Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Jessica Lee Prince Guerra Date

HIV-1 Replicative Capacity Predicts Disease Progression and is Associated with Early Pathogenesis

By

Jessica Lee Prince Guerra
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Microbiology and Molecular Genetics

_____

Eric Hunter, Ph.D., Advisor

_____ _____

Cynthia Derdeyn, Ph.D.
Committee Member

David Steinhauer, Ph.D.
Committee Member

_____ _____

Arash Grakoui, Ph.D.
Committee Member

Samuel Speck, Ph.D.
Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

HIV-1 Replicative Capacity Predicts Disease Progression and is Associated with Early Pathogenesis

By

Jessica Lee Prince Guerra
B.S., Pacific University, 2007

Advisor: Eric Hunter, Ph.D.

An abstract of a dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Graduate Division of Biological and Biomedical Science
Microbiology and Molecular Genetics
2014

Abstract

HIV-1 infection is characterized by a gradual decline in peripheral CD4+ T cells and generalized immune dysfunction. While a majority of HIV-infected individuals eventually progress to AIDS, they do so at varying rates. Determining the host and viral factors that shape the trajectory of HIV-1 pathogenesis is key for developing rational prevention strategies. To date, research has focused on elucidating host factors associated with disease progression while viral factors have been relatively understudied.

The following study utilizes a well-characterized cohort of subtype C Zambian seroconverters in order to investigate how replicative capacity of the transmitted virus, as defined by Gag, contributes to HIV-1 disease progression and pathogenesis. From over 200 Zambian seroconverters, we created Gag-chimeric HIV-1 viruses in which the patient-derived *gag* gene was isolated from acute time points and cloned into MJ4. In individuals recently infected with HIV-1 subtype C, low viral replicative capacity as defined by the transmitted Gag sequence, was associated with a delayed loss of CD4+ T cells independent of set point VL and host immunogenetic factors. We hypothesized that early viral replication might initiate crucial events during early infection that influence HIV-1 pathogenesis.

Chronic immune activation is a hallmark of HIV-1 infection and predicts disease progression better than viral load. We show that attenuated replicative capacity leads to dampened levels of immune activation. Furthermore, we found that in individuals infected with low replicating viruses, CD8+ T cells were less exhausted and more cytotoxic perhaps leading to a more functional immune response. Moreover, replicative capacity was positively correlated with CD4+ T cell proliferation and with levels of HIV DNA in CD4+ central memory T cells, a population highlighted to be integral for the maintenance of latency and preferentially spared in non-pathogenic SIV infection. Consistent with previous studies, we observed that all of these measures of immune dysfunction were associated with the rate of disease progression in this cohort. Collectively, this provides a mechanistic link between replicative capacity and CD4+ T cell decline.

This study highlights the previously unrecognized role that replicative capacity of the transmitted virus plays in defining several facets of HIV-1 immunopathology and disease progression.

HIV-1 Replicative Capacity Predicts Disease Progression and is Associated with Early Pathogenesis

By

Jessica Lee Prince Guerra
B.S., Pacific University, 2007

Advisor: Eric Hunter, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Graduate Division of Biological and Biomedical Science
Microbiology and Molecular Genetics
2014

Dedication

In loving memory of my mom, Georgia Ann Prince. Her kindness, humor, and strength shaped the woman I have become. She will always be in my heart guiding me through this adventure we call life.

Acknowledgements

I would like to thank:

- My dad for unlimited love, support, and patience, and for always indulging in my childish curiosities.

- For my brother, who has always shown me what it means to be resilient in life and who reminds me of the important things in life.

- To my aunt, Pam, who has been like a second-mom to me. I appreciate all of your love and support throughout the years.

- To my husband, Leo. You have been with me through all the ups and downs both in grad school and in life; for that I cannot thank-you enough. You have inspired me to become a better person and have given my life so much meaning. I cannot wait for this next chapter in our lives.

- To my high school Spanish teacher and life-mentor, Señora Judy Sprunger, who taught me to take risks, to enjoy life's smaller moments, and to go after my dreams with my heart and mind open. She always encouraged me to think outside the box and challenged me to combine my love of languages and cultures with my passion for science.

- To Dan: I couldn't have asked for a better partner in lab. Working with you has made lab and grad school fun, exciting, and has allowed me to always see the light at the end of the tunnel even when things weren't working out. I have become a better scientist and person thanks to your influence. We are a team and I couldn't have done all this without you!

- To my boss, Eric, who has always given me endless opportunities and has shaped my career in ways I probably still don't comprehend yet. I only hope that I will become half the mentor you are one day.

- To Debby, for always being such a great friend, no matter what the circumstance and for all of the dessert breaks!

- To all the other Hunter lab members (who are getting too numerous to count!). Paul, Jon, Effie, and Sheng for doing all the little things that are so very critical to our projects and really maintain sanity in the working environment. To all the other grad students past and present for always being willing to lend a hand, who have taught me experiences, and with whom I can share in the grad school experience – Martin, Jasmine, Ted, Zach, Debby, Dan, and Liang

- To Dr. Susan Allen and all who work at the Zambia Emory HIV Research Project for the countless dedication towards founding and establishing this research.

- To all of the human volunteers in our study, because without them, none of this would have been possible.

Table of Contents

Page

Chapter 1: Introduction

Global burden of HIV/AIDS

Acquired Immune Deficiency Syndrome (AIDS) was first recognized as a new emerging human disease in a case-report study published by the CDC in 1981 [1,2]. This report documented five cases of homosexual men who had succumbed to unusual opportunistic infections and malignancies that were more frequently manifested in individuals with a compromised immune system. Soon thereafter in 1983, a retrovirus termed Lymphadenopathy Associated Virus (LAV)/Human T-cell Lymphotropic Virus III (HTLV-III) – later to be named Human Immunodeficiency Virus type 1 (HIV-1) – was isolated from an AIDS patient and identified as the etiological agent of the disease [3,4]. HIV/AIDS has since been described as one of the most devastating infectious diseases in history. The rapid spread of HIV/AIDS has resulted in a substantial burden of health worldwide.

According to the most recent UNAIDS report in 2012, an estimated 35.5 million individuals are living with HIV worldwide which includes around 2.3 million newly documented cases that year alone. Despite a significant reduction in the growth of the epidemic over the past few decades, this decline is not uniformly distributed across all regions and groups [5]. A majority of HIV-1 infections worldwide, 26.6 million, occur in Sub-Saharan Africa with over 80% of new infections attributed to heterosexual transmission worldwide. Zambia is one of the most urbanized countries of Sub-Saharan Africa, and as a result, has one of the highest per capita rates of HIV-1 infection in the world. In 2009, approximately 15% of the estimated Zambian population of 13 million was infected with HIV-1 (UNAIDS 2012) [5]. Despite the fact that over 95% of the Sub-Sahara African HIV-1 infections are from subtype C HIV-1 viruses [6] the majority of

HIV research has focused on subtype B viruses. As such, future HIV research efforts should likely focus on studying subtype C HIV-1 viruses, because this subtype represents over 50% of HIV infections worldwide [7,8].

The development of anti-retroviral treatment (ART) was regarded as a major turning point in combating the HIV/AIDS epidemic. Its success comes from the ability to reduce an individual's viral load, often times to undetectable levels, and therefore prolongs life expectancy and improves the quality of life [9]. Despite this major advancement, ART by itself has proven unable to significantly impact the epidemic. Access to ART within the US is relatively high compared to developing countries, and yet, infection rates have not drastically declined to manageable levels as was originally hoped [10]. Adherence issues are of particular concern because intermittent discontinuation of ART can lead to development of resistance, which can then lead to an increase in viral load. Transmission of ART-resistant variants within populations would have dire consequences on the efforts to treat the epidemic with drug therapy alone. Prolonged and/or lifelong ART therapy can result in adverse effects in treated individuals. Furthermore, there is also the question of whether such continued use of ART for an ever-growing population of HIV infected individuals eligible for treatment is economically sustainable [10,11]. It is therefore imperative to investigate other preventative and therapeutic strategies. Identifying the components of a vaccine that would prevent infection is a key goal [12], but alternatively, if sterilizing immunity cannot be achieved through vaccination in all cases, therapeutic strategies aimed at functionally curing infected individuals through eradication of viral reservoirs and treatment of immune dysfunction should also be explored.

Origin, classification, and viral diversity of HIV-1

HIV can be subdivided into two related, but antigenically distinct viruses that can both cause AIDS: HIV-1 and HIV-2. HIV-2 was discovered in 1986 [13] and is more closely related to simian immunodeficiency virus (SIV) than to HIV-1 [14,15]. Various species of primates harbor their own genetically distinct SIV viruses [16]. Surprisingly, natural SIV infection of nonhuman primates is generally nonpathogenic, which perhaps indicates significant co-evolution and adaptation between host and virus [17].

Sequencing and phylogenetic analysis of HIV and SIV isolates led to significant advances in elucidating the origins and classification of HIV [18-21]. HIV can be categorized into four distinct lineages termed group M or Main, N, O, and P, each the result of independent cross-species transmission events from non-human primates [18,20,22]. HIV-1 group M was transmitted by SIVcpz infected *Pan troglodytes troglodytes* chimpanzees in West Central Africa during the early 1900s [20,22-25]. HIV-1 group M is the most prevalent group worldwide and is now considered to be responsible for the global pandemic of HIV-1 that emerged in the early 1980's. HIV-1 group M is further subcategorized into 9 phylogenetically distinct subtypes (A, B, C, D, F, G, H, J, K), which have varying distributions worldwide [8,21,26,27]. HIV-1 subtype C accounts for slightly more than 50% of infections worldwide and predominantly occurs within sub-Saharan Africa while subtype B infection occurs primarily in the Americas and Europe [28]. The other HIV-1 groups (N, O, P) are much less prevalent and are geographically limited to Cameroon and neighboring countries [29-31]. More in-depth studies of SIVcpz infected chimpanzees have provided insights into the pathogenicity of this virus. Interestingly, it was demonstrated that SIVcpz infected chimpanzees had an increased risk of mortality in comparison to uninfected chimpanzees and displayed signs of CD4+T cell decline [32].

HIV-2 emerged as a result of cross-species transmission from SIVsmm infected sooty mangabeys [33] and can be subdivided into distinct lineages termed groups A-H, although only groups A and B have spread significantly [21,34,35]. HIV-2 is much less prevalent than HIV-1 and is largely restricted to West Africa [36]. HIV-2 is generally less pathogenic than HIV-1. Overall, individuals infected with HIV-2 tend to have lower viral loads, which might explain the lower transmission rate [37]. HIV-2 also causes slower rates in CD4+T cell decline, longer periods of asymptomatic infection, and is associated with decreased mortality rates [38].

A defining characteristic of all retroviruses is the unique mechanism by which they replicate their genome. Retroviruses, like HIV-1, use a reverse transcriptase enzyme in order to copy their positive sense single stranded RNA genome into double stranded DNA, which is then integrated into the host's chromosomal DNA. This process is highly error-prone with an estimated 3×10^{-5} mutations/nucleotide/replication cycle [39]. This allows for rapid viral evolution and adaption to selective pressures present within the host.

The extraordinary amount of HIV-1 genetic diversity can be attributed to several factors including the highly error-prone reverse transcription of +ssRNA into dsDNA [39], high frequencies of recombination [40,41], the high rate of viral replication [42,43], and large numbers of infected individuals worldwide. Intra-subtype and inter-subtype diversity is estimated to be upwards of 8-17% and 17-35% in the *env* gene, respectively [28,41,44]. Additionally, significant viral evolution occurs within a single host leading to a HIV-1 quasispecies with estimates as high as 10% for intra-host viral diversity in the *env* gene [45]. Using advanced sequencing techniques, such as single-genome amplification and 454 deep sequencing, it has been shown that viral sequence diversification occurs early post-infection synchronous with the establishment of a steady state set point viral load [46,47]. During this time, an accumulation of both

synonymous and non-synonymous nucleotide changes can be attributed to viral escape from both neutralizing antibody (Nab) and cytotoxic T lymphocyte (CTL) pressure [12,46-52]. Thus, the inherent propensity for HIV-1 viral evolution and immune escape makes designing a successful preventative vaccine able to block infection of all known genetic variants extremely difficult [12,41,48].

HIV-1 transmission

HIV is most often transmitted through exposure at mucosal surfaces. Common routes of infection include perinatal (from mother-to-child), sexual, injection drug use, and from blood transfusions [53]. Despite the fact that the risk of HIV-1 infection from heterosexual exposure is the lowest in comparison to other routes (between 1/200 to 1/3000) [53,54], heterosexual transmission accounts for around 70% of new HIV-1 infections worldwide. Additionally, a significant percentage of infections due to heterosexual transmission within sub-Saharan Africa occur between serodiscordant couples [55]. Serodiscordant couples are characterized by the fact that one individual in the partnership is HIV+ while the other partner remains HIV seronegative. There is increased precedence for focusing research efforts on understanding heterosexual transmission in the context of the HIV-1 subtype C epidemic.

The Zambia Emory HIV Research Project (ZHERP), which was initiated in Lusaka, Zambia in 1994 by Dr. Susan Allen of Emory University, is one of the largest discordant couples cohorts worldwide. Following identification, discordant couples are offered enrollment in a long-term follow up study. Through participation in the study, counseling and condom provision are given to the couple at one-month intervals in addition to monitoring of the HIV seronegative partner's status. Despite this highly effective public health intervention, 8% of the couples transmit and become concordant

positive each year. Samples are collected regularly from both the HIV+ and HIV- partner before and after seroconversion. This allows for early identification of acute infection within weeks of the transmission event as well as subsequent longitudinal follow-up [56]. Consequently, the study of discordant couples from Zambia provides a unique opportunity to explore the contribution of viral characteristics to the biology of HIV-1 transmission as well as understanding how early events following transmission dictate subsequent disease progression and long-term pathogenesis.

Several known factors have been identified that modulate the transmission efficiency of HIV-1 outside of the route of transmission [57,58]. For example, the presence of sexually transmitted diseases or genital ulcers in either partner has been shown to increase transmission perhaps through increasing viral shedding at the genital mucosa, inducing inflammation, or by disruption of the mucosal barrier [59]. In addition, high viral loads in the transmitting partner [60,61] as well the acute stage of infection was associated with increased transmission frequency [62]. Social factors such as the seroprevalence of one's sexual network and lower economic status are associated with increased risk [63]. Finally, several host immunogenetic factors have been identified that potentially modulate transmission, although more research is needed in order to more conclusively demonstrate the direct mechanisms [64,65]. Conversely, several factors have been identified that reduce transmission. Male circumcision has been shown to reduce the risk of transmission by 60% and has since been implemented as a public health prevention measure [66]. Additionally, couples' voluntary counseling and testing (CVCT) [55,56] and antiretroviral treatment of HIV+ partners in serodiscordant couples significantly reduces the rate of HIV-1 transmission [67].

Despite high genetic diversity within the viral quasispecies of a chronically infected individual, a severe genetic bottleneck occurs during heterosexual transmission resulting in the transmission of a single/transmitted founder (T/F) variant [68,69]. This

phenomenon has been demonstrated in several independent studies and has been recapitulated in experimental SIV infection of rhesus macaques [70]. The concept of a genetic bottleneck raises the possibility that selection pressures during transmission play a role in determining the virus that is transmitted. If this is the case, characterization of the genetic and phenotypic properties that distinguish T/F viruses from other viruses present in chronic infection could be used to inform prevention strategies aimed at preventing transmission [71,72]. A study of transmitted vs. non-transmitted viruses present in the blood and genital compartment of epidemiologically linked transmission pairs from Zambia demonstrated that the T/F is often represented as a minor variant within the chronically infected donor quasispecies [73] providing supporting evidence that transmission involves a selective process rather than being entirely stochastic. Several studies have described genetic and phenotypic properties of T/F viruses: preferential CD4/CCR5 co-receptor tropism [52,69], shorter variable loops and fewer N-linked glycosylation sites in the Envelope (Env) glycoprotein [52,68,69,74-77], increased Env neutralization sensitivity to the transmitting partner's antibodies, tier 2-3 neutralization sensitivity (i.e. not particularly neutralization sensitive) [68,71] and greater replication efficiency in CD4+T cells as opposed to replication in macrophages [52,78].

One overall limitation to the majority of the above mentioned studies were the focus on Env properties rather than studying full-length infectious molecular clones. A recent study using subtype B and subtype C T/F infectious molecular clones (IMC) demonstrated that T/F IMCs as compared to chronic viruses exhibited modest increases in infectivity, greater Env incorporation, more efficient interaction with dendritic cells, and a higher resistance to interferon-alpha, which might be important in controlling viral replication during the earliest stages of infection by inducing an anti-viral innate immune response [79]. Despite studying full-length IMCs, this study as well others to

date fail to directly compare T/F viruses to matched non-transmitting viruses derived from the chronically infected donor in a transmission pair. Therefore, it is plausible that the inconsistencies observed within the field are due to the experimental limitations of comparing chronic viruses vs. T/F viruses derived from different individuals, and that a study of viruses derived from epidemiologically linked heterosexual transmission pairs will be ideal for further elucidating the role of viral characteristics in defining the selection bias during transmission.

HIV-1 genome and viral lifecycle

The HIV-1 replication cycle begins when viral particles adsorb to a target cell surface via interactions between the viral glycoprotein (Env) and the cell surface receptor, CD4. In the case of HIV-1, interaction with a co-receptor, which is most often one of the chemokine receptors CCR5 or CXCR4, is essential for membrane fusion and entry. Following entry, viral particles are partially uncoated in the cytoplasm and reverse transcription of their positive sense single stranded RNA genome ensues. The double stranded DNA product in the form of the preintegration complex (PIC) is then subsequently transported into the nucleus. Integration of viral DNA into the host cell's chromosome is facilitated by the virally encoded integrase (IN). The integrated viral DNA (provirus) serves as a template for production of mRNA via the host's DNA-dependent RNA polymerase. Splicing of viral mRNA occurs and the products are transported to the cytoplasm where they are transcribed into viral proteins. The Env protein and the Gag-Pro-Pol polyprotein are trafficked through separate pathways to the surface of the cell where progeny virus begins the budding process with the aid of a myriad of host cellular proteins. Immature viral particles are released from the cell

surface, and proteolytic cleavage of the immature viral particle by the viral protease forms a mature viral particle that is fully infectious [80].

Human Immunodeficiency Virus (HIV) is a member of the *Lentivirus* genus and *Retroviridae* family. HIV is an enveloped virus with a positive sense single stranded RNA genome present in two copies within the virion (pseudodiploid). The HIV-1 genome comprises genes that are important for major structural and enzymatic functions (g*ag, pol,* and *env*), regulatory proteins (*tat* and *rev*) as well as several accessory proteins (*vpu, vpr, nef,* and *vif*) [80]. The main function of the viral Env protein is to bind the cellular receptors/co-receptors and to facilitate membrane fusion reaction between the viral and host membranes [81-83]. In addition, the Env protein is immunogenic and contains epitopes that elicit antibody as well as CTL responses [84].

The Gag and Pol proteins serve critical functions in the viral lifecycle. Pol consists of three viral enzymes: protease (PR), reverse transcriptase (RT), and integrase (IN). These viral enzymes are contained within the domains of the Gag-Pro-Pol polyprotein, which is generated via a ribosomal frame-shifting event near the 3' end of the *gag* open reading frame. Three major folded domains, the matrix (MA), capsid (CA) and nucleocapid (NC) within Gag each mediate essential events during virion assembly. Each domain is separated by flexible linker regions, which contain protease cleavage sites [85]. MA, which is found on the N-terminus, functions by binding to the plasma membrane and recruits the Env glycoprotein [86-89]. CA is the central most protein domain and functions as the main structural protein that assembles to form both the immature and mature cores of the virion [90,91]. NC binds and captures the two copies of the viral RNA genome during assembly, and also plays a role during tRNA primer annealing during reverse transcription [92,93]. Two spacer regions, SP1 and SP2, aid in regulating the conformational changes necessary for the transition from the immature to the mature virion following cleavage [94]. Finally, the carboxy-terminal p6 region has

been shown to be important for late assembly events and binds to ESCRT-I proteins [95,96].

The pathogenesis of HIV-1 infection

Traditional hallmarks of AIDS pathogenesis were recognized early on during the epidemic as involving the progressive depletion of CD4+T cell populations and global immune dysfunction. This in turn leads to immunodeficiency marked by an increased susceptibility to a plethora of opportunistic infections and tumors uncommon in healthy individuals. Unfortunately, a majority of HIV infected individuals succumb to these and other AIDS related co-morbidities [97-100]. Successful ART treatment of HIV-1 infected individuals reduces viremia to undetectable levels, restores CD4+ T cells counts to some degree, and significantly prolongs life [42,43]. Despite this, current therapeutic strategies cannot fully eradicate the virus, and viral rebound occurs upon treatment interruption [101-103]. Research has also shown that even in ART treated individuals with successfully suppressed viral loads, immune function is not always completely restored, and there is an increased morbidity and mortality compared to healthy individuals [104-106].

The last three decades of HIV-1 research have been in part dedicated towards understanding the intricate interaction between the HIV-1 virus and the immune system. Such research has made great strides through the study of the pathogenic infection of humans as well as experimentally infected Rhesus Macaques (RMs) but also through comparative studies with two natural SIV hosts that do not develop AIDS, sooty mangabeys (SMs) and African green monkeys (AGMs). HIV/AIDS research has shown us that the interplay between virus and host is multifaceted and much more complicated than initially thought [107,108]. It is the ultimate hope that a deeper understanding of

HIV-1 pathogenesis will translate into more effective interventions to treat, prevent, improve quality of life, and hopefully end the HIV/AIDS epidemic.

Several important facets of pathogenic HIV-1 infection have been identified including 1) rapid depletion of activated mucosal memory CD4+ T cells during acute infection followed by a slower progressive depletion of CD4+ T cells in the periphery 2) disruption of mucosal immunity and microbial translocation 3) chronic immune activation 4) functional exhaustion of immune responses 4) failure of critical CD4+ T cell homeostasis and regenerative capacity 5) establishment of a latent viral reservoir.

Depletion of CD4 T cells in HIV Infection:

Damage to the immune system occurs very early after transmission when HIV-1 is able to directly infect and deplete CD4+ T cells. However, the model of HIV-1 pathogenesis is much more complicated than viral replication, and direct infection and killing of CD4+ T cells alone [97,109-111]. Natural SIV hosts do not progress to AIDS despite high levels of viral replication [112-114]. Thus, although damage does ensue following direct infection and apoptosis of virally infected CD4+ T cells, the cellular targets of viral infection and the populations that are spared in pathogenic HIV and SIV infection of RM vs. SIV infection of natural hosts sheds light on the immunopathology beyond direct cytopathic effects of the virus [107-109,115].

One of the most important discoveries in understanding HIV-1 immunopathology came with elucidating the cellular targets of HIV-1 viral infection. HIV-1 utilizes CD4 as a primary receptor, however engagement with one the chemokine receptors - CCR5 or CXCR4 - is necessary for Env mediated fusion and cellular entry [116,117]. CCR5 and CXCR4 are expressed on different CD4+ T cell subsets with CXCR4 being highly expressed on a vast majority of CD4+ T cells while CCR5 expression is mostly limited to

memory CD4+ T cells [110,115,118,119]. Additionally, it was demonstrated that HIV-1 preferentially infects activated T cells that have already encountered antigen [120-122]. This includes HIV-specific T cells, as well as other pathogen specific CD4+ memory cells, such as those specific for Mycobacterium tuberculosis [123], but to a lesser extent cytomegalovirus-specific CD4+ T cells [124,125]. The hierarchy of preferential infection of pathogen specific CD4+ T cells partially explains the tempo with which opportunistic infections occur during the course of the disease [126,127]. Moreover, the infection of HIV-specific CD4+ T cells renders the immune system ill-equipped to control HIV replication, because the generation of any antigen specific CD4+ T cell response results in more cellular targets for the virus [128].

The human memory T-cell pool can be heterogeneous, but is often delineated by either the presence or absence of a combination of host of cell surface markers such as CD45RA, CD45RO, CD11alpha, CCR7, CD27, CD28, CD57, and CD62L [129,130]. Memory CD4+ T cells can be generally classified into two distinct groups based on their homing capabilities and effector functions: effector memory (T_{EM}) and central memory (T_{CM}). T_{CM} are characterized by an increased capacity to survive and proliferate after activation and home to secondary lymphoid organs through the expression of the chemokine receptor, CCR7 [129-132]. In contrast, T_{EM} lack expression of CCR7 and reside outside of lymph node organs in peripheral effector sites such as the lamina propria of the intestinal mucosa [115,133,134]. T_{EM} are classified by more rapid production of cytokines and cytolytic enzymes upon antigen stimulation, which is important for an immediate recall response to pathogens encountered at mucosal and peripheral sites - hence their name “effector memory” [129-132].

HIV-1 co-receptor usage reflects the dynamics of CD4+ T cell depletion in HIV-1 disease progression and pathogenesis [115,117]. In experimental infection of RMs with CXCR4 or dual trophic SIV/HIV hybrid virus (SHIV), global CD4+ T cell depletion

occurs, and AIDS quickly ensues, highlighting the fact that massive destruction of CD4+ T cells can result in end-stage disease [110]. However, almost all recently transmitted viruses are dependent on CCR5, while Env utilization of CXCR4 typically evolves very late during HIV-1 subtype B chronic infection, but very rarely evolves during subtype C infection [71,116,135,136]. During the earliest stages of viral infection, the virus encounters large pools of target cells, mainly CD4+ T_{EM}. CD4+T cells are rapidly depleted at mucosal tissues mostly in the gastrointestinal tract where the majority of the initial cellular targets reside [137,138]. This drastic depletion does not result in overt immunodeficiency immediately, but rather results in a more gradual loss of CD4+ T cells in the periphery that occurs throughout the course of infection [108,115,139,140]. Importantly, it has also been shown that not all CD4+ T cell depletion occurs due to direct infection and apoptosis of virally infected cells, but can also occur through other indirect mechanisms [141-143].

Targeting of activated T_{EM} cells through preferential use of CCR5 during early viral replication at mucosal sites is of a great advantage to the virus, because although a large number of target cells are depleted early on during infection, additional T_{EM} target cells can be partially regenerated through the proliferation and differentiation of T_{CM}. This regenerative process does not fully restore immunity and is not stable over time likely owing to many factors such as the direct infection and loss of T_{CM} cells, immune activation, exhaustion, increased proliferation/cellular turn-over, and destruction of critical lymph node architecture [139,140,144,145].

As for the direct infection of T_{CM}, this subset of memory cells were found to be the largest contributor to the viral reservoir, and levels of infected T_{CM} *in vivo* correlate with disease progression. HIV-1 and pathogenic SIV infection of RMs results in highly proliferating T_{CM} that have a high turn over rate thereby constantly maintaining the pool of T_{EM} target cells. This high level of proliferation, activation, and cellular turn over in

combination with apoptosis due to direct and indirect mechanisms results in a progressive failure of T_{CM} cell homeostasis and regenerative potential, eventually leading to the slow decline of CD4+ T cells [139,140,144,146-149].

In contrast, during acute infection of natural SIV hosts, there is an initial depletion of mucosal CD4+ T cells similar to pathogenic infection, but the numbers of mucosal CD4+ T cells soon stabilize or even recover, and importantly they maintain healthy levels of peripheral CD4+ T cells throughout the course of infection [107,112,114]. Natural SIV hosts have circumvented the progressive depletion of peripheral CD4+T cells through several complex mechanisms [107]. One such mechanism is the preferential sparing of the T_{CM} CD4+T cell subset, which is critical for homeostasis, by exclusively infecting T_{EM} cells, which are more dispensable. *Ex vivo* experiments in which sorted memory populations from both SMs and RMs were experimentally infected with SIV demonstrated an intrinsic resistance to infection in SM T_{CM} cells, which was shown to be linked to CCR5 expression. CD4+ T_{CM} from SMs express lower levels of CCR5 than T_{EM} from SMs or even T_{CM} from RM. This dampened expression of CCR5 was perhaps even more striking after *in vitro* activation of cells [150,151]. Thus, it appears from both human as well as non-human primate studies, that cellular tropism, co-receptor usage, and preferential infection of certain CD4+ T cell subsets plays an integral role in the immunopathology of HIV/SIV.

Immune Activation in HIV Infection:

Chronic immune activation is a hallmark of progressive HIV infection and is a strong predictor of CD4+ T cell decline independent of viral load [152-155]. Chronic immune activation cannot be solely attributed to continuous viral replication, but is a complex phenomenon that plays a major role in many of the aspects of HIV

immunopathology [155]. In ART treated individuals with suppressed viral loads, immune activation persists and predicts morbidity and mortality. Additionally, reduced immune activation in ART treated individuals predicts CD4+ T cell gains better than suppression of viral load [156,157]. Additional evidence for the pathogenic role of immune activation comes from comparative studies of pathogenic SIV infection of RMs vs. nonpathogenic natural SIV hosts such as AGMs and SMs [107]. Despite high viremia, SIV infection of SMs and AGMs is associated with low levels of chronic immune activation [158,159].

One of the first immunological abnormalities described in HIV infection in 1983 was poly-clonal B-cell activation [160]. Immune activation in HIV infection is characterized by a myriad of other immunological abnormalities including high T cell turnover of both CD4+ and CD8+ T cells that express activation markers [161,162], as well as high levels of circulating pro-inflammatory cytokines and chemokines [163,164]. Innate cells such as dendritic cells (DCs), macrophages (MO), and natural killer (NK) cells also show signs of immune activation and dysfunction. Mucosal DC subsets are lost with an overabundance of pDCs predominating in chronic HIV-1 infection [165-167]. MOs become dysfunctional and display a reduced capability for phagocytizing bacteria [168] and decreased NK cell cytotoxicity has also been observed [169].

T-cell activation is traditionally characterized by the cell surface expression of HLA-DR and CD38, which was first recognized as important in HIV pathogenesis in the late 1980s [170]. HLA-DR is constitutively expressed on antigen presenting cells (APCs) and is integral for the presentation of antigen to CD4+ T cells. Although not highly expressed on all T-cells, it is unregulated on certain subsets of activated cells during an immune response [130,171]. In contrast, CD38 is constitutively expressed on naïve T-cells, down regulated on resting memory T-cells, but is upregulated during the earliest stages of activation. CD38 is a transmembrane glycoprotein and its expression has been

associated with increased cell-cell adhesion and increased cytokine production [172]. It is therefore most common to analyze these markers in combination; however, the distribution of CD38 on T-cell subsets in HIV infection by itself has also proved informative [173,174]. Additionally, markers of proliferation can also be measured as surrogates for activation. Assessment of the intracellular protein Ki-67 provides a measure of cellular proliferation and essentially reflects the rate of T-cell destruction and replacement (turnover). Ki-67 traditionally correlates with activation markers such as CD38 and HLA-DR, but provides additional information on proliferation and cellular turnover, which cannot be elucidated through activation markers alone [130,175].

Although HIV immune activation has been implicated in HIV pathogenesis since the beginning of the epidemic [160,170], the causes of it have not been fully elucidated. A seminal study by Deeks et al. showed that immune activation was apparent even during acute infection. More importantly, a steady-state level of activation termed the "immune activation set-point" was reached during early HIV infection, varied among different individuals, and was predictive of longer-term CD4+ T cell decline independent of viral load [152]. This particular study demonstrated that early events during the virus-host interaction play a role in defining the trajectory of immune activation, immunopathology and ultimately disease progression.

The causes and consequences of chronic immune activation in HIV-1 infection are not completely understood however significant advances have elucidated several key features. Some of the potential mechanisms of chronic immune activation that have been proposed include: 1) direct effects of viral gene products and viral replication 2) innate and adaptive immune response to the virus 3) dysfunction of the regulation system that dampens immune responses 4) bystander activation and the production of pro-inflammatory cytokines and 5) microbial translocation [176]. Chronic immune activation is thought to disrupt CD4+ T cell homeostasis, causes overall immune dysfunction [176],

and more recently has been associated with non-AIDS related comorbidities such as HIV-associated neurocognitive disorder, adverse cardiovascular events, liver damage, and pre-mature aging [177-179].

One possible factor that can contribute to immune activation is direct viral replication and the innate recognition of viral products [164,166]. HIV might induce immune activation through GP160/120 binding to CD4 and/or CCR5 on pDCs causing downstream signaling, IFN production, and induction of apoptosis in CD4+ T cells [180]. In addition, the accessory protein Nef, has been implicated in influencing levels of activation. A wide variety of Nef functions have been identified including down-modulation of CD4 and HLA-I molecules, disruption of CD3+ TCR signaling, induction of apoptosis, as well as transcriptional factor translocation and increased expression of IL-2 [181,182].

Additionally, innate recognition of HIV-associated pathogen associated molecular patterns (PAMPs) has been implicated in the production of pro-inflammatory cytokines and subsequent activation of adaptive immunity [183]. Various pattern recognition sensors for HIV have been identified: Toll-like receptors (TLR7/8, and TLR9) that can recognize viral RNA and DNA CpG motifs within endosomes [184,185]; RIG-I homologues act as cytosolic sensors of viral RNA species [186]; and more recently, cytosolic DNA sensors that induce IFN in a TLR-independent manner have been identified [187]. Moreover, host restriction factors such as TRIM5 alpha, which recognizes three-dimensional capsid lattice structures, appear to function like PAMPS and can induce innate signaling pathways [188-191].

Another potential mechanism involved in immune activation is bystander activation and cell death of B and T-lymphocytes caused by the increased production of pro-inflammatory cytokines, and potentially by the up-regulation of apoptosis related molecules [192,193]. Interestingly, a recent study by Doitsh et al. [194] demonstrated

that the majority of CD4+ T cells lost in HIV-1 infection undergo caspase-1 mediated pyroptosis triggered by abortive viral infection. Pyroptosis is a form of programmed cell death in which high levels of inflammatory cytokines such as IL-1β are released [195]. Still yet another factor potentially involved in HIV-associated immune activation is the role of regulatory T-cells (Tregs), which can dampen chronic immune activation, but on the other hand could also inhibit necessary immune responses that target viral replication. Owing to these factors, the role of Tregs in inducing immune activation remains controversial in the field [196].

One common theme that seems to be tightly associated with the level of chronic immune activation is microbial translocation. Microbial translocation during HIV infection was first described by Brenchely et al. in 2006 when they demonstrated that microbial products such as bacterial lipopolysaccharide (LPS) are elevated in circulating plasma in HIV-1 infected individuals as well as in the pathogenic SIV infection of RMs as compared to uninfected healthy controls. Levels of LPS also correlated with the extent of activation on both adaptive and innate immune cells [197]. Microbial translocation is not limited to pathogenic HIV/SIV infection, but has also been described in inflammatory bowel disease, hepatitis C infection, and cardiovascular disease [198]. Microbial translocation is described as passage of the naturally occurring gastrointestinal microflora through the intestinal epithelial barrier and lamina propria. Translocated gut microflora then travel to local mesenteric lymph nodes and eventually circulate throughout the body and can be found systemically Normally, translocated microbial products are phagocytosed within the lamina propria, and never make it to circulation; however, when host mucosal immunity is compromised, this host defense fails, and allows egress of the microbial products [199]. Translocated bacteria as well as other viral and parasite components stimulate immune cells through TLR-recognition pathways on antigen presenting cells and contribute to increased circulating pro-inflammatory

cytokines such as IL-6, IL-1, TNF-alpha, D-dimer, and C-reactive protein [200-202]. D-dimer is involved in the coagulation cascade, while C-reactive protein is an acute phase liver enzyme. Levels of D-dimer and CRP have been implicated in the link between HIV infection and risk of adverse cardiovascular incidents [203]. In addition to measuring plasma LPS, levels of soluble CD14, which is the cell surface receptor for LPS on monocytes/macrophages and is indicative of monocyte activation, can also be utilized to indirectly measure the extent of microbial translocation [202,204]. More recently, plasma levels of intestinal fatty acid binding protein (iFABP), which is a marker of enterocyte damage, has been utilized [205]. In support of this role of microbial translocation in HIV pathogenesis, the levels of many of the above mentioned markers of microbial translocation have been associated with disease progression both in treated as well as untreated HIV infection [201].

Perhaps even more intriguing, a key feature that distinguishes pathogenic SIV infection of RM and HIV infection of humans with nonpathogenic SIV infection of SM and AGMs is the lack of microbial translocation with low levels of immune activation observed in the latter [107,114,197,206,207]. Moreover, increased immune activation and viral replication can be observed following experimental administration of LPS to natural hosts of SIV [208]. While it is generally accepted that translocation of bioreactive microbial products from the lumen of the gut into systemic circulation induces immune activation, it remains unclear if immune activation is the initial driver in causing the mucosal immune dysfunction that leads to microbial translocation. It is plausible that microbial translocation is just a secondary component in a feed-forward loop that amplifies immune activation [176,201].

Although the exact causes of microbial translocation in HIV infection are still being fully investigated, microbial translocation seems to be driven by destruction and impairment of mucosal immunity that leads to a breach in the mucosal barrier during

early infection [197,198]. The early depletion of mucosal and gut CD4+CCR5+ T cells is thought to be an important driver of microbial translocation; however, mucosal depletion of CD4+ T cells without microbial translocation occurs in natural SIV infection of SMs and AGMs [112,114], thus it is not simply depletion of gastrointestinal CD4+ T cells that influences gut integrity. CD4+ memory T cell populations are not all functionally equivalent, and several distinct subsets have been identified and can be categorized based on their function and cytokine profile. Some of the main subsets identified to date include Th1, Th2, Th17, and T-follicular helper subsets [196].

Dysregulation of Th17 cells, in particular, has been found to contribute to microbial translocation [206,209,210]. This subset is important for the maintenance of mucosal immunity, structural integrity, and is crucial for defense against extracellular pathogens and bacteria in the gastrointestinal tract. Th17 cells are delineated from other subsets by their secretion of the cytokines IL-17 and IL-22. These cytokines promote neutrophil recruitment, induce production of antibacterial defensins, and stimulate tissue repair by inducing the survival and proliferation of epithelial cