Identification and Characterization of Immunologically Relevant CD8+ T cell Epitopes of HIV-1 Nef Protein

By

Elnaz Shadabi

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the degree of Doctor of Philosophy

Department of Medical Microbiology and Infectious Diseases
University of Manitoba
Winnipeg

ABSTRACT

Background:

Human Immunodeficiency Virus-1 (HIV-1) relies on the host cells for its replication and survival. Escape mutations in the viral genome arise in response to the selective pressure by Human Leukocyte Antigen (HLA) class I-restricted CD8+ T cell responses. Escape mutations that are accumulated at the host population level as a result of immune selection pressure may indirectly result in different consequences for the host due to fitness cost to the virus. Some are associated with rapid disease progression, while others that diminish viral fitness are associated with slower CD4 decline. Nef is one of the accessory proteins of HIV-1 and plays an important role in viral pathogenesis and replication. Its major functions include down-regulation of CD4 and HLA Class I expression on the host cell surface to avoid recognition by immune cells, and modification of host cell signal transduction pathways to ensure a persistent state of infection. Between 50-90% of CD8+ T cell responses during the acute phase of HIV infection are targeted against the Nef protein, and the highest Interferon-γ (IFN-γ) production of CD8+ T cells is directed towards Nef during the chronic phase. In both human and macaque models of infection some Nef-specific CD8+ T cell responses that drive viral escape mutations have been associated with better immune control.

Hypothesis:

I hypothesize that immunologically relevant CD8+ T cell epitopes of HIV-1 Nef protein contain positively selected mutations (PSMs) that are associated with different clinical outcomes in the infected host such as disease progression profiles.

To test this hypothesis, the following specific objectives were carried out: 1) identification of PSMs in HIV-1 Nef sequences using Bioinformatics tools (Quasi analysis) and determining their association with disease progression; 2) validation of epitopes that contain PSMs, associated with differential disease outcome, by ELISPOT; and 3) characterization of CD8+ T cell populations restricted by the HLA class I epitopes containing these PSMs, using flow cytometry.

Major Findings:

1) PSMs were identified in both the variable and conserved regions of HIV-1 Nef. Several of the PSMs were associated with different disease outcomes. Three PSMs were associated with faster CD4 decline (E63D p = 0.028, log rank= 4.799; I101V p=0.003, log rank=8.667 and I168M p=0.042, log rank=4.150), while two PSMs were associated with slower CD4 decline (H116N p= 0.00011, log rank=14.891 and K182M p=0.03, log rank=4.753)

2) These PSMs are within the CD8+ T cell epitopes recognized by several HLA class I alleles common in the Kenyan population.

3) The CD8+ T cell responses restricted by the same HLA class I allele can drive PSMs associated with different clinical consequences. This is demonstrated by the example of A*02:01. A*02:01 was associated with I101V and H116N (p=0.028 and 0.021, respectively). I101V was associated with faster CD4 decline (p=0.003, log rank= 8.667) and H116N was associated with slower CD4 decline (p= 0.00011, log rank= 14.891).

4) There was no difference in the frequency of HLA class I restricted CD8+ T cells that recognized epitope variants containing PSMs or consensus amino acids.

5) There was no significant difference between the frequency of IFN-γ producing CD8+ T cells that target epitopes with PSMs associated with slow or fast CD4 decline. There was no significant difference in the frequency of antigen-specific CD8+ T cells with antiviral intracellular cytokines, their proliferation and exhaustion characteristics when peptides with different PSMs were tested by flow cytometry. However, the frequency of CD8+ T cells restricted by A*02:01- ILDLWVY**N**T (IT9-N) epitope, containing a PSM associated with slower CD4 decline, was higher ($p \leq 0.0001$) than CD8+ T cells restricted by A*02:01-ILDLWVY**H**T (IT9-H) epitope, containing consensus amino acid associated with faster CD4 decline.

Conclusions:

Identification of PSMs associated with different clinical outcomes and characterization of CD8+ T cell populations specific to epitopes associated with differential disease progression can help determine potentially effective immungens of a HIV-1 vaccine candidate.

DEDICATIONS:

I dedicate this work to my parents Roya Kholas Ardestani and Heshmatollah Shadabi, my brother Moein Shadabi and sister Saghar Shadabi, without whose love and support I would not have accomplished much.

ACKNOWLEDGEMENTS

This work has been made possible through the support of many individuals and I would like to use this opportunity to thank everyone involved.

First, I would like to thank my advisors Dr. Ma Luo and Dr. Frank Plummer for having provided me with the opportunity to participate in this project. Thank you to Dr. Ma Luo for her guidance and for motivating me to learn and grow at every opportunity. Thanks to Dr. Plummer for his vision and extraordinary leadership.

Thank you to the members of my advisory committee for their feedback and insights. To Dr. Blake Ball, thank you for always helping me build on strength instead of focusing on weakness and for your expert input all along. Thank you Dr. Brian Fristensky and Dr. Gary Van Domselaar for your keen interest in my project and for your guidance and the challenging questions that have pushed me to learn.

I would also like to thank all the staff at the Department of Medical Microbiology for their support and assistant. Thank you, Sharon and Eva. Thank you, Jude, for bringing so much kindness to your work and thank you to Angie for knowing the answers to every question a grad student could ask. You truly know them all!

This work was made possible through the support of others at the Department of Medical Microbiology and JC Wilt Infectious Disease Research Centre. I would like to thank Sue Ramdahin, Dr. Adrienne Meyers, Dr. Julie Lajoie, and Dr. Ruey Su and Bernie Abrenica for providing support whenever I have needed it. Thank you to Christine Mesa, Dr. Sandra Kiazyk, Dr. Jillian Waruk and especially Dr. Jennifer Juno for your assistance during my experience with flow cytometry.

Thank you to Dr. Paul Sandstrom for being such an excellent director of JC Wilt and for all your efforts to create the best work environment possible. I would also like to thank Dr. Matthew Gilmour, Scientific Director General of the National Microbiology Laboratory, for taking our lab under his wings in the recent months and providing the necessary support.

I owe a lot of what I have learned to the support and mentorship of past and present members of Luo/Plummer lab. Thank you, John Ho, for being my first teacher in the lab and for your patience throughout my learning. Thank you to Rupert Capina and Jeff Tuff for accompanying me during the Nairobi experience. I'm also grateful to Chris Czarnecki, David Tang, David La, Philip Lacap, Christina Daniuk and Thomas Bielawney for their support and friendship throughout the years. Thank you Dr. Binhua Liang and Dr. Raghavan Sampathkumar for your help with the HIV sequencing project. Most recently, thank you to Dr. Hongzhao Li, Dr. Lin Li, Dr. Were Omange and Lukas for your positivity and encouragement.

I am also particularly thankful to Dr. Meika Richmond for her genuine eagerness to teach, to help and to be a friend in all circumstances. Your support has truly meant a lot to me during the completion of the last part of my research project.

To my fellow graduate students Lewis and Kashem, thank you for your friendship, for chats and for returning emails at late hours of the night when I needed your help in last minute preparations for talks, conferences and assignments. Thank you and welcome to the newest addition to our lab, Nikki.

To all those at the University of Nairobi, my deepest appreciation for making my experience of visiting and working with you so pleasant and for continuing to provide support in accessing study samples for my project. I am particularly thankful to Dr. Joshua Kimani, Dr. Julius Oyugi, Tony, Irungu and Anne.

I am also grateful to all the trainees and mentors of the International Infectious Disease and Global Health Training Program for the unique experience of learning together in Colombia, India, Kenya and Canada, which truly enriched my PhD experience. Thank you to Dr. Keith Fowke for your leadership and Natasha for coordination of the program. I would also like to thank my friend and travel companion, Laura Thompson, for conversations, laughter and opportunities to reflect on career and life goals.

I am forever indebted to my mentor and friend, “Amou Redwan”, also known as Dr. Moqbel, who encouraged me to pursue my PhD despite my endless doubt in my abilities to do so. More importantly I am grateful to him for lovingly and joyfully displaying in every one of his interactions what it means to channel one’s work into service to humanity. Though you left us so soon and I miss you tremendously and your unique ability to empower and encourage everyone you touched, I knew you were helping me all along while on your journey in the next world.

I will always be grateful to the love and support of my friends. Romie, Loie, Jill P, Siavash, Jill D, Anis, Natasha, Christie, Meleyna, Jordan, Mel, Ryan and Zele, one way or another you have each supported and sustained my heart in this journey. To my dear friend Shar, thank you for not only being the best listener, but also for being an example of calm determination and reliance on God’s Will. To my childhood best friends, Sogol, Jelveh, Azadeh and Mahshid, whom I miss every day, thank you for always keeping in touch despite the distance and the many years passed. Thank you Karmil, for always making me laugh and for ignoring time zone differences when calling me. You were always the best distraction! To my dearest Sonia, I will never be able to thank you enough for the light that you are in my life.

Finally, to my family, my pillar through many challenges and the joy of my heart at every moment, thank you from the bottom of my heart. My brother Moein and my sister Saghar, you will never know how much I love you each. Thank you for still being my friend despite my science related comments at every conversation. To my parents, your sacrifices before and after our relocation to Canada have made so much possible for me and our home has always been my refuge. I love you!

TABLE OF CONTENTS:

List of Figures

List of Tables

1. INTRODUCTION

1.1 The HIV-1 Pandemic

1.1.1 Global Pandemic and Burden

After nearly three decades of research Human Immunodeficiency Virus-1 (HIV-1) continues to be a major public health threat globally(1). According to the most recent data from the World Health Organization (WHO), in the year 2016, 36.7 million people around the globe were living with HIV, with 1.8 million (1.6 million-2.1 million) new infections and 1.0 million (830,000-1.2 million) deaths(2,3). From the beginning of the epidemic in 1981(4) to the end of 2015, it is estimated that nearly 78 million people have become infected by the virus and 35 million have died as a result of AIDS (Acquired immunodeficiency syndrome)-related illnesses(5). Access to antiretroviral therapy (ART) has improved over the years, with nearly 18.2 million people receiving treatment in June 2016 compared to 15.8 million in 2015(5). This has led to an increase in life expectancy and therefore, an increase in the prevalence of HIV-1 infection globally, from 31 million in 2002 to 36.7 million in 2015(6). Despite the progress made due to ART accessibility the number of new HIV infections among adults has not declined since 2010 with 1.9 million new infections(7).

1.1.2 HIV-1 Epidemic in Sub-Saharan Africa, Kenya

Kenya is among one of the most severely affected countries in Africa since the first HIV/AIDS diagnosis was made in 1984(8). Within a decade after this initial discovery, between 1995-96, the prevalence of HIV-1 infection was approximately 10.5%. Gradually by the year 2003, a 40% decline led to 6.7% prevalence countrywide and finally by the year

2012 the prevalence had reached 5.6% (9). Factors that have contributed to this decline are AIDS related mortality and scale up of ART. Despite this decline women and girls in Kenya remain more vulnerable than men and boys for risk of HIV-1 infection. The results of several surveys in Kenya have shown that the prevalence among women is nearly twice that in men. Among people living with HIV (PLHIV) 57% are women and 43% are men. Among young people between the ages of 15-24, the prevalence of HIV-1 infection for boys was lower (1.1-1.5%) compared to girls (3.0-5.9%) between 2003-2012(9).

Aside from the generalized epidemic described above, Kenya also has a concentrated epidemic, where certain population demographics continue to remain disproportionally affected; these include sex workers and their clients, injection drug users (IDU), and men who have sex with men (MSM). HIV prevalence is highest among female sex workers (FSW) at approximately 29.3%(9). Factors such as poverty, socioeconomic status, stigma and discrimination have all played a role in leaving certain segments of the population more vulnerable than others. However, HIV educational programs, counseling, testing and treatment are in place to reduce HIV-1 incidence in Kenya.

1.1.3 Origins of HIV

The first cases of AIDS were reported in the United States in 1981 and 1982 (10,11) with large numbers of homosexual men suffering from opportunistic infections and rare malignancies such as Kaposi's sarcoma(12). Finally, in 1984, the causative agent behind AIDS, HIV, was isolated and discovered as a retrovirus(8). The HIV-1 epidemic initially began in heterosexual populations in Zaire and Rawanda in 1984(13,14). The epidemic is a

result of multiple zoonotic transmissions, where simian immunodeficiency virus (SIV) crossed from non-human primates to humans, through exposure to blood with possibly the bushmeat hunters in Central and West Africa as the first group to be infected(1). There are two strains of the virus; HIV-1, the more pathogenic form, is traced back to chimpanzees and gorillas while HIV-2, originates from sooty mangabeys(1). Unlike HIV-1, HIV-2 is mostly limited to West Africa and while it also causes a similar immunodeficiency, it has a lower transmission ability and results in slower disease progression(1).

There are four major groups of HIV-1 that represent independent transmission events: groups M (major), N (non-M or non-O) and O (outlier) were first transmitted to humans from chimpanzees while, group P was transmitted from gorillas(15-17). Today's global HIV-1/AIDS pandemic is the result of group M virus that has diversified into nine distinct subtypes based on genetic variability which include subtype A, B, C, D, F, G, H, J and K(18) and sub-subtypes A1-A4 and F1 and F2 (19,20). There are also up to 54 circulating recombinant forms (CRF). The criteria for a CRF is to characterize identical mosaic structure from full genome sequences of at least three epidemiologically unlinked individuals. If this criterion is not met the virus is classified a unique recombinant form (URF)(21) The predominant subtype responsible for HIV-1 global pandemic is subtype C with approximate prevalence of 50%, followed by subtype A, 12%, and subtype B, 11%(5).

1.1.4 Impact of HIV-1 Diversity and Subtype Variation on Disease Progression and Viral Control

The error prone reverse transcriptase (RT) enzyme of HIV-1 lacks proofreading ability and is responsible for high rates of mutations and recombination in the virus that has resulted in HIV's genetic diversity(12). Genetic variation in virus within the same subtypes is between 8-17% and in virus from different subtypes approximately 17-35%(22). Naturally HIV-1 diversity leads to challenges with vaccine development, response to treatment, and has an impact on transmission, host immunity and pathogenesis (23). An example of the difference of transmission rates among subtypes was demonstrated in an Ugandan study which revealed that heterosexual transmission rates of subtype A is higher than subtype D(24). Higher rates of subtype C heterosexual transmission in South Africa and India maybe due to the stronger localization of virus of this subtype in the female genital tract compared to other subtypes(25-27). In terms of pathogenesis several studies have shown that within the same population patients infected with subtype D virus progressed to AIDS more rapidly than those infected with subtype A(28,29). Subtype D infected patients also had lower CD4 T cell counts compared to subtype A patients(29,30). However, this issue remains controversial, since it has also been shown that disease progression rates were not different among individuals infected with subtype A or D (31). A study that compared sequences from subtype B and non-subtype B infected patients on ART showed that drug resistant mutational patterns were differ between the two groups(32). Although HIV-1 diversity presents these challenges, host genetic factors such as Human Leukocyte Antigen (HLA) class I alleles drive the evolution of various forms of the virus and thus provide researchers the opportunity to predict evolutionary pathways. This will in turn better inform therapeutic strategies(33). The relationship between HLA alleles and viral evolution is outlined in more detail in section 1.5.2.

1.1.5 HIV-1 Transmission

HIV transmission can occur by three major exposure routes: sexual intercourse (vaginal or anal), vertical transmission (mother to child through breast milk or birth) and parenteral transmission (injection drug use needle sharing, percutaneous needle injury or blood transfusions). Sexual intercourse and parenteral transmission are considered horizontal routes of transfer(5,34) . Worldwide the most common transmission route is through unprotected sex, followed by IDU needle sharing and mother to child transmission (MTCT) (5,35). Anal sex carries a greater risk of transmission compared to vaginal sex, accounting for the epidemic of HIV-1 infection in the MSM population globally(36). The risk of HIV-1 transmission is 18 times higher by receptive anal sex compared to vaginal sex (37,38). This can be attributed to three major physiological and immunological factors(5). First, rectal epithelial layer is more vulnerable to abrasions than the vaginal mucosal layer. Second, the microfold cells (M) that line the lymphoid follicles in the rectal mucosa are highly specialized to bind any foreign antigens like HIV. Finally the rectum, which is an extension of the intestinal epithelium tends to be populated by HIV-1 target cells such as CD4+ lymphocytes, dendritic cells (DC), monocytes and macrophages(38).

Several biological and behavioral factors contribute to the increased risk in the transmission of HIV-1. Increased plasma viral load (VL) is associated with increased risk of transmission; with every 1 log10 increase in VL the risk of transmission is increased by approximately 2.4 times(39). Since the plasma VL is highest during the acute phase of the infection (first 12 weeks), transmission of the virus is most effective during this time(40).

Increased endocervical and seminal VL have also been linked to higher transmission rates(41). Other biological factors including, pregnancy and sexually transmitted infections (STI) can contribute to increased transmission risks. Behavioral factors that contribute to increased risk of transmission include, number of sexual partners(42), concurrent partners(42), sexual abuse/violence, injection drug use and needle sharing while male circumcision and condom use can reduce the risk (34). Social factors such as stigma and discrimination represent obstacles to access to treatment and care and thus contribute to increased risk of transmission(43).

1.2 Treatments and Prevention

1.2.1 Treatment of HIV-1

Although currently there is no effective vaccine to prevent HIV-1 infection, one of the major achievements in the field of HIV research has been the discovery of drugs that treat the infection, making it a more manageable condition compared to a previously fatal one. Initially in early 1980s, acyclic nucleoside analogue antiviral agents used against infections caused by herpes simplex virus (HSV) were found to also serve as potent inhibitors of HIV-1(44,45). Today there are nearly 30 ART drugs available and approved to treat HIV-1 infection(46). These drugs disrupt the life cycle of the virus by targeting its replication cycle at various steps. They include nucleoside and nucleotide reverse transcriptase inhibitors (NRTIs), Non-nucleotide reverse transcriptase inhibitors (NNRTIs), integrase inhibitors, protease inhibitors, CCR5 (chemokine co-receptor 5) inhibitors and fusion inhibitors (46,47). The combination of two or more ARV drugs is a stronger regiment that is more effective against the virus(46). This form of therapy is referred to as highly active

antiretroviral therapy (HAART). ARV therapy reduces the risk of viral transmission to an uninfected individual (e.g. in serodiscordant couples), as it lowers peak viraemia during the acute phase of the infection, and reduces latent viral reservoirs and infection of other cells (46).

Previously the WHO guidelines recommend initiation of ARV treatment when the CD4 T cells of an infected patient reach 350 cells per µl of blood. In a healthy uninfected adult, the CD4 T cell count is approximately 500 cells per µl and when an individual infected with HIV develops AIDS the CD4 T cell count drops below 200 cells per µl of blood. Thus, a threshold of 350 CD4 cells per µl was determined as the optimal point to benefit from treatment well in advance of suffering from symptoms associated with AIDS(47). As of 2015 however, new WHO guidelines are in place that recommend initiation of ART for all infected individuals immediately after a diagnosis has been made irrespective of patient CD4 T cell count(6). It is expected that CD4 T cell counts increase rapidly for up to three months after ART initiation and ideally remain steady as the overall count reaches the normal 500, although this may not be achieved in all patients, especially those presenting with low CD4 counts at start of treatment (47).

Presidents Emergency Preparedness for AIDS Relief (PEPFAR) is an initiative established in 2003 by American President George W. Bush in an effort to respond to the HIV/AIDS global epidemic by providing ART to those living in resource limited settings(48). Nearly 7.7 million PLHIV worldwide have received lifesaving treatment through PEPFAR. Although, initially between 2003 and 2007 PEPFAR I adopted an emergency response approach,

PEPFAR II, between 2008 and 2012, utilized a more targeted approach with greater emphasis on country ownership. The success of PEPFAR will continue to rely on strengthening existing health systems, engaging the private sector and support from various donors(48).

1.2.2 HIV-1 Prevention Strategies

WHO recommends *Pre-exposure prophylaxis* (PrEP), to those at increased risk to maximize prevention of HIV-1 infection(6). PrEP is a prevention strategy that uses combination of ART drugs(49,50). Uninfected persons who are at high risk of exposure, such as commercial sex-workers (CSW) or IDUs (50), can administer PrEP orally or topically (in the form of a gel as a microbicide), prior to exposure to HIV to reduce chance of infection. The first successful PrEP study, called CAPRISA 004, was conducted in 2010 in South Africa and it enrolled 889 high-risk women (51). The drug tested was 1% tenofovir gel that was administered vaginally 12 hours before and up to 12 hours after sexual intercourse. The result of this study was encouraging; it was reported that HIV acquisition was reduced by 39% and acquisition was inversely correlated with traces of tenofovir in vaginal secretion. Several other randomized control studies investigating the effectiveness of PrEP have since taken place and included various population demographics such as MSMs in the United States, IDUs in Thailand and serodiscordant couples in Kenya and Uganda(52-54). The effectiveness of PrEP has ranged from 39-75% overall reduction in HIV acquisition. However, the effectiveness of this powerful prevention tool relies on patient adherence. Several studies have shown that low adherence was correlated with no protection against HIV while daily adherence reduced risk of HIV infection by 90%(49).

Post exposure prophylaxis (PEP) uses the same concept as PrEP but is administered after exposure to HIV, for example, in incidences of needle injury for health care professionals, victims of sexual assault or other unprotected sexual encounters(55). PEP must be administered within 72 hours after exposure and daily for up to 28 days. Scale up of PEP is likely not practical or sustainable and therefore, other prevention strategies must be in place to effectively reduce risk of infection.

Prevention of mother to child transmission (PMTCT) is an indispensable strategy to drastically reduce the risk of transmission to a newborn infant(43,56). WHO has recommended four different approaches that are effective in achieving this goal. These include: 1) HIV prevention in women during childbearing age 2) prevention of unwanted pregnancy in infected women 3) prevention of transmission from mother to infant using ART and 4) continued care and support of mothers with HIV infection. Of these strategies, one of the most successful has been the scale up of ART treatment in low and middle income countries which reached 53% of HIV-positive women in 2009, compared to 45% in 2008 and only 15% in 2005(57).

Other prevention approaches include behavioral strategies such as condom use, counseling and treatment of STIs. *Voluntary Medical Male Circumcision* (VMMC) has also been shown to reduce the risk of transmission and therefore prevent infection(58). As stated above there are several approaches for the prevention of HIV-1 infection and a combination of these approaches will be the most effective way in reducing the risk of transmission and

overall incidence. However, the single most effective prevention strategy is through immunization that features a vaccine platform with effective immunogen.

1.2.3 HIV-1 Vaccine Development

Although the development of ARV treatment has been successful in combating the deleterious impact of HIV-1/AIDS globally, challenges associated with cost, drug accessibility and long term side effects of treatment continue to push the research community towards an innovative preventative strategy. Nearly 30 different vaccine candidates have demonstrated some ability to induce immune responses, though many are not protective(59). The major challenge for a vaccine candidate is to elicit effective T cell responses while also generating high-titres of broadly neutralizing antibodies. Early vaccine strategies were focused on inducing B cells that through antibody-mediated neutralization may prevent HIV acquisition(60). Two phase III gp120 vaccine trials failed to demonstrate protection: Vax004 was tested for efficacy in MSM populations in North America and Europe (61) and Vax003 was tested in an IDU population in Thailand(62). With discouraging outcomes of the B-cell targeted AIDSVAX® trials T cell based vaccine strategies that aimed to reduce viral set point (e.g. plasma VL) and therefore, delay disease progression, quickly became popular(60). Two phase IIb T cell based vaccine trials in heterosexual and MSM populations in the Americas and Australia (STEP) and heterosexual men and women in South Africa (Phambilit) were halted due to poor efficacy(63,64). In fact, the STEP trial showed evidence for increased risk of acquisition in the vaccinated vs. placebo group (4.6% vs. 3.1% respectively, *p*-value= 0.07). Both vaccine candidates were DNA based Ad5 vector expressing *gag, pol*, *nef* as the main viral components(59).

To date the most effective vaccine candidate is the RV144 Thai trial, which combined both antibody and T cell strategy. The vaccine consisted of ALVAC-HIV prime using a canarypox vector, which was followed by AIDSVAX boost (gp120 subunit). In a randomized double blind trial among 16,402 participants, efficacy first reached 60% but over 42 months it dropped to 31.2% (95% confidence interval, 1.1 to 52.1; p= 0.04)(65). Two main correlates of protection among vaccinated individuals were identified. First, there was an inverse correlation between plasma concentration of IgG antibody specific for the V1V2 loop region of gp120 and risk of HIV acquisition. Second, IgA antibody to HIV-1 Env was shown to be correlated with acquisition of infection(66). Additionally, low levels of Env-specific CD4+ T cells, Env-specific IgA antibodies, IgG avidity, non-neutralizing antibodies and antibody-dependent cell-mediated cytotoxicity were also inversely correlated with risk of infection (66-68). Although the results of RV144 trial were not durable, the combined T cell mediated and antibody neutralization approach provides hope and insight for further innovative strategies.

1.3 HIV-1 VIROLOGY

1.3.1 Genetic Structure

HIV-1 is approximately 9 kilobases (kb) long and belongs to the *Retroviridae* family, *Lentivirinae* subfamily and *Lentivirus* genus. The viral gene products consist of three major proteins, Gag, Pol and Env, two regulatory proteins, Tat and Rev and four accessory proteins, Vif, Vpr, Vpu, and Nef (Figure 1.1). Gag and Env play a structural role for the virus. The Gag polyprotein is cleaved by the protease to give rise to matrix (p17), capsid (p24),

NucleoCapsid (p7), and p6 protein, all of which contribute to inner and outer structure of the virus. The Env protein is cleaved into gp120 and gp41. gp120 binds to CD4 and other receptors on host cells, and gp41, allows for viral fusion into the host cell upon interaction with CD4 receptors. Pol is a catalytic protein of HIV that is cleaved into three viral enzyme proteins, *protease*, *integrase*, and *reverse trascriptase*. Tat and Rev contribute to viral replication early on during infection and the accessory proteins of HIV (Vif, Vpr, Vpu, and Nef) each play an important role in the pathogenesis and virulence of the virus mainly through immune evasion strategies(69). (*Note: structure and function of Nef, the main HIV-1 protein studied in this work, is discussed in more detail in section 1.6.1).*

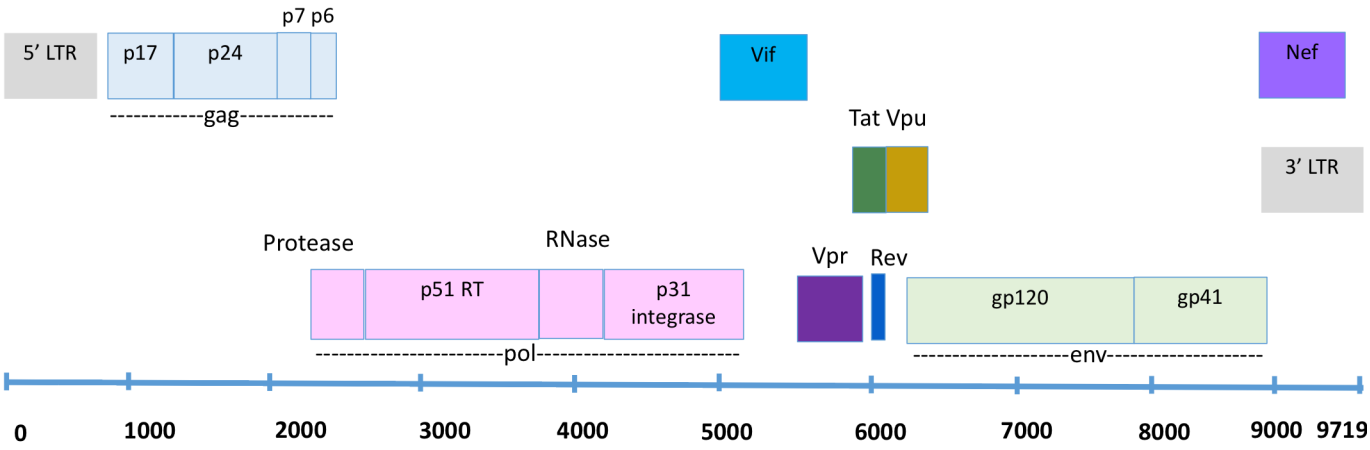


Figure 1.1. Structure of HIV-1 genome. Individual genes are drawn based on the nucleotide base pairs and arranged vertically by translational frame.

1.3.2 HIV-1 Life Cycle

The first step in the viral life cycle of HIV-1 is binding or attachment between gp41 and gp120 of the viral Env protein and host cell receptors, CD4 and chemokine receptors, CCR5 or CXCR4(70). The primary targets of the virus are cells that express the CD4 receptor such as CD4+ lymphocytes, monocytes, macrophages and DCs. Although resting CD4+ T cells are also infected, activated CD4+ T cells are most susceptible due to the high surface expression levels of viral receptor and the ability to support viral replication while proliferating(71). HIV-1 is also able to infect non-CD4 expressing cells such as renal epithelial cells(72) and brain astrocytes(73), which can result in nephropathy and neurocognitive disorder complications respectively(43). Once binding is established the virus will enter the host cell through the fusion of its Env protein and the host cell membrane. Inside the cell, HIV releases the RT enzyme that converts viral RNA into double stranded DNA. In the host nucleus, viral enzyme *integrase* will facilitate the integration of viral and host cellular genome. Following integration, the HIV provirus can either undergo active transcription resulting in high population of viral progeny or enter proviral latency as a result of silencing mechanisms that repress transcription(74,75). While ARV treatment can reduce viral load (by halting replication processes that lead to viral progeny), viral latency represents a great obstacle in treatment and eradication of HIV-1. For transcriptionally capable virus, following the integration step the virus uses the host machinery for the transcription of viral mRNA. Within host cell cytoplasm viral mRNA is translated to polyproteins that are assembled at the cell surface and released via the cell membrane. These immature virions undergo a maturation process during which the

polyproteins are cleaved by the HIV *protease*. Only mature virions are capable of infecting CD4+ T cells to continue the replication cycle.

1.3.3 HIV-1 Pathogenesis and Immune Dysfunction

Once HIV establishes itself within a host cell it is transmitted to other neighboring cells across the mucosal membrane that also express CCR5 or CXCR4(76,77) . DCs and Langerhan cells become infected in the submucosal layer(43). The acute phase of the infection, which often lasts about 12 weeks, is characterized by rapid viral replication. During this period cellular immunity signals the production of inflammatory and antiviral cytokines. A peak in HIV-specific CD8+ T cells can be detected 1-2 weeks after increase in viremia. Eventually VL reaches a setpoint as a result of potent killing of infected cells by CD8+ T cells, an important component of adaptive immunity. Approximately, three months after the infection neutralizing antibodies, released by B cells, begin to appear. In nearly 20% of patients broadly neutralizing antibodies (bNab) have been shown to directly kill the virus with the ability to neutralize HIV from multiple subtypes (78).

A distinctive feature of HIV-1 infection is the depletion of CD4+ T cells as a result of their destruction, which ultimately leads to immune dysfunction. Several mechanisms contribute to the elimination of CD4+ T cells. These include: direct killing of CD4 T cells, bystander effect of syncytia formation, senescence and apoptosis(43). During the acute phase of the infection, CD4 T cells circulating in the blood are rapidly lost from over 1000 to approximately 500 cells per µl blood. This is followed by recovery of CD4 T cells to normal concentrations, which occurs concurrent with the reduction of pVL due to killing by CD8+ T

cells. However, in untreated infections CD4 count continues to decrease slowly throughout the chronic phase and once below 200 cells per μl blood, the AIDS stage of disease begins, followed by death due to AIDS related complications. A major cause of CD4+ T cell depletion is the infection of activated memory T cells of the gut-associated lymphoid tissue (GALT)(79). GALT cells within the gastrointestinal (GI) tract contain high levels of CCR5 expressing activated memory T cells (80,81), therefore, servings as targets for HIV infection. Gradually the depletion of these lymphoid T cells increases permeability of the gut tissue and subsequently results in high concentrations of microbial products that further activate both innate and adaptive immune systems, causing activation of HIV target cells. Preferential loss of T-helper-17 cells (Th17) and mucosal-associated invariant T cells, which play a role in controlling bacterial infections result in weakened immunity (82).

Increased immune activation is also another feature of HIV-1 pathogenesis(43). One major factor of immune activation is the stimulation of host Toll-like receptors (TLR)- 7 and -8 by the virus (83) and subsequent release of pro-inflammatory cytokines IFN-α, Interleukin (IL) -6 and -8, macrophage inflammatory protein (MIP)-1α, and adhesion molecules (84). Other TLRs are also activated by bacterial products such as lipopolysaccharides (LPS) and bacterial DNA that accumulate after damage to mucosal intestinal barriers (85). Co-infection with other viruses such as cytomegalovirus (CMV) can also induce expansion of activated T cells (86). Overall, these factors contribute to enhanced immune proliferative state that further mediates infection of HIV target cells and disease progression. ART drugs are successful in diminishing pVL levels and thus reducing immune activation, however, baseline levels that characterize an uninfected individual's immune profile are never

achieved. Currently, some treatment plans seek to reduce T cell proliferation and activation in order to eliminate HIV target of replication(87).

1.4 HIV-1 and Host Immune Responses

The various components of the host immune response against HIV-1 have been studied extensively in the global effort to develop an effective vaccine. The host immune response consists of innate and adaptive immunity. The components of the innate immunity allow for rapid and immediate response against the infecting pathogens, while the adaptive immunity provides long-term protection through antigen recognition and antigenic memory. In the following sections the major players of the immune response against HIV-1 will be discussed.

1.4.1 Innate Immune Responses

The components of the innate immune response form the first line of defense against HIV-1 infection(88). The main purpose of innate immunity against HIV-1 is to limit the infection by controlling viral spread and to activate and modulate the antiviral effects of the adaptive immunity. Once mucosal resident CD4+ T cells, DCs and macrophages are infected, viral recognition by innate immunity is initiated. Pathogen-recognition receptors (PRRs) on host cells have the ability to detect viral particles called pathogen-associated molecular patterns (PAMPs). Among PRRs, host cell TLRs can recognize viral proteins. TLR-7 and TLR-8 commonly recognize HIV-1 genomic RNA and signal the activation of plasmacytoid dendritic cells (pDCs), which produce high levels of type I interferons (IFN) that have antiviral properties (89-91).

Activation of the innate immune cells results in the release of cytokines and chemokines, such as IFN-γ and IL-2, which eventually recruit other immune cells to the site of the infection(91). Natural Killers (NK) cells for examples are recruited by IL-2, IL-15 and IL-21 and have potent antiviral activity(92). Killer immunoglobulin-like receptors (KIRs) on the surface of NK cells are highly polymorphic and play an important role in the activation and inhibition of these cells. KIRs also interact with HLA class I molecules, thereby initiating NK cell antiviral effector functions. These include antibody-dependent cellular cytotoxity (ADCC)- mediated and direct killing of HIV infected cells. Numerous studies have demonstrated that certain KIR receptors like KIR2DL1, KIR2DL2 and KIR2DL3 are expanded in early infection which begin antiviral activity against HIV-1 infection (93).

The innate immune response has been of interest for researchers in the recent years since it has been shown to impact the adaptive immune response that follows the initial line of defense towards viral infection. T cell mediated immunity as well as antibody production by B cells are linked to initial activation of innate immune pathways. Therefore, the component of innate immune response, activated during the early phase of the infection, play a crucial role in determining the subsequent control of disease course, such as CD4 depletion and VL (93).

1.4.2 Adaptive Immune Responses

The adaptive arm of the immune response consists of *humoral* response that tend to neutralize or kill the virus thereby preventing infection of new cells, and *cellular* responses that kill infected cells(94).

1.4.2.1 Humoral Responses

The antiviral activity of humoral immune response against HIV-1 is initiated by B cells, which form plasma cells after antigen stimulation and produce potent antibodies (Ab) capable of virus neutralization(95). bNAbs are specialized in recognition of gp120 and gp41 epitopes of the Env protein. Their major function is to prevent infection of uninfected cells through direct virus neutralization. Immunoglobulins (Ig) can mediate cytotoxicity and viral inhibition. IgG proteins for example bind to viral epitopes and facilitate killing of infected cells by recruiting other effectors such NK cells, monocytes and neutraphills. In HIV uninfected individuals IgA and IgM have demonstrated catalytic activity against the virus (96).

1.4.2.2 CD4+ T cells

The evidence of the role that CD4+ T cells play in control against HIV has been documented in numerous studies (97). In long-term non-progressors (LTNP) and elite controllers (EC), a subset of HIV-1 infected individuals who maintain stable CD4 count for at least 7 years without treatment, HIV-1 specific CD4+ T cells were enriched in the mucosal tissue. These cells have also demonstrated polyfunctionality (i.e. secretion of multiple cytokines) and proliferation capacity(98,99). Activation and frequency of CD4+ T cells have also been linked with lower VL (100,101). In LTNPs, CD4+ T cells demonstrated low expression

levels of exhaustion markers and co-inhibitory molecules, T-cell immunoglobulin and mucin-domain containing-3 (TIM-3), program cell-death protein 1(PD-1) and Lymphocyte-activation gene 3 (LAG3) (102,103). Proliferating CD4+ T cells maintain long-term activation of CD8+ T cells, a hallmark of effective immune response in LTNPs and ECs(103). Furthermore, studies have shown that loss of CD4+ T cells in the chronic phase of the infection can diminish the effector function of CD8+ T cells (104,105). CD4+ T cells are able to potentiate the function of other immune cells(106); CD4+ T cells secret IL-2 to activate CD8+ T cells, and IL-4 to upregulate B cells.

1.4.2.3 CD8+ T cells

Cytotoxic T Lymphocytes (CTL), also known as Killer or CD8+ T cells play a major role in the elimination of infected or malignant cancer cells in the host (106-108). In HIV infection, numerous studies have outlined the crucial role of CTLs in controlling viremia in both acute and chronic phases of infection. An inverse correlation between the number of CD8+ T cells and viremia in infected hosts has been shown (109). Depleting CD8+ T cells of SIV infected rhesus macaques has been shown to accelerate disease progression with lowered CD4+ T cells and increased VL (110). Beside a *quantitative* measure of the important role of CD8+ T cells, several studies have also demonstrated their *qualitative* impact (111). For example, CD8+ T cells of elite viral controllers (ECs) and LTNPs were found to be polyfunctional, defined by their ability to secret multiple cytokines such as IFN-γ, TNF-α, MIP-1β, IL-2 and CD107a (193, 216). Such responses can effectively inhibit viral replication and promote the killing of HIV-1 infected cells (112,113). Additionally, in highly exposed seronegatives (HESN), a subset of individuals who remain HIV-negative for several years despite frequent

exposure to infectious virus (114,115), polyfunctional CD8+ T cells may play a role in protection against HIV-1 infection(116).

Killing by CTLs is also accomplished through the production and release of two protein families, perforins and granzymes(117) (118). Perforins create pores in the membrane of HIV-1 infected cells mediating the transport of granzymes into the cell. Granzymes, which are serine proteases, interact with cellular proteins to promote apopotosis (self-killing) of the infected cells, thereby diminishing the spread of infection to other cells. Studies have shown that the granzyme B-mediated killing of HIV-1 infected cells are more frequent in ECs than in disease progressors(119,120).

In an HIV-1 infected cell's cytoplasm the enzyme protease breaks down the viral protein into smaller peptides of approximately 9 amino acids long (121). Peptides are then passed through transport-associated antigen processing (TAP) protein into the cell's endoplasmic reticulum (ER) where they are coupled with the host's HLA molecule. This peptide-HLA complex is further processed through the Golgi apparatus and is finally expressed on the surface where the viral peptide is presented to the TCR on CD8+ T cells (122). It is through antigen recognition by TCR that CD8+ T cells can initiate their cytotoxic (release of granzymes and perforins) and cytolytic (release of cytokines like IL-2 etc) activities (Figure 1.2).

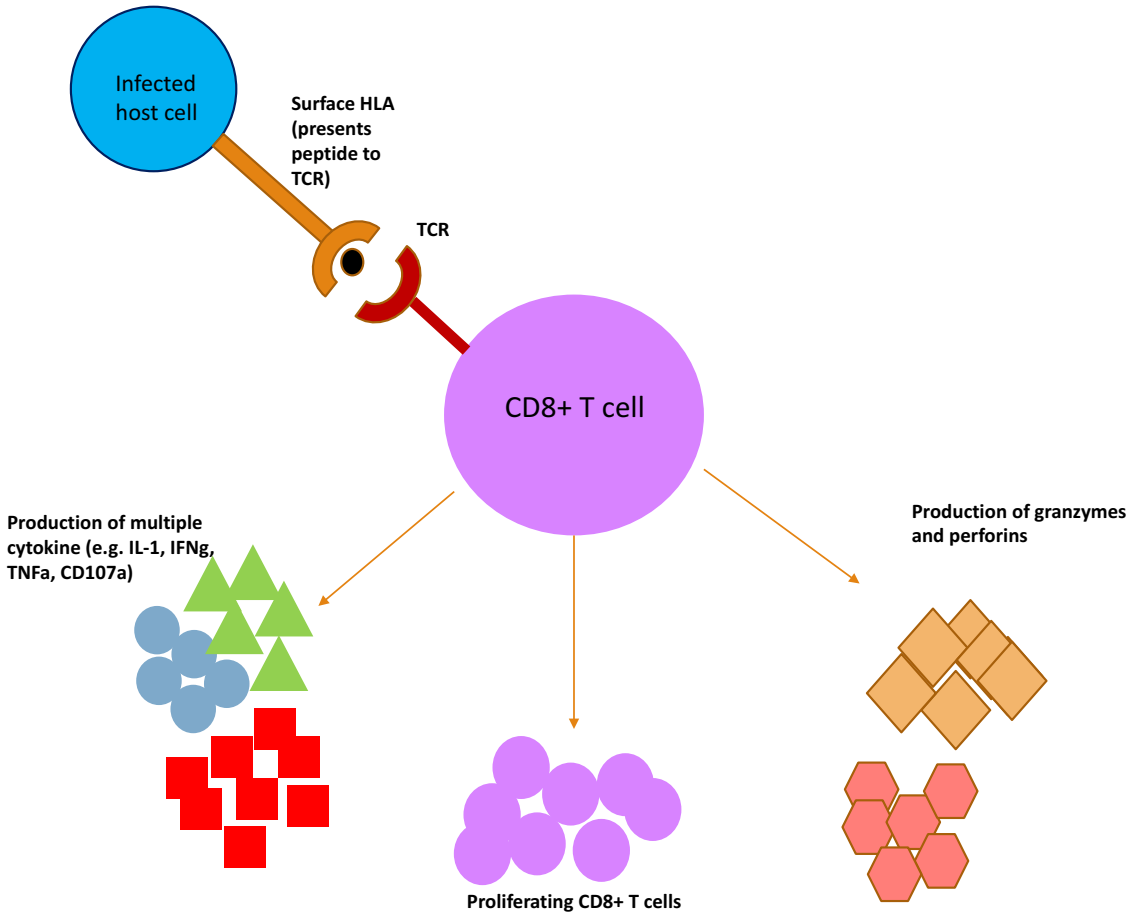


Figure 1.2. HLA peptide presentation to CD8+ T cells via the TCR and subsequent activation of CD8+ T cells.

1.5 HIV-1 Host Immune Evasion

Overtime, HIV-1 has developed strategies to escape the host immune response to survive and enhance its replication capacity(123), and to overcome host immune barriers (124). The host immune response represents several lines of defense through innate and adaptive immunity as well as host restriction factors to effectively reduce viral replication and halt infection(103). The evolving interactions between the virus and the host immune response take place at different stages during the course of untreated infection and the host immune response ultimately drives viral evolution (125). Below some mechanisms are described.

1.5.1 CTL Escape Mutations

HLA-restricted CTL responses represent major selective pressures that drive viral evolution in an infected host. At transmission, viral genetic bottleneck results in infection of the new host by a founder virus(34,38,126). Upon entering the host nucleus the virus quickly replicates using an error prone RT enzyme. It is estimated that $3x10^{-5}$ mutations per base pair per replication cycle are introduced as a result of RT's lack of DNA proofreading ability(127,128). This leads to the generation of closely related but diverse population of viral sequences referred to as *quasi species*(129,130). Immediately the cellular immune response initiates the killing of HIV-infected cells through the recognition of viral epitopes presented by HLA class I molecules to T cell receptors (TCR). Then with virus-specific CTL responses during the acute phase of the infection a dramatic decline in viremia occurs simultaneously (131-133). These early CD8+ T cell responses select for mutations within viral

epitopes that make immune escape possible. It has been shown that escape mutations result in the generation of new viral sequences in the first 7-21 days of infection(134,135) and continue to take place even during the chronic phase(136,137). Escape mutation was first described in 1991 among a cohort of HIV-seropositive haemophiliac patients where a fluctuation in the specificity of CTL responses was matched with proviral *gag* epitope DNA sequence variability. Two variants of p17 *gag* epitopes containing arginine to lysin (R->K) amino acid substitution, present in the blood of some donors were not recognized by B-lymphoblastoid T cell lines or autologous activated T cells. This study showed that amino acid substitutions in Gag epitopes diminish CTL recognition and therefore, clearly demonstrated HIV-1 escape(138). Subsequently, escape was described in 1997 in a study that examined viral-specific CD8+ T cell responses against an immunodominant epitope of gp160 protein in Env(139). Sequence analysis using RT-PCR revealed amino acid replacement mutations in the Env protein that weakened the peptide-HLA binding, therefore, demonstrating CD8+ T cell escape mutation. Numerous studies since then have described the role of HLA-restricted C8+ T cells in driving the evolution of the virus by selecting escape mutations in viral epitopes of multiple genes such as gag, pol, env, nef and other accessory and regulatory proteins of HIV (140-144).

Epitope processing is a complex mechanism and CTL escape mutations can occur at three different stages (Reviewed in Carlson et al. 2015, Trends in Microbiology)(145). The first is when viral protein mutations interfere with

intracellular processing of peptides by the proteasome in the infected cell(146). This is called *antigen-processing escape*" and can take place within or nearby an epitope(147,148) . An example of this type of mutation is A146P in Gag of B*57:03 individuals that results in improper trimming of the epitope (149). The second type of mutation, called "*HLA-binding escape*" affects the epitope-HLA binding step(145,150). In this scenario epitopes are processed by the proteasome, however, a replacement mutation, often in one of the peptide's anchor positions, 2, 8 or 9, diminishes binding ability between that epitope and the HLA molecule(150,151). R264K mutation in B*27 restricted Gag KK10 epitope is an example of this(137). The third possible stage for escape mutation can take place at the step of antigen presentation to TCR. While some mutations do not impact antigen processing or HLA-binding steps, "*TCR escape mutations*" reduce epitope recognition by the TCR clonotypes of the selecting CTLs(138,152). These mutations are typically selected at the middle position of epitopes, often position 6 (152). CTL driven escape mutations have been shown in both HIV and SIV models. (123,153). Since escape mutations can occur both within and outside an epitope, effective methods to survey the emergence of mutations is crucial in determining regions that are under immune selection pressure. Further, identifying these key residues and correlating them with clinical outcomes of disease progression or survival can help us select effective vaccine immunogens.

1.5.2 HIV-1 and HLA Class I

Major Histocompatibility Complex (MHC) is the most polymorphic coding region within the human genome(154). Human Leukocyte antigens (HLA) are a group of genes within the MHC coding region in chromosome 6 in humans (154,155) and their major function is to recognize and distinguish between self and foreign antigens. Upon recognition of non-self antigens HLA activates host immune cells to respond to infection. There are two classes of HLA. HLA Class I genes code for HLA-A, -B and -C that present antigens to cytotoxic T lymphocytes (also known as CD8+ T cells). By presenting antigen peptides to TCR, HLA-peptide complex initiates the activation of $CD8^+$ T cells (Figure 1.2.). This process was first established by Zinkernagel and Doherty in 1974 and was called "MHC class I restriction"(156). HLA Class II genes which consist of HLA-DR, -DP, and -DQ present extracellular antigens to CD4+ T cells(157). The vast polymorphism of the HLA coding region is the result of its interactions with pathogens encountered throughout human evolutionary history(154). In HIV infection, HLA is one of the major host genetic factors that may determine the rate of disease progression and may play an important role in viral control, through the recognition of certain immunodominant epitopes(158,159). For example, in HLA-B*27-expressing individuals strong CTL responses to the immunodominant KK10 epitope of Gag was observed(137,160-162). Several cohort and genome-wide association studies (GWAS) have depicted the role of HLA alleles on disease susceptibility and control on a large scale. The first of these studies was conducted nearly two decades ago in 1996 where Kaslow and colleagues determined that HLA class I alleles B*57 and B*27 in a subtype B infected Caucasian cohort were associated with delayed onset of AIDS from the time of infection(160) .

In a South African study involving subtype C infected patients, B*58:01 was associated with slower disease progression(163); while the results of a GWAS study showed that in a Caucasian cohort of subtype B infected patients, B*07:02 was associated with susceptibility to progression to AIDS(162). Overall, these studies demonstrate the important role that HLA alleles play in both disease control and susceptibility. Determining host genetic factors that contribute to HIV control is therefore, an important approach that can help in the design of protective vaccine strategies.

Epitope recognition by HLA class I alleles restricted CD8+ T cells is a major driving force in the evolution of the virus(158) . To evade recognition by HLA, mutations may arise within an epitope or in the flanking regions, which impair peptide processing and presentation(164). While the emergence of escape mutants may present a challenge to the immune response and in designing protective therapeutics, the mutational patterns are largely predictable at the individual and population level based on the host HLA characteristics(158). For examples, in individuals who are B*57 positive, a T to N mutation at position three of TW10 epitope of p24 protein is selected(123,165). Similarly, in B*27 individuals, in the Gag KK10 epitope, M is replaced by L(166). The understanding that immune selection pressure drives HIV evolution within a population is based on the "HLA footprinting" hypothesis, which suggests that common HLA alleles influence the viral adaptation patterns that eventually determine the circulating consensus sequence within a population(167,168) . For example, A*24:02 is highly prevalent in

Japan, appearing in approximately 60-70% of the population(169). The common Nef Y135F mutation, where tyrosine (Y) is replaced by phenylanaline (F) at position 135 of the protein is selected by A*24:02. In the infected Japanese population F is the consensus at position 135, rather than Y, the global subtype B consensus(170-172). This illustrates how viral sequence diversity is determined and shaped by HLA alleles most common in the population. Therefore, viral sequence diversity is the result of selective forces imposed by CTLs restricted by common HLAs(158). The predictability of viral mutation pathways based on the host immune selection pressure provides the opportunity to determine regions of the virus that are highly targeted.

One of the major consequences of escape mutation is the loss of fitness incurred by the virus. Although the emergence of escape mutations allow the virus to survive under host CTL responses, many escape mutations result in a weakened and replication incompetent virus (158). Therefore, the consequences of viral escape mutations could be different. Viral fitness cost associated with immune escape is evidenced by several documented observations. First, *in vivo* studies have revealed that when a mutant virus is transmitted to a new host that lacks the restricting HLA allele in the original host, the virus reverts back to its wildtype form, indicating the wildtype is likely the more fit version of the virus and is thus preferable(140,173-177). For example, CTL responses during the acute phase of infection in B*57/B*5801 individuals select for an escape mutation within the TW10 epitope of gag, which reverts to wildtype upon transmission to B*57/B*580:1 negative hosts

(176). Second line of evidence for fitness loss, is reduced replication capacity in virus harboring escape mutation compared to a wildtype virus (158). In HLA-B57* individuals, while the T242N escape mutation in TW10 epitope results in reduced peptide-HLA binding and ultimately escape from host CTL, the mutant virus also loses its replication capacity measured by *in vitro* assays(178). In another study when major gag escape mutations, restricted by protective HLA B*27 and B*57 were engineered into the NL4-3 HIV-1 reference strain, the replication capacity of the virus was reduced *in vitro* (142). Finally, emergence of compensatory mutations that help to restore viral fitness also showed that HLA-associated polymorphisms confer a substantial replicative cost(179). In B*27+ individuals the common R264K mutation in the KK10 gag epitope results in reduced viral replication measured by *in vitro* assays(140). However, this mutation is often accompanied by the S173A substitution that takes place upstream of the KK10 epitope and in NL4-3 strains restores viral replication to that of the original wildtype virus(140).

