# MECHANISMS UNDERLYING EXOSOMAL TRAFFICKING OF THE ANTIRETROVIRAL FACTOR APOBEC3G: IMPLICATIONS FOR THE CONTAINMENT AND ERADICATION OF HIV-1 INFECTION

**AFONSO MACHADO MENDES**

**A dissertation submitted in partial fulfillment of the requirements for the Degree of Masters in Biomedical Research**

***Dissertação para obtenção do grau de Mestre em Investigação Biomédica***

**at Faculdade de Ciências Médicas | NOVA Medical School of NOVA University Lisbon**

**September 2019**

# MECHANISMS UNDERLYING EXOSOMAL TRAFFICKING OF THE ANTIRETROVIRAL FACTOR APOBEC3G: IMPLICATIONS FOR THE CONTAINMENT AND ERADICATION OF HIV-1 INFECTION

**Afonso Machado Mendes**

**Supervisors: Paulo C. Pereira & Vasco M. Barreto, Principal Investigators at the Center for Chronic Diseases (CEDOC/NOVA Medical School)**

**A dissertation submitted in partial fulfillment of the requirements for the Degree of Masters in Biomedical Research**

***Dissertação para obtenção do grau de Mestre em Investigação Biomédica***

**September, 2019**

## ABSTRACT

The first reports of patients with acquired immunodeficiency syndrome (AIDS) and the discovery of the Human Immunodeficiency Virus type 1 (HIV-1) as the causative agent of AIDS occurred around three decades ago. Africa is, by far, the most affected region in the world, registering over 1 million new infections in 2017. Despite the enormous progress made with the introduction of relatively effective preventive and therapeutic interventions, 37 million people are still infected and without a cure. The natural antiretroviral protein APOBEC3G catalyzes C-to-U deamination in the viral DNA and is capable of restricting HIV-1 replication in the absence of Vif, a viral protein that interacts with APOBEC3G and promotes its ubiquitylation and subsequent degradation the 26S proteasome, rendering the natural antiviral response incapable of restricting the replication of wild-type HIV-1. The presence of APOBEC3G in the exosomes (small nano-sized extracellular vesicles originated in multivesicular bodies (MVBs) and secreted by fusion of the MVBs with the cell membrane) of several cell lines has been shown. Furthermore, A3G-containing exosomes were shown to effectively restrict HIV-1 replication in cell models. However, the mechanisms underlying the import of A3G into exosomes remain unclear. In this study, the exosomes of the HEK293T cell line were found to be enriched in a EGFP-tagged version of A3G when compared to EGFP, higher than those of EGFP, a protein that is not selectively incorporated in exosomes. Furthermore, we show for the first that this enrichment depends on the presence of a specific amino acid sequence in the primary structure of A3G called the KFERQ-like motif. In the future, these mechanisms may be used to design new strategies in preventing, treating and even curing HIV-1 infection.

## PREFACE

This study is above all the materialization of multiple new friendships. About two years ago, Dr. Vasco M. Barreto (friendship #1) taught me for the first time about the existence of a natural DNA-editing enzyme (Activation-Induced Deaminase, or AID). In simplistic terms, by catalyzing a single point mutation (C-to-U) in genomic DNA, AID would conduct the activity of an authentic orchestra of DNA-repairing enzymes and cell-cycle regulators which, by acting upon the DNA on the antibody Immunoglobulin (Ig) locus, ultimately resulted in random genomic rearrangements and culminated in the "wonderful sound" of a B-cell producing an affinity-maturated antibody with improved function. In parallel, Dr. Paulo C. Pereira (friendship #2) convinced me that most, if not all cells, continuously secrete (among other things) a population of nano-sized vesicles called exosomes, with enormous overall biomedical and scientific potential regarding their use as biomarkers and delivery vectors for several molecules. He was particularly interested in a specific amino acid sequence (the KFERQ-like motif) that seemed to induce selective inclusion of soluble proteins into exosomes. These apparently unrelated subjects quickly grew on me.

On a certain day, Dr. Manuel Vicente (friendship #3), with whom I had immediately connected ever since my first stay at Dr. Vasco Barreto's lab shared most of my scientific thoughts and interests, gave me an unexpected call. He told me about APOBEC3 proteins, natural antiretroviral enzymes belonging to the same family as AID. He knew about my deep interest regarding exosomes, and so he sent me a study that he had stumbled upon, describing the presence of APOBEC3G in the exosomes of several cell lines and showing that these exosomes were capable of restricting HIV-1 in infected cells from AIDS patients (1). He suggested I could further investigate this matter. I must confess, when I saw the two subjects ("cytosine deaminases" and "exosomes") together, I was immediately ecstatic, but I was unsure about how to properly tackle the subject, since so many elaborate studies had been made in the fields of HIV and cellular trafficking. In a posterior article from the same authors (2), I found the following sentence: - "Another unanswered question is how A3 proteins are packaged into exosomes." – I had found my MSc thesis project. Thus, I designed a project intentionally seeking to join the laboratories of Drs. Paulo Pereira and Vasco Barreto. I quickly applied for a rotation at Dr. Paulo Pereira's lab, where I learned to purify exosomes from the supernatants of cell cultures and to analyze them by Western blot. I was privileged to have my project accepted by Paulo and Vasco, who agreed to join efforts. Since then, many projects for the future have been emerging between all of us. I hope I can continue to learn by exploring these ideas with the help and guidance of my supervisors.

This manuscript is entirely dedicated to the people with whom I've had the privilege of sharing space-time during the last two years: to everyone that was part of my surroundings, who knowingly or unknowingly

played an important role in defining the setting under which I was fortunate to develop myself through the present study. The works in this here presented were developed thanks to the unconditional support of my two dear supervisors, Dr. Paulo C. Pereira and Dr. Vasco M. Barreto. An equal contribution was given by the resident "Post-Docs" and Ph.D. students of both laboratories, Dr. Ana Soares and Dr. Nadiya Kubasova (Ph.D. students and friendships #4 and #5), and Dr. João Ferreira and Dr. João Proença (resident "Post-Docs" and friendships #6 and #7). Finally, and because the last to be mentioned is often the most remembered, a huge thanks to Manuel Vicente for his irreplaceable mentoring and friendship.

# INTRODUCTION

## *The rise of the HIV/AIDS epidemic*

In the late 1970s, the main Public Health institutions of the United States of America received an abnormally high number of reports regarding patients who displayed unusual opportunistic infectious diseases, mainly of Kaposi's sarcoma (KS), *Pneumocystis carinii* pneumonia (PCP) and extensive mucosal candidiasis (3). The Centers for Disease Control and Prevention (CDC) had only received a single request for pentamidine (a drug typically used to treat PCP) between 1976 and 1980, while in 1982 there were already at least 42 requests. These cases were also of high severity: around 40% of PCP patients died from initial infections, and around 20% of the survivors died soon after due to recurrent infections. In 1981, the CDC created a team dedicated to the surveillance of this outbreak through epidemiologic and laboratory investigations.

The first epidemiologic studies revealed that the vast majority of patients were men averaging 35 years of age. Most of them were white, but several were Hispanic or black. Haitian immigrants who had traveled to the USA recently seemed to be particularly affected. Notably, around 92% of the patients were either homosexual or bisexual, and all of them had been healthy before suddenly developing refrained immune responses (anergy) and contracting opportunistic infections known to be usually fought by cell-mediated immunity. Such an outbreak of these rare diseases was unlikely to happen; they were usually linked to elders and seldom caused death (4).

Medical doctors and scientists were quick to point that the incidence of opportunistic infections was known to be higher in patients with compromised immune systems, such as those receiving immunosuppressive therapy (3). Therefore, it was hypothesized that the outbreak of opportunistic infections could indicate an epidemic of underlying immunosuppression caused by an unknown source or agent. Thus, the immune functions of these patients were tested routinely. Testing consistently showed severe defects in T-cell

numbers (lymphopenia), as well as diminished T-cell proliferative responses to previously established mitogens (e.g., phytohemaglutinin (PHA)) and antigens. Furthermore, it appeared that CD4+ helper T-cells were the most affected T-cell subset by this newly acquired immunodeficiency syndrome (AIDS) was the CD4+ helper subset, since the CD8+ suppressor subset showed either normal or increased levels of cell numbers. Thus, AIDS patients showed a markedly decreased T4 helper/T8 suppressor cell ratio.

The socio-cultural characteristics of AIDS patients increased the difficulty in understanding the etiology of the syndrome because most patients hosted more than one infection and reported an overall incautious lifestyle. Many possible confounding infections were ruled out, such as CMV and Epstein-Barr virus (EBV). In parallel, laboratory tests continued to show that decreased T-cell numbers and proliferative responses to mitogens and antigens were the hallmarks of AIDS, hence why serious opportunistic infections would flourish in these patients. However, how AIDS spread between individuals was still not understood (5).

***Increased incidence in homosexuals and IV drug-abusers***

In July of 1981, the CDC reported that the majority of patients affected by AIDS were homosexual men with a history of IV drug-abuse, who had multiple sexual partners in a short period. Moreover, statistical analyses showed that variables such as “meeting random sexual partners in bathhouses” and behaviors related to anal sex were strikingly significant. The sexual orientation of AIDS patients was likely not a coincidence, as for many years immunosuppressive sexually-transmitted diseases (STDs) such as cytomegalovirus (CMV) and herpes simplex virus (HSV) infections also prevailed in this group (and still do). Notably, studies done using mice had shown that infection with CMV increased immunosuppression and mortality from doses of bacteria or fungus that are usually non-lethal. *In vitro*, human CMV was also shown to decrease cellular immune function by decreasing lymphocyte proliferative responses and interferon production in response to mitogens and antigens. Accordingly, many of these mysterious patients at that time tested positive for CMV. Although CMV infection with subsequent immunosuppression remained a candidate cause, the severity of this new epidemic was far higher than any reported case of CMV (6).

***Transmission by transfusion of blood and blood products***

In 1982, a 20-month-old infant from San Francisco was hospitalized due to opportunistic infections and unexplained immunodeficiency (7). A series of complications during birth, including asphyxia with endotracheal intubation and hyperbilirubinemia led to multiple blood transfusions. Although the infant was discharged after one month of treatment, he was hospitalized again three months later due to enlarged liver and spleen (hepatosplenomegaly), and acute infections (including extensive oral candidiasis) that persisted after antibiotic therapy. At 9 months old, the child developed anorexia, vomiting and jaundice. He

tested negative for hepatitis A and B, as well as CMV, and was diagnosed with "non-A non-B hepatitis." This patient continued to develop other symptoms, including high levels of immunoglobulins (IgG, IgA and IgM), low T-cell numbers and decreased T-cell function and proliferative responses *in vitro*. Meanwhile, it was confirmed that one of the 19 blood donors who contributed to the transfusions given to the infant had later been diagnosed with AIDS and died tragically after experiencing severe illness. In another instance, the CDC had received three reports of patients with both PCP and hemophilia A, of which two had died and one was critically ill. These patients had received multiple injections of irradiated factor VIII concentrate (a blood product) due to their hemophilia and started to develop AIDS-like symptoms (8). These reports, and those showing increased rates of AIDS among adult IV drug abusers, suggested that AIDS was a blood-borne infectious disease that could be transmitted by blood transfusion (7).

### *Transmission to infants during pregnancy or birth*

In parallel to the ongoing AIDS epidemic in adults, the CDC received several reports of infants less than 2 years of age with unexplained cellular immunodeficiency and opportunistic infections. Most of them displayed hepatosplenomegaly, bacterial and fungal infections, high immunoglobulin levels, low T-cell count and impaired proliferative responses of T-cells to mitogens and antigens. Unlike the "San Francisco baby" mentioned previously (7), none of these cases were known to have received blood transfusions; however, their parents' health was either unknown or they had a history of IV drug abuse and possibly AIDS. These reports raised the possibility that the causative agent of AIDS, if any, could have been transmitted from mother to child during pregnancy, or during birth, by the transmission of blood (9–11).

### *Isolation of the causative agent of AIDS*

Although some evidence suggested that AIDS patients also showed abnormalities in B-cell function (12), it was clear that the main hallmark of AIDS was a markedly decreased T4 helper/T8 suppressor cell ratio. Fahey et al. (13) determined that the main factor influencing this ratio was a decrease in T4 helper cells, rather than an increase in the T8 suppressor subset; the latter, although detected, occurred more likely due to the simultaneous incidence of Kaposi's sarcoma and opportunistic infections in AIDS patients.

At this point, several hints pointed towards the causative agent of AIDS being a blood borne retrovirus of the Human T-cell Leukemia Virus (HTLV), which had been previously isolated from patients with T-cell leukemia and that, similarly to AIDS, could be transmitted by intimate contact and blood products (14). Serological and epidemiological studies showed that HTLV was endemic in the Caribbean regions (15). Accordingly, Haiti, a Caribbean country, had been shown to be a possible focus of AIDS, many of the first North American AIDS patients being Haitian immigrants. Moreover, the members of the HTLV family of retroviruses displayed a tropism for CD4+ T-cells (14,16,17) and induced many cytopathic changes also

observed in T-cells of AIDS patients, such as syncytia formation (18) and cell death (19). Finally, another retrovirus, the feline leukemia virus, was known to cause immune deficiency in cats (20–22).

In 1983, future Nobel Laureates Barré-Sinoussi and Montagnier, along with their colleagues, reported the isolation of a novel retrovirus from a lymph node of an AIDS patient. This patient was also positive for CMV, Epstein-Barr virus (EBV) and herpes simplex virus (HSV). He also had had more than 50 sexual partners in the past year and had recently traveled to many countries including North Africa, Greece and India. The new viral isolate was related to the HTLV family of retroviruses and could be propagated to healthy T cells simply by incubation with the supernatant from an infected culture (23). Furthermore, the same group showed that the virus had a high tropism towards the T4 helper subset (24). In parallel, other groups conducted similar studies that further confirmed these results. Gallo et al. (25) isolated mature viral particles from blood samples of an AIDS patient and showed that these isolates were related (but distinct) to both HTLV-I and HTLV-II. The new HTLV isolates were collectively called HTLV-III, although it was still unclear if they were identical. Notably, the HTLV-III variants differed from their HTLV-I and –II counterparts in their biological effects and morphology. Nonetheless, they presented similarities such as a high T4 lymphotropism, similar reverse transcriptase (RT) optimal conditions, syncytia formation and expression of the p19 and p24 core proteins typical of this type of retroviruses. The same group continued to detect these variants in a vast number of AIDS and pre-AIDS patients (26). The presence of antibodies against HTLV-III in blood samples of hemophiliacs receiving blood products (27) and AIDS patients (28,29) was further reported. The viruses isolated from these experiments were collectively called lymphadenopathy-associated viruses (LAVs) or AIDS-related viruses (ARVs) (30–32).

Attempts at studying the properties of HTLV-III/LAV/ARV required a system of continuous virus production that allowed for the propagation, detection and isolation of viral isolates extracted from AIDS patients. A major obstacle in designing such a system was the acute or total lymphopenia that characterizes AIDS, allowing only for transient and short-term cultivation of infected T-cells. To tackle this problem, Popovic et al. (32) tested several cell lines for their susceptibility to HTLV-I, HTLV-II and the HTLV-III variants isolated in the previous studies (25,26). One of these cell lines (the HuT T-cell line) was found to be susceptible to infection with the HTLV-III variants and to produce viral particles in amounts that allowed for comparisons with other isolates using established methodologies. Continuous production of HTLV-III was obtained after repeated exposure of parental HuT cells to concentrated fluids harvested from short-term cultures of T-cells from AIDS and pre-AIDS patients. After validation by analysis of RT activity and cross-reactivity with serum from a patient with hemophilia A and AIDS, the parental HuT cells were then extensively cloned to select for those lineages that preserved the highest rates of proliferation and virus production. Finally, to generate a cell line stably producing HTLV-III particles, the most proliferative and susceptible clones (H4

and H9) were selected and infected with the different HTLV-III isolates. The levels of infection were measured by RT activity of the culture supernatant at days 6 and 14 after infection. After a 14-day period of cell number decline, both the total number of cells and the proportion of infected cells started to increase. Light and electron microscopy revealed the presence of syncytia and a high rate of virus secretion, respectively. The infected cells were monitored over a period of 5 months: virus production was stable, and cell viability was consistently in the 65%-85% range. Many studies with tremendous relevance were built upon this work by Popovic et al. (32).

In 1985, several articles unveiling the complete nucleotide sequences of the HTLV-III, LAV and ARV isolates were published (33–36). Gonda *et al.* (37) reported that HTLV-III was more closely related to the *visna* virus, a lentivirus, than to any other retrovirus. Accordingly, expression of the lentiviral p24 protein in isolates from AIDS patients had been previously reported. Finally, Ratner *et al.* (38) published a study comparing the nucleotide sequences of the isolated variants and determined that HTLV-III, LAV and ARV isolates were simply variants of the same virus, and thus, it was at this moment that the cause of AIDS was pinpointed to a single, defined group of retroviruses, collectively called Human Immunodeficiency Virus type 1 (HIV) thereafter.

As a side note, the 2008 Nobel Prize in Physiology/Medicine was awarded to Montagnier and Barré-Sinoussi for the discovery of HIV-1 and shared with Harald Zur Hausen for finding the link between infection with Human Papilloma Virus (HPV) and cervical cancer.

### *The HIV-1 genes and their functional capabilities*

At this stage, although the major consequences of HIV infection were already described at both the cellular and organismal level, the functionality of HIV regarding molecular and genetic aspects was poorly understood. In a hallmark study, Fisher et al. (39) produced plasmid constructs encoding an HIV proviral sequence isolated from an H9/HTLV-III library and showed that transfection of these plasmids into phytohemagglutinin (PHA)-activated lymphocytes resulted in expression of unintegrated provirus and production of infectious viral particles that further infected other cells in culture. This was the first study showing that the major hallmarks of HIV infection could be induced simply by expressing the HIV genome, paving the way for a multitude of studies that relied on similar techniques in order to unveil the function of the HIV genes.

Nucleotide sequence analysis of the different isolates obtained from AIDS patients revealed three long open reading frames (ORFs) that, due to their high homology with regions in other retroviruses (namely HTLV-I and II), were quickly attributed to the *gag* and *env* genes, which encode structural proteins, as well

as to the *pol* gene required for virus replication (40–42). The same studies also showed that the HIV genome had at least three more ORFs that were distinctive from other HTLV isolates: ORFs *A*, *B* and *3'-ltr*. Allan et al. (43) discovered that ORF *3'-ltr* (located between *env* and the 3'-LTR) encoded a 27KDa protein, which reacted with antibodies from sera of AIDS patients. ORF *3'-ltr* was later named *nef* (as in "negative regulatory factor") and was found to encode a regulatory protein that induces numerous changes in the host cell that ultimately lead to increased viral titers. ORF *B*, later called *tat* (as in "trans-activator of transcription"), includes two exons, one located just 5' of *env* and the other in an alternative ORF of *env*, and encodes a 14KDa protein that induces high expression of genes under the HIV LTR promoter, increasing virus replication (44–46). The product of *vpr* regulates the nuclear import of the HIV pre-integration complex and that of *vpu* enchances the levels of viral particles secreted. The *rev* product localizes in the nucleus and helps by exporting defective viral RNAs that would otherwise exist without being able to be expressed. For an updated review on HIV-1 genes and proteins see (47). A diagram showing the current understanding on the arrangement of the HIV genes and their products is shown in Figure 1.

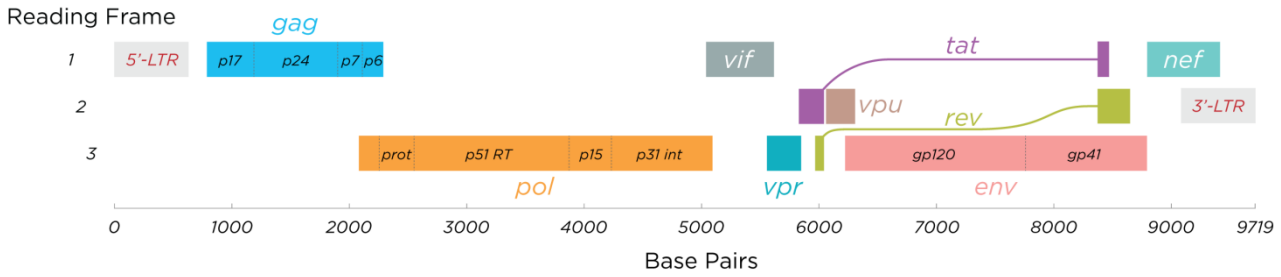


**Figure 1 – The HIV-1 genome as perceived from the current knowledge.** The *gag*, *pol* and *env* genes encode functional and structural proteins commonly found in other retroviruses. *tat* encodes a regulatory protein that, among other roles, induces high levels of expression from the HIV-1 promoter. *nef* induces numerous changes in the host cell, ultimately leading to increased viral titers. *vpr* regulates the nuclear import of the HIV pre-integration complex. *vpu* enhances the levels of viral particles secreted. The *rev* product localizes in the nucleus and helps by exporting defective viral RNAs that would otherwise exist without being able to be expressed. (Author of diagram: Thomas Splettstoesser (www.scistyle.com))

### *The elusive function of the HIV 'A' gene product, Vif*

In 1986, with the exception of ORF *A*, at least one major function for each HIV gene had been described. The HIV *A* gene, also called "*sor*" (as in "short open reading frame") is located in the HIV genome between *pol* and *env* (partially overlapped with a 3' region of the first). The first antibodies that allowed for specific detection of the *A* gene product were produced by Sodroski et al. (48). In this study, rabbits were immunized using tetanus toxin and an oligo-peptide mimicking a region of the amino acid sequence of the predicted protein product. Antiserum from the immunized rabbits was then reacted with metabolically [$^{35}$S] methionine-radiolabeled cell extracts of Popovic's H9/HTLV-III cell line (32), which precipitated a 23 kDa protein. The absence of the 23KDa protein in cells infected with HIV mutants bearing deletions in ORF *A* confirmed that this region encoded the predicted protein product. To determine the importance of the

HIV *A* gene in virus replication and cytopathogenicity, the same deletion mutants were transfected into a B-cell line (Raji-$tat_{III}$ cells) followed by co-culture with a recipient T-cell line (C8166 cells). Although the levels of other viral proteins were similar to the wild-type control, the spread of infection with mutant HIV was delayed; while 95% of cells displayed virus-specific membrane fluorescence after four days of infection with prototype virus, the same level of fluorescence was achieved by cells infected with mutant HIV only ten days post-infection. The emergence of cytopathic changes and a marked decrease in cell viability followed the same "delayed" pattern. These results suggested that the HIV *A* gene was not required for virus replication and the emergence of cytopathic effects, instead acting as a promoter of infectivity. In a different study, Kan et al. (49) used bacteria to synthesize the product of the HIV *A* gene. The synthetic 23 kDa protein was metabolically radiolabeled with either [$^{35}$S] methionine or [$^{35}$S] cysteine and resolved on sodium dodecyl sulfate-polyacrylamide gels (SDS-PAGE). Western blot analysis revealed that this protein reacted with antisera from several patients with AIDS and pre-AIDS.

