells are missing infectivity factors required for the viral infectivity cycle following cell entry. Further studies are needed to identify the mechanism of inhibition as well as the infectivity factor or factors which HERC6 R10G effects. While we did not find that neutralization of CADM1 resulted in reduced HIV infectivity the ability of HERC6 to affect this attachment molecule may have significant implications in a different tissue where both of these molecules are highly expressed, the testis. CADM1 is believed to play a role in the normal differentiation of spermatids and spermatocytes into mature spermatozoa, including in their adhesion to Sertoli cells [49-52] while HERC6 knockout mice had severe defects in their sperm sac morphology [8]. Further studies are required to discern if the effect of these protein on male fertility is related.

To better understand the breadth of HERC5s antiviral activity, we researched its ability to inhibit EBOV and its structural protein VP40. Further studies are needed to determine the protagonist-antagonist relationship of HERC5 with other pathogenic viruses. HERC5 is a broadly acting antiviral protein with multiple mechanisms of viral protein restriction. HERC5 can be used as a template for the development of small molecule inhibitors and mimics with the capacity to fight a multitude of viruses both known and yet undiscovered. The use of one antiviral molecule as opposed to the induction of the interferon system can avoid the systemic inflammation and cell damage associated with dysregulation of the interferon response. Knowledge of HERC5s inhibitory mechanisms as well as its antagonism by viral proteins can inform future drug development.

5.3 Concluding Remarks

Enhanced awareness of the small HERC family of proteins will aid in the understanding of diverse cellular function in both healthy and diseased states. Further research implementing novel technologies such as RNA sequencing, single cell transcriptomic and proteomic profiling, as well as high-throughput high-resolution mass spectrometry will enhance the comprehension of this fascinating and important family of proteins and how their regulation effects cellular processes. Knowledge regarding the antiviral activities of the small HERC family of proteins from the effect of HERC4 on fish immunity to the understanding of HERC5 and HERC6 regulation and antagonism by various viruses can lead to the development of small molecule inhibitors and mimics which can aid the immune system in combating viral infection.

5.4 References

1.   Chen Y, Li Y, Peng Y, et al. DeltaNp63alpha down-regulates c-Myc modulator MM1 via
      E3 ligase HERC3 in the regulation of cell senescence. *Cell Death Differ*.
      2018;25(12):2118-2129. doi:10.1038/s41418-018-0132-5

2.   Zheng X, Chen L, Jin S, et al. Ultraviolet B irradiation up-regulates MM1 and induces
      photoageing of the epidermis. *Photodermatol Photoimmunol Photomed*. 2021;37:395-
      403. doi:10.1111/phpp.12670

3.   Zheng Y, Li J, Pan C, et al. HERC4 Is Overexpressed in Hepatocellular Carcinoma and
      Contributes to the Proliferation and Migration of Hepatocellular Carcinoma Cells. *DNA
      Cell Biol*. 2017;36(6):490-500. doi:10.1089/dna.2016.3626

4.   Li H, Li J, Chen L, et al. HERC3-Mediated SMAD7 Ubiquitination Degradation
      Promotes Autophagy-Induced EMT and Chemoresistance in Glioblastoma. *Clin Cancer
      Res*. 2019;25(12):3602-3616. doi:10.1158/1078-0432.CCR-18-3791

5.   Aerne BL, Gailite I, Sims D, Tapon N. Hippo Stabilises Its Adaptor Salvador by
      Antagonising the HECT Ubiquitin Ligase Herc4. *PLoS One*. 2015;10(6):e0131113.
      doi:10.1371/journal.pone.0131113

6.   Woods MW, Tong JG, Tom SK, et al. Interferon-induced HERC5 is evolving under
      positive selection and inhibits HIV-1 particle production by a novel mechanism targeting
      Rev/RRE-dependent RNA nuclear export. *Retrovirology*. 2014;11:27. doi:10.1186/1742-
      4690-11-27

7.   Hochrainer K, Pejanovic N, Olaseun VA, Zhang S, Iadecola C, Anrather J. The ubiquitin
      ligase HERC3 attenuates NF-kappaB-dependent transcription independently of its
      enzymatic activity by delivering the RelA subunit for degradation. *Nucleic Acids Res*.
      2015;43(20):9889-9904. doi:10.1093/nar/gkv1064

8.   Arimoto K ichiro, Hishiki T, Kiyonari H, et al. Murine Herc6 Plays a Critical Role in Protein ISGylation In Vivo and Has an ISGylation-Independent Function in Seminal Vesicles. *J Interferon Cytokine Res*. 2015;35(5):351-358. doi:10.1089/jir.2014.0113

9.   Sala-Gaston J, Martinez-Martinez A, Pedrazza L, et al. HERC Ubiquitin Ligases in Cancer. *Cancers (Basel)*. 2020;12(6). doi:10.3390/cancers12061653

10.   Zhou H, Shi R, Wei M, Zheng WL, Zhou JY, Ma WL. The expression and clinical significance of HERC4 in breast cancer. *Cancer Cell Int*. 2013;13(1):113. doi:10.1186/1475-2867-13-113

11.   Mao X, Sethi G, Zhang Z, Wang Q. The Emerging Roles of the HERC Ubiquitin Ligases in Cancer. *Curr Pharm Des*. 2018;24(15):1676-1681. doi:10.2174/1381612824666180528081024

12.   Zeng WL, Chen YW, Zhou H, Zhou JY, Wei M, Shi R. Expression of HERC4 in lung cancer and its correlation with clinicopathological parameters. *Asian Pac J Cancer Prev*. 2015;16(2):513-517. doi:10.7314/apjcp.2015.16.2.513

13.   Zhu L, Wu J, Liu H. Downregulation of HERC5 E3 ligase attenuates the ubiquitination of CtBP1 to inhibit apoptosis in colorectal cancer cells. *Carcinogenesis*. 2021;42(8):1119-1130. doi:10.1093/carcin/bgab053

14.   Wrage M, Hagmann W, Kemming D, et al. Identification of HERC5 and its potential role in NSCLC progression. *Int J Cancer*. 2015;136(10):2264-2272. doi:10.1002/ijc.29298

15.   Wei M, Zhang YL, Chen L, Cai CX, Wang HD. [RNA interference of HERC4 inhibits proliferation, apoptosis, and migration of cervical cancer Hela cells]. *Nan Fang Yi Ke Da Xue Xue Bao*. 2016;37(2):232-237.

16.   Woods MW, Kelly JN, Hattlmann CJ, et al. Human HERC5 restricts an early stage of HIV-1 assembly by a mechanism correlating with the ISGylation of Gag. *Retrovirology*. 2011;8:95. doi:10.1186/1742-4690-8-95

17.   Munoz O, Banga R, Perreau M. Host Molecule Incorporation into HIV Virions, Potential Influences in HIV Pathogenesis. *Viruses*. 2022;14(11):2523. doi:10.3390/V14112523

18.   Saphire ACS, Gallay PA, Bark SJ. Proteomic analysis of human immunodeficiency virus using liquid chromatography/tandem mass spectrometry effectively distinguishes specific incorporated host proteins. *J Proteome Res*. 2006;5(3):530-538. doi:10.1021/PR050276B/ASSET/IMAGES/LARGE/PR050276BF00004.JPEG

19.   Linde ME, Colquhoun DR, Mohien CU, et al. The Conserved Set of Host Proteins Incorporated into HIV-1 Virions Suggests a Common Egress Pathway in Multiple Cell Types. *J Proteome Res*. 2013;12(5):2045. doi:10.1021/PR300918R

20.   Chertova E, Chertov O, Coren L v., et al. Proteomic and Biochemical Analysis of Purified Human Immunodeficiency Virus Type 1 Produced from Infected Monocyte-Derived Macrophages. *J Virol*. 2006;80(18):9039. doi:10.1128/JVI.01013-06

21.   Chan EY, Sutton JN, Jacobs JM, Bondarenko A, Smith RD, Katze MG. Dynamic Host Energetics and Cytoskeletal Proteomes in Human Immunodeficiency Virus Type 1-Infected Human Primary CD4 Cells: Analysis by Multiplexed Label-Free Mass Spectrometry. *J Virol*. 2009;83(18):9283-9295. doi:10.1128/JVI.00814-09/SUPPL_FILE/TABLE_S2.XLS

22.   Brégnard C, Zamborlini A, Leduc M, et al. Comparative Proteomic Analysis of HIV-1 Particles Reveals a Role for Ezrin and EHD4 in the Nef-Dependent Increase of Virus Infectivity. *J Virol*. 2013;87(7):3729-3740. doi:10.1128/JVI.02477-12/SUPPL_FILE/ZJV999097418SO1.XLS

23.   Santos S, Obukhov Y, Nekhai S, Bukrinsky M, Iordanskiy S. Virus-producing cells determine the host protein profiles of HIV-1 virion cores. *Retrovirology*. 2012;9(1):1-28. doi:10.1186/1742-4690-9-65/FIGURES/4