Fitness cost associated with escape mutation is further evidenced by several biological disadvantages for the virus and clinical advantages for the infected host (145). Mutations that undermine host immunity may result in weakened virus with lower replication capacity. The clinical implications of a less fit virus is illustrated by reduced viral load(178) and higher CD4 counts in the infected host(180). Reduced pVL associated with escape mutant sequences has been shown in both early and chronic infection(181-183). Additionally, transmission of virus that harbours CTL escape mutations are overall less likely than those without(184). In cases where

mutant virus is transmitted to an HLA-mismatched individual plasma viral load is lower in the recipient host (185,186). In EC individuals, the selected escape mutations result in diminished viral protein function as demonstrated by *in vitro* studies(187) . Furthermore, protective HLA class I select for mutations that are associated with a higher fitness cost than non-protective alleles(142,188). These findings demonstrate the role of HLA-restricted CTL responses that select for escape mutants and viral mechanisms that restore overall fitness. A comprehensive understanding of escape mutations in HIV-1, their fitness cost and impact on disease progression is critical in informing therapeutic strategies.

1.6 HIV-1 NEF PROTEIN

1.6.1 HIV-1 Nef Protein Structure and Function

Negative Effector (Nef) protein is an accessary protein of HIV-1 that is approximately between 200-220 amino acids long, with 206 as the most common length(189). The HIV-1 Nef protein is Myristoylated, meaning that it contains a myristoly fatty acid group attached to its N- terminus, allowing Nef to interact with host cell signaling pathways. Nef consists of a variable domain, amino acids 1-65, an internal conserved domain, amino acids 70-200 and is flanked by an N-terminus, amino acids 2-61 and the Carboxy-terminus, amino acids 201-206(189).

Nef is a polyfunctional accessory protein of HIV that through its different functions contributes to HIV's infectivity and virulence. Nef is vital for increased pathogencity

of HIV and its role in clinical outcomes have been previously reported(190-192). A well-documented case which took place prior to donor screening guidelines were strictly implemented, is described in the Sydney Blood Bank Cohort (SBBC) study (193,194). A group of six patients received blood from an asymptomatic HIV-1 infected donor; some of these patients showed very slow progression to AIDS while others remained asymptomatic with normal CD4 counts and undetectable viral load levels for over two decades. Viral sequencing analysis revealed that the virus from the donor and the recipients lacked *nef* and *nef-LTR* region. This was the first case that showed HIV-1 Nef plays an important role in viral pathogenesis and its absence will result in favorable clinical conditions. Another example was seen in a group of Elite Controllers(195); Nef clones isolated from these patients were inefficient in CD4 and MHC downregulation and in their overall infectivity abilities compared to Nef isolates from those of rapid progressors(196). Nef-deficient or Nef-inefficient virus is therefore, less capable of infection and overall contribution to disease progression. These findings were further corroborated when rhesus macaque monkeys infected with attenuated SIVmac239 virus, lacking the *nef* gene, showed decreased levels of plasma viral load and slow disease progression (197).

1.6.2 Major Functions of Nef as a Virulence factor of HIV-1

Nef's ability to enhance viral pathogenicity and replication is accomplished by a number of functions it carries out early on during the viral life cycle and throughout chronic infection. These include downregulation of host CD4 and HLA molecules

from the surface and interfering with cell signaling pathways. Below each are briefly described.

CD4 downregulation

The heterotetrameric clathrin-associated adaptor protein 2 (AP-2), is responsible for the internalization and transport of material into the host cell through endocytosis. HIV-1's Nef can hijack this host cellular function by binding to the AP-2 protein in the cytoplasm and forming clathrine-coated pits that eventually internalize and removes CD4 receptor from the cell surface(195). CD4 receptors bind to Env protein of the virus initiating viral life cycle. However, during early infection this binding can result in ineffective budding and release of non-infectious (not fully assembled) virions. Nef's removal of CD4 from the cell surface prevents this from happening, therefore, allowing for the formation of fully assembled virions that can be released at a later time(198,199). Nef-induced down regulation of CD4 also contributes to reduced immunity as CD4+ T cells play a central role in cellular immunity(200).

MHC down-regulation

One of the major ways Nef contributes to host immune evasion is through the downregulation of MHC molecules from the surface of the infected cells (195,201). The mechanism involves the interaction of Nef protein and AP-1 and their binding to the cytoplasmic tail of MHC in the *trans*-Golgi network (TGN) of the cell(202,203). This interaction disrupts the normal pathway of MHC from the nucleus to the cell

surface via the TGN. With reduced expression of surface MHC, fewer CD8+ T cells are engaged and thus the infected cells evade the host immune response.

Activation of cell signaling pathways

Another function of Nef is to reduce the motility and migration of infected CD4 T lymphocytes. This localizes the infection in an environment that is rich in HIV-1 target cells such as other nearby CD4+ T lymphocytes, DC and macrophages and also enhances host immune evasion(195). Reduced motility of infected host cell is accomplished by the ability of the Nef protein to interact with host cell signaling pathways. Nef interacts with various kinases within CD4 T cells such as P21 activated kinase (Pak2) family. Nef-mediated Pack2 activation results in the degranulation of cofilin, a host cell protein involved in cellular organization and motility(204). In this way Nef utilizes yet another cellular mechanism to the advantage of the virus. In HIV-1 infection CD4 T cell activation is associated with disease progression(205,206). Therefore, the deleterious effects of Nef interfering with cellular signaling pathways that result in CD4 T cell activation have implications for clinical outcomes.

1.6.3 Nef immunogenicity

The immunogenicity of Nef protein has been studied in both human and NHP models where Nef specific immune responses have been shown to mediate effective immune control of infection(190). Nef is among the first proteins to be generated during the HIV-1 life cycle, due to its crucial role in increasing the infectivity of the

virus from the outset, and is the most immunogenic among all other HIV-1 accessory proteins(207). Nef is the most targeted protein in acute infection and accounts for 50-90% of CD8 T cell responses during the acute phase. During the chronic phase, Nef is responsible for the highest magnitude of IFN-γ responses compared to other viral proteins(190). Some studies have even shown that the sequence variability of Nef is likely driven by HLA Class I selection pressure(143). In fact, HLA selected mutations in Nef have been associated with viremia suppression. For example, in HLA-B*57 and -B*58:01 LTNP individuals, selection of HW9 (HTQGYFPDW) epitope in Nef was associated with sustained and frequent CD8+ T cell responses(196). In SIV models Nef specific CD8+ T cell responses were associated with disease control. Elite viremia control was observed in Mamu-B*08+ rhesus macaques targeting Nef RL10 epitope (RRHRILDIYL) (208) and in cynomolgus monkeys whose CD8+ T cells targeted epitopes in the carboxy terminus of Nef (209). In a large cohort study CD8+ T cell responses to specific Nef regions were associated with lower viral set point(190). These studies showed that targeting specific regions of Nef could contribute to disease control and the potential of Nef epitopes as a T-cell vaccine candidate.

1.7 Project Rationale, Hypothesis and Objectives

1.7.1 Rational and Gaps in Knowledge:

Development of an effective vaccine against Human Immunodeficiency Virus-1 remains one of the grand challenges in global health despite three decades of study

and research. The major obstacles are the insufficient understanding of the protective immunity of the host in response to HIV-1 infection and the rapid mutation and diversity of the virus. Since HIV-1 depends on the host to survive and the effect of host immune response is imprinted in the sequence of the virus, characterization of the interactions between virus and the host immune system can provide clues to which immunogens should be included in an effective HIV vaccine candidate. Several studies have shown that Human Leukocyte Antigen restricted Cytotoxic T Lymphocyte responses play a major role for immune control of HIV infection(113) (210) (162). HIV-1 mutates to escape CTL recognition and to survive under host immune pressure. Mutations that allow the virus to survive overtime are referred to as positively selected mutations (PSMs)(157,211). However, the effects of PSMs are not always equal; while they allow the virus to survive, there is sometimes a fitness cost associated with mutations(142,178,212,213). Identifying and characterizing these mutations can help select immunogens for candidate vaccines. Since the PSMs reflect viral escape from host immune responses in the context of HLA class I restricted CD8+ T cells, correlating PSMs in HIV-1 with specific host HLA alleles and disease progression status is a promising approach to identify best immunogens to be included in a candidate HIV vaccine.

HIV-1 Nef plays an important role in the pathogenesis of the virus(192,214). Recently some studies have identified certain regions of Nef that may were associated with protective clinical outcomes, such as decreased VL(143,190,196,215,216). However, in majority of these studies treatment naïve

samples are limited to one time point since in most developed countries ART treatment is initiated immediately after a patient is diagnosed HIV-positive, making it challenging to determine host-driven mutational patterns overtime. More importantly, these studies also lack patient disease progression data, such as longitudinal CD4 counts before the initiation of ART. Therefore, it is not possible to determine if HLA driven PSMs are correlated with disease outcome. This is a major knowledge gap: to determine the clinical consequences of HIV-1 Nef PSMs in disease outcome and ultimately to define best immunogens in Nef that may be included in an effective HIV-1 vaccine.

1.7.2 Hypothesis and Objective:

This thesis has four major hypotheses:

1. The population frequency of PSMs in HIV-1 Nef from subtypes A1, B, C and D remain stable overtime and PSMs in HIV-1 Nef motifs are consistent among subtypes A1, B, C and D since functional motifs represent conserved regions of the virus.
2. HIV-1 Nef CD8+ T cell epitopes contain positive selected mutations (PSMs) that differentially influence disease progression. These epitopes can be identified using correlation with host HLA class I alleles.
3. HLA class I restricted CD8+ T cell epitopes containing PSMs will elicit CD8+ T cell responses that can be measured by ELISPOT assay.

4. CD8+ T cell populations specific to epitopes that contain PSMs associated with differential CD4 decline will differ in their proliferation, expression of antiviral intracellular cytokines and exhaustion characteristics.

Hypotheses will be tested by the following Objectives:

1. **To identify and compare PSMs in four major M group subtypes, A1, B, C and D and analyze their mutational pattern overtime (Chapter 3)**

   HIV-1 Nef sequences from four major M group subtypes (A1, B, C and D) will be obtained from Los Alamos National HIV Laboratory Database and PSMs in Nef will be identified by a bioinformatics approach called quasi analysis selection mapping program.

2. **To amplify and sequence HIV-1 Nef and identify PSMs in Nef using bioinformatic analysis (Chapter 4)**

   HIV-1 *nef* gene will be amplified from 508 HIV-1 infected, antiretroviral drug naïve women using a nested PCR approach (Section 2.7.1). The amplified *nef* will be sequenced using the 454 pyrosequencing technology. PSMs in HIV-1 Nef will be identified by quasi analysis.

3. **To correlate PSMs with disease outcome and HLA class I alleles (Chapter 4)**

   PSMs will be correlated with HLA class I alleles and CD4+ T cell counts. A database consisting of all the Nef PSMs will be established and merged with a

database of host HLA class I alleles of all study participants. Fisher's exact test will be used to correlate PSMs in Nef with HLA class I alleles; Kaplan-Meier survival analysis will be used to determine PSMs associated with different rates of CD4 decline.

4. **To confirm epitopes containing PSMs that correlate differentially with CD4 decline (Chapter 5)**

   IFN-γ ELISPOT assays with peptides overlapping the identified PSMs will be used to confirm potential HLA-restricted CD8+ T cell epitopes identified through correlation analysis. For PSMs that significantly correlate with specific HLA alleles, 9mer peptides will be synthesized. These epitopes will contain the identified PSMs and consensus amino acids at anchor positions 2, 8 or 9 and middle position 5, which have been shown to impact peptide-HLA binding.

5. **To characterize immunological signature of CD8+ T cell responses induced by epitopes containing beneficial or detrimental PSMs (Chapter 6)**

   Tetramers containing peptides with PSM or consensus amino acids, from beneficial and detrimental categories will be obtained (NIH Tetramer Core Facility). The tetramer positive cells will be stained using surface and intracellular CD8+ T cell marker panels. Flow cytometry analysis will be performed using multicolor BD LSR II flow cytometer to assess frequency,

proliferation, expression of antiviral intracellular cytokines and exhaustion characteristics of tetramer-specific CD8+ T cells.

2. General Materials and Methods

2.1 Study Ethics

All patients enrolled in the study provided informed consent. The Ethics Committee of the University of Manitoba and the Ethics and Research Committee of Kenyatta National Hospital approved this study

2.2 Study Cohort and Participants

Samples used in this study were from the Majengo (ML) female sex-worker cohort, from Pumwani district of Nairobi, Kenya. The cohort was established in 1985, through collaboration between the University of Manitoba and University of Nairobi. Since its establishment, approximately 4,000 individuals have been enrolled, with close to 600 participants in active follow-up. At biannual resurveys, serology and PCR tests are used to determine the HIV-1 status of enrollees. From each participant, 21 ml of heparinized blood was collected for immunological assays, and 7ml additional of blood in ethylenediamine tetraacetic acid (EDTA) was collected to determine CD4 count and viral load. Study participants also completed a behavioral questionnaire, were screened for STIs and received medical treatment, including ARV after 2003 for HIV+ participants with CD4 counts below 250/mm^3, counseling and free condoms.

The members of the study cohort are divided into four major epidemiological categories based on participants' clinical features:

HIV infected: participants who were tested HIV positive (PCR and serology) at enrolment.

HIV-seroconvertors: participants who were negative at enrollment but seroconverted at later follow up

HIV Exposed Sero-negative (*HESNs):* participants who remained HIV-negative for ≥7 years despite high risk of exposure to HIV due to commercial sex-work.

HIV-susceptible: participants who were HIV-negative at enrolment, but have not been followed up long enough to meet the time requirement for HESN.

For this study, the samples used were either from HIV-infected or seroconverted participants.

To optimize experimental conditions samples from local donors (Winnipeg, Manitoba) were also obtained. The Ethics Committee at the University of Manitoba has approved the local donor program.

2.3 HIV-1 Testing

HIV-1 testing in the Majengo cohort consisted of two tests; The first test, was the Recombigen enzyme-linked immunosorbent assay (ELISA) (Trinity Biotech, Ireland), samples are tested for HIV Antibodies (Ab) at the Pumwani clinic. If this test is positive, i.e. HIV Ab detected, then a second test, Detect HIV1/2 Immunoassay (Adaltis Inc.) was used to confirm positive results in the lab. Only samples tested positive for both tests were considered HIV-positive.

2.4 CD4 T Cell Count

To measure CD4+ T cells for each cohort participant, whole blood collected in EDTA tubes, were analyzed using Tritest reagent CD3/CD4/CD8 (BD Bioscience) by flow

cytometry assay. Absolute CD4+ T cell counts were calculated by multiplying the lymphocyte counts to the percentage of CD4+ T cells.

2.5 HLA Class I typing

DNA was isolated from blood using the QIA Amp DNA mini Kit and QIA EZ1Blood Robot (QIAgen Inc, Mississauga, ON, Canada) and used to type the HLA Class I allele for each participant. HLA A, B, and C genes were amplified by PCR with specific primers(217-219). After purification the PCR products were sequenced using BigDye cycle sequencing kit (Applied Biosystems, Foster City, CA) with specific primers. Each sequence was then analyzed with AB13100 or ABI3130xl Prism Genetic Analyzer. Ambiguous combinations were resolved with allele specific primers. To assign HLA-A, HLA-B and HLA-C, a taxonomy-based sequencing software, called CodonExpress™ was used(220,221). The peptide binding region of HLA molecule are encoded by exon 2 and 3 of the HLA chromosome, therefore, only these regions were analyzed. Alleles were typed to 4-digit resolution. This type of resolution was also used for the assignment of homozygotes for a specific allele(222).

2.6 Sample Preparation

2.6.1 Isolation of Peripheral Blood Mononuclear Cells (PBMCs) from Blood

Isolation of PBMCs from whole blood, was done using the Ficoll-Hypaque density gradient isolation technique. In summary, whole blood was first centrifuged at 1200 rpm for 10 minutes to separate plasma from other blood components. The remaining part of the blood was diluted 1:1 with 2%FBS-PBS solution (Life Technologies, Cat. No.12483020), and carefully layered over approximately 10 ml of Ficoll (Cedarlane, Cat. No. 1114547). This layered mixture was then centrifuged at 1400 rpm for 25 minutes, with the centrifuge's break completely off to avoid mixing of the layers. The PBMC layer was carefully collected, then washed with 2%FBS-PBS by centrifugation at 1600 rpm for 10 minutes. Cells were resuspended in 10ml of R10 culture media and counted using hemocytometer, and the Trypan blue exclusion dye (ThermoFisher) used to determine cellular viability. The purified PBMCs were aliquoted at 1.0 x10^7 cells/ml in freezing medium (8%DMSO and 92% FBS) for long-term storage in cryovial tubes and store in liquid nitrogen tank (LN2) tank or -80°C following the standard protocol.

2.6.2 Thawing and Resting Cryopreserved PBMCs

Freeze-thawed-rested PBMC samples were used in ELISPOT and flow cytometry assays. . In brief, cryovials containing frozen PBMCs, were placed in 37°C water bath until the content had mostly thawed. Then, R-10 media at 37°C was added drop-wise slowly to the PBMC suspension in thawed freezing media, to prevent cell death due to osmotic pressure. Fully thawed PBMCs were mixed with10 ml of R-10 media and centrifuged at 1600 rpm for 10 minutes. The wash and centrifugation steps were repeated to ensure removal of any remaining dimethyl sulfoxide (DMSOs).

Finally, cells were allowed to equilibrate at 37°C, 5% CO_2 overnight (6-12 hours), in plastic round bottom loose cap tubes(VWR).

2.7 Viral Amplification and Sequencing

2.7.1 HIV-1 Full-length genome amplification

Genomic DNA extracted from PBMCs, buffy coat and whole blood using QIAamp DNA Mini Kit (Qiagen Inc, Mississauga Ontario) was used for amplification of HIV proviral DNA. To amplify HIV-1 provirus, published primers MSF12b and ofm19 (HXB2 locations 623-649 and 9632-9604 respectively) were used to generate an amplicon approximately 9 kb. The Roche Expand Long Range PCR kit was used for the primary PCR. The optimized conditions were as follow: two minute initial denaturation cycle at 94°C, followed by 25 amplifications cycles that consisted of 10 seconds of further denaturation at 94°C (10x), 30 seconds of primer annealing at 60°C , 8 minutes of DNA elongation at 68°C. An additional 25 extension cycles that consisted of 15 second denaturation at 94°C, 8 minutes plus 20 sec/cycle of elongation and final 7 minute elongation at 68°C concluded the amplification and finally the reaction was cooled to 4°C.

To amplify the *gag*, *pol* and *env* amplicons, the Roche Expand High Fidelity PCR kit was used with forward and reverse primers listed in Table 2.1. A 2 kb *gag* amplicon was generated using forward primer- alnf1 and reverse primer -GagRT. We used the following optimized PCR amplification conditions: two minute initial denaturation cycle at 94°C to activate Taq polymerase, followed by 10 amplifications cycles

consisting of a 10 second denaturation at 94°C, 30 second primer annealing at 62°C and a 3 minute DNA elongation at 72°C. Additionally, 25 extension cycles that consisted of 15 seconds denaturation at 94°C, 30 seconds of annealing at 60°C, 2 minutes and 10 seconds/cycle of elongation at 72°C and a final 7 minute elongation at 68°C concluded the amplification. The reaction was concluded by cooling to 4°C.

To amplify the second amplicon, *pol*, which is approximately 3.7 kb, forward primer alpf1 and reverse primer alpr1 were used. The following conditions were applied: two minute initial denaturation at 94°C, followed by 10 amplifications cycles that consisted of 10 seconds of further denaturation at 94°C, 30 seconds of primer annealing at 59°C, 8 minutes of DNA elongation at 68°C. An additional 25 extension cycles that consisted of 15 second denaturation at 94°C, 30 seconds of annealing at 57°C, 3 minutes and 10 seconds/cycle of elongation at 68°C and a final 7 minute elongation at 68°C concluded the amplification. The reaction was concluded by cooling to 4°C.

Finally, 3.6 kb *env* was amplified using forward primer Envpcrf and reverse primer env reverse. The env amplification reaction was as follows: a two minute initial denaturation cycle at 94°C, followed by 10 amplifications cycles that consisted of, an additional10 second denaturation at 94°C, a 30 second primer annealing at 60°C, 8 minute DNA elongation at 68°C and 25 extension cycles of 15 second denaturation at 94°C, 30 seconds annealing at 60°C, 3 minutes and 10 seconds/cycle of elongation at 68°C and a final 7 minute elongation at 68°C concluded the amplification. The reaction was concluded by cooling to 4°C.

Table 2.1 primers used for the amplification of HIV provirus, *gag*, *pol* and *env* amplicon.

Gene product	Primer name	Primer sequence
full genome, forward	msF12B	5'-AAATCTCTAGCAGTGGCGCCCGAACAG-3'
full genome, reverese	ofm19	5'-GCACTCAAGGCAAGCTTTATTGAGGCTTA-3'
gag, forward	alnf1	5'-GCCCGAACAGGGACYYGAAAGCGAAAG-3'
gag, reverse	GagRT	5'-CCATTGTTTAACCTTTGGGCCATCCA-3'
pol, forward	alpf1	5'-TAGGACCTACACCTGTCAACATAATTG-3'
pol, reverse	alpr1	5'-TCATTGCCACTGTCTTCTGCTCTTTC-3'
env, forwrad	Envpcrf	5'-GGCTTAGGCATCTCCTATGGCAGGAAGAAG-3'
env, reverse	*env* reverse	5'-GGCAAGCTTTATTGAGGCTTAAGCAGTG-3'

2.7.2 Purification of PCR products

Purification of amplified PCR products was done using the Quantum Prep Freeze 'N Squeeze DNA Gel Extraction Spin Column (Bio-Rad, Cat. No. 432-6166) technique (223). First, the PCR products were separated using agarose gel electrophoresis with 1% agrose gel stained with SYBR™ Green I dye. The bands of interest were then excised and trimmed using sterile razor blades. Gel slices were placed inside filter cups of the Quantum Prep Spin Columns and stored at -20^{O}C for 5-10 minutes. The columns were then spun at 13,000 x g for 3 minutes at room temperature and purified DNA was collected. Ethanol precipitation method was used to further purify DNA, which was then stored in 30-50 µl of 10mM Tris Buffer pH 8.5 for pyrosequencing.

2.7.3 454 Pyrosequencing

All library preparation, emulsion PCR and sequencing kits were purchased from *Roche* Life Sciences.

The following kits were used:

- GS Rapid Library Rgt/Adaptors Kit (Cat. No. 5619203001)
- GS Rapid Library MID Adaptors Kit (Cat. No. 5619211001)
- GS Titanium MV emPCR Kit (Lib-L) v2 (Cat. No. 5618436001)
- GS Titanium emPCR Breaking Kits LV/MV 12pc (Cat. No. 5233658001)
- GS Titanium Sequencing Kit XLR70 (Cat. No. 5233526001)
- GS Titanium PicoTiterPlate Kit 70x75 (Cat. No. 5233682001)
- GS Titanium Control Bead Kit (Cat. No. 596643001)

Equipment used:

- Genome Sequencer FLX (GS FLX), 454 Life Sciences

2.7.3.1 Library Preparation

One microgram of the purified DNA product obtained following gag, pol, env PCR, was dissolved in 100 µl Tris-EDTA buffer (10 mM Tris-HCl, pH7.5-8.0; 0.1 mM EDTA) and used for DNA fragmentation by the nebulization method. This created shorter double-stranded DNA fragments of approximately 400 to 600 bp long according to double-helix DNA ladder. Nebulized DNA samples were purified on column using MinElute PCR purification kit (Qiagen). Fragmented DNA ends were then repaired by ligation using the DNA End Repair mix (2.5 µl RL 10x PNK Buffer, 2.5 µl RL ATP, 1 µl RL dNTP, 1 µl RL T4 Polymerase, 1 µl RL PNK, 1 µl RL Taq Polymerase), followed by the addition of adaptors to bind to each fragment. Next, the repaired DNA-adaptor complexes were added to AMPure beads for removal of small fragments of DNA. Double stranded DNA was denatured to form single stranded DNA to be used in emulsion PCR.

2.7.3.2 Emulsion PCR

The single stranded adaptor carrying DNA fragments, were added to DNA capture beads and PCR reaction reagents in a water mixture followed by mixing with emulsion oil (a synthetic oil). This created droplets around each bead carrying a DNA library ("micro-reactor") in the water mixture; each DNA fragment can be amplified using PCR to approximately 1.0×10^7 identical copies of DNA. After the completion of the PCR step, capture beads were cleaned from the emulsion oil. First

oil was collected from each well using a syringe followed by a wash step with 100µl of isopropanol and discarding of oil through emPCR filters (GS Titanium, Roche). Next, the filter was removed from syringe and bead suspension was dispensed into a 1.7 ml collection tube, which was spun for 10 seconds using a bench top mini centrifuge. Supernatant was discarded and with filter attached to syringe Enhancing Fluid XT (GS Titatnium, Roche) was drawn up to 0.5 ml. Filter was then removed and bead suspension was collected in the same 1.7 ml collection tube. Beads not attached to DNA were eliminated using enrichment beads and centrifugation and wash steps. For annealing sequence primers to DNA, 48 µl of primers were added to the mixture followed by placing collection tube at $65^{O}C$ for 5 minutes and then cooled on ice for 2 minutes. Annealing buffer was added at 500 µl, followed by a vortex and spin-rotate-spin step and supernatant was discarded. Finally 100 µl of annealing buffer was added to the pellet and vortexed. Bead enrichment was calculated as percent of number of enriched beads recovered by total input bead ($19.2x10^{6}$). Beads were stored at 2-8 ^{O}C and sequenced within 2 weeks.

2.7.3.3 Sequencing

Capture beads with amplified DNA from the emulsion PCR step, were loaded unto the pico-titre plate of the sequencer FLX, which contains $1.6\ x10^{6}$ wells. The small wells size (approximately 44 µm), enabled the loading of one bead per well. Enzyme beads necessary for light conversion and detection, were then added to the mixture. Single nucleotides A (Adenine), C (Cytosine), G (Guanine) and T (Thymine), were then washed over the pico-titre plate sequentially 4 times, 1 per nucleotide,

allowing for the formation of the second template of DNA. Each chemical reaction between a nucleotide and the DNA template strand was converted to a light signal owing to luciferase enzyme activity. The light generated was then captured by the charge-coupled device (CCD) camera on the Genome Sequencer FLX. The signal detected was proportional to the number of complementary nucleotides incorporated. The intensity of the signal was proportionally stronger when more of the same nucleotide was incorporated consecutively on the single-stranded DNA fragment. For each run the sequencing software of the instrument generates signals that can further be analyzed.

2.8. Bioinformatics Analyses

2.8.1. Generation of Consensus Sequence

Pyrosequencing of HIV full-genomes was performed at the Genomic Core of National Microbiology Laboratory (NML) using Roche GS FLX by standard protocol. The raw HIV sequence read-outs obtained by pyrosequencing, were first decoded by sample IDs then classified into the corresponding sample groups. In-house quality control (QC) was then conducted on the labeled sequence reads to filter out the poor-quality reads either in short length (less than 100 bps) or with low average quality scores (less than 20). The sequence read-outs, which passed this QC step, were mapped to the HXB.2. reference sequence using BLASTN (BioPerl Toolkit). Only sequence reads with calculated average identities (proportion of aligned read length to total length of read) $\geq$ 65% and percent identities (the similarity between a sequence read and a reference sequence) > 75% were remained in "pileup" alignment for generating

consensus sequence. The nucleotides of consensus sequences were resolved one base at a time based on the most abundant nucleotide(s) that appeared at each column. IUPAC DNA codes were finally assigned to the nucleotides. Insertions were reported if their abundance was $\geq$ 50% of the coverage of the preceding position and deletions were reported as dashes (-).

2.8.2. Subtype Classification

The qualified sequence reads from each study participant, were blasted to HIV subtypes A1, A2, B, C, D, and G sequences which correspond to nine HIV-gene references (*Env, Gag, Nef, Pol, Vif, Vpr, Vpu, Rev, Tat*) from Los Alamos National HIV Laboratory (LANL) database as described in consensus sequence generation. Average percent identities from each "pileup" alignments were compared to determine the best matching clade reference per gene. Subtypes were assigned to genes based on the highest percent identity.

2.8.3 Identification of PSMs using Quasi Analysis

Positively selected mutations (PSMs) were identified using a selection mapping program called, Quasi, which calculates the overabundance of non-synonymous (replacement) mutations compared to synonymous (silent) mutations at every cordon (129). Positive selected (PS) sites were recorded along the consensus sequences. The frequency of PS sites in each subgroup was calculated as the percentage of the number of PS sites out of the total number of amino acids of the consensus sequence.

For a set of viral nucleotide sequences, we determine the variants that confer selective advantage by measuring the empirical replacement to silent mutation ratio (R:S) of each possible amino acid replacement and then comparing this observed ratio to that which would be expected if mutation were unselected. An R:S that is found to be higher than expected indicates that the replacement mutation tested is positively selected, while a lower-than-expected observed R:S indicates that the tested replacement mutation is negatively selected.

2.9 Enzyme-Linked ImmunoSpot (ELISPOT)

Ninety-six well nitrocellulose plates were coated with anti-IFN-γ monoclonal antibody (mAb; Mabtech, Nacka Strand, Sweden) and left overnight at 4° C followed by blocking the reaction using R10 media. Peptides were diluted to a final concentration of 10 μg/ml in R10 media and used to stimulate 1.0 $x10^5$ patient PBMCs resuspended in RPMI media (Life Technologies, Cat. No. 11875119). The peptide-PBMC mixture was immediately transferred into pre-coated ELISPOT plates, and PBMCs were stimulated overnight in duplicates at 37°C and 5% CO2 for between 12-16 hours, with either phytohemagglutinin (PHA) or media as positive and negative controls, respectively. Plates were washed using commercially available wash buffer and, biotinylated anti-IFN-γ mAb added to each plate, followed by a 2 hour incubation at room temperature. Next, streptavidin-conjugated alkaline phosphatase was added to each well of the plate and incubated for 1 hour at room temperature. Detection was done using alkaline-phosphatase substrate kit (Bio-Rad, Cat. No. 1721063). An ELISPOT Reader (Autoimmun Diagnostika GmbH

AID iSpot) was used to read the plates by counting spot-forming units (SFU). Positive responses were determined as those with at least 50 SFUs/million PBMCs after background subtraction and a successful positive control reaction for the same patient.

2.10 Flow Cytometry

Surface and intracellular markers of CD8+ T cells were detected using antibodies listed in Table 2.2 and Table 2.3.

Table 2.2. Flowy cytometry panel of surface markers of CD8+ T cells

Fluorochrome	CD8+ T cell Marker	Volume used (µl)	Manufacturer	Cat. No.
AmCyan (Fixable Far Red)	Live/Dead	6.25	ThermoFisher	L23102
APC-H7	CD3	2	BD Bioscience	560176
BV605	CD4	1	Biolengend	300556
BV650	CD8	1	Biolengend	301042
BV786	Ki67	5	BD Bioscience	563756
PeCy7	CCR7	3	BD Bioscience	557648
BB515	CD45RA	3	BD Bioscience	564552
BV421	PD-1	1	Biolengend	329920
BV510	TIM3	1	Biolengend	345030
PE	CD160	1	BD Bioscience	562118
APC	Tetramers	0.125 µg/ml	NIH	--

Table 2.3. Flow cytometry anel of intracellular (ICS) markers of CD8+ T cells

Fluorochrome	CD8+ T cell Marker	Volume used (ul)	Manufacturer	Cat. No.
PE-CF594 (Fixable Aqua)	Live/Dead	1	ThermoFisher	L34957
APC-H7	CD3	2	BD Bioscience	560176
BV605	CD4	1	Biolengend	300556
BV650	CD8	1	Biolengend	301042
BV421	IFNγ	1	Biolengend	502532
Alexa700	TNFα	2	BD Bioscience	557996
PE	IL-2	15	BD Bioscience	559334
pe daz	GM-CSF	1	BD Bioscience	562857
BV786	CD107a	2	BD Bioscience	563869
APC	Tetramers	0.125 µg/ml	NIH	--

2.10.1 Titrations

To conduct accurate flow cytometry analysis all antibody and tetramers conjugated with a fluorochrome (see Table 2.2 and 2.3) were titrated at varying volume (for Abs) and concentration (for tetramers). A brief description is provided below:

Antibody Titration

Each Ab was titrated at several varying volumes, including the manufacturer's recommendations. PBMCs were isolated from local blood donor and used for surface marker and intracellular staining with respective antibodies of different volumes. We used a staining index calculation; MFI of Ab-negative population was subtracted from MFI of Ab-positive population and divided by two times the standard deviation of MFI of the Ab-negative populations to determine indices for each experimental volume and for comparison(224).

Tetramer Titrations

Two-fold dilutions ranging from 15 ug/ml and 1.875 ug/ml (manufacturer's recommendation) of tetramer were used to determine optimal tetramer concentration that resulted in high fluorescent intensity with low background staining. PBMCs from specific local donors with HLA alleles known to restrict epitope of interest were used for tetramer staining protocol. We determined that the best concentration of tetramer was 0.125 ug/ml for IT9-N, KY9-I and KY9-V and 0.5 ug/ml for IT9-H; these were used as final concentration for respective

tetramers-staining protocols. (Note: tetramer and peptide names are described in chapter 5, Figure 5.1)

2.10.2 Fluorescent Minus One

Flourescent Minus One (FMO) is a technique used to determine the best gating strategy for a specific flurochrome where multiple markers are examined through a flow cytometry panel (e.g. Table 2.2 & 2.3), and fluorescent spillovers between some flurochromes are inevitable. In brief, several experimental conditions each containing all antibodies in a panel, except for one, were carried out. Two additional experimental tubes, one with all antibodies and another one without any of the antibodies were also included in the experimental conditions. Through individual exclusion of one antibody per each tube, all fluorescent spillovers into all channels were taken into account.

2.10.3 Compensation Controls

In flow cytometry the compensation technique is essential when using multiple fluorochromes. Its purpose is to take into account and remove fluorescent spillover from multiple markers that usually occurs as a result of overlap between each marker's emission spectra. This is accomplished with the use of commercially available compensation beads (BD Biosciences). Flow data acquisition was done using BD FACSDIVA software (BD bioscience).

A drop of either the positive and negative control CompBeads (BD bioscience), was added to FACS Tube (BD bioscience) already containing 300 µl of PBS , which was

mixed by vortexing. Anti-mouse and anti-rat CompBeads were used for mouse-anti-human and rat-anti-human antibodies. Then, 1 µl of fluorescently labelled antibody was added to the appropriate tube containing 300 µl of the diluted CompBeads, which was incubated for approximately 20 minutes at room temperature in the dark.

The ArC™ Amine Reactive Compensation Bead Kit (Thermofisher) was used for preparing compensation controls for LIVE/DEAD aqua and vivid red dye (ThermoFisher). Here, 1 drop of arc beads and 1 µl of antibody (aqua or red LIVE/DEAD) were incubated for 30 minutes at room temperature, then washed with 2 ml FACS wash before centrifugation at 300 xg for 5 minutes. Beads were then washed with 300 µl of FACS wash and negative control beads were added.

2.10.4 Flow Cytometry: Surface Marker Staining

Approximately $5.0x10^5$ PBMCs per study participant sample, was used in surface staining for flow cytometry analysis. PBMCs were aliquoted into FACS tubes (BD Bioscience), and washed using FACS Wash media. Next 40 µl of tetramer was added at final concentration of 0.125 µg/ml to each tube, followed by 30 minutes of incubation in 4°C in dark to protect the photosensitive APC flourochrome on tetramers. Next surface antibody cocktail (Table 2.2.) was added with the final volume of 50 µl. Cells were then incubated for 30 minutes at 4°C in dark. Next, cells were washed using 2 ml FACS wash and centrifuged at 1600 rpm for 10 minutes. The supernatant was discarded and the pellets were immediately processed with the Ki-67 staining protocol (described below).

2.10.5 Flow Cytometry: Intracellular Marker Staining

PMBCs were thawed and rested overnight at 37°C (Section 2.6.2). The rested cells were washed with 2 ml FACS wash media, centrifuged at 1600 rmp for 10 min. Supernatant was discarded and the pellets were resuspended in R10 media. Approximately $1.0x10^6$ PBMCs were aliquoted into sterile FACS tubes. 40 µl of tetramers were added to appropriate tubes and cells were incubated for 30 min at 4°C in dark. The tetramer-staining step must be done prior to cell stimulation (by peptide or mitogen) since TCRs on the surface of CD8+ T cells are downregulated with stimulation(225). Therefore, a pre-stimulation tetramer staining ensures that TCRs can bind to peptides presented by the tetramer complex prior to downregulation. After the tetramer staining step, a pre-titrated amount (2 µl) of the CD107a antibody was added to tubes and cells were incubated for 20-30 minutes at 4°C protected from light. CD107a is a marker of degranulation that appears transiently on T cell surface and therefore, it needs to be stained for prior to intracellular staining(226).

Following tetramer and CD107a staining steps cells were stimulated with the appropriate peptides at the final concentration of 5µg/ml. For positive control stimulation, a combined mixture of phorbol 12-myristate 13-acetate (PMA, Sigma) and ionomycin (Sigma), at final concentration of 25ng/ml and 500 ng/ml respectively, were used. To enhance stimulation the co-stimulatory reagent, CD28/CD49d (BD FastImmune), was added to all stimulation conditions and cells were incubated for 1 hour at 37°C(227). When cells are stimulated with either

peptide or a mitogen (e.g. PMA/Ionomycin), the intracellular protein transport process needs to be blocked in order for cytokines and other proteins to be accumulated in the Golgi complex; this increased level of cytokines allows for an enhanced detection with flow cytometry(227). To achieve this, 1 µl of protein transport inhibitors Golgi Plug (brefeldin A, BD Biosciences) and Golgi Stop (monesin, BD Biosciences) were added to all stimulation tubes. Next cells were incubated for an additional 6 hours at at 37°C and protected from light.

At the end of the 6-hour stimulation, the surface staining protocol (Section 2.10.4) was carried out without fixing with PFA. Instead, 250 µl of Cytofix/Cytoperm (BD Biosciences) solution was added to all tubes for cells to become permeabilized so that intracellular antigens can be stained with their respective fluorochrome-conjugated anti-cytokine antibodies. Cells were then incubated for 20 minutes at 4°C in dark, followed by a wash step with 1 ml BD PermWash buffer and centrifugation at 1600 rpm for 10 min. To the supernatant, 50 µl of antibody cocktail of intracellular markers (Table 2.3) was added and cell were incubated for 30-40 minutes at 4°C in dark. Cells were washed with 1 ml PermWash buffer and centrifuged at 1600 rpm for 10 minutes and were resuspended in 300 µl of FACS wash. Tubes were then read on BD LSR II flow cytometer.

2.10.6. Flow Cytometry: Intracellular Staining of Ki67

The intracellular staining of Ki67 is carried out after the PBMCs were stained with all surface markers. The procedure is as follows: after the completion of surface marker staining (Section 2.10.4) PBMCS were washed and centrifuged at 1600 rpm

for 10 minutes and 1 ml of transcription factor Fix/Perm buffer (BD Bioscience TF Fix/Perm) was added to the pellet while vortexing gently. Cells were then incubated at 4°C for 40-50 minutes, followed by two wash steps with 1ml (first wash) and then 2 mls of TF Perm/Wash buffer (BD Bioscience), and centrifugation at 350 xg at 2-84°C for 6 minutes. A pre-titrated amount of Ki-67 (5 µl), in a total volume of 100 µl with TF Perm/Wash buffer was used to stain cells. PBMCs were incubated for 50 minutes in dark at 4°C. Following several wash and centrifugation steps, 281 µl of FACS wash buffer and 19 µl of 16% paraformaldehyde (PFA) were added to cells to fix and prepare cells prepare cells to be analyzed on BD LSR II flow cytometer.

2.11. General Reagents

The following solutions were made in-house for cell preparation (PBMC isolation and cryopreserving) and flow cytometry. For all other techniques, reagents and solutions were purchased from manufacture (listed throghout).

- FBS (Gibco, Invitrogen)
    - heat inactivated at 56°C for 1 hour
- Freezing Media
    - 10% dimethyl sulfoxide (DMSO, tissue culture grade, Sigma) and 90% heat-inactivated FBS
- Phosphate-buffered saline (PBS) (Gibco, Invitrogen)
    - 48.5g PBS powder, 137.93mM NaCl, 2.67mM KCl, 8.1mM Na_2HPO_4, 1.47mM KH_2PO_4 (Gibco, Invitrogen), dissolved in 1L of ddH_20.
- Fluorescence-activated cell sorting wash (FACS wash)

- PBS supplemented with 2% FBS (Gibco, Invitrogen)

- 16% Paraformaldehyde (PFA)
  - 48ml ddH_20, 2ml 5M NaCl, 5g PFA (Electron Microscopy Sciences). H
  - heated for 1 minute and 20ul of 10N NaOH

- R10 Cell culture media (R10)
  - Roswell Park Memorial Institute (RPMI) 1640 media (HyClone, Thermo Scientific) supplemented with 10% heat inactivated FBS and 1% Penicillin/Streptomycin (Gibco, Invitrogen)

- Perm wash
  - 1x of the stock buffer prepared in ddH_2O.

3. Identification and Characterization of PSMs in Nef of Four HIV-1 Major Subtypes from Los Alamos National Laboratory

3.1 Rationale:

As a member of the Retroviridae family, HIV-1 is highly diverse. Group M is the major HIV-1 group and is responsible for the current global pandemic. The M group is classified into 9 subtypes based on phylogenetic differences(5). It is estimated that the Env amino acid sequence of the virus is nearly 35% different among the major subtypes and 20% different within a subtype (22). Subtype variability can impact HIV-1 transmission(24), drug resistance patterns(32) and disease progression(28,29). HIV-1 diversity has widely been viewed as one of the major challenges in the development of an effective vaccine (23,145,228). However, identification of sites of positively selected mutations (PSM) can help predict mutational patterns that point out to immunologically significant regions of the virus.

HIV-1 depends on the host to survive and replicate. The host HLA class I restricted CD8+ T cell responses shape viral diversity and evolution. Previous studies showed that HIV-1 escape pathways are predictable based on the HLA profile of the host population (229). PSMs are amino acid substitutions that arise as a result of immune selection pressure and allow the virus to evade host immune responses(129,230,231). Identifying these PSMs and analyzing their change over time at population level can provide important insights on HIV-1 evolution under host immune selection pressure. A useful approach for the identification and

characterization of PSMs is by defining the ratio of observed nonsynonymous (nucleotide changes that result in amino acid replacement) and synonymous (nucleotide changes that do not alter amino acids) mutations at individual codons across the viral genome(232). The d_N/d_S ratio (also called R:S for replacement vs. silent mutation) determines the direction of viral evolution. When d_N/d_S is equal to 1, amino acid changes are neither selected for or against, therefore, indicating "neutral" selection. A $d_N/d_S<1$ means that amino acid changes may be deleterious and therefore, are selected against; this is called "negative selection". Finally, when $d_N/d_S>1$, amino changes are favourable since they may increase viral fitness. This is referred to as "positive selection" (233). It has previously been shown that strong CTL responses contribute to the emergence of PSMs in HIV-1 Nef epitope encoding regions(234), therefore, providing evidence that immune selection pressure shapes the overall evolution of the virus. What remains unclear is to what extent do PSMs within Nef change overtime and how do Nef PSMs vary between major subtypes of HIV-1. Previous studies analyzing PSMs of Env and Gag proteins of HIV-1 have shown that frequency of PSMs in a subtype tends to stabilize overtime despite initial increase over several time periods(230,235). In this study PSMs in Nef of major HIV-1 subtypes and their changes over time were analyzed.

3.2 Hypothesis

1) The frequency of PSMs in HIV-1 Nef from subtypes A1, B, C and D remain stable overtime

2) PSMs in HIV-1 Nef motifs will be consistent among subtypes A1, B, C and D since functional motifs represent conserved regions of the virus.

3.3 Objectives

To identify and characterize PSMs in Nef sequences of four major subtypes of HIV-1 M group, A1, B, C and D

3.4 Methods

3.4.1 HIV-1 Nef sequences

The HIV Nef sequences were downloaded from the HIV Database of Los Alamos National Laboratory (LANL) (http://www.hiv.lanl.gov/content/index). A total of 4011 sequences from clades A1, B, C and D were examined. The sequences of each clade were divided into three time periods (prior to 1995, 1995-1999, after 2000). Table 3.1 shows the number of HIV-1 Nef sequences extracted from LANL and analyzed in this study, sorted by year and subtypes. Any duplicate sequences or those without labels of year or clade were removed from the analysis. The Nef sequences of each subgroup were aligned using ClustalW, followed by manual editing and converted to codon format using MEGA 5.0 (http://megasoftware.net/). All Sequences were aligned according to the HXB2CG, a subtype B HIV-1 isolate commonly used as reference strain for numbering HIV-1 sequences(22). PSMs were identified using Quasi analysis (details described in section 2.8.2).

3.4.2. Construction of Phylogenetic Tree:

The evolutionary history tree was constructed using the Maximum Likelihood method in the MEGA 5.0 software(236). All positions containing gaps and missing data in the aligned sequences were eliminated for tree construction.

3.4.3. Statistical Analysis:

Two-tailed contingency Fisher's exact t-test was conducted in GraphPad Prism v7.0a (GraphPad Software, La Jolla, CA) to determine association between PSM frequencies and regions of Nef within each subtype. *p*-values of <0.05 were considered significant.

3.5 Results

3.5.1 Distribution and nature of PSMs across 4 major HIV-1 subtypes

PSMs were predicted using Quasi analysis and mapped on the consensus sequence for each subtype (Figure 3.1). Quasi calculates the ratio of replacement to silent mutations (d_N/d_S) to define positively selected amino acids. An observed d_N/d_S ratio that is greater than the expected 1 is considered positively selected. Nef variable domain spans from amino acid 1-65 and the conserved core domain is from amino acid 66-206. In order to determine whether PSMs occur more frequently in one region compared to the other, the frequency of PSMs within each was calculated as the percentage of the number of PSMs out of the total amino acids in each region (65 in anchor and 141 in core domain). For subtype A1 and C there was a higher proportion of PSMs within the variable anchor domain, although the differences between PSMs for anchor and core domain were not significant for either subtype (p

= 0.1391 and 0.3697, Fisher's exact test, respectively). For subtype B and D the frequency of PSMs between anchor and core domains were similar with no significant differences (Figure 3.2). Functional constraints in the conserved region of Nef are likely the reason for fewer PSMs. Among the subtypes analyzed, some regions of Nef conserved core domain are completely free of PSMs. The conserved hydrophobic pocket residues F90 and W113 remain free of PSMs in all four subtypes examined (Figure 3.1). In the typically conserved proline rich motif PxxPxR (a.a. 72-77) of Nef (189,237-240) a PSM was observed in subtypes B and C with a lys→val substitution. However, there were no PSMs in this region of subtype A and D (Figure 3.1). HLA class I allele variations of infected host population could be accounted for the PSM patterns among different subtypes.

Table 3.2 shows the location of PSMs that take place within Nef's major functional motifs. Mutational patterns appear similar among motifs of four different subtypes. For example in all subtypes $DDPxxE_{(174\text{-}179)}$, involved in cell signaling, contains at least one PSM. Cell trafficking motif $DL_{(111\text{-}112)}$ does not have any PSM in any of the four subtypes. For all other motifs PSMs were observed in at least two of the four subtypes. This indicates the similarity of PSM patterns across clades. Additionally, conserved and non-conserved substitutions were assessed based on the shared chemical and physical features of individual amino acids**(241)**, such as size, hydrophobic/hydrophilic properties and amino acid charge. If the substituting and consensus amino acids shared any of these characteristics that substitution was considered conserved. For all four major subtypes the number of conserved amino

acid substitutions in Nef were higher than non-conserved substitutions ranging from 54-59% compared to 41-46% (Table 3.2).

Table 3.1. Summary of number of HIV-1 Nef sequences analyzed in the study, sorted by year and subtype.

Period	A1	B	C	D
before 1995	18	776	16	16
1995-1999	27	605	157	41
after 2000	116	1709	472	58
Total	161	3090	645	115

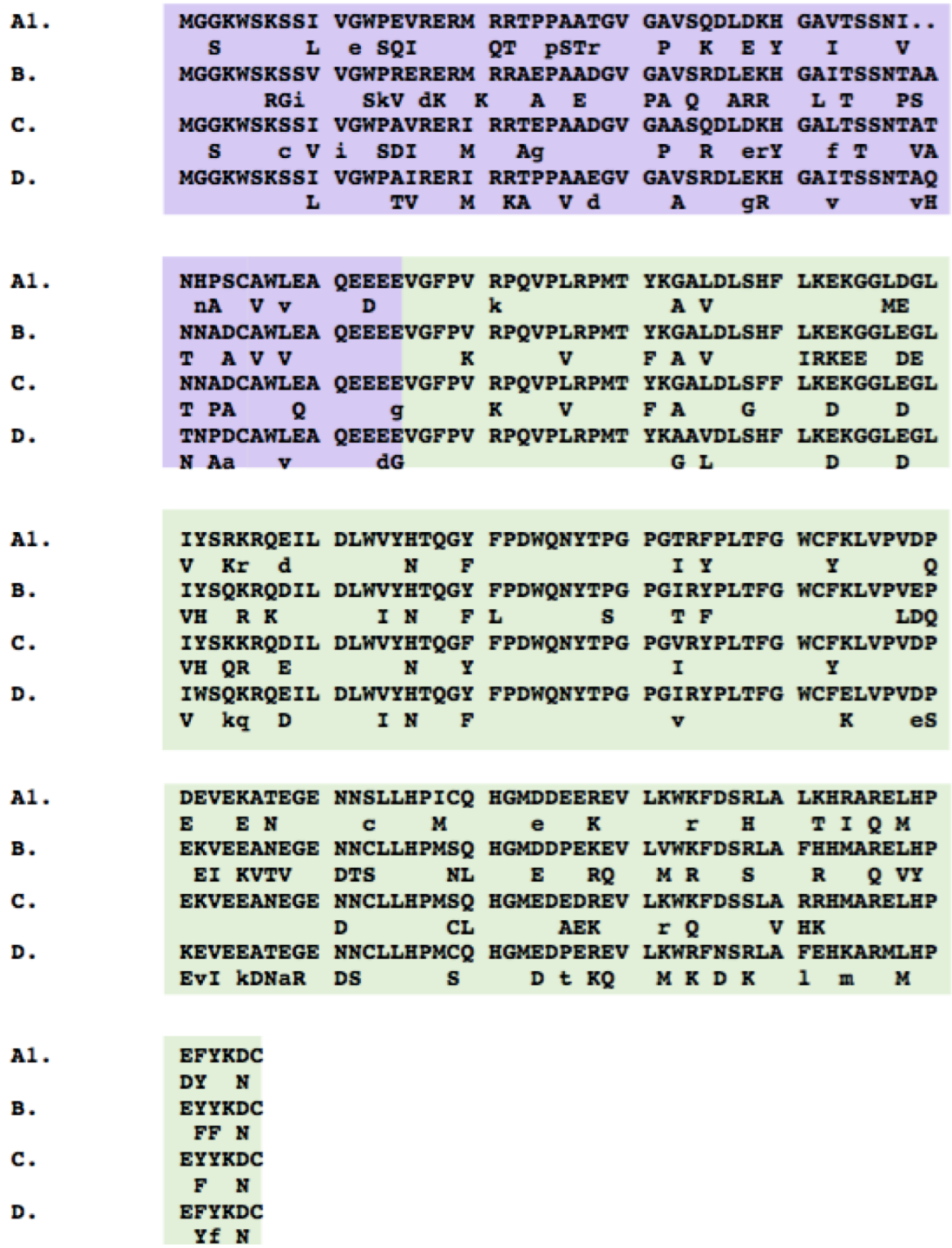


Figure 3.1. A map of positively selected mutations across HIV-1 Nef generated by QUASI for clades A1, B, C, and D sequences. The consensus sequence is shown as a single line of residues, in increments of 50, with positively selected amino acids underneath the consensus at each site in uppercase letters (lowercase letters indicate mutations where the neutral drift hypothesis is not ruled out). Numbering of amino acid positions in subtype sequence is relative to Nef HXB2CG strain. Dots (".") indicate deletions relative to HXB2CG sequences. Purple shaded area, variable anchor region (a.a. 1-65), green shaded area, conserved core domain (66-206).