Upon identification of the HIV *A* gene product, several studies attempting to elucidate its function were published. Strebel et al. (50) used a colon carcinoma cell line (SW480 cell line) to produce and extract HIV viral particles by transfecting plasmids encoding either wild-type provirus or an ORF *A* deletion mutant generated by digestion with restriction enzymes. Co-culture of producer cells with a $CD4^+$ T-cell line (A3.01 cells) yielded infected recipient cells in all cases; however, while wild-type virus resulted in a peak of infection between days 7 and 9 post-infection, which persisted for an additional 10 days, the peak of infection in A3.01 cells infected with ORF *A* mutant virus was delayed for about 10 days and persisted at least for 82 days, similar to what had been described by Sodroski et al. (48). The same experiment done by infecting A3.01 cells with equal amounts of viral extracts (as measured by the RT activity of the extracts) instead of co-culturing with a virus-producing cell line resulted in almost undetectable levels of infection by the ORF *A* mutant when compared to the wild-type, suggesting that these viruses contained an intrinsic defect affecting early events of the infectious cycle. Furthermore, trans-complementation by co-transfection of the producer cells with both the ORF *A* deletion mutant and a plasmid encoding an intact ORF *A* abrogated this effect, confirming the previous result. Moreover, when higher amounts of virus were used to infect cells, although the ORF *A* mutant yielded detectable levels of infection, the levels of viral progeny production were low, and the relative infectivity of that progeny was around 1000-fold less than that of wild-type particles. Luciw et al. (51) generated an ORF *A* "frameshift" mutant expected to produce a protein encoded by the first 28 codons of the ORF *A* gene and 23 additional codons of a different translation frame. In contrast with previous results using C8166 and SW480 cells, transfection of a T-cell line (HuT 78 cells) with a plasmid encoding the ORF *A* mutant HIV yielded low levels of virus (as measured by RT activity) and no detectable cytopathic changes, suggesting that the ORF A mutant was not able to replicate in this specific cell line.

In a more elaborate study, Fisher et al. (52) used Popovic's H9 cell line (32), as well as another CD4+ T-cell line (Molt4) and PHA-stimulated peripheral blood mononuclear cells (PBMCs), to express similar ORF *A* mutants to those produced by Sodroski et al. (48), in addition to mutants containing premature stop codons introduced by site-directed mutagenesis. Monitoring these cultures at approximately weekly intervals by reactivity with antibodies against viral proteins, the RT activity of culture supernatants and electron microscopy revealed that the ORF *A* mutant strains failed to achieve stable infection. The authors sought to determine if this was caused by impaired virus production or by the production of defective viral particles. To do this, the Cos-1 cell line of simian kidney fibroblasts was used; because this cell line is transformed with the SV40 provirus it has the capacity to promote episomal replication of plasmids carrying the SV40 origin of replication such as those encoding the ORF *A* mutants used in this study, perpetuating plasmid expression through culture expansion. Transfection of Cos-1 cells with the ORF *A* mutant or wild-type HIV plasmids resulted in the production of similar levels of viral particles that were morphologically indistinguishable from wild-type particles, indicating that ORF *A* was not required for the assembly of normal levels of morphologically intact viral particles. However, the supernatant from these cultures failed to establish stable infection in H9 cells, even after repeated attempts, suggesting that the ORF *A* mutant viruses are transmissible, although inefficiently. Surprisingly, when the transfected Cos-1 producer cells were co-cultured with H9 cells, the recipient population showed a 0.4% to 1.2% peak incidence of infection at 10 days of co-culture for the ORF A strains, compared to the >50% observed for wild-type virus. Interestingly, the small percentage of cells infected by the ORF A mutant viruses persisted in the cultures at low levels (0.1% to 0.4%) for more than 70 days; since these cultures divided rapidly, maintaining such levels of infection would require the production of infectious particles, thus suggesting that ORF A mutant viruses are inefficiently transmissible. Surprisingly, when a different T-cell line (Molt3) was co-cultured with the transfected Cos-1 cells, the transmission of infection by ORF *A* mutant strains was only slightly delayed when compared to the wild-type strain (peak of infection at 21-28 days vs. 14-15 days of co-culture, respectively). Thus, it seemed that the *A* gene of HIV was an enhancer of HIV transmission *in vitro* whose effect varied according to cell type and it was hypothesized that intrinsic cellular factors could play a role in these mechanisms. When viral particles isolated from infected Molt3 cultures were used to infect H9 cells, the infectivity of those particles was at least 100 times inferior to the wild-type strain. Northern blot analysis revealed that viral RNA and proteins produced by infected Molt3 cells were similar in amount and quality between ORF A mutant and wild-type strains, suggesting that the product of the *A* gene did not exert its effect at the level of viral transcription or translation. This was the first study showing that the HIV *A* gene product acted post-translationally, likely during virus maturation.

At least five years passed with scientists not knowing the true function of the HIV *A* gene. Given the apparent influence of its product in determining the infectivity of the assembled HIV-1 virions, the name "Virion infectivity factor" (or Vif) was adopted to refer to this protein and its corresponding gene (*vif*). Using a series of reciprocal recombinants between a wild-type (N1T-A) and a specific strain of HIV-1 (N1T-E) obtained after characterizing a family of HIV-1 isolates from a patient with lymphadenopathy (26,53,54), Volsky et al. (55) isolated a natural *vif* deletion mutant. As previously observed in other *vif* mutants, N1T-E viruses isolated from H9 and CEM T-cell lines displayed slow infection kinetics and were not cytopathic; just like prototypic HIV-1, its viral particles were able to enter cells and, in some cases, achieve stable infection with an apparently highly productive phenotype (based on virus protein levels), which did not correlate with infection kinetics and cytopathic effects.

Sodroski et al. (56) reported that after transfection of the SupT1, Jurkat and C8166 T-cell lines with *Δvif* HIV-1, viral infection would ultimately reach the same peak levels in these cultures, although with an average 10-day delayed onset when compared to the wild-type HIV-1, suggesting that *Δvif* HIV was only slightly less infectious than wild-type HIV-1. However, the same experiment done in H9 and CEM T-cells would result in failure of *Δvif* HIV-1 to achieve stable infection, indicating that either different cell lines were differently permissive to *Δvif* HIV-1 infection, or that the cell line in which the virus was produced played a role in determining its infectivity. In favor of the latter hypothesis, Δ*vif* virus produced in the simian Cos-1 cell line infected all five human T-cell lines with similar efficiencies; in turn, viruses isolated from Δ*vif* HIV-expressing SupT1, Jurkat and C8166 cell cultures was unable to achieve stable infection in PHA-activated PBMCs. However, these studies were unable to properly distinguish if the differences in the infectivity of *Δvif* HIV-1 were caused by the virus-producing cell line or by the recipient cell line, as viruses produced were simply allowed to infect the transfected cell culture. Thus, the authors followed with a series of experiments in which viruses produced in one cell line were used to infect another cell line. When HIV-1-expressing SupT1 cells were co-cultured with either CEM or SupT1 cells, *Δvif* HIV-1 was about 1.7-fold less infectious than wild-type HIV-1. However, when HIV-1-expressing CEM cells were co-cultured with either SupT1 or CEM cells, *Δvif* HIV-1 was about 7-fold less infectious than the wild-type control. Thus, it seemed that the Δ*vif* phenotype was much more dependent on the virus-producing cells used, in spite of the recipient cells. These effects were much more pronounced when cells were infected with viral stocks instead of using co-cultures; *Δvif* HIV-1 produced in SupT1, Jurkat and C8166 cells was only 1.5-, 3.2- and 9-fold less infectious than the wild-type, while *Δvif* HIV-1 produced in CEM cells was about 75-fold less infectious than the wild-type. All of this evidence also pointed towards a late-stage effect of Vif in viral replication since cells expressing *Δvif* HIV-1 seemed to produce normal levels of morphologically intact virus. However, this virus was able to infect other cells with different efficiencies when produced in different cell lines.

In a more elucidative study, Trono et al. (57) first infected the different cell lines with virus stocks produced in the simian Cos-1 cell line and measured the spread of viral infection. Three major clusters were found: (i) one containing primary CD4+ T-cells, macrophages and H9 cells, which did not support the spread of *Δvif* HIV-1 infection ("non-permissive" or "restrictive" to *Δvif* HIV-1 infection); (ii) a second cluster containing the CEM-SS line (a "syncytia-susceptible" derivative of the CEM line, hence the name SS) which displayed an intermediate phenotype of delayed infection kinetics; and (iii) a third cluster including SupT1, Jurkat and C8166 cells, which totally supported *Δvif* HIV-1 replication and to which the *Δvif* mutation had no visible consequence ("permissive" to *Δvif* HIV-1 infection). By then harvesting the viral progenies of the infected cultures and posteriorly infecting other cells with them, the authors showed that the "restrictive" or "permissive" status to infection with *Δvif* HIV-1 depended, at least to a major extent, on the cell line originally used to produce the virus stocks. Furthermore, trans-complementation by co-transfection of Vif-expressing plasmid only abrogated the *Δvif* phenotype when done in the producer cells, confirming the previous hypothesis. Thus, the *Δvif* virus produced in "restrictive" cells is non-infectious; in turn, the *Δvif* virus produced in "permissive" cells is infectious. Moreover, while virus-entry assays showed no differences between wild-type and *Δvif* HIV-1, PCR analysis of HIV-1 reverse transcripts showed that reverse transcription was highly affected in recipient cells, indicating that the intrinsic defect of *Δvif* virions produced in "restrictive" cells was at the level of reverse transcription and not virus entry. The virions, however, contained normal amounts of RT, which was completely functional. In the same year, Volsky et al. (58) also published a study relating the infectivity of *Δvif* HIV-1 when produced in "permissive" or "restrictive" cell lines with their efficiency of initiating reverse transcription, despite containing normal levels of functional RT units.

The two possible explanations for the results showed in the previous studies were that (i) permissive cells expressed a factor that could compensate for Vif, or that (ii) non-permissive cells expressed a factor that inhibits HIV-1 infectivity in the absence of Vif. Simon et al. (59) provided insights to this problem by analyzing the infectivities of virus produced in heterokaryons formed by fusion between permissive and non-permissive cells. When both cell components of the heterokaryons were permissive, the infectivities of the virus produced in these cells were always high in the absence of Vif and did not change in its presence. In contrast, when one component of the heterokaryons was non-permissive, the infectivity of *Δvif* virions was decreased by a factor of 10 to 20 in the absence of Vif but increased as a function of Vif levels. In parallel, Kabat et al. (60) confirmed those results using very similar experimental approaches based on the production of heterokaryons between either permissive or non-permissive cells and the fully permissive HeLa-CD4 cell line: in other words, the non-permissive phenotype was dominant over the permissive phenotype. Thus, it was clear that non-permissive cells expressed some factor that in the absence of Vif would decrease the infectivity of HIV-1 viral particles produced in these cells.

### *The natural antiretroviral factor inducing the Δvif phenotype*

In 2002, Sheehy et al. (61) used individual subtracted cDNAs to probe RNA extracted from a panel of non-permissive and permissive cells. A particular cDNA (CEM15) allowed for the consistent detection of a 1.5 kilobase transcript in non-permissive cells but not in permissive cells. Furthermore, overexpression of CEM15 in the permissive CEM-SS cell line was sufficient to induce a non-permissive phenotype. In the same study, bioinformatic analyses showed that the N- and C-terminals of CEM15 had a high degree of homology with a protein called Apolipoprotein B mRNA-editing enzyme 1 (APOBEC1) (62), which was shown to act upon Apolipoprotein B mRNA, where it catalyzes a C-to-U mutation that inserts a premature translation stop codon, impairing the synthesis of full-length and functional Apolipoprotein B (63,64). CEM15 encoded a similar zinc-binding domain (65) previously shown to be critical for such enzymatic activity (66,67). Thus, it was hypothesized that CEM15 was the cellular target of Vif and that when the *Δvif* virus is produced in non-permissive cells, CEM15 produces the *Δvif* phenotype by affecting viral RNA components, ultimately inhibiting the synthesis of reverse transcripts and the establishment of infection in recipient cells. Thus, CEM15 was part of a newly found innate antiretroviral resistance that, in this case, was counteracted by the viral protein Vif. Since CEM15 is also highly expressed in activated primary CD4+ T-cells display a non-permissive phenotype towards *Δvif* HIV-1, it was inferred that Vif is required in order for wild-type HIV-1 to efficiently infect its natural host cells.

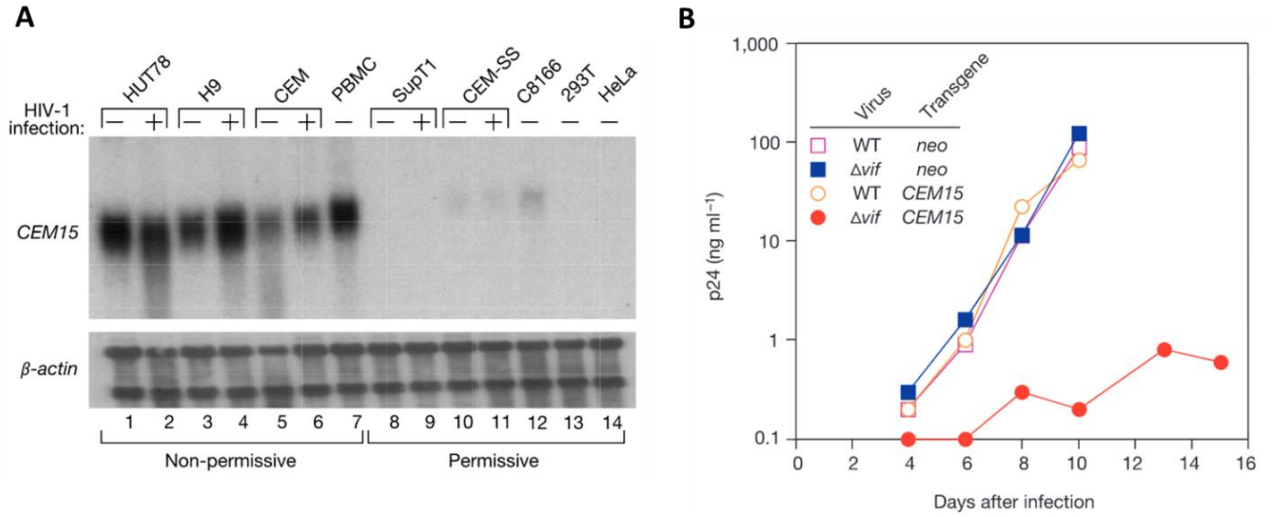


**Figure 2 – Key observations from Sheehy et al. (61) that established CEM15 as a natural antiretroviral factor. (A)** Western blot analysis showing that the CEM15 protein is exclusively expressed in non-permissive cell lines. **(B)** Overexpression of CEM15 is sufficient to induce a non-permissive phenotype in the permissive CEM-SS cell line, as measured by the levels of p24 (*gag*) in the culture supernatants.

In 2002, Jarmuz et al. (68) discovered that APOBEC1 is part of a larger family of deaminases now collectively called AID/APOBEC, whose corresponding genes reside at 3 different *loci* on the human genome: APOBEC1

and the B-cell-specific Activation-Induced Deaminase (AID) form a gene cluster in chromosome 12; APOBEC2 is in chromosome 6; and a third cluster of genes and pseudogenes found in chromosome 22 is called APOBEC3A through G. It was soon widely accepted that CEM15 was, in fact, APOBEC3G (A3G). Because the AID/APOBEC superfamily shared the catalytic domain of the RNA-editing APOBEC1, the other APOBECs were temporarily though to also act upon RNA.

In parallel, the seminal works of Muramatsu et al. (69), Scharff et al. (70) and Neuberger et al. (71) as well as posterior studies including those of my supervisor Dr. Vasco Barreto (72) and others, unveiled the particularities of the B-cell-specific expression of AID in the germinal centers of the spleen and secondary lymphoid organs that is required for somatic hypermutation (SHM) and class switch recombination (CSR) of the antibody variable and constant regions. To initiate these processes, AID catalyzes C-to-U deamination in the immunoglobulin (Ig) *locus* to elicit a "controlled" error-prone DNA repair response involving multiple enzymes that ultimately results in point mutations (the SHM process) and the rearrangement of the heavy-chain locus (CSR). Subsequent clonal expansion of the B-cells producing the antibodies with the highest affinity for the epitopes in the germinal centers is favored in detriment of other cells, resulting in the selection B-cells producing antibodies with improved specificity. Some of these cells then become memory B-cells that remain dormant in circulation, ready to be re-activated and express their already rearranged and mutated antibody genes.

Motivated by the finding that a member of the AID/APOBEC family acts upon DNA rather than RNA, Lecossier et al. (73) tested the hypothesis that A3G could hypermutate the "minus" chain of the HIV-1 DNA reverse transcript and in that way inhibit the production of infectious virions. PCR analysis using primers for the "plus" strand of reverse transcripts from HIV-1 virions produced in non-permissive H9 cells detected a high level of G-to-A hypermutation compared to wild-type virions or those generated in the permissive HeLa cell line. This, in turn, could be explained by C-to-U hypermutation of the "minus" DNA chain upon reverse transcription, in which the U introduced by A3G are read as T during the synthesis of the "plus chain". If A3G acted upon viral RNA, C-to-T was expected in the DNA "plus" strand, instead of the observed G-to-A; furthermore, PCR analysis of cDNA resulting directly from *in vitro* reverse transcription of isolated viral RNA did not show substantial G-to-A hypermutation.

A series of studies (74,75) were also released in which a particularity of the A3G:Vif interaction emerged: in non-permissive cells, A3G was readily incorporated inside virions only in the absence of Vif. Moreover, incorporation of A3G into virions was the major contributor to the non-permissive phenotype, which is consistent with early observations showing that the decreased infectivity of the *Δvif* viral particle was "imprinted" during its assembly in the cells in which it was originally produced. Since the viral single-

stranded DNA substrate of A3G is only available upon reverse transcription in the infected cell, A3G appears to be incorporated in virions specifically in order to travel from cell to cell and become available at the specific regions where RT is about to be initiated. In fact, producing *Δvif* virions in non-permissive cells is sufficient to generate virions that are unable to infect permissive cells, indicating the marked antiviral activity that virion-incorporated A3G can have; in turn, when *Δvif* virions produced in permissive cells are used to infect non-permissive cells, the onset of stable infection is slightly delayed but ultimately achieved at levels similar to those observed using wild-type virions, indicating the small role played by non-incorporated A3G in restricting viral replication. The mechanisms by which Vif excludes A3G from incorporation into virions involve post-translational ubiquitination of A3G and its subsequent degradation in the 26S proteasome by a multi-protein complex recruited by Vif itself upon binding to A3G (76–79). Despite these demonstrations, the mechanisms underlying the sorting of A3G into virions are not fully understood. For instance, it is unknown if the incorporated A3G is either passively sorted into virions as a consequence of high expression in activated CD4+ T-cells and subsequent cytosol engulfment during virus assembly, or if (and which) selective mechanisms underlie this phenomenon. The present work will give some insight into that specific subject.

The biology of APOBEC:Vif was, to my knowledge, only exploited by scientists in the field of biomedical research for its dependency on the ubiquitin-proteasome system (UPS) with the development of proteasome inhibitors (80). However, since the UPS plays a crucial role in several other degradative pathways and an overall proteostasis is required for normal cellular function, inhibition of the 26S proteasome appears to be an over-intrusive approach to counter HIV-1 infection compared to the more specific HIV protease inhibitors that started to be released in the market prior to the APOBEC:Vif studies. An obvious but apparently unexplored strategy is the development of a specific inhibitor of the A3G:Vif interaction. It seems intuitive that A3G and Vif co-exist in a strictly regulated stoichiometric equilibrium in non-permissive cells stably infected with the wild-type HIV-1. Not surprisingly when dealing with lymphocytes, this equilibrium seems to involve the sharing of endogenous factors between cells. Because the immune response was shown to work efficiently in inhibiting HIV-1 replication in the absence of Vif (as demonstrated by the aforementioned studies in which *Δvif* HIV-1 particles produced in non-permissive cells were unable to infect activated primary CD4+ T-cells), modulation of the A3G:Vif stoichiometric equilibrium appears to be a promising strategy in preventing, treating or even curing HIV-1 infection. These could be achieved by increasing the amounts of functional A3G or decreasing the amounts of functional Vif in the target cells. The development of therapeutic strategies based on these mechanisms and pathways would be an improvement to the set of therapeutic options already available to refrain HIV infection. More importantly, the synergistic effects of a combinatorial therapy that includes such alternatives is also a relevant topic.

***Remarks regarding APOBEC enzymes***

Following the discovery of APOBEC-1 as the mutator underlying the C-to-U transitions observed in Apolipoprotein B mRNA (62–67), and of A3G as the major endogenous factor restricting the replication of *Δvif* HIV-1 (61), several studies regarding the AID/APOBEC family were published, giving more insights on this peculiar family of enzymes. As shown in Figure 3A, it is currently accepted that the AID/APOBEC family is composed of 10 highly-related primary gene products: A1, AID, A2, A3A-D, A3G, A3H and A4. Furthermore, genetic polymorphisms and alternative splicing give rise to numerous isoforms with varying degrees of activity (81–83). Several lines of evidence suggest that the AID/APOBEC genes originated from the ancestral AID gene, which is present in all vertebrates, unlike the APOBECs (Fig. 3B). Orthologs of these cytosine deaminases have been found in jawless fish (84) and even invertebrates (85). The AID/APOBEC genes are thought to have been originated as a result of gene duplications and translocations during evolutionary history. This diversity was apparently favored along the mammalian lineage, and most notably in primates, which could imply an underlying selective pressure, such as immunity to retroviruses in individuals in which such genetic events would have happened (86).

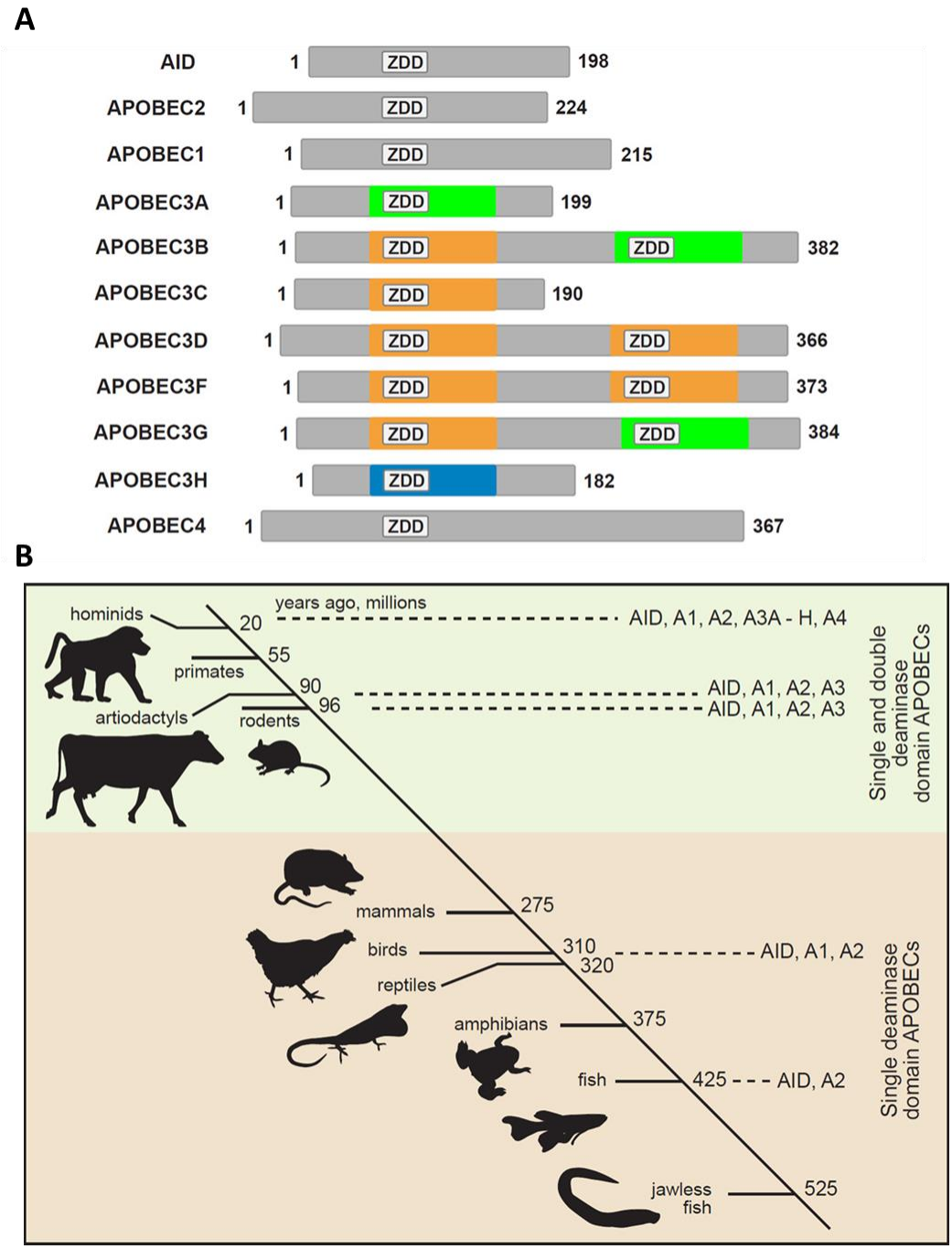


**Figure 3 - Alignment of members of the APOBEC family and their predicted phylogeny (adapted from Salter et al. (86)). (A)** APOBEC protein lengths (amino acid) are depicted to scale with the total length shown at the C terminus and with proteins aligned to the (first) zinc-binding motif (ZDD). **(B)** "A cartoon depicting the hypothesis for APOBEC paralog evolution. The phylogenetic relationships between several vertebrates and the emergence of APOBEC family members are shown on a timeline in millions of years that is not drawn to scale. AID and A2 emerged in jawed fish whereas A1 may have first appeared in ancestors of reptiles and amphibians. The timing of gene duplications and diversification events that led to the A3 subgroup (brown to green transition) is unclear, but likely occurred around 200 million years ago. Humans and other primates express the largest diversity of A3 proteins. A4 is the least studied APOBEC protein. Although A4 is present in mammals, chickens and frogs, it is absent in fish. Its gene was found by computational analyses (87), and its function is poorly understood. In a study by Lada et al. (88), A4 was shown to be incapable of catalyzing C-to-U deaminations in DNA, suggesting that it may instead interact with RNA. Notably, Marino et al. (89) recently reported that A4 enhances HIV replication by deaminase-independent mechanisms.

Yang et al. (90) demonstrated that the antiviral activity of APOBEC3G is, at least, partially inhibited if the expression of the enzymes uracil DNA glycosylase-2 (UNG2) and apurinic/apyrimidinic endonuclease (APE) is decreased using either specific chemical inhibitors or siRNA-mediated down-regulation of expression. Furthermore, similarly to A3G, UNG2 appears to be incorporated into virions, making both proteins available at the specific sites of infection before reverse transcription even commences. An outline of the processes that are predicted to occur as a consequence of the enzymatic activity of DNA mutators such as A3G is shown in Figure 4. Briefly, upon virus entry, viral RNA is released into the cytosol. The viral "plus" strand of RNA is quickly reversely transcribed into a "minus" chain of DNA upon which cytosolic DNA mutators such as A3G can act. This activity results in C-to-U transitions that can have more than one consequence: (i) the mutations can be kept if DNA polymerase mistakenly reads the U as T, inserting A instead of G (G-to-A hypermutation); (ii) the presence of U in DNA strands is picked up by uracil DNA-glycosylase enzymes that cleave off the U, leaving behind abasic sites which either impair second-strand synthesis or further recruit the degradative apurinic/apyrimidinic endonucleases. Notably, a concern regarding the hypothesis that G-to-A hypermutation of proviruses may increase provirus variability and infectious properties is often mentioned in the literature. In fact, Armitage et al. (91) showed that the mutational signatures of A3G were present in hypermutated proviral sequences of HIV-1 found in a database of infected patients. However, Delviks-Frankenberry et al. (92) estimated that the contribution of G-to-A hypermutation for the variability of HIV-1 is minimal and even smaller than that of "error-prone" viral replication.

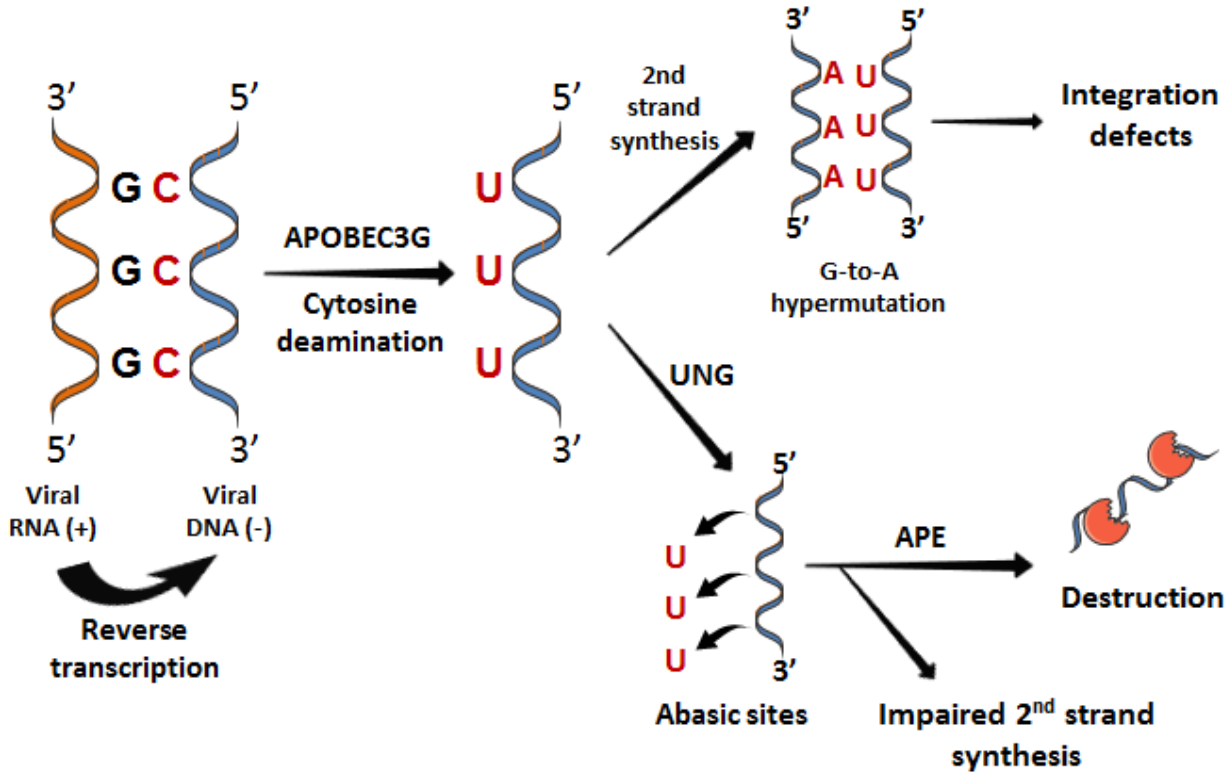


**Figure 4 – Outline of the predicted mechanisms underlying A3 antiviral activity.** Upon reverse transcription of a viral "plus" RNA strand, a "minus" chain of DNA is generated, upon which cytosolic DNA mutators, such as A3G, can act by generating C-to-U transitions with multiple consequences. The mutations can be kept if DNA polymerase mistakenly reads the U as T, inserting A instead of G (G-to-A hypermutation). In turn, the presence of U in DNA strands can be picked up by uracil DNA-glycosylase (UNG) enzymes that cleave off the U in the DNA, leaving behind abasic sites which either directly impair second-strand synthesis or further recruit the degradative apurinic/apyrimidinic endonucleases (APEs).

Another interesting characteristic of AID/APOBEC proteins is their subcellular distribution. Because these proteins, with the exception of A1 and possibly A4 (88), are DNA-editing enzymes, their containment must be ensured in order to avoid unwanted mutations in the cell's genome (i.e., genotoxicity). Naturally, this becomes an even bigger problem during mitosis, when the nuclear envelope is broken down and chromosomes are exposed to the cytoplasm. AID appears to be mostly cytoplasmic but during SHM and CSR clearly shuttles between the cytoplasm and the nucleus (93–96). Some APOBEC proteins are mostly cytoplasmic, while others also shuttle between the cytoplasmic and nuclear compartments, with some evidence suggesting that even when the nuclear envelope is broken down during mitosis, the cytoplasmic A3 proteins are kept from interacting with genomic DNA due to chromatin condensation (44). Moreover, it appears that A3 protein export is not exclusive to the nucleus since A3 has been detected in small nano-sized extracellular vesicles (called exosomes) from the non-permissive H9 cell line (1,2). In these studies, simple overexpression of some A3 proteins was sufficient to detect them in exosomes of both permissive and non-permissive cells. Importantly, these exosomes containing A3 proteins revealed potential in

preventing and treating infection with both *Δvif* and wild-type HIV-1 in permissive and non-permissive cells, including PHA-activated PBMCs from HIV-seropositive patients. However, to our knowledge, this therapeutic potential was not further explored.

### *Exosome biogenesis and function*

In 1983, Harding et al. (97) and Pan et al. (98) described the cellular mechanisms underlying the loss of cell membrane-associated transferrin receptor (TfR), a hallmark of the reticulocyte-to-erythrocyte transition. In these studies, TfR in reticulocytes was labeled with either gold particles or an anti-TfR antibody and the intracellular trafficking of the receptor was monitored using electron microscopy and Western blot analysis, respectively (Fig 5A-D). Receptors present at the cell membrane are internalized in endosomes/multivesicular bodies (MVBs), more specifically in the small vesicles (~50 nm) present in their lumens called intra-luminal vesicles (ILVs). ILVs are formed by inward budding of the endosome/MVB membrane and posteriorly released by fusion of the LEs/MVBs with the cell membrane. The term "exosome" was coined by Johnstone et al. (99). Due to the demonstrations regarding TfR, it was widely accepted that the major function of ILV/exosome formation and secretion was to remove unwanted contents from the cell.

A well-studied pathway for the formation of MVBs/ILVs, which is conserved from yeast to mammals, is driven by the endosomal sorting complex required for transport (ESCRT), a group of around thirty proteins that are usually grouped into four major functional complexes (ESCRT-0 to ESCRT-III). The ESCRT complexes are not only involved in the formation of ILVs but also in the inclusion of specific proteins in their lumens during early ILV biogenesis (100–103). However, some ESCRT-independent mechanisms have been reported. For instance, in human melanoma cells, exosome formation and release appear to be much more influenced by the levels of CD63, a membrane-associated protein of the tetraspanin family that is often used as a proxy for the presence/amount of exosomes in a sample (104). Also, in a B cell line, CD63 appears to be critical for the inclusion of the Epstein-Barr Virus-encoded LMP1 protein in ILVs/exosomes (105). Other tetraspanins, such as CD9, CD81 and CD82 have been implicated in the formation and release of ILV/exosomes in different models (106).

Ten years after the discovery of exosomes, two important studies showed that these vesicles carry functional cellular factors, which can be internalized and elicit a response in recipient cells. Raposo et al. (107) showed that a B cell line is capable of effectively presenting antigens to T-cells by releasing exosomes with a high content of major histocompatibility complex II (MHC-II) molecules. Furthermore, Zitvogel et al. (108) reported that exosomes derived from tumor peptide-pulsed dendritic cells (DCs) are able to mediate

an internal response that suppresses the growth of an established murine tumor. Since then, many studies showing the involvement of exosomes in other disease models have been published (e.g. (1)).

***KFERQ-like motifs in the sorting of soluble protein into exosomes***

An important characteristic of exosomes that was demonstrated by Raposo et al. (107) and Zitvogel et al. (108) was that the protein composition of these vesicles was different from that of the cell membrane and cytosol. Several studies have shown that the lumen of exosomes contains both soluble proteins and nucleic acids, which can be included by either selective or non-selective mechanisms during ILV formation (106). Non-selective inclusion of molecular cargo in ILVs is theoretically stochastic by nature and occurs as a consequence of cytosol engulfment during ILV formation. Thus, virtually every molecule present in the cytosol should be detectable in exosomes given the correct conditions, even if only in trace amounts. An important argument in favor of this model is that simple overexpression of a soluble protein is often enough to detect it in exosome fractions by Western blot analysis (1,2,109). In contrast, selective mobilization of molecular cargo into ILVs is a deterministic process in which specific molecules are sequestered for possessing biochemical properties that enable interactions with their molecular sequesters. In a hallmark study by Sahu et al.(109), proteomic analysis of exosomes from dendritic cells revealed an enrichment (75% in exosomes vs. 37% in the cytosol) of proteins containing a specific amino acid sequence commonly referred to as the "KFERQ-like motif", which had been shown to promote the translocation of soluble proteins into the lumen of lysosomes by enabling interactions with the chaperone Hsc70 (or HSPA8) during the process of chaperone-mediated autophagy (CMA) (110). The so-called "canonical" KFERQ-like motifs consist of a glutamine (Q) flanked on either side by four amino acids made up of a single acidic amino acid (E, D), a basic amino acid (K, R), a bulky hydrophobic amino acid (I, L, F, V), and a repeated basic or hydrophobic amino acid. Non-canonical KFERQ-like motifs are slight variations of the canonical rule including the presence of an N or an acetylated K instead of the terminal Q, and the presence of a phosphorylated amino acid (S, T, Y) instead of an acidic amino acid (111)(Fig. 5E).

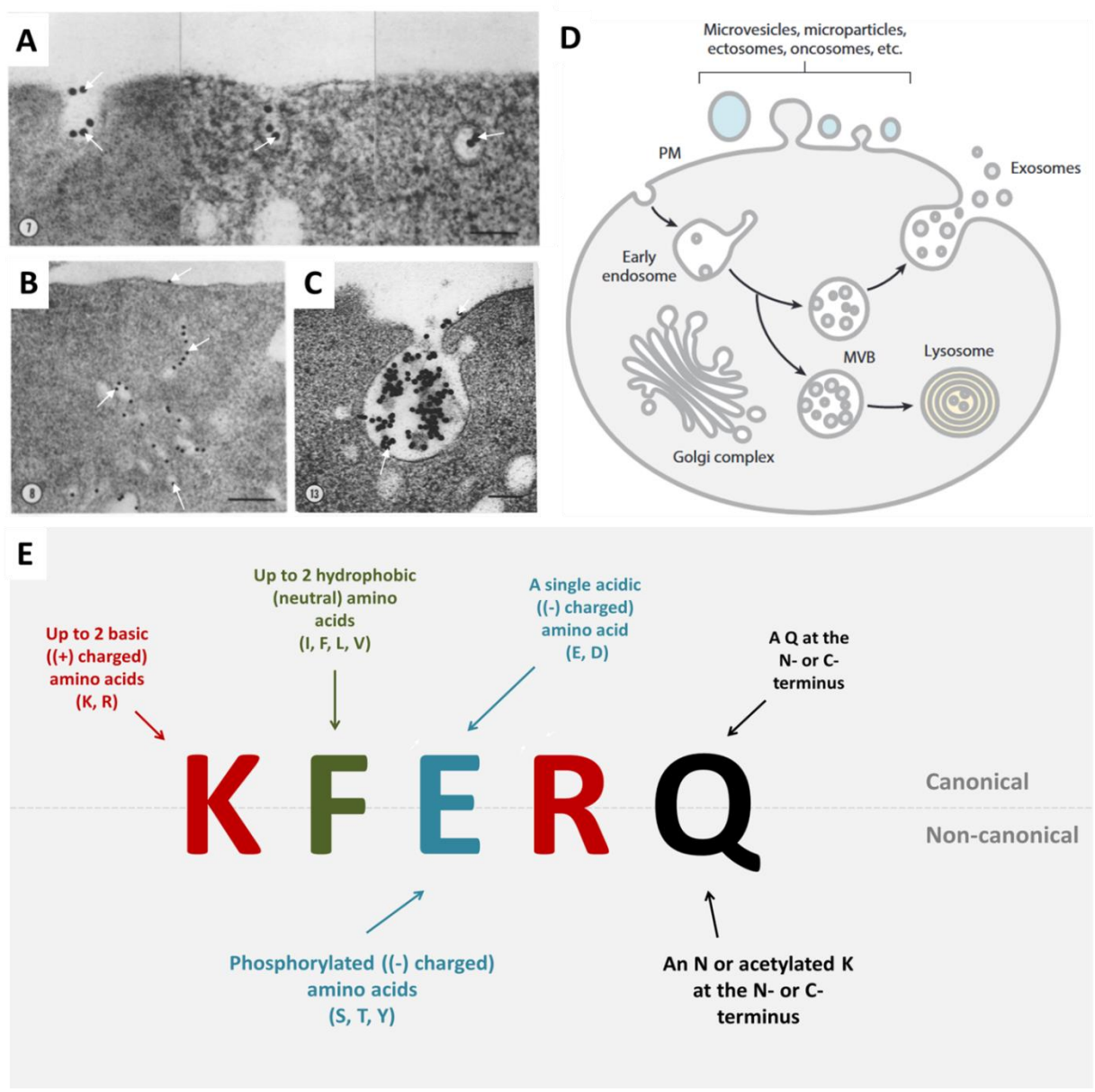


**Figure 5 – Exosome biogenesis, protein sorting, and secretion. (A-C)** EM results from Harding et al. (97) showing **(A)** the internalization of TfR (the dark spots pointed with white arrows) into endosomes at the level of the cell membrane, **(B)** their subsequent migration and re-arrangement (e.g. formation of MVBs/ILVs) in the cytoplasm and **(C)** release of exosomes by fusion of MVBs with the cell membrane. (D) Diagram from Colombo et al. (106) depicting the model for exosome biogenesis. **(E)** The "rules" for the formation of canonical and non-canonical KFERQ-like motifs. Briefly, "canonical" KFERQ-like motifs consist of a glutamine (Q) flanked on either side by four amino acids made up of a single acidic amino acid (E, D), a basic amino acid (K, R), a bulky hydrophobic amino acid (I, L, F, V), and a repeated basic or hydrophobic amino acid. Non-canonical KFERQ-like motifs are slight variations of the canonical rule including the presence of an N or an acetylated K instead of the terminal Q, and the presence of a phosphorylated amino acid (S, T, Y) instead of an acidic amino acid (111).

Selective microautophagy of soluble proteins containing a KFERQ-like motif into late endosomes/MVBs requires both a KFERQ-like motif and Hsc70 (110). The amount of GAPDH (which contains a KFERQ-like motif) in LEs/MVBs was significantly reduced by introducing a double point mutation that switched the "Q"

and the following amino acid in the KFERQ-like motif into two alanines (e.g., KFERQ -» KFEAA)(109). Furthermore, in contrast to CMA, which requires an unfolded substrate for translocation across the lysosomal membrane in a LAMP2A-dependent manner (112), *in vitro* assays using an engineered form of dihydrofolate reductase (DHFR) containing a KFERQ-like motif in its N-terminus and whose folding is stabilized in the presence of methotrexate showed that Hsc70-mediated selective inclusion of DHR in LEs did not require an unfolded substrate nor the LAMP2A receptor (109).

The influence of the CMA receptor LAMP2A in the loading of proteins into ILVs remains unclear. Mass spectrometry analysis of isolated LEs did not reveal putative receptors of Hsc70; however, thin-layer chromatography revealed the presence of lipids, and multidimensional mass spectrometry scan indicated the presence of phosphatidylserine (PS) in an ATP-dependent manner. Mutations in specific amino acids of Hsc70 putatively involved in electrostatic interactions with PS resulted in a loss of binding between Hsc70 and PS, indicating that the anchoring of Hsc70 to the membrane of LEs requires those regions to be intact and providing more evidence for such interactions. Thus, the selective inclusion of cytosolic soluble proteins based on KFERQ-like motifs could rely on the presence of “lipidic rafts” with an increased PS content that promotes binding of Hsc70 and its substrates to the regions where LEs/MBVs undergo inward budding to form ILVs.

It is noteworthy that a functional interaction between KFERQ-like motifs and Hsc70 is shared between CMA and the selective sorting of soluble cytosolic proteins into ILVs, which may seem intriguing. It remains unclear if and how Hsc70 “distinguishes” between the two pathways and “chooses” one in detriment of the other. However, researchers in the field of exosomes are widely aware that serum deprivation (or starvation), a well-established inductor of macroautophagy and CMA in cell lines and liver hepatocytes (113–116), decreases the rate of exosome secretion; in turn, autophagy inhibitors such as the lysosome inhibitor Bafilomycin A1 either increase exosome secretion (117). This is likely due to the fact that exosomes, as they were first described (98,118), are used by cells as “shuttling” vectors to remove contents from the cytosol. Furthermore, the contrasting effects of starvation and lysosome/proteasome inhibition in exosome release indicate the presence of a compensatory relationship between these pathways (i.e., over-accumulation of cytosolic proteins resulting from lysosome/proteasome inhibition is compensated by an increase in proteins released via exosomes) that together operate in cells to keep proteostasis in check. Thus, the common usage of molecular players such as the Hsc70 chaperone and KFERQ-like motifs creates an overlap between the two pathways (CMA and exosome biogenesis/secretion) that is ultimately beneficial in coping with proteostatic stress.

## RESULTS AND DISCUSSION

### Conservation of KFERQ-like motifs along the AID/APOBEC evolutionary history

The presence of several A3 proteins in exosomes was first reported by Khatua et al. (1,2). However, the mechanisms by which soluble cytosolic proteins are mobilized into ILVs remain unclear. To determine if mechanisms of protein sorting into exosomes that depend on the presence of KFERQ-like motifs could be involved, the amino acid sequences of the human AID/APOBEC protein family were screened for KFERQ-like motifs. Since the AID/APOBEC genes originate from ancestral genes thought to date back to the "jawless fish" lineage, and even invertebrates, we were also interested in knowing if the found KFERQ-like motifs display some degree of evolutionary conservation. Thus, we expanded our analysis to account for the murine, bovine and primate homologs of these proteins.

The main studies describing KFERQ-like motifs were used as a reference (110,115,119). Briefly, the so-called "canonical" KFERQ-like motifs consist of a glutamine (Q) flanked on either side by four amino acids made up of a single acidic amino acid (E, D), a basic amino acid (K, R), a bulky hydrophobic amino acid (I, L, F, V), and a repeated basic or hydrophobic amino acid. Non-canonical KFERQ-like motifs, such as those containing an N instead of a Q (120) and those generated from a putative motif by post-translational modifications (111), do not strictly obey the canonical rule but yet seem to retain functional properties (Fig. 5E). For this study, our approach is conservative, as it only accounts for KFERQ-like motifs those encoded in the amino acid sequences of proteins, leaving out putative motifs.

The results of this analysis are shown in Table 1. At least one KFERQ-like motif is present in each of the A1 homologs analyzed. Interestingly, a specific motif that is present in the murine and bovine A1 homologs (65NFLEK69) is not present in the primate and human sequences; however, the motif present in the primate and human A1 homologs (132RDLVN136) seems to derive from one of the motifs found in the bovine homolog (132KDLVQ136). Human A3D encodes a motif (199EILRN203) that appears to have been conserved along the primate lineage and possible the bovine and murine lineages. Human A3F encodes a similar motif (186EILR190) that also appears to have been conserved along at least the primate lineage. Finally, human A3G encodes a single motif (168QRELF172; QRELF$^{168\text{-}172}$ hereafter) that also seems to have been conserved along the mammalian lineage. Supposing that the conservation of such motifs is positively selected along evolutionary history for their functionality, these results suggest that some of the KFERQ-like motifs in A3 proteins might be functional. The incidence of numerous viral infections along the human lineage may have influenced the conservation of these motifs. This goes in line with the finding that the majority of A3 proteins diverged from their ancestral genes during the primate lineage (86).

**Table 1**

***Genetically-encoded KFERQ-like motifs found in proteins of the AID/APOBEC family***

| Species | Protein | Length[a] | KFERQ-like motifs |
|---|---|---|---|
| *Mus musculus* (house mouse) | AID | 198 | None |
| | A1 | 229 | 65NFLEK69 |
| | A2 | 224 | None |
| | A3 | 440 | 167KVFDN171<br>194QEILR198 |
| *Bos taurus* (cattle) | AID | 199 | None |
| | A1 | 236 | 65NFIEK69<br>102REFLN106<br>132KDLVQ136 |
| | A2 | 224 | None |
| | A3Z1 | 185 | 97FLKEN101 |
| | A3Z2 | 202 | 190QRLVE194 |
| | A3Z3 | 206 | 27NLLRE31<br>85IDKIN89 |
| *Pan troglodytes* (chimpanzee) | AID | 192 | None |
| | A1 | 236 | 132RDLVN136 |
| | A2 | 230 | None |
| | A3A | 199 | None |
| | A3B | 394 | 207EIIRN211 |
| | A3C | 190 | None |
| | A3D | 386 | 199EILRN203 |
| | A3F | 373 | 186EILRN190 |
| | A3G | 384 | 168QRELF172 |
| | A3H (SV183) | 183 | 178ERIKQ182 |
| *Homo sapiens* (human) | AID | 198 | None |
| | A1 | 236 | 132RDLVN136 |
| | A2 | 224 | None |
| | A3A | 199 | None |
| | A3B | 382 | None |
| | A3C | 190 | None |
| | A3D | 386 | 73FRFEN77<br>199EILRN203<br>382REILQ386 |
| | A3F | 373 | 186EILRN190 |
| | A3G | 384 | 168QRELF172 |
| | A3H[b] | 200 | None |

Note: Amino acids are represented using single-letter code. The numbers flanking each KFERQ-like motif represent the positions of the initial and final amino acids of that sequence in the corresponding protein. [a]Given in number of amino acids. [b]Each described functional splice variant of this protein (SV182, SV183 and SV200) was taken into account.

We focused our next observations on human A3G for three main reasons: (i) it contains a single and canonical KFERQ-like motif that is located in a region of the protein not linked to any known function (Fig. 6A), (ii) it is an endogenous antiretroviral factor whose functions and activity are well-established, enabling the design of functional assays to test our hypothesis, and (iii) its presence in exosomes of at least two cell

lines of different origins and in primary T-cells has been reported previously (1,2). The JPred4 algorithm was used to predict the exposure of the A3G KFERQ-like motif (Fig. 6B)(121). The motif appears to form a region of higher solvent accessibility compared to the surrounding 20 (or so) amino acids. Moreover, >50% of the motif itself appears to be exposed, with only the L171 and F172 amino acids being classified by the algorithm as "buried"; however, L and F are hydrophobic amino acids by nature and the influence of this property on the results is unclear. Because the JPred4 algorithm accounts for amino acid properties to determine secondary structures and calculate solvent accessibilities and predict the exposure of those structures, analysis of tridimensional structures is generally more insightful. The predicted tridimensional structure of A3G (Fig. 6C and D) suggests that the KFERQ-like motif of A3G forms a highly accessible region that is seemingly available for molecular interactions without requiring an unfolded version of the protein, in contrast to what is thought to happen in the case of some CMA substrates, whose KFERQ-like motifs are embedded in the protein's core and become available for interactions with Hsc70 as a consequence of protein misfolding/unfolding (122). Thus, it was tentative to hypothesize that the KFERQ-like motif of A3G is a conserved functional region that promotes the shuttling of A3G into exosomes.

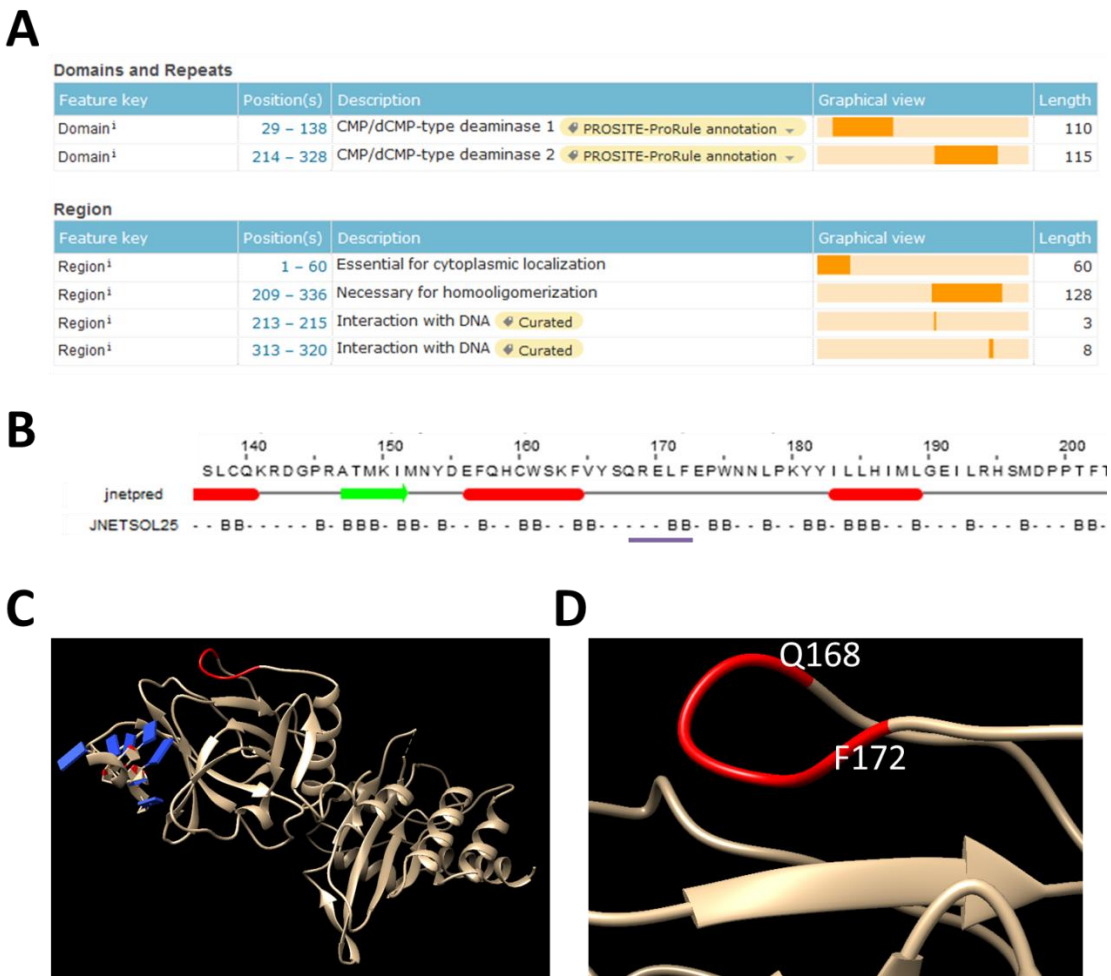

**Figure 6 – In silico characterization of the KFERQ-like motif of A3G. (A)** The functional regions and domains of A3G described in the literature (adapted from: https://www.uniprot.org/uniprot/Q9HC16). The QRELF$^{168-172}$ region is not linked to any known function. **(B)** The JPred4 algorithm predictions on the exposure of the KFERQ-like motif of A3G (highlighted with a purple line). The red and green arrows in the "jnetpred" row represent predicted alpha-helixes and beta-sheets in the structure. A "B" in the "JNETSOL25" row represents an amino acid that is predicted to be "buried", or likely inaccessible, for having a score <0.25. **(C and D)** Predicted tridimensional structure of A3G based on models of X-ray crystallography and NMR. The region in red represents the QRELF$^{168-172}$ motif.

**A3G is enriched in exosomes**

The presence of A3G proteins in exosomes isolated from the non-permissive H9 T-cell line was first reported by Khatua et al. (1,2). Moreover, transfection of A3G-negative cell lines (SupT1, Jurkat and HEK293T) with an A3G-GFP-expressing plasmid also resulted in the presence of A3G-GFP in exosomes isolated from these cultures. However, these studies were not designed to clearly distinguish between selective and non-selective sorting of proteins into exosomes, as detection of A3G by Western blot analysis in itself does not provide a measure of enrichment without a term of comparison, such as a protein that is known not to be selectively included in exosomes. Furthermore, A3 proteins expressed from exogenous DNA templates could be included in exosomes in amounts that seem higher than vestigial due to synthesis of large amounts of protein, as it often occurs with common plasmid-based expression vectors relying on strong constitutive transcription promoters. To determine if A3G is selectively included in exosomes, the HEK293T cell line was used, which (i) is easily transfected with plasmid-based expression vectors using the "calcium phosphate technique" (123) and (ii) upon expression of A3G releases A3G-positive exosomes (1,2). If A3G is specifically sorted into ILVs, then exosome fractions isolated from A3G-expressing cells should be enriched in A3G when compared to a protein that lacks functional KFERQ-like motifs and is not selectively sorted into ILVs, such as EGFP (unpublished data).

Western blot analysis of exosome fractions isolated by ultracentrifugation from the SN of HEK293T cell cultures transiently expressing either an A3G-EGFP fusion protein or EGFP alone first revealed that the levels of exosomal A3G-EGFP were only slightly higher than those of EGFP (Fig. 7A). However, the levels of cellular EGFP were about 5±1 times higher than that of A3G-GFP (unpaired *t*-test; $p<0.01$; n=2) as measured by the cellular levels of both proteins (Fig. 7B). This difference was also observed under a fluorescence microscope (data not shown) and can result from an indefinite number of reasons, of which the shorter reading frame, ease of expression and high protein stability of EGFP are especially noteworthy. For quantification purposes, exosomal A3G-EGFP or EGFP were normalized to their cellular levels and then to the levels of exosomal CD63, a marker typically used to control for the amount of exosomes in each sample (Fig. 7C). After these normalization steps, an average 6.5±0.5-fold enrichment in exosomal A3G-EGFP when compared to EGFP was found (unpaired *t-test*; $p<0.01$; n=2).

An alternative approach was explored to validate the previous result. Exosomes were immunoisolated from cleared SNs of transfected cells using anti-CD63-conjugated latex beads and analyzed by flow cytometry. To control for the amount of exosomes present in each sample of beads, the FM 4-64 dye was used, which according to the manufacturer, effectively stains membrane lipids in mammalian cells and emits fluorescence in a different channel than EGFP (FL3 for FM 4-64 and FL1 for EGFP). The gating strategy used is depicted in Figure 7D. Briefly, bead populations in each sample were gated in a "Forward scatter (FSC) vs. Side scatter (SSC)" plot using FACS buffer as a negative control. Then, an FL3(+) population of beads was gated using stained beads samples incubated with PBS instead of cleared SN as a negative control. Finally, FL1(+) beads populations inside the FL3(+) subsets were gated using samples coming from cells transfected without any plasmid. At this point, it is important to note that while the Western blot approach was used to detect A3G-EGFP and EGFP protein amounts, this alternative strategy accounts for the numbers of positive/negative events. Measures such as the "median fluorescence intensity" may be used as better predictors of protein amounts in each event, however in these conditions it was not possible to detect significant differences between treated and control samples (data not shown), possibly because the antibody-to-beads and/or beads-to-exosomes ratios during the incubation steps were not ideal (the first was not-tested due to low availability of primary antibody and the second was only tested for fractions of SN smaller but not higher than 1). Nonetheless, populations of beads incubated with cleared SNs were visually distinguishable in FL3 histograms from unstained samples and, more importantly, samples incubated with PBS instead of SN. With this methodology, the average cellular level of EGFP was about 1.8±0.2-fold higher than that of A3G-EGFP (Fig. 7E), as observed by the percentage of EGFP-positive and -negative cells (ratio paired *t* test; $p<0.01$; n=2). The percentage of EGFP-positive beads: exosomes conjugates was not significantly different from that of samples coming from cells transfected without any plasmid, suggesting that EGFP is either excluded from exosomes or that it is included in amounts that fall below the sensitivity threshold of these experimental conditions (Fig. 7F). In contrast, the percentage of EGFP/FL1-positive beads: exosomes inside the FM 4-64/FL3-positive subset (normalized to the corresponding levels of cellular EGFP) was 2.1±0.1 times greater for A3G-EGFP when compared to EGFP (ratio paired *t*-test; $p<0.001$; n=2).

The discrepancy between the fold change enrichment values obtained using different techniques is likely to be caused by differences in the sensitivity and overall concept underlying each method. In the Western blot approach, the values for cellular EGFP are likely to be biased due to a marked difference in cellular A3G-EGFP and EGFP protein levels, leading to a scenario where if the exposure to chemiluminescence is adjusted so that cellular A3G-EGFP is visible, EGFP easily becomes saturated; in turn, if the exposure is adjusted so that EGFP in not saturated, A3G-EGFP becomes undetectable. For the purpose of relative quantification, it was necessary to select an exposure time that resulted in both proteins being visible at

the same time. Thus, for these reasons, underestimation of cellular EGFP levels is a predicted caveat of these assays, likely impacting the final values to some extent. However, in that case, since by definition the exosomal-to-cellular EGFP ratio is inversely proportional to the amount of cellular EGFP, the corresponding "real" (i.e., unbiased) ratios should be even smaller than the calculated ones, resulting in an even more pronounced enrichment of A3G-EGFP when compared to EGFP. In theory, this technical problem can be overcome if the levels of cellular A3G-EGFP and EGFP are more closely matched, for instance, by using different amounts of plasmid DNA during transfections (i.e., less for EGFP and/or more for A3G-EGFP); however, in that case, exosomal EGFP should become undetectable if exposed side-by-side with exosomal A3G-EGFP; this, ironically, could turn into an even greater setback for relative quantification. Furthermore, while in the Western blot approach, the levels of exosomal A3G-EGFP and EGFP were normalized to the levels of their cellular counterparts, the flow cytometry approach accounts for the percentage of A3G-EGFP(+) and EGFP(+) cells, which is a completely different variable and somewhat equivalent to the "median fluorescence intensity" in flow cytometry. However, in the present study, the variance associated with the median fluorescence intensity was too high between replicates to be used as an estimator of protein amounts. Thus, while this study allows one to conclude that A3G-EGFP is enriched in exosomes, the true dimension of that enrichment is yet to be clearly demonstrated. Nonetheless, the data here presented indicates that A3G-EGFP is enriched in exosomes by a factor estimated between 2 and 6 when compared to EGFP.

Altogether, these results demonstrate for the first time that A3G is enriched in exosome fractions of at least the HEK293T cell line when compared to a protein that is not selectively included in exosomes (EGFP), further suggesting that selective mechanisms underlie the sorting of A3G into ILVs.

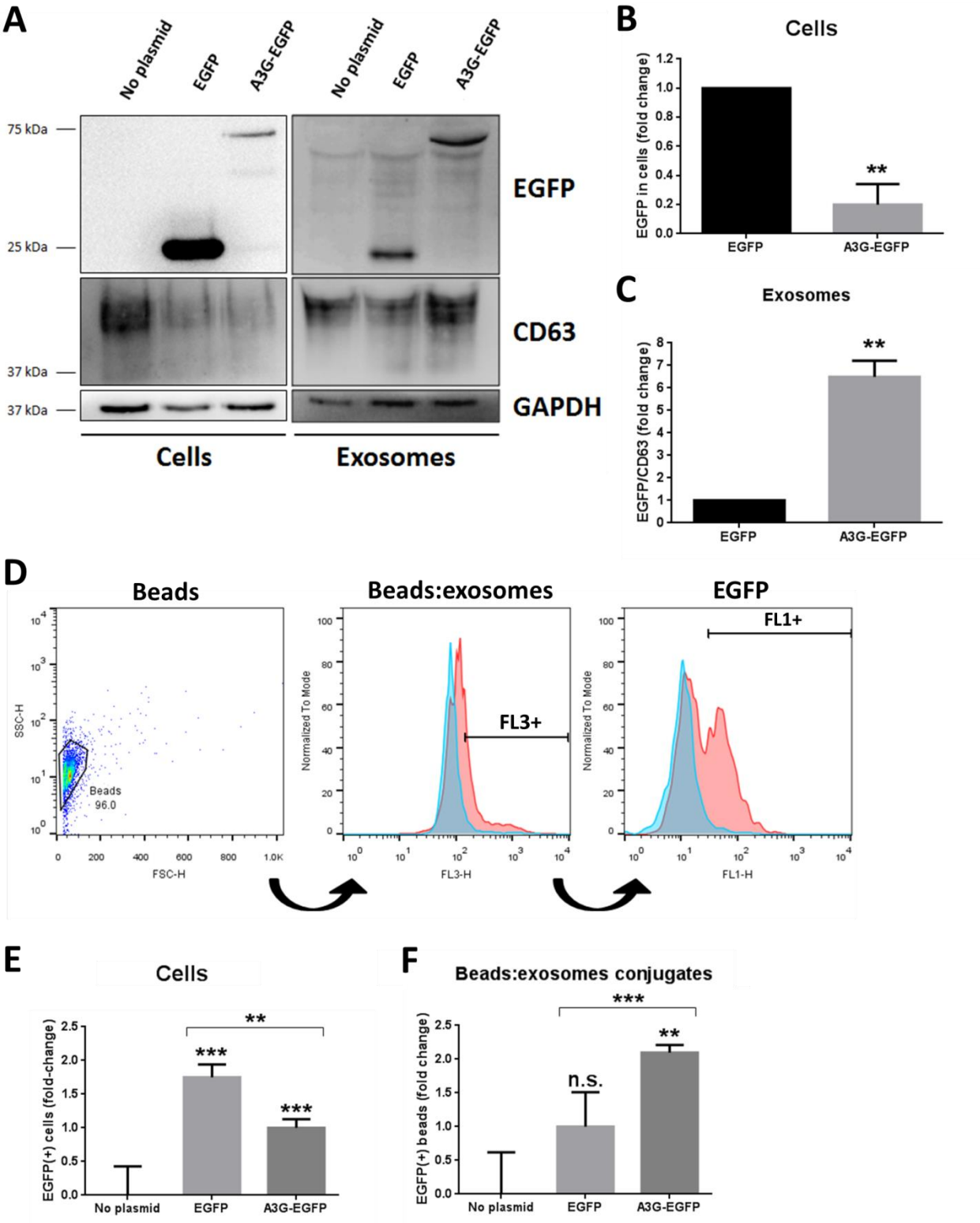


**Figure 7- An A3G-EGFP fusion protein is enriched in exosomes of HEK293T cells when compared to EGFP. (A)** Western blot analysis of HEK293T cells transiently expressing either A3G-EGFP or EGFP and exosomes isolated by ultracentrifugation from the SN of those cultures. **(B)** Cellular levels of A3G-EGFP and EGFP were estimated in fold change of EGFP as measured by Western blot quantification of cell lysates. An average 5±1-fold decrease in cellular A3G-EGFP relative to EGFP was calculated, although equal amounts of plasmid DNA were used to transfect cells (unpaired *t-test*; $p<0.01$; n=2). **(C)** The levels of exosomal A3G-EGFP relative to EGFP were calculated by normalizing the levels of each protein to their cellular levels and then to the levels of exosomal CD63 (exosome quantity marker) of the corresponding samples using the following formula: Exosomal EGFP(adjusted) = ((exosomal EGFP/cellular EGFP)/exosomal CD63). An average 6.5±0.5-fold enrichment in exosomal A3G-EGFP relative to EGFP was found in these conditions (unpaired *t-test*; $p<0.01$; n=2). **(D)** Gating strategy used in flow cytometry analysis of bead:exosomes conjugates

immunoisolated from cleared SNs of HEK293T cells expressing either A3G-EGFP or EGFP. Briefly, bead populations were gated in an FSC vs. SSC plot using FACS buffer as a negative control; a gate based on the levels of FL3 fluorescence was then applied to each bead population using a stained sample of beads incubated with PBS instead of SN as a negative control; finally, FL1(+) beads populations inside the FL3(+) subsets were gated using samples coming from cells transfected without any plasmid. The negative controls and treated samples correspond to the populations represented in blue and red, respectively. All samples were tested in triplicates. **(E)** Cellular levels of A3G-EGFP and EGFP were measured as the percentage of FL1/EGFP(+) or (-) cells. A 75±5% increase in cellular EGFP relative to A3G-EGFP was calculated (ratio paired *t-test*; p<0.01; n=2). **(F)** The levels of exosomal A3G-EGFP relative to EGFP were calculated by normalizing the percentage of FL1(+) beads in each sample to the corresponding cellular levels. A 2.1±0.1-fold increase in exosomal A3G-EGFP relative to EGFP was calculated (ratio paired *t-test*; p<0.001; n=2).

**The KFERQ-like motif of A3G is required for its enrichment in exosomes**

An enrichment of exosomal A3G-EGFP vs. EGFP by itself does not imply that active sorting mechanisms underlie this observation, as it is theoretically possible that a given protein is enriched in exosomes when compared to another due to passive mechanisms whose efficiency depends, for instance, of protein size, structure and/or electromagnetic properties. In fact, EGFP is a much smaller (~25 kDa) protein than A3G-EGFP (~75 kDa). Active sorting of A3G into ILVs would require some sort of interaction between A3G and (an)other molecular player(s) promoting its presence in exosomes at the cost of energy. I have previously shown that A3G contains a KFERQ-like motif comprising amino acids 168 to 172 ($QRELF^{168-172}$), which appears to be highly exposed in the protein's tri-dimensional structure (Fig. 6B-D) and could serve as a platform for interactions with molecular players linked to the active sorting of soluble cytosolic proteins into ILVs during exosome biogenesis, such as the chaperone Hsc70.

To determine if the KFERQ-like motif of A3G promotes the selective sorting of A3G into exosomes, an EGFP-tagged version of A3G bearing a double point mutation (Q168A/R169A), which has been shown to neutralize functional KFERQ-like motifs in other proteins (116), was generated by PCR using primers bearing the desired mutation. Given the corresponding migration rates in a SDS-PAGE (Fig. 9A), the overall integrity of the protein appeared to be unaffected. Moreover, as expected, A3G appears to be excluded from the nucleus of transfected cells when observed by fluorescence microscopy, as previously described (124) (Fig. 8B). Importantly, infection of HEK293T cells with lentivirus produced in the presence of wild-type or mutant A3G-EGFP (Fig. 8A) revealed that the antiretroviral properties of A3G are conserved upon introduction of both an EGFP tag and the Q168A/R169A mutations when compared to lentivirus produced in the presence of EGFP alone (Fig. 8C)(Two-way ANOVA; n = 2).