24. Ott DE, Coren L v., Johnson DG, et al. Actin-Binding Cellular Proteins inside Human Immunodeficiency Virus Type 1. *Virology*. 2000;266(1):42-51. doi:10.1006/VIRO.1999.0075

25. Malakhova OA, Zhang DE. ISG15 inhibits Nedd4 ubiquitin E3 activity and enhances the innate antiviral response. *J Biol Chem*. 2008;283(14):8783-8787. doi:10.1074/jbc.C800030200

26. Okumura A, Pitha PM, Harty RN. ISG15 inhibits Ebola VP40 VLP budding in an L-domain-dependent manner by blocking Nedd4 ligase activity. *Proc Natl Acad Sci U S A*. 2008;105(10):3974-3979. doi:10.1073/pnas.0710629105

27. Paradis F, Yue S, Grant JR, Stothard P, Basarab JA, Fitzsimmons C. Transcriptomic analysis by RNA sequencing reveals that hepatic interferon-induced genes may be associated with feed efficiency in beef heifers. *J Anim Sci*. 2015;93(7):3331-3341. doi:10.2527/JAS.2015-8975

28. Thery F, Eggermont D, Impens F. Proteomics Mapping of the ISGylation Landscape in Innate Immunity. *Front Immunol*. 2021;12:3089. doi:10.3389/FIMMU.2021.720765/BIBTEX

29. Liu X, Speranza E, Muñoz-Fontela C, et al. Transcriptomic signatures differentiate survival from fatal outcomes in humans infected with Ebola virus. *Genome Biol*. 2017;18(1):4. doi:10.1186/s13059-016-1137-3

30. al Kalaldeh M, Gibson J, Lee SH, Gondro C, van der Werf JHJ. Detection of genomic regions underlying resistance to gastrointestinal parasites in Australian sheep. *Genet Sel Evol*. 2019;51(1). doi:10.1186/S12711-019-0479-1

31. Kim JJ, Lee YHK& MY, *. Proteomic characterization of differentially expressed proteins associated with no stress in retinal ganglion cells. *BMB Rep*. 2009;42(7):456-461. Accessed January 4, 2022. https://www.bmbreports.org/journal/view.html?spage=456&volume=42&number=7

32. Cheng J, Cao X, Hanif Q, et al. Integrating Genome-Wide CNVs Into QTLs and High Confidence GWAScore Regions Identified Positional Candidates for Sheep Economic Traits. *Front Genet*. 2020;11. doi:10.3389/FGENE.2020.00569

33. Khuder SA, Al-Hashimi I, Mutgi AB, Altorok N. Identification of potential genomic biomarkers for Sjögren's syndrome using data pooling of gene expression microarrays. *Rheumatol Int*. 2015;35(5):829-836. doi:10.1007/s00296-014-3152-6

34. Dai L, Bai L, Lin Z, et al. Transcriptomic analysis of KSHV-infected primary oral fibroblasts: The role of interferon-induced genes in the latency of oncogenic virus. *Oncotarget*. 2016;7(30):47052-47060. doi:10.18632/oncotarget.9720

35. Ariyarathne HBPC, Correa-Luna M, Blair HT, Garrick DJ, Lopez-Villalobos N. Identification of Genomic Regions Associated with Concentrations of Milk Fat, Protein, Urea and Efficiency of Crude Protein Utilization in Grazing Dairy Cows. *Genes (Basel)*. 2021;12(3). doi:10.3390/GENES12030456

36. Al-Mamun HA, Kwan P, Clark SA, Ferdosi MH, Tellam R, Gondro C. Genome-wide association study of body weight in Australian Merino sheep reveals an orthologous region on OAR6 to human and bovine genomic regions affecting height and weight. *Genet Sel Evol*. 2015;47(1). doi:10.1186/S12711-015-0142-4

37. Forde N, Duffy GB, McGettigan PA, et al. Evidence for an early endometrial response to pregnancy in cattle: both dependent upon and independent of interferon tau. *Physiol Genomics*. 2012;44(16):799-810. doi:10.1152/PHYSIOLGENOMICS.00067.2012

38. Rise ML, Hall JR, Rise M, et al. Impact of asymptomatic nodavirus carrier state and intraperitoneal viral mimic injection on brain transcript expression in Atlantic cod (Gadus morhua). *Physiol Genomics*. 2010;42(2):266-280. doi:10.1152/physiolgenomics.00168.2009

39.   Ohnishi J, Ayuzawa S, Nakamura S, et al. Distinct transcriptional and metabolic profiles associated with empathy in Buddhist priests: A pilot study. *Hum Genomics*. 2017;11(1). doi:10.1186/s40246-017-0117-3

40.   Gémez-Mata J, Labella AM, Bandín I, Borrego JJ, García-Rosado E. Immunogene expression analysis in betanodavirus infected-Senegalese sole using an OpenArray® platform. *Gene*. 2021;774:145430. doi:10.1016/J.GENE.2021.145430

41.   Eslamloo K, Xue X, Booman M, Smith NC, Rise ML. Transcriptome profiling of the antiviral immune response in Atlantic cod macrophages. *Dev Comp Immunol*. 2016;63:187-205. doi:10.1016/j.dci.2016.05.021

42.   Eslamloo K, Inkpen SM, Rise ML, Andreassen R. Discovery of microRNAs associated with the antiviral immune response of Atlantic cod macrophages. *Mol Immunol*. 2018;93:152-161. doi:10.1016/j.molimm.2017.11.015

43.   Lu MY, Huang CI, Hsieh MY, et al. Dynamics of PBMC gene expression in hepatitis C virus genotype 1-infected patients during combined peginterferon/ribavirin therapy. *Oncotarget*. 2016;7(38):61325-61335. doi:10.18632/oncotarget.11348

44.   Suen WW, Imoda M, Thomas AW, et al. An Acute Stress Model in New Zealand White Rabbits Exhibits Altered Immune Response to Infection with West Nile Virus. *Pathogens*. 2019;8(4). doi:10.3390/PATHOGENS8040195

45.   Li HK, Kaforou M, Rodriguez-Manzano J, et al. Discovery and validation of a three-gene signature to distinguish COVID-19 and other viral infections in emergency infectious disease presentations: a case-control and observational cohort study. *Lancet Microbe*. 2021;2(11):e594. doi:10.1016/S2666-5247(21)00145-2

46.   Dong Z, Yan Q, Cao W, Liu Z, Wang X. Identification of key molecules in COVID-19 patients significantly correlated with clinical outcomes by analyzing transcriptomic data. *Front Immunol*. 2022;13. doi:10.3389/FIMMU.2022.930866

47.   Surace EI, Strickland A, Hess RA, Gutmann DH, Naughton CK. Tslc1 (nectin-like molecule-2) is essential for spermatozoa motility and male fertility. *J Androl*. 2006;27(6):816-825. doi:10.2164/JANDROL.106.000398

48.   van der Weyden L, Arends MJ, Chausiaux OE, et al. Loss of TSLC1 causes male infertility due to a defect at the spermatid stage of spermatogenesis. *Mol Cell Biol*. 2006;26(9):3595-3609. doi:10.1128/MCB.26.9.3595-3609.2006

49.   Wakayama T, Ohashi K, Mizuno K, Iseki S. Cloning, and characterization of a novel mouse immunoglobulin superfamily gene expressed in early spermatogenic cells. *Mol Reprod Dev*. 2001;60(2):158-164. doi:10.1002/MRD.1072

50.   Yamada D, Yoshida M, Williams YN, et al. Disruption of spermatogenic cell adhesion and male infertility in mice lacking TSLC1/IGSF4, an immunoglobulin superfamily cell adhesion molecule. *Mol Cell Biol*. 2006;26(9):3610-3624. doi:10.1128/MCB.26.9.3610-3624.2006

51.   Terada N, Ohno N, Saitoh S, et al. Involvement of a membrane skeletal protein, 4.1G, for Sertoli/germ cell interaction. *Reproduction*. 2010;139(5):883-892. doi:10.1530/REP-10-0005

52.   Wakayama T, Iseki S. Role of the spermatogenic-Sertoli cell interaction through cell adhesion molecule-1 (CADM1) in spermatogenesis. *Anat Sci Int*. 2009;84(3):112-121. doi:10.1007/S12565-009-0034-1

Appendix

Appendix 1: A portion of section 1.3.5 was previously published in Viruses, a MDPI open access journal. Chapter 4 was published in Cells, a MDPI open access journal.

Mathieu NA, Paparisto E, Barr SD, Spratt DE. HERC5 and the ISGylation Pathway: Critical Modulators of the Antiviral Immune Response. Viruses. 2021 Jun 9;13(6):1102. doi: 10.3390/v13061102

Paparisto E, Hunt NR, Labach DS, Coleman MD, Di Gravio EJ, Dodge MJ, Friesen NJ, Côté M, Müller A, Hoenen T, Barr SD. Interferon-Induced HERC5 Inhibits Ebola Virus Particle Production and Is Antagonized by Ebola Glycoprotein. Cells. 2021 Sep 13;10(9):2399. doi: 10.3390/cells10092399.