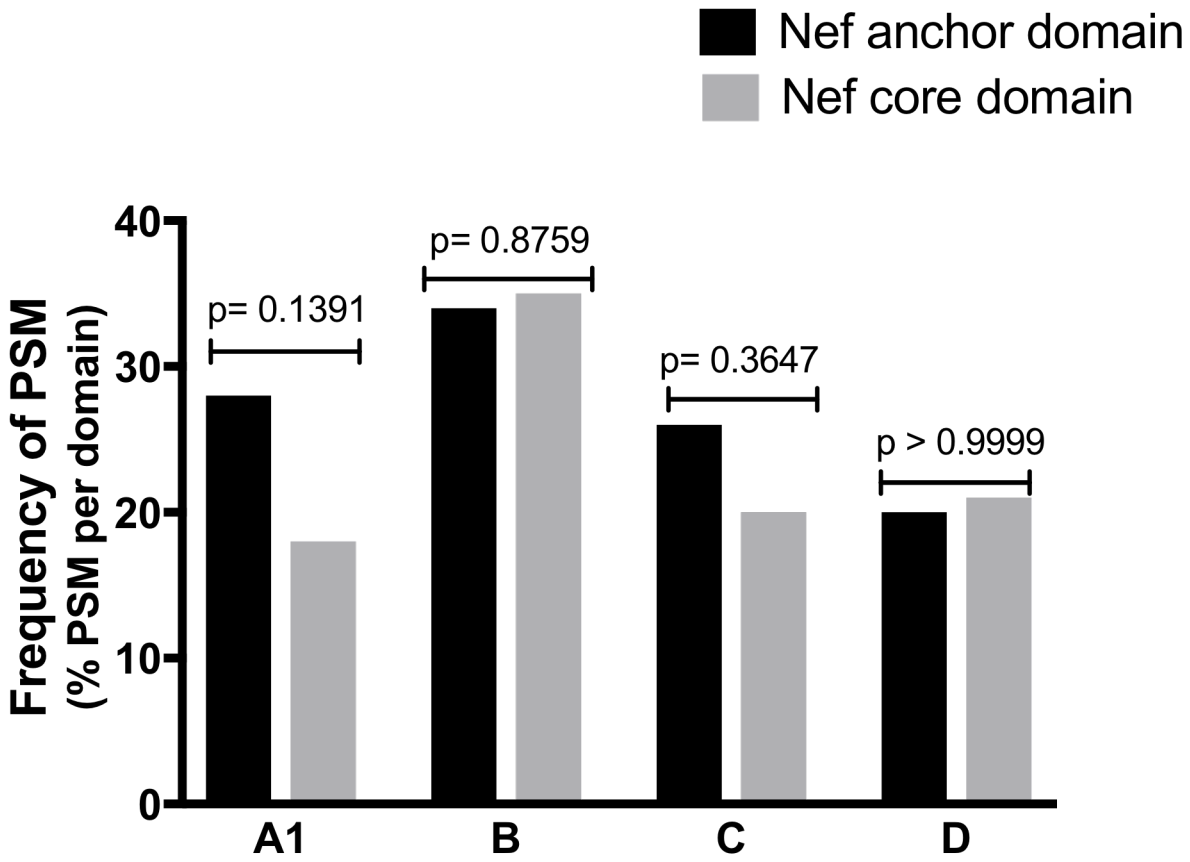


Figure 3.2. Comparison of PSM frequencies for the variable anchor domain and the conserved core domain of Nef in subtypes A1, B, C and D. Frequencies were calculated as the percentage of the number of PSMs out of the total amino acids in each region (65 in anchore and 141 in core domain). There is no significant difference between the frequencies of PSMs in the two domains of Nef in any of the subtypes (Fisher's exact test).

Table 3.2. Summary of PSMs within functional motifs of HIV-1 Nef in clades A1, B, C and D.

Region in Nef	Clade A1	Clade B	Clade C	Clade D
Protein Modification				
$MGxxxS_{(1-6)}$	+	-	+	-
$CAWLEA_{(55-60)}$	+	+	+	-
Signaling				
$PxxPxR_{(72-77)}$	-	+	+	-
$DDPxxE_{(174-179)}$	+	+	+	+
Trafficking				
$WL_{(57-58)}$	+	+	-	-
$DL_{(111-12)}$	-	-	-	-
$EEEE_{(62-65)}$	+	-	-	+
$FPD_{(121-123)}$	-	+	-	-
$EE_{(154155)}$	+	+	-	-
$ExxxLL_{(160-165)}$	-	+	+	+
$DD_{(174-175)}$	-	+	-	+

Numbers within brackets indicate the position of motif in Nef
"+" indicates presence of PSM; "-" indicates no PSM
Table was adapted from: Geyer et al. Structure-function relationships in HIV-1 Nef. (2001), 2(7), 580–585.

Table 3.3. Proportion of conserved and non-conserved amino acid substitutions in HIV-1 Nef clade A1, B, C and D. Conserved and non-conserved substitutions were based on similarities or differences in amino acid size, charge, and hydrophobicity or hydrophilicity.

Subtype	Total # of PSMs	# of Conserved mutation (%)	# of non-conserved mutations (%)
A1	34	20 (59)	14 (41)
B	50	27 (54)	23 (46)
C	40	22 (55)	18 (45)
D	35	20 (57)	15 (43)

3.5.2 Changes in frequency of PSMs overtime

To determine whether there is a relationship between subtype and PSM frequencies, the percentage of the number of mutation sites out of the total number of amino acids (206) in the consensus sequence of Nef for each subtype was calculated. The PSM frequencies ranged from 20.1 %, for subtype A1, and 35% for subtype B (Figure 3.3A). While subtypes A1, C and D have similar PSM frequencies, subtype B represents the highest PSM frequency in Nef. There was no association between subtypes and PSM frequencies (two-tailed Fisher's exact test). PSM frequencies at different time periods (pre 1995, 1995-1999 and after 2000) were calculated for each subtype. Frequency of PSMs increased overtime for subtypes A1, C and D, whereas in subtype B it remains constant (Figure 3.3 B). This difference may be due to the larger sample size of subtype B sequences compared to other subtypes.

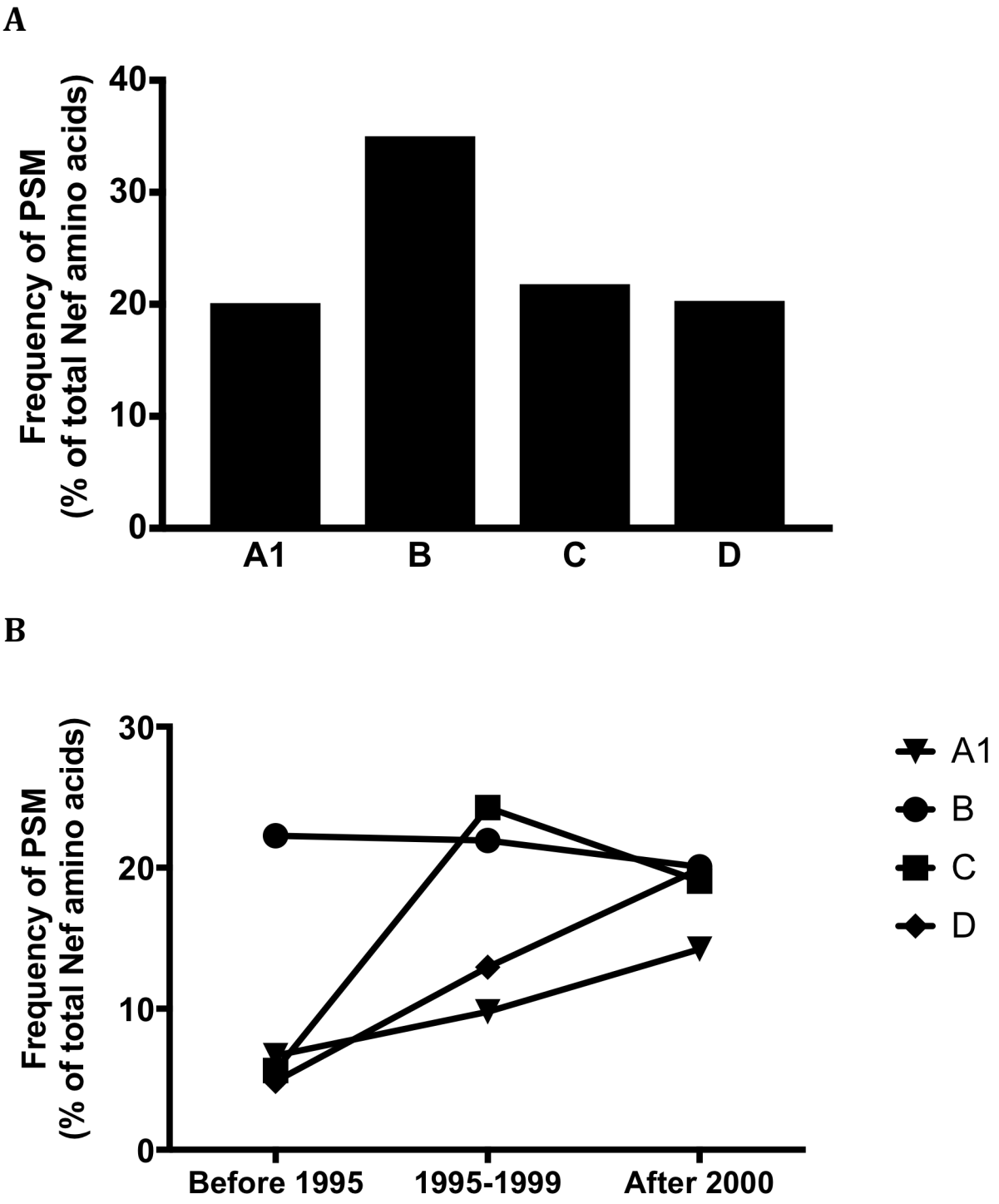


Figure 3.3. Subtype specific PSM frequencies of HIV-1 Nef A) Frequencies of PSMs in HIV-1 Nef sequences from subtypes A1, B, C, and D using consensus sequences of all time points. Frequencies were measured as the percentage of the number of PSMs out of the total number of amino acids for each subtype. **B)** Frequencies of PSMs over three time periods for each subtype are shown. An overall increase in PSM frequencies for the entire Nef sequence was observed over time for all subtypes except for subtype B.

3.5.3 Nef amino acid changes and mutational patterns over time

To examine Nef amino acid changes of the consensus sequence overtime among the four major subtypes, sequences were compared from three different time periods with those of ancestral and subtype consensus sequences obtained from Los Alamos Database (Figure 3.4). The differences between three time periods and Nef consensus or ancestral sequences are represented by closely related amino acids, most likely to preserve the specific function of Nef within a specific domain. The non-core region of Nef (residues 1-65) contains the highest number of variability among each time point. This is particularly true for subtypes A1 and B (Figure 3.4 A-B), whereas for subtypes C and D this appears to be more dispersed across the Nef protein (Figure 3.4 C-D).

Because host HLA class I restricted CD8+ T cell responses and viral functional constraint are two major factors shaping viral evolution, we expect that CD8+ T cell responses restricted by the same HLA class I alleles would have similar effect on PSMs of same functional motifs of Nef in different HIV-1 subtypes. At these sites HIV subtypes may become more similar to each other. To test this, we selected three major functional Nef motifs that contained PSMs in three or all subtypes (Table 2). These include: CAWLEA (protein modification), DDPxxE (cell signaling), and ExxxLL (cell trafficking). We conducted phylogentic analysis to determine the relationships between the sequences containing PSMs and the consensus of each time period. Maximum Likelihood trees were constructed using MEGA 5 software and the distances between sequences were examined. Three types of phylogenetic

relationships were observed. In the first type, two PSM containing sequences from the same subtype group and from two different time periods cluster together, but are separated from other sequences within the same subtype. The PSM containing sequences, DDEEKE (A1*, T1) and DDEEKE (A1*, T3) are clustered together on one branch of the phylogenetic tree, and are separated from other A1 sequences (Figure 3.5B). A similar pattern is seen in subtype C, PSM containing EDNCLL motifs from 1995-1999 and 2000 time periods (EDNCLL C*T2 and T*3), which are also clustered together (Figure 3.5C). A second type of phylogenetic relationship shows that sequences containing PSMs cluster with sequences of other subtypes. An example of this is observed in subtype B, PSM containing EDPEKE (B*. T1), which appears on the same branch as subtype D, EDPEKE (D*. T3)(Figure 3.5B). Finally, the third type of divergence is when PSM sequences cluster together with consensus sequence of the same subtype but different time period. For example, subtype C, PSM containing CAWLQA (C*. T2) sequence appears together on the same branch of the evolutionary tree as subtype C consensus sequence CAWLQA (C. T3) (Figure 3.5A). These patterns indicate that emergence of PSMs in subtypes can drive the evolution of the virus towards closely related sequences at both inter- and intra- subtype levels.

A) SUBTYPE A1 NEF

PERIOD/ Seq. Origin	NEF SEQUENCE	HXB2CG #
Nef HXB2.	MGGKWSKSSV IGWPTVRERM RRAEPAADRV GAASRDLEKH GAITSSNTAA	50
Cons. A1	MGGKWSKSSI VGWPEVRERM RRTPPAATGV GAVSQDLDKH GAVTSSNI..	
Anc. A1	MGGKWSKSSI VGWPEVRERM RRTPPAAKGV GAVSQDLDKH GAVTSSNTAA	
Pre-1995	MGGKWSKSSI VGWPEVRERM RRAPPAARGV GAVSQDLDKH GAVTSSNI..	
	:EQTRPTAAEGQAPTAAK (R28 & G29)	
1995-99	MGGKWSKSSI VGWPEVRERM RQT.PAPPGV GAVSQDLDKH GAVTSSNI..	
	:HPTPAAT (P28 & G29)	
POST 2000	MGGKWSKSSI VGWPEVRERI RRAAPAATGV GAVSQDLDKH GAVTSSNV..	
	:P (A23 & A24)	
	:PPAAT (T28 & G29)	
Nef HXB2.	TNAACAWLEA QEEEEVGFPV TPQVPLRPMT YKAAVDLSHF LKEKGGLEGL	100
Cons. A1	NHPSCVWLEA QEEEEVGFPV RPQVPLRPMT YKGALDLSHF LKEKGGLDGL	
Anc. A1	NNPGCAWLEA QEEEEVGFPV RPQVPLRPMT YKGAFDLSHF LKEKGGLDGL	
Pre-1995	NHPSCAWLEA QEEEEVGFPV RPQVPLRPMT YKGAVDLSHF LKEKGGLDGL	
1995-99	NHPSCVWLEA QEEEEVGFPV RPQVPLRPMT YKGALDLSHF LKEKGGLDGL	
POST 2000	NHPSCAWLEA QEEEEVGFPV RPQVPLRPMT YKGALDLSHF LKEKGGLDGL	
Nef HXB2.	IHSQRRQDIL DLWIYHTQGY FPDWQNYTPG PGVRYPLTFG WCYKLVPVEP	150
Cons. A1	IYSRKRQEIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVDP	
Anc. A1	IYSKKRQEIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVDP	
Pre-1995	IYSRKRQEIL DLWVYHTQGY FPDWQNYTPG PGTRFPLTFG WCFKLVPVDP	
1995-99	IYSRKRQEIL DLWVYNTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVDP	
POST 2000	IYSRKRQEIL DLWVYHTQGY FPDWQNYTPG PGIRFPLTFG WCFKLVPVDP	
Nef HXB2.	DKIEEANKGE NTSLLHPVSL HGMDDPEREV LEWRFDSRLA FHHVARELHP	200
Cons. A1	DEVEKATEGE NNSLLHPICQ HGMDDEEREV LKWKFDSRLA LKHRAQELHP	
Anc. A1	AEVEEATEGE NNSLLHPICQ HGMDDEEREV LMWKFDSRLA LKHRARELHP	
Pre-1995	DEVEKATEGE NNSLLHPICQ HGMDDEEKET LMWKFDSRLA LKHRA-ELHP	
1995-99	DEVEKATEGE NNSLLHPICQ HGMDDEERET LMWKFDSRLA LKHRARELHP	
POST 2000	DEVEKATEGE NNSLLHPICQ HGMDDEEREV LMWKFDSRLA LKHRARELHP	
Nef HXB2.	EYFKNC	206
Cons. A1	EYFKNC	
Anc. A1	EYFKNC	
Pre-1995	EYFKNC	
1995-99	EFYKDC	
POST 2000	EYFKNC	

B) SUBTYPE B NEF

PERIOD/ Seq. Origin	NEF SEQUENCE	HXB2CG #
Nef HXB2.	MGGKWSKSSV IGWPTVRERM RRAEPAADRV GAASRDLEKH GAITSSNTAA	50
Cons. B	MGGKWSKRSV VGWPTVRERM RRAEPAADGV GAVSRDLEKH GAITSSNTAA	
Anc. B	MGGKWSKSSM GGWPAVRERM KRAEPAADGV GAVSRDLEKH GAITSSNTAA	
Pre-1995	MGGKWSKRSV VGWPTVRERM RRAEPAAEGV GAVSRDLEKH GAITSSNTAA	
	:RGTATRTEPAADGVTD (E28 & G29)	
1995-99	MGGKWSKSSV VGWPTVRERM RRAEPAATGV GAVSRDLEKH GAITSSNTAT	
	:RRRTRSRAAANNLGEPAAD (T28 & G29)	
POST 2000	MGGKWSKRSV AGWPAVRERM RRTEPAARGV GAVSRDLEKH GAITSSNTAA	
	:RRRDKRREPPAAAGMEPAAE (R28 & G29)	
Nef HXB2.	TNAACAWLEA QEEEEVGFPV TPQVPLRPMT YKAAVDLSHF LKEKGGLEGL	100
Cons. B	NNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKGALDLSHF LKEKGGLEGL	
Anc. B	TNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKAALDLSHF LKEKGGLEGL	
Pre-1995	NNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKAALDLSHF LKEKGGLEGL	
1995-99	NNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKAALDLSHF LKEKGGLEGL	
POST 2000	NNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKGALDLSHF LKEKGGLEGL	
Nef HXB2.	IHSQRRQDIL DLWIYHTQGY FPDWQNYTPG PGVRYPLTFG WCYKLVPVEP	150
Cons. B	IYSQKRQDIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP	
Anc. B	IYSQKRQDIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP	
Pre-1995	IYSQKRQDIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP	
1995-99	IYSQKRQDIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP	
POST 2000	IYSQKRQDIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP	
Nef HXB2.	DKIEEANKGE NTSLLHPVSL HGMDDPEREV LEWRFDSRLA FHHVARELHP	200
Cons. B	EKVEEANEGE NNSLLHPMSL HGMDDPEREV LVWKFDSRLA FHHMARELHP	
Anc. B	EKVEEATEGE NNSLLHPMCQ HGMDDPEKEV LVWKFDSRLA FHHMARELHP	
Pre-1995	EKVEEANEGE NNSLLHPMSL HGMDDPEREV LVWKFDSRLA FHHMARELHP	
1995-99	EKVEEANEGE NNCLLHPMSQ HGMDDPEKEV LVWKFDSRLA FHHMARELHP	
POST 2000	EKVEEANEGE NNCLLHPMSQ HGMDDPEREV LVWKFDSRLA FHHVAREIHP	
Nef HXB2.	EYFKNC	206
Cons. B	EYYKDC	
Anc. B	EYYKDC	
Pre-1995	EYYKDC	
1995-99	EYYKDC	
POST 2000	EYYKDC	

C) SUBTYPE C NEF

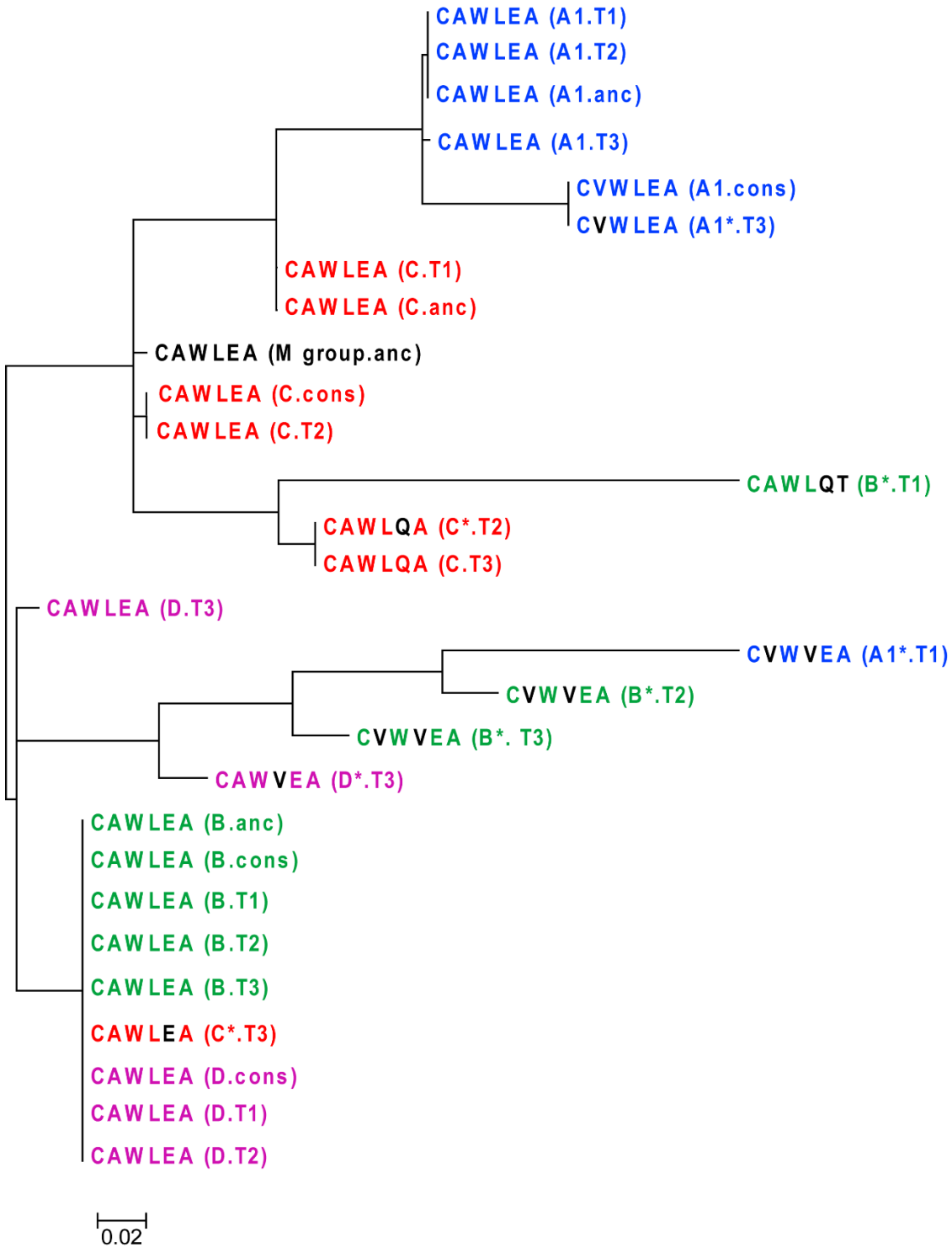
PERIOD/ Seq. Origin	NEF SEQUENCE	HXB2CG #
Nef HXB2.	MGGKWSKSSV IGWPTVRERM RRAEPAADRV GAASRDLEKH GAITSSNTAA	50
Cons. C	MGGKWSKRSV VGWPTVRERM RRAEPAADGV GAVSRDLEKH GAITSSNTAA	
Anc. C	MGGKWSKSSM GGWPAVRERM KRAEPAADGV GAVSRDLEKH GAITSSNTAA	
Pre-1995	MGGKWSKRSV VGWPTVRERM RRAEPAAEGV GAVSRDLEKH GAITSSNTAA	
	:RGTATRTEPAADGVTD (E28 & G29)	
1995-99	MGGKWSKSSV VGWPTVRERM RRAEPAATGV GAVSRDLEKH GAITSSNTAT	
	:RRRTRSRAAANNLGEPAAD (T28 & G29)	
POST 2000	MGGKWSKRSV AGWPAVRERM RRTEPAARGV GAVSRDLEKH GAITSSNTAA	
	:RRRDKRREPPAAAGMEPAAE (R28 & G29)	
Nef HXB2.	TNAACAWLEA QEEEEVGFPV TPQVPLRPMT YKAAVDLSHF LKEKGGLEGL	100
Cons. C	NNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKGALDLSHF LKEKGGLEGL	
Anc. C	TNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKAALDLSHF LKEKGGLEGL	
Pre-1995	NNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKAALDLSHF LKEKGGLEGL	
1995-99	NNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKAALDLSHF LKEKGGLEGL	
POST 2000	NNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKGALDLSHF LKEKGGLEGL	
Nef HXB2.	IHSQRRQDIL DLWIYHTQGY FPDWQNYTPG PGVRYPLTFG WCYKLVPVEP	150
Cons. C	IYSQKRQDIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP	
Anc. C	IYSQKRQDIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP	
Pre-1995	IYSQKRQDIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP	
1995-99	IYSQKRQDIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP	
POST 2000	IYSQKRQDIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP	
Nef HXB2.	DKIEEANKGE NTSLLHPVSL HGMDDPEREV LEWRFDSRLA FHHVARELHP	200
Cons. C	EKVEEANEGE NNSLLHPMSL HGMDDPEREV LVWKFDSRLA FHHMARELHP	
Anc. C	EKVEEATEGE NNSLLHPMCQ HGMDDPEKEV LVWKFDSRLA FHHMARELHP	
Pre-1995	EKVEEANEGE NNSLLHPMSL HGMDDPEREV LVWKFDSRLA FHHMARELHP	
1995-99	EKVEEANEGE NNCLLHPMSQ HGMDDPEKEV LVWKFDSRLA FHHMARELHP	
POST 2000	EKVEEANEGE NNCLLHPMSQ HGMDDPEREV LVWKFDSRLA FHHVAREIHP	
Nef HXB2.	EYFKNC	206
Cons. C	EYYKDC	
Anc. C	EYYKDC	
Pre-1995	EYYKDC	
1995-99	EYYKDC	
POST 2000	EYYKDC	

D) SUBTYPE D NEF

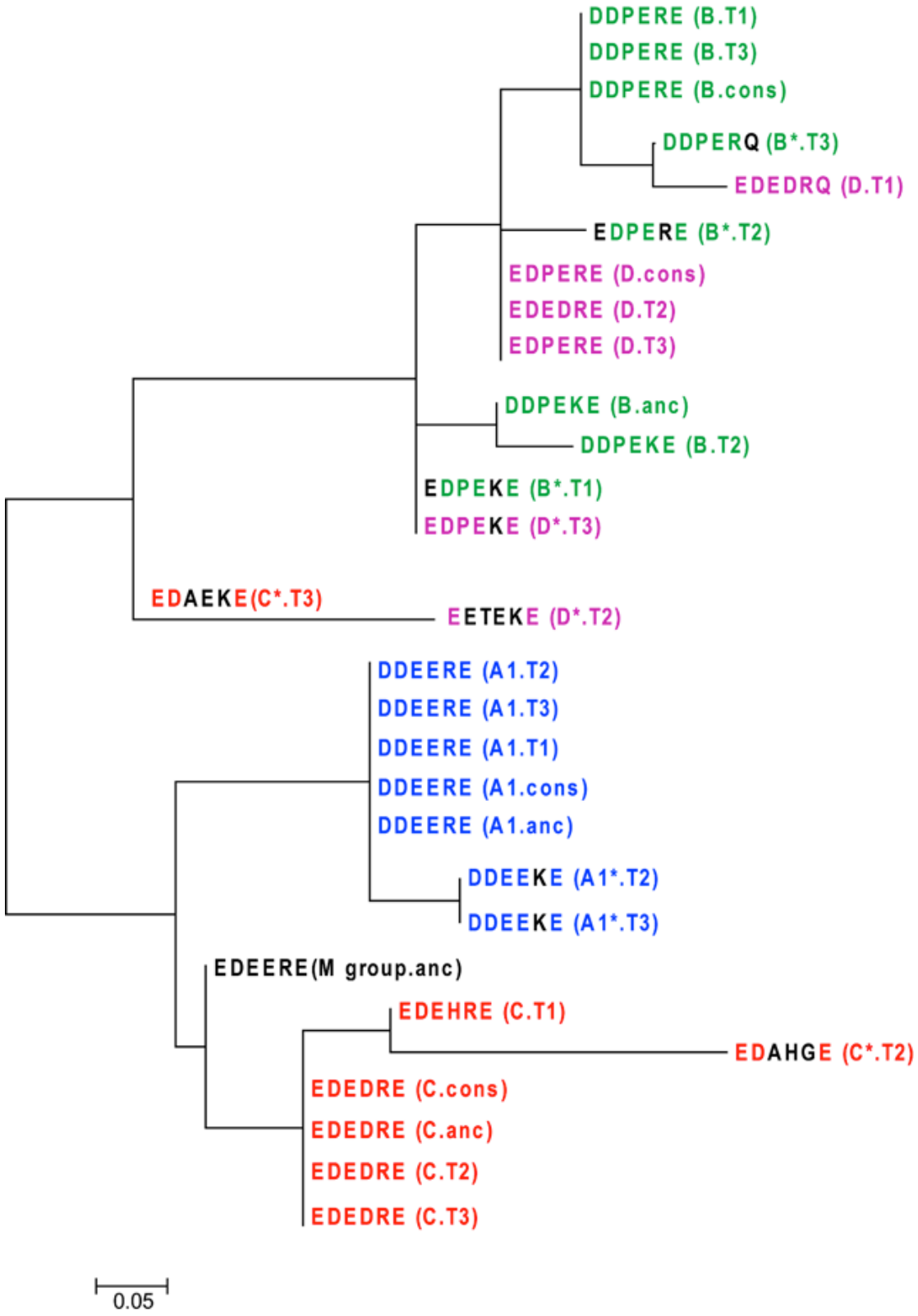
PERIOD/ Seq. Origin	NEF SEQUENCE	HXB2CG #
Nef HXB2.	MGGKWSKSSV IGWPTVRERM RRAEPAADRV GAASRDLEKH GAITSSNTAA	50
Cons. D	MGGKWSKSSI VGWPAIRERI RRTEPAADGV GAVSRDLEKH GAITSSNTAA	
Anc. B*	MGGKWSKSSM GGWPAVRERM KRAEPAADGV GAVSRDLEKH GAITSSNTAA	
Pre-1995	MGGKWSKSSI VGWPAIRERI RRTDAAAAGV GAVSRDLEKH GAITSSNTAS	
	:PRERRRPDPAADG (D24 & A25)	
1995-99	MGGKWSKSSI VGWPAIRERI RRTDAAAEGV GAVSRDLEKH GAITSSNTAH	
	:PARTEPAAEGVR (D24 & A25)	
POST 2000	MGGKWSKSSI VGWPAIRERI RRTDPAAEGV GAASRDLEKH GAITSSNTAH	
	:PAAERARIRSTAAEGE (D24 &P25)	
Nef HXB2.	TNAACAWLEA QEEEEVGFPV TPQVPLRPMT YKAAVDLSHF LKEKGGLEGL	100
Cons. D	TNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKAALDLSHF LKEKGGLEGL	
	:E (E63 & E64)	
Anc. B*	TNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKAALDLSHF LKEKGGLEGL	
Pre-1995	TNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKAAVDLSHF LKEKGGLEGL	
	:S (E63 & E64)	
1995-99	TNPDCAWLEA QEEEEVGFPV RPQVPLRPMT YKGAVDLSHF LKEKGGLEGL	
	:D (E63 & E64)	
POST 2000	TNPDCAWLEA QEEEEVGFPV RPQVPLRPMT YKAAVDLSHF LKEKGGLEGL	
	:E (E63 & E64)	
Nef HXB2.	IHSQRRQDIL DLWIYHTQGY FPDWQNYTPG PGVRYPLTFG WCYKLVPVEP	150
Cons. D	VWSQKRQEIL DLWVYNTQGF FPDWQNYTPG PGIRYPLTFG WCFELVPVDP	
Anc. B*	IYSQKRQDIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP	
Pre-1995	IWSKKRQEIL DLWVYNTQGF FPDWQNYTPG PGIRYPLTFG WCFELVPVDP	
1995-99	IWSQKRQEIL DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFELVPVDP	
POST 2000	IWSQKRQDIL DLWVYHTQGF FPDWQNYTPG PGIRYPLTFG WCFELVPVDP	
Nef HXB2.	DKIEEANKGE NTSLLHPVSL HGMDDPEREV LEWRFDSRLA FHHVARELHP	200
Cons. D	EEVEEATEGE NNCLLHPMCQ HGMEDPEREV LMWRFNSRLA FEHKARVLHP	
Anc. B*	EKVEEATEGE NNSLLHPMCQ HGMDDPEKEV LVWKFDSRLA FHHMARELHP	
Pre-1995	QEVEEATEGE NNCLLHPMCQ HGMEDPERQV LKWRFNSRLA FEHKARELHP	
1995-99	KEVEEDTEGE NNCLLHPMCQ HGMEDPEREV LKWRFNSRLA FEHKARVMHP	
POST 2000	EEVEEDTEGE NNCLLHPMCQ HGMEDPEREV LKWRFNSRLA FEHKARMLHP	
NEF HXB2.	EFYKDC	206
Cons. D	EYFKNC	
Anc. B*	EYYKDC	
Pre-1995	EFYKDC	
1995-99	EFYKDC	
POST 2000	EFYKDC	

Figure 3.4 Map of consensus sequence of HIV-1 Nef protein from three different time points compared to clade ancestral sequences. Subtypes A1, B, C and D (A-D). Numbering of amino acid positions in subtype sequences is relative to Nef HXB2CG. Insertions are shown using ":" directly under the amino acid after which insertion occurs followed by the insertion sequence and the amino acids before and after the sequences in brackets. Residues that change between time periods are shown in red. *For subtype D, no ancestral sequences are available on LANL; instead ancestral subtype B sequence was used since these two subtypes are most closely related to each other compared to other subtypes(242). Ancestral sequences obtained are based on the Complete Genome M-group Ancestral sequence and phylogenetic trees provided in LANL.

A


CAWLEA (A1.T1)
CAWLEA (A1.T2)
CAWLEA (A1.anc)
CAWLEA (A1.T3)
CVWLEA (A1.cons)
CVWLEA (A1*.T3)
CAWLEA (C.T1)
CAWLEA (C.anc)
CAWLEA (M group.anc)
CAWLEA (C.cons)
CAWLEA (C.T2)
CAWLQT (B*.T1)
CAWLQA (C*.T2)
CAWLQA (C.T3)
CAWLEA (D.T3)
CVWVEA (A1*.T1)
CVWVEA (B*.T2)
CVWVEA (B*. T3)
CAWVEA (D*.T3)
CAWLEA (B.anc)
CAWLEA (B.cons)
CAWLEA (B.T1)
CAWLEA (B.T2)
CAWLEA (B.T3)
CAWLEA (C*.T3)
CAWLEA (D.cons)
CAWLEA (D.T1)
CAWLEA (D.T2)
0.02

B


DDPERE (B.T1)
DDPERE (B.T3)
DDPERE (B.cons)
DDPERQ (B*.T3)
EDEDRQ (D.T1)
EDPERE (B*.T2)
EDPERE (D.cons)
EDEDRE (D.T2)
EDPERE (D.T3)
DDPEKE (B.anc)
DDPEKE (B.T2)
EDPEKE (B*.T1)
EDPEKE (D*.T3)
EDAEKE(C*.T3)
EETEKE (D*.T2)
DDEERE (A1.T2)
DDEERE (A1.T3)
DDEERE (A1.T1)
DDEERE (A1.cons)
DDEERE (A1.anc)
DDEEKE (A1*.T2)
DDEEKE (A1*.T3)
EDEERE(M group.anc)
EDEHRE (C.T1)
EDAHGE (C*.T2)
EDEDRE (C.cons)
EDEDRE (C.anc)
EDEDRE (C.T2)
EDEDRE (C.T3)
0.05

C

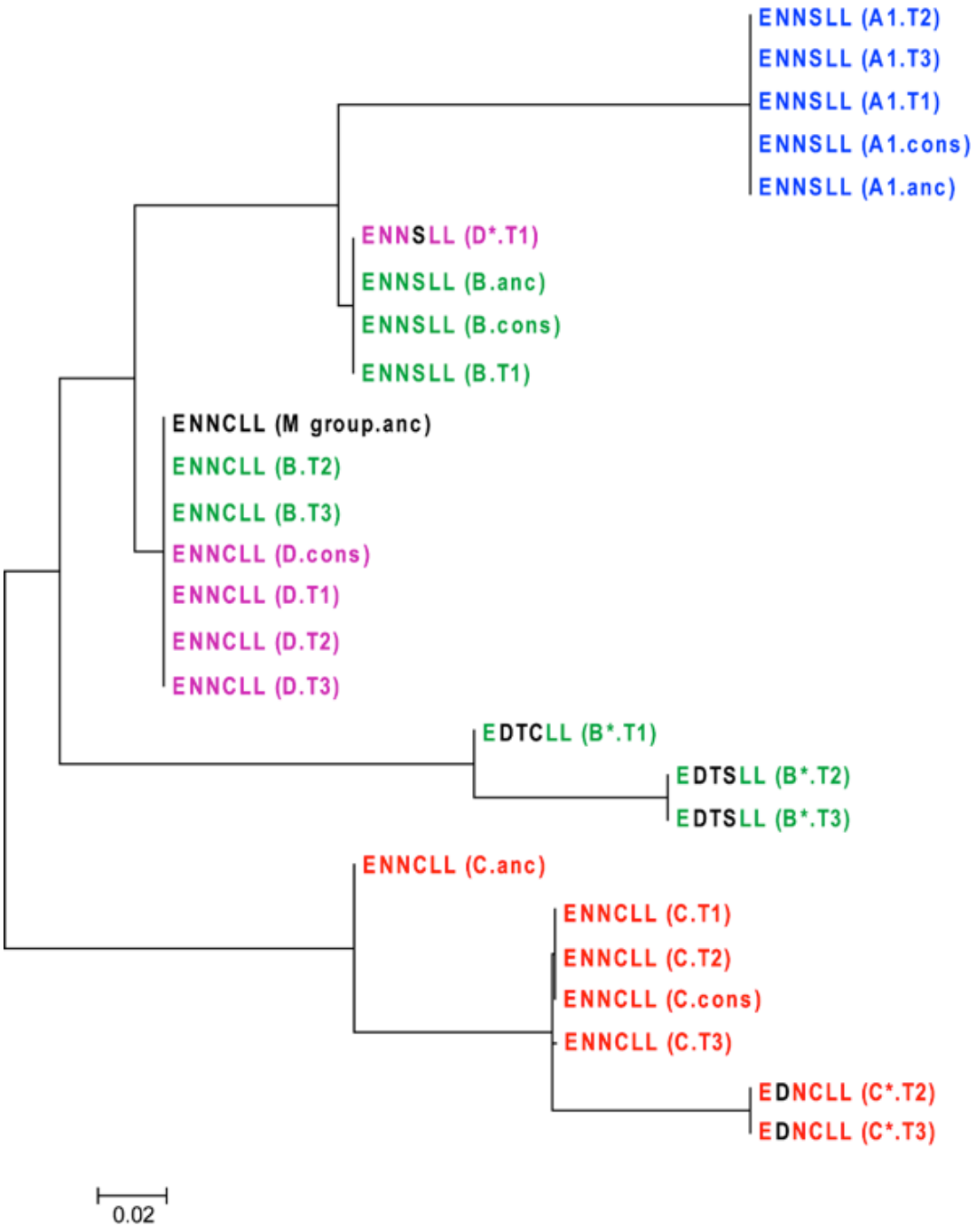


Figure 3.5. Evolutionary tree of HIV-1 Nef comparing sequences of three functional motifs CAWLEA (A), DDPxxE (B) and ExxxLL (C). Trees were constructed using the Maximum Likelihood method in MEGA5(243). Subtype A is shown in blue, B in green, C in red and D in pink. Each sequence is followed by subtype and time point (T1: pre-1995, T2: 1995-1999 and T3: post 2000) inside brackets. The asterisk sign, "*", indicates the sequence contains PSM, which are shown in black letters. The tree with the highest log likelihood (-82.2210, -125.5229 and -67.1081 for A, B and C, respectively) are shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

3.6 Discussion

In this study, quasi analysis was used to identify PSMs in HIV-1 Nef in four major M group subtypes, A1, B, C and D (Figure 3.1). QUASI is an effective selection mapping algorithm that identifies the ratio of replacement and silent mutations (d_N/d_S) at every codon (129). Several other methods are used to infer selective advantage of viral mutation, by determining the overabundance of replacement mutations compared to silent mutations at the protein level. However, the disadvantage of these methods is failing to accurately identify which replacement mutations are positively selected at individual codons. Additionally, combining all mutations together at the protein level may mask positively selected mutations within background or negatively selected mutations. Therefore, QUASI overcomes these disadvantages through its *per* codon analysis. The QUASI program is also capable of identification of PSMs in large number of sequences, unlike other software programs(129).

The data from this study showed that Nef contains mutations in both the variable and conserved regions (Figure 3.2). Although the distribution of PSMs in both regions was similar, many key motifs of Nef within the conserved region did not have any PSMs. This is consistent with previous findings that conserved regions of Nef are important for viral survival and functions and thus fewer PSMs were observed (240,244). Mutations within the conserved regions of Nef can affect its function and diminish viral replication capacity(245). The functional significance of proline rich domain (PxxPxR), hydrophobic residues F90, W113 and LL (a.a. 164-165) region, several motifs in which no or very few PSMs were observed, have been described in previous studies (246). The proline rich

domain is particularly crucial for Nef's interference with cell signaling pathways, leading to reduced surface expression of CD4 and HLA class I molecules, two of Nef's major functions that ultimately contribute to evasion from host immune responses(207,247) . The observed substitutions in PxxPxR (a.a. 72-77) motif in subtypes B and C may have implications for functions and interaction with signaling proteins. In a study it was demonstrated that among HLA B*35:01 patients an arginine to threonine substitution in PxxPxR was linked with impaired down-regulation of CCR5 from cell surface (245). In SIV models of infection, mutations in this Nef motif were linked to impaired activation of Pak2 signaling molecule, resulting in altered T cell activation(192,197,239,240). Therefore, regions of Nef that play a functional significance may remain free of PSM to preserve key functions. These results are also consistent with our previous findings where we identified PSMs in the Env protein of HIV-1(230). There, it was also shown that evolutionary conserved regions of Env, which are critical for the survival of the virus, had fewer PSMs or were completely free of them. The overall variable distribution of PSMs across Nef indicates that some regions are under greater selection pressure compared to others and functional constraints may not accommodate PSMs if costly to survival.

In this study, the frequencies of PSMs in Nef over three consecutive time periods in the four major subtypes were examined. Our data showed that the frequencies of PSMs in HIV-1 Nef increase over time in subtypes A1, C and D but not in subtype B. The pattern observed in subtypes A1, C and D may be due to a small pool of sequences available for the final analysis (A1,161; C, 647; and D, 115), which likely does not cover the HLA genetic diversity of the infected host population. Given the evidence from numerous studies which have

shown that polymorphisms in HIV-1 are associated with specific host HLA class I alleles (98,135,167,171,190,229), we speculate that the observed consistent PSM frequencies in subtype B is due to a large sample size of sequences which better represent HIV's adaptation overtime to the HLA-restricted immune responses at the population level. In fact a study of RT polymorphisms showed that viral sequences of a large cohort of HLA-typed individuals could reveal escape mutations reliably(167) allowing to identify selection pressure at the population level based on the HLA footprint (159). Therefore, we would expect a similar pattern as subtype B in the PSM frequencies overtime for the other three subtypes if a larger number of sequences were analyzed.

It was observed that for all subtypes, the amino acid variations that exist between time points are closely related either by amino acid size, function and or structure, most likely to preserve the functions of Nef within a specific domain (Figure 3.4). For each subtype the changes that occur within the three time points represent a dynamic that exists between consensus and the ancestral sequence, where a group of amino acids with similar properties tend to replace one another in each time periods. In some cases, the substitutions are amino acids that revert to the ancestral sequence, further showing the tendency to preserve function by replacement with a closely related amino acid. An example of this is shown in the case of subtype A1, where alanine (a.a. 56) is the dominant residue in sequences before 1995, changes to valine in 1995-2000 sequences and then reverts to alanine in the latest reported sequences. Alanine is also the dominant amino acid in the ancestral sequence at this residue. Therefore, the observations of this study seem to suggest that HIV-1 Nef amino acid changes reflect a preference for the ancestral sequence

and minimum deviation between closely related amino acids. This demonstrates that despite HIV-1 diversity changes are limited by the function and fitness of virus.

Since HIV depends on the host to survive, its genetic variation largely depends on the host immune selection pressure, which ultimately determines the evolutionary pathways of the virus. Host immune responses such as cytotoxic CD8+ T cells select for escape mutations within the viral sequence but in order for the virus to survive it can only accommodate amino acid replacement mutations that lead to minimal fitness cost and loss of function(178,179,212). We speculated that within each subtype, PSMs from various time points result in similar sequences since the HLA allele frequencies at the population level remain constant overtime (248). The evolutionary analysis of our study indicates that PSM containing sequences tend to cluster together when examining three of the major Nef motifs: $CAWLEA_{(55\text{-}60)}$, $ExxxLL_{(160\text{-}165)}$ and $DDPxxE_{(174\text{-}179)}$ (Figure 3.5A-C). In some cases, PSM sequences of multiple subtypes appear closer to each other than they do to the consensus sequences of the same subtype (e.g. CVWVEA A1*T1, B*T3 and D* T3, Figure 3.5A). These inter and intra-subtype similarities of PSM sequences observed in Nef demonstrate that host immune selection pressure restricted by the same HLA class I alleles (in different populations) force the virus to mutate to more closely related sequences. Because HIV depends on the host to survive, the constant host immune selection pressures shape and restrict viral diversity and evolution.

In addition to CD8+ T cells, multiple other host factors play a role in selecting viral mutations that drive the evolution of the virus. Natural killer (NK) cells, Antibody-

dependent cellular cytotoxicity (ADCC) and cytotoxic CD4+ T cells all exert immune selection pressure on the virus(249), which may also contribute to the emergence of PSMs. While $CAWLEA_{(55-60)}$ and $DDPxxE_{(174-179)}$ are restricted by several HLA alleles in different populations(250-254), $ExxxLL_{(160-165)}$ appears to not be targeted by any HLA groups, as reported in LANL. However, this motif is targeted by monoclonal antibodies (mAbs) (255) that may be responsible for the observed PSMs in this region in subtypes B, C and D (Table 3.2). The flanking region of this motif is targeted by HLA class II DRB1*0101, DRB1*0401, DRB1*0701, DRB1*1101, DRB1*1302, DRB1*1501, DRB5*0101 (256,257) and HLA class I B*35 and B*54:01 (258,259). However, the fact that HLA Class I and II alleles do not recognize the major component of this motif may be the result of its 3D position within Nef's protein structure. $ExxxLL_{(160-165)}$ lies within the anti-parallel beta-sheet conserved core domain of Nef,(260), which may partially determine whether it could be processed and presented to HLA class I molecules as an epitope (261).

In summary, Quasi analysis was used to map and determine PSMs in HIV-1 Nef sequences from four major global subtypes, A1, B, C and D obtained from LANL. We have found that both conserved and variable regions of Nef contain PSMs, however certain sites within the conserved region, such as the proline rich domain, which play a crucial role in major Nef function, remain free of PSMs in various subtypes. This study showed that PSMs in Nef is relatively stable when a large viral population is analyzed, such as subtype B, perhaps due to stability of HLA allele frequencies of the host population. However, the number of sequences obtained from LANL for subtypes A1, C and D were quite small and limited our analysis. Obviously, further analysis is needed to determine the frequencies of mutations

within these subtypes and to accurately determine changes over time. We have also shown that PSMs, likely driven by HLA class I restricted CD8 T cell responses, in Nef drivee the evolving virus of different subtypes to be more similar. Therefore, the global trend in viral mutation and host population genetic factors need to be considered and evaluated for the development of an effective HIV-1 vaccine strategy.

3.7 Limitations and opportunities

Although HIV-1 Nef sequences on LANL public database are numerous, host genetic factors, such as HLA alleles of patients from whom sequences are obtained, are not provided. Therefore, one disadvantage of this study is that the association of PSMs identified using Quasi and specific HLA alleles could not be validated. Additionally, the number of observations for subtype B and C were substantially larger than those of subtypes A1 and D, which could skew the results of the study. A larger number of sequences will likely show a more realistic representation of PSM changes overtime at the population level, therefore, for subtypes with low sequence availabilities, future comprehensive evolutionary analyses are needed to draw better conclusions. Nonetheless, the identification of regions of the virus that experience selective pressure and undergo PSM, is important in understanding the interactions between the host and the virus that can help in defining therapeutic strategies.

4. Positively Selected Mutations in Subtype A1 HIV-1 Nef Protein in a Kenyan Population

4.1 Rationale

HIV-1 rapidly adapts to the host immune selective pressure by undergoing mutations that may result in immune evasion(140,153,262,263). While some mutations allow the virus to escape CTL activity, other escape mutations are associated with viral fitness cost(113,142,178,212). HLA-restricted CTL responses are a crucial aspect of cellular immunity against HIV-1 during both acute and chronic phase of the infection (141,149,264). Several CTL targets of HIV-1, that are associated with reduced viral fitness and overall immune control have previously been described(113,265,266). Multiple studies have identified HLA-associated polymorphisms (HLA-AP) in different HIV-1 genes within several distinct populations around the world (143,180,267,268) . However, HIV-1 Nef escape mutations in subtype A1 from a treatment-naïve large cohort population, with long-term CD4 count profiles have not yet been characterized. In this study PSMs in subtype A1 HIV-1 Nef sequences are identified and correlated with HLA class I alleles and longitudinal CD4 counts in a Kenyan female sex-worker population. Association of PSM with disease progression data and host HLA profiles can help identify protective and non-protective immune targets within Nef.

4.2 Hypothesis

HIV-1 Nef contains immunologically relevant T cell epitopes with mutations that are beneficial and detrimental for the host. These epitopes can be identified and characterized through analysis of the interactions between host and virus.

4.3 Objective

To characterize PSMs in HIV-1 Nef protein using Quasi analysis and to determine their association with clinical outcomes of disease (CD4 count) and patient HLA class I profile

4.4 Methods

4.4.1 Study cohort

Genomic DNA samples of 508 HIV-1 positive ART-naïve female sex-workers from the Nairobi Majengo cohort, obtained between 1980-2003, were used for amplification using PCR (detailed description in section 2.7.1). By sequencing virus from treatment-naÏve patients, possible drug driven mutations were avoided. CD4 counts were determined using Becton Dickinson Tritest reagent for all women in the cohort since 1990 and tracked on a biannual basis (section 2.4).

4.4.2 Statistical Analysis

HLA class I association with Nef PSMs- Fisher's exact test was used to analyze the number of patients with and without each HLA allele and the amino acid sequence with and without each PSM. *p*- value of <0.05 was considered significant. Statistical analyses were conducted in SPSS version 13.0. For the multiple comparisons of PSM and alleles, the classic one-stage method of False Discovery Rate (FDR) was used.

CD4 count decline and association with Nef PSMs- For each PSM individually, Kaplan-Meier survival analysis was conducted to determine the probability of CD4 count above 200 cells/mm^3 of blood over the number of days patients status was HIV+. *p*-values <0.05 were considered statistically significant. Statistical analyses were conducted in SPSS version 13.0.

PSM frequency comparison- Two-tailed contingency Fisher's exact t-test was conducted in GraphPad Prism v7.0a (GraphPad Software, La Jolla, CA) and *p*-values of <0.05 was considered significant.

4.5 Results

4.5.1 Location and Distribution of PSMs in HIV-1 Nef Protein

In the previous chapter, Nef PSMs of subtype A1 from the LANL public database were identified. However, sequences extracted from LANL may be from patients receiving ART, making it difficult to conclude whether PSMs are indeed driven by host immune selection pressure. In this chapter, we only used HIV+ treatment naïve samples to accurately determine PSMs across the Nef protein that could be selected by HLA-restricted CD8+ T cells. Of the 508 HIV+ treatment naïve samples amplified, 326 belonged to subtype A1 and PSMs were predicted using Quasi analysis (129). PSM sites were mapped on the Nef consensus sequence (Figure 4.1). There were a total of 48 amino acid substitutions identified in Nef in comparison to 43 identified for subtype A1 from LANL sequences. Overall, PSMs were not uniformly distributed, with the variable anchor region of Nef,

residues 1-65, displaying a higher frequency of PSMs than the conserved core region, although the difference was not significant (Fisher's exact test) (Figure 4.2).

As with LANL subtype A1 sequences, analysis of Nef amplified from Kenyan samples also showed that some regions of the core domain (amino acids 66-206) are completely free of PSMs (Table 4.1). For example, the proline rich motif, PxxPxR, did not have any PSMS in subtype A1 Nef sequences of the study cohort, similar to the sequences obtained from LANL (Table 3.2). Table 4.1 shows all major Nef functional motifs and the presence or absence of PSMs within each, with 50% (6/12) of motifs containing PSMs. The varying distribution of PSMs across Nef sequence suggest that selection pressure is different at each region, which can be attributed to the function of each motif or differential recognition by host HLA.

SUBTYPE A1 KENYAN NEF SEQUENCE **HXB2CG #**

```
MGGKWSKSSI VGWPEVRERM RRA.PAAAGV GAVSQDLDKH GAVTSSNI..      50
  S      K      I       PT  SP    RE     A Y R I    V

NHPSCVWLEA QEEEEVGFPV RPQVPLRPMT YKGAFDLSHF LKEKGGLDGL      100
 N   A       D   S    k           F A               ER

IYSRKRQEIL DLWVYHTQGY FPDWQNYTPG PGTRYPLTFG WCFKLVPVDP      150
V  K            N               R  EIF            E   Q

DEVEKATEGE NNSLLHPICQ HGMDDEEREV LKWKFDSRLA LKHRAQELHP      200
E    DN R    C    M    R     K    M R   H    R I R M
                                      S
EFYKDC                                                      206
 Y  N
```

Figure 4.1. **Map of PSMs in Nef protein of HIV-1 identified by Quasi analysis from 326 Kenyan subtype A1 sequences**. The consensus sequence is shown as a single line of residues, in increments of 50, with positively selected amino acids underneath the consensus at each site. Numbering of amino acid positions in the consensus sequences is relative to Nef HXB2CG strain. Dots (".") indicate deletions relative to HXB2CG sequences. PSMs were identified for 48 residues. R188 can be replaced by two PSMs: R188H or R188S. All other amino acids had one PSM replacement.

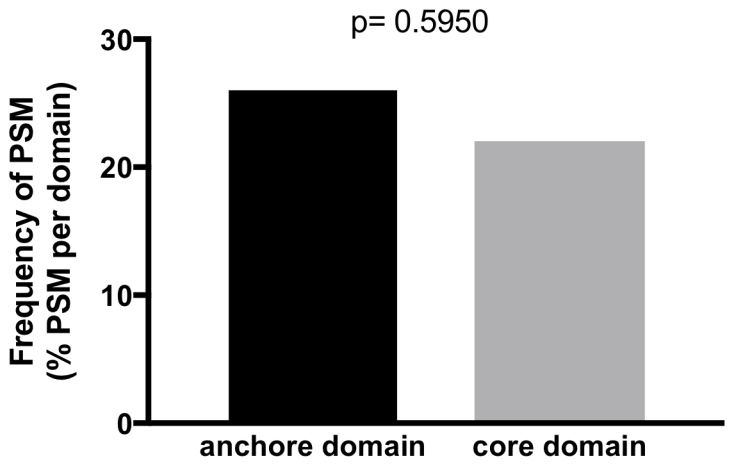


Figure 4.2. Comparison of the percent frequency of PSMs within the variable region (residue 1-65) and the conserved core region (66-206) of subtype A1 HIV-1 Nef protein from a Kenyan cohort. Frequencies were calculated as the percentage of the number of PSMs out of the total amino acids in each region (65 in anchor and 141 in core domain). There is no significant difference in the frequency of PSMs in the two domains of Nef (p= 0.5950 Fisher's exact test).

Table 4.1. Characteristics of PSMs in well-defined functional motifs of HIV-1 Nef of subtype A1 sequences from treatment naïve HIV+ female sex worker cohort in Nairobi, Kenya.

Nef Functional Motifs	Subtype A1 PSMs	PSM position (HXB2CG)	amino acid substitution
Protein Modification			
$MGxxxS_{(1-6)}$	Yes	3	G->S
$CAWLEA_{(55-60)}$	Yes	56	V->A
Cell Signalling			
$PxxPxR_{(72-77)}$	--- (conserved)	---	---
$RR_{(112)}$	---(conserved)	---	---
$DDPxxE_{(174-179)}$	Yes	178	R->K
Cell Trafficking			
$WL_{(57-58)}$	--- (conserved)	---	---
$LL_{(111-112)}$	--- (conserved)	---	---
$EEEE_{(62-65)}$	Yes	63	E->D
$FPD_{(121-123)}$	--- (conserved)	---	---
$EE_{(154-155)}$	--- (conserved)	---	---
$ExxxLL_{(160-165)}$	Yes	163	S->C
$DD_{(174-175)}$	--- (conserved)	---	---

Numbers within brackets indicate the position of motif in Nef
Table was adapted from: Geyer et al. Structure-function relationships in HIV-1 Nef. (2001), 2(7), 580–585

4.5.2 Effect of PSMs on the rate of CD4 decline

CD4 T cell count is an important indicator of disease state in HIV-1 infection. In an uninfected individual the normal CD4 count is approximately 500-1600 cell per µl of blood(47). A person infected with HIV-1 not receiving treatment will gradually lose healthy CD4 T cells as they become infected with the virus; CD4 T cell counts below 200 is indicative of disease state, AIDS(47,269). We hypothesized that PSMs and consensus amino acids in Nef will associate differently with the rate of disease progression. To test this, Kaplan-Meier survival analysis was used to compare CD4 decline over time for PSMs and consensus amino acids for all 48 sites where a PSM was identified, in patients whose *nef* gene was sequenced (Figure 4.3; only significant associations are shown). Two distinct types of correlations between PSMs and disease progression rates were observed; 3 PSMs were associated with faster CD4 decline in comparison to the consensus amino acid (E63D p = 0.028, log rank= 4.799; I101V p= 0.003, log rank=8.667 and I168M p=and 0.042, log rank=4.150) (Figure 4.3 A-C) and 2 PSMs were associated with slower CD4 decline overtime compared to the consensus amino acid (H116N p= 0.00011, log rank=14.891 and K182M p=0.03, log rank=4.753) (Figure 4.3 D-E).

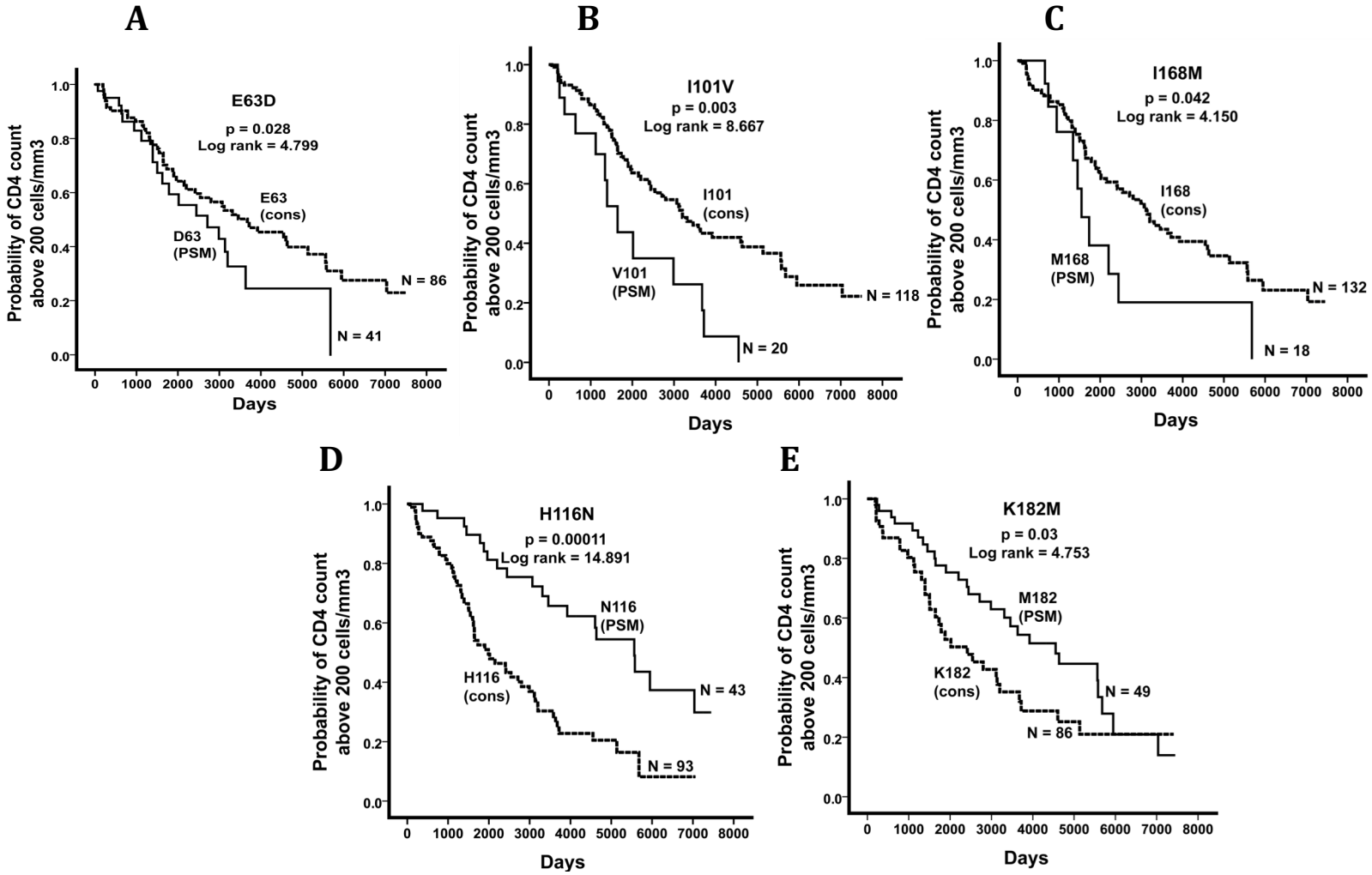


Figure 4.3. Kaplan-Meier survival analysis of the rate of CD4+ T cell decline with PSMs. CD4 decline was determined chronologically from the time of enrollment as the probablity of counts remaining above or equal to 200 cell/mm^3. E63D, I101V, I168M are correlated with faster disease progression (detrimental epitopes) (A-C), and H116N and K182M are correlated with slower disease progression (beneficial epitopes) (D and E). Solid lines are PSM and dotted lines are consensus amino acids.

4.5.3 HLA Class I correlation in the Pumwani Sex-worker cohort with HIV-1 Nef PSMs

We used Fisher's Exact test to determine the correlation between the five PSMs, that were associated with faster or slower CD4 decline, and all the HLA alleles of the patients in the Pumwani sex-worker cohort. E63D PSM is correlated with Class I alleles A*2301 and B*3501 (p= 0.002 and 0.003, respectively), I101V correlates with HLA-A*0201 and HLA-B*5802 (p= 0.028 and 0.048, respectively) and I168M correlated with HLA-A*2301, B*0702, B*5101 and C*1601 (p= 0.002, 0.006, 0.037 and 0.045, respectively). Of the PSMs that were associated with slow CD 4 decline, H116N correlated with HLA-A*0201, A*3002, B*5703 and C*0201 (p= 0.021, 0.026, 0.00038 and 0.011, respectively) and K182M was correlated with HLA-A*34 and C*0602 (p = 0.026 and 0.037, respectively) (Table 4.2). For the correction of multiple comparisons a classic FDR test was used to confirm p-values.

Table 4.2. Correlation of HLA class I alleles with PSMs in HIV-1 Subtype A1 Nef. The classic one-stage FDR method was used for the multiple comparisons of HLA alleles with PSMs.

Nef PSM	Association with CD4 decline (compared with consensus a.a.)	HLA class I allele	Fisher's Exact test *p*-value	FDR corrected *q*-value
E63D	Faster	A*2301	0.002	0.009
		B*350101	0.003	0.011
I101V	Faster	A*0201	0.028	0.039
		B*5801	0.048	0.048
I168M	Faster	A*2301	0.002	0.009
		B*070201	0.006	0.017
		B*510101	0.037	0.043
		Cw*160101	0.045	0.048
H116N	Slower	A*0201	0.021	0.039
		A*3002	0.026	0.039
		B*570301	0.00038	0.005
		Cw*0210	0.011	0.026
K182M	Slower	A*34	0.026	0.039
		Cw*0602	0.037	0.043

4.6 Discussion

Early *in-vitro* studies of CTL-specific escape mutations in HIV-1 suggest that alterations in Nef are generally not favorable and mainly constrained(215), arguing that Nef's two most vital functions, MHC and CD4 downregulation from cell surface, are essential in protecting the newly infected cell from the host immune response(270). Other studies have shown that Nef is a heavily targeted region of the virus and that mutations in Nef arise very rapidly in the early phase of the infection(229). Using Quasi analysis, we have demonstrated that PSMs in the Nef protein of subtype A1 HIV-1 from a cohort of women in Nairobi, Kenya, are highly frequent (Figure 4.1). We have also shown in this study that Nef contains PSMs in both the variable (anchor) and the conserved (core domain) region (Figure 4.1 and 4.2). We have demonstrated that several PSMs in Nef occur in the central region (a.a. 66-148) and carboxyl terminal region of Nef (a.a. 182-206), which have previously been reported as targets for CTL responses(271-273). Overall, we show that the PSMs in Nef can occur in both conserved and variable region of the protein. Therefore, our findings that the Nef protein accommodates amino acid substitutions at multiple sites contradicts with early studies that argue Nef mutations are constrained(215). There are likely several reasons for this. First, Nef has been shown to be targeted by multiple HLA class I alleles(229); therefore, it is possible that the many observed mutations are in the flanking regions of targeted epitopes. One study for example showed that three specific PSMs (R69K, A81G, and H87R) occur in the flanking region of a HLA-B35 epitope in Nef (274), which houses the functionally significant PxxPxR domain necessary for viral entry and cell signaling pathways. Any changes to this motif could potentially limit the virus replication capacity or fitness. Therefore, HLA-restricted CTL mutations in the flanking

regions of the B35 epitope (eg. R69K and A81G), instead of those within it, can help maintain Nef function(274).

Another explanation for the high number of PSMs present in Nef is the emergence of secondary mutations that compensate for the loss of fitness in another region of Nef (180,275,276). We also speculate that PSMs within Nef, although frequent (48/206 amino acids, Figure 4.1), may not all result in functional consequences for the virus, due to similar chemical and physical properties between the consensus and PSM amino acids, because new variants that cannot support viral function are unlikely to survive(178,277). Longitudinal studies have revealed that mutations in Nef are accumulated early on during the acute phase of the infection(278). Nef is the first protein to be expressed in the viral life cycle and is immediately targeted by CD8+ T cells (279). Therefore, we speculate that the high number of PSMs observed in Nef may be the result of accumulated mutations overtime in the study enrollees. PSMs are likely to arise as a result of HLA-restricted CTL responses(279). This is supported by findings from other groups that have shown Nef, compared to all other regulatory proteins of HIV, has the highest number of HLA-Class I associated polymorphisms, in one case up to 40% (229).

In this study, we have also examined the association between PSMs in HIV-1 Nef and disease progression in the infected hosts. Studies of CTL-driven mutations in HIV have analyzed the effects of mutations in the major genes of the virus, mainly Gag and Pol(178,280-282). While some studies have also surveyed the impact of mutations in Nef

protein, they did not include long-term treatment-naïve patient samples (190)or have included recombinant Nef sequences(280) . Having 326 subtype A1 Nef sequences from an ARV naïve Kenyan population of a FSW cohort and their longitudinal CD4 count data, we were able to analyze the associations of all the identified PSMs in Nef with rates of disease progression (Figure 4.3 shows significant Kaplan-Meier estimates). We show that PSMs have different impact on patient clinical outcomes. Some are associated with faster CD4 decline in comparison to their consensus counterpart (E63D, I101V, I168M), while others are associated with slower CD4 decline (H116N and K182M). Therefore, PSMs can be categorized into "*detrimental*" and "*beneficial*" groups according to their associated clinical outcomes manifested in the host. It may at first seem counterintuitive that "beneficial" mutations are positively selected and it raises the following question: how can such mutations associate with slower CD4 decline, presumably through limiting viral fitness, be positively selected for when there is a negative consequence for the virus itself? To answer this question, it is worthwhile to review features of viral evolution that select for the best fit virus under host immune selection pressures. HIV-1 depends on the host to survive, and much of its evolution depends on the components of host immunity such as antibodies, NK cells and CTLs (165,249). Under these pressures there are two possible outcomes for the virus at the population level. First, in individuals without CTL selection pressure restricted by a specific HLA class I allele, the virus may remain unchanged(164,176), where the consensus amino acid represents the majority in the viral population . Second, in individuals with specific HLA class I alleles, CTL responses kill CD4 T cells infected by the wildtype virus, while cells infected by a mutant virus will survive because they cannot be recognized by patient HLA class I alleles. This process selects for the most fit virus (i.e.

positively selected) under hostile host immune conditions. With host selection pressure, only the PSM containing virus can survive, although it might be less fit than the virus carrying consensus amino acid (in patients without specific HLA alleles), thus potentially leading to slower disease progression. Therefore, positive selection refers to the process where the fittest virus under host immune pressures is selected due to having a better chance at survival. Our data shows that some Nef PSMs were associated with slower CD4 decline compared to the consensus amino acids at the cohort population level. This demonstrates that in those individuals infected by PSM containing virus, CD4 count decline was slower than individuals carrying viruses without PSMs (i.e. infected instead with virus that carries the consensus amino acid). Other Nef PSMs are associated with faster CD4 decline in comparison to the consensus amino acids. In this scenario, the PSM-harboring viral population possibly has a reproductive advantage over the consensus virus.

As shown by the data in this study, the impact of PSMs that arise within the virus are varied. Traditionally mutations were known to benefit the virus by helping it to evade the host immune response, thereby allowing the virus to replicate efficiently and maintain high levels of VL(283). However, mutations that arise in the virus as a result of HLA-restricted CTL pressure may also lead to a lesser functionally fit virus, and thus indirectly result in higher CD4 counts and lower VL(113,140,178,185,186). We have shown that some PSMs in Nef are associated with detrimental clinical outcomes, presumably via a functionally fit virus, while others are associated with slower disease progression, likely as the result of viral fitness cost. Viral fitness and disease progressions associated with HIV-1 mutations have been examined in a handful of recent studies (144,190,280). While HLA associated

mutations in Gag have been shown to alter replication capacity of the virus (113,284,285) one study showed that in recombinant HIV NL4-3 virus, containing Gag-protease targets of HLA-B*52 and B*67 alleles from a Japanese population, there was no association between replication capacity and pVL(144). This aligns with the findings from our study that some mutations may not have a significant fitness cost to the virus, as they are associated with faster CD4 decline in the infected host (E63D, I101V, I168M). Another study showed that the proportion of HLA-restricted escape mutations in Pol was negatively correlated with pVL while the same was not true in the case of Gag(171). Data from a study of subtype C Nef sequences from a large South African cohort has shown that HLA-B Nef polymorphisms are associated with fitness cost to the virus and HLA-B alleles that drive selection mutation in Nef are correlated with lower viral loads(190). Studies of HLA-associated HIV mutations with clinical outcomes have demonstrated the link between proportion of mutations and disease progression(280). One study showed that several HLA-associated mutations in Gag and Pol from subtype A/E infected patients were correlated with high pVL and others were associated with low CD4 T cell counts, both of which are unfavorable clinical outcomes for the infected hosts(280). In the same study no correlations were reported between HLA-associated polymorphisms in Nef and CD4 T cell count and pVL. These studies highlight the differential consequences of viral mutations and our findings corroborate with them in demonstrating that not all PSMs equally impact disease progression. We showed that some PSMs in HIV-1 Nef are associated with rapid CD4 decline, while other are associated with slower CD4 decline. Thus, our study extends on the findings of previous work, which have examined the impact of viral mutations on disease progression and overall viral fitness.

In this chapter, we identified PSMs using quasi analysis of subtype A1 Nef sequences from a cohort of commercial FSW in Nairobi, Kenya. We showed that PSMs within Nef were associated differentially with disease progression. Given that the Nef protein of HIV-1 is heavily targeted by multiple HLA Class I alleles(229), we correlated major HLA alleles in this cohort with the identified PSMs (Table 4.2). Our data showed that in some cases more than one allele is associated with a PSM, indicating that mutations are likely the result of multiple HLA class I alleles targeting Nef epitopes. Additionally, some HLA class I alleles can be correlated with PSMs that are associated with different disease progression rates. For example, HLA-A*02:01 restricted CD8+ T cells appear to drive the selection for H116N that is associated with slower CD4 decline, while also selecting for I101V, which is associated with faster CD4 decline. This finding has implications for choosing the best immunogen for a vaccine. HLA-A*02:01 expressing individuals could drive two types of PSMs with opposing outcomes, which can ultimately cancel out their respective effects. Therefore, to design an effective T cell based vaccine, it is critical to identify specific amino acid targets of host immune selection pressure and determine their association with clinical outcomes, to avoid potential opposing effects (252,286,287).

In conclusion, our study shows that PSMs in HIV-1 Nef are not equal in their impact on host disease progression. Understanding the characteristics Nef PSMs and their associated clinical outcomes may help to optimize immunogens for effective HIV vaccine candidates. CD8+ T cells that target regions of Nef associated with better outcomes may play a role in viral suppression and contribute to the enhanced immunity against the virus. Therefore, it

is worthwhile to investigate the characteristics of such CD8+ T cells, which will be the theme of the later part of this study.

4.7 Limitations and Opportunities

While Nef gene was amplified from more than 500 HIV-positive study participants, nearly one-third of sequences belonged to subtypes C, D and B (data not shown). Only 326 were subtype A1. Correlating a larger samples size of viral sequences with cohort HLA alleles could help identify associations with alleles that are present at lower frequencies in the population. Additionally, PSMs can be correlated with VL data to help draw insights about any possible links with viral suppression. However, since many samples used in this study were from archived specimen collected as far back as early 1980s, the poor storage conditions do not allow for accurate viral load assays. Host clinical parameters, consisting of both CD4 count and pVL, are most informative and we suspect that pVL data could provide additional information to the disease progression data associated with PSMs.

5. Confirmation and characterization of HIV-1 Nef epitopes containing PSMs that correlate with different disease outcome

5.1 Rationale

In the previous chapter PSMs in HIV-1 subtype A1 Nef sequence were identified using Quasi analysis. Of forty-eight PSMs identified in HIV-1 Nef three were associated with faster CD4 decline and two were associated with slower CD4 decline. HLA class I restricted cytotoxic T cells play an important role in driving HIV-1 mutations, which has been well documented in many studies(140,276,288). Mutations can accumulate within and outside peptide sequences of approximately 9 amino acids (9mer) to avoid recognition by HLA-restricted CD8+ T cells (274). Studies have shown that within 9mer HLA class I restricted CD8+ T cell epitopes, escape mutations often occur in HLA-peptide anchor positions 2, 8, and 9(151,288), and sometimes in the middle position of the 9mer peptide (e.g. position 5) (138). These mutations can result in diminished recognition by HLA-restricted CTLs and reduced TCR binding(151,289). In the previous chapter the HLA Class I alleles associated with PSMs in HIV-1 Nef were identified (Table 4.2). To confirm whether the identified mutations are within peptides targeted by the specific HLA class I alleles, the ability of such peptides in generating CD8+ T cell responses must be tested. Immune assays that detect peptide-specific CD8+ T cells in response to stimulation are useful for this purpose. For example, enzyme-linked immunosorbent spot (ELISPOT) assay can be used for direct *ex vivo* quantification of CD8+ T cell cytokines (e.g. IFN-γ) in PBMCs(290). In this chapter, the results of ELISPOT assay used to confirm the PSMs within HLA Class I restricted CD8+ T cell epitopes are shown.

Data from previous studies suggest that emergence of new mutants result in weakened CTL responses(105,274,291). However, studies have also shown that new CD8+ T cells can arise concomitant with new virus variants despite ongoing viral evolution (113,135,291-298). It is therefore, important to confirm whether HLA class I restricted CD8+ T cells can recognize epitope variants with PSMs. Identification and characterization of these HLA Class I restricted CD8+ T cell epitopes in HIV-1 Nef can provide insights about vulnerable regions of the virus that may be important for vaccine design.

5.2 Hypothesis

Major Hypothesis

HLA class I restricted CD8+ T cell epitopes containing the identified PSMs will elicit CD8+ T cell responses that can be measured by ELISPOT assay.

Sub Hypothesis

The frequency of IFN-γ-producing HLA Class I restricted CD8+ T cells stimulated by epitopes harbouring PSM or subtype A1 consensus amino acids is similar.

5.3 Objective

To confirm the HLA class I restricted CD8+ T cell epitopes in Nef, containing consensus or PSM amino acids, using IFN-γ ELISPOT assay.

5.4. Methods

5.4.1. Study Cohort

The study participants consisted of 168 patients enrolled in the Pumwani female sex-worker cohort who were typed for HLA class I genes and whose HLA Class I alleles correlated with PSMs (Table 4.2).

5.4.2. IFN-γ ELISPOT Assay and Peptide design

IFN-γ ELISPOT assay was performed using fresh patients' PBMCs according to methodology described in section 2.9, page 51. Briefly, in pre-anti-IFN-γ monoclonal antibody-coated plates, PBMCs were stimulated overnight with synthesized 9mer peptides (Sigma-Aldrich, Custom Peptide Library | PEPscreen®) containing specific PSM or subtype A1 consensus amino acids in anchor positions 2, 5, 8 or 9 (Figure 5.1 and Table 5.1). Biotinylated anti-IFN-γ mAb and streptavidin-conjugated alkaline phosphatase were added sequentially for the detection of IFN- γ. After wash steps and color reaction of 15-20 minutes with alkaline-phosphatase substrate kit, IFN-γ-producing cells were detected as dark spots counted by ELISPOT reader. Values ≥50 SFUs/million PBMCs, after subtracting negative control for each reaction were considered as positive responses.

5.4.3 Statistical analysis

Mann Whitney non-parametric tests and non-parametric multiple comparison Kruskal-Wallis tests were performed using GraphPad Prism v7.0a (GraphPad Software, La Jolla, CA) and p values <0.05 were considered statistically significant.

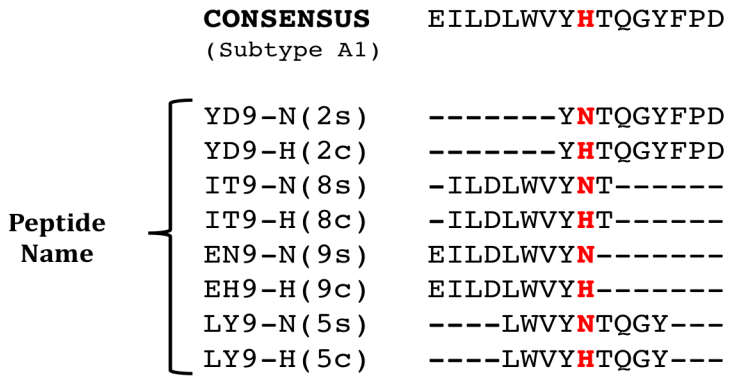


Figure 5.1. Schematic diagram of Nef peptide design strategy with PSM or subtype A1 consensus amino acids at anchor positions 2, 8, 9 and middle position 5. Peptides names contain the first and last amino acids of the 9mer, followed by 9 indicating the peptide length, "-", and amino acid of interest, its position within the 9mer and whether PSM or consensus within brackets.

Table 5.1. Sequences and names of all Nef peptides used in ELISPOT assay with PSMs located at anchor positions 2, 8, 9 and middle position 5.

PSM	Peptide Name	Peptide Sequence
E63D	EV9-D(2s)	E**D**EEVGFPV
	EV9-E(2c)	E**E**EEVGFPV
	VE9-D(8s)	VWLEAQE**D**E
	VE9-E(8c)	VWLEAQE**E**E
	CD9-D(9s)	CVWLEAQE**D**
	CE9-E(9c)	CVWLEAQE**E**
	EG9-D(5s)	EAQE**D**EEVG
	EG9-E(5c)	EAQE**E**EEVG
I101V	LE9-V(2s)	L**V**YSRKPQE
	LE9-I(2c)	L**I**YSRKPQE
	KY9-V(8s)	KGGLDGL**V**Y
	KY9-I(8c)	KGGLDGL**I**Y
	EV9-V(9s)	EKGGLDGL**V**
	EI9-I(9c)	EKGGLDGL**I**
	LK9-V(5s)	LDGL**V**YSRK
	LK9-I(5c)	LDGL**I**YSRK
I168M	PD9-M(2s)	P**M**CQHGMDD
	PD9-I(2c)	P**I**CQHGMDD
	NC9-M(8s)	NNSLLHP**M**C
	NC9-I(8c)	NNSLLHP**I**C
	EM9-M(9s)	ENNSLLHP**M**
	EI9-I(9c)	ENNSLLHP**I**
	LG9-M(5s)	LLHP**M**CQHG
	LG9-I(5c)	LLHP**I**CQHG
H116N	YD9-N(2s)	Y**N**TQGYFPD
	YD9-H(2c)	Y**H**TQGYFPD
	IT9-N(8s)	ILDLWVY**N**T
	IT9-H(8c)	ILDLWVY**H**T
	EN9-N(9s)	EILDLWVY**N**
	EH9-H(9c)	EILDLWVY**H**
	LY9-N(5s)	LWVY**N**TQGY
	LY9-H(5c)	LWVY**H**TQGY
K182M	LL9-M(2s)	L**M**WKFDSRL
	LL9-K(2c)	L**K**WKFDSRL
	DW9-M(8s)	DEEREVL**M**W
	DW9-K(8c)	DEEREVL**K**W
	DM9-M(9s)	DDEEREVL**M**
	DK9-K(9c)	DDEEREVL**K**
	RD9-M(5s)	REVL**M**WKFD
	RD9-K(5c)	REVL**K**WKFD

5.5 Results

5.5.1. Multiple HIV-1 Nef Epitopes containing PSMs associated with differential rates of disease progression confirmed by ELISPOT assay

In the previous chapter, it was shown that three PSMs were associated with faster CD4 decline compared to the consensus amino acids, and two were associated with slower CD4 decline (Figure 4.3). Additionally, each of these PSMs were correlated with multiple HLA Class I alleles common in Kenya (Table 4.2 and 5.2). To confirm that these alleles can present the 9mer peptides with the identified PSMs to CD8+ T cells, IFN-γ ELISPOT assay was used. For each of the five PSMs and their consensus counterpart (E63D, I101V, H116N, I168M and K182M), 9mer peptides were synthesized with the PSM or consensus amino acids placed in positions 2, 5, 8 and 9 (Figure 5.1 and Table 5.1). PBMCs from patients with specific HLA class I alleles that correlated with each of the PSMs were stimulated with corresponding peptides and the frequency of IFN-γ producing CD8+ T cells were measured. Table 5.2 shows the detailed results of all 40 peptides tested and lists the results for each responding patient. IFN-γ ELISPOT responses were detected to peptides containing PSMs and consensus amino acid (Table 5.2). For example, IFN-γ responses to peptides VWLEAQE**D**E (VE9-D(8s)) and VWLEAQE**E**E (VE9-E(8c)) were detected in PBMCs of at least one A*23:01+ patient. In this case, it appears that there are CD8+ T cells that can recognize both peptide variants. Therefore, peptide-HLA binding was not compromised by the amino acid variation introduced by the PSM. However, no IFN-γ responses were detected to peptide E**D**EEVGFPV containing the same PSM at anchor position 2 (EV9-D(2s)) in PBMCs of any of the seven A*2301+ patients

tested. Whereas, IFN-γ responses to peptide E**E**EVGFPV (EV9-E(2c)) containing the consensus amino acid "E" were detected in at least one A*23:01+ patient. This likely indicates that the PSM amino acid "D" interferes with peptide-HLA binding.

Positive IFN-γ ELISPOT responses (≥50 SFU/million PBMCs after background subtraction) to 27 out of 40 peptides tested were detected in patient PBMCs (Table 5.3). Figure 5.2 shows a map of subtype A1 Nef protein with all confirmed epitopes. Out of eight peptides tested for each PSM/consensus group of peptides (Table 5.1), by stimulating PBMCs of patients whose HLA class I alleles correlated with PSMs (Table 4.2), ELISPOT assay confirmed six E63D peptides (CVWLEAQED, CVWLEAQEE, VWLEAQEDE, VWLEAQEEE, EAQEEEEVG, EEEEVGFPV), eight I101V peptides (EKGGLDGLV, EKGGLDGLI, KGGLDGLVY, KGGLDGLIY, LDGLVYSRK, LDGLIYSRK, LVYSRKPQE, LIYSRKPQE), eight H116N peptides (EILDLWVYN, EILDLWVYH, ILDLWVYNT, ILDLWVYHT, LWVYNTQGY, LWVYHTQGY, YNTQGYFPD YHTQGYFPD) and five K182M peptides (DDEEREVLK, DEEREVLMW, DEEREVLKW, REVLMWKFD, LMWKFDSRL). None of the eight I168M peptides were confirmed by ELISPOT assay in PBMCs of patients with A*2301, B*07:02, B*58:01 or Cw*16:01 alleles. The "unrecognizability" of I168M peptides, at least by these HLA molecules, may be due to functional role of this region of Nef in viral pathogenesis and the need to remain conserved in order to preserve function. The confirmation of other peptides tested by ELISPOT in this study shows that they are indeed targets of CD8+ T cells restricted by HLA alleles that were statistically correlated with PSMs (Table 4.2). This confirms our hypothesis that peptides containing PSMs identified through

bioinformatics analyses can be confirmed as T cell epitopes by ELISPOT. Therefore, CD8+ T cell functional profiles specific to these epitopes can be assessed using other immune assays. Table 5.3 depicts the detailed results of these 27 confirmed epitopes with results for each responding patient.

We also used dose-dependent ELISPOT assay to further confirm a subset of peptides for which archived PBMCs from patients with specific HLA alleles were available. Two HLA A*02:01 epitopes, ILDLWVYHT and ILDLWVYNT, and two HLA B*58:01 epitopes, KGGLDGLIY and KGGLDGLVY, were tested with two-fold dilutions at various concentrations (10 ug/ml- 0.625 ug/ml) in PBMCs of 10 patients for each. The results showed that for both HLA-peptide groups the frequency of IFN-γ+CD8+ T cells were decreased with reduction in peptide concentrations (Figure 5.3). Therefore, despite a low number of patients tested using fresh PBMCs (4 A*02:01 and 6 B*58:01 patients, Table 5.2), does-dependent ELISPOT results for these four HLA-A*02:01 and B*58:01 peptides, confirm their binding with the respective HLA molecules and validate them as true CD8+ T cell epitopes.

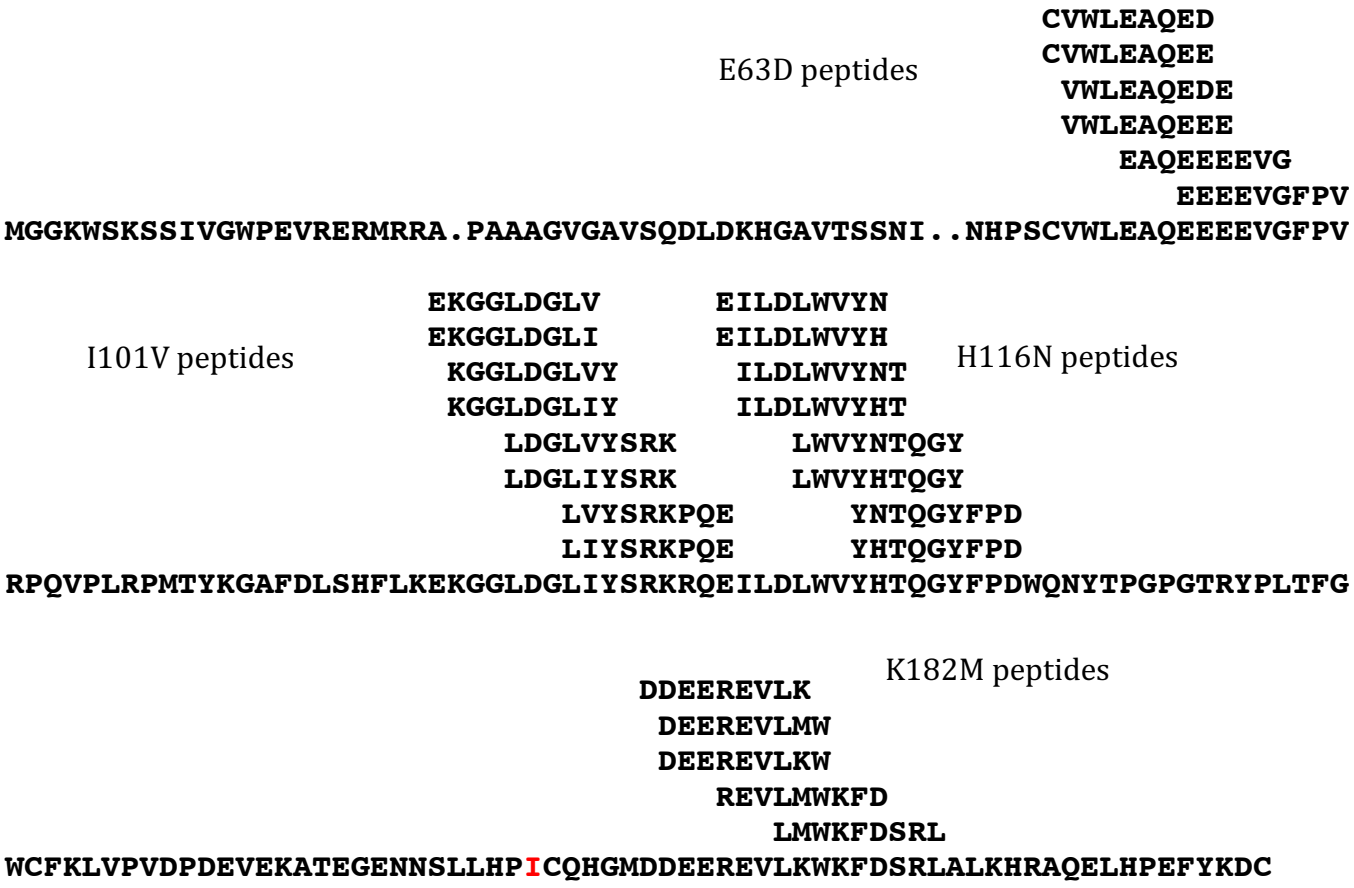


Figure 5.2. Map of HIV-1 subtype A1 Nef amino acids with epitopes confirmed using ELISPOT assay. Peptide sequences are shown above the consensus sequence and labeled accordingly (*note*: the 8 peptides tested for I168M, shown in the consensus map of Nef in red, did not result in positive IFN-γ responses and therefore, were not confirmed as true epitopes).

Table 5.2. Summary of ELISPOT data for all 40 HIV-1 Nef peptides containing PSMs or consensus amino acids tested in PBMCs from patients whose HLA alleles were correlated with a specific PSM.

Allele	PSM	Peptide Code*	Epitope Sequence	# of Patients tested	# of patients responded	Response to Both s/c	Mean SFU/10^6*
A*23:01	E63D	EV9-D(2s)	E**D**EEVGFPV	8	0	No	-
		EV9-E(2c)	E**E**EEVGFPV	8	1	No	100.00
		VE9-D(8s)	VWLEAQE**D**E	8	2	Yes	72.50
		VE9-E(8c)	VWLEAQE**E**E	8	1	Yes	50.00
		CD9-D(9s)	CVWLEAQE**D**	8	3	Yes	73.33
		CE9-E(9c)	CVWLEAQE**E**	8	1	Yes	115.00
		EG9-D(5s)	EAQE**D**EEVG	8	0	No	-
		EG9-E(5c)	EAQE**E**EEVG	8	1	No	85.00
B*35:01		EV9-D(2s)	E**D**EEVGFPV	2	0	-	-
		EV9-E(2c)	E**E**EEVGFPV	2	0	-	-
		VE9-D(8s)	VWLEAQE**D**E	2	0	-	-
		VE9-E(8c)	VWLEAQE**E**E	2	0	-	-
		CD9-D(9s)	CVWLEAQE**D**	2	0	-	-
		CE9-E(9c)	CVWLEAQE**E**	2	0	-	-
		EG9-D(5s)	EAQE**D**EEVG	2	0	-	-
		EG9-E(5c)	EAQE**E**EEVG	2	0	-	-
A*02:01	I101V	LE9-V(2s)	L**V**YSRKPQE	6	1	Yes	110
		LE9-I(2c)	L**I**YSRKPQE	6	1	Yes	125
		KY9-V(8s)	KGGLDGL**V**Y	6	1	Yes	125
		KY9-I(8c)	KGGLDGL**I**Y	6	1	Yes	80
		EV9-V(9s)	EKGGLDGL**V**	6	1	Yes	75
		EI9-I(9c)	EKGGLDGL**I**	6	1	Yes	75
		LK9-V(5s)	LDGL**V**YSRK	6	1	Yes	135
		LK9-I(5c)	LDGL**I**YSRK	6	1	Yes	50
B*58:01		LE9-V(2s)	L**V**YSRKPQE	6	0	-	-
		LE9-I(2c)	L**I**YSRKPQE	6	0	-	-
		KY9-V(8s)	KGGLDGL**V**Y	6	2	Yes	375
		KY9-I(8c)	KGGLDGL**I**Y	6	1	Yes	700
		EV9-V(9s)	EKGGLDGL**V**	6	0	No	-
		EI9-I(9c)	EKGGLDGL**I**	6	1	No	75
		LK9-V(5s)	LDGL**V**YSRK	6	0	No	50
		LK9-I(5c)	LDGL**I**YSRK	6	1	No	75
A*23:01	I168M	PD9-M(2s)	P**M**CQHGMDD	4	0	-	-
		PD9-I(2c)	P**I**CQHGMDD	4	0	-	-
		NC9-M(8s)	NNSLLHP**M**C	4	0	-	-
		NC9-I(8c)	NNSLLHP**I**C	4	0	-	-

		EM9-M(9s)	ENNSLLHP**M**	4	0	-	-
		EI9-I(9c)	ENNSLLHP**I**	4	0	-	-
		LG9-M(5s)	LLHP**M**CQHG	4	0	-	-
		LG9-I(5c)	LLHP**I**CQHG	4	0	-	-
B*07:02		PD9-M(2s)	P**M**CQHGMDD	0	0	-	-
		PD9-I(2c)	P**I**CQHGMDD	0	0	-	-
		NC9-M(8s)	NNSLLHP**M**C	0	0	-	-
		NC9-I(8c)	NNSLLHP**I**C	0	0	-	-
		EM9-M(9s)	ENNSLLHP**M**	0	0	-	-
		EI9-I(9c)	ENNSLLHP**I**	0	0	-	-
		LG9-M(5s)	LLHP**M**CQHG	0	0	-	-
		LG9-I(5c)	LLHP**I**CQHG	0	0	-	-
B*58:01		PD9-M(2s)	P**M**CQHGMDD	2	0	-	-
		PD9-I(2c)	P**I**CQHGMDD	2	0	-	-
		NC9-M(8s)	NNSLLHP**M**C	2	0	-	-
		NC9-I(8c)	NNSLLHP**I**C	2	0	-	-
		EM9-M(9s)	ENNSLLHP**M**	2	0	-	-
		EI9-I(9c)	ENNSLLHP**I**	2	0	-	-
		LG9-M(5s)	LLHP**M**CQHG	2	0	-	-
		LG9-I(5c)	LLHP**I**CQHG	2	0	-	-
Cw*16:01		PD9-M(2s)	P**M**CQHGMDD	2	0	-	-
		PD9-I(2c)	P**I**CQHGMDD	2	0	-	-
		NC9-M(8s)	NNSLLHP**M**C	2	0	-	-
		NC9-I(8c)	NNSLLHP**I**C	2	0	-	-
		EM9-M(9s)	ENNSLLHP**M**	2	0	-	-
		EI9-I(9c)	ENNSLLHP**I**	2	0	-	-
		LG9-M(5s)	LLHP**M**CQHG	2	0	-	-
		LG9-I(5c)	LLHP**I**CQHG	2	0	-	-
A*02:01	H116N	YD9-N(2s)	Y**N**TQGYFPD	4	1	Yes	85
		YD9-H(2c)	Y**H**TQGYFPD	4	1	Yes	75
		IT9-N(8s)	ILDLWVY**N**T	4	1	Yes	110
		IT9-H(8c)	ILDLWVY**H**T	4	2	Yes	87.50
		EN9-N(9s)	EILDLWVY**N**	4	1	Yes	55.00
		EH9-H(9c)	EILDLWVY**H**	4	2	Yes	75.00
		LY9-N(5s)	LWVY**N**TQGY	4	2	Yes	80.00
		LY9-H(5c)	LWVY**H**TQGY	4	1	Yes	55.00
A*30:02		YD9-N(2s)	Y**N**TQGYFPD	2	0	-	-
		YD9-H(2c)	Y**H**TQGYFPD	2	0	-	-
		IT9-N(8s)	ILDLWVY**N**T	2	1	No	65.00
		IT9-H(8c)	ILDLWVY**H**T	2	0	No	-
		EN9-N(9s)	EILDLWVY**N**	2	0	-	-
		EH9-H(9c)	EILDLWVY**H**	2	0	-	-

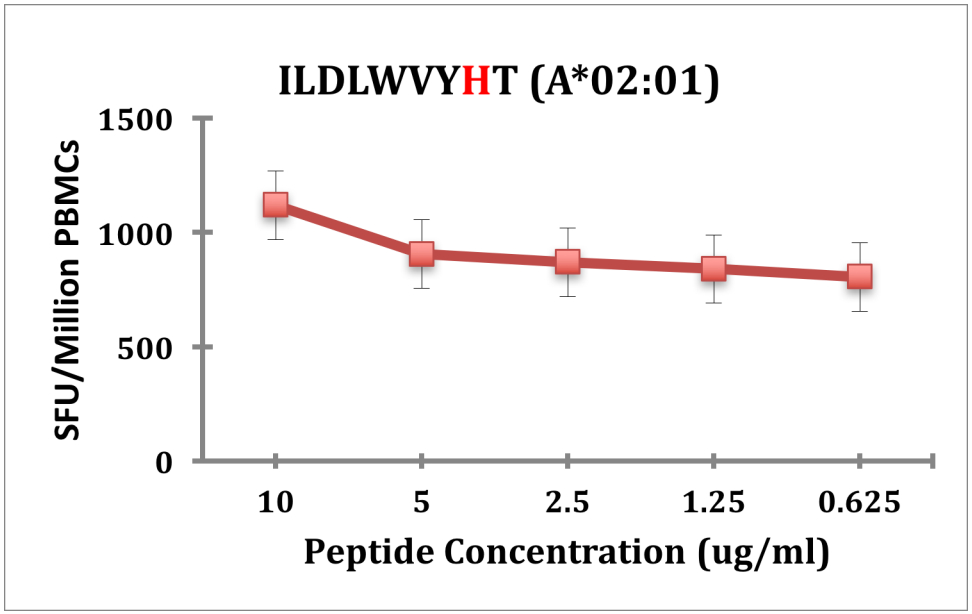
		LY9-N(5s)	LWVY**N**TQGY	2	0	No	-
		LY9-H(5c)	LWVY**H**TQGY	2	1	No	60.00
B*57:03		YD9-N(2s)	Y**N**TQGYFPD	3	0	-	-
		YD9-H(2c)	Y**H**TQGYFPD	3	0	-	-
		IT9-N(8s)	ILDLWVY**N**T	3	1	Yes	175.00
		IT9-H(8c)	ILDLWVY**H**T	3	1	Yes	90.00
		EN9-N(9s)	EILDLWVY**N**	3	0	No	-
		EH9-H(9c)	EILDLWVY**H**	3	1	No	70.00
		LY9-N(5s)	LWVY**N**TQGY	3	1	No	115.00
		LY9-H(5c)	LWVY**H**TQGY	3	0	No	-
Cw*02:10		YD9-N(2s)	Y**N**TQGYFPD	2	0	-	-
		YD9-H(2c)	Y**H**TQGYFPD	2	0	-	-
		IT9-N(8s)	ILDLWVY**N**T	2	0	-	-
		IT9-H(8c)	ILDLWVY**H**T	2	0	-	-
		EN9-N(9s)	EILDLWVY**N**	2	1	No	60.00
		EH9-H(9c)	EILDLWVY**H**	2	0	No	-
		LY9-N(5s)	LWVY**N**TQGY	2	0	-	-
		LY9-H(5c)	LWVY**H**TQGY	2	0	-	-
A*34	K182M	LL9-M(2s)	L**M**WKFDSRL	1	0	-	-
		LL9-K(2c)	L**K**WKFDSRL	1	0	-	-
		DW9-M(8s)	DEEREVL**M**W	1	0	-	-
		DW9-K(8c)	DEEREVL**K**W	1	0	-	-
		DM9-M(9s)	DDEEREVL**M**	1	0	-	-
		DK9-K(9c)	DDEEREVL**K**	1	0	-	-
		RD9-M(5s)	REVL**M**WKFD	1	0	-	-
		RD9-K(5c)	REVL**K**WKFD	1	0	-	-
Cw*16:01		LL9-M(2s)	L**M**WKFDSRL	9	1	No	50.00
		LL9-K(2c)	L**K**WKFDSRL	9	0	No	-
		DW9-M(8s)	DEEREVL**M**W	9	1	Yes	50.00
		DW9-K(8c)	DEEREVL**K**W	9	1	Yes	55.00
		DM9-M(9s)	DDEEREVL**M**	9	0	-	-
		DK9-K(9c)	DDEEREVL**K**	9	0	-	-
		RD9-M(5s)	REVL**M**WKFD	9	1	No	70.00
		RD9-K(5c)	REVL**K**WKFD	9	0	No	-

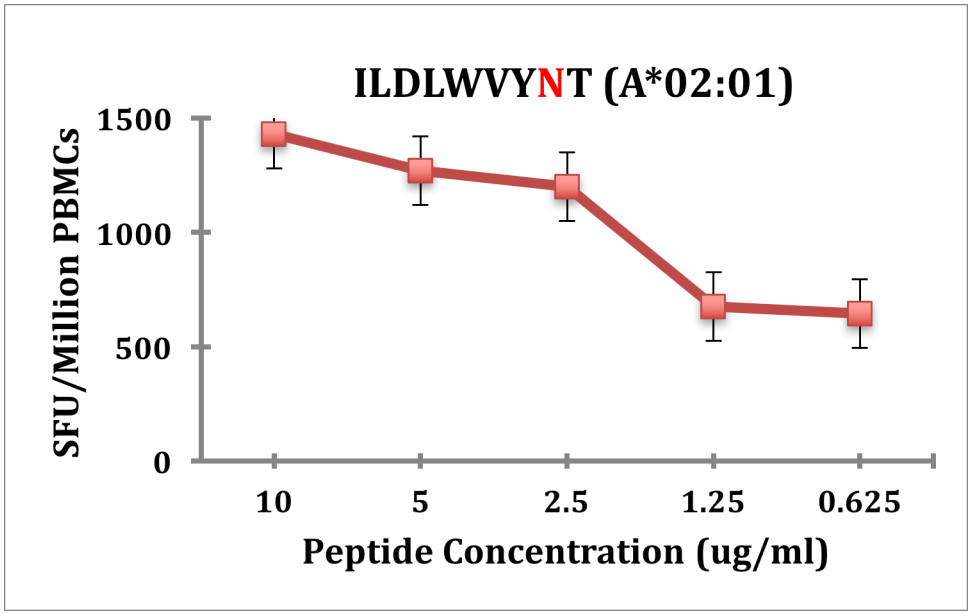
**Mean SFU/10^6 is the mean of number of IFNγ-producing CD8+ T cells per million of PBMCs of all patients who have responded to a given peptide*

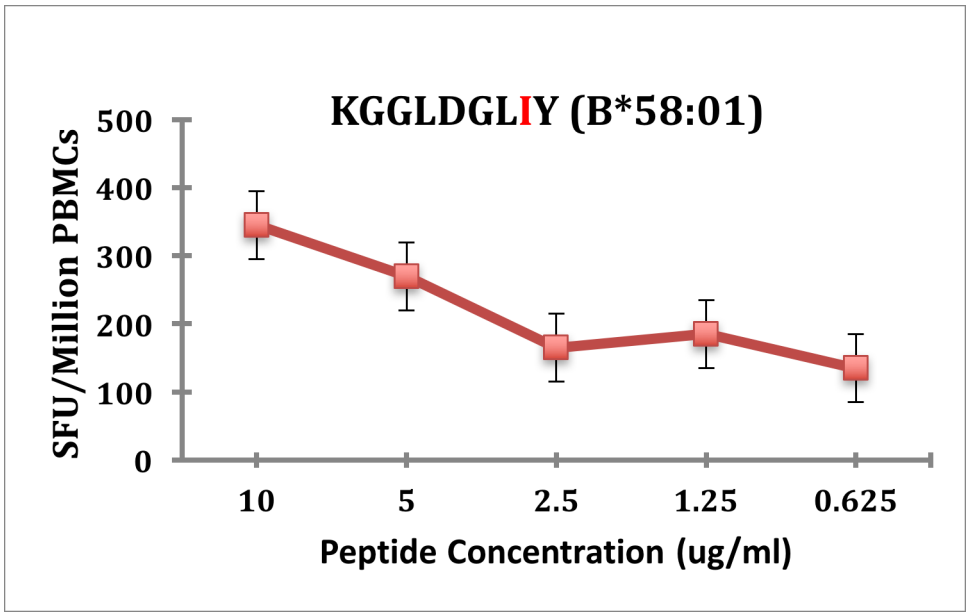
Table 5.3. Summary of ELISPOT results for 27 confirmed peptides of HIV-1 Nef containing PSM or consensus amino acids correlated with various HLA class I alleles. Peptides for which no values are listed did not meet the ≥50 SFU/10^6 cells threshold.

PSM	Allele	ML I.D.*	Frequency of IFNg+ CD8+ T cells (SFU/10^6 cells)							
E63D			**EV9-D(2s)**	**EV9-E(2c)**	**VE9-D(8s)**	**VE9-E(8c)**	**CD9-D(9s)**	**CE9-E(9c)**	**EG9-D(5s)**	**EG9-E(5c)**
	A* 23:01	612			55	50	70	115		85
		1490		100	90		100			
		3268					50			
I101V			**LE9-V(2s)**	**LE9-I(2c)**	**KY9-V(8s)**	**KY9-I(8c)**	**EV9-V(9s)**	**EI9-I(9c)**	**LK9-V(5s)**	**LK9-I(5c)**
	A* 02:01	3277	110	125	125	80	75	75	135	50
	B* 58:01	3156								75
		1947						50		
		915			685	700				
		2493			60					
H116N			**YD9-N(2s)**	**YD9-H(2c)**	**IT9-N(8s)**	**IT9-H(8c)**	**EN9-N(9s)**	**EH9-H(9c)**	**LY9-N(5s)**	**LY9-H(5c)**
	A* 02:01	2877				60	55	50	70	
		3277	85	75	110	115		95	90	50
		2049	75	95	130	60	85	95	95	70
		59			445	315				
	A* 3002	2257			65					
	B* 57:03	3009			175	90		70	115	
	C* 02:01	3268								
K182M			**LL9-M(2s)**	**LL9-K(2c)**	**DW9-M(8s)**	**DW9-K(8c)**	**DM9-M(9s)**	**DK9-K(9c)**	**RD9-M(5s)**	**RD9-K(5c)**
	C* 06:02	2680							70	
		3434	50		55	50			50	
		1665						50		
		1778						60		

**"ML I.D." is Majengo cohort member identifier*







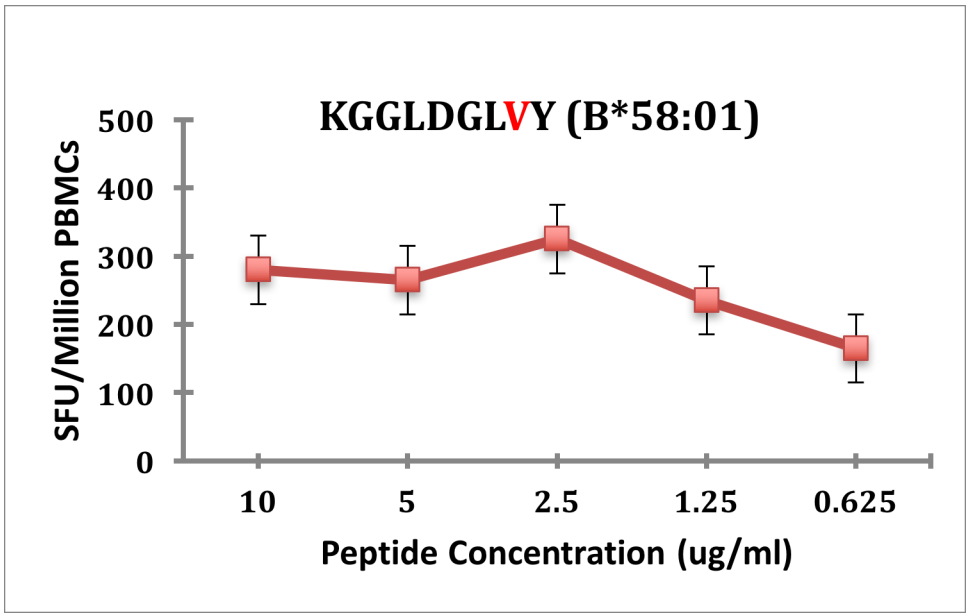


Figure 5.3. ELISPOT results for dose-dependent assay with peptide concentration using two-fold dilution (10 ug/ml- 0.625 ug/ml), for two HLA A*02:01 epitopes: ILDLWVYHT and ILDLWVYNT, and for two HLA B*58:01 epitopes: KGGLDGLIY and KGGLDGLVY

5.5.2 Similar IFNγ ELISPOT responses to peptides containing PSMs associated with slow or fast CD4 T cell decline

ELOISPOT assays were conducted to determine whether there is a difference in the frequency of IFNγ+ CD8+ T cells when stimulated by peptides that contain PSMs or consensus amino acids. Among the detectable responses there were no differences in frequency of IFNγ+ CD8+ T cells when PBMCs were stimulated by peptides that contained PSM or consensus amino acid, irrespective of association with different CD4 decline outcomes ($p = 0.2979$ Mann-Whitney test, Figure 5.4 A). When comparing the frequency of IFNγ+ CD8+ T cells in PBMCs stimulated by peptides containing amino acids associated with slower or faster CD4 decline, there were no significant differences ($p= 0.4660$ Mann-Whitney test, Figure 5.4 B). This was irrespective of amino acid of interest being a consensus or PSM. The difference in the proportion of CD8+ T cells that express IFNγ in PBMCs when elicited by peptides encompassing consensus or selected amino acid at various anchor positions was also assessed to determine if amino acid position plays a role (Figure 5.5). Although, peptides with PSM or consensus amino acid placed in anchor positions 8 and 9 were more frequently recognized by host HLA (20 and 17 responses, respectively), there was no significant difference in the frequency of IFNγ-producing CD8+ T cells among the groups (nonparametric multiple comparison Kruskal-Wallis test). This included peptides with both consensus and PSM, regardless of association with CD4 decline. Finally, when separating these responses based on peptide categories, i.e. those with beneficial or detrimental amino acids, there were also no differences in the number of IFNγ producing CD8+ T cells (Mann-Whitney test, Figure 5.6 A-D). In

summary, these results support the hypothesis that there are no differences in the frequency of IFNγ+ CD8+ T cells when PBMCs are stimulated with epitopes containing PSM or consensus amino acid.

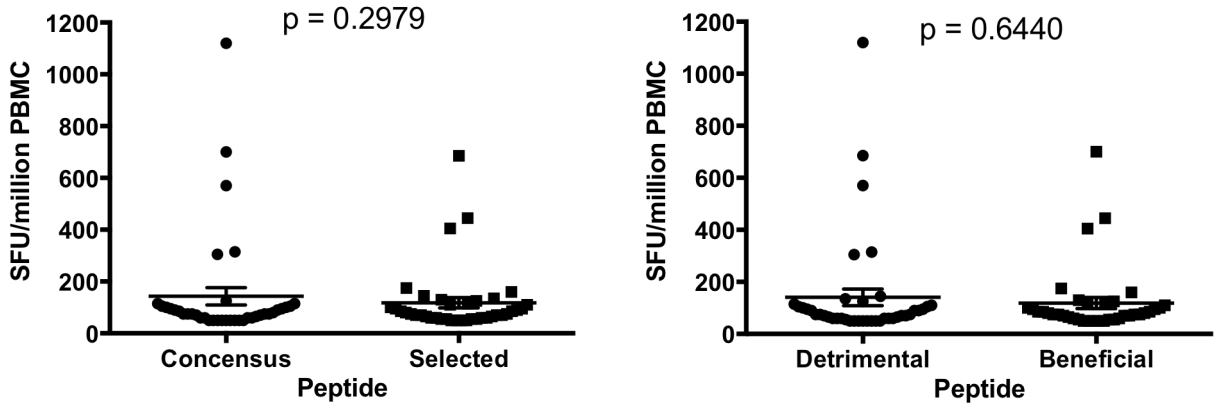


Figure 5.4. The frequencies of IFNγ producing CD8+ T cells in PBMCs after stimulation with peptides containing PSMs associated with different clinical outcomes of disease. (A) shows the comparison between all peptides containing consensus or selected (PSM) amino acids. (B) comparison of the frequencies of IFNγ producing CD8+ T cells in PBMCs after stimulation with peptides containing PSMs associated with beneficial disease outcome (slower CD4 decline) and peptides with amino acids associated with detrimental disease outcomes (faster CD4 decline).

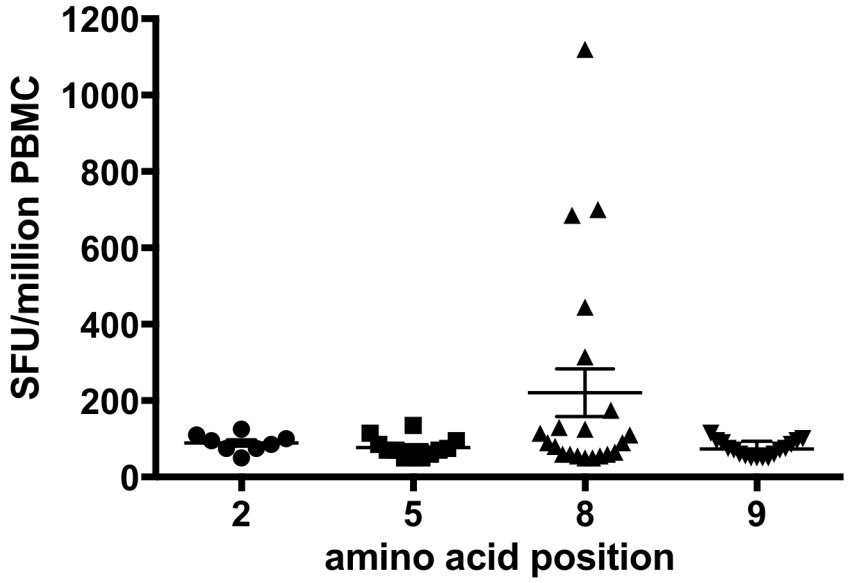


Figure 5.5. The frequencies of IFNγ producing CD8+ T cells in PBMCs after stimulation with peptides with PSMs placed at various positions along the 9mer peptide.

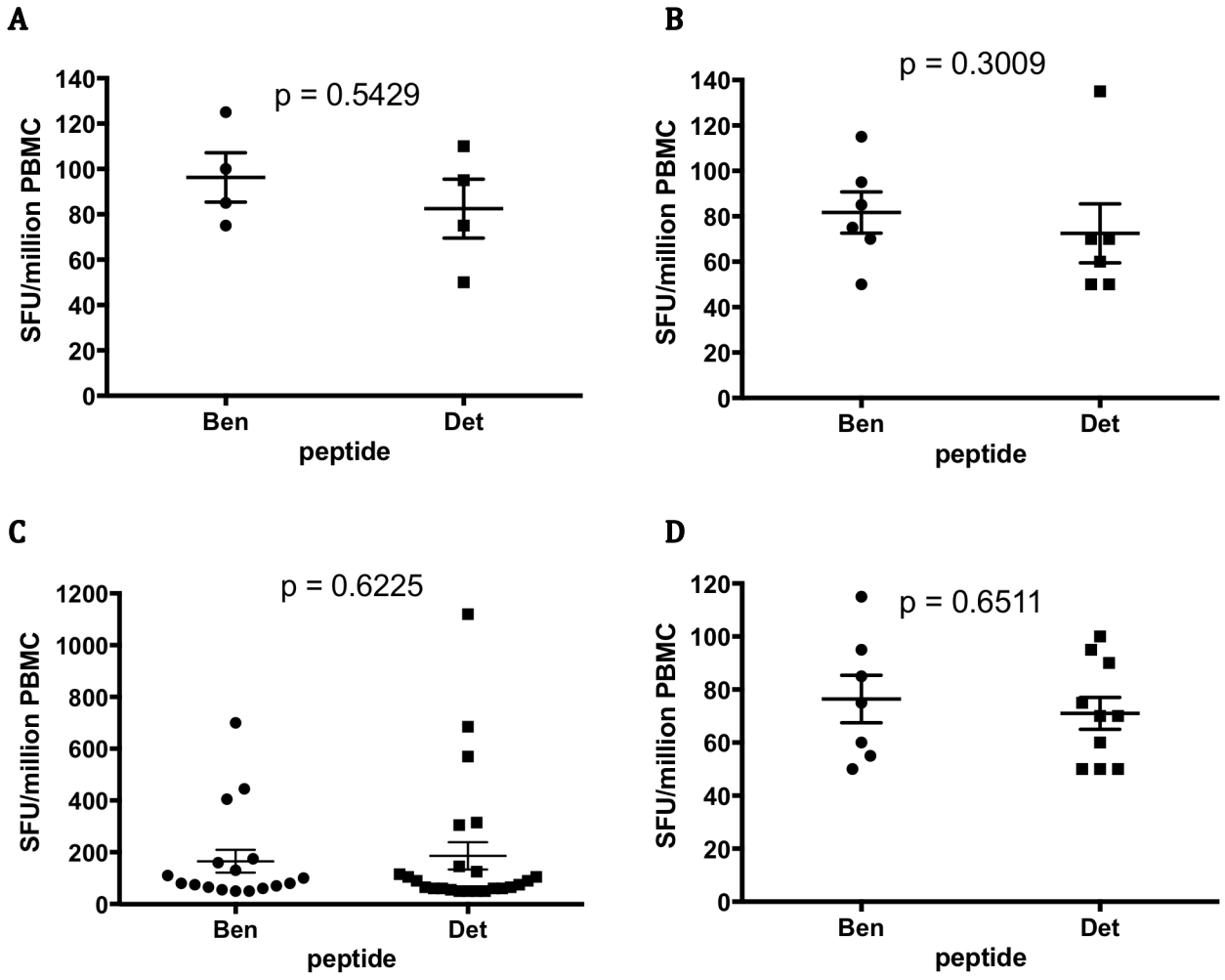


Figure 5.6. The frequencies of IFNγ producing CD8+ T cells in PBMCs after stimulation with peptides containing amino acids associated with beneficial ("Ben") and detrimental ("Det") outcomes when PSMs were in 9mer positions 2, 5, 8 and 9 (A-D respectively).

5.6 Discussion

Selective pressure on HIV-1 exerted by HLA class I restricted CD8+ T cells has been illustrated in detail in multiple studies(134,140,143,145,158,295,299). To validate the observed HLA associations, previously described in chapter 4 (Table 4.2), with specific Nef PSMs, INF-γ ELISPOT assays were conducted using overlapping peptides of the targeted epitopes to stimulated PBMCs of patients with specific HLA class I alleles. The 9-mer peptides with either the consensus or selected amino acid at anchor positions 2, 8 and 9 and middle position 5 were synthesized and used to stimulate patient PBMCs. Amino acid substitutions at these positions have been shown to play a role in peptide-HLA binding,(145) therefore, we were interested to investigate if position could play a role in epitope confirmation. Of the 40 overlapping peptides tested, 27 were confirmed with positive IFN- γ ELISPOT response. Although the number of respondents for several peptides were low (Table 5.2 and Table 5.3), two sets of peptides (A*02:01 epitopes ILDLWVYHT, ILDLWVYNT and B*58:01 epitopes KGGLDGLIY, KGGLDGLVY) were further tested using dose-dependent ELISPOT assay in PBMCs from two groups of HLA specific patients (10 of each of A*02:01 and B*58:01). The data from this analysis shows effective peptide-HLA binding at various concentrations (Figure 5.3), which validates each peptide as an HLA targeted epitope. Indeed, one of the challenges of conducting ELISPOT with fresh PBMC samples was the low number of patients with specific HLA alleles from whom samples were obtained. However, we believe that through dose-dependent ELISPOT analysis using archived frozen PBMCs we have confidently proven predicted epitopes as true epitopes.

Our data showed that there is no significant difference between the frequency of IFN-γ producing CD8+ T cells in response to stimulation by peptide with consensus or PSM amino acids (Figure 5.4). Previous studies have shown that CTL activity tends to diminish in response to new mutant virus, demonstrating that viral mutations leads to escape from CD8+ T cells (146,291,293). However, numerous studies have also indicated that CD8+ T cells adapt to the emerging new viral variants despite ongoing viral evolution, through the generation of de novo CD8+ T cell population specific to the mutant epitopes(113,135,291-296). Our results are in agreement with reports that CD8+ T cell responses generated towards consensus and PSM epitopes can be equal, as measured by IFN-γ producing CD8+T cells(293). CD8+ T cell adaptation to ongoing evolution of Nef has been described previously (135). One study examining viral escape mutations in HIV-1 clones, obtained longitudinally from a single patient, showed that mutations in Nef are recognized to the same extent as the wild type strain and Nef mutations did not result in reduced CD8+ T cell reactivity when compared to Gag, Pol and accessory proteins of HIV-1(291). Additionally, even Nef mutant epitopes that result in minimal CD8+ T cell reactivity, such as the Y120F epitope still had comparable responses similar to their wild type counterpart(136). The emergence of novel CD8+ T cell populations specific to escape mutant epitopes in other HIV-1 proteins have also been described previously. In HLA-A*11, HLA-B*57 and HLA-B*27 individuals, the frequencies of mutant epitope-specific and wild type epitope-specific CD8+ T cells were similar (113,166,300).

Overall, the findings of our study agree with several previous reports that first, escape mutations do not necessarily lead to loss of CTL responses, and second, PSM-specific CD8+ T cell frequencies are comparable to CD8+ T cells against consensus epitope. Although initially viral mutation leads to lower and weaker CTL responses this is likely a feature of the acute phase of the infection when the CD8+ T cell targeting of viral sequences is at its peak(301). As the viral variants begin to emerge novel CD8+ T cells that target mutant epitopes rise immediately (291), likely as a result of recruitment of new TCR clonotypes, which are also capable of recognizing the mutant viral sequences(293). Therefore, a new clone of CD8+ T cells can recognize the mutant epitope, leading to equal levels of CTL reactivity as previously to wildtype epitope (302) (303).

There are also opposing views to the findings of our study. For example, it has been shown that SIV escape variants lead to poor epitope recognition by CD8+ T cells in naïve but immunocompetent macaques(153). The study showed that CTL escape mutants did not generate CD8+ T cell responses since the mutations were in positions that were crucial for TCR binding. In this study, I found that position of PSMs in the HLA-specific epitopes did not affect CD8+ T cell recognition of peptides that associated differentially with CD4 decline (Figure 5.5). It can be speculated that similar recognition of HLA/TCR of different epitopes is due to similarities of PSM and consensus amino acids.

My study also showed that there was no significant difference in the numbers of IFN-γ producing CD8+ T cells after being stimulated by peptides containing PSMs associated with slow or fast CD4 decline (Figure 5.4 B, Figure 5.6). One explanation could be that patient HLA alleles are able to recognize both epitopes containing consensus and PSMs. Multiple studies have shown that escape mutations may not affect peptide-HLA binding and TCR presentation(113,300,304). The observation that both epitopes are recognized equally by HLA-restricted CD8+ T cells is an indication that to escape, the virus is not able to deviate from one or two closely related amino acids, perhaps due to the importance of maintaining the function of certain motifs. This was demonstrated in a study which analyzed CTL responses to mutant epitopes in p1 spacer protein of HIV-1 Gag (305). The HLA-B*13:02 driven K4R mutation in p1 spacer epitopes did not impact CD8+ T cell responses. The physical and chemical properties of lysine and arginine, such as positive charge and hydrophilic characteristics of both amino acids were thought to play a role in the consistent CD8+ T cell outcomes(305). It is likely that to maintain the function of Nef or viability of the virus few options are available for virus to escape host HLA class I restricted CD8+ T cell responses.

In this study, it was also observed that none of the 8 variants of I168M peptides could elicit IFNγ CD8+ T cells, and therefore, they were not confirmed as true epitopes (Figure 5.2). We speculate that the specific motifs in this region of Nef play an important functional role and I168M peptide variants manage to escape binding to HLA or TCR to avoid CD8+ T cell responses that would otherwise interfere with

viral function. In fact, I168M lies near the ExxxLL domain of Nef, which is an important viral motif involved in cell trafficking pathways(306,307). Therefore, consensus and selected amino acids both appear to evade CD8+ T-cell responses to preserve viral function. Viral fitness assays could also determine whether the I168M PSM can compensate for other mutations in Nef. Further T cell assays, such as flow cytometry could better characterize CD8+ T cell responses to I168M containing peptides.

Overall, this study has shown the validation of 27 out of 40 HLA-restricted CTL epitopes of HIV-1 Nef that contained consensus or PSMs at various anchor positions. The frequency of CD8+ T cells that recognized epitopes with PSM or consensus amino acids associated with different rates of CD4 decline were similar. This indicates that viral mutation is limited to amino acids that are closely related to the consensus sequence to maintain the fitness and function of virus. In addition, our study extends on the current knowledge that de novo CTL reactivity appears with emerging viral variants. The findings from this study help design an effective candidate vaccine, which likely needs to include both consensus and PSM epitopes to generate the optimal CD8+ T cell responses.

5.7 Limitations and Opportunities

Although ELISPOT assay is very sensitive in the detection of CD8+ T cells upon stimulation with peptides, it only measures IFN-γ producing CD8+ T cells. Other cytokines such as IL-2, TNFα and MIP1β are also secreted from CD8+ T cells when

they are stimulated in response to a peptide, which can be measured by two colour ELISPOT assays (e.g. IL-2 and IFNγ detecting ELISPOT). We were only able to carry out one colour ELISPOT assay in this study, and it is likely that peptides that did not result in the induction of IFNγ+CD8+ T cells, could elicit CTLs that express other cytokines. Therefore, we may have missed peptides that could be considered true CD8+ T cell epitopes. In the future, other CD8+ T cell parameters could be detected using two-colour ELISPOT assays or by flow cytometry. Another setback of our study is that for most peptides that correlated with specific HLA class I alleles, PBMCs from only few patients with that allele were available. A larger sample size will help to better confirm epitopes in multiple patients with the same HLA allele. Additionally, among those tested the number of patients who responded to a given epitope were small, ranging from 1 out of 9 to 3 out of 8 (K182M and E63D epitopes respectively). Therefore, the statistical power of this study can be improved with further T cell assays with a larger sample size.

6. Characterization of CD8+ T cell populations targeting two A*02:01 restricted and two B*58:01 restricted HIV-1 Nef epitopes

6.1 Rationale

CD8+ T cells play an important role in viral control and several of their characteristics are associated with protective outcomes(288). These include frequency of antigen-specific CD8+ T cells, proliferation capacity(308,309), polyfunctionality (310), effector memory(105,311), efficient killing of infected CD4+ T cells(287) and the absence of exhaustion(312). It has been shown that proliferation of CD8+ T cells was correlated with decreased viral load and increased CD4+ T cell count(313). Conversely, the expression of exhaustion markers on HIV-specific CD8+ T cells was correlated with disease progression(312,314). It has been shown that HIV-1-specific CD8+ T cells with higher expressions of PD-1 were correlated with impaired CTL responses, lower CD4 count and higher viral load(315). Additionally, elite controllers were shown to have lower levels of PD-1 expression compared to typical disease progressors (316). These studies provide evidence that the quality of CD8+ T cells is an important contributor to effective viral control.

A question that still remains is what are specific targets of CD8+ T cells and which epitopes in HIV-1 are best associated with protective outcomes. Many studies have shown that targeting HIV-1 Gag epitopes is correlated with viral control(180,317,318) and in the recent years Nef-specific CD8+ T cell responses correlated with protective outcome (low viral load) have also been defined(190). In

addition to evidence which shows that viral replication is mediated through CTL targeting of specific epitopes, *in silico* analyses have indicated that viral control may be the result of mutations within epitopes that change stability of protein structure(319). Therefore, mutations within CTL epitopes can have an impact on viral control. We have identified several PSMs in Nef that are associated with either slower or faster CD4 decline compared to their consensus counterpart. Characterization of CD8+ T cell populations restricted by epitopes containing PSM or consensus amino acids can help in defining protective immune targets that contribute to HIV-1 control. In this study, we examined the expression of antiviral intracellular cytokines, proliferation and exhaustion characteristics of the CD8+ T cell populations restricted by two A*02:01 epitopes (ILDLWVYHT and ILDLWVYNT) and two B*58:01 epitopes (KGGLDGLIY and KGGLDGLVY) (refer to section 6.4.2. for epitope selection and tetramer design).

6.2 Hypothesis

CD8+ T cell populations restricted by epitopes containing PSMs associated with differential CD4 decline are different in their proliferation, , expression of antiviral intracellular cytokines and exhaustion characteristics.

6.3 Objective

- To determine and compare frequencies of CD8+ T cell populations targeting peptides that contain PSMs associated with faster or slower CD4 decline in HLA-A*02:01+ and HLA-B*58:01+ Kenyan sex workers

- To determine and compare proliferation and frequency of antiviral intracellular cytokine expressing CD8+ T cell populations targeting peptides that contain PSMs associated with faster or slower CD4 decline in HLA-A*02:01+ and HLA-B*58:01+ Kenyan sex workers
- To determine and compare exhaustion characteristics of CD8+ T cell populations targeting peptides that contain PSMs associated with faster or slower CD4 decline in HLA-A*02:01+ and HLA-B*58:01+ Kenyan sex workers

6.4 Methods

6.4.1 Study Population

Samples used in this study were cryopreserved PBMCs obtained from 11 A*02:01 positive and 10 B*58:01 positive HIV-1 infected patients from the Pumwani sex-worker cohort, Nairobi Kenya. The selection criterion was to include patients that expressed one of the above HLA alleles.

6.4.2 Tetramer selection strategy and design

Tetramers are peptide-HLA complexes that allow for direct *ex vivo* characterization of antigen-specific CD8+ T cell and have commonly been used to characterize HIV-1 specific CD8+ T cell profiles(320). In chapter 5, HLA class I restricted CD8+ T cell Nef epitopes containing PSM or consensus amino acids associated with differential CD4 decline were validated using ELISPOT assay (refer to Table 5.3 for summary of all epitopes confirmed). To design tetramers for flow cytometry experiments two peptides that generated the highest frequency of IFN-γ producing CD8+ T cells,

measured by ELISPOT assays, from each of the CD4 decline categories (beneficial and detrimental) were selected. From the "detrimental" category, I101V peptides, KY9-I (KGGLDGL**I**Y) and KY9-V (KGGLDGL**V**Y), restricted by B*58:01 were selected since they resulted in the highest frequency of IFN-γ producing CD8+ T cells in PBMCs (ML0915, 700 and 685 SFU/10^6 PBMCs, respectively Table 5.3). From the "beneficial" category, the two peptides that generated the highest frequency of IFN-γ producing CD8+ T cells in PBMCs were the H116N peptides, IT9-H (ILDLWVY**H**T) and IT9-N (ILDLWVY**N**T) of A*02:01patients (315 and 445 SFU/10^6 PBMCs in ML 0059, respectively) (Table 5.3).

The selected peptides were further tested by does-dependent ELISPOT assays to confirm binding with their respective alleles at varying peptide concentrations (2-fold dilution ranging 0.625 -10 μg/ml). Both peptide groups demonstrated binding ability with HLA alleles at different concentrations (See chapter 5, Figure 5.2 for dose-dependent ELISPOT responses). This served as evidence of peptide-HLA binding, a criterion required by the National Institute of Health (NIH) Tetramer Core Facility, for design and synthesis of tetramers. All tetramers were provided by the Tetramer Core Facility of NIH in lyophilized format and were diluted to the desired concentration in DMSO (section 2.10.1).

6.4.3 Background subtraction for peptide stimulations

To calculate the frequency of tetramer-specific CD8+ T cells, background levels of unstimulated tetramer-stained negative controls were first subtracted from

tetramer-positive cells. If tetramer staining was twice the background level and >0.01% of the total CD8+ T cells, the response was considered positive. To calculate cytokine responses, background levels were first determined using unstimulated tetramer stained negative controls. If the frequency of tetramer-positive CD8+ T cells was twice the background, the response was considered to be positive for the specific marker.

6.4.4 Statistical Analysis

Statistical analysis included Mann-Whitney test for comparison of cytokines and frequency of tetramer-specific CD8+ T cells. All tests were conducted in GraphPad Prism v7.0a (GraphPad Software, La Jolla, CA) and p value of <0.05 was considered significant.

6.5 Results

6.5.1 *Ex vivo* quantification of tetramer-specific CD8+ T cell populations

To determine whether CD8+ T cells from study participants can recognize HLA A*02:01 and HLA B*58:01 epitopes, the frequency of surface expression of tetramer-positive CD8+ T cells were measured. First, PBMCs from 11 HLA A*02:01 and 10 HLA B*58:01 patients were stained with IT9-H and IT9-N (ILDLWVY**H**T, ILDLWVY**N**T) and KY9-I and KY9-V (KGGLDGL**I**Y, KGGLDGL**V**Y) tetramers, respectively. Then using flow cytometry, the direct *ex vivo* surface quantification of tetramer-positive CD8+ T cells were measured through a series of gating strategies depicted in Figure 6.1. Table 6.1 summarizes the frequency of tetramer-positive CD8+ T cells reported as percentage of total CD8+ T. If tetramer staining was twice the background level and >0.01% of the total CD8+ T cells, the response was considered positive. The frequency of A*02:01 tetramer specific CD8+ T cells ranged from 0.292% (tetramer IT9-H, ML 2280) to 25.689% (tetramer IT9-N, ML 1847). The frequency of B*58:01 tetramer specific CD8+ T cells ranged from 0.133% (tetramer KY9-V, ML 1802) to 3.684% (tetramer KY9-V, ML 2194). PBMCs from all individuals in this study were able to recognize peptides selected for further characterization by flow cytometry. Since these peptide variants are recognizable by CD8+ T cells from HLA A*02:01 and HLA B*58:01 patients, they can be further characterized by flow cytometry analysis, which will be the focus of the remainder of this chapter.

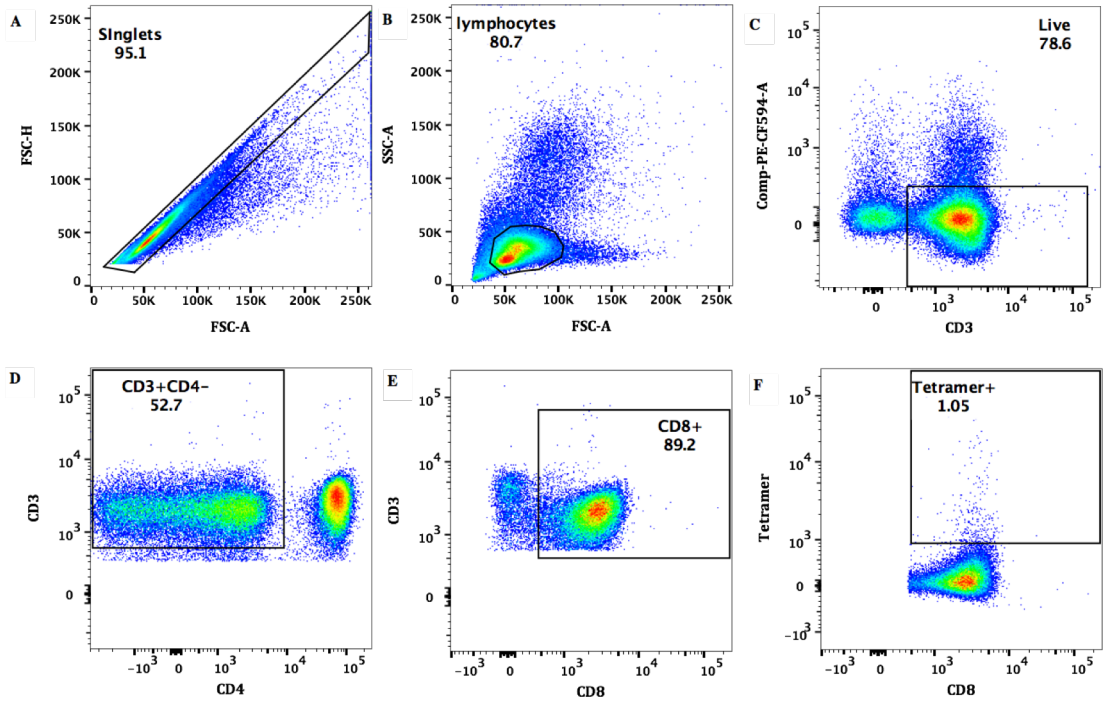


Figure 6.1. Representative gating strategy and screening for *ex vivo* tetramer staining of CD8+ T cells in PBMCs from HLA A*02:01 and HLA B*58:01 patients (example provided is from ML1425, HLA A*02:01 positive patient) A) Singlets are selected by gating on FSC-H vs FSC-A with equal size B) Lymphocyte population of cells are selected on SSC-A vs. FSC-A from singlets C) live cells are gated on D) CD3+ T cells are gated on from the live cell population E) CD8+ T cell are gated on from the CD3+ T cell population. F) Tetramer+ cells are gated on from the CD8+ T cell population (*Numbers shown in top corners, A-C, or within boxes, D-F, represent frequency of the specific cell population that is gated on as the percentage of the total cells in each plot*).

Table 6.1: **Frequency of tetramer positive CD8+ T cells reported as percentage of total CD8+ T cells.** BPMCs from all 11 HLA A*02:01 and 10 HLA B*58:01 patients were stained with respective tetramers either containing consensus or PSM peptides. The frequency of all tetramer specific CD8+ T cells were >0.01% of total CD8+ T cells.

	Patient ID	**Tetramers**	
HLA A*02:01 patients		ILDLWVY**H**T	ILDLWVY**N**T
	1424	0.682	5.651
	1592	1.150	6.569
	1847	4.827	25.689
	1852	1.222	3.119
	1917	1.841	20.911
	2280	0.292	4.282
	2414	2.228	4.838
	2503	0.902	6.969
	2619	5.661	17.566
	2726	6.365	1.733
	3004	1.175	4.707
	Average	2.395	9.276
HLA B*58:01 patients		KGGLDGL**I**Y	KGGLDGL**V**Y
	915	0.148	0.200
	1592	0.984	0.539
	1597	0.773	0.426
	1802	0.363	0.133
	1852	1.074	0.592
	2014	1.390	0.486
	2194	2.165	3.684
	2244	1.990	1.225
	2321	3.025	1.326
	2620	0.307	0.255
	Average	1.222	0.887

6.5.2 The frequency of tetramer-specific CD8+ T cell populations post antigen stimulation

To assess the cellular responsiveness of CD8+ T cells in A*02:01 and B*58:01 patients, PBMCs were first stained with tetramers, then stimulated by peptides. PBMCs from A*02:01+ patients were stimulated with IT9-H and IT9-N (ILDLWVY**H**T and ILDLWVY**N**T) peptides, and PBMCs of B*58:01+ patients were stimulated with KY9-I and KY9-V (KGGLDGL**I**Y and KGGLDGL**V**Y) peptides. After 6 hours of stimulation, the frequency of tetramer-specific CD8+ T cells were measured by flow cytometry (Figure 6.2 shows the representative gating strategy for the detection of tetramer-positive CD8+ T cells.). In A*02:01+ patients, stimulation of PBMCs with ILDLWVY**N**T, resulted in higher frequency of tetramer-specific CD8+ T cells compared to stimulation with ILDLWVY**H**T ($p< 0.0001$, Mann-Whitney) (Figure 6.3 A). The higher frequency of ILDLWVY**N**T-specific CD8+ T cells compared to ILDLWVY**H**T-specific cells may indicate that a larger proportion of CTLs can kill cells that are infected with viruses harboring ILDLWVY**N**T sequences. Interestingly, the PSM "asparagine (N)" at this position, was associated with slower CD4 decline compared to the consensus "histidine (H)" (Figure 4.3). However, there was no significant difference in the frequency of antigen-specific CD8+ T cells in PBMCs stimulated with the two B*58:01 specific peptides, KGGLDGL**I**Y and KGGLDGL**V**Y ($p=0.6445$, Mann-Whitney)(Figure 6.3 B). In this case, the KGGLDGL**V**Y epitopes contains the PSM "valine (V)", which was associated with faster CD4 decline (Figure 4.3), but the proportion of tetramer-specific CD8+ T cells was not affected.

Figure 6.4 shows the comparison of CD4 counts of HLA-A*02:01 and B*58:01 patients at the time of sample collection. There is no significant difference between the two patient

groups with respect to their CD4 counts, which indicates that this is not a factor for the differences observed between the two patient groups.

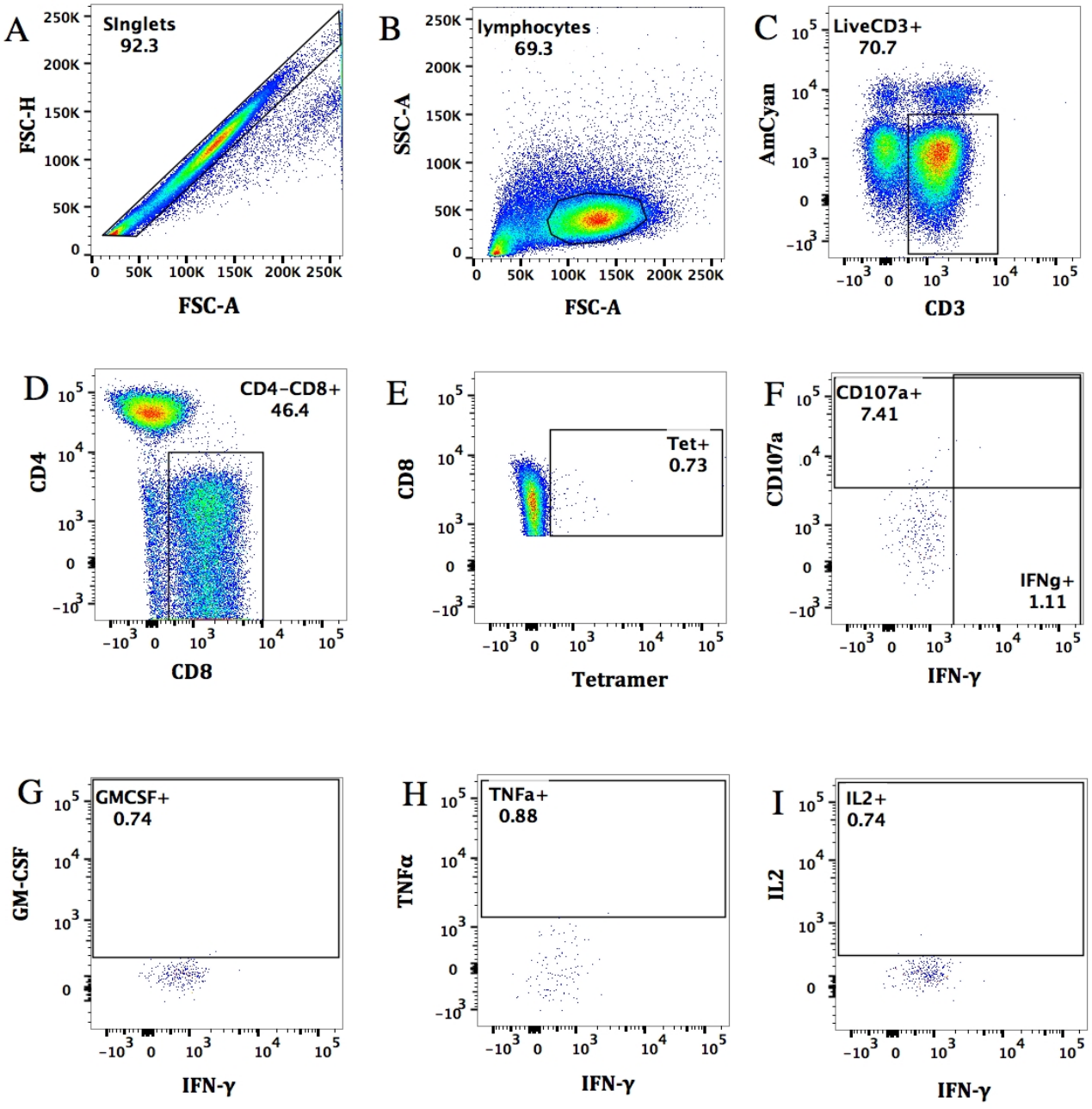


Figure 6.2. Representative gating strategy for characterization of CD8+ T cells in PBMCs from A*02:01 and B*58:01 patients (example provided is from ML1425, A*02:01 positive patient). A) Singlets are selected by gating on FSC-H vs FSC-A with equal size, B) Lymphocyte population of cells are selected on SSC-A vs. FSC-A from singlets, C) live-CD3+ T cells are selected from lymphocytes, D) CD8+ T cells are gated on from live CD3+ T cell population, E) Tetramer+ cells are selected from CD8+ T cells and F-G) shows the gating strategy for tetramer+ CD8+ T cells for CD107a, GM-CSF, TNFα and IL-2 gated against IFNγ, respectively. (*Numbers shown in top corners, A-D, or within boxes, E-F, represent frequency of the specific cell population that is gated on as the percentage of the total cells in each plot*)

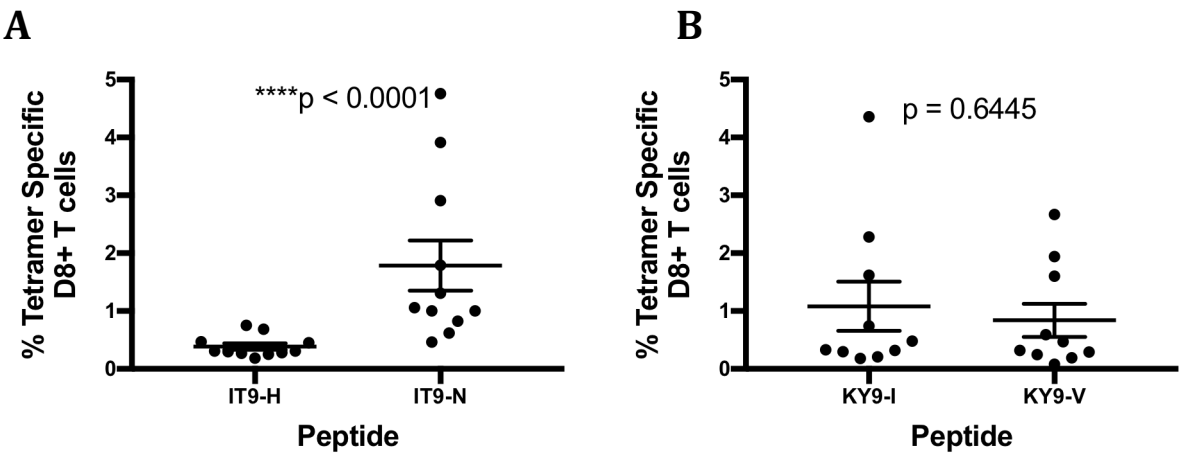


Figure 6.3. Frequency of antigen-specific CD8+ T cells in PBMCs from A*02:01 patients and B*58:01 patients (A and B, respectively). A*02:01 epitope, IT9-N (ILDLWVY**N**T) is more frequently recognized by CD8+ T cells than IT9-H (ILDLWVY**H**T) ($p<0.0001$, Mann-Whitney). There was no difference in the frequency of antigen-specific CD8+ T cells in PBMCs stimulated with B*58:01 peptides, KY9-I (KGGLDGL**I**Y) and KY9-V (KGGLDGL**V**Y).

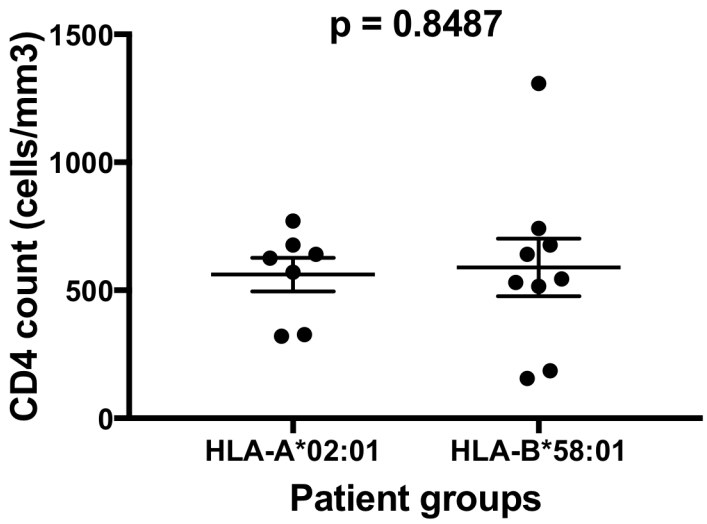


Figure 6.4. Comparison of absolute CD4 counts in A*02:01 and B*58:01 patients. CD4 data represent the counts at the time of specimen collection. Data from four A*02:01 and one B*58:01 patients were missing.

6.5.3. Expression of intracellular markers by CD8+ T cells post antigen stimulation

To determine the antiviral properties of CD8+ T cells after stimulation with peptides containing PSMs or consensus amino acids associated with different rates of disease progression, PBMCs were stimulated by peptide variants (ILDLWVY**H**T/ ILDLWVY**N**T in A*02:01 patients and KGGLDGL**I**Y/ KGGLDGL**V**Y in B*58:01 patients) for 6 hours, followed by measuring the frequency of cytokine secreting total CD8+ T cells using flow cytometry. In HLA A*02:01 patients, there was no significant difference in the frequency of cytokine producing CD8+ T cells when stimulated by IT9- H, associated with faster CD4 decline, or IT9-N, associated with slower CD4 decline (Figure 6.5). Likewise, among B*58:01 patients, we found no significant difference in the frequencies of CD8+ T cells that produce antiviral cytokines, when PBMCs were stimulated by KY9-V, associated with faster CD4 decline and KY9-I, associated with slow CD4 decline (Figure 6.6). Therefore, these results showed that total CD8+ T cells stimulated by different peptide variants, associated with distinct disease progression rates, have similar cytokine producing profiles, when considering CD107a, GM-CSF, IFN-γ, IL-2 and TNFα. Overall, the frequencies of CD8+ T cells expressing different cytokines were low. Because stimulation with PMA/Ionomycin induced high frequencies of TNFα and IL2 producing CD8+ T cells (data not shown), the low expression observed with peptide stimulations may be peptide specific.

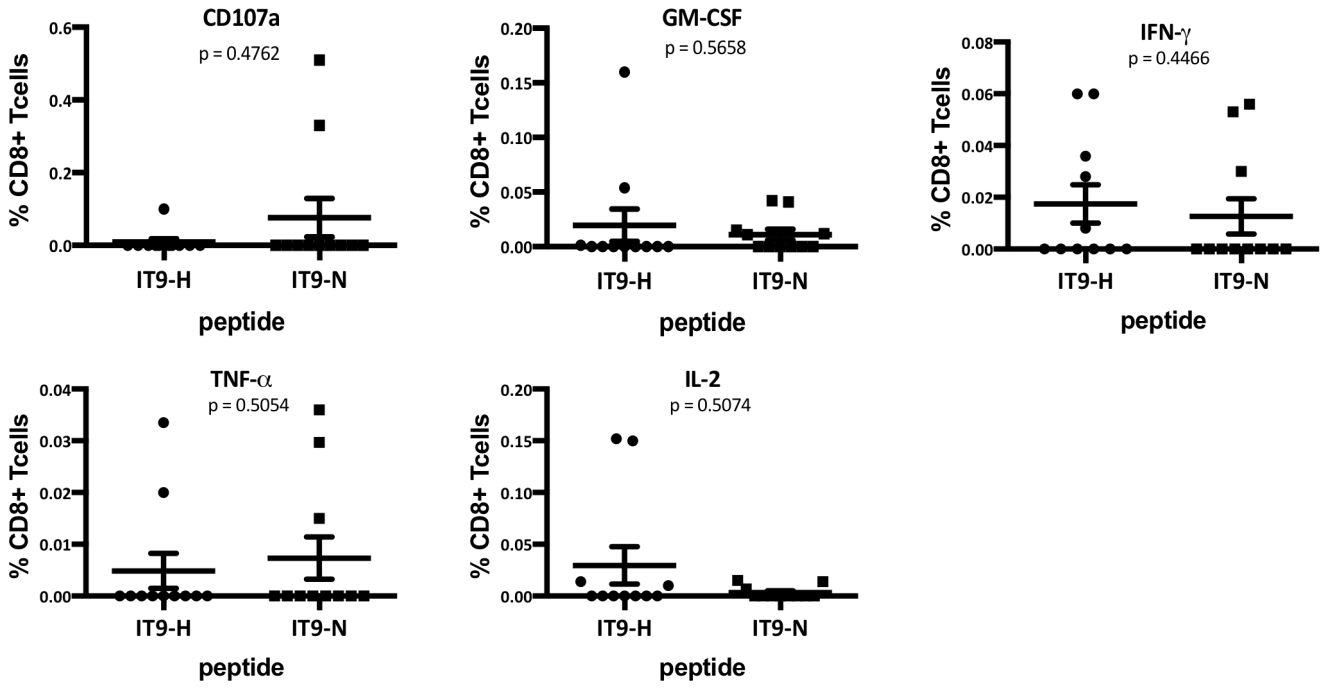


Figure 6.5. Frequency of total CD8+ T cells expressing effector intracellular cytokines in response to stimulation by A*02:01 peptide variants: IT9-H (ILDLWVY**H**T), associated with faster CD4 decline, and IT9-N (ILDLWVY**N**T), associated with slower CD4 decline.

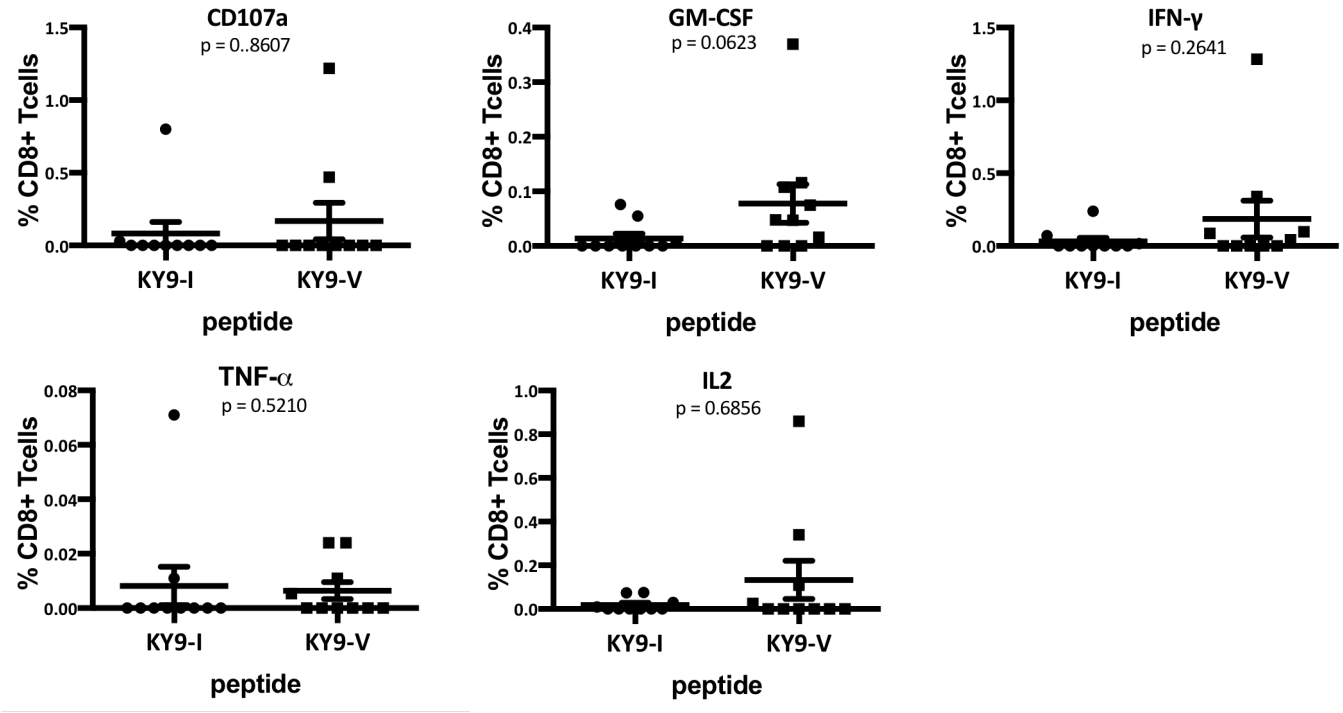


Figure 6.6. Frequency of total CD8+ T cells expressing effector intracellular cytokines in response to stimulation by B*58:01 peptide variants: KY9-I (KGGLDGLIY), associated with slower CD4 decline and KY9-V (KGGLDGLVY), associated with faster CD4 decline.

6.5.4. Expression of intracellular markers by tetramer-specific CD8+ T cells post antigen stimulation

To determine differences in the characteristics of tetramer-specific CD8+ T cells stimulated by peptides that contain amino acids associated with faster or slower CD4 decline, the frequency of CD8+ T cells that produce antiviral intracellular cytokines, such as CD107a, IFN-γ, IL-2 and TNFα were measured and compared in each group. PBMCs were stimulated by peptide variants (ILDLWVY**H**T/ ILDLWVY**N**T in A*02:01 patients and KGGLDGL**I**Y/ KGGLDGL**V**Y in B*58:01 patients) for 6 hours, and cytokine producing tetramer-specific CD8+ T cells were assessed by intracellular cytokine staining using flow cytometry. Figure 6.2 F-G shows a representative gating strategy for each intracellular cytokine. The frequencies of tetramer-specific CD8+ T cells expressing different cytokines after peptide stimulations are shown in Figure 6.7 and Figure 6.8. There were no significant differences in the frequencies of CD8+ T cells expressing CD107a, TNFα, and IL-2 after stimulation with peptide variants, IT9-H and IT9-N (ILDLWVY**H**T/ ILDLWVY**N**T) restricted by A*02:01. A higher frequency of IFNγ expressing CD8+ T cells were observed when PBMCs were stimulated with the IT9-H peptide, which contains amino acid associated with fast CD4 decline, compared to stimulation with IT9-N peptide, associated with slower decline (p= 0.0045, Mann-Whitney test) (Figure 6.7B). Similarly, stimulation of PBMCs from B*58:01+ patients with KY9-I and KY9-V peptides, with consensus or PSM amino acids associated with different CD4 decline rates, did not result in significant difference in the frequency of antigen-specific CD8+ T cells expressing intracellular markers (Figure 6.8 A-D).

With low frequencies of IL2-expressing tetramer-specific CD8+ T cells (Figures 6.7C and 6.8C), we also used Ki67, a nuclear protein, which is used indirectly as a marker of proliferation(321), to determine if tetramer-specific CD8+ T cells proliferate differently, when stimulated by peptides harbouring consensus or PSM amino acids associated with differential CD4 decline. Ki67 is present during the active phase of the cell cycle but it is not detected when cells are at their resting stage(322). Therefore, recently they have been used as a surrogate marker for proliferation to differentiate dividing and non-dividing CD8+ T cells(323,324). In both A*02:01 or B*58:01 patient groups, we found no significant differences in the frequency of tetramer-positive CD8+ T cells that express Ki67, when comparing *ex-vivo* expression by tetramer groups that associate differently with CD4 decline (Figure 6.9 A-B). This indicates that CD8+ T cells that recognize Nef tetramers, with mutations associated with different rates of CD4 decline, are not different in their proliferation ability. Overall, these results do not agree with the study hypothesis that peptides associated with beneficial clinical outcomes induce higher frequency of CD8+ T cells that express antiviral and proliferation markers.

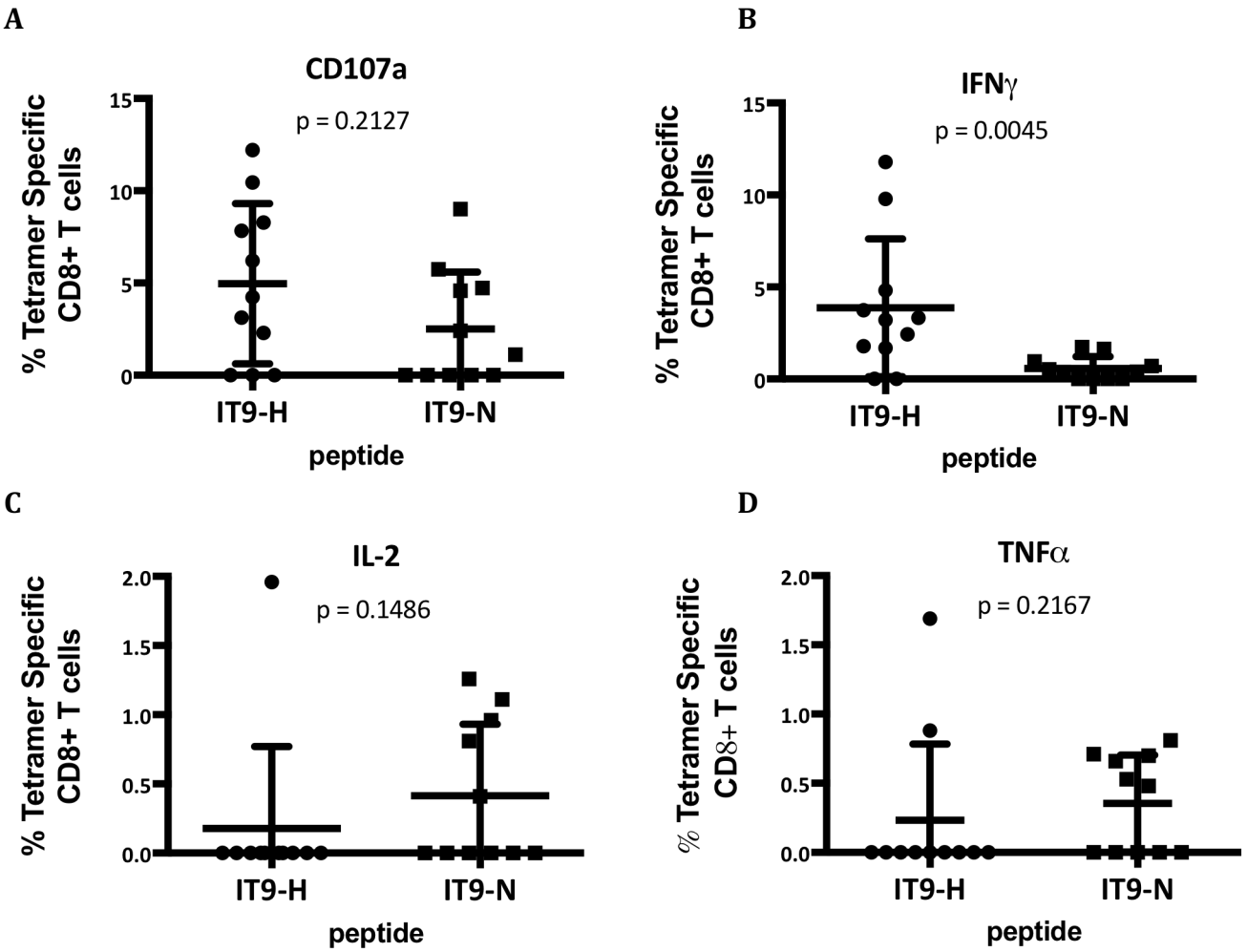


Figure 6.7. Frequency of tetramer-specific CD8+ T cells expressing effector intracellular cytokines: CD107a (A), IFNγ (B), IL-2 (C) and TNFα (D) in A*02:01 PBMCs after stimulation by IT9-H (ILDLWVYHT), associated with faster CD4 decline, and IT9-N (ILDLWVYNT), associated with slower CD4 decline.

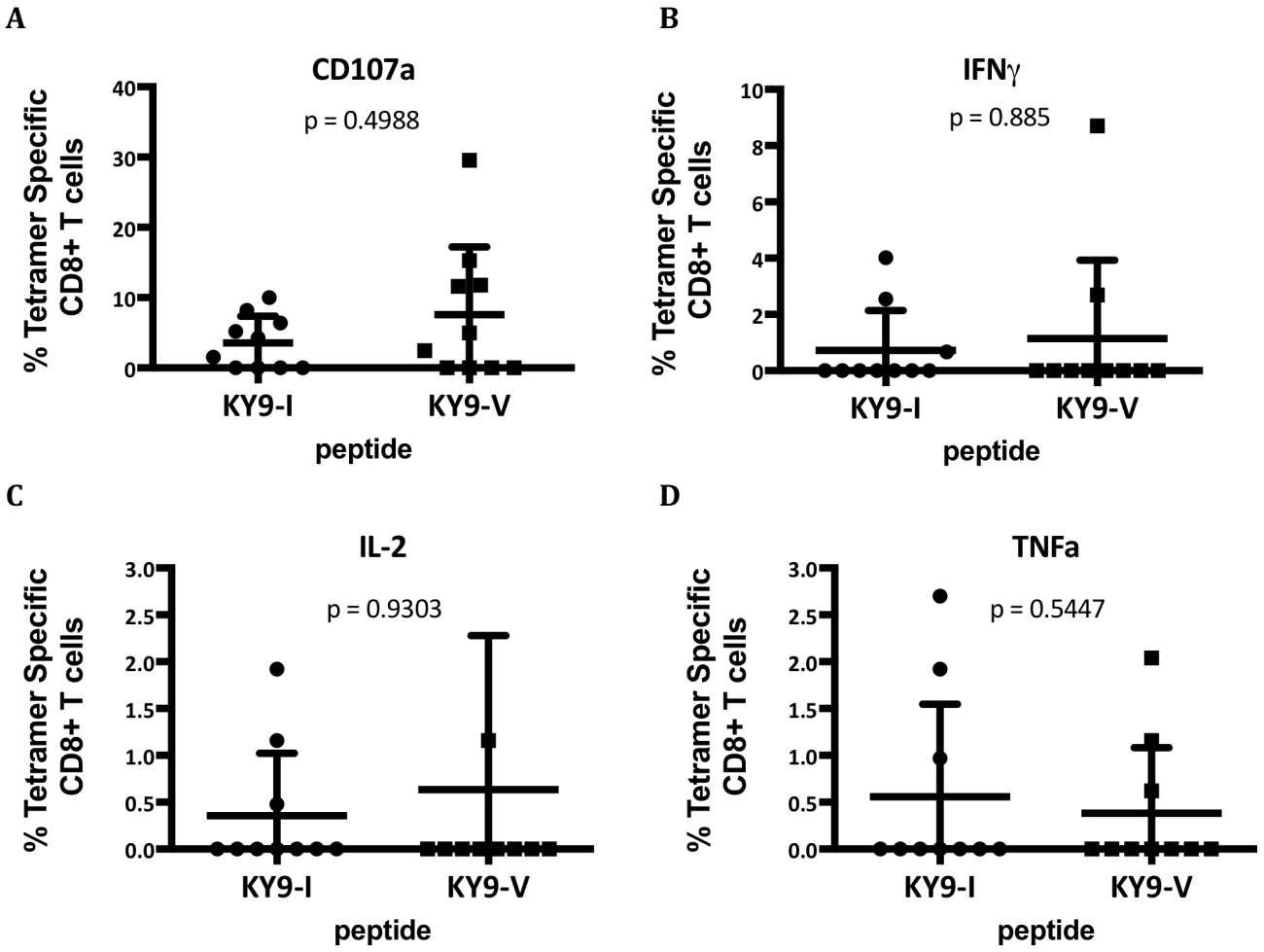


Figure 6.8. Frequency of tetramer-specific CD8+ T cells expressing effector intracellular cytokines: CD107a (A), IFNγ (B), IL-2 (C) and TNFα (D) in B*58:01 PBMCs after stimulation by peptide variants, KY9-I (KGGLDGLIY), associated with slower CD4 decline and KY9-V (KGGLDGLVY), associated with faster CD4 decline.

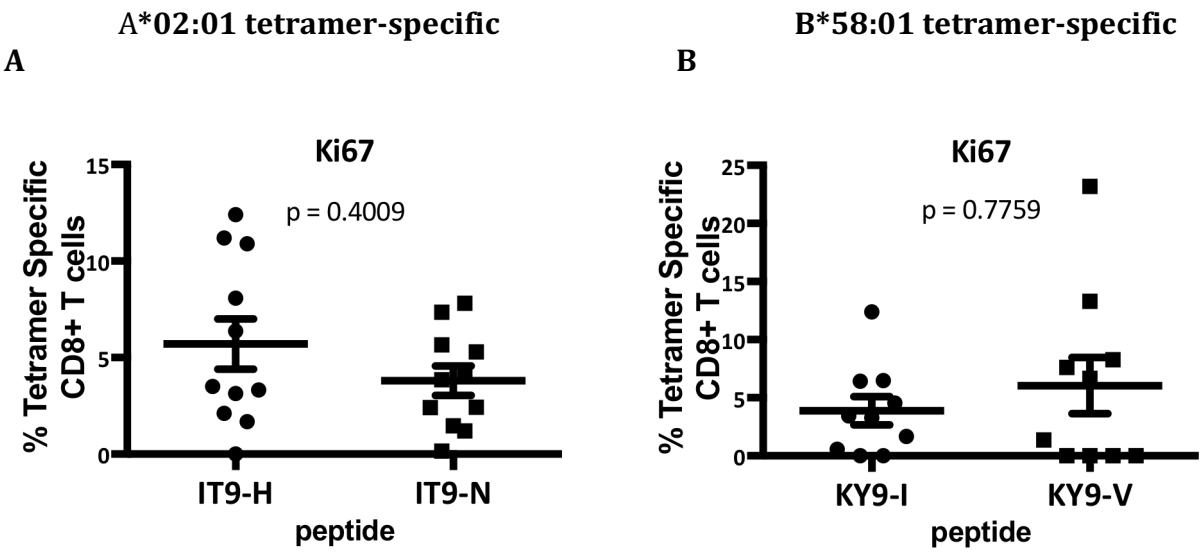


Figure 6.9. Frequency of Ki67+ tetramer-specific CD8+ T cells in A*02:01 and B*58:01 patients. PBMCs were stained with tetramers followed by staining with Ki67 marker. There was no significant difference in the frequency of Ki67+ CD8+ T cells specific to IT9-H (ILDLWVYHT), associated with faster CD4 decline, and IT9-N (ILDLWVYNT), associated with slower CD4 decline (A) or to KY9-I (KGGLDGLIY), associated with slower CD4 decline and KY9-V (KGGLDGLVY), associated with faster CD4 decline (B).

6.5.5. Ex-vivo expression of exhaustion markers by tetramer-specific CD8+ T cells

Surface expression of inhibitory molecules such as CD160, PD-1 and Tim3 has been shown to define virus-specific CD8+ T-cell exhaustion and immune dysfunction in HIV infection (312). Since peptides with consensus or PSM amino acids were found to associate differently with CD4 decline rates, we wanted to determine whether CD8+ T cell exhaustion, a feature of rapid disease progression (325), was epitope-specific. To do this, the frequencies of CD8+ T cells that express CD160, PD-1 and Tim3 were measured and compared between IT9-H and IT9-N-specific CD8+ T cells in PBMCs of A*02:01 patients, and between KY9-I and KY9-V-specific CD8+ T cells in PBMCs of B*58:01 patients. Our data shows that there are no significant differences in the frequency of tetramer-positive CD8+ T cells that express CD160, PD-1 and Tim3, when comparing *ex vivo* expression by tetramer groups that associate differently with CD4 decline, in neither A*02:01 or B*58:01 patient groups (Figure 6.11 A-B). Therefore, epitopes with consensus or PSM amino acids that associated with faster CD4 decline, do not seem to be recognized more frequently by exhausted CD8+ T cells compared to those associated with slower CD4 decline. Figure 6.10 shows the gating strategy for the tetramer-specific CD8+ T cells with various surface exhaustion markers.

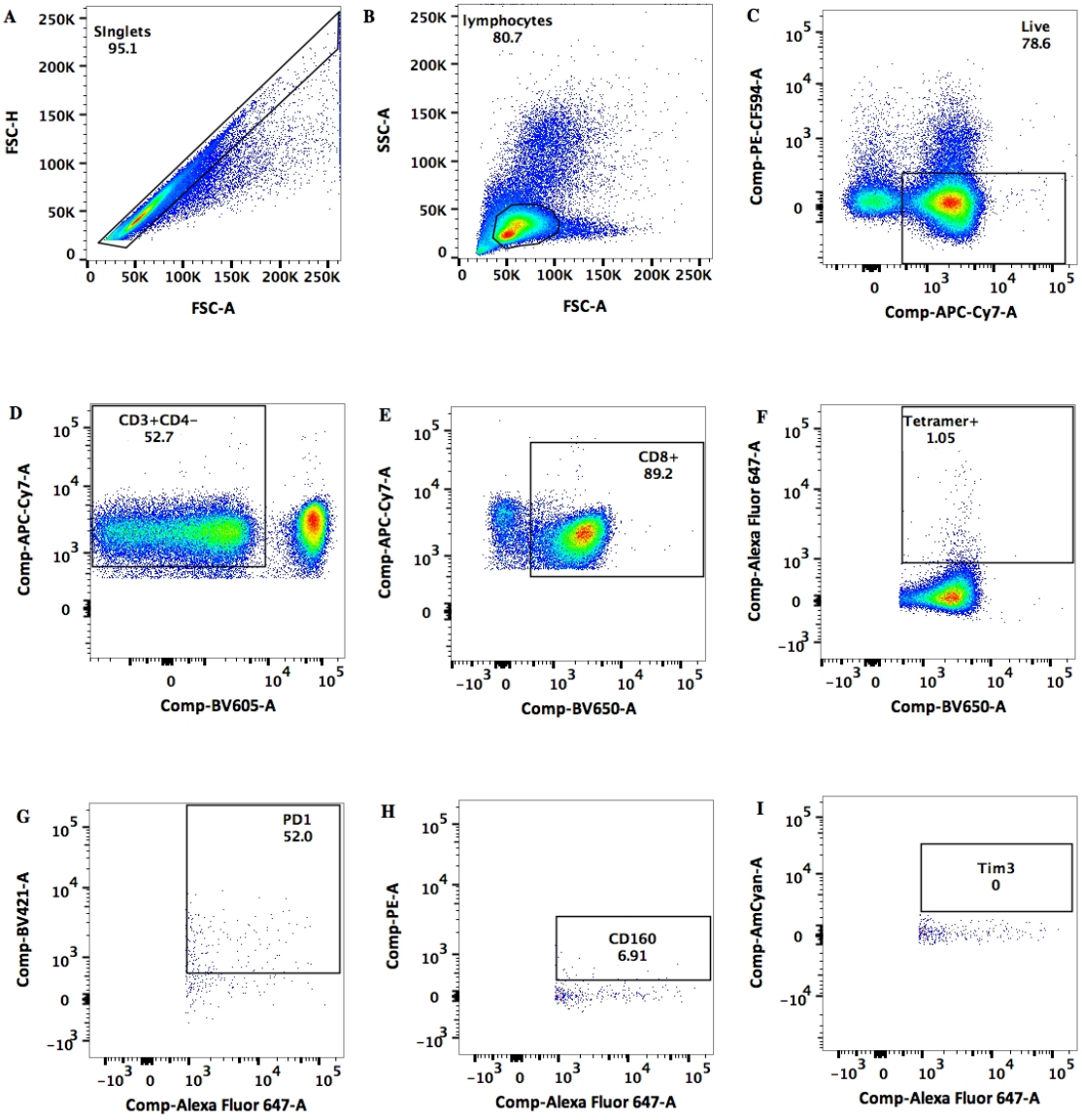
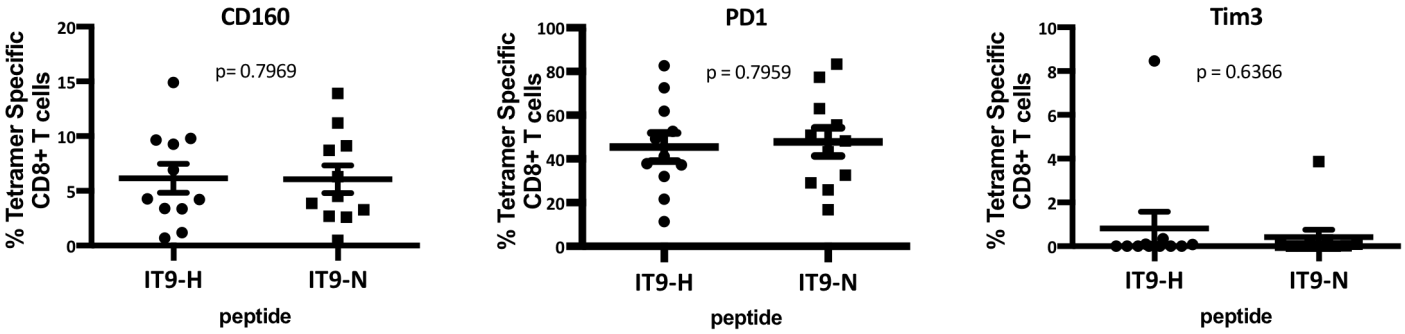


Figure 6.10. Representative gating strategy for characterization of tetramer-specific CD8+ T cells that express surface exhaustion markers A) Singlets are selected by gating on FSC-H vs FSC-A with equal size B) Lymphocyte population of cells are selected on SSC-A vs. FSC-A from singlets C) live CD3+ T cells are gated D) CD3+CD4- T cells are gated on from the live lymphocyte population E) CD8+ T cell are gated on from the CD3+CD4- T cell population F) tetramer+ CD8+ T cell are selected G-I) PD1, CD160 and Tim3 are gated within the tetramer+CD8+ T cell population. (*Numbers shown in top corners, A-C, or within boxes, D-I, represent frequency of the specific cell population that is gated on as the percentage of the total cells in each plot*)

A) HLA-A*02:01 tetramers



B) HLA-B*58:01 tetramers

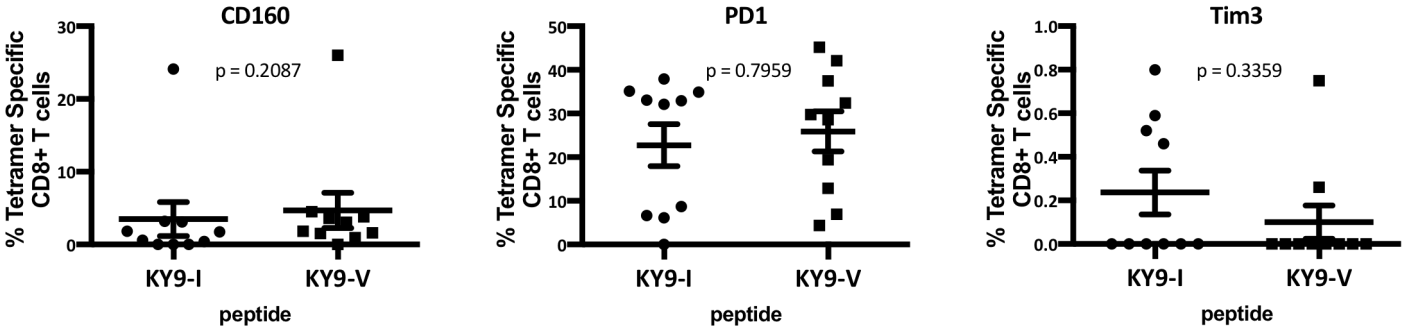


Figure 6.11. Frequency of tetramer-specific CD8+ T cells in A*02:01 and B*58:01 patients that express exhaustion markers, CD160, PD-1 and Tim 3. Cells were stained with tetramers followed by staining with surface markers. For each parameter tested, there were no significant differences in the frequency of CD8+ T cells specific to IT9-H (ILDLWVYHT), associated with faster CD4 decline, and IT9-N (ILDLWVYNT), associated with slower CD4 decline (A) or to KY9-I (KGGLDGLIY), associated with slower CD4 decline and KY9-V (KGGLDGLVY), associated with faster CD4 decline (B) (Mann-Whitney test). *Error bars are displayed as SEM*

6.5.6 Ex-vivo memory phenotyping of tetramer-specific CD8+ T cells in A*02:01 and B*58:01 patients

Memory phenotype of CD8+ T cells can be defined using surface markers such as CCR7 and CD45RA. These phenotypes include naïve (N), central memory (T_{CM}), effector memory (T_{EM}) and terminally differentiated effector memory (T_{EMRA}). The various stages of CD8+ T cell differentiation may indicate their effector function or lack thereof(326-328). Mature and differentiated CD8+ T cells such as T_{EM} and T_{EMRA} with effector function have been shown to play a role in viremia control and reduced viral set point(329,330), whereas less differentiated cells like naïve and T_{CM} have slower effector activity. Since peptides with consensus or PSM amino acids were found to associate differently with CD4 decline, surface memory phenotyping of CD8+ T cells were characterized to determine if CD8+ T cell differentiation was epitope-specific. To do this, CCR7 and CD45RA surface markers were used to characterize the memory phenotype and subset of CD8+ T cells specific for A*02:01 and B*58:01 tetramers, harboring consensus or PSM amino acids associated with varying CD4 decline rates. Figures 6.12 and 6.13 summarize the results for this analysis. All four phenotypes were observed in IT9-H (associated with faster CD4 decline) and IT9-N (associated with slower CD4 decline)-specific CD8+ T cells of A*02:01 patients. For both tetramers a greater proportion of CD8+ T cells belonged to T_{EM} category (CCR7-CD45RA-; 50.6%, IT9-H and 59%, IT9-N), followed by T_{EMRA} cells (CCR7-CD45RA+; 24.1%, IT9-H and 27.6%, IT9-N), naïve (CCR7+CD45RA+; 20.4%, IT9-H and 9.9%, IT9-N) and T_{CM} (CCR7+CD45RA-; 4.9%, IT9-H and 3.2%, IT9-N) (Figure 6.12). There was a higher frequency of naïve CD8+ T cells specific for IT9-H compared to those specific for IT9-N ($p = 0.0233$, Kolmogorov-Smirnov test). The lower frequency of naïve IT9-N-specific CD8+ T

cells concomitant with higher frequency of T_{EM}-and T_{EMRA}-IT9-N cells, although not statistically significant (Figure 6.12), may support slow disease progression associated with this epitope. Figure 6.13 summarizes the memory phenotype results for B*58:01 tetramer-specific CD8+ T cells. Like A*02:01 restricted CD8+ T cells, all four memory phenotypes were observed in B*58:01 cells. However, majority of cells were naïve (CCR7+CD45RA+; 40.8%, KY9-I and 47.2%, KY9-V). There was a higher proportion of T_{EM} cells (CCR7-CD45RA-) specific for KY9-I (29.7%) compared to KY9-V (17.2%), although this difference was not significant (p= 0.7591, Kolmogorov-Smirnov test). The remaining phenotypes included T_{EMRA} (23.5%, KY9-I and 25.8%, KY9-V) followed by T_{CM} (6.0%, KY9-I and 9.7%, KY9-V).

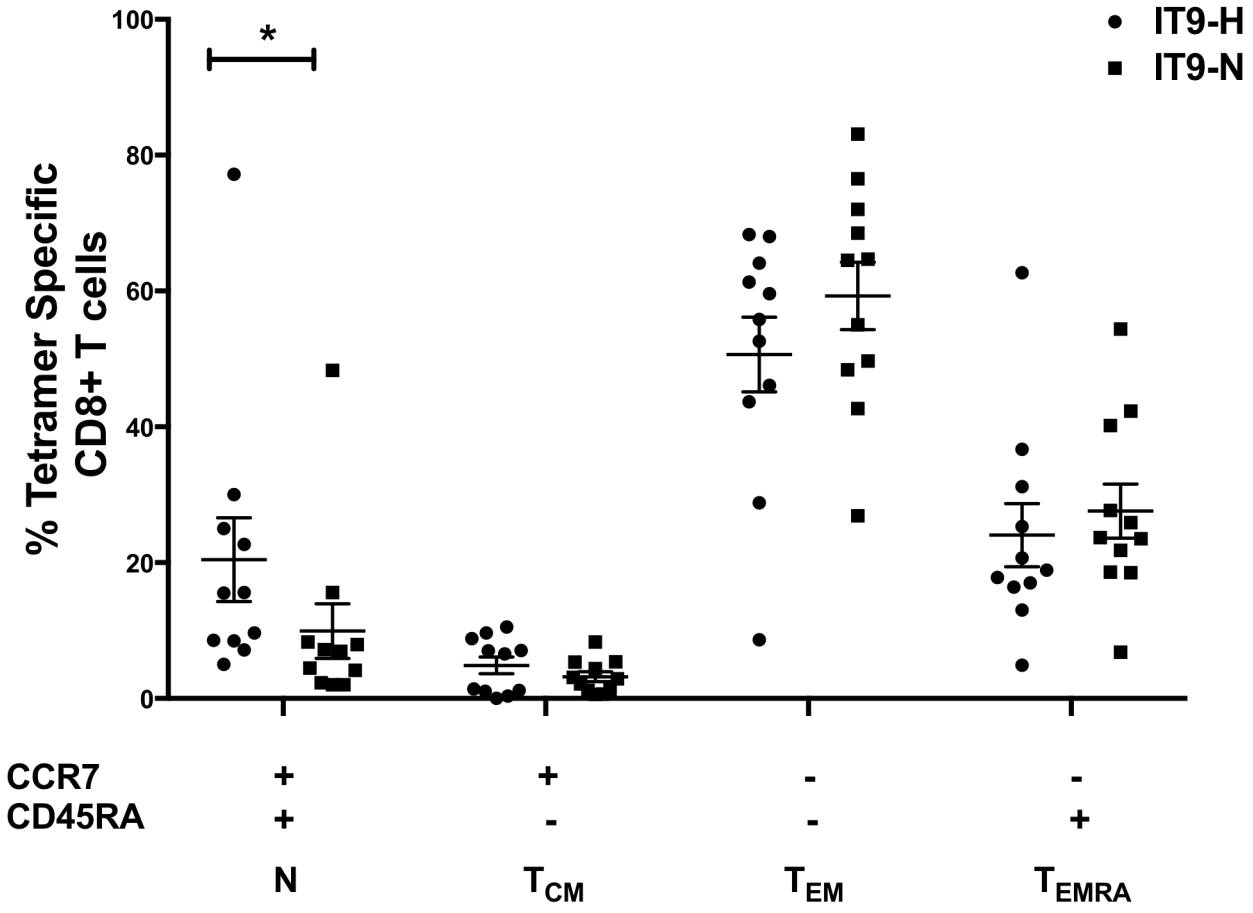


Figure 6.12. Expression of four memory phenotypic categories on tetramer-specific CD8+ T cells of A*02:01 patients. Cells were stained with tetramers, followed by staining with surface markers. Comparison of IT9-H and IT9-N specific CD8+ T cells in PBMCs of A*02:01 patients did not reveal any significant difference except in the proportion of naïve CD8+ T cells which were higher in the group specific for IT9-H tetramer (p= 0.0233, Kolmogorov-Smirnov test). * p-value < 0.05.

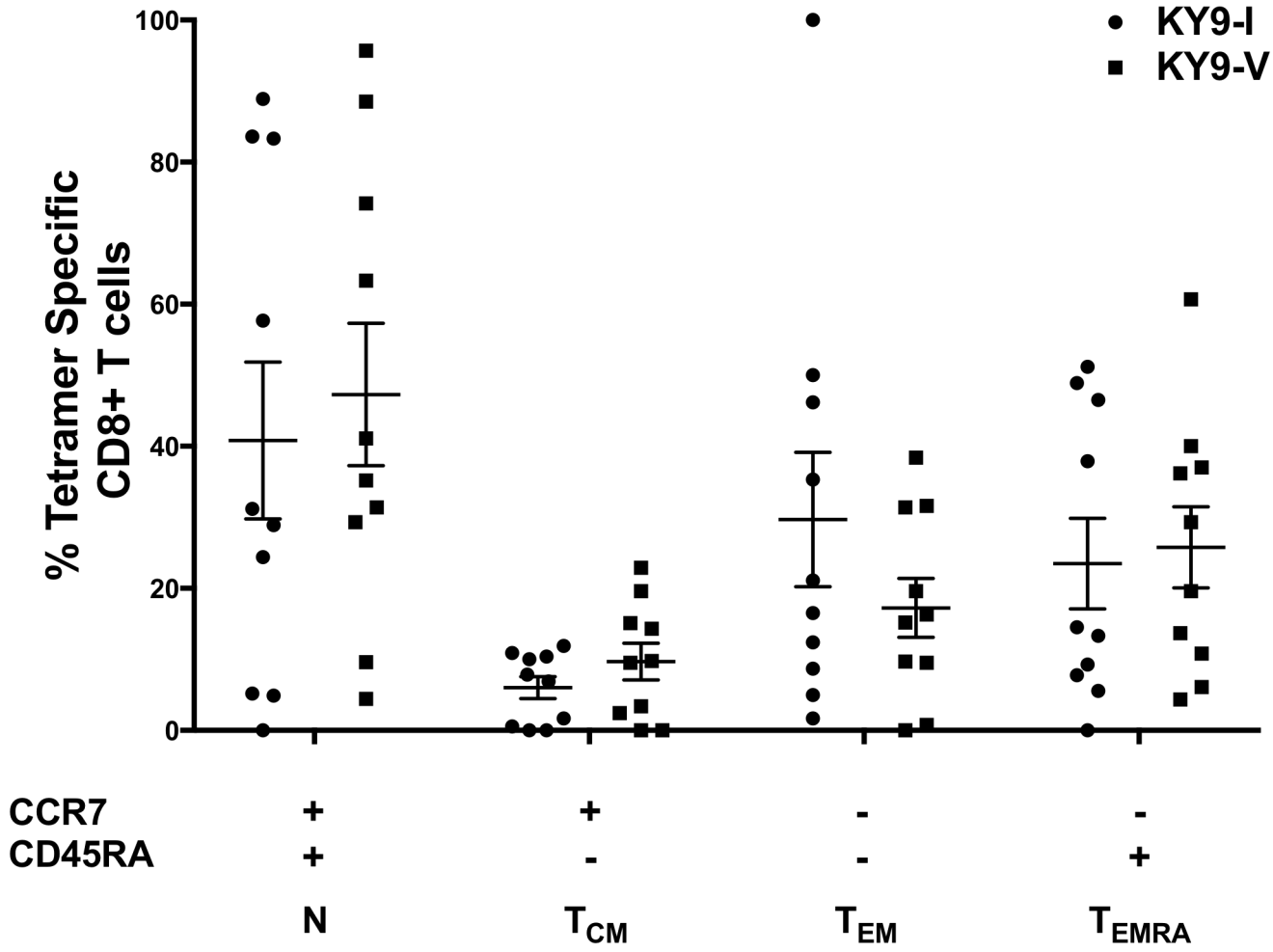


Figure 6.13. Expression of four memory phenotypic categories on tetramer specific CD8+ T cells of B*58:01 patients. Cells were stained with tetramers, followed by staining with surface markers. Comparison of KY9-I and KY9-V specific CD8+ T cells in PBMCs of B*58:01 patients did not reveal any significant difference.

6.6 Discussion:

In this study, we examined and compared the frequency of antigen-specific CD8+ T cells, their antiviral intracellular cytokines and proliferation characteristics when stimulated with two A*02:01-specific or B*58:01-specific peptide variants containing consensus and PSM amino acids of Nef associated with different rates of CD4 decline. We observed a significant difference between the frequencies of CD8+ T cells stimulated with A*02:01-specific peptides, ILDLWVY**H**T(IT9-H) and ILDLWVY**N**T (IT9-N), with higher frequency of CD8+ T cells recognizing the PSM containing epitope (ILDLWVY**N**T). We showed previously that in this epitope, ILDLWVY**N**T, amino acid asparagine was a PSM that correlated with slower CD4 T cell decline. The higher frequency CD8+ T cell recognition of this epitope (ILDLWVY**N**T) compared to the consensus (ILDLWVY**H**T) may explain the better clinical outcome associated with this PSM. It is possible that more CD8+ T cells recognizing this epitope may lead to more CTL killing of infected cells. Polymorphisms in Nef that are associated with better disease outcomes for the infected host have previously been reported. A study of a cohort of South African Clade C infected individuals showed that Nef mutations targeted by HLA-B alleles were associated with greater fitness cost to the virus and lower viral set-point(190). Several other studies have demonstrated that targeting specific regions of the virus including Nef may contribute to viremia suppression(196,277,288,331). Additionally, protection may be linked to the number of CD8+ T cell responses restricted by certain HLA alleles and that frequency of virus-specific CD8+ T cell responses may play a role in determining disease progression(332). Therefore, the immunological benefits of targeting ILDLWVY**N**T, i.e. robust recognition by a higher

number of CD8+ T cells, could be one of the factors contributing to the PSM in this epitope being associated with favorable clinical outcomes.

Unlike A*02:01 Nef epitopes, no difference was observed in the frequencies of CD8+ T cells recognizing the two B*58:01 epitopes, KGGLDGL**I**Y (consensus) and the KGGLDGL**V**Y (PSM). In this case the PSM valine is associated with faster CD4 decline than the consensus, isoleucine. The lack of difference in the frequencies of epitope-specific CD8+ T cells may indicate that HLA-peptide binding and subsequent presentation to TCR is not affected by the PSM in KGGLDGL**V**Y epitope. Since there is no significant difference between the frequencies of CD8+ T cells that target these two B*58:01 epitopes, the I101V mutation in KGGLDGL**V**Y epitope could represents a "true" viral escape that leads to rapid CD4 decline and poor clinical outcomes.

We also investigated the frequency of tetramer-specific CD8+ T cell populations that expressed intracellular cytokines, CD107a, GM-CSF, IFN-γ, IL-2 and TNFα, after stimulation with A*02:01 and B*58:01 specific peptides that associated differently with disease progression. We found that among both A*02:01 and B*58:01 groups, the frequency of epitope-specific CD8+ T cells that expressed intracellular cytokines were not significantly different despite stimulation with peptides associated with slower or faster CD4 decline. Although, in PBMCs from A*02:01 patients the frequency of IFN-γ+ CD8+ T cells that targeted ILDLWVY**H**T (consensus epitope associated with faster CD4 decline) was higher than those targeting ILDLWVY**N**T (PSM epitope association with slower CD4 decline). IFN-γ plays an important role in CTL antiviral activities and contributes to the differentiation of

memory CD8+ T cells(333). The high frequency of IFN-γ+ CD8+ T cells that are specific to ILDLWVY**H**T, associated with poor clinical outcomes, may be surprising. However, despite this difference higher killing of infected cells may be mediated by a larger pool of tetramer-specific CD8+ T cells, which was not observed against this epitope. Therefore, we postulate that even with higher levels of IFN-γ+ CD8+ T cells, CTL killing may not be effective due to a lower frequency of CD8+ T cells that target ILDLWVY**H**T. Overall, there was no significant difference in the frequency of intracellular cytokine+ tetramer-specific CD8+ T cells that targeted epitopes containing amino acids associated with slow CD4 decline (ILDLWVY**N**T and KGGLDGL**I**Y) or with rapid CD4 decline (ILDLWVY**H**T and KGGLDGL**V**Y) in PBMCs of A*02:01 and B*58:01 patients. This contradicts with our hypothesis that a higher proportion of CD8+ T cells that produce antiviral intracellular cytokines target epitopes associated with slower disease progression. Therefore, the slower CD4 decline associated with the PSM in ILDLWVY**N**T can be attributed to higher frequency of CD8+ T cells that recognize this epitope and not to effector cytokines of CD8+ T cells targeting it. The observed pattern of higher proportion (although not significant) of T_{EM} and T_{EMRA} CD8+ T cells in A*02:01+ patients that recognize this epitope may also be related to better viral control and the observed slow disease progression. More detailed memory phenotyping of tetramer-specific CD8+ T cells that take into account the expression of other markers, such as CD27, CD28 and CD62L are needed to better determine this relationship.

CD8 T cell polyfunctionality, the ability to produce several cytokines in response to stimulation by an antigen, has been defined as a protective feature of these cells(103,300,334). In HIV-1 infection CD8 T cell polyfunctionality has been shown to play a

role in viremia control in LTNPs(335). Further, slow progression to AIDS is associated with polyfunctional profile in CD8 T cells(336). In another study, polyfunctional Gag-specific responses were associated with low viral load, whereas oligofunctional responses were associated with poor viral control (337). In our study majority of tetramer-specific CD8+ T cells were oligofunctional for both PSM and consensus epitopes of A*02:01 and B*58:01 and there was no significant difference in the polyfunctionality of CD8+ T cells specific to epitopes associated with differential CD4 decline (data not shown.)

The lack of difference in cytokine production between consensus and PSM epitopes has recently been demonstrated in another study. Stimulation by two HIV-1 Gag epitopes, APPEESFRS (AS9) and APPEESFRF (AF9), containing consensus and variant amino acids respectively, targeted by B*35:01 allele, resulted in similar frequencies of functional CD8+ T cells(338). It can be speculated that the strength of binding between the two stimulating epitopes and the HLA allele is not unique to each peptide, therefore, leading to similar CD8+ T cell characteristics. In fact, in the same study, docking stimulation assays showed that HLA-peptide binding energies for the consensus and variant Gag epitopes were similar. Therefore, regardless of amino acid substitution, peptide presentation to TCR via the MCH molecule does not seem to be affected, resulting in the subsequent parallel CD8+ T cell profiles. This is consistent with the observation of our study, that there were no significant differences in the frequency of tetramer-specific CD8+ T cells expressing intracellular cytokines in response to stimulation with different peptides. Although the data from our analysis does not support our hypothesis that stimulation with epitopes correlated with slow CD4 decline, result in higher frequency of functional CD8+ T cell, we

speculate that the similar amino acid properties between the consensus and PSM epitopes result in recognition by TCR of CD8+ T cells that have similar antiviral properties.

In general, we observed low frequency of tetramer-specific CD8+ T cells that expressed intracellular cytokines when PBMCs were stimulated with Nef specific peptides. There may be several reasons for this observation. First, hypo-responsiveness of CD8+ T cells, also known as anergy, may be a factor in low CD8+ T cell cytokine production(105,339). Anergy is often accompanied by low IL-2 production and incomplete activation of CD8+ T cells(339). The low frequencies of IL-2 producing tetramer-specific CD8+ T cells demonstrated by our data suggest that these cells may in fact be unresponsive to stimulation by Nef peptides. Second, low frequency of CD8+ T cells that express intracellular cytokines, may be due to the inability of Nef peptides to induce such responses since stimulation by PMA/ionomycin resulted in measurable expression of cytokines. Better optimization techniques will be needed to further investigate these factors.

The lack of difference in the exhaustion and functional characteristics between A*02:01-IT9-N specific and A*02:01-IT9H specific CD8+ T cells, and between B*58:01-KY9-I specific and B*58:01-KY9-V specific CD8+ T cells, may be due to differences in virological factors. The immune pressure exerted by the HLA class I restricted CD8+ T cells and viral fitness are two important factors when assessing PSMs associated with differential disease progression. The slower disease progression observed in association with amino acids of certain epitopes in our study (IT9-N and KY9-I) is likely due to a weakened or fitness-

compromised virus and cellular immunity may not be the only factor in determining the observed clinical outcome differences. A functionally unfit virus has a lower replication capacity, infecting fewer target cells and therefore, resulting in more favorable outcomes to the host, one of which might be slower CD4 decline.

6.7 Limitations and Opportunities

One major limitation of this study was the small number of patients for each HLA allele group. The tetramer-specific CD8+ T cell populations could be better characterized with a larger sample size and more optimized procedures.

The data from this study showed that A*02:01-IT9-N specific CD8 T cells are more frequent than A*02:01-IT9H specific CD8 T cells. However, the frequency of B*58:01-KY9-I and B*58:01-KY9-V restricted CD8+ T cells were not significantly different. There is no difference in the other functional characteristics between A*02:01-IT9-N specific and A*02:01-IT9H specific CD8 T cell populations, and between B*58:01-KY9-I specific and B*58:01-KY9-V specific CD8+ T cells. Thus, the differential CD4 decline cannot be explained by CD8+ T cell killing. Because both host and viral factors influence disease progression it is possible that the virus carrying PSM (IT9-N) or consensus (KY9-I) associated with slower CD4 decline is less fit than the virus with IT9H or KY9-V. This can be tested using viral replication fitness assays (178,212).

7. Final Discussion

7.1 Major Findings of the thesis

This study investigated HIV-1 and host interactions by a) identifying PSMs through analysis of HIV-1 Nef sequences from Los Alamos National Laboratory public database and from 326 subtype A1 HIV-1 Nef sequences from treatment naïve patients of a female sex-worker cohort in Nairobi, Kenya, b) correlating the identified PSMs with patient CD4 profile and HLA class I and c) characterizing the antigen-specific CD8+ T cell responses to selected PSMs associated with different rates of CD4 decline. The major findings of this research work are listed below:

- Among major subtypes of HIV-1 the mutational patterns are consistent within major functional motifs of HIV-1 Nef and substitutions at each residue are limited to one or two closely related amino acids that share similar chemical or physical properties. The change of PSMs in each subtype over time displays a shift back and forth between a limited numbers of amino acids that sometimes represent the ancestral sequences.
- HIV-1 Nef protein contains PSMs in both the variable and the conserved regions, which have CD8 T cell epitopes restricted by several HLA class I alleles.
- Several PSMs in Nef were associated with different rates of disease progression. Three PSMs were associated with faster CD4 decline (E63D, I101V, I168M) and 2 PSMs were associated with slower CD4 decline (H116N and K182M).
- Each PSM was correlated with at least one HLA class I allele common in the Kenyan population. HLA alleles associated with protective outcome(300,304) restrict Nef

epitopes containing PSMs associated with slower disease progression (e.g. B*57:03 and H116N). HLA A*02:01, the Class I allele with the highest prevalence in Kenya(222), restricts epitopes containing PSMs associated with slower CD4 decline, *and* epitopes containing PSMs associated with faster CD4 decline.

- There was no difference between the frequency of IFNγ+ CD8+ T cells stimulated with peptides containing consensus or PSM amino acids, irrespective of anchor position and differential associations with CD4 count decline.
- CD8+ T cells restricted by A*02:01-ILDLWVY**N**T (with PSM N associated with slower CD4 decline) are significantly more frequent than those restricted by A*02:01-ILDLWVY**H**T (with consensus H associated with faster CD4 decline). There is no significant difference in the frequency of CD8+ T cells that target the two B*58:01 epitope variants, even though one of these epitopes contains PSM associated with faster CD4 decline.
- There is no difference in the frequency of epitope-specific CD8+ T cells with respect to antiviral intracellular cytokines, proliferation, and exhaustion characteristics when comparing epitopes containing PSM or consensus amino acids that associate differently with CD4 decline rates.

Taken together, these findings demonstrated that: a) host HLA restricted CD8+ T cell responses and viral functional constraints shape and limit HIV evolution; b) PSMs driven by HLA class I restricted CD8+ T cell responses with similar functionality can be associated with different disease outcomes; c) CD8+ T cells restricted by the same HLA class I allele can drive PSMs with opposite effect on disease progression. These findings extend the

understandings of HIV-1 and host interactions and viral evolution in the context of host immunity and can contribute to the design of therapeutic strategy against HIV-1.

7.2 Implications for rational design of an effective vaccine

Traditional vaccine approaches including those using whole pathogens, live-attenuated or inactivated full pathogens have been successful in preventing infection from viruses such as polio, measles, and mumps (60,340). However, these approaches cannot be used for vaccines against HIV-1 due to safety reasons(60,341). Current HIV candidate vaccines tested have been designed to evoke strong and broad cellular immune responses (342) and/or neutralizing antibodies to block or abort infection (68,343,344). However, studies have shown that the generated strong and broad spectrum cellular immune responses had no effect in preventing HIV infection in clinical trials(63,345) and neutralizing antibodies were not the correlate of protection of the RV144 clinical trial which presented modest efficacy(346-348).