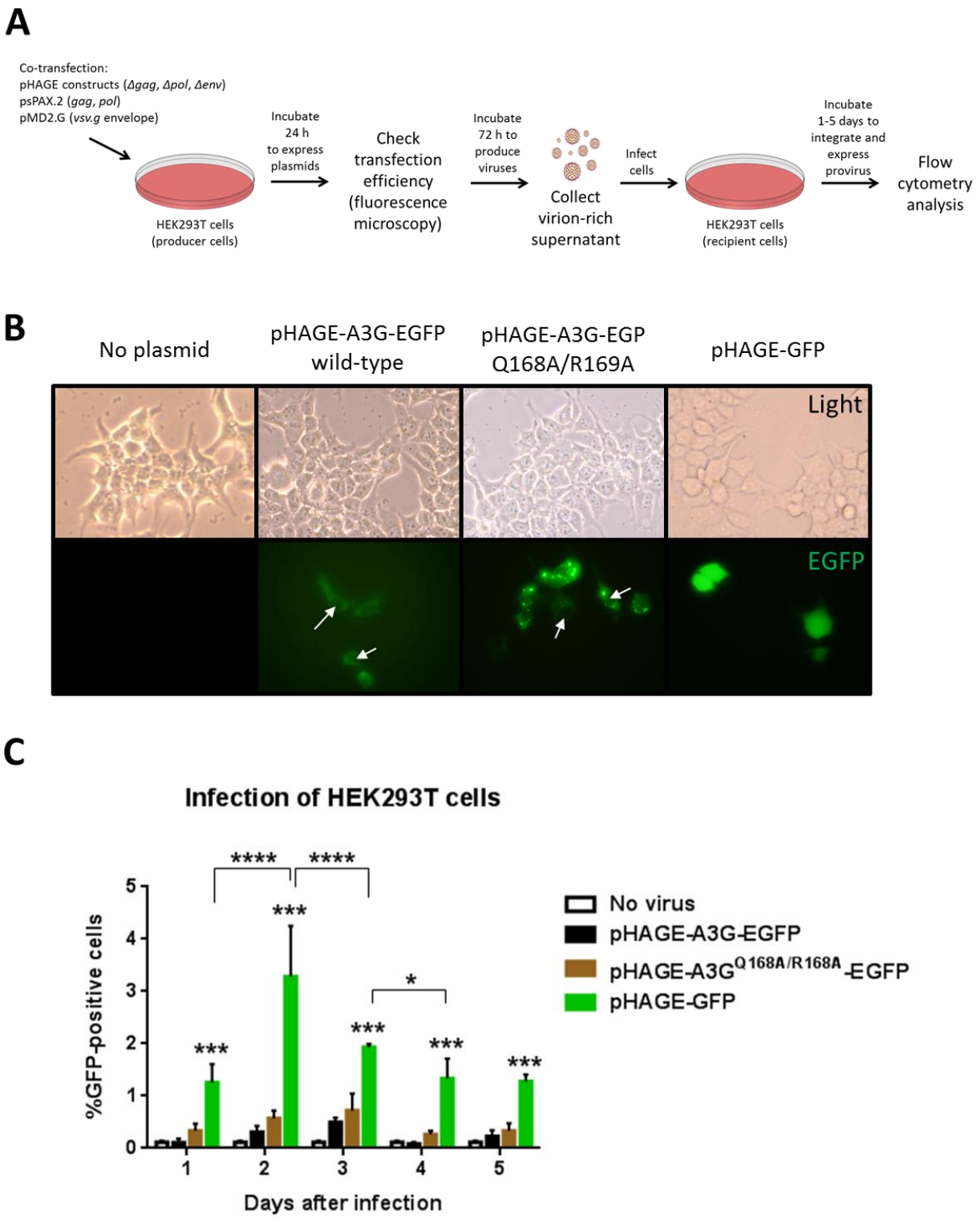


**Figure 8 - Preliminary characterization of the newly discovered A3G Q168A/R169A mutant. (A)** An outline of the protocol used to produce viral particles and infect cells. Wild-type, mutant A3G-EGFP, or EGFP alone, were cloned into the "pHAGE_puro" lentiviral backbone. To produce VSV.G pseudotyped lentiviral particles, HEK293T cells were co-transfected with the pHAGE constructs along with equimolar amounts of the psPAX.2 (*gag+pol*) and pMD2.G (*vsv.g*) accessory plasmids. After 24 hours of incubation, plasmid expression was verified using a fluorescence microscope, and the transfection medium was changed to fresh medium. The producer cells were then incubated for 72 hours to release viral particles, and the virus-rich supernatants were harvested and used to infect new cell cultures. The infected cells were collected at different time points and analyzed by flow cytometry. All samples were processed in

triplicates. **(B)** Fluorescence microscopy of HEK293T cell cultures transfected with different lentiviral-based plasmids encoding either wild-type, mutant A3G-EGFP, or EGFP alone. The white arrows point to the nuclei, where both wild-type and mutant A3G are not visible in substantial amounts, whereas in EGFP-expressing cells the nucleus is indistinguishable from the cytoplasm. **(C)** This system results in the production of viral particles assembled in the presence of whichever gene product is expressed from the lentiviral plasmid upon transfection of producer cells. Any differences in the infectivity of the virions produced should then result from the presence or absence of those gene products. Because A3G inhibits viral replication in the absence of Vif, it was expected that viral particles produced from the pHAGE-A3G plasmids would have been non-infectious while those produced from the pHAGE-EGFP plasmid would be infectious. In this case, while pHAGE-EGFP virus infected around 2% of cells ("repeated measures" two-way ANOVA with Tukey's multiple comparisons; $p<0.001$; n=2), both wild-type and Q168A/R169A mutant pHAGE-A3G-EGFP viruses were unable to achieve stable infection after five days post-inoculation. These results indicate that the Q168A/R169A mutation does not abrogate A3G's antiviral activity when compared to the wild-type.

Upon transfection with the different plasmids, no significant differences were found between the cellular steady-state protein levels of wild-type and mutant A3G-EGFP (ratio paired *t-test*; n=3)(Fig. 9A and B). In turn, Western blot analysis of exosomes isolated by ultracentrifugation from the SN of HEK293T cultures transfected with either wild-type or mutant A3G-EGFP revealed that the Q168A/R169A mutation significantly decreases the amount of A3G-EGFP present in exosomes when compared to wild-type A3G-EGFP (Fig. 9C). Quantification of protein levels in exosomes normalized to their cellular counterparts and then to exosomal marker CD63 revealed that the levels of mutant A3G-EGFP in exosomes was only 16±3% of the wild-type control (ratio paired *t-test*; $p<0.1$; n=3). Although CD63 is a protein that is present and enriched at the membrane of intracellular vesicles, it does not necessarily represent the amount of exosomes in a purified sample. For instance, when the levels of CD63 in two different samples are evaluated by Western blot and appear to be similar, it is possible that one sample contains more/less exosomes than the other but those exosomes are less/more enriched in CD63 for some reason. It is also possible that CD63 in a sample comes from intracellular structures other than exosomes, in the case where these samples are improperly cleared of cells and other vesicles before analysis. As a consequence of these and other arguments, some concern has risen over the years among researchers in the field of exosomes regarding the use of CD63 as a marker of exosome quantity in purified samples of exosomes. A recent paper showing that a subpopulation of exosomes that do not display enrichment in CD63 exists in typical exosome isolates has only increased the skepticism on the value of CD63 as a proxy for the quantity of exosomes (125). For these reasons, alternative exosome markers were used for quantification purposes. Flotilin-1 (FLOT1), which appears to be involved in the formation of caveolae and is also enriched in exosomes (126), was chosen as a representative of early exosome biogenesis, an indirect measure of exosome quantity. In turn, glyceraldehyde 3-phosphate dehydrogenase (GAPDH), a cytosolic soluble protein that is highly expressed in cells and easily observed in exosome extracts was chosen as a marker of exosome cargo, another indirect measure of exosome quantity. As shown in Figure 9C, when exosomal FLOT1 levels were taken into account, the levels of exosomal mutant A3G-EGFP were only 15±1% (ratio paired *t-test*; $p<0.1$; n=3) of the wild-type control. No significant differences were found when exosomal

GAPDH levels were taken into account (ratio paired *t-test*; n=3). The results using FLOT1 as an exosome quantity marker did not differ significantly from those obtained using CD63, supporting the notion that, unless the steady-state levels of the CD63 protein are expected to change as a consequence of treatments, at least in our settings, it can be used to demonstrate that the amounts of exosomes do not differ substantially between samples.

To further validate the observation that the Q168A/R169A of A3G inhibits its sorting into exosomes, the alternative immunoisolation- and flow cytometry-based experimental approach previously established was also undertaken in these conditions. Similar to the Western blot analysis shown in Figure 9A, no significant differences in cellular mutant A3G-EGFP were found, when compared to the wild-type A3G-EGFP control, upon transfection with equal amounts of plasmid DNA (Fig. 9D). In turn, wild-type A3G-EGFP(+) bead: exosomes conjugates were detected in significantly higher numbers than in the non-transfected control as well as in mutant A3G-EGFP(+) (ratio paired *t-test*; $p<0.0001$; n=2)(Fig. 9E).

The fact that exosomal mutant A3G-EGFP was detected by Western blot but not by flow cytometry analysis was likely due to the inferior sensitivity of our flow cytometry assay; supporting this idea is the fact that the exosomal mutant A3G-EGFP was only slightly observable in the Western blot assay (in which around 20 times more exosomes are used when compared to the immunoisolation/flow cytometry protocol). Accordingly, it is possible that by increasing the amount of antibody used to coat beads, the volume of cell culture SN used and/or the incubation time of beads with SN could ultimately increase the amount of exosomes isolated, thus increasing the amount of EGFP/FL1(+) exosomes detected. Nonetheless, an astonishing reduction in exosomal mutant vs. wild-type A3G-EGFP was observed (Fig. 9E), despite the inability to produce a respectable estimate of its magnitude.

We have screened the amino acid sequences of the human AID/APOBEC proteins and their murine, bovine and primate orthologs for KFEREQ-like motifs and found some degree of conservation of these motifs in the family, including a KFERQ-like motif in human A3G ($QRELF^{168-172}$) that is also present in the primate ortholog (Table 1). More importantly, using two complementary experimental approaches, we demonstrated for the first time an enrichment of exosomal A3G-EGFP when compared to EGFP, a protein that is not selectively included in exosomes (Fig. 7), indicating that selective mechanisms underlie the sorting of A3G-EGFP into exosomes. Furthermore, this enrichment is precluded by the Q168A/R169A mutation (Fig. 9), a region of A3G that was not previously linked to any function (Fig. 6). The stability of the A3G Q168A/R169A mutant appears to similar to that of its wild-type counterpart, as observed by their migration on SDS-PAGE and their subcellular localization (Fig. 8)

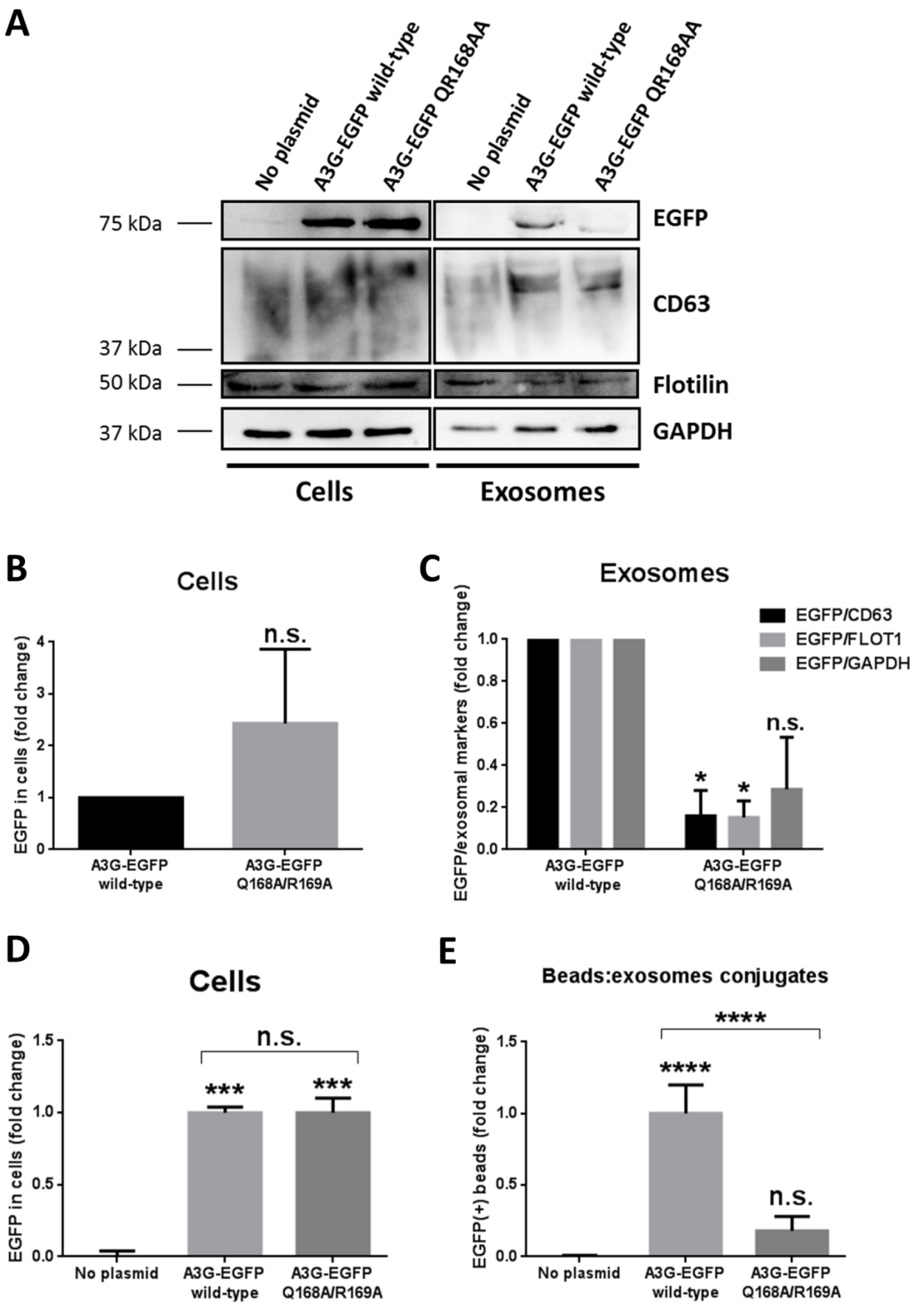
A
No plasmid
A3G-EGFP wild-type
A3G-EGFP QR168AA
No plasmid
A3G-EGFP wild-type
A3G-EGFP QR168AA
75 kDa
37 kDa
50 kDa
37 kDa
EGFP
CD63
Flotilin
GAPDH
Cells
Exosomes
B
Cells
EGFP in cells (fold change)
n.s.
A3G-EGFP wild-type
A3G-EGFP Q168A/R169A
C
Exosomes
EGFP/CD63
EGFP/FLOT1
EGFP/GAPDH
EGFP/exosomal markers (fold change)
n.s.
*
*
A3G-EGFP wild-type
A3G-EGFP Q168A/R169A
D
Cells
EGFP in cells (fold change)
n.s.
***
***
No plasmid
A3G-EGFP wild-type
A3G-EGFP Q168A/R169A
E
Beads:exosomes conjugates
EGFP(+) beads (fold change)
****
****
n.s.
No plasmid
A3G-EGFP wild-type
A3G-EGFP Q168A/R169A

**Figure 9 – The KFERQ-like motif of A3G is required for its enrichment in exosomes. (A)** Western blot analysis of HEK293T cells transiently expressing either A3G-EGFP or EGFP and exosomes isolated by ultracentrifugation from the SN of those cultures. **(B)** Relative cellular levels of A3G-EGFP wild-type and A3G-EGFP Q168A/R169A as measured by Western blot quantification of EGFP of cell lysates. No significant differences were found between the two averages (ratio paired *t-test*; n=3). **(C)** Relative amounts of wild-type and mutant A3G-EGFP in exosomes when normalized to cellular EGFP and the different exosome markers. An overall marked decrease was found in exosomal mutant A3G-EGFP when compared to the wild-type (ratio paired t-tests; n=3). **(D)** Cellular levels of wild-type and Q168A/R169A-mutant A3G-EGFP in transfected HEK293T cells as measured by the percentage of FL1/EGFP(+) or (-) cells. While both groups were significantly distinct from the "no plasmid" control (ratio paired *t-test*; $p<0.001$; n=2), no significant differences were found between the two transfected groups. **(E)** The relative levels of exosomal wild-type and mutant A3G-EGFP were calculated by normalizing the percentage of FL1(+) beads in each sample to the corresponding cellular levels. Wild-type A3G-EGFP(+) beads were readily detected (ratio paired *t-test*; $p<0.0001$; n=2) while no significant differences were detected between mutant A3G-EGFP and the "no plasmid" negative control.

## FUTURE DIRECTIONS

The next step in this line of research will be to determine if A3G interacts or binds with the chaperone Hsc70, previously linked to the sorting of cytosolic KFERQ-containing proteins into lysosomes and now to exosomes. If so, and if this is the mechanism impaired by our A3G Q168A/R169A mutant, we expect that protein-protein interaction assays such as co-immunoprecipitations and proximity assays yield a positive result for interactions between Hsc70 and wild-type A3G and a negative result for the Q168A/R169A mutant. Furthermore, the influence of the LAMP2A receptor will also be determined, following undergoing works in Dr. Paulo Pereira's laboratory.

I am currently optimizing a medium-throughput functional assay that allows for the quantification of virus infectivity and antiretroviral potential of specific treatments with at least two different outputs. This assay mostly consists of pre- or post-treating cells and infecting them with the different HIV reporter constructs in our possession, such as a Luciferase-expressing HIV plasmid (pNL4-3-Luciferase) or the classical GFP-expressing pNL4-3-GFP. Furthermore, we expect to increase the scope of our analysis to "virus latency vs. virus activity" by using the doubly fluorescent pRGH-WT plasmid, an HIV latency/activity reporter-construct that expresses mCherry under a constitutive CMV promoter upon integration in the host's genome and GFP under the 3'LTR promoter (127). I am also currently working on three different A3G "knockout" variants of the non-permissive H9 cell line using a CRISPR/Cas9-based approach, as well as an shRNA-based A3G "knockdown" variant. Moreover, I have also recently generated the A3G E67Q, E259Q and E67Q/E259Q "null" mutants. These tools will not only serve as controls for most of our further experiments, but will also allow for the confirmation of previous results in the literature that were never tested using this type of

models. To further validate our results, the enrichment of exosomal A3G compared to EGFP, as well as the impact of the Q168A/R169A mutation in the levels of exosomal A3G will be tested using non-EGFP-tagged versions of the protein and A3G-specific antibodies. Additionally, the amounts of exosomal A3G will be compared to other proteins that either have or lack functional KFERQ-like motifs.

In the medium term, we are interested in designing strategies that allow for the enrichment of certain molecules in exosomes and in maximizing the potential benefits of naturally- and synthetically-enriched exosomes as a delivery vector for the ingredients involved in a multitude of cellular and molecular processes. The links between the exosome and virion biogenesis pathways will be explored under the scope of intra- and intercellular trafficking of the natural endogenous antiretroviral factors.

Our entire sequence analysis will eventually be expanded to KFERQ-like motifs generated post-translationally and our KFERQ-like motif functional analysis will be expanded to the remaining members of the AID/APOBEC family containing at least one of those motifs. We are also naturally interested in further characterizing the A3G Q168A/R169A mutant, a task that simultaneously fits on both a “DNA-oriented” lab such as that of Dr. Vasco Barreto’s and a “protein-oriented” lab such as that of Dr. Paulo Pereira’s. Also, in the near future, with the help of Dr. Manuel Vicente, we will release a free online tool similar to that described in Cuervo et al. (111), which allows for high throughput detection of putative and encoded KFERQ-list motifs in amino acid sequences.

## ON THE RELEVANCE OF THIS WORK

### Current situation and challenges in the field of HIV infection

Presently, around 37 million people worldwide live with HIV, of which about 25% do not know their HIV status. Out of those who know their HIV status, only around 60% are on ART, with a success rate slightly under 50% (i.e. fewer than 50% of people receiving ART progress to a state of undetectable levels of HIV)(128). From 2008 through 2014, the CDC estimated an annual HIV incidence decline of 18% in the USA. However, these declines occurred especially among heterosexuals, people who inject drugs (PWID), and white men who have sex with men (MSM); in contrast, no decline was observed among black MSM, and an increase was even documented among Latino MSM(129).

Africa is, by far, the most affected region by HIV/AIDS in the world. In 2017, the East and Southern Africa regions reported 800,000 new infections, which already represents a decrease of 30% since 2010. In the

same period, West and Central Africa reported 370 000 new infections, 8% less than in 2010. Only in 2017, 660,000 Africans died of AIDS-related illnesses (130).

Portugal exhibits one of the highest annual rates of new HIV and AIDS diagnoses in the European Union (EU). In a country with close to 11 million inhabitants, around 60 000 (0.005%) people had been diagnosed with HIV infection by the end of 2017, of which, more importantly, around 40% progressed to AIDS. From January 1 until June 30 of 2017, 1,068 new HIV infections were reported (10.4 new cases/$10^5$ inhabitants). Around 70% of the new patients were men with the highest rate of infection found in individuals between 25 to 29 years old. MSM accounted for 64.1% of cases in those under the age of 30. The predominant mode of transmission was unprotected sexual contact, accounting for 60% of cases in heterosexual patients and 40% in MSM. Only 1.8% of the new diagnoses corresponded to PWID. Around 50% of the new reported cases already had a CD4+ T-cell count <350 cells/$mm^3$ of blood. Despite these cases, a decrease of 40% for HIV cases and 60% for AIDS cases was reported in Portugal between 2007 and 2016 (131).

A number of (inter)governmental projects regarding the fight against the HIV/AIDS epidemic have arisen recently. One such project is the "HIV Epidemic Initiative", launched by the US government in 2019. The main goals of this project are to reduce the incidence of HIV infections nationally by 75% in five years and by 90% in ten years (132). Another well-known initiative is the "90-90-90" project founded by the United Nations, whose ultimate goals are to achieve (i) 90% of all infected people knowing their HIV status, (ii) 90% of all the people diagnosed with HIV infection receiving sustained ART, and (iii) 90% of the people receiving ART to have viral suppression (128) These projects seem to be more directed towards public awareness and behavioral/social intervention, and less towards advances in biomedical research. Ironically, the ambitious goals established are based on predictions made from estimates of the rate at which advances in biomedical research occur, among other statistics. Moreover, despite the "true" accuracy of such predictions, projects like these are naturally thought to promote progress in the field by creating goals that break important barriers and which are often not met in due time but nonetheless promoted important advances. Unlike what such optimistic goals may suggest, the AIDS epidemic is still highly incident and prevalent in many populations.

Concerning the possible contributions from the field of biomedical research, efforts to refrain or end the HIV/AIDS epidemic can be broadly grouped into three categories, as it often occurs for other infectious diseases: (i) decreasing the incidence of new infections, (ii) stabilizing patients with acute infections, and (iii) curing acquired infections.

**Decreasing HIV incidence: Preexposure prophylaxis (PrEP)**

Prevention of new HIV infections in individuals at risk can be achieved to some extent using preexposure prophylaxis (PrEP). The only medication approved by the Food and Drug Administration (FDA) for PrEP in healthy adults at risk of acquiring HIV infection is a daily dose of tenofovir disoproxil fumarate with or without emtricitabine (TDF/FTC) (133). TDF and FTC belong to a class of antiretroviral drugs known as nucleotide analogue reverse transcriptase inhibitors (nRTIs) and are thought to prematurely terminate reverse transcription upon binding to the HIV reverse transcriptase (134,135). PrEP using TDF/FTC effectively decreased HIV transmission in multiple populations with disproportionally high rates of HIV incidence, including homosexual men (136), transgender women (136), IV drug users (137), heterosexuals engaging in high-risk behaviors (138) and HIV-serodiscordant partnerships(139). Despite this, out of the 1.1 million North Americans with indications for PrEP (140) only about 77000 (7%) received prescriptions in 2016 (141). Moreover, although the majority of new HIV infections in the USA occur in Black (43%) and Latino/Hispanic (26%) populations (129), only 11% and 13% of patients receiving PrEP were Black and Latino, respectively.

Although the substantial effectiveness of PrEP with TDF/FTC in reducing the incidence of HIV infection has been demonstrated in several clinical trials and observational studies (142–145), this strategy has not yet been sufficient to counter the HIV epidemic completely. In one hand, patient adherence to treatments, which can be influenced by a wide range of variables including insurance barriers or scheduling challenges (144), is highly limiting; as an example, a study on the efficacy of PrEP in 50 patients from Providence (Rhode Island, USA) at risk of acquiring HIV infection resulted in only 19 patients (38%) attending a 6-month follow-up consultation (143). On the other hand, as is the case with most drugs, the efficiency of TDF/FTC can decrease due to several genetic and environmental factors such as drug-resistant HIV strains (143,144) and feminizing hormone treatments(146). There is also a concern that PrEP may positively select drug-resistant HIV strains, which could decrease its efficiency; for an updated review on this subject see (147).

**Stabilization of patients with acute infection: Antiretroviral Therapy (ART)**

When a patient acquires an HIV infection, preventive measures are rendered useless and active mitigation of viral replication is required to minimize the onset of AIDS. Several drugs can be used to refrain HIV-1 replication, of which protease inhibitors, which bind HIV-1 protease and block the proteolytic cleavage of protein precursors, are the most well-known. Other drugs can be effective in ART, such as reverse transcriptase inhibitors, fusion inhibitors and integrase inhibitors. For a detailed list of ART drugs and their mechanisms of action see (148). These drugs have been highly effective at decreasing the levels of HIV replication in patients in the patients that either seek them or have access to them.

Despite the great success of ART, this strategy is not sufficient to completely eliminate HIV in infected individuals. Among other possible reasons, this occurs due to the existence of treatment-resistant self-replenishing viral reservoirs in infected patients, mainly consisting of CD4(+) T cells infected with replication-competent HIV that remains innocuous (or even dormant) at low levels during therapy (149–151) but can rapidly propagate if ART is ceased (152–154). Notably, the timing and duration of ART induction appear to influence the dimensions of the viral reservoir and the temporal extent of transient aviremia after ART cessation. Patients who started ART when the infection had already reached a chronic state retained significantly more replication-competent HIV proviral copies than those who started ART when the infection had not reached a chronic state. In other words, the HIV reservoir is mostly established in the early phases of infection, which goes in line with an important study showing that the majority of integrated HIV-1 is latent shortly after infection (127). Importantly, ART at that stage seems to abrogate part of the establishment of such reservoir. This may contribute to the fact that "late" ART appears to result in much shorter viral rebound times than "early" ART, with an average aviremia time after ART cessation of 9 days vs. 50 days, respectively (155). Other studies using different populations and methodologies to correlate the phase of infection at which ART was started with the dynamics of viral rebound clearly demonstrate the same trends whilst pointing to an even more accentuated difference between these two treatment groups (156–158). These and other studies clearly demonstrate that infected patients receiving ART will experience viral rebound back to pre-treatment levels regardless of the level of HIV in their CD4+ compartment at the time of ART cessation. Consequently, this also increases the onward transmission risk (159). Thus, despite the remarkable success of ART, it is by itself insufficient to eradicate HIV from populations. Adding to this, a substantial prevalence of drug-resistant HIV strains in both untreated (160) and treated (161,162) patients from low- and medium-income countries is also a problem.

### Curing HIV infection

Although ART by itself is incapable of curing HIV infection (at least in a timely manner), it can efficiently reduce the number of infected cells and HIV copies down to vestigial levels, remarkably turning HIV infection into a chronic condition rather than a lethal one. As stated previously, the existence of drug-resistant HIV reservoirs in patients undergoing ART is a major contributor to viral rebound after ART cessation (152–154). Naturally, specifically targeting these cell populations is a promising strategy to completely eliminate HIV from patients. This, however, is yet to be achieved. Although it might seem somewhat trivial to target a small population of HIV-infected cells that remain in the host upon ART induction, the elusiveness and overall heterogeneity of these cells, and the fact that, in theory, even a very small percentage of the HIV reservoir is sufficient to result in viral rebound after ART cessation, have turned this seemingly simple task into an unresolved one. Although the biological characteristics of the HIV reservoir remain to be fully understood, some light has been shed on the topic. Rebound viremia appears

to originate from a small subset of the total infected cells that persist during ART (163,164), of which an even smaller subset seems to be maintained through clonal expansion, as observed by the lineage dynamics of rebound-associated cells. Also, most of these cells appear to express HIV RNA during ART, suggesting that they specifically resist treatment (164).

Few other strategies attempting to cure HIV infection in the setting of ART treatment (and outside of it) have been explored with promising results. These were mostly designed around knowledge regarding the "so-called" HIV co-receptors, CCR5 and CXCR4, which are chemokine receptors belonging to the vast superfamily of G protein-coupled receptors (GPCRs)(165–167). As such, binding of CCR5 and CXCR4 to their cognate ligands induces a conformational change that results in phosphorylation of intracellular portions of the receptor at serine and threonine residues, in turn enabling receptor coupling with an intracellular G protein (168). The G protein propagates the signal by regulating intracellular enzymes. This mechanism of cellular signaling is conserved throughout eukaryotic evolution, taking part in numerous molecular pathways. Consequently, GCPR dysfunction is the cause of several diseases, and about half of the pharmaceutical drugs on the market target GCPRs (166). CCR5 was first isolated as a GPCR that is antagonized by three different chemokines produced by CD8(+) T cells (RANTES MIP-1α and MIP-1β)(169,170), which had been previously shown to suppress HIV replication (171). In mice, despite resulting in normal development, a "CCR5 -/-" genotype has been linked to impaired macrophage function and enhanced T cell-dependent immune responses (172). More importantly, an estimated 2-3% of white Europeans possess a "CCR5 -/-" genotype (the *ccr5 Δ32* deletion), all without any obvious associated phenotypes (173); these people appear to be highly protected against HIV infection, although this protection is not absolute, supposedly due to HIV strains found in patients that mostly use CXCR4 (X4 HIV) or both CCR5 and CXCR4 (along with CD4) as entry co-receptors (dualtropic R5X4 HIV)(174–180). CXCR4 was first isolated back in 1993 (181,182) and three years later was found to be an HIV co-receptor (183), followed by the discovery of its endogenous ligand, SDF-1/CXCL12 (184,185). Of the different isoforms of CXCL12, some display more potent anti-HIV properties than others (186). The high degree of conservation of the CXCR4/CXCL12 axis is demonstrated by the phenotypes associated with its absence: mice that have a "CXCR4 -/-" genotype die *in utero* due to defects in vascular development, hematopoiesis, and cardiogenesis (187); in turn, mice with a "CXCL12 -/-" genotype display deficient B-lymphopoiesis and myelopoiesis, and incomplete neuronal and cardiovascular development. During HIV-1 entry, CCR5 and CXCR4 interact with CD4-activated gp120, triggering a cascade of reactions that ultimately results in the fusion of viral particles with the host cell's membrane (188–191).

The most polemic experiment of the last years was the CRISPR/Cas9-based genetic modification of two human embryos that were posteriorly developed *in utero.* In these experiments, the Chinese scientist He

Jiankui claimed to have successfully "knocked out" the *ccr5 locus*, generating two supposedly HIV-immune babies. A more consented case in which the ccr5 genotype of a patient was changed in order to control HIV-1 infection was that of the "so-called" Berlin patient", Tymothy Ray Brown, a 40-year-old man who was diagnosed with acute myeloid leukemia (AML) and previously with HIV infection. As a treatment for his leukemia, he received two allogeneic stem-cell transplantation with CD34(+) peripheral blood stem cells from a donor who was homozygous for the *ccr5 Δ32* deletion (192–194). ART was stopped on the day of the first transplantation. Surprisingly, even after 20 months of ART cessation, no active HIV could be detected in this patient. In 2009, the patient was officially announced to be cured HIV infection, and for as long as viral loads remained undetectable, no further ART would be required(192–194). Ten years after, the "Berlin patient" remains HIV-free, confirming the potential of this type of treatment to eradicate HIV infection (194).

A second, similar case of success was that of the so-called "London patient" (195), who remains anonymous to this day. This patient was diagnosed with Hodgkin's lymphoma in 2012, nine years after receiving a positive HIV test. Having also received a hematopoietic stem-cell transplant from a donor that was purposely selected for homozygosity for the *ccr5* Δ32 variant, he achieved lymphoma remission soon after. Analytical ART interruption was initiated less than two years after transplantation. Weekly analyses of plasma viral load were done for the first three months after ART interruption and monthly thereafter, with HIV being undetected in every instance. Several other tests relying on different markers confirmed these results as well as the absence of CCR5 expression. Moreover, while hematopoietic stem-cell transplantation by itself significantly reduces the HIV reservoir (which is not enough to cure HIV), the potential of the *ccr5* Δ32 mutation in these cases is demonstrated by the viral rebounds observed in three HIV patients receiving transplantations from wild-type *ccr5* donors(196,197).

**Final remarks**

The works in this study follow a "fresh" line of thinking about the current challenges in the eradication of HIV-1 infection. Having found that the presence of KFERQ-like motifs in APOBEC proteins seems to display some degree of conservation along the mammalian lineage, and that the major antiretroviral A3 protein, A3G, is selectively packaged into exosomes in a KFERQ-like motif-dependent manner, it is tentative to think that the mechanisms underlying these intra- and extracellular trafficking routes play an influential role in the maintenance of proteostasis, and consequently, in the efficiency of the immune response to viral infections. If at least some enveloped viruses, such as HIV-1, evolved to "hi-jack" the endocytic pathways in order to replicate and assemble infectious virions, selective shuttling of endogenous antiviral factors through the same pathways would maximize the efficiency of those factors by increasing their availability

at the specific locations where they are most required. An obvious example that follows this model is the packaging of A3G in HIV-1 virions, which substantially increases the concentrations of this DNA-editing enzyme at the specific regions in the cytoplasm where reverse transcription is about to take place; because the infected cell (an activated CD4+ T-cell) also expresses high amounts of A3G and, consequently, the virions it produces also contain A3G, viral bypass of the immune response and successful establishment of stable infection in these cells implies the perpetuation of the antiviral mechanisms described.

We intend to thoroughly understand the mechanisms underlying the proteostasis of endogenous antiretroviral factors, the natural APOBEC:Vif equilibrium and the connections between the shuttling of A3G into exosomes and into viral particles. Although we have demonstrated that the Q168A/R169A mutation of A3G does not significantly decrease its antiviral properties using a lentivirus-based vector in the absence of Vif, it is unknown if this mutation has any impact in the interactions between A3G and Vif and/or in the packaging of A3G into viral particles. If Q168A/R169A also prevents the packaging of A3G into HIV virions, a parsimonious hypothesis would be that the pathways shuttling A3G into exosomes and viral particles are intertwined or even completely shared. Thus, the fraction of viral particles containing A3G would positively correlate with the influence of the “endosomal virus assembly line” (as opposed to its cell-membrane counterpart) in stabilizing that particular infection. In this case, new therapeutic targets with the potential to restrict viral replication could be unveiled. If, in turn, Q168A/R169A precludes the A3G:Vif interaction, this would result in failure of HIV-1 to shuttle A3G for degradation in the 26S proteasome, consequently increasing its steady-state concentrations and antiviral efficiency. In that case, simple overexpression of mutant A3G in both permissive and non-permissive cells would result in a “dominant-negative” Vif-resistant phenotype. The impact of such manipulations may even project to other related topics, such as the establishment and maintenance of an HIV-1 reservoir in infected patients, or infections with other viruses. With this knowledge, we intend to design strategies that effectively modulate cellular and molecular processes taking place before and after viral infection, in order to substantially increase the efficiency of the natural antiviral immune response, ultimately increasing the availability of effective therapeutic alternatives and respond to the current lack of viable strategies to effectively cure HIV-1 infection.

## MATERIALS AND METHODS

### Screening of amino acid sequences for KFERQ-like motifs

The NCBI Sequence Database accession codes of the amino acid sequences analyzed are listed: A3A (NP_663745.1), A3B (NP_004891.4), A3C (NP_055323.2), A3D (NP_689639.2), A3F (NP_660341.2), A3G

(NP_068594.1), A3H (NP_001159475.2). The predicted tertiary structures of these proteins were analyzed using the models available at UniProtKB (http://www.uniprot.org/) and SWISS-MODEL (http://www.swissmodel.expasy.org/) with the following accession codes: Q96AK3 (A3D), Q8IUX4 (A3F), Q9HC16 (A3G). The tridimensional model in Fig. 6 was done using PyMol and the A3G accession previously mentioned. The solvent accessibilities were calculated using the JPred4 algorithm (121) (available at http://www.compbio.dundee.ac.uk/jpred/index.html).; an amino acid was considered "accessible" when its solubility score was >0.25, as suggested by (111,121).

**Plasmids and genetic constructs**

All A3G constructs were based on a template cDNA of A3G kindly provided by Reuben S. Harris, which we further cloned into the pGEM-t® Easy cloning plasmid vector using the EcoRV restriction site. A FLAG tag encoding-sequence (5'-GATTACAAGGACGACGATGACAAG-3') was inserted 3' of the A3G sequence (pGEM-A3G-FLAG) by PCR. The C-terminus EGFP-tagged A3G constructs (pCMV-A3G-EGFP) were produced by direct restriction cloning using the EcoRI/KpnI sites of both pGEM-A3G-FLAG and pEGFP-N3 (Clontech, California, USA). The FLAG tag is not conserved in these constructs. The lentivirus-based A3G plasmid constructs (pHAGE-A3G) were produced by restriction cloning using the EcoRI/NotI restriction sites of both pCMV-A3G-EGFP and pHAGE_puro (pHAGE_puro was a gift from Christopher Vakoc (Addgene plasmid # 118692 ; http://n2t.net/addgene:118692 ; RRID:Addgene_118692)(87). All restriction enzymes were bought from NEB and used according to the manufacturer's instructions unless when stated otherwise. The T4 DNA Ligase (Thermofisher) was used for the ligation steps following the brand's instructions unless stated otherwise. Ligation products were transformed into DH5α *E. coli* using the "heat-shock" technique (198) and positive colonies were selected by colony-PCR (199).

**Production of A3G mutants**

To generate A3G mutants, the PCR technique was used. All mutagenesis reactions were done over the A3G cDNA previously cloned into pGEM-t Easy™ (Promega, Wisconsin, USA), which contains T7/SP6 sites flanking the cloned inserts. To minimize errors introduced by DNA polymerase in these reactions, the High-Fidelity Phusion™ DNA Polymerase was used according to the manufacturer's protocol (Thermofisher, Massachusetts, USA). Complementary primer pairs (Fw and Rv) bearing the desired mutations were designed. In the first step of mutagenesis, two partially complementary DNA fragments were produced from 75 ng of template DNA by combining the mutagenic primers with primers that bind to the flanking regions (Fw/SP6 and T7/Rv respectively). The reaction conditions were 30 seconds at 98ºC, 30 cycles of 10 seconds at 98ºC, 30 seconds at 53ºC and 1 minute per kilobase of amplified DNA, followed by 7 minutes at 72ºC. The amplified fragments were run on an agarose gel and the DNA was purified using the PCR & Gel

Band Purification Kit (GRiSP). In a second PCR step, equimolar amounts of each partially complementary fragment were used to amplify a fragment corresponding to the full-length mutated A3G cDNA flanked by T7/SP6 regions. The reaction conditions were similar to those stated previously but lasting 20 cycles using adjusted extension times. Finally, T7/SP6 primers were added to the prior reaction, the extension times were adjusted to enable the polymerization of the full T7-A3G-SP6 fragment (1 minute per kilobase of DNA) and the reaction was run for another 30 cycles. The mutated A3G inserts were purified by agarose gel electrophoresis and cloned into pGEM-t Easy™ (Promega). To do this, A-overhangs were added to the blunt-ended inserts by incubating for 20 minutes at 72ºC with GoTaq® DNA Polymerase and 200 mM of dNTPs (as suggested in NEB's web protocol). Further, ligation of inserts to the pGEM-t Easy™ backbone previously digested with EvoRV (NEB, Massachusetts, USA) was done using T4 DNA Ligase (Thermofisher) following the brand's protocol. The ligation products were transformed into competent DH5alpha *E. coli* and the resulting colonies were screened by colony-PCR.

**Transfections and viral infections**

All transfections in this study were done using the "calcium phosphate method" (REF) and either treated and analyzed until 48 hours after transfection or maintained for longer to collect the supernatant and produce virus stocks.

To produce VSV-G-pseudotyped viral particles, HEK293T cells were seeded in 6-well plates, cultured until 80% of confluence, and transfected using the "calcium phosphate method" (123) with the desired plasmids. For pHAGE-based viral particles, both pMD2.G (containing *VSV-G*) and psPAX.2 (containing *gag* and *pol*) were simultaneously co-transfected along with the pHAGE constructs at equimolar amounts. For pNL4-3 Δenv GFP and pRGH-WT only pMD2.G co-transfection is required to produce infectious virions. At 24 hours post-transfection, cells were checked for plasmid expression on a fluorescence microscope and incubated with a minimal amount of medium for 72 hours to collect viral particles. After this, the conditioned medium was frozen at -80ºC for at least 24 hours and the transfection efficiency of virus-producing cells was evaluated by flow cytometry.

To infect HEK293T cells, ~200 000 cells/well (24-well plates) were seeded 24 hours in advance. A desired volume of viral extract (the same for all samples) was added to the culture medium along with polybrene at a final concentration of 1 ug/mL. Then, cell plates were centrifuged at 1200 xg (25ºC) for 90 minutes and the conditioned medium was replaced by fresh medium. After 1 to 5 days of incubation, cells were fixed in 2% of paraformaldehyde (final concentration) for 30 minutes at 4ºC and analyzed by flow cytometry.

**Exosome purification by ultracentrifugation**

Exosomes were purified as described previously (200). Cells were cultured for 24 hours in exosome-depleted medium prepared by supplementation with fetal bovine serum previously centrifuged overnight at 120,000 xg to remove bovine exosomes and other contaminants. The cell supernatants were harvested, centrifuged at 300 xg for 10 minutes to pellet cells and then at 16,500 xg for 20 minutes to pellet microvesicles and other debris. Additionally, to thoroughly purify exosomes, supernatants were filtered through 0.22-µm syringe filters. To pellet exosomes, pre-cleared supernatants were centrifuged at 120,000 xg for 70 minutes. The supernatant was discarded and the pellet was washed with PBS. After a second round of centrifugation at 120,000 xg, the pelleted exosomes were resuspended in a small volume of PBS (usually 30 µL of PBS to resuspend the exosomes extracted from ~40 mL of supernatant) and either used immediately or stored at -20ºC for up to a week.

**Western blot analysis**

The total protein concentration of samples was determined using the Pierce™ BCA Protein Assay Kit (Thermofisher). Cell lysates and exosome extracts were denatured with Laemmli buffer, heated at 95ºC for 5 minutes and separated by SDS-PAGE using a running buffer containing 0.2 M glycine, 0.025 M Tris and 0.1% SDS. The resolved samples were then "wet"-transferred into 0.45-µm Immobilon-P® PVDF membranes (Milipore) using transfer buffer containing 0.2 M glycine, 0.025 M Tris, 20% methanol and 0.25% SDS. These procedures were done using a Mini-Protean® II Electrophoresis Cell (BioRad). Membranes were then blocked for 1 hour with TBS-T (140 mM NaCl, 3 mM KCl, 20 mM Tris and 0.005% Tween-20) containing 5% skim milk and probed using the following polyclonal primary antibodies: goat anti-GFP (SICGEN, Cat. No.: AB0020-200; 1:1000 dilution), goat anti-CD63 (SICGEN, Cat. No.: AB0047-200), rabbit anti-Flotillin-1 (Santa Cruz Biotechnology, Cat. No.: sc-25506), goat anti-GAPDH (SICGEN, Cat. No.: AB0049-200); all primary antibodies were diluted in blocking solution at 1:1000. After overnight incubation at 4ºC, membranes were washed thrice for 5 minutes in TBS-T with gentle horizontal shaking. To visualize the probed membranes, a polyclonal rabbit anti-goat HRP-conjugated secondary antibody (DAKO) was used at a 1:5000 dilution in blocking solution. Finally, the probed membranes were visualized in a ChemiDoc™ Touch Imaging System (BioRad). Stripping of antibodies for reprobing of membranes was done by incubating with stripping buffer containing 6M GnHCl, 0.2% Nonidet P-40 (substitute), 0.1M β-mercaptoethanol and 20mM Tris-Hcl (pH 7.5) twice for 5 minutes with gentle horizontal shaking(201).

**Immunoisolation of exosomes with anti-CD63-conjugated beads and flow cytometry analysis**

Aldehyde/sulfate latex beads (4% w/v, 4 µm of diameter) (Thermofisher) were used for immunoisolation. To conjugate beads with antibody, the Thermofisher's "Passive Adsorption Protocol" was followed thoroughly with the exception that MOPS buffer (0.02 M MOPS, 0.005 M sodium acetate, 0.001 M

Na2EDTA, pH 7) replaced the MES buffer; since the MES buffer was not available, care was taken to avoid the use of phosphate-based buffers before antibody conjugation, which according to the brand can negatively impact the process. Briefly, 2.5 mL (40 mg/mL) of beads were washed twice in 10 mL of MOPS buffer by centrifuging at 3 000 xg for 20 minutes and resuspended in 5 mL of MOPS buffer. To coat beads with antibody, 5 mL of MOPS buffer containing 1.25 mg of goat anti-CD63 (SICGEN, Cat. No.: AB0047-200) antibody was added to the 5 mL of prepared beads and incubated with gentle mixing at room temperature overnight. The antibody-conjugated beads were then washed thrice in 10 mL of PBS to remove unbound protein and finally resuspended in 10 mL of Storage Buffer containing 0.1 M PBS (pH 7), 0.1% glycine and 0.1% sodium azide.