Excerpt from https://journals.asm.org/author-self-archiving-permissions

Authors May Republish/Adapt Portions of Their Articles

ASM also grants authors the right to republish discrete portions of their article in any other publication (including print, CD-ROM, and other electronic formats), provided that proper credit is given to the original ASM publication. ASM authors also retain the right to reuse the full article in their dissertation or thesis. "Proper credit" means the copyright notice shown on the initial page of the article.

Appendix 2: Chapter 2 was published in Journal of Virology, an ASM journal.

Paparisto E, Woods MW, Coleman MD, Moghadasi SA, Kochar DS, Tom SK, Kohio HP, Gibson RM, Rohringer TJ, Hunt NR, Di Gravio EJ, Zhang JY, Tian M, Gao Y, Arts EJ, Barr SD. Evolution-Guided Structural and Functional Analyses of the HERC Family Reveal an Ancient Marine Origin and Determinants of Antiviral Activity. J Virol. 2018 Jun 13;92(13):e00528-18. doi: 10.1128/JVI.00528-18.
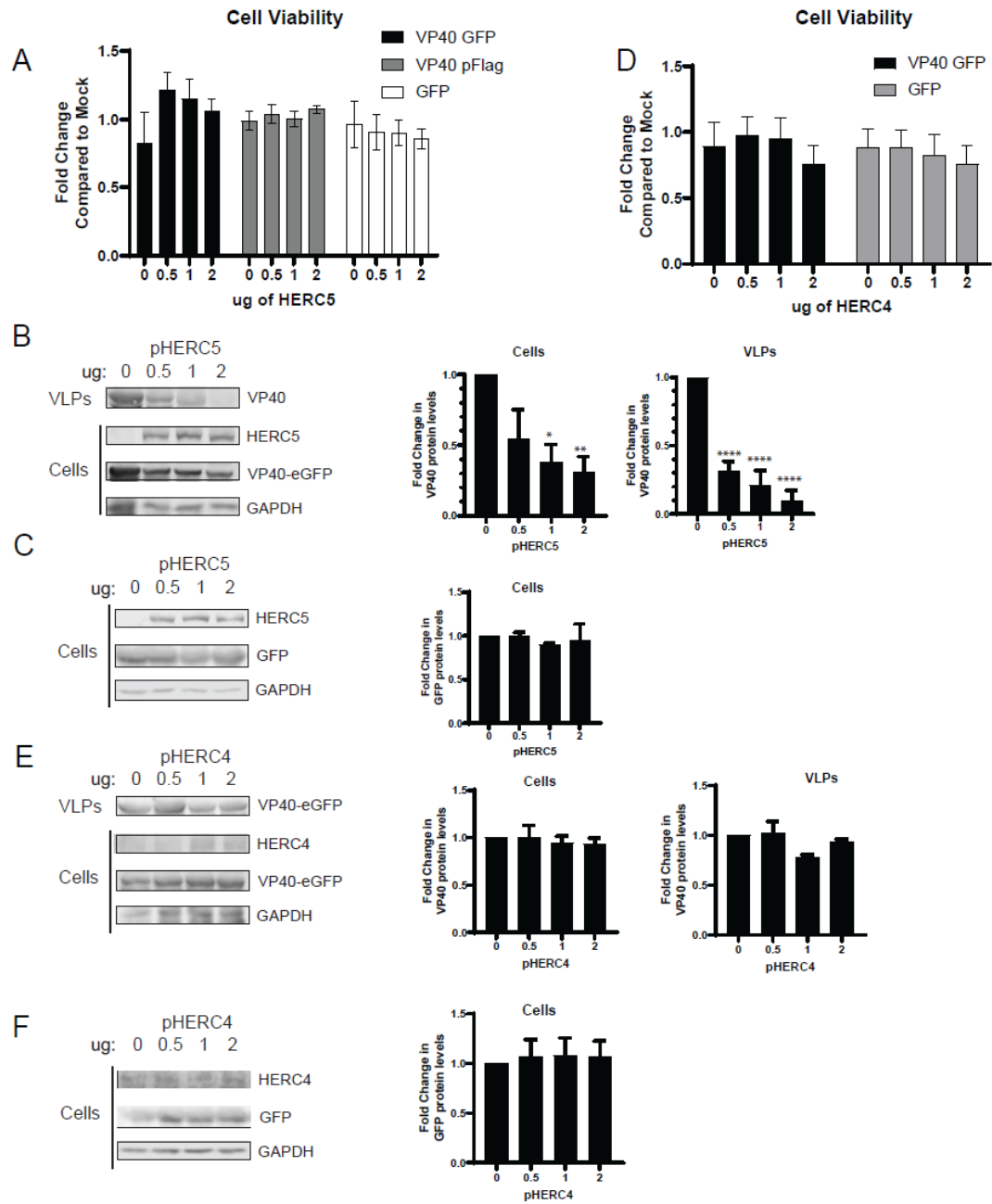
Appendix 3: HERC5 but not HERC4 depletes FLAG-tagged VP40 protein in a dose dependent manner

A) 293T cells were co-transfected with plasmids carrying FLAG-tagged VP40 (pFLAG-VP40), VP40-eGFP or just GFP plasmid and increasing concentrations of FLAG-tagged HERC5 (pFLAG-HERC5). Empty vector plasmid was transfected in the condition with 0 ug of HERC5 plasmid and used to ensure equal amounts of DNA were transfected in each condition Forty-eight hours post-transfection cell viability was measured using the CCK8 assay. Fold change in absorbance compared to untransfected cells was calculated. The results of 4 independent experiments are shown **(B)** 293T cells were co-transfected with plasmids carrying VP40-eGFP or just GFP plasmid and increasing concentrations of FLAG-tagged HERC4 (pFLAG-HERC4). Forty-eight hours post-transfection cell viability was measured using CCK8 assay. Fold change in absorbance compared to untransfected cells was calculated. Results of 4 independent experiments are shown. **(C and D)** Cells were transfected as in (A) purified VLPs released into the cell supernatant and intracellular protein were subjected to Western blot analysis using anti-FLAG, anti-VP40 and anti-GAPDH. The average densitometric quantification of VP40 protein bands is shown to the right after normalization to GAPDH levels (+/- S.E.M.). Representative Western blot of four independent experiments is shown. **(E and F)** Cells were transfected as in (B) purified VLPs released into the cell supernatant and intracellular protein were subjected to Western blot analysis using anti-HERC4, anti-VP40 and anti-GAPDH. The average densitometric quantification of VP40 protein bands is shown to the right after normalization to GAPDH levels (+/- S.E.M.). Representative Western blot of four independent experiments is shown.
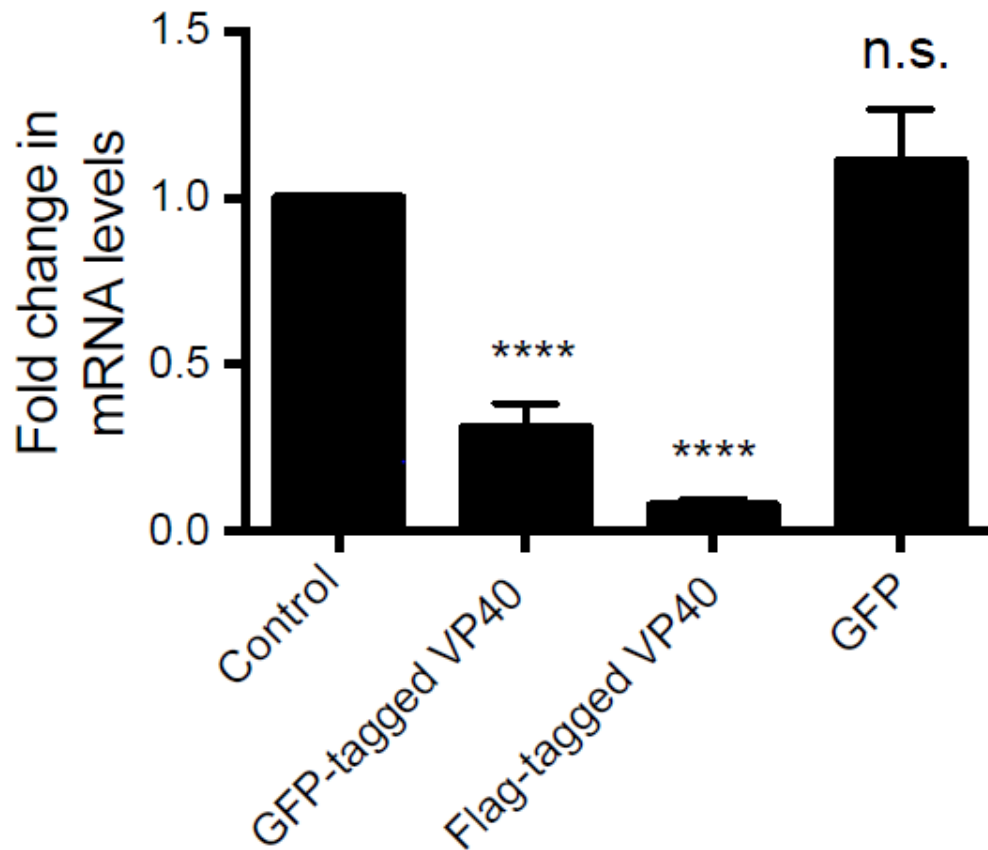
Appendix 4: HERC5 depletes GFP- and FLAG-tagged VP40 mRNA but not GFP mRNA