Studies have shown that viral mutations driven by CD8+ T cell responses restricted by HLA class I alleles can have an impact on viral fitness(142,178,179,212,349), and reduce viral replication capacity(140,144,298), resulting in slow disease progression. Viral epitopes that are associated with protective cellular immunity have been identified(150,190,252,335,350). Therefore, the identification of regions of virus that contain PSMs associated with reduced viral fitness and better disease outcome can provide insights for vaccine design. Because HIV-1 depends on the host to survive, analysis of host-viral interactions can help develop such strategies. With advances in bioinformatics and

next generation sequencing technology, it is now possible to analyze a large number of viral sequences that represent the impact of host-viral interactions and regions of the virus associated with favorable disease outcomes.

The major findings of this study can be used to rationally design a potentially effective HIV-1 vaccine. How can this be done? An effective T cell based HIV-1 vaccine should be able to elicit T cell responses that destroy infected cells and drive viral mutations that result in reduced viral fitness. The vaccine elicited CD8+ T cell responses should drive the virus to mutate to PSMs that are associated with slower CD4 decline (beneficial to host), while avoiding PSMs that are associated with faster CD4 decline (detrimental to host). Figure 7.1 shows a simplified strategy of selecting the best immunogens based on the identified association of consensus or PSM amino acids with CD4 decline profile. For example, in HIV-1 Nef, we identified H116**N** and K182**M** as the PSMs associated with slower CD4 decline. Based on this finding, an effective vaccine should include epitopes that contain the consensus amino acids **H** and **K** as immunogens to elicit CD8+ T cell responses that drive viral mutations towards *PSM* amino acids **N** and **M**. On the other hand, E63**D**, I101**V** and I168**M** were associated with faster CD4 decline. In this case the vaccine immunogen should incorporate the PSM amino acids **D**, **V** and **M** to elicit CD8+ T cell responses that drive viral mutations towards the *consensus* amino acids, **E**, **I** and **K**, respectively. Because our study showed that PSMs tend to not deviate from few amino acid substitutions (Figure 3.4 and 3.5), the immune driven selection of the viral mutations to the specific amino acids (e.g. selecting N at position 116 when the immunogen is H or selecting I at position 101 when the immunogen is V) can be expected based on the host HLA class I type.

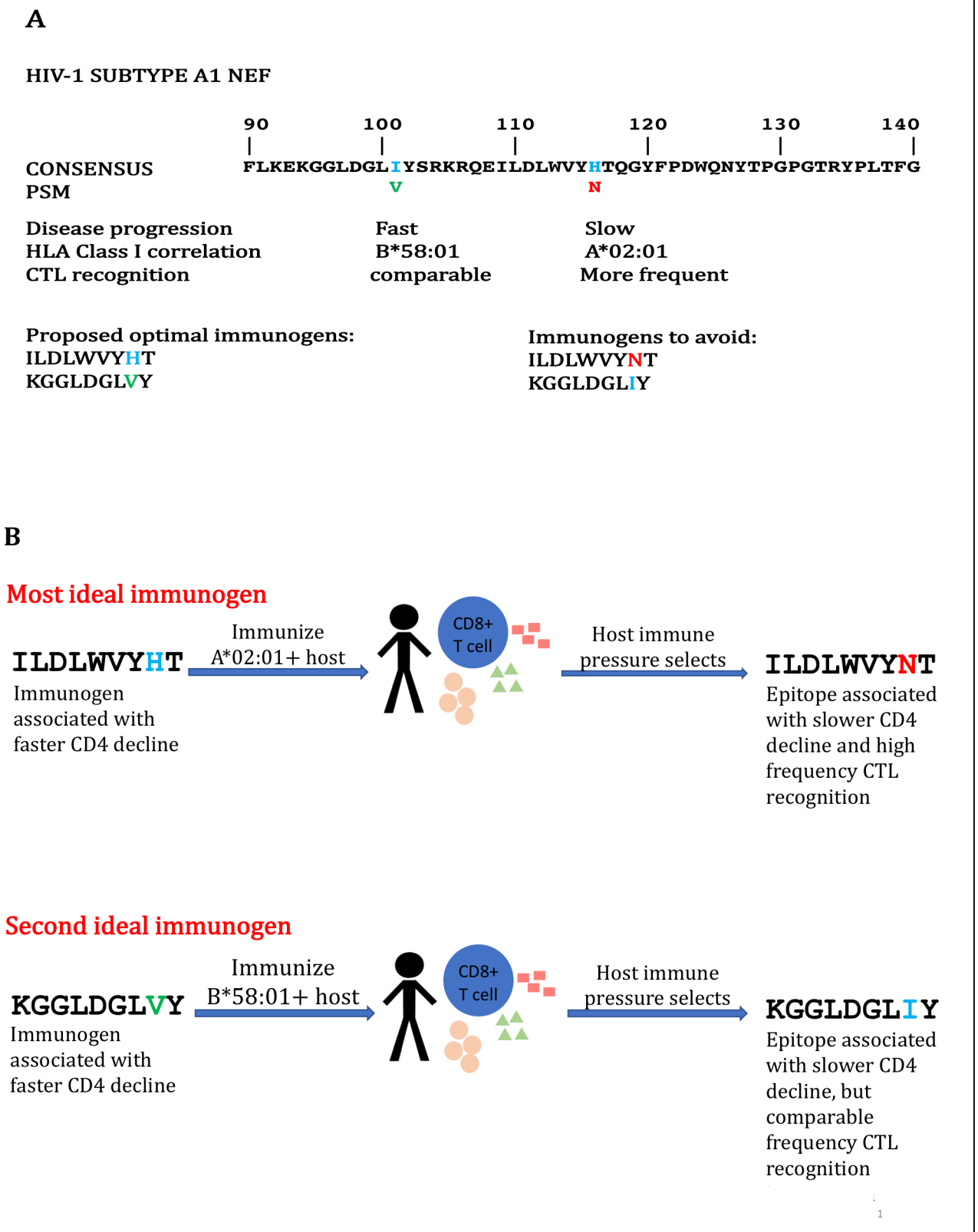


Figure 7.1. A schematic example of using optimal immunogens to design a therapeutic HIV-1 vaccine based on association of consensus and PSM amino acids with disease progression and HLA-restricted CTL recognition frequencies (A). This model allows for design of a vaccine that can generate immune responses to drive mutations to the disadvantage of the virus (B). The most ideal immunogen drives mutation to epitope associated with slow disease progression and high frequency of CTL recognition. Note: immunogen can consist of epitope with either consensus (**ILDLWVYHT)** or PSM (**KGGLDGLVY)** amino acid.

Our study showed that A*02:01 epitope, ILDLWVY**N**T containing the PSM "N" associated with slower CD4 decline, was also targeted by a higher frequency of CD8+ T cells compared to the epitope ILDLWVY**H**T containing the consensus amino acid "H". Therefore, vaccine immunogen consisting of ILDLWVY**H**T epitope may generate CD8+ T cell responses that drive viral mutation to ILDLWVY**N**T. The higher frequency of CD8+ T cells targeting the mutant ILDLWVY**N**T could mediate killing of more infected cells.

Our study also showed that PSMs associated with differential rates of disease progression (E63D, I101V, H116N, I168M and K182M) were all located within the conserved region of the Nef protein. The inclusion of conserved regions of HIV as vaccine immunogen is an attractive approach because it is based on the assumption that viral mutations within such regions will likely result in fitness cost for the virus. Multiple studies in the recent years have evaluated the conserved-element vaccine strategy(351-353). It has been shown that vaccines that target conserved epitopes resulted in high levels of effector T cells and functional HIV-1-specific CD8+ T cells that inhibited HIV-1 replication (354). Another study involving untreated Clade B infected individuals, showed HIV-1-specific T cell responses that targeted conserved regions of Gag protein were associated with lower levels of viral load (252). *In silico* analyses have identified regions of the virus where a mutation is likely to result in protein instability, and it was shown that protective epitopes lied within these region(319). Our study has demonstrated that major PSMs associated with significant clinical outcomes (both beneficial and detrimental) are within the structurally conserved region of Nef. The results of our study extend the existing knowledge on the conserved-element vaccine strategy by demonstrating that even within the conserved regions of Nef,

immunogens need to be optimized to drive viral mutations to the disadvantage of the virus (Fig. 7.1).

Similarly, our approach of immunogen optimization based on PSM identification and characterization can be applied to improve mosaic vaccine design (355), where a polyvalent immunogen that contains several major variations from different strains of the virus, as well as multiple proteins of HIV-1 are combined to form an immunogen (356-358). Based on the findings of this study, I propose that a vaccine containing protective epitopes of Nef as immunogen could be effective. Additionally, epitope variants (PSM or consensus) from other HIV-1 proteins and other subtypes could be taken into consideration to improve this type of vaccine immunogen design (Section 7.3).

7.3 Future Directions

The proposed mechanisms above based on the major findings of this study can be expanded by future studies:

- HIV-1 Nef sequences were obtained by amplifying full HIV-1 genome. We postulate that PSMs correlated with different disease progression and viral suppression also exist in HIV-1 major genes, *gag*, *pol*, and *env* and in other accessory proteins. Identification of all possible protective and non-protective epitopes within HIV-1 can contribute to the selection of the most effective immunogens of a T cell based vaccine.
- HLA-associated PSMs in other HIV-1 clades prevalent in the Kenyan population need to be evaluated. Within Kenya Clade A1 virus is the most prevalent

subtypes(359). In this study nearly 82% of Nef sequences were Clade A1, followed by 10% Clade D, 7% Clade C and 1% Clade B. Evaluation of HLA-associated polymorphisms across multiple clades of HIV-1 can better inform a global vaccine strategy.

- Compensatory mutations in HIV-1 Nef that contribute to viral fitness need to be identified and correlated with clinical outcomes such as CD4 count and VL, to better understand viral dynamics that define mutational patterns.
- Viral fitness assays need to be conducted to assess the impact of PSMs on the replication capacity of virus and overall viral fitness.
- In this study the identification of protective and non-protective epitopes was limited by having no reliable viral load counts. Any future study would benefit from VL data to fully correlate PSMs with viral suppression, which could supplement disease progression associations.
- In this study we mapped the consensus sequence of HIV-1 Nef and PSMs were identified by Quasi analysis. These were based on the ratio of nonsynonymous and synonymous mutations detected among the study population, from 326 subtype A1 sequences. PSMs at the individual patient level also need to be identified to determine alternate amino acid replacement mutations that may not be detected at the population level. 454 pyrosequencing technology allows for this since a high coverage of HIV-1 full genome sequences can be obtained.

7.4 Concluding Remarks

This study identified immunologically relevant CD8+ T cell epitopes of HIV-1 Nef protein by analyzing sequences from a population of 326 subtype A1, ARV naïve HIV+ Kenyan sex workers. PSMs under host immune selection pressures were identified by correlation with patient HLA class I alleles and classified with CD4 T cell count profiles. Three PSMs were associated with faster CD4 decline and two were associated with slower CD4 decline. HLA class I epitopes containing these PSMs were validated and epitope-specific CD8+ T cell responses were examined. We showed that while HLA class I restricted CD8+ T cells targeting epitopes containing beneficial or detrimental amino acids are similar in proliferation, polyfunctionality and exhaustion properties, epitopes associated with slower disease progression were targeted by a higher frequency of CD8+ T cells. Our study not only identified the immunologically relevant CD8+ T cell epitopes in HIV-1 Nef protein, but also established a method to evaluate immunogens to be included for an effective T cell based vaccine for HIV-1 and improve other candidate HIV vaccines.

8. References

1. Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med. Cold Spring Harbor Laboratory Press; 2011 Sep;1(1):a006841–1.

2. UNAIDS. Get on the Fast-Track — The life-cycle approach to HIV. 2016 Dec 1;:1–140.

3. UNAIDS. UNAIDS Data 2017. 2017 Sep 6;:1–248.

4. Gottlieb MS, Schroff R, Schanker HM, Weisman JD, Fan PT, Wolf RA, et al. Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. N Engl J Med. Massachusetts Medical Society; 1981 Dec 10;305(24):1425–31.

5. Becerra JC, Bildstein LS, Gach JS. Recent Insights into the HIV/AIDS Pandemic. Microb Cell. 2016 Sep 5;3(9):451–75.

6. World Health Organization. Guideline on When to Start Antiretroviral Therapy and on Pre-Exposure Prophylaxis for HIV. Geneva: World Health Organization; 2015 Sep.

7. UNAIDS. GLOBAL AIDS UP DATE 2016. 2016 Jun 1;:1–16.

8. Broder S, Gallo RC. A pathogenic retrovirus (HTLV-III) linked to AIDS. N Engl J Med. 1984 Nov 15;311(20):1292–7.

9. UNAIDS. Kenya AIDS Response Progress Report 2014 Progress Towards Zero. March 2014. 2014 Jul 18;:1–41.

10. Hymes KB, Cheung T, Greene JB, Prose NS, Marcus A, Ballard H, et al. Kaposi's sarcoma in homosexual men-a report of eight cases. Lancet. 1981 Sep 19;2(8247):598–600.

11. Masur H, Michelis MA, Greene JB, Onorato I, Stouwe RA, Holzman RS, et al. An outbreak of community-acquired Pneumocystis carinii pneumonia: initial manifestation of cellular immune dysfunction. N Engl J Med. 1981 Dec 10;305(24):1431–8.

12. Greene WC. A history of AIDS: looking back to see ahead. Eur J Immunol. WILEY-VCH Verlag; 2007 Nov;37 Suppl 1(S1):S94–102.

13. Piot P, Quinn TC, Taelman H, Feinsod FM, Minlangu KB, Wobin O, et al. Acquired immunodeficiency syndrome in a heterosexual population in Zaire. Lancet. 1984 Jul 14;2(8394):65–9.

14. Van de Perre P, Rouvroy D, Lepage P, Bogaerts J, Kestelyn P, Kayihigi J, et al.

Acquired immunodeficiency syndrome in Rwanda. Lancet. 1984 Jul 14;2(8394):62–5.

15. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, et al. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. Science. American Association for the Advancement of Science; 2006 Jul 28;313(5786):523–6.

16. Van Heuverswyn F, Li Y, Neel C, Bailes E, Keele BF, Liu W, et al. Human immunodeficiency viruses: SIV infection in wild gorillas. Nature. 2006 Nov 9;444(7116):164–4.

17. Plantier J-C, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemée V, et al. A new human immunodeficiency virus derived from gorillas. Nature Medicine. 2009 Aug;15(8):871–2.

18. Hemelaar J, Gouws E, Ghys PD, Osmanov S, WHO-UNAIDS Network for HIV Isolation and Characterisation. Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS. 2011 Mar 13;25(5):679–89.

19. Becerra JC, Bildstein LS, Gach JS. Recent Insights into the HIV/AIDS Pandemic. Microb Cell. 2016 Sep 5;3(9):451–75.

20. Peeters M, D'Arc M, Delaporte E. Origin and diversity of human retroviruses. AIDS Rev. NIH Public Access; 2014 Jan;16(1):23–34.

21. Liu Y, Li L, Bao Z, Li H, Zhuang D, Liu S, et al. Identification of a novel HIV type 1 circulating recombinant form (CRF52_01B) in Southeast Asia. AIDS Res Hum Retroviruses. 2012 Oct;28(10):1357–61.

22. Korber B, Gaschen B, Yusim K, Thakallapally R, Kesmir C, Detours V. Evolutionary and immunological implications of contemporary HIV-1 variation. Br Med Bull. 2001;58:19–42.

23. Butler IF, Pandrea I, Marx PA, Apetrei C. HIV genetic diversity: biological and public health consequences. Curr HIV Res. 2007 Jan;5(1):23–45.

24. Kiwanuka N, Laeyendecker O, Quinn TC, Wawer MJ, Shepherd J, Robb M, et al. HIV-1 subtypes and differences in heterosexual HIV transmission among HIV-discordant couples in Rakai, Uganda. AIDS. 2009 Nov 27;23(18):2479–84.

25. John-Stewart GC, Nduati RW, Rousseau CM, Mbori-Ngacha DA, Richardson BA, Rainwater S, et al. Subtype C Is associated with increased vaginal shedding of HIV-1. Journal of Infectious Diseases. 2005 Aug 1;192(3):492–6.

26. Iversen AKN, Learn GH, Skinhøj P, Mullins JI, McMichael AJ, Rambaut A. Preferential detection of HIV subtype C' over subtype A in cervical cells from a dually infected woman. AIDS. 2005 Jun 10;19(9):990–3.

27. Walter BL, Armitage AE, Graham SC, de Oliveira T, Skinhøj P, Jones EY, et al. Functional characteristics of HIV-1 subtype C compatible with increased heterosexual transmissibility. AIDS. 2009 Jun 1;23(9):1047–57.

28. Kiwanuka N, Laeyendecker O, Robb M, Kigozi G, Arroyo M, McCutchan F, et al. Effect of human immunodeficiency virus Type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. Journal of Infectious Diseases. Oxford University Press; 2008 Mar 1;197(5):707–13.

29. Baeten JM, Chohan B, Lavreys L, Chohan V, McClelland RS, Certain L, et al. HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. Journal of Infectious Diseases. Oxford University Press; 2007 Apr 15;195(8):1177–80.

30. Kiwanuka N, Robb M, Laeyendecker O, Kigozi G, Wabwire-Mangen F, Makumbi FE, et al. HIV-1 viral subtype differences in the rate of CD4+ T-cell decline among HIV seroincident antiretroviral naive persons in Rakai district, Uganda. J Acquir Immune Defic Syndr. 2010 Jun;54(2):180–4.

31. Alaeus A, Lidman K, Björkman A, Giesecke J, Albert J. Similar rate of disease progression among individuals infected with HIV-1 genetic subtypes A-D. AIDS. 1999 May 28;13(8):901–7.

32. Kantor R, Katzenstein DA, Efron B, Carvalho AP, Wynhoven B, Cane P, et al. Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration. Ho DD, editor. PLoS Med. Public Library of Science; 2005 Apr;2(4):e112.

33. Carlson JM, Du VY, Pfeifer N, Bansal A, Tan VYF, Power K, et al. Impact of pre-adapted HIV transmission. Nature Medicine. 2016 Jun;22(6):606–13.

34. Shaw GM, Hunter E. HIV transmission. Cold Spring Harb Perspect Med. Cold Spring Harbor Laboratory Press; 2012 Nov 1;2(11):a006965–5.

35. Morison L. The global epidemiology of HIV/AIDS. Br Med Bull. 2001;58:7–18.

36. Beyrer C, Sullivan P, Sanchez J, Baral SD, Collins C, Wirtz AL, et al. The increase in global HIV epidemics in MSM. AIDS. 2013 Nov 13;27(17):2665–78.

37. Grulich AE, Zablotska I. Commentary: probability of HIV transmission through anal intercourse. Int J Epidemiol. 2010 Aug;39(4):1064–5.

38. Keele BF, Estes JD. Barriers to mucosal transmission of immunodeficiency viruses. Blood. American Society of Hematology; 2011 Jul 28;118(4):839–46.

39. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, Wabwire-Mangen F, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1.

Rakai Project Study Group. N Engl J Med. Massachusetts Medical Society; 2000 Mar 30;342(13):921–9.

40. Cohen MS, Shaw GM, McMichael AJ, Haynes BF. Acute HIV-1 Infection. N Engl J Med. 2011 May 19;364(20):1943–54.

41. Baeten JM, Kahle E, Lingappa JR, Coombs RW, Delany-Moretlwe S, Nakku-Joloba E, et al. Genital HIV-1 RNA predicts risk of heterosexual HIV-1 transmission. Sci Transl Med. American Association for the Advancement of Science; 2011 Apr 6;3(77):77ra29–9.

42. Tanser F, Bärnighausen T, Hund L, Garnett GP, McGrath N, Newell M-L. Effect of concurrent sexual partnerships on rate of new HIV infections in a high-prevalence, rural South African population: a cohort study. Lancet. Elsevier; 2011 Jul 16;378(9787):247–55.

43. Maartens G, Celum C, Lewin SR. HIV infection: epidemiology, pathogenesis, treatment, and prevention. Lancet. Elsevier; 2014 Jul 19;384(9939):258–71.

44. Mitsuya H, Weinhold KJ, Furman PA, St Clair MH, Lehrman SN, Gallo RC, et al. 3"-Azido-3-"deoxythymidine (BW A509U): an antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus in vitro. Proc Natl Acad Sci USA. National Academy of Sciences; 1985 Oct;82(20):7096–100.

45. Mitsuya H, Broder S. Inhibition of the in vitro infectivity and cytopathic effect of human T-lymphotrophic virus type III/lymphadenopathy-associated virus (HTLV-III/LAV) by 2",3-"dideoxynucleosides. Proc Natl Acad Sci USA. National Academy of Sciences; 1986 Mar;83(6):1911–5.

46. Lange JMA, Ananworanich J. The discovery and development of antiretroviral agents. Antivir Ther (Lond). 2014;19 Suppl 3(Suppl 3):5–14.

47. Volberding PA, Deeks SG. Antiretroviral therapy and management of HIV infection. Lancet. Elsevier; 2010 Jul 3;376(9734):49–62.

48. Vogus A, Graff K. PEPFAR Transitions to Country Ownership: Review of Past Donor Transitions and Application of Lessons Learned to the Eastern Caribbean. Glob Health Sci Pract. Johns Hopkins University- Global Health. Bloomberg School of Public Health, Center for Communication Programs; 2015 Jun 17;3(2):274–86.

49. Silapaswan A, Krakower D, Mayer KH. Pre-Exposure Prophylaxis: A Narrative Review of Provider Behavior and Interventions to Increase PrEP Implementation in Primary Care. J Gen Intern Med. Springer US; 2017 Feb;32(2):192–8.

50. Beyrer C. Pre-exposure prophylaxis for people who inject drugs and their sex partners. Addiction. 2017 Apr;112(4):579–80.

51. Abdool Karim Q, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, Mansoor LE, et al. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science. 2010 Sep 3;329(5996):1168–74.

52. Mayer KH, Safren SA, Elsesser SA, Psaros C, Tinsley JP, Marzinke M, et al. Optimizing Pre-Exposure Antiretroviral Prophylaxis Adherence in Men Who Have Sex with Men: Results of a Pilot Randomized Controlled Trial of "Life-Steps for PrEP". AIDS Behav. 2nd ed. 2017 May;21(5):1350–60.

53. Ying R, Sharma M, Heffron R, Celum CL, Baeten JM, Katabira E, et al. Cost-effectiveness of pre-exposure prophylaxis targeted to high-risk serodiscordant couples as a bridge to sustained ART use in Kampala, Uganda. J Int AIDS Soc. The International AIDS Society; 2015;18(4 Suppl 3):20013.

54. Choopanya K, Martin M, Suntharasamai P, Sangkum U, Mock PA, 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. Elsevier; 2013 Jun 15;381(9883):2083–90.

55. Sultan B, Benn P, Waters L. Current perspectives in HIV post-exposure prophylaxis. HIV AIDS (Auckl). Dove Press; 2014;6:147–58.

56. Kourtis AP, Lee FK, Abrams EJ, Jamieson DJ, Bulterys M. Mother-to-child transmission of HIV-1: timing and implications for prevention. The Lancet Infectious Diseases. Elsevier; 2006 Nov;6(11):726–32.

57. World Health Organization. MONITORING & EVALUATION FRAMEWORK FOR ANTIRETROVIRAL TREATMENT FOR LIVING WITH HIV AND THEIR INFANTS PREGNANT AND BREASTFEEDING WOMEN. 2015 Mar 18;:1–69.

58. Gray RH, Wawer MJ, Kigozi G. Programme science research on medical male circumcision scale-up in sub-Saharan Africa. Sex Transm Infect. 2013 Aug;89(5):345–9.

59. Shin SY. Recent update in HIV vaccine development. Clin Exp Vaccine Res. 2016 Jan;5(1):6–11.

60. Fauci AS, Marston HD. PUBLIC HEALTH. Toward an HIV vaccine: A scientific journey. Science. American Association for the Advancement of Science; 2015 Jul 24;349(6246):386–7.

61. Flynn NM, Forthal DN, Harro CD, Judson FN, Mayer KH, Para MF, et al. Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection. Journal of Infectious Diseases. 2005 Mar 1;191(5):654–65.

62. Pitisuttithum P, Gilbert P, Gurwith M, Heyward W, Martin M, van Griensven F, et al. Randomized, double-blind, placebo-controlled efficacy trial of a bivalent

recombinant glycoprotein 120 HIV-1 vaccine among injection drug users in Bangkok, Thailand. Journal of Infectious Diseases. 2006 Dec 15;194(12):1661–71.

63. Buchbinder SP, Mehrotra DV, Duerr A, Fitzgerald DW, Mogg R, Li D, et al. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. The Lancet. Elsevier; 2008 Nov 29;372(9653):1881–93.

64. Duerr A, Huang Y, Buchbinder S, Coombs RW, Sanchez J, Del Rio C, et al. Extended follow-up confirms early vaccine-enhanced risk of HIV acquisition and demonstrates waning effect over time among participants in a randomized trial of recombinant adenovirus HIV vaccine (Step Study). J Infect Dis. 2012 Jul 15;206(2):258–66.

65. Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, Paris R, et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N Engl J Med. Massachusetts Medical Society; 2009 Dec 3;361(23):2209–20.

66. Haynes BF, Gilbert PB, McElrath MJ, Zolla-Pazner S, Tomaras GD, Alam SM, et al. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. N Engl J Med. Massachusetts Medical Society; 2012 Apr 5;366(14):1275–86.

67. Zolla-Pazner S, deCamp AC, Cardozo T, Karasavvas N, Gottardo R, Williams C, et al. Analysis of V2 antibody responses induced in vaccinees in the ALVAC/AIDSVAX HIV-1 vaccine efficacy trial. Ensoli B, editor. PLoS ONE. Public Library of Science; 2013;8(1):e53629.

68. Karasavvas N, Billings E, Rao M, Williams C, Zolla-Pazner S, Bailer RT, et al. The Thai Phase III HIV Type 1 Vaccine trial (RV144) regimen induces antibodies that target conserved regions within the V2 loop of gp120. AIDS Res Hum Retroviruses. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA; 2012 Nov;28(11):1444–57.

69. Malim MH, Emerman M. HIV-1 accessory proteins--ensuring viral survival in a hostile environment. Cell Host Microbe. Elsevier; 2008 Jun 12;3(6):388–98.

70. Wilen CB, Tilton JC, Doms RW. Molecular mechanisms of HIV entry. Adv Exp Med Biol. Boston, MA: Springer US; 2012;726(Chapter 10):223–42.

71. Pan X, Baldauf H-M, Keppler OT, Fackler OT. Restrictions to HIV-1 replication in resting CD4+ T lymphocytes. Cell Res. Nature Publishing Group; 2013 Jul;23(7):876–85.

72. Chen P, Chen BK, Mosoian A, Hays T, Ross MJ, Klotman PE, et al. Virological synapses allow HIV-1 uptake and gene expression in renal tubular epithelial cells. J Am Soc Nephrol. 2011 Mar;22(3):496–507.

73. Liu Y, Liu H, Kim BO, Gattone VH, Li J, Nath A, et al. CD4-independent infection of astrocytes by human immunodeficiency virus type 1: requirement for the human mannose receptor. Journal of Virology. American Society for Microbiology (ASM); 2004 Apr;78(8):4120–33.

74. Dahabieh MS, Battivelli E, Verdin E. Understanding HIV latency: the road to an HIV cure. Annu Rev Med. Annual Reviews; 2015;66(1):407–21.

75. Lusic M, Siliciano RF. Nuclear landscape of HIV-1 infection and integration. Nature Publishing Group. 2017 Feb;15(2):69–82.

76. Stevenson M. HIV-1 pathogenesis. Nature Medicine. 2003 Jul;9(7):853–60.

77. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, et al. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc Natl Acad Sci USA. 2008 May 27;105(21):7552–7.

78. Walker LM, Huber M, Doores KJ, Falkowska E, Pejchal R, Julien J-P, et al. Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature. 2011 Sep 22;477(7365):466–70.

79. Costiniuk CT, Angel JB. Human immunodeficiency virus and the gastrointestinal immune system: does highly active antiretroviral therapy restore gut immunity? Mucosal Immunol. Nature Publishing Group; 2012 Nov;5(6):596–604.

80. Brenchley JM, Schacker TW, Ruff LE, Price DA, Taylor JH, Beilman GJ, et al. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. Journal of Experimental Medicine. Rockefeller University Press; 2004 Sep 20;200(6):749–59.

81. Mehandru S, Poles MA, Tenner-Racz K, Horowitz A, Hurley A, Hogan C, et al. Primary HIV-1 infection is associated with preferential depletion of CD4+ T lymphocytes from effector sites in the gastrointestinal tract. Journal of Experimental Medicine. Rockefeller University Press; 2004 Sep 20;200(6):761–70.

82. Prendergast A, Prado JG, Kang Y-H, Chen F, Riddell LA, Luzzi G, et al. HIV-1 infection is characterized by profound depletion of CD161+ Th17 cells and gradual decline in regulatory T cells. AIDS. 2010 Feb 20;24(4):491–502.

83. Chang JJ, Altfeld M. Immune activation and the role of TLRs and TLR agonists in the pathogenesis of HIV-1 infection in the humanized mouse model. J Infect Dis. 2013 Nov;208 Suppl 2(suppl_2):S145–9.

84. Sokoya T, Steel HC, Nieuwoudt M, Rossouw TM. HIV as a Cause of Immune Activation and Immunosenescence. Mediators Inflamm. Hindawi; 2017;2017(3):6825493–16.

85. Brenchley JM, Price DA, Schacker TW, Asher TE, Silvestri G, Rao S, et al. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. Nature Medicine. 2006 Dec;12(12):1365–71.

86. Hsue PY, Hunt PW, Sinclair E, Bredt B, Franklin A, Killian M, et al. Increased carotid intima-media thickness in HIV patients is associated with increased cytomegalovirus-specific T-cell responses. AIDS. 2006 Nov 28;20(18):2275–83.

87. Hileman CO, Funderburg NT. Inflammation, Immune Activation, and Antiretroviral Therapy in HIV. Curr HIV/AIDS Rep. Springer US; 2017 Jun;14(3):93–100.

88. Rustagi A, Gale M. Innate antiviral immune signaling, viral evasion and modulation by HIV-1. Journal of Molecular Biology. 2014 Mar 20;426(6):1161–77.

89. Newton K, Dixit VM. Signaling in innate immunity and inflammation. Cold Spring Harb Perspect Biol. Cold Spring Harbor Lab; 2012 Mar 1;4(3):a006049–9.

90. Beignon A-S, McKenna K, Skoberne M, Manches O, DaSilva I, Kavanagh DG, et al. Endocytosis of HIV-1 activates plasmacytoid dendritic cells via Toll-like receptor-viral RNA interactions. J Clin Invest. American Society for Clinical Investigation; 2005 Nov;115(11):3265–75.

91. Borrow P, Shattock RJ, Vyakarnam A, EUROPRISE Working Group. Innate immunity against HIV: a priority target for HIV prevention research. Retrovirology. BioMed Central; 2010 Oct 11;7(1):84.

92. Bostik P, Takahashi Y, Mayne AE, Ansari AA. Innate immune natural killer cells and their role in HIV and SIV infection. HIV Ther. Future Medicine Ltd London, UK; 2010 Jul 1;4(4):483–504.

93. Altfeld M, Gale M. Innate immunity against HIV-1 infection. Nature. 2015 Jun;16(6):554–62.

94. Bowen DG, Walker CM. Adaptive immune responses in acute and chronic hepatitis C virus infection. Nature. 2005 Aug 18;436(7053):946–52.

95. Mouquet H. Antibody B cell responses in HIV-1 infection. Trends Immunol. Elsevier; 2014 Nov;35(11):549–61.

96. Baum LL. Role of humoral immunity in host defense against HIV. Curr HIV/AIDS Rep. 2010 Feb;7(1):11–8.

97. Thèze J, Chakrabarti LA, Vingert B, Porichis F, Kaufmann DE. HIV controllers: a multifactorial phenotype of spontaneous viral suppression. Clin Immunol. 2011 Oct;141(1):15–30.

98. Pereyra F, Addo MM, Kaufmann DE, Liu Y, Miura T, Rathod A, et al. Genetic and

immunologic heterogeneity among persons who control HIV infection in the absence of therapy. Journal of Infectious Diseases. Oxford University Press; 2008 Feb 15;197(4):563–71.

99. Migueles SA, Connors M. Success and failure of the cellular immune response against HIV-1. Nature. 2015 Jun;16(6):563–70.

100. Malhotra U, Holte S, Dutta S, Berrey MM, Delpit E, Koelle DM, et al. Role for HLA class II molecules in HIV-1 suppression and cellular immunity following antiretroviral treatment. J Clin Invest. American Society for Clinical Investigation; 2001 Feb;107(4):505–17.

101. Julg B, Pereyra F, Buzón MJ, Piechocka-Trocha A, Clark MJ, Baker BM, et al. Infrequent recovery of HIV from but robust exogenous infection of activated CD4(+) T cells in HIV elite controllers. Clin Infect Dis. Oxford University Press; 2010 Jul 15;51(2):233–8.

102. Deeks SG, Walker BD. Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. Immunity. Elsevier; 2007 Sep;27(3):406–16.

103. Perreau M, Levy Y, Pantaleo G. Immune response to HIV. Current Opinion in HIV and AIDS. 2013 Jul;8(4):333–40.

104. Matloubian M, Concepcion RJ, Ahmed R. CD4+ T cells are required to sustain CD8+ cytotoxic T-cell responses during chronic viral infection. Journal of Virology. American Society for Microbiology (ASM); 1994 Dec;68(12):8056–63.

105. Zajac AJ, Blattman JN, Murali-Krishna K, Sourdive DJ, Suresh M, Altman JD, et al. Viral immune evasion due to persistence of activated T cells without effector function. Journal of Experimental Medicine. The Rockefeller University Press; 1998 Dec 21;188(12):2205–13.

106. Kaech SM, Ahmed R. Immunology. CD8 T cells remember with a little help. Science. American Association for the Advancement of Science; 2003 Apr 11;300(5617):263–5.

107. Mohan T, Bhatnagar S, Gupta DL, Rao DN. Current understanding of HIV-1 and T-cell adaptive immunity: progress to date. Microb Pathog. 2014 Aug;73:60–9.

108. Munier CML, Kelleher AD, Kent SJ, De Rose R. The role of T cell immunity in HIV-1 infection. Curr Opin Virol. 2013 Aug;3(4):438–46.

109. Jiao Y, Xie J, Li T, Han Y, Qiu Z, Zuo L, et al. Correlation between gag-specific CD8 T-cell responses, viral load, and CD4 count in HIV-1 infection is dependent on disease status. J Acquir Immune Defic Syndr. 2006 Jul;42(3):263–8.

110. Jin X, Bauer DE, Tuttleton SE, Lewin S, Gettie A, Blanchard J, et al. Dramatic rise in plasma viremia after CD8(+) T cell depletion in simian immunodeficiency virus-infected macaques. Journal of Experimental Medicine. The Rockefeller University Press; 1999 Mar 15;189(6):991–8.

111. Tomiyama H, Matsuda T, Takiguchi M. Differentiation of human CD8(+) T cells from a memory to memory/effector phenotype. The Journal of Immunology. 2002 Jun 1;168(11):5538–50.

112. Owen RE, Heitman JW, Hirschkorn DF, Lanteri MC, Biswas HH, Martin JN, et al. HIV+ elite controllers have low HIV-specific T-cell activation yet maintain strong, polyfunctional T-cell responses. AIDS. 2010 May 15;24(8):1095–105.

113. Miura T, Brockman MA, Schneidewind A, Lobritz M, Pereyra F, Rathod A, et al. HLA-B57/B*5801 human immunodeficiency virus type 1 elite controllers select for rare gag variants associated with reduced viral replication capacity and strong cytotoxic T-lymphocyte [corrected] recognition. Journal of Virology. American Society for Microbiology; 2009 Mar;83(6):2743–55.

114. Shacklett BL. Understanding the “lucky few”: the conundrum of HIV-exposed, seronegative individuals. Curr HIV/AIDS Rep. 2006 Feb;3(1):26–31.

115. Horton RE, McLaren PJ, Fowke K, Kimani J, Ball TB. Cohorts for the study of HIV-1-exposed but uninfected individuals: benefits and limitations. J Infect Dis. 2010 Nov 1;202 Suppl 3(S3):S377–81.

116. Erickson AL, Willberg CB, McMahan V, Liu A, Buchbinder SP, Grohskopf LA, et al. Potentially exposed but uninfected individuals produce cytotoxic and polyfunctional human immunodeficiency virus type 1-specific CD8(+) T-cell responses which can be defined to the epitope level. Clin Vaccine Immunol. American Society for Microbiology; 2008 Nov;15(11):1745–8.

117. Gulzar N, Copeland KFT. CD8+ T-cells: function and response to HIV infection. Curr HIV Res. 2004 Jan;2(1):23–37.

118. Demers KR, Reuter MA, Betts MR. CD8(+) T-cell effector function and transcriptional regulation during HIV pathogenesis. Immunol Rev. 2013 Jul;254(1):190–206.

119. Mendoza D, Migueles SA, Rood JE, Peterson B, Johnson S, Doria-Rose N, et al. Cytotoxic capacity of SIV-specific CD8(+) T cells against primary autologous targets correlates with immune control in SIV-infected rhesus macaques. Silvestri G, editor. PLoS Pathog. 2013 Feb;9(2):e1003195.

120. Blankson JN, Siliciano RF. Elite suppression of HIV-1 replication. Immunity. Elsevier; 2008 Dec 19;29(6):845–7.

121. Yewdell JW, Reits E, Neefjes J. Making sense of mass destruction: quantitating MHC class I antigen presentation. Nat Rev Immunol. 2003 Dec;3(12):952–61.

122. Hansen TH, Bouvier M. MHC class I antigen presentation: learning from viral evasion strategies. Nat Rev Immunol. 2009 Jul;9(7):503–13.

123. Goulder PJR, Watkins DI. HIV and SIV CTL escape: implications for vaccine design. Nat Rev Immunol. 2004 Aug;4(8):630–40.

124. Johnson WE, Desrosiers RC. Viral persistence: HIV's strategies of immune system evasion. Annu Rev Med. 2002;53(1):499–518.

125. Altman JD, Feinberg MB. HIV escape: there and back again. Nature Medicine. 2004 Mar;10(3):229–30.

126. Hill AL, Rosenbloom DIS, Nowak MA. Evolutionary dynamics of HIV at multiple spatial and temporal scales. J Mol Med. 2012 May 3;90(5):543–61.

127. Mansky LM, Temin HM. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. Journal of Virology. American Society for Microbiology (ASM); 1995 Aug;69(8):5087–94.

128. Freed EO. HIV-1 replication. Somat Cell Mol Genet. 2001 Nov;26(1-6):13–33.

129. Stewart JJ, Watts P, Litwin S. An algorithm for mapping positively selected members of quasispecies-type viruses. BMC Bioinformatics. BioMed Central; 2001;2(1):1.

130. M M, J E, J Q. HepatitisCVirus(HCV) circulates as a population of different but closely related Genomes: Quasispecies Nature of HCV Genome Distribution. 2014 Jul 18;:1–5.

131. Koup RA, Safrit JT, Cao Y, Andrews CA, McLeod G, Borkowsky W, et al. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. Journal of Virology. American Society for Microbiology (ASM); 1994 Jul;68(7):4650–5.

132. Borrow P, Lewicki H, Hahn BH, Shaw GM, Oldstone MB. Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. Journal of Virology. American Society for Microbiology (ASM); 1994 Sep;68(9):6103–10.

133. Goonetilleke N, Liu MKP, Salazar-Gonzalez JF, Ferrari G, Giorgi E, Ganusov VV, et al. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. J Exp Med. 2009 Jun 8;206(6):1253–72.

134. Ferrari G, Korber B, Goonetilleke N, Liu MKP, Turnbull EL, Salazar-Gonzalez JF, et al.

Relationship between functional profile of HIV-1 specific CD8 T cells and epitope variability with the selection of escape mutants in acute HIV-1 infection. Desrosiers RC, editor. PLoS Pathog. 2011 Feb 10;7(2):e1001273.

135. Roider J, Kalteis A-L, Vollbrecht T, Gloning L, Stirner R, Henrich N, et al. Adaptation of CD8 T cell responses to changing HIV-1 sequences in a cohort of HIV-1 infected individuals not selected for a certain HLA allele. Le Gall S, editor. PLoS ONE. Public Library of Science; 2013;8(12):e80045.

136. Bailey JR, Brennan TP, O'Connell KA, Siliciano RF, Blankson JN. Evidence of CD8+ T-cell-mediated selective pressure on human immunodeficiency virus type 1 nef in HLA-B*57+ elite suppressors. Journal of Virology. 2009 Jan;83(1):88–97.

137. Goulder PJ, Phillips RE, Colbert RA, McAdam S, Ogg G, Nowak MA, et al. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. Nature Medicine. 1997 Feb;3(2):212–7.

138. Phillips RE, Rowland-Jones S, Nixon DF, Gotch FM, Edwards JP, Ogunlesi AO, et al. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. Nature. Nature Publishing Group; 1991 Dec 12;354(6353):453–9.

139. Borrow P, Lewicki H, Wei X, Horwitz MS, Peffer N, Meyers H, et al. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. Nature Medicine. 1997 Feb;3(2):205–11.

140. Schneidewind A, Brockman MA, Yang R, Adam RI, Li B, Le Gall S, et al. Escape from the dominant HLA-B27-restricted cytotoxic T-lymphocyte response in Gag is associated with a dramatic reduction in human immunodeficiency virus type 1 replication. Journal of Virology. 2007 Nov;81(22):12382–93.

141. Peters HO, Mendoza MG, Capina RE, Luo M, Mao X, Gubbins M, et al. An integrative bioinformatic approach for studying escape mutations in human immunodeficiency virus type 1 gag in the Pumwani Sex Worker Cohort. Journal of Virology. American Society for Microbiology; 2008 Feb;82(4):1980–92.

142. Boutwell CL, Carlson JM, Lin T-H, Seese A, Power KA, Peng J, et al. Frequent and variable cytotoxic-T-lymphocyte escape-associated fitness costs in the human immunodeficiency virus type 1 subtype B Gag proteins. Journal of Virology. 2013 Apr;87(7):3952–65.

143. Brumme ZL, John M, Carlson JM, Brumme CJ, Chan D, Brockman MA, et al. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. Nixon DF, editor. PLoS ONE. Public Library of Science; 2009 Aug 19;4(8):e6687.

144. Sakai K, Chikata T, Brumme ZL, Brumme CJ, Gatanaga H, Gatanag H, et al. Lack of a significant impact of Gag-Protease-mediated HIV-1 replication capacity on clinical

parameters in treatment-naive Japanese individuals. Retrovirology. BioMed Central; 2015 Nov 19;12(1):98.

145. Carlson JM, Le AQ, Shahid A, Brumme ZL. HIV-1 adaptation to HLA: a window into virus–host immune interactions. Trends in Microbiology. Elsevier Ltd; 2015 Jan 19;:1–13.

146. Yokomaku Y, Miura H, Tomiyama H, Kawana-Tachikawa A, Takiguchi M, Kojima A, et al. Impaired processing and presentation of cytotoxic-T-lymphocyte (CTL) epitopes are major escape mechanisms from CTL immune pressure in human immunodeficiency virus type 1 infection. Journal of Virology. American Society for Microbiology (ASM); 2004 Feb;78(3):1324–32.

147. Cardinaud S, Consiglieri G, Bouziat R, Urrutia A, Graff-Dubois S, Fourati S, et al. CTL escape mediated by proteasomal destruction of an HIV-1 cryptic epitope. Luban J, editor. PLoS Pathog. 2011 May;7(5):e1002049.

148. Kim V, Green WR. The role of proximal and distal sequence variations in the presentation of an immunodominant CTL epitope encoded by the ecotropic AK7 MuLV. Virology. 1997 Sep 29;236(2):221–33.

149. Draenert R, Le Gall S, Pfafferott KJ, Leslie AJ, Chetty P, Brander C, et al. Immune selection for altered antigen processing leads to cytotoxic T lymphocyte escape in chronic HIV-1 infection. Journal of Experimental Medicine. 2004 Apr 5;199(7):905–15.

150. Carlson JM, Brumme CJ, Martin E, Listgarten J, Brockman MA, Le AQ, et al. Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. Journal of Virology. American Society for Microbiology; 2012 Dec;86(24):13202–16.

151. Bronke C, Almeida C-AM, McKinnon E, Roberts SG, Keane NM, Chopra A, et al. HIV escape mutations occur preferentially at HLA-binding sites of CD8 T-cell epitopes. AIDS. 2013 Mar 27;27(6):899–905.

152. Iglesias MC, Almeida JR, Fastenackels S, van Bockel DJ, Hashimoto M, Venturi V, et al. Escape from highly effective public CD8+ T-cell clonotypes by HIV. Blood. American Society of Hematology; 2011 Aug 25;118(8):2138–49.

153. Friedrich TC, McDermott AB, Reynolds MR, Piaskowski S, Fuenger S, De Souza IP, et al. Consequences of cytotoxic T-lymphocyte escape: common escape mutations in simian immunodeficiency virus are poorly recognized in naive hosts. Journal of Virology. American Society for Microbiology; 2004 Sep;78(18):10064–73.

154. Goulder PJR, Walker BD. HIV and HLA class I: an evolving relationship. Immunity. Elsevier; 2012 Sep 21;37(3):426–40.

155. Marmor M, Hertzmark K, Thomas SM, Halkitis PN, Vogler M. Resistance to HIV infection. J Urban Health. 2006 Jan;83(1):5–17.

156. Zinkernagel RM, Doherty PC. Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. Nature. 1974 Apr 19;248(5450):701–2.

157. Carrington M, O'Brien SJ. The influence of HLA genotype on AIDS. Annu Rev Med. 2003;54(1):535–51.

158. Carlson JM, Brumme ZL. HIV evolution in response to HLA-restricted CTL selection pressures: a population-based perspective. Microbes Infect. 2008 Apr;10(5):455–61.

159. Kløverpris HN, Leslie A, Goulder P. Role of HLA Adaptation in HIV Evolution. Front Immunol. Frontiers; 2015;6:665.

160. Kaslow RA, Carrington M, Apple R, Park L, Muñoz A, Saah AJ, et al. Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. Nature Medicine. 1996 Apr;2(4):405–11.

161. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, et al. A whole-genome association study of major determinants for host control of HIV-1. Science. 2007 Aug 17;317(5840):944–7.

162. International HIV Controllers Study, Pereyra F, Jia X, McLaren PJ, de Bakker PIW, Walker BD, et al. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science. 2010 Dec 10;330(6010):1551–7.

163. Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C, Chetty S, et al. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature. 2004 Dec 9;432(7018):769–75.

164. Allen TM, Altfeld M, Yu XG, O'sullivan KM, Lichterfeld M, Le Gall S, et al. Selection, transmission, and reversion of an antigen-processing cytotoxic T-lymphocyte escape mutation in human immunodeficiency virus type 1 infection. Journal of Virology. 2004 Jul;78(13):7069–78.

165. Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, et al. Antibody neutralization and escape by HIV-1. Nature. 2003 Mar 20;422(6929):307–12.

166. Kelleher AD, Long C, Holmes EC, Allen RL, Wilson J, Conlon C, et al. Clustered mutations in HIV-1 gag are consistently required for escape from HLA-B27-restricted cytotoxic T lymphocyte responses. Journal of Experimental Medicine. The Rockefeller University Press; 2001 Feb 5;193(3):375–86.

167. Moore CB, John M, James IR, Christiansen FT, Witt CS, Mallal SA. Evidence of HIV-1

adaptation to HLA-restricted immune responses at a population level. Science. 2002 May 24;296(5572):1439–43.

168. Leslie A, Kavanagh D, Honeyborne I, Pfafferott K, Edwards C, Pillay T, et al. Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA. Journal of Experimental Medicine. 2005 Mar 21;201(6):891–902.

169. Itoh Y, Mizuki N, Shimada T, Azuma F, Itakura M, Kashiwase K, et al. High-throughput DNA typing of HLA-A, -B, -C, and -DRB1 loci by a PCR-SSOP-Luminex method in the Japanese population. Immunogenetics. 2005 Nov;57(10):717–29.

170. Furutsuki T, Hosoya N, Kawana-Tachikawa A, Tomizawa M, Odawara T, Goto M, et al. Frequent Transmission of Cytotoxic-T-Lymphocyte Escape Mutants of Human Immunodeficiency Virus Type 1 in the Highly HLA-A24-Positive Japanese Population. Journal of Virology. 2004 Jul 27;78(16):8437–45.

171. Chikata T, Carlson JM, Tamura Y, Borghan MA, Naruto T, Hashimoto M, et al. Host-Specific Adaptation of HIV-1 Subtype B in the Japanese Population. Journal of Virology. 2014 Apr 4;88(9):4764–75.

172. Han C, Kawana-Tachikawa A, Shimizu A, Zhu D, Nakamura H, Adachi E, et al. Switching and emergence of CTL epitopes in HIV-1 infection. 2014 May 21;11(1):1–15.

173. Duda A, Lee-Turner L, Fox J, Robinson N, Dustan S, Kaye S, et al. HLA-associated clinical progression correlates with epitope reversion rates in early human immunodeficiency virus infection. Journal of Virology. 2009 Feb;83(3):1228–39.

174. Li B, Gladden AD, Altfeld M, Kaldor JM, Cooper DA, Kelleher AD, et al. Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. Journal of Virology. 2007 Jan;81(1):193–201.

175. Treurnicht FK, Seoighe C, Martin DP, Wood N, Abrahams M-R, Rosa D de A, et al. Adaptive changes in HIV-1 subtype C proteins during early infection are driven by changes in HLA-associated immune pressure. Virology. 2010 Jan 20;396(2):213–25.

176. Leslie AJ, Pfafferott KJ, Chetty P, Draenert R, Addo MM, Feeney M, et al. HIV evolution: CTL escape mutation and reversion after transmission. Nature Medicine. 2004 Mar;10(3):282–9.

177. Brumme ZL, Brumme CJ, Carlson J, Streeck H, John M, Eichbaum Q, et al. Marked epitope- and allele-specific differences in rates of mutation in human immunodeficiency type 1 (HIV-1) Gag, Pol, and Nef cytotoxic T-lymphocyte epitopes in acute/early HIV-1 infection. Journal of Virology. American Society for Microbiology; 2008 Sep;82(18):9216–27.

178. Martinez-Picado J, Prado JG, Fry EE, Pfafferott K, Leslie A, Chetty S, et al. Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. Journal of Virology. American Society for Microbiology; 2006 Apr;80(7):3617–23.

179. Gijsbers EF, Feenstra KA, van Nuenen AC, Navis M, Heringa J, Schuitemaker H, et al. HIV-1 Replication Fitness of HLA-B*57/58:01 CTL Escape Variants Is Restored by the Accumulation of Compensatory Mutations in Gag. Mammano F, editor. PLoS ONE. 2013 Dec 5;8(12):e81235–10.

180. Huang K-HG, Goedhals D, Carlson JM, Brockman MA, Mishra S, Brumme ZL, et al. Progression to AIDS in South Africa is associated with both reverting and compensatory viral mutations. Chauhan A, editor. PLoS ONE. 2011 Apr 22;6(4):e19018.

181. Prince JL, Claiborne DT, Carlson JM, Schaefer M, Yu T, Lahki S, et al. Role of transmitted Gag CTL polymorphisms in defining replicative capacity and early HIV-1 pathogenesis. Lifson J, editor. PLoS Pathog. 2012;8(11):e1003041.

182. Brockman MA, Brumme ZL, Brumme CJ, Miura T, Sela J, Rosato PC, et al. Early selection in Gag by protective HLA alleles contributes to reduced HIV-1 replication capacity that may be largely compensated for in chronic infection. Journal of Virology. 2010 Nov;84(22):11937–49.

183. Wright JK, Brumme ZL, Carlson JM, Heckerman D, Kadie CM, Brumme CJ, et al. Gag-protease-mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. Journal of Virology. American Society for Microbiology; 2010 Oct;84(20):10820–31.

184. Carlson JM, Schaefer M, Monaco DC, Batorsky R, Claiborne DT, Prince J, et al. HIV transmission. Selection bias at the heterosexual HIV-1 transmission bottleneck. Science. 2014 Jul 11;345(6193):1254031–1.

185. Chopera DR, Woodman Z, Mlisana K, Mlotshwa M, Martin DP, Seoighe C, et al. Transmission of HIV-1 CTL escape variants provides HLA-mismatched recipients with a survival advantage. Koup RA, editor. PLoS Pathog. 2008 Mar 21;4(3):e1000033.

186. Goepfert PA, Lumm W, Farmer P, Matthews P, Prendergast A, Carlson JM, et al. Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. J Exp Med. 2008 May 12;205(5):1009–17.

187. Miura T, Brumme ZL, Brockman MA, Rosato P, Sela J, Brumme CJ, et al. Impaired replication capacity of acute/early viruses in persons who become HIV controllers. Journal of Virology. 2010 Aug;84(15):7581–91.

188. Matthews PC, Prendergast A, Leslie A, Crawford H, Payne R, Rousseau C, et al.

Central role of reverting mutations in HLA associations with human immunodeficiency virus set point. Journal of Virology. 2008 Sep;82(17):8548–59.

189. Geyer M, Fackler OT, Peterlin BM. Structure--function relationships in HIV-1 Nef. EMBO Rep. EMBO Press; 2001 Jul;2(7):580–5.

190. Adland E, Carlson JM, Paioni P, Kløverpris H, Shapiro R, Ogwu A, et al. Nef-specific CD8+ T cell responses contribute to HIV-1 immune control. Goepfert PA, editor. PLoS ONE. Public Library of Science; 2013;8(9):e73117.

191. Ren X, Park SY, Bonifacino JS, Hurley JH. How HIV-1 Nef hijacks the AP-2 clathrin adaptor to downregulate CD4. Elife. eLife Sciences Publications Limited; 2014;3:e01754.

192. Lang SM, Iafrate AJ, Stahl-Hennig C, Kuhn EM, Nisslein T, Kaup FJ, et al. Association of simian immunodeficiency virus Nef with cellular serine/threonine kinases is dispensable for the development of AIDS in rhesus macaques. Nature Medicine. 1997 Aug;3(8):860–5.

193. Dyer WB, Geczy AF, Kent SJ, McIntyre LB, Blasdall SA, Learmont JC, et al. Lymphoproliferative immune function in the Sydney Blood Bank Cohort, infected with natural nef/long terminal repeat mutants, and in other long-term survivors of transfusion-acquired HIV-1 infection. AIDS. 1997 Nov;11(13):1565–74.

194. McIntyre LB, Geczy AF, Dyer WB, Learmont JC, Sullivan JS. The Sydney Blood Bank Cohort: a case-control study using a transfused HIV-1 seronegative group. Ann Epidemiol. 1999 Oct;9(7):436–40.

195. Foster JL, Garcia JV. HIV-1 Nef: at the crossroads. Retrovirology. 2008;5(1):84–13.

196. Navis M, Schellens IMM, van Swieten P, Borghans JAM, Miedema F, Kootstra NA, et al. A nonprogressive clinical course in HIV-infected individuals expressing human leukocyte antigen B57/5801 is associated with preserved CD8+ T lymphocyte responsiveness to the HW9 epitope in Nef. Journal of Infectious Diseases. 2008 Mar 15;197(6):871–9.

197. Carl S, Iafrate AJ, Lang SM, Stolte N, Stahl-Hennig C, Mätz-Rensing K, et al. Simian immunodeficiency virus containing mutations in N-terminal tyrosine residues and in the PxxP motif in Nef replicates efficiently in rhesus macaques. Journal of Virology. American Society for Microbiology (ASM); 2000 May;74(9):4155–64.

198. Lama J, Mangasarian A, Trono D. Cell-surface expression of CD4 reduces HIV-1 infectivity by blocking Env incorporation in a Nef- and Vpu-inhibitable manner. Curr Biol. 1999 Jun 17;9(12):622–31.

199. Ross TM, Oran AE, Cullen BR. Inhibition of HIV-1 progeny virion release by cell-surface CD4 is relieved by expression of the viral Nef protein. Curr Biol. 1999 Jun

17;9(12):613–21.

200. Skowronski J, Parks D, Mariani R. Altered T cell activation and development in transgenic mice expressing the HIV-1 nef gene. EMBO J. European Molecular Biology Organization; 1993 Feb;12(2):703–13.

201. Greenberg ME, Iafrate AJ, Skowronski J. The SH3 domain-binding surface and an acidic motif in HIV-1 Nef regulate trafficking of class I MHC complexes. EMBO J. EMBO Press; 1998 May 15;17(10):2777–89.

202. Wonderlich ER, Williams M, Collins KL. The tyrosine binding pocket in the adaptor protein 1 (AP-1) mu1 subunit is necessary for Nef to recruit AP-1 to the major histocompatibility complex class I cytoplasmic tail. J Biol Chem. 2008 Feb 8;283(6):3011–22.

203. Atkins KM, Thomas L, Youker RT, Harriff MJ, Pissani F, You H, et al. HIV-1 Nef binds PACS-2 to assemble a multikinase cascade that triggers major histocompatibility complex class I (MHC-I) down-regulation: analysis using short interfering RNA and knock-out mice. J Biol Chem. American Society for Biochemistry and Molecular Biology; 2008 Apr 25;283(17):11772–84.

204. Stolp B, Abraham L, Rudolph JM, Fackler OT. Lentiviral Nef proteins utilize PAK2-mediated deregulation of cofilin as a general strategy to interfere with actin remodeling. Journal of Virology. 2010 Apr;84(8):3935–48.

205. Giorgi JV, Hultin LE, McKeating JA, Johnson TD, Owens B, Jacobson LP, et al. Shorter survival in advanced human immunodeficiency virus type 1 infection is more closely associated with T lymphocyte activation than with plasma virus burden or virus chemokine coreceptor usage. Journal of Infectious Diseases. 1999 Apr;179(4):859–70.

206. Sousa AE, Carneiro J, Meier-Schellersheim M, Grossman Z, Victorino RMM. CD4 T cell depletion is linked directly to immune activation in the pathogenesis of HIV-1 and HIV-2 but only indirectly to the viral load. The Journal of Immunology. 2002 Sep 15;169(6):3400–6.

207. Das SR, Jameel S. Biology of the HIV Nef protein. Indian J Med Res. 2005 Apr;121(4):315–32.

208. Mudd PA, Ericsen AJ, Burwitz BJ, Wilson NA, O'Connor DH, Hughes AL, et al. Escape from CD8(+) T cell responses in Mamu-B*00801(+) macaques differentiates progressors from elite controllers. J Immunol. American Association of Immunologists; 2012 Apr 1;188(7):3364–70.

209. Budde ML, Greene JM, Chin EN, Ericsen AJ, Scarlotta M, Cain BT, et al. Specific CD8+ T cell responses correlate with control of simian immunodeficiency virus replication in Mauritian cynomolgus macaques. Journal of Virology. American

Society for Microbiology; 2012 Jul;86(14):7596–604.

210. Betts MR, Exley B, Price DA, Bansal A, Camacho ZT, Teaberry V, et al. Characterization of functional and phenotypic changes in anti-Gag vaccine-induced T cell responses and their role in protection after HIV-1 infection. Proc Natl Acad Sci USA. National Acad Sciences; 2005 Mar 22;102(12):4512–7.

211. Baker BM, Block BL, Rothchild AC, Walker BD. Elite control of HIV infection: implications for vaccine design. Expert Opinion on Biological Therapy. 2009 Jan;9(1):55–69.

212. Troyer RM, McNevin J, Liu Y, Zhang SC, Krizan RW, Abraha A, et al. Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. Walker CM, editor. PLoS Pathog. 2009 Apr;5(4):e1000365.

213. Bernstein A. AIDS and the next 25 years. Science. American Association for the Advancement of Science; 2008 May 9;320(5877):717–7.

214. Narute PS, Smithgall TE. Nef Alleles from All Major HIV-1 Clades Activate Src-Family Kinases and Enhance HIV-1 Replication in an Inhibitor-Sensitive Manner. Vartanian J-P, editor. PLoS ONE. 2012 Feb 29;7(2):e32561–12.

215. Ali A, Pillai S, Ng H, Lubong R, Richman DD, Jamieson BD, et al. Broadly increased sensitivity to cytotoxic T lymphocytes resulting from Nef epitope escape mutations. The Journal of Immunology. 2003 Oct 15;171(8):3999–4005.

216. Ueno T, Motozono C, Dohki S, Mwimanzi P, Rauch S, Fackler OT, et al. CTL-mediated selective pressure influences dynamic evolution and pathogenic functions of HIV-1 Nef. The Journal of Immunology. 2008 Jan 15;180(2):1107–16.

217. Cereb N, Yang SY. The regulatory complex of HLA class I promoters exhibits locus-specific conservation with limited allelic variation. The Journal of Immunology. 1994 Apr 15;152(8):3873–83.

218. Cereb N, Maye P, Lee S, Kong Y, Yang SY. Locus-specific amplification of HLA class I genes from genomic DNA: locus-specific sequences in the first and third introns of HLA-A, -B, and -C alleles. Tissue Antigens. 1995 Jan;45(1):1–11.

219. Cereb N, Hughes AL, Yang SY. Locus-specific conservation of the HLA class I introns by intra-locus homogenization. Immunogenetics. 1997;47(1):30–6.

220. Luo M, Blanchard J, Pan Y, Brunham K, Brunham RC. High-resolution sequence typing of HLA-DQA1 and -DQB1 exon 2 DNA with taxonomy-based sequence analysis (TBSA) allele assignment. Tissue Antigens. 1999 Jul;54(1):69–82.

221. Luo M, Blanchard J, Brunham K, Pan Y, Shen CX, Lu H, et al. Two-step high resolution sequence-based HLA-DRB typing of exon 2 DNA with taxonomy-based

sequence analysis allele assignment. Hum Immunol. 2001 Nov;62(11):1294–310.

222. Peterson TA, Kimani J, Wachihi C, Bielawny T, Mendoza L, Thavaneswaran S, et al. HLA class I associations with rates of HIV-1 seroconversion and disease progression in the Pumwani Sex Worker Cohort. Tissue Antigens. Blackwell Publishing Ltd; 2013 Feb;81(2):93–107.

223. Thuring RW, Sanders JP, Borst P. A freeze-squeeze method for recovering long DNA from agarose gels. Anal Biochem. 1975 May 26;66(1):213–20.

224. Lamoreaux L, Roederer M, Koup R. Intracellular cytokine optimization and standard operating procedure. Nat Protoc. 2006;1(3):1507–16.

225. San José E, Borroto A, Niedergang F, Alcover A, Alarcón B. Triggering the TCR complex causes the downregulation of nonengaged receptors by a signal transduction-dependent mechanism. Immunity. 2000 Feb;12(2):161–70.

226. Chan KS, Kaur A. Flow cytometric detection of degranulation reveals phenotypic heterogeneity of degranulating CMV-specific CD8+ T lymphocytes in rhesus macaques. J Immunol Methods. 2007 Aug 31;325(1-2):20–34.

227. Gauduin M-C. Intracellular cytokine staining for the characterization and quantitation of antigen-specific T lymphocyte responses. Methods. 2006 Apr;38(4):263–73.

228. Gao F, Korber BT, Weaver E, Liao H-X, Hahn BH, Haynes BF. Centralized immunogens as a vaccine strategy to overcome HIV-1 diversity. Expert Rev Vaccines. 2004 Aug;3(4 Suppl):S161–8.

229. Brumme ZL, Brumme CJ, Heckerman D, Korber BT, Daniels M, Carlson J, et al. Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1. PLoS Pathog. Public Library of Science; 2007 Jul;3(7):e94.

230. Liang B, Luo M, Ball TB, Plummer FA. QUASI analysis of the HIV-1 envelope sequences in the Los Alamos National Laboratory HIV sequence database: pattern and distribution of positive selection sites and their frequencies over yearsThis paper is one of a selection of papers in this Special Issue, entitled International Symposium on Recent Advances in Molecular, Clinical, and Social Medicine, and has undergone the Journal's usual peer-review process. Biochem Cell Biol. 2007 Apr;85(2):259–64.

231. Pan C, Kim J, Chen L, Wang Q, Lee C. The HIV positive selection mutation database. Nucleic Acids Res. 2007 Jan;35(Database issue):D371–5.

232. Li WH. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J Mol Evol. 1993 Jan;36(1):96–9.

233. Chen L, Perlina A, Lee CJ. Positive selection detection in 40,000 human immunodeficiency virus (HIV) type 1 sequences automatically identifies drug resistance and positive fitness mutations in HIV protease and reverse transcriptase. Journal of Virology. American Society for Microbiology (ASM); 2004 Apr;78(7):3722–32.

234. Price DA, Goulder PJ, Klenerman P, Sewell AK, Easterbrook PJ, Troop M, et al. Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. Proc Natl Acad Sci USA. National Academy of Sciences; 1997 Mar 4;94(5):1890–5.

235. Liang B, Luo M, Ball TB, Jones SJM, Plummer FA. QUASI analysis of host immune responses to Gag polyproteins of human immunodeficiency virus type 1 by a systematic bioinformatics approachThis paper is one of a selection of papers published in this special issue entitled “Second International Symposium on Recent Advances in Basic, Clinical, and Social Medicine” and has undergone the Journal's usual peer review process. Biochem Cell Biol. 2010 Aug;88(4):671–81.

236. Merani S, Petrovic D, James I, Chopra A, Cooper D, Freitas E, et al. Effect of immune pressure on hepatitis C virus evolution: Insights from a single-source outbreak. Hepatology. 2011 Jan 18;53(2):396–405.

237. Saksela K, Cheng G, Baltimore D. Proline-rich (PxxP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef+ viruses but not for down-regulation of CD4. EMBO J. European Molecular Biology Organization; 1995 Feb 1;14(3):484–91.

238. Raney A, Shaw AY, Foster JL, Garcia JV. Structural constraints on human immunodeficiency virus type 1 Nef function. Virology. 2007 Nov 10;368(1):7–16.

239. Khan IH, Sawai ET, Antonio E, Weber CJ, Mandell CP, Montbriand P, et al. Role of the SH3-ligand domain of simian immunodeficiency virus Nef in interaction with Nef-associated kinase and simian AIDS in rhesus macaques. Journal of Virology. American Society for Microbiology (ASM); 1998 Jul;72(7):5820–30.

240. Agopian K, Wei BL, Garcia JV, Gabuzda D. A Hydrophobic Binding Surface on the Human Immunodeficiency Virus Type 1 Nef Core Is Critical for Association with p21-Activated Kinase 2. Journal of Virology. 2006 Feb 24;80(6):3050–61.

241. Bordo D, Argos P. Suggestions for “safe” residue substitutions in site-directed mutagenesis. Journal of Molecular Biology. 1991 Feb 20;217(4):721–9.

242. Buonaguro L, Tornesello ML, Buonaguro FM. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. Journal of Virology. American Society for Microbiology; 2007 Oct;81(19):10209–19.

243. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution. 1993 May;10(3):512–26.

244. Crystal Structure of the Conserved Core ofHIV-1 Nef Complexed with a Src Family SH3 Domain. 1997 Mar 25;:1–12.

245. Mwimanzi P, Hasan Z, Tokunaga M, Gatanaga H, Oka S, Ueno T. Naturally arising HIV-1 Nef variants conferring escape from cytotoxic T lymphocytes influence viral entry co-receptor expression and susceptibility to superinfection. Biochemical and Biophysical Research Communications. Elsevier Inc; 2010 Dec 17;403(3-4):422–7.

246. Lee CH, Saksela K, Mirza UA, Chait BT, Kuriyan J. Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. Cell. 1996 Jun 14;85(6):931–42.

247. Piguet V, Gu F, Foti M, Demaurex N, Gruenberg J, Carpentier JL, et al. Nef-induced CD4 degradation: a diacidic-based motif in Nef functions as a lysosomal targeting signal through the binding of beta-COP in endosomes. Cell. 1999 Apr 2;97(1):63–73.

248. Alter I, Gragert L, Fingerson S, Maiers M, Louzoun Y. HLA class I haplotype diversity is consistent with selection for frequent existing haplotypes. Weissing FJ, editor. PLoS Comput Biol. 2017 Aug;13(8):e1005693.

249. Isitman G, Stratov I, Kent SJ. Antibody-Dependent Cellular Cytotoxicity and NK Cell-Driven Immune Escape in HIV Infection: Implications for HIV Vaccine Development. Adv Virol. Hindawi Publishing Corporation; 2012;2012(5):637208–8.

250. Apps R, Qi Y, Carlson JM, Chen H, Gao X, Thomas R, et al. Influence of HLA-C Expression Level on HIV Control. Science. 2013 Apr 4;340(6128):87–91.

251. Wang Z, Liu H-W, Hong K-X, Yu Z-J, Chen J-P, Ruan Y-H, et al. Complete human immunodeficiency virus-1 specific T lymphocyte response to Chinese human immunodeficiency virus-1 B/C chronic infectors. Biomed Environ Sci. 2009 Dec;22(6):522–8.

252. Mothe B, Llano A, Ibarrondo J, Daniels M, Miranda C, Zamarreño J, et al. Definition of the viral targets of protective HIV-1-specific T cell responses. Journal of Translational Medicine. BioMed Central; 2011 Dec 7;9(1):208.

253. Kunwar P, Hawkins N, Dinges WL, Liu Y, Gabriel EE, Swan DA, et al. Superior control of HIV-1 replication by CD8+ T cells targeting conserved epitopes: implications for HIV vaccine design. Goepfert PA, editor. PLoS ONE. 2013;8(5):e64405.

254. Bernardin F, Kong D, Peddada L, Baxter-Lowe LA, Delwart E. Human

immunodeficiency virus mutations during the first month of infection are preferentially found in known cytotoxic T-lymphocyte epitopes. Journal of Virology. 2005 Sep;79(17):11523–8.

255. Ranki A, Nyberg M, Ovod V, Haltia M, Elovaara I, Raininko R, et al. Abundant expression of HIV Nef and Rev proteins in brain astrocytes in vivo is associated with dementia. AIDS. 1995 Sep;9(9):1001–8.

256. Kaufmann DE, Bailey PM, Sidney J, Wagner B, Norris PJ, Johnston MN, et al. Comprehensive analysis of human immunodeficiency virus type 1-specific CD4 responses reveals marked immunodominance of gag and nef and the presence of broadly recognized peptides. Journal of Virology. American Society for Microbiology (ASM); 2004 May;78(9):4463–77.

257. Fonseca SG, Coutinho-Silva A, Fonseca LAM, Segurado AC, Moraes SL, Rodrigues H, et al. Identification of novel consensus CD4 T-cell epitopes from clade B HIV-1 whole genome that are frequently recognized by HIV-1 infected patients. AIDS. 2006 Nov 28;20(18):2263–73.

258. Altfeld M, Rosenberg ES, Shankarappa R, Mukherjee JS, Hecht FM, Eldridge RL, et al. Cellular immune responses and viral diversity in individuals treated during acute and early HIV-1 infection. Journal of Experimental Medicine. The Rockefeller University Press; 2001 Jan 15;193(2):169–80.

259. Kitano M, Kobayashi N, Kawashima Y, Akahoshi T, Nokihara K, Oka S, et al. Identification and characterization of HLA-B*5401-restricted HIV-1-Nef and Pol-specific CTL epitopes. Microbes Infect. 2008 Jun;10(7):764–72.

260. Piguet V, Trono D. The Nef protein of primate lentiviruses. Rev Med Virol. 1999 Apr;9(2):111–20.

261. Li T, Steede NK, Nguyen H-NP, Freytag LC, McLachlan JB, Mettu RR, et al. Comprehensive analysis of contributions from protein conformational stability and major histocompatibility complex class II-peptide binding affinity to CD4+ epitope immunogenicity in HIV-1 envelope glycoprotein. Journal of Virology. American Society for Microbiology; 2014 Sep 1;88(17):9605–15.

262. Vossen M, Westerhout E, S x000F6 derberg-Naucl x000E9 r CXC, Wiertz E. Viral immune evasion: a masterpiece of evolution. Immunogenetics. 2002 Nov 1;54(8):527–42.

263. Schweighardt B, Wrin T, Meiklejohn DA, Spotts G, Petropoulos CJ, Nixon DF, et al. Immune escape mutations detected within HIV-1 epitopes associated with viral control during treatment interruption. J Acquir Immune Defic Syndr. 2010 Jan;53(1):36–46.

264. Day CL, Shea AK, Altfeld MA, Olson DP, Buchbinder SP, Hecht FM, et al. Relative

dominance of epitope-specific cytotoxic T-lymphocyte responses in human immunodeficiency virus type 1-infected persons with shared HLA alleles. Journal of Virology. American Society for Microbiology; 2001 Jul;75(14):6279–91.

265. Gudmundsdotter L, Bernasconi D, Hejdeman B, Sandstrom E, Alaeus A, Lidman K, et al. Cross-clade immune responses to Gag p24 in patients infected with different HIV-1 subtypes and correlation with HLA class I and II alleles. Vaccine. 2008 Sep 19;26(40):5182–7.

266. Prado JG, Honeyborne I, Brierley I, Puertas MC, Martinez-Picado J, Goulder PJR. Functional consequences of human immunodeficiency virus escape from an HLA-B*13-restricted CD8+ T-cell epitope in p1 Gag protein. Journal of Virology. 2009 Jan;83(2):1018–25.

267. Rousseau CM, Daniels MG, Carlson JM, Kadie C, Crawford H, Prendergast A, et al. HLA class I-driven evolution of human immunodeficiency virus type 1 subtype c proteome: immune escape and viral load. Journal of Virology. American Society for Microbiology; 2008 Jul;82(13):6434–46.

268. Avila Rios S, Ormsby CE, Carlson JM, Valenzuela Ponce H, Blanco-Heredia J, Garrido-Rodriguez D, et al. Unique features of HLA-mediated HIV evolution in a Mexican cohort: a comparative study. Retrovirology. BioMed Central; 2009 Aug 10;6(1):72.

269. Mocroft A, Phillips AN, Gatell J, Ledergerber B, Fisher M, Clumeck N, et al. Normalisation of CD4 counts in patients with HIV-1 infection and maximum virological suppression who are taking combination antiretroviral therapy: an observational cohort study. Lancet. Elsevier; 2007 Aug 4;370(9585):407–13.

270. Goldsmith MA, Warmerdam MT, Atchison RE, Miller MD, Greene WC. Dissociation of the CD4 downregulation and viral infectivity enhancement functions of human immunodeficiency virus type 1 Nef. Journal of Virology. American Society for Microbiology (ASM); 1995 Jul;69(7):4112–21.

271. Hadida F, Parrot A, Kieny MP, Sadat-Sowti B, Mayaud C, Debre P, et al. Carboxyl-terminal and central regions of human immunodeficiency virus-1 NEF recognized by cytotoxic T lymphocytes from lymphoid organs. An in vitro limiting dilution analysis. J Clin Invest. American Society for Clinical Investigation; 1992 Jan;89(1):53–60.

272. Culmann B, Gomard E, Kieny MP, Guy B, Dreyfus F, Saimot AG, et al. Six epitopes reacting with human cytotoxic CD8+ T cells in the central region of the HIV-1 NEF protein. The Journal of Immunology. 1991 Mar 1;146(5):1560–5.

273. Culmann-Penciolelli B, Lamhamedi-Cherradi S, Couillin I, Guegan N, Lévy JP, Guillet JG, et al. Identification of multirestricted immunodominant regions recognized by cytolytic T lymphocytes in the human immunodeficiency virus type 1 Nef protein. Journal of Virology. American Society for Microbiology (ASM); 1994

Nov;68(11):7336–43.

274. Milicic A, Price DA, Zimbwa P, Booth BL, Brown HL, Easterbrook PJ, et al. CD8+ T cell epitope-flanking mutations disrupt proteasomal processing of HIV-1 Nef. The Journal of Immunology. 2005 Oct 1;175(7):4618–26.

275. Crawford H, Prado JG, Leslie A, Hué S, Honeyborne I, Reddy S, et al. Compensatory mutation partially restores fitness and delays reversion of escape mutation within the immunodominant HLA-B*5703-restricted Gag epitope in chronic human immunodeficiency virus type 1 infection. Journal of Virology. American Society for Microbiology; 2007 Aug;81(15):8346–51.

276. Shahid A, Olvera A, Anmole G, Kuang XT, Cotton LA, Plana M, et al. Consequences of HLA-B*13-Associated Escape Mutations on HIV-1 Replication and Nef Function. Kirchhoff F, editor. Journal of Virology. 2015 Nov;89(22):11557–71.

277. Mwimanzi P, Markle TJ, Martin E, Ogata Y, Kuang XT, Tokunaga M, et al. Attenuation of multiple Nef functions in HIV-1 elite controllers. Retrovirology. Retrovirology; 2013 Jan 7;10(1):1–1.

278. Klotman ME, Kim S, Buchbinder A, DeRossi A, Baltimore D, Wong-Staal F. Kinetics of expression of multiply spliced RNA in early human immunodeficiency virus type 1 infection of lymphocytes and monocytes. Proc Natl Acad Sci USA. National Academy of Sciences; 1991 Jun 1;88(11):5011–5.

279. Lichterfeld M, Yu XG, Cohen D, Addo MM, Malenfant J, Perkins B, et al. HIV-1 Nef is preferentially recognized by CD8 T cells in primary HIV-1 infection despite a relatively high degree of genetic diversity. AIDS. 2004 Jul;18(10):1383–92.

280. Van Tran G, Chikata T, Carlson JM, Murakoshi H, Nguyen DH, Tamura Y, et al. A strong association of human leukocyte antigen-associated Pol and Gag mutations with clinical parameters in HIV-1 subtype A/E infection. AIDS. 2016 Mar 13;30(5):681–9.

281. Honda K, Zheng N, Murakoshi H, Hashimoto M, Sakai K, Borghan MA, et al. Selection of escape mutant by HLA-C-restricted HIV-1 Pol-specific cytotoxic T lymphocytes carrying strong ability to suppress HIV-1 replication. Eur J Immunol. WILEY-VCH Verlag; 2011 Jan;41(1):97–106.

282. Song H, Pavlicek JW, Cai F, Bhattacharya T, Li H, Iyer SS, et al. Impact of immune escape mutations on HIV-1 fitness in the context of the cognate transmitted/founder genome. Retrovirology. BioMed Central; 2012 Oct 30;9(1):89.

283. Smith SM. HIV CTL escape: at what cost? Retrovirology. BioMed Central; 2004 May 12;1(1):8.

284. Wright JK, Novitsky V, Brockman MA, Brumme ZL, Brumme CJ, Carlson JM, et al.

Influence of Gag-protease-mediated replication capacity on disease progression in individuals recently infected with HIV-1 subtype C. Journal of Virology. 2011 Apr;85(8):3996–4006.

285. Juarez-Molina CI, Payne R, Soto-Nava M, Avila Rios S, Valenzuela Ponce H, Adland E, et al. Impact of HLA selection pressure on HIV fitness at a population level in Mexico and Barbados. Journal of Virology. 2014 Sep;88(18):10392–8.

286. Kiepiela P, Ngumbela K, Thobakgale C, Ramduth D, Honeyborne I, Moodley E, et al. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. Nature Medicine. 2007 Jan;13(1):46–53.

287. Sáez-Cirión A, Lacabaratz C, Lambotte O, Versmisse P, Urrutia A, Boufassa F, et al. HIV controllers exhibit potent CD8 T cell capacity to suppress HIV infection ex vivo and peculiar cytotoxic T lymphocyte activation phenotype. Proc Natl Acad Sci USA. National Acad Sciences; 2007 Apr 17;104(16):6776–81.

288. Carlson JM, Listgarten J, Pfeifer N, Tan V, Kadie C, Walker BD, et al. Widespread impact of HLA restriction on immune control and escape pathways of HIV-1. Journal of Virology. 2012 May;86(9):5230–43.

289. Kløverpris HN, Cole DK, Fuller A, Carlson J, Beck K, Schauenburg AJ, et al. A molecular switch in immunodominant HIV-1-specific CD8 T-cell epitopes shapes differential HLA-restricted escape. Retrovirology. BioMed Central; 2015 Feb 20;12(1):20.

290. Schmittel A, Keilholz U, Bauer S, Kuhne U, Stevanovic S, Thiel E, et al. Application of the IFN-gamma ELISPOT assay to quantify T cell responses against proteins. J Immunol Methods. 2001 Jan 1;247(1-2):17–24.

291. Geels MJ, Cornelissen M, Schuitemaker H, Anderson K, Kwa D, Maas J, et al. Identification of sequential viral escape mutants associated with altered T-cell responses in a human immunodeficiency virus type 1-infected individual. Journal of Virology. American Society for Microbiology (ASM); 2003 Dec;77(23):12430–40.

292. Rolland M, Frahm N, Nickle DC, Jojic N, Deng W, Allen TM, et al. Increased breadth and depth of cytotoxic T lymphocytes responses against HIV-1-B Nef by inclusion of epitope variant sequences. Brusic V, editor. PLoS ONE. 2011 Mar 28;6(3):e17969.

293. Allen TM, Yu XG, Kalife ET, Reyor LL, Lichterfeld M, John M, et al. De novo generation of escape variant-specific CD8+ T-cell responses following cytotoxic T-lymphocyte escape in chronic human immunodeficiency virus type 1 infection. Journal of Virology. American Society for Microbiology; 2005 Oct;79(20):12952–60.

294. O'Connell KA, Hegarty RW, Siliciano RF, Blankson JN. Viral suppression of multiple escape mutants by de novo CD8(+) T cell responses in a human immunodeficiency virus-1 infected elite suppressor. Retrovirology. BioMed Central; 2011 Aug

3;8(1):63.

295. Bailey JR, Williams TM, Siliciano RF, Blankson JN. Maintenance of viral suppression in HIV-1-infected HLA-B*57+ elite suppressors despite CTL escape mutations. Journal of Experimental Medicine. 2006 May 15;203(5):1357–69.

296. Navis M, Schellens I, van Baarle D, Borghans J, van Swieten P, Miedema F, et al. Viral replication capacity as a correlate of HLA B57/B5801-associated nonprogressive HIV-1 infection. The Journal of Immunology. 2007 Sep 1;179(5):3133–43.

297. Streeck H, Brumme ZL, Anastario M, Cohen KW, Jolin JS, Meier A, et al. Antigen load and viral sequence diversification determine the functional profile of HIV-1-specific CD8+ T cells. Rowland-Jones S, editor. PLoS Med. Public Library of Science; 2008 May 6;5(5):e100.

298. O'Connell KA, Brennan TP, Bailey JR, Ray SC, Siliciano RF, Blankson JN. Control of HIV-1 in elite suppressors despite ongoing replication and evolution in plasma virus. Journal of Virology. American Society for Microbiology; 2010 Jul;84(14):7018–28.

299. Allen TM, Altfeld M, Geer SC, Kalife ET, Moore C, O'sullivan KM, et al. Selective escape from CD8+ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution. Journal of Virology. American Society for Microbiology; 2005 Nov;79(21):13239–49.

300. Pohlmeyer CW, Buckheit RW, Siliciano RF, Blankson JN. CD8+ T cells from HLA-B*57 elite suppressors effectively suppress replication of HIV-1 escape mutants. Retrovirology. BioMed Central; 2013 Dec 12;10(1):152.

301. Lichterfeld M, Yu XG, Le Gall S, Altfeld M. Immunodominance of HIV-1-specific CD8(+) T-cell responses in acute HIV-1 infection: at the crossroads of viral and host genetics. Trends Immunol. Elsevier; 2005 Mar;26(3):166–71.

302. Singh MK, Janvier G, Calvez V, Coulaud P, Rivière Y. A long-term follow-up of an HIV type 1-infected patient reveals a coincidence of Nef-directed cytotoxic T lymphocyte effectors and high incidence of epitope-deleted variants. AIDS Res Hum Retroviruses. Mary Ann Liebert, Inc; 2001 Sep 1;17(13):1265–71.

303. Haas G, Plikat U, Debre P, Lucchiari M, Katlama C, Dudoit Y, et al. Dynamics of viral variants in HIV-1 Nef and specific cytotoxic T lymphocytes in vivo. The Journal of Immunology. 1996 Nov 1;157(9):4212–21.

304. Migueles SA, Laborico AC, Imamichi H, Shupert WL, Royce C, McLaughlin M, et al. The differential ability of HLA B*5701+ long-term nonprogressors and progressors to restrict human immunodeficiency virus replication is not caused by loss of recognition of autologous viral gag sequences. Journal of Virology. American

Society for Microbiology (ASM); 2003 Jun;77(12):6889–98.

305. Semeniuk CA, McKinnon L, Peters HO, Gubbins M, Mao X, Ball TB, et al. Multiple T-cell epitopes overlap positively-selected residues in the p1 spacer protein of HIV-1 gag. AIDS. 2009 Apr 27;23(7):771–7.

306. Martínez-Bonet M, Palladino C, Briz V, Rudolph JM, Fackler OT, Relloso M, et al. A Conserved GPG-Motif in the HIV-1 Nef Core Is Required for Principal Nef-Activities. Wu Y, editor. PLoS ONE. Public Library of Science; 2015;10(12):e0145239.

307. Foster JL, Denial SJ, Temple BRS, Garcia JV. Mechanisms of HIV-1 Nef Function and Intracellular Signaling. J Neuroimmune Pharmacol. 2011 Feb 19;6(2):230–46.

308. Malhotra U, Nolin J, Horton H, Li F, Corey L, Mullins JI, et al. Functional properties and epitope characteristics of T-cells recognizing natural HIV-1 variants. Vaccine. 2009 Nov;27(48):6678–87.

309. Migueles SA, Laborico AC, Shupert WL, Sabbaghian MS, Rabin R, Hallahan CW, et al. HIV-specific CD8+ T cell proliferation is coupled to perforin expression and is maintained in nonprogressors. Nature. 2002 Nov;3(11):1061–8.

310. Cossarizza A, Bertoncelli L, Nemes E, Lugli E, Pinti M, Nasi M, et al. T cell activation but not polyfunctionality after primary HIV infection predicts control of viral load and length of the time without therapy. Gray CM, editor. PLoS ONE. Public Library of Science; 2012;7(12):e50728.

311. Cao J, McNevin J, Malhotra U, McElrath MJ. Evolution of CD8+ T cell immunity and viral escape following acute HIV-1 infection. The Journal of Immunology. 2003 Oct 1;171(7):3837–46.

312. Yamamoto T, Price DA, Casazza JP, Ferrari G, Nason M, Chattopadhyay PK, et al. Surface expression patterns of negative regulatory molecules identify determinants of virus-specific CD8+ T-cell exhaustion in HIV infection. Blood. American Society of Hematology; 2011 May 5;117(18):4805–15.

313. Honeyborne I, Prendergast A, Pereyra F, Leslie A, Crawford H, Payne R, et al. Control of human immunodeficiency virus type 1 is associated with HLA-B*13 and targeting of multiple gag-specific CD8+ T-cell epitopes. Journal of Virology. American Society for Microbiology; 2007 Apr;81(7):3667–72.

314. Blackburn SD, Wherry EJ. IL-10, T cell exhaustion and viral persistence. Trends in Microbiology. 2007 Apr;15(4):143–6.

315. Day CL, Kaufmann DE, Kiepiela P, Brown JA, Moodley ES, Reddy S, et al. PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. Nature. 2006 Sep 21;443(7109):350–4.

316. Zhang J-Y, Zhang Z, Wang X, Fu J-L, Yao J, Jiao Y, et al. PD-1 up-regulation is correlated with HIV-specific memory CD8+ T-cell exhaustion in typical progressors but not in long-term nonprogressors. Blood. American Society of Hematology; 2007 Jun 1;109(11):4671–8.

317. Rolland M, Heckerman D, Deng W, Rousseau CM, Coovadia H, Bishop K, et al. Broad and Gag-biased HIV-1 epitope repertoires are associated with lower viral loads. Nixon D, editor. PLoS ONE. Public Library of Science; 2008 Jan 9;3(1):e1424.

318. Zuniga R, Lucchetti A, Galvan P, Sanchez S, Sanchez C, Hernandez A, et al. Relative dominance of Gag p24-specific cytotoxic T lymphocytes is associated with human immunodeficiency virus control. Journal of Virology. 2006 Mar;80(6):3122–5.

319. Pereyra F, Heckerman D, Carlson JM, Kadie C, Soghoian DZ, Karel D, et al. HIV control is mediated in part by CD8+ T-cell targeting of specific epitopes. Journal of Virology. 2014 Nov;88(22):12937–48.

320. Gallimore A, Glithero A, Godkin A, Tissot AC, Plückthun A, Elliott T, et al. Induction and exhaustion of lymphocytic choriomeningitis virus-specific cytotoxic T lymphocytes visualized using soluble tetrameric major histocompatibility complex class I-peptide complexes. Journal of Experimental Medicine. The Rockefeller University Press; 1998 May 4;187(9):1383–93.

321. Shedlock DJ, Talbott KT, Morrow MP, Ferraro B, Hokey DA, Muthumani K, et al. Ki-67 staining for determination of rhesus macaque T cell proliferative responses ex vivo. Cytometry A. Wiley Subscription Services, Inc., A Wiley Company; 2010 Mar;77(3):275–84.

322. Scholzen T, Gerdes J. The Ki-67 protein: from the known and the unknown. J Cell Physiol. John Wiley & Sons, Inc; 2000 Mar;182(3):311–22.

323. Amu S, Lantto Graham R, Bekele Y, Nasi A, Bengtsson C, Rethi B, et al. Dysfunctional phenotypes of CD4+ and CD8+ T cells are comparable in patients initiating ART during early or chronic HIV-1 infection. Medicine (Baltimore). 2016 Jun;95(23):e3738.

324. Massanella M, Gianella S, Schrier R, Dan JM, Pérez-Santiago J, Oliveira MF, et al. Methamphetamine Use in HIV-infected Individuals Affects T-cell Function and Viral Outcome during Suppressive Antiretroviral Therapy. Sci Rep. Nature Publishing Group; 2015 Aug 24;5(1):13179.

325. Hoffmann M, Pantazis N, Martin GE, Hickling S, Hurst J, Meyerowitz J, et al. Exhaustion of Activated CD8 T Cells Predicts Disease Progression in Primary HIV-1 Infection. Swanstrom R, editor. PLoS Pathog. 2016 Jul;12(7):e1005661.

326. Jones N, Eggena M, Baker C, Nghania F, Baliruno D, Mugyenyi P, et al. Presence of distinct subsets of cytolytic CD8+ T cells in chronic HIV infection. AIDS Res Hum

Retroviruses. Mary Ann Liebert, Inc. 2 Madison Avenue Larchmont, NY 10538 USA; 2006 Oct;22(10):1007–13.

327. Addo MM, Draenert R, Rathod A, Verrill CL, Davis BT, Gandhi RT, et al. Fully differentiated HIV-1 specific CD8+ T effector cells are more frequently detectable in controlled than in progressive HIV-1 infection. Zhang L, editor. PLoS ONE. 2007 Mar 28;2(3):e321.

328. Mahnke YD, Brodie TM, Sallusto F, Roederer M, Lugli E. The who's who of T-cell differentiation: human memory T-cell subsets. Eur J Immunol. 2013 Nov;43(11):2797–809.

329. Killian MS, Johnson C, Teque F, Fujimura S, Levy JA. Natural suppression of human immunodeficiency virus type 1 replication is mediated by transitional memory CD8+ T cells. Journal of Virology. 2011 Feb;85(4):1696–705.

330. Northfield JW, Loo CP, Barbour JD, Spotts G, Hecht FM, Klenerman P, et al. Human immunodeficiency virus type 1 (HIV-1)-specific CD8+ T(EMRA) cells in early infection are linked to control of HIV-1 viremia and predict the subsequent viral load set point. Journal of Virology. 2007 Jun;81(11):5759–65.

331. Matthews PC, Koyanagi M, Kløverpris HN, Harndahl M, Stryhn A, Akahoshi T, et al. Differential clade-specific HLA-B*3501 association with HIV-1 disease outcome is linked to immunogenicity of a single Gag epitope. Journal of Virology. 2012 Dec;86(23):12643–54.

332. Dinges WL, Richardt J, Friedrich D, Jalbert E, Liu Y, Stevens CE, et al. Virus-specific CD8+ T-cell responses better define HIV disease progression than HLA genotype. Journal of Virology. American Society for Microbiology; 2010 May;84(9):4461–8.

333. Stoycheva D, Deiser K, Stärck L, Nishanth G, Schlüter D, Uckert W, et al. IFN-γ regulates CD8+ memory T cell differentiation and survival in response to weak, but not strong, TCR signals. J Immunol. American Association of Immunologists; 2015 Jan 15;194(2):553–9.

334. Turk G, Gherardi MM, Laufer N, Saracco M, Luzzi R, Cox JH, et al. Magnitude, breadth, and functional profile of T-cell responses during human immunodeficiency virus primary infection with B and BF viral variants. Journal of Virology. 2008 Mar;82(6):2853–66.

335. López M, Peris A, Soriano V, Lozano S, Vicario JL, Rallón NI, et al. The expansion ability but not the quality of HIV-specific CD8(+) T cells is associated with protective human leucocyte antigen class I alleles in long-term non-progressors. Immunology. Blackwell Publishing Ltd; 2011 Nov;134(3):305–13.

336. Betts MR, Nason MC, West SM, De Rosa SC, Migueles SA, Abraham J, et al. HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells.

Blood. American Society of Hematology; 2006 Jun 15;107(12):4781–9.

337. Riou C, Burgers WA, Mlisana K, Koup RA, Roederer M, Abdool Karim SS, et al. Differential impact of magnitude, polyfunctional capacity, and specificity of HIV-specific CD8+ T cell responses on HIV set point. Journal of Virology. American Society for Microbiology; 2014 Feb;88(3):1819–24.

338. Acevedo-Sáenz L, Carmona-Pérez L, Velilla-Hernández PA, Delgado JC, Rugeles L MT. The APPEESFRS Peptide, Restricted by the HLA-B*35:01 Molecule, and the APPEESFRF Variant Derived from an Autologous HIV-1 Strain Induces Polyfunctional Responses in CD8+ T Cells. Biores Open Access. 2015;4(1):115–20.

339. Crespo J, Sun H, Welling TH, Tian Z, Zou W. T cell anergy, exhaustion, senescence, and stemness in the tumor microenvironment. Curr Opin Immunol. 2013 Apr;25(2):214–21.

340. Rosa DS, Ribeiro SP, Fonseca SG, Almeida RR, Santana VC, Apostólico J de S, et al. Multiple Approaches for Increasing the Immunogenicity of an Epitope-Based Anti-HIV Vaccine. AIDS Res Hum Retroviruses. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA; 2015 Nov;31(11):1077–88.

341. Rappuoli R, Medaglini D. Big science for vaccine development. Vaccine. 2014 Aug 20;32(37):4705–7.

342. McEnery R. HVTN 505 trial expanded to see if vaccine candidates can block HIV acquisition. IAVI Rep. 2011 Jul;15(4):17.

343. Gottardo R, Bailer RT, Korber BT, Gnanakaran S, Phillips J, Shen X, et al. Plasma IgG to linear epitopes in the V2 and V3 regions of HIV-1 gp120 correlate with a reduced risk of infection in the RV144 vaccine efficacy trial. Chen Z, editor. PLoS ONE. Public Library of Science; 2013;8(9):e75665.

344. Chung AW, Ghebremichael M, Robinson H, Brown E, Choi I, Lane S, et al. Polyfunctional Fc-effector profiles mediated by IgG subclass selection distinguish RV144 and VAX003 vaccines. Sci Transl Med. 2014 Mar 19;6(228):228ra38–8.

345. Hammer SM, Sobieszczyk ME, Janes H, Karuna ST, Mulligan MJ, Grove D, et al. Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine. N Engl J Med. Massachusetts Medical Society; 2013 Nov 28;369(22):2083–92.

346. de Souza MS, Ratto-Kim S, Chuenarom W, Schuetz A, Chantakulkij S, Nuntapinit B, et al. The Thai phase III trial (RV144) vaccine regimen induces T cell responses that preferentially target epitopes within the V2 region of HIV-1 envelope. J Immunol. American Association of Immunologists; 2012 May 15;188(10):5166–76.

347. Dommaraju K, Kijak G, Carlson JM, Larsen BB, Tovanabutra S, Geraghty DE, et al. CD8 and CD4 epitope predictions in RV144: no strong evidence of a T-cell driven

sieve effect in HIV-1 breakthrough sequences from trial participants. Goepfert PA, editor. PLoS ONE. Public Library of Science; 2014;9(10):e111334.

348. Liao H-X, Bonsignori M, Alam SM, McLellan JS, Tomaras GD, Moody MA, et al. Vaccine induction of antibodies against a structurally heterogeneous site of immune pressure within HIV-1 envelope protein variable regions 1 and 2. Immunity. 2013 Jan 24;38(1):176–86.

349. Brown RJP, Hudson N, Wilson G, Rehman SU, Jabbari S, Hu K, et al. Hepatitis C Virus Envelope Glycoprotein Fitness Defines Virus Population Composition following Transmission to a New Host. Journal of Virology. 2012 Oct 18;86(22):11956–66.

350. Horton H, Frank I, Baydo R, Jalbert E, Penn J, Wilson S, et al. Preservation of T cell proliferation restricted by protective HLA alleles is critical for immune control of HIV-1 infection. The Journal of Immunology. 2006 Nov 15;177(10):7406–15.

351. Rolland M, Manocheewa S, Swain JV, Lanxon-Cookson EC, Kim M, Westfall DH, et al. HIV-1 conserved-element vaccines: relationship between sequence conservation and replicative capacity. Journal of Virology. American Society for Microbiology; 2013 May;87(10):5461–7.

352. Hu X, Valentin A, Rosati M, Manocheewa S, Alicea C, Chowdhury B, et al. HIV Env Conserved Element DNA Vaccine Alters Immunodominance in Macaques. Hum Vaccin Immunother. 2017 Jul 5;:0–00.

353. Kulkarni V, Valentin A, Rosati M, Rolland M, Mullins JI, Pavlakis GN, et al. HIV-1 conserved elements p24CE DNA vaccine induces humoral immune responses with broad epitope recognition in macaques. Bansal GP, editor. PLoS ONE. Public Library of Science; 2014;9(10):e111085.

354. Borthwick N, Ahmed T, Ondondo B, Hayes P, Rose A, Ebrahimsa U, et al. Vaccine-elicited human T cells recognizing conserved protein regions inhibit HIV-1. Mol Ther. 2014 Feb;22(2):464–75.

355. Corey L, McElrath MJ. HIV vaccines: mosaic approach to virus diversity. Nature Medicine. 2010 Mar;16(3):268–70.

356. Barouch DH, O'Brien KL, Simmons NL, King SL, Abbink P, Maxfield LF, et al. Mosaic HIV-1 vaccines expand the breadth and depth of cellular immune responses in rhesus monkeys. Nature Medicine. 2010 Mar;16(3):319–23.

357. Ondondo B, Murakoshi H, Clutton G, Abdul-Jawad S, Wee EG-T, Gatanaga H, et al. Novel Conserved-region T-cell Mosaic Vaccine With High Global HIV-1 Coverage Is Recognized by Protective Responses in Untreated Infection. Mol Ther. 2016 Apr;24(4):832–42.

358. Abdul-Jawad S, Ondondo B, van Hateren A, Gardner A, Elliott T, Korber B, et al.

Increased Valency of Conserved-mosaic Vaccines Enhances the Breadth and Depth of Epitope Recognition. Mol Ther. 2016 Feb;24(2):375–84.

359. Khoja S, Ojwang P, Khan S, Okinda N, Harania R, Ali S. Genetic analysis of HIV-1 subtypes in Nairobi, Kenya. Carr JK, editor. PLoS ONE. Public Library of Science; 2008 Sep 11;3(9):e3191.

9. Appendices

9.1 Abbreviations

Ab: antibody

ADCC: antibody-dependent cellular cytotoxity

AIDS: Acquired Immune Deficiency Syndrome

AP-2: Adaptor protein 2

APCs: Antigen presenting cells

ART: Antiretroviral therapy

ARV: Antiretroviral

bNab: Broadly neutralizing antibody

CCR5: chemokine co-receptor 5

CD: cluster differentiation

CMV: Cytomegalovirus

CRF: circulating recombinant form

CSWs: Commercial sex workers

CTL: Cytotoxic lympohcytes

DCs: Dendritic cells

DMSO: Dimethyl sulfoxide

EC: Elite controller

EDTA: Ethylenediamine tetraacetic acid

ELISA: Enzyme-linked immunosorbent assay

ELISPOT: Enzyme Linked ImmunoSpot

Env: Envelope

FACS: Fluorescent activated cell sorting

FBS: fetal bovine serum

Gag: Group-specific antigen

GALT: Gut associated lymph tissue

GI: Gastrointestinal

GM-CSF: Granulocyte macrophage costimulatory factor

GWAS: Genome wide association study

HESN: HIV Exposed Sero-negative

HIV-1: Human Immunodeficiency Virus-1

HLA: Human leukocyte antigen

HSV: Herpes Simplex Virus

ICS: Intracellular cytokine staining

IDUs: Injection drug users

IFNγ: Interferon gamma

Ig: Immunoglobulin

IL: Interleukin

KIR: Killer immunoglobulin-like receptors

LAG: Lymphocyte-activation gene 3

LANL: Los Alamos National Library

LPS: Lipopolysaccharides

LTNP: Long-term non-progressor

MHC: Major histocompatibility complex

MIP1: Macrophage inflammatory protein

MSM: Men who have sex with men

MTCT: Mother to child transmission

Nef: Negative effector

NIH: National Institute of Health

NK cells: Natural killer cells

NNRTI: Non- nucleotide reverse transcriptase inhibitors

NRTI: Nucleotide reverse transcriptase inhibitors

Pak2: P21 activated kinase

PAMP: Pathogen-associated molecular patterns

PBMC: Peripheral blood mononuclear cells

PBS: Phosphate buffered saline

PD-1: Programmed death 1

PEP: post exposure prophylaxis

PEPFAR: Presidents Emergency Preparedness for AIDS Relief

PFA: paraformaldehyde

PHA: Phytohemagglutanin

PMA: Phorbol 12-myristate 13-acetate

PMTCT: Prevention of mother to child transmission

Pol: Polymerase

PrEP: Pre-exposure prophylaxis

PRR: Pattern Recognition Receptors

PSM: Positively selected mutations

pVL: Plasma viral load

RP: Rapid progressor

RPMI: Roswell Park Memorial Institute

RT: Reverse transcriptase

SBBC: Sydney Blood Bank Cohort

SFU: spot forming units

SIV: Simian Immunodeficiency Virus

STIs: Sexually transmitted infections

TGN: Trans golgi network

Th17: T-helper-17

Tim3: T-cell immunoglobulin and mucin-domain containing-3

TLR: Toll-like receptor

TNF: Tumour necrosis factor

URF: Unique recombinant form

VL: Viral load

VMMC: Voluntary male medical circumcision

WHO: World Health Organization