The immunoisolation protocol used in this study was designed/optimized by the author, mostly based on the protocol provided by Thermofisher and those developed by Théry et al. (202). To immunoisolate exosomes, HEK293T cells were cultured in 6-well plates until ~80% of confluence and transfected with the corresponding expression plasmids. At 24 hours after transfection, the medium was changed to 1 mL of exosome-depleted medium and the cell cultures were incubated for another 24 hours. The exosome-enriched medium was harvested, centrifuged at 300 xg for 10 minutes to remove cells and then at 16 500 xg for 20 minutes to remove microvesicles and other unwanted entities. To further purify exosomes, the culture supernatants were filtered through 0.22-µm syringe filters. Then, the anti-CD63-conjugated beads were incubated overnight with the pre-cleared supernatants at ~200 000 beads/mL of supernatant with gentle shaking. The beads:exosomes conjugates were washed twice with PBS by centrifuging at 3000 xg for 20 minutes and finally resuspended in 50 to 100 µL of PBS. Staining with the FM 4-64 dye (Thermofisher) at 5 µg/mL was done in the final analysis volume by following the directions in the brand's protocol. The final samples were analyzed using a FACScan™ flow cytometer (B&D).

### Statistical analyses

All statistical analyses were done in GraphPad Prism 6. The type of analysis in each case is described in both the "Results and Discussion" section and in the figure legends of the corresponding experiment. The asterisks in the graphs represent the "p-values" for each comparison (e.g. "*" means "p<0.1", "**" means "p<0.01", and so on).

## ACKNOWLEDGMENTS

I am particularly grateful for the support of my supervisors, Dr. Vasco Barreto and Dr. Paulo Pereira, who believed and invested in me. To Dr. Manuel Vicente, whose irreplaceable intelligence, enthusiasm and dedication have been crucial for the makings of this study. To the resident Ph.D. students and "Post-Docs" -

Dr. Ana Soares and Dr. Nadiya Kubasova (Ph.D. students) – and - Dr. João Ferreira and Dr. João Proença (resident "Post-Docs"). To all the people who for the last decades or so have dedicated their lives to tapping into the knowledge explored in this study. Finally, infinite thanks to my "unknowingly psychologist", Mr. Maynard J. Keenan.

## REFERENCES

1. Khatua A, Taylor H, Hildreth J, Popik W. Exosomes packaging APOBEC3G confer Human Immunodeficiency Virus resistance to recipient cells. J Virol. 2009;83(2):512–21.

2. Khatua A, Taylor H, Hildreth J, Popik W. Inhibition of LINE-1 and Alu retrotransposition by exosomes encapsidating APOBEC3G and APOBEC3F. Virology. 2010;400(1):68–75.

3. Masur H, Michelis M, Greene J, Onorato I, Stouwe R, Holzman R, et al. An outbreak of community-acquired Pneumocystis carinii pneumonia. N Engl J Med. 1981;305(24):1431–8.

4. Centers for Disease Control and Prevention. Epidemiologic aspects of the current outbreak of Kaposi's sarcoma and oportunistic infections. N Engl J Med. 1982;306(4):248–52.

5. Siegal F, Lopez C, Hammer G, Brown A, Kornfeld S, Gold J, et al. Severe acquired immunodeficiency in male homosexuals, manifested by chronic perianal ulcerative herpes simplex lesions. N Engl J Med. 1981;305(24):1439–44.

6. Gottlieb M, Schroff R, Schanker H, Weisman J, Fan P, Wolf R, et al. Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men. N Engl J Med. 1981;305(24):1425–31.

7. Centers for Disease Control and Prevention. Possible transfusion-associated Acquired Immune Deficiency Syndrome (AIDS) - California. MMWR Wkly. 1982;31(48):652–4.

8. Centers for Disease Control and Prevention. Pneumocystis carinii pneumonia among persons with Hemophilia A. MMWR Wkly. 1982;31(27):365–7.

9. CDC C for DC and P. Unexplained immunodeficiency and oportunistic infections in infants - New York, New Jersey, California. MMWR Wkly. 1982;31(49):665–7.

10. Oleske J, Minnefor A, Cooper R, Thomas K, Cruz A, Ahdieh H, et al. Immune Deficiency Syndrome in Children. JAMA J Am Med Assoc. 1983;249(17):2345–9.

11. Rubinstein A, Sicklick M, Gupta A, Bernstein L, Klein N, Rubinstein E, et al. Acquired Immunodeficiency With Reversed T4/T8 Ratios in Infants Born to Promiscuous and Drug-Addicted Mothers. JAMA J Am Med Assoc. 1983;249(17):2350–6.

12. Lane H, Masur H, Edgar L, Whalen G, Rook A, Fauci A. Abnormalities of B-cell activation and immunoregulation in patients with the Acquired Immunodeficiency Syndrome. N Engl J Med. 1983;309(8):453–8.

13. Fahey J, Prince H, Weaver M, Groopman J, Schwartz K, Detels R. Quantitative changes in T helper or T supressor/cytotoxic lymphocyte subsets that distinguish Acquired Immune Deficiency Syndrome from other immune subset disorders. Am J Med. 1984;76(January):95–100.

14. Popovic M, Sarin P, Robert-Gurroff M, Kalyanaraman V, Mann D, Minowada J, et al. Isolation and Transmission of Human Retrovirus ( Human T-Cell Leukemia Virus ). Science (80- ). 1982;219(4586):856–9.

15. Blattner W, Kalyanaraman V, Robert-guroff M, Lister T, Greaves M, Galton D, et al. The human C-type retrovirus, HTLV, in black from the Caribbean region, and relationship to adult T-cell leukemia/lymphoma. Int J Cancer. 1982;30(3):257–64.

16. Gallo R, Kalyanaraman V, Sarngadharan M, Sliski A, Vonderheid E, Maeda M, et al. Association of the Human Type C Retrovirus with a subset of adult T-Cell cancers. Cancer Res. 1983;43(August):3892–9.

17. Wong-staal F, Hahn B, Manzari V, Colombini S, Franchini G, Gelmann E, et al. A survey of human leukaemias for sequences of a human retrovirus. Nature. 1983;302(5909):626–8.

18. Nagy K, Weiss R. Human T-cell Leukemia Virus type I: Induction of syncytia and inhibition by patients' sera. Int J Cancer. 1983;32(3):321–8.

19. Mitsuya H, Guo H, Megson M, Trainor C, Reitz M, Broder S. Transformation and cytopathogenic effect in an immune human T-Cell clone infected by HTLV-I. Science (80- ). 1983;223(4642):1293–6.

20. Hardy W, Hess P, Macewen E, Mcclelland A, Zuckerman E, Cotter S, et al. Biology of Feline Leukemia Virus in the natural environment. Cancer Res. 1976;36(February):582–8.

21. Essex M, Hardy W, Cotter S, Jakowski R, Sliski A, Hess P, et al. Naturally occurring persistent Feline Oncornavirus infections in the absence of disease. Infect Immun. 1975;11(3):470–5.

22. Anderson L, Jarrett W, Jarrett O, Laird H. Feline Leukemia-Virus infection of kittens : Mortality associated with atrophy of the thymus and lymphoid depletion. J Natl Cancer Inst. 1971;47(4):807–17.

23. Barré-Sinoussi F, Chermann J, Rey F, Nugeyre M, Chamaret S, Gruest J, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science (80- ). 1983;220(4599):868–71.

24. Klatzmann D, Barré-Sinoussi F, Nugeyre M, Dauguet C, Vilmer E, Griscelli C, et al. Selective tropism of lymphadenopathy associated virus (LAV) for helper-inducer T lymphocytes. Science (80- ). 1984;225(4657):59–63.

25. Gallo R, Sarin P, Gelmann E, Robert-Guroff M, Richardson E. Isolation of Human T-cell Leukemia Virus in Acquired Immune Deficiency Syndrome (AIDS). Science (80- ). 1983;220(4599):865–7.

26. Casareale D, Dewhurst S, Sonnabend J, Sinangil F, Purtilo DT, Volsky DJ. Prevalence of AIDS-associated retrovirus and antibodies among male homosexuals at risk for AIDS in Greenwich Village. AIDS Res. 1(6):407–21.

27. Kitchen L, Barin F, Sullivan J, McLane M, Brettler D, Levine P, et al. Aetiology of AIDS - antibodies to human T-cell leukaemia virus (type III) in haemophiliacs. Nature. 1984;312(5991):237–42.

28. Schupbach J, Popovic M, Gilden R, Gonda M, Sarngadharan M, Gallo R. Serological analysis of a subgroup of human T-lymphotropic retroviruses (HTLV-III) associated with AIDS. Science (80- ). 1984;224(4648):503–5.

29. Sarngadharan M, Popovic M, Bruch L, Schupbach J, Gallo R. Antibodies reactive with human T-lymphotropic retroviruses (HTLV-III) in the serum of patients with AIDS. Science (80- ). 1984;224(4648):506–8.

30. Vilmer E, Rouzioux C, Vezinet Brun F, Fischer A, Chermann J, Barre-Sinoussi F, et al. Isolation of new lymphotropic retrovirus from two siblings with Haemophilia B, one with AIDS.

Lancet. 1984;323(8380):753–7.

31. Levy J, Hoffman A, Kramer S, Landis J, Shimabukuro J, Oshiro L. Isolation of lymphocytopathic retrovirus from San Francisco patients with AIDS. Science (80- ). 1983;225(4664):840–2.

32. Popovic M, Sarngadharan M, Read E, Gallo R. Detection , isolation , and continuous production of cytopathic retroviruses ( HTLV-III) from patients with AIDS and Pre-AIDS. 1984;224(4648):497–500.

33. Ratner L, Haseltine W, Patarca R, Livak K, Starcich B, Josephs S, et al. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature. 1985;313(6000):277–84.

34. Sanchez-Pescador R, Power M, Barr P, Steimer K, Stempien M, Brown-Shimer S, et al. Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2). Science (80- ). 1985;227(4686):484–92.

35. Wain-hobson S, Sonigo P, Danos O, Cole S, Alizon M. Nucleotide sequence of the AIDS virus , LAV. Cell. 1985;40(1):9–17.

36. Muesing M, Smith D, Cabradilla C, Benton C, Lasky L, Capon D. Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. Nature. 1985;313(6002):450–8.

37. Gonda MA, Wong-Staal F, Gallo R, Clements J, Narayan O, Gilden R. Sequence homology and morphologic similarity of HTLV-III and visna virus, a pathogenic lentivirus. Science (80- ). 1985;227(4683):173–7.

38. Ratner L, Gallo R, Wong-Staal F. HTLV-IIII, LAV, ARV are variants of same AIDS virus. Nature. 1985;313(6004):636–7.

39. Fisher A, Collalti E, Ratner L, Gallo R, Wong-Staal F. A molecular clone of HTLV-III with biological activity. Nature. 1985;316(6025):262–5.

40. Lemoine P, Harrousseau H, Borteyru J, Claren S, Alvord E, Sumi S, et al. Characterization of envelope and core structural products of HTLV-III with sera from AIDS patients. Science (80- ). 1985;228(4699):593–5.

41. Veronese F, DeVico A, Copeland T, Oroszlan S, Gallo R, Sarngadharan M. Characterization of gp41 as the transmembrane protein coded by the HTLV-III/LAV envelope gene. Science (80- ). 1985;229(4720):1402–5.

42. Allan J, Coligan J, Barin F, McLane M, Sodroski J, Rosen C, et al. Major glycoprotein antigens that induce antibodies in AIDS patients are encoded by HTLV-III . Science (80- ). 1985;228(4703):1091–4.

43. Allan J, Coligan J, Lee T, McLane M, Kanki P, Groopman J, et al. A new HTLV-III/LAV encoded antigen detected by antibodies from AIDS patients. Science (80- ). 1985;230(4727):810–3.

44. Goh W, Sodroski J, Rosen C, Essex M, Haseltine W. Subcellular localization of the product of the long open reading frame of Human T-Cell Leukemia Virus Type I. Science (80- ). 1985;227(4691):1227–8.

45. Arya S, Guo C, Josephs S, Wong-Staal F. Trans-activator gene of Human T-Lymphotropic Virus type III (HTLV-III). Science (80- ). 1985;229(4708):69–73.

46. Sodroski J, Patarca R, Rosen C, Wong-Staal F, Haseltine W. Location of the Trans-activating region on the genome of Human T-Cell Lymphotropic Virus Type III. Science (80- ). 1985;229(4708):74–7.

47. Li G, Clercq E. HIV Genome-Wide Protein Associations : a Review of 30 Years of Research. Microbiol Mol Biol Rev. 2016;80(3):679–731.

48. Sodroski J, Goh W, Rosen C, Tartar A, Portetelle D, Burny A, et al. Replicative and cytopathic potential of HTLV-III/LAV with sor gene deletions. Science (80- ). 1986;231(4745):1549–53.

49. Kan N, Franchini G, Wong-Staal F, DuBois G, Robey W, Lautenberger J, et al. Identification of HTLV-III/LAV sor gene product and detection of antibodies in human sera. Science (80- ). 1986;231(4745):1553–5.

50. Strebel K, Daugherty D, Clouse K, Cohen D, Folks T, Martin M. The HIV “A” (sor) gene product is essential for virus infectivity. Nature. 1987;328(6132):728–30.

51. Luciw P, Cheng-mayert C, Levy J. Mutational analysis of the human immunodeficiency virus : The orf-B region down-regulates virus replication. Proc Natl Acad Sci USA.

1987;84(March):1434–8.

52. Fisher A, Ensoli B, Ivanoff L, Chamberlain M, Petiteway S, Ratner L, et al. The sor Gene of HIV- 1 is required for efficient virus transmission in vitro. Science (80- ). 1987;237(4817):888–93.

53. Sakai K, Dewhurst S, Ma X, Volsky DJ. Differences in cytopathogenicity and host cell range among fnfectious molecular clones of Human Immunodeficiency Virus type 1 simultaneously isolated from an individual. J Virol. 1988;62(11):4078–85.

54. Sakai K, Ma X, Volsky DJ. Low-cytopathic infectious clone of human immunodeficiency virus type I (HIV-I). FEBS Lett. 1988;238(2):257–61.

55. Sakai K, Ma X, Gordienko I, Volskyl DJ. Recombinational analysis of a natural noncytopathic Human Immunodeficiency Virus type 1 (HIV-1) isolate : role of the vif gene in HIV-1 infection kinetics and cytopathicity. J Virol. 1991;65(11):5765–73.

56. Gabuzda DH, Lawrence K, Langhoff E, Terwilliger, Dorfman T, Haseltine WA, et al. Role of vif in replication of Human Immunodeficiency Virus type 1 in CD4 + T lymphocytes. J Virol. 1992;66(11):6489–95.

57. Schwedler U, Song J, Aiken C, Trono D. vif is crucial for Human Immunodeficiency Virus type 1 proviral DNA synthesis in infected cells. J Virol. 1993;67(8):4945–55.

58. Sova P, Volsky DJ. Efficiency of viral DNA synthesis during infection of Permissive and Nonpermissive cells with vif-negative Human Immunodeficiency Virus type 1. J Virol. 1993;67(10):6322–6.

59. Simon JHM, Gaddis NC, Fouchier RAM, Malim MH. Evidence for a newly discovered cellular anti-HIV-1 phenotype. Nat Med. 1998;4(12):1397–400.

60. Madani N, Kabat D. An endogenous inhibitor of Human Immunodeficiency Virus in human lymphocytes is overcome by the viral Vif protein. J Virol. 1998;72(12):10251–5.

61. Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature. 2002;418(6898):646–50.

62. Teng B, Burant CF, Davidson N. Molecular Cloning of an Apolipoprotein B Messenger RNA

Editing Protein. Science (80- ). 1993;260(5115):1816–9.

63. Chen S, Habib G, Yang C, Gu Z, Lee BR, Weng S, et al. Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in-frame stop codon. Science (80- ). 1987;238(4825):363–6.

64. Powell LM, Wallis SC, Edwards YH, Knott TJ, Scott J. A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. Cell. 1987;50(6):831–40.

65. Bhattacharya S, Navaratnam N, Morrison JR, Scott J, Taylor WR. Cytosine nucleoside/nucleotide deaminases and apolipoprotein B mRNA editing. Trends Biochem Sci. 1994;19(3):105–6.

66. MacGinnitie AJ, Anant S, Davidson NO. Mutagenesis of apobec-1, the catalytic subunit of the mammalian Apolipoprotein B mRNA Editing enzyme, reveals distinct domains that mediate cytosine nucleoside deaminase, RNA binding and RNA editing activity. J Biol Chem. 1995;270(24):14768–75.

67. Smith AA, Carlow DC, Wolfenden R, Short SA. Mutations affecting transition-state stabilization by residues coordinating zinc at the active site of cytidine deaminase. Biochemistry. 1994;33(21):6468–74.

68. Jarmuz A, Chester A, Bayliss J, Gisbourne J, Dunham I, Scott J, et al. An anthropoid-specific locus of orphan C to U RNA-Editing enzymes on chromosome 22. Genomics. 2002;79(3):285–96.

69. Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, et al. Specific expression of Activation-induced Cytidine Deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. J Biol Chem. 1999;274(26):18470–6.

70. Martin A, Scharff MD. Somatic hypermutation of the AID transgene in B and non-B cells. Proc Natl Acad Sci USA. 2002;99(19):12304–8.

71. Petersen-mahrt SK, Harris RS, Neuberger MS. AID mutates E . coli suggesting a DNA deamination mechanism for antibody diversification. Nature. 2002;418(6893):99–104.

72. Barreto V, Reina-san-martin B, Ramiro AR, Mcbride KM, Nussenzweig MC. C-terminal

deletion of AID uncouples Class Switch Recombination from Somatic Hypermutation and Gene Conversion. Mol Cell. 2003;12(2):501–8.

73. Lecossier D, Bouchonnet F, Hance AJ. Hypermutation of HIV-1 DNA in the absence of the Vif protein. Science (80- ). 2003;300(5622):1112.

74. Stopak K, Noronha C, Yonemoto W, Greene WC, Francisco S. HIV-1 Vif Bbocks the antiviral activity of APOBEC3G by impairing both its translation and itracellular stability. Mol Cell. 2003;12:591–601.

75. Kao S, Khan MA, Miyagi E, Plishka R, Buckler-white A, Strebel K. The Human Immunodeficiency Virus type 1 Vif protein reduces intracellular expression and inhibits packaging of APOBEC3G (CEM15), a cellular inhibitor of virus infectivity. J Virol. 2003;77(21):11398–407.

76. Marin M, Rose KM, Kozak SL, Kabat D. HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. Nat Med. 2003;9(11):1398–403.

77. Sheehy AM, Gaddis NC, Malim MH. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. Nat Med. 2003;9(11):1404–7.

78. Yu X, Yu Y, Liu B, Luo K, Kong W, Mao P, et al. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. Science (80- ). 2013;302(5647):1056–60.

79. Liu B, Yu X, Luo K, Yu Y, Yu XF. Influence of primate lentiviral Vif and proteasome inhibitors on Human Immunodeficiency Virus type 1 virion packaging of APOBEC3G. J Virol. 2004;78(4):2072–81.

80. Miller LK, Kobayashi Y, Chen C, Russnak TA, Ron Y, Dougherty JP. Proteasome inhibitors act as bifunctional antagonists of human immunodeficiency virus type 1 latency and replication. Retrovirology [Internet]. 2013;10(120). Available from: Retrovirology

81. Mohammadzadeh N, Follack TB, Love RP, Stewart K. Polymorphisms of the cytidine deaminase APOBEC3F have different HIV-1 restriction efficiencies. Virology [Internet]. 2019;527(September):21–31. Available from: https://doi.org/10.1016/j.virol.2018.11.004

82. Zhen A, Wang T, Zhao K, Xiong Y, Yu X. A Single amino acid difference in human APOBEC3H

variants determines HIV-1 Vif sensitivity. J Virol. 2010;84(4):1902–11.

83. Ebrahimi D, Richards CM, Carpenter MA, Wang J, Ikeda T, Becker JT, et al. Genetic and mechanistic basis for APOBEC3H alternative splicing, retrovirus restriction, and counteraction by HIV-1 protease. Nat Commun [Internet]. 2018;9:4137. Available from: http://dx.doi.org/10.1038/s41467-018-06594-3

84. Rogozin IB, Iyer LM, Liang L, Glazko G V, Liston VG, Pavlov YI, et al. Evolution and diversification of lamprey antigen receptors : evidence for involvement of an AID-APOBEC family cytosine deaminase. Nat Immunol. 2007;8(6):647–56.

85. Liu M, Liao W, Buckley KM, Yang SY, Fugmann SD. AID/APOBEC-like cytidine deaminases are ancient innate immune mediators in invertebrates. Nat Commun [Internet]. 2018;9(1948):1–11. Available from: http://dx.doi.org/10.1038/s41467-018-04273-x

86. Salter JD, Bennett RP, Smith HC. The APOBEC protein family: united in structure, divergent in function. Trends Biochem Sci. 2017;41(7):578–94.

87. Rogozin IB, Basu MK, Jordan IK, Pavlov YI, Koonin E V. APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases predicted by computational analysis. Cell Cycle. 2005;4(9):1281–5.

88. Lada AG, Krick CF, Kozmin SG, Mayorov VI, Karpova TS, Rogozin IB, et al. Mutator effects and mutation signatures of editing deaminases produced in Bacteria and Yeast. Biochem. 2011;76(1):131–46.

89. Marino D, Perkovi M, Hain A, Vasudevan AAJ. APOBEC4 enhances the replication of HIV-1. PLOS(ONE). 2016;11(6):e0155422.

90. Yang B, Chen K, Zhang C, Huang S, Zhang H. Virion-associated Uracil DNA Glycosylase-2 and Apurinic / Apyrimidinic Endonuclease are involved in the degradation of APOBEC3G-edited nascent HIV-1 DNA. J Biol Chem. 2007;282(16):11667–75.

91. Armitage AE, Katzourakis A, Oliveira T, Welch JJ, Belshaw R, Bishop KN, et al. Conserved footprints of APOBEC3G on hypermutated Human Immunodeficiency Virus type 1 and Human Endogenous Retrovirus HERV-K (HML2) sequences. J Virol. 2008;82(17):8743–61.

92. Delviks-Frankenberry KA, Nikolaitchik OA, Burdick RC, Gorelick RJ, Keele BF, Hu WS, et al. Minimal contribution of APOBEC3-induced G-to-A hypermutation to HIV-1 recombination and genetic variation. PLOS Pathog. 2016;12(5):e1005646.

93. Lackey L, Demorest ZL, Land AM, Hultquist JF, Brown WL, Harris RS. APOBEC3B and AID have similar nuclear import mechanisms. J Mol Biol. 2012;419(5):301–14.

94. Patenaude A, Orthwein A, Hu Y, Campo VA, Kavli B, Buschiazzo A, et al. Active nuclear import and cytoplasmic retention of activation-induced deaminase. Nat Struct Mol Biol. 2009;16(5):517–27.

95. Mcbride KM, Barreto V, Ramiro AR, Stavropoulos P, Nussenzweig MC. Somatic Hypermutation is limited by CRM1-dependent nuclear export of Activation-induced Deaminase. J Exp Med. 2004;199(9).

96. Ito S, Nagaoka H, Shinkura R, Begum N, Muramatsu M, Nakata M, et al. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. Proc Natl Acad Sci. 2004;101(7):1975–80.

97. Harding C, Heuser J, Stahl P. Receptor-mediated endocytosis of transferrin and recycling of the Transferrin Receptor in rat reticulocytes. J Cell Biol. 1983;97(2):329–39.

98. Pan B, Johnstone R. Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro : Selective externalization of the receptor. Cell. 1983;33(3):967–78.

99. Johnstone RM, Adam M, Hammonds JR, Turbide C. Vesicle formation during reticulocyte maturation. J Biol Chem. 1987;262(1):9412–20.

100. Katzmann DJ, Stefan CJ, Babst M, Emr SD. Vps27 recruits ESCRT machinery to endosomes during MVB sorting. J Cell Biol. 2003;162(3):413–23.

101. Katzmann DJ, Babst M, Emr SD. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex , ESCRT-I. Cell. 2001;106(2):145–55.

102. Babst M, Katzmann DJ, Snyder WB, Wendland B, Emr SD. Endosome-associated complex, ESCRT-II, recruits transport machinery for protein sorting at the multivesicular body. Dev

Cell. 2002;3(8):283–9.

103. Babst M, Katzmann DJ, Estepa-sabal EJ, Meerloo T, Emr SD. ESCRT-III : an endosome-associated heterooligomeric protein complex required for MVB sorting. Dev Cell. 2002;3(8file:///C:/Users/User/Downloads/nihms328759.pdf):271–82.

104. Van Niel G, Charrin S, Simoes S, Romao M, Rochin L, Saftig P, et al. The tetraspanin CD63 regulates ESCRT-independent and dependent endosomal sorting during melanogenesis. Dev Cell. 2011;21(4):708–21.

105. Verweij FJ, Van MAJ, Hopmans ES, Vendrig T, Wurdinger T, Cahir-mcfarland E, et al. LMP1 association with CD63 in endosomes and secretion via exosomes limits constitutive NF- j B activation. EMBO J [Internet]. 2011;30(11):2115–29. Available from: http://dx.doi.org/10.1038/emboj.2011.123

106. Colombo M, Raposo G, Théry C. Biogenesis , secretion , and intercellular interactions of exosomes and other extracellular vesicles. Annu Rev Cell Dev Biol. 2014;30(1):255–92.

107. Raposo G, Nijman HW, Stoorvogel W, Leijendekker R, Harding C V, Melief CJM, et al. B lymphocytes secrete antigen-presenting vesicles. J Exp Med. 1996;183(3):1161–72.

108. Zitvogel L, Regnault A, Lozier A, Wolfers J, Flament C, Tenza D, et al. Eradication of established murine tumors using a novel cell-free vaccine: dendritic cell-derived exosomes. Nat Med. 1998;4(5):594–600.

109. Sahu R, Kaushik S, Clement C, Cannizzo E, Scharf B, Follenzi A, et al. Microautophagy of cytosolic proteins by late endosomes. Dev Cell. 2011;20(1):131–9.

110. Dice J. Peptide sequences that target cytosolic proteins for lysosomal proteolysis. Trends Biochem Sci. 1990;15(8):305–9.

111. Kirchner P, Bourdenx M, Madrigal-matute J, Tiano S, Diaz A, Bartholdy BA, et al. Proteome-wide analysis of chaperone- mediated autophagy targeting motifs. PLOS Biol. 2019;17(5):e3000301.

112. Salvador N, Aguado C, Horst M, Knecht E. Import of a cytosolic protein into lysosomes by Chaperone-Mediated Autophagy depends on its folding state. J Biol Chem.

2000;275(35):27447–59.

113. Cuervo A, Knecht E, Terlecky S, Dice J. Activation proteolysis of a selective pathway of lysosomal in rat liver by prolonged starvation. Am J Physiol Physiol. 1995;269(5):C1200–8.

114. Cuervo A, Dice J, Knecht E. A population of rat liver lysosomes responsible for the selective uptake and degradation of cytosolic proteins. J Biol Chem. 1997;272(9):5606–15.

115. Cuervo A, Dice J. A receptor for the selective uptake and degradation of proteins by lysosomes. Science (80- ). 1996;273(5274):501–3.

116. Ferreira J, Fôfo H, Bejarano E, Bento C, Ramalho J, Girão H, et al. STUB1 / CHIP is required for HIF1A degradation by chaperone-mediated autophagy. Autophagy. 2013;9(9):1349–66.

117. Xu J, Camfield R, Gorski SM. The interplay between exosomes and autophagy – partners in crime. J Cell Sci. 2018;131(jcs215210).

118. Harding C, Stahl P. Transferrin receptor in reticulocytes: pH and iron are important determinants of ligand binding and processing. Biochem Biophys Res Commun. 1983;113(2):650–8.

119. Majeski A, Dice J. Mechanisms of chaperone-mediated autophagy. Int J Biochem Cell Biol. 2004;36(12):2435–44.

120. Cuervo AM, Terlecky SR, Dice JF, Knecht E. Selective binding and uptake of Ribonuclease A and Glyceraldehyde-3-phosphate Dehydrogenase by isolated rat liver lysosomes. J Biol Chem. 1994;269(42):26374–80.

121. Drozdetskiy A, Cole C, Procter J, Barton GJ. JPred4 : a protein secondary structure prediction server. Nucleic Acids Res. 2015;43(April):389–94.

122. Kiffin R, Christian C, Knecht E, Cuervo A. Activation of chaperone-mediated autophagy during oxidative stress. Mol Biol Cell. 2004;15(November):4829–40.

123. Jordan M, Christiane K, Wurm FM. Calcium-phosphate mediated DNA transfer into HEK-293 cells in suspension : control of physicochemical parameters allows transfection in stirred media Transfection and protein expression in mammalian cells. Cytotechnology. 1998;26(1):39–47.

124. Lackey L, Law EK, Brown WL, Harris RS. Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. Cell Cycle. 2013;12(5):762–72.

125. Zhang H, Freitas D, Kim HS, Fabijanic K, Li Z, Chen H, et al. Identification of distinct nanoparticles and subsets of extracellular vesicles by asymmetric flow field-flow fractionation. Nat Cell Biol [Internet]. 2018;20(March):332–43. Available from: http://dx.doi.org/10.1038/s41556-018-0040-4

126. Soares AR, Martins-marques T, Ribeiro-rodrigues T, Vasco J, Catarino S, Pinho MJ, et al. Gap junctional protein Cx43 is involved in the communication between extracellular vesicles and mammalian cells. Sci Rep [Internet]. 2015;5(13243). Available from: http://dx.doi.org/10.1038/srep13243

127. Dahabieh MS, Ooms M, Simon V, Sadowski I. A doubly fluorescent HIV-1 reporter shows that the majority of integrated HIV-1 is latent shortly after infection. Am Soc Microbiol. 2013;87(8):4716–27.

128. UNAIDS. 90-90-90: treatment for all [Internet]. [cited 2019 Jun 30]. Available from: https://www.unaids.org/en/resources/909090

129. CDC C for DC and P. HIV in the United States and Dependent Areas. [Internet]. 2017. Available from: https://www.cdc.gov/hiv/statistics/overview/ataglance.html

130. HIV/AIDS WHO Regional Office for Africa [Internet]. Available from: https://www.afro.who.int/health-topics/hivaids

131. Martins HC, Aldir I. Infeção VIH e SIDA : a situação em Portugal a 31 de dezembro de 2017. Inst Nac Saúde Dr Ricardo Jorge. 2018;Documento.

132. United States Government. Ending the HIV Epidemic (overview). [Internet]. 2019 [cited 2019 Jun 30]. Available from: https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/overview

133. Centers for Disease Control and Prevention. Preexposure prophylaxis for the prevention of HIV infection in the United States - 2017 update: a clinical practice guideline. 2017.

134. Tenofovir disoproxil - DrugBank [Internet]. [cited 2019 Jun 30]. Available from: https://www.drugbank.ca/drugs/DB00300

135. Emtricitabine - DrugBank [Internet]. [cited 2019 Jun 30]. Available from: https://www.drugbank.ca/drugs/DB00879

136. Grant R, Lama J, Anderson P, McMahan V, Liu A, Vargas L, et al. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med. 2010;363(27):2587–99.

137. Choopanya K, Martin M, Suntharasamai P, Sangkum U, Mock P, Leethochawalit M, et al. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok , Thailand ( the Bangkok Tenofovir Study ): a randomised , double-blind , placebo-controlled phase 3 trial. Lancet [Internet]. 2012;381(9883):2083–90. Available from: http://dx.doi.org/10.1016/S0140-6736(13)61127-7

138. Thigpen M, Kebaabetswe P, Paxton L, Smith D, Rose C, Segolodi T, et al. Antiretroviral preexposure prophylaxis for heterosexual HIV transmission in Botswana. N Engl J Med. 2012;367(5):423–34.

139. Baeten J, Donnell D, Ndase P, Mugo N, Campbell J, Wangisi J, et al. Antiretroviral prophylaxis for HIV-1 prevention among heterosexual men and women. 2012;367(5):399–410.

140. Smith D, Handel M, Grey J. Estimates of adults with indications for HIV preexposure prophylaxis by jurisdiction, transmission risk group, and race/ethnicity, United States, 2015. Ann Epidemiol [Internet]. 2018; Available from: https://doi.org/10.1016/j.annepidem.2018.05.003

141. Huang Y, Zhu W, Smith D, Harris N, Hoover K. HIV preexposure prophylaxis , by Race and Ethnicity — United States, 2014-2016. Morb Mortal Wkly Rep. 2018;67(41):2014–6.

142. Marcus J, Hurley L, Hare C, Nguyen D, Phengrasamy T, Silverberg M, et al. Preexposure prophylaxis for HIV prevention in a large integrated Health Care system: Adherence, renal safety, and discontinuation. J Acquir Immunodefic Syndr. 2016;73(5):540–6.

143. Montgomery M, Oldenburg C, Nunn A, Mena L, Anderson P, Liegler T, et al. Adherence to

pre-exposure prophylaxis for HIV prevention in a clinical setting. 2016;1–10.

144. Mayer K. Decreased HIV incidence among PrEP users compared to non-users in a Boston Community Health Center, 2012-2017. In: HIV research for prevention conference. 2018.

145. Grulich A, Guy R, Amin J, Jin F, Selvey C, Holden J, et al. Population-level effectiveness of rapid , targeted , high-coverage roll-out of HIV pre-exposure prophylaxis in men who have sex with men : the EPIC-NSW prospective cohort study. Lancet HIV. 2018;5(11):629–37.

146. Cottrell M, Prince H, Schauer A, Sykes C, Dellon S, Adams J, et al. Decreased tenofovir diphosphate concentrations in a transgender female cohort: Implications for HIV pre-exposure prophylaxis (PrEP). Clin Infect Dis. 2019;ciz290:1–16.

147. Powell V, Gibas K, Dubow J, Krakower D. Update on HIV preexposure prophylaxis : Effectiveness , Drug Resistance , and Risk Compensation. Curr Infect Dis Reports. 2019;21(8):21–8.

148. Kemnic TR, Gulick PG. HIV Antiretroviral Therapy [Internet]. 2019. Available from: https://www.ncbi.nlm.nih.gov/books/NBK513308/

149. Wong JK, Hezareh M, Günthard HF, Havlir D V, Ignacio CC, Spina CA, et al. Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. Scinece. 1997;278(5341):1291–5.

150. Chun TW, Stuyver L, Mizell SB, Ehler LA, Mican JAM, Baseler M, et al. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. Proc Natl Acad Sci. 1997;94(November):13193–7.

151. Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, et al. Identification of a reservoir for HIV-1 in patients on Highly Active Antiretroviral Therapy. Science (80- ). 1997;278(5341):1295–300.

152. Chun TW, Davey RT, Engel D, Lane HC, Fauci AS. Re-emergence of HIV after stopping therapy Microscopic chaos from. Nature. 1999;401(6756):874–5.

153. Davey RT, Bhat N, Yoder C, Chun TW, Metcalf JA, Dewar R, et al. HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of

sustained viral suppression. Proc Natl Acad Sci. 1999;96(26):15109–14.

154. Chun TW, Nickle DC, Justement JS, Large D, Semerjian A, Curlin ME, et al. HIV-infected individuals receiving effective antiviral therapy for extended periods of time continually replenish their viral reservoir. J Clin Invest. 2005;115(11):3250–5.

155. Chun TW, Justement JS, Murray D, Hallahan CW, Maenza J, Collier AC, et al. Rebound of plasma viremia following cessation of antiretroviral therapy despite profoundly low levels of HIV reservoir: implications for eradication. AIDS. 2010;24(18):2803–8.

156. Hamlyn E, Ewings FM, Porter K, Cooper DA, Tambussi G, Schechter M, et al. Plasma HIV viral rebound following protocol-indicated cessation of ART commenced in primary and chronic HIV infection. PLoS One. 2012;7(8):e43754.

157. Skiest DJ, Su Z, Havlir D V, Robertson KR, Coombs RW, Cain P, et al. Interruption of antiretroviral treatment in HIV- infected patients with preserved immune function is associated with a low rate of clinical progression : a prospective study by AIDS Clinical Trials Group 5170. J Infect Dis. 2007;195(10):1426–36.

158. Steingrover R, Pogány K, Garcia EF, Jurriaans S, Brinkman K, Schuitemaker H, et al. HIV-1 viral rebound dynamics after a single treatment interruption depends on time of initiation of highly active antiretroviral therapy. AIDS. 2008;22(13):1583–8.

159. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, Mangen-Wabwire F, et al. Viral load and heterosexual transmission of Human Immunodeficiency Virus type 1. N Engl J Med. 2000;342(13):921–9.

160. Gupta RK, Gregson J, Parkin N, Haile-selassie H, Tanuri A, Forero LA, et al. HIV-1 drug resistance before initiation or re-initiation of first-line antiretroviral therapy in low-income and middle-income countries : a systematic review and meta-regression analysis. Lancet Infect Dis [Internet]. 2018;18(3):346–55. Available from: http://dx.doi.org/10.1016/S1473-3099(17)30702-8

161. Group TTS. Global epidemiology of drug resistance after failure of WHO recommended first-line regimens for adult HIV-1 infection : a multicentre retrospective cohort study. Lancet Infect Dis. 2016;16(5):565–75.

162. Barth RE, Loeff MFS, Schuurman R, Hoepelman AIM, Wensing AMJ. Virological follow-up of adult patients in antiretroviral treatment programmes in sub-Saharan Africa : a systematic review. Lancet Infect Dis [Internet]. 2010;10(3):155–66. Available from: http://dx.doi.org/10.1016/S1473-3099(09)70328-7

163. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DIS, et al. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell [Internet]. 2013;155(3):540–51. Available from: http://dx.doi.org/10.1016/j.cell.2013.09.020

164. Kearney MF, Wiegand A, Shao W, Coffin JM, Mellors JW, Lederman M. Origin of rebound plasma HIV includes cells with identical proviruses that are transcriptionally active before stopping Antiretroviral Therapy. J Virol. 2015;90(3):1369–76.

165. Allen SJ, Crown SE, Handel TM. Chemokine : receptor structure, interactions, and antagonism. Annu Rev Immunol. 2007;25(1):787–820.

166. Schlyer S, Horuk R. I want a new drug : G-protein-coupled receptors in drug development. Drug Discov Today. 2006;11(11–12):481–93.

167. Yeagle PL, Albert AD. G-protein coupled receptor structure. Biochim Biophys Acta. 2007;1768(4):808–24.

168. Marchese A, Paing MM, Temple BRS, Trejo J. G protein-coupled receptor sorting to endosomes and lysosomes. Annu Rev Pharmacol Toxicol. 2008;48(1):601–29.

169. Samson M, Labbe O, Mollereau C, Vassart G, Parmentier M. Molecular cloning and functional expression of a new human CC-chemokine receptor gene. Biochemistry. 1996;35(11):3362–7.

170. Combadiere C, Ahuja SK, Tiffany L, Murphy PM. Cloning and functional expression of CC CKR5, a human monocyte CC chemokine receptor selective for MIP-1alpha, MIP-1beta, and RANTES. J Leukoc Biol. 1996;60(1):147–52.

171. Cocchi F, Devico AL, Garzino-demo A, Arya SK, Gallo RC, Lussot P. Identification of RANTES, MIP-1alpha, and MIP-1beta as the major HIV-suppressive factors produced by CD8+ T cells. Science (80- ). 1995;270(5243):1811–5.

172. Zhou Y, Kurihara T, Ryseck R, Yang Y, Ryan C, Loy J, et al. Impaired macrophage function and enhanced T cell-dependent immune response in mice lacking CCR5, the mouse homologue of the major HIV-1 coreceptor. J Immunol. 1998;160(8):4018–25.

173. Brien SJ, Moore JP. The effect of genetic variation in chemokines and their receptors on HIV transmission and progression to AIDS. Immunol Rev. 2000;177(1):99–111.

174. Brien TR, Winkler C, Dean M, Nelson JAE, Carrington M, Michael NL, et al. HIV-1 infection in a man homozygous for CCR5 delta32. Lancet. 1997;349(9060):1219.

175. Balotta C, Bagnarelli P, Violin M, Ridolfo AL, Zhou D, Berlusconi A, et al. Homozygous Δ32 deletion of the CCR-5 chemokine receptor gene in an HIV-1-infected patient. AIDS. 1997;11(10):F67–71.

176. Biti R, French R, Young J, Bennets B, Stewart G, Liang T. HIV-1 infection in an individual homozygous for the CCR5 deletion allele. Nat Med. 1997;3(3):252–3.

177. Gorry PR, Zhang C, Wu S, Kunstman K, Trachtenberg E, Phair J, et al. Persistence of dual-tropic HIV-1 in an individual homozygous for the CCR5delta32 allele. Lancet. 2002;359(9320):1832–4.

178. Michael NL, Nelson JAE, Kewalramani VN, Chang G, Brien SJO, Mascola JR, et al. Exclusive and persistent use of the entry coreceptor CXCR4 by Human Immunodeficiency Virus Type 1 from a subject homozygous for CCR5 delta32. J Virol. 1998;72(7):6040–7.

179. Sheppard HW, Celum C, Michael NL, O'Brien S, Dean M, Carrington M, et al. HIV-1 infection in individuals with the CCR5-delta32/delta32 genotype: acquisition of syncytium-inducing virus at seroconversion. J Acquir Immune Defic Syndr. 2002;29(3):307–13.

180. Naif HM, Cunningham AL, Alali M, Li S, Nasr N, Buhler MM, et al. A Human Immunodeficiency Virus Type 1 isolate from an infected person homozygous for CCR5 delta32 exhibits dual tropism by infecting macrophages and MT2 cells via CXCR4. J Virol. 2002;76(7):3114–24.

181. Federsppiel B, Melhado IG, Duncan AM V, Delaney A, Schappert K, Clark-Lewis I, et al. Molecular cloning of the cDNA and chromosomal localization of the gene for a putative seven-transmembrane segment (7-TMS) receptor isolated from human spleen. Genomics.

1993;16(3):707–12.

182. Herzog H, Hort YJ, Shine J, Selbie LA. Molecular cloning , characterization , and localization of the human homolog to the reported bovine NPY Y3 receptor: lack of NPY binding and activation. DNA Cell Biol. 1993;12(6):465–71.

183. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor : functional cDNA cloning of a seven-transmembrane , G Protein-Coupled Receptor. Science (80- ). 1996;272(5263):872–7.

184. Bleul CC, Farzan M, Choe H, Parolin C, Clark-Lewis I, Sodroski J, et al. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. Nature. 1996;382(6594):829–33.

185. Oberlin E, Amara A, Bachelerie F, Bessia C, Virelizier JL, Arenzana-Seisdedos F, et al. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. Nature. 1996;382(6594):833–5.

186. Altenburg JD, Broxmeyer HE, Jin Q, Cooper S, Basu S, Alkhatib G. A naturally occurring splice variant of CXCL12/stromal cell-derived factor 1 is a potent Human Immunodeficiency Virus type 1 inhibitor with weak chemotaxis and cell survival activities. J Virol. 2007;81(15):8140–8.

187. Tachibana K, Hirota S, Iizasa H, Yoshida H, Kawabata K, Kataoka Y, et al. The chemokine receptor CXCR4 is essential for vacularization of the gastrointestinal tract. Nature. 1998;393(6685):591–4.

188. Bandres JC, Wang QF, O'Leary J, Baleuaux F, Amara A, Hoxie JA, et al. Human Immunodeficiency Virus (HIV) Envelope binds to CXCR4 independently of CD4, and binding can be enhanced by interaction with soluble CD4 or by HIV Envelope deglycosylation. J Virol. 1998;72(3):2500–4.

189. Lapham CK, Ouyang J, Chandrasekhar B, Nguyen NY, Dimitrov DS, Golding H. Evidence for cell-surface association between Fusin and the CD4-gp120 complex in human cell lines. Science (80- ). 1996;274(5287):602–5.

190. Trkola A, Dragic T, Arthos J, Binley JM, Olson WC, Allaway GP, et al. CD4-dependent,

antibody-sensistive interactions between HIV-1 and its co-receptor CCR-5. Nature. 1996;384(6605):184–7.

191. Salzwedel K, Smith ED, Dey B, Berger EA. Sequential CD4-coreceptor interactions in Human Immunodeficiency Virus type 1 Env function: soluble CD4 activates Env for coreceptor-dependent fusion and reveals blocking activities of antibodies against cryptic conserved epitopes on gp120. J Virol. 2000;74(1):326–33.

192. Hütter G, Nowak D, Mossner M, Ganepola S, Mübig A, Allers K, et al. Long-term control of HIV by CCR5 delta32/delta32 stem-cell transplantation. N Engl J Med. 2009;360(7):692–8.

193. Brown TR. I am the Berlin Patient : A personal reflection. AIDS Res Hum Retroviruses. 2015;31(1):2–3.

194. Lederman MM, Pike E. Ten years HIV free : an interview with " The Berlin Patient." Pathog Immun. 2017;2(3):422–30.

195. Gupta RK, Abdul-jawad S, Mccoy LE, Mok HP, Peppa D, Salgado M, et al. HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation. Nature [Internet]. Available from: http://dx.doi.org/10.1038/s41586-019-1027-4

196. Henrich TJ, Hanhauser E, Marty FM, Sirignano MN, Keating S, Lee T, et al. Antiretroviral-free HIV-1 remission and viral rebound after allogeneic stem cell transplantation. Ann Intern Med. 2014;161(5):319–28.

197. Cummins NW, Rizza S, Litzow MR, Hua S, Lee Q, Einkauf K, et al. Extensive virologic and immunologic characterization in an HIV-infected individual following allogeneic stem cell transplant and analytic cessation of antiretroviral therapy : A case study. PLOS Med. 2017;14(11):e1002461.

198. Froger A, Hall JE. Transformation of plasmid DNA into E . coli using the Heat Shock method. J Vis Exp. 2007;6:253.

199. Azevedo F, Pereira H, Johanssonm B. Colony PCR. Methods Mol Biol. 2017;1620:129–39.

200. Lässer C, Eldh M, Lötvall J. Isolation and characterization of RNA-containing exosomes. J Vis Exp. 2012;59(e3037).

201. Yeung Y, Stanley ER. A solution for stripping antibodies from PVDF immunoblots for multiple reprobing. An Biochem. 2009;389(1):89–91.

202. Théry C, Amigorena S, Raposo G, Clayton A. Isolation and characterization of exosomes from cell culture supernatants. Curr Protoc Cell Biol. 2006;30(1):3.22.1-3.22.29.