293T cells were transfected with HERC5 or control plasmid and GFP-tagged VP40, Flag-tagged VP40, or GFP plasmid. mRNA levels for VP40 and GFP were measured 24 hours post transfection using qRT-PCR. The data shown represent the average (+/− S.E.M.) of 4 independent experiments. **** $p < 0.0001$, * $p \leq 0.05$, ns (not significant) $p > 0.05$

Region	Observed gold count, G_0	Point count, P	Expected gold count, G_e	G_0/G_e	X^2	X^2 as %
Particles + plasma membrane	30	15	4	7.75	176.43	66.94
Cytoplasm + nucleus	35	49	13	0.71	39.62	15.03
Non-particle + non-cell	1	192	53	0.02	47.52	18.02
TOTAL	70	256	70		263.57	100

Appendix 5: Quantification of 5nm gold particle-labeled anti-GFP in cells expressing HERC5 and VP40-eGFP.

For X2= 176.43 and df=2, P <0.0001 (X2 analysis). The gold labeling distribution is significantly different from random. Only the particles + plasma membrane region (Go/Ge= 7.75, X2= 66.9% of total) meets the two criteria for being preferentially labeled ((Go/Ge) $\geq$ 1 and X2 > 10% of total).
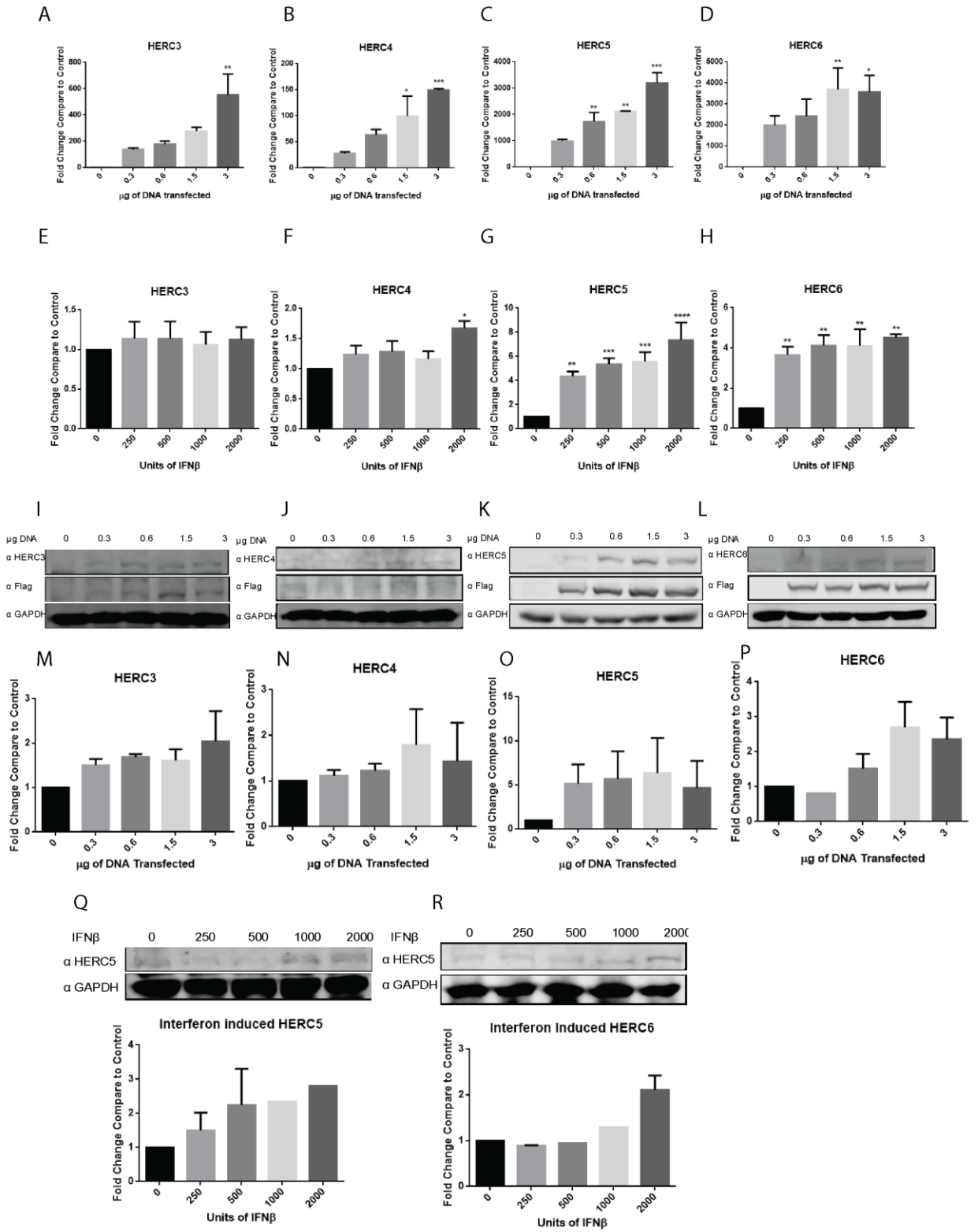
Appendix 6: Quantification of Small HERC mRNA and protein expression when overexpressed through transfection or induced by IFNβ.

293T cells were either transfected **(A to D and I to P)** with increasing concentrations of HERC plasmid as indicated or treated with increasing amounts of IFNβ **(E to H, Q and R)**. mRNA levels (A to H) were measured via qRT-PCR. I to L, Q and R, shows representative western blot of protein expression. (**M to P**) quantification of 3 trials of the HERC protein.

Curriculum Vitae

Name: Ermela Paparisto

Post-secondary Education and Degrees:
University of Toronto (Scarborough)
Toronto, Ontario, Canada
2009-2013 Honors BSc. Human Biology and Psychology

The University of Western Ontario
London, Ontario, Canada
2014-2023 Ph.D. Microbiology & Immunology

Honors and Awards:
RGE Murray Scholarship in Microbiology and Immunology
2017-2019

Graduate (>6 Months) Poster Presentation Award
Infection and Immunity Research Forum
2015

Related Work Experience
Teaching Assistant (MNI 2500, MNI3820, BIO1001)
The University of Western Ontario
2015-2019

Let's Talk Science Representative
The University of Western Ontario
2014-2019

Infection and Immunity Research Forum Organizational
Committee Chair
The University of Western Ontario
2015-2016

Graduate Studies Committee Student Representative
Department of Microbiology and Immunology
The University of Western Ontario
2017-2019

Graduate Student Representative to the Appointments Committee
Department of Microbiology and Immunology
The University of Western Ontario
2016-2019

Publications:

Paparisto E, Woods MW, Coleman MD, Moghadasi SA, Kochar DS, Tom SK, Kohio HP, Gibson RM, Rohringer TJ, Hunt NR, Di Gravio EJ, Zhang JY, Tian M, Gao Y, Arts EJ, Barr SD. **Evolution-Guided Structural and Functional Analyses of the HERC Family Reveal an Ancient Marine Origin and Determinants of Antiviral Activity**. J Virol. 2018 Jun 13;92(13):e00528-18. doi: 10.1128/JVI.00528-18. PMID: 29669830; PMCID: PMC6002735.

Paparisto E, Hunt NR, Labach DS, Coleman MD, Di Gravio EJ, Dodge MJ, Friesen NJ, Côté M, Müller A, Hoenen T, Barr SD. **Interferon-Induced HERC5 Inhibits Ebola Virus Particle Production and Is Antagonized by Ebola Glycoprotein**. Cells. 2021 Sep 13;10(9):2399. doi: 10.3390/cells10092399. PMID: 34572049; PMCID: PMC8472148.

Mathieu NA, **Paparisto E**, Barr SD, Spratt DE. **HERC5 and the ISGylation Pathway: Critical Modulators of the Antiviral Immune Response**. Viruses. 2021 Jun 9;13(6):1102. doi: 10.3390/v13061102. PMID: 34207696; PMCID: PMC8228270.

Ajoge HO, Kohio HP, **Paparisto E**, Coleman MD, Wong K, Tom SK, Bain KL, Berry CC, Arts EJ, Barr SD. **G-Quadruplex DNA and Other Non-Canonical B-Form DNA Motifs Influence Productive and Latent HIV-1 Integration and Reactivation Potential.** Viruses. 2022 Nov 11;14(11):2494. doi: 10.3390/v14112494. PMID: 36423103; PMCID: PMC9692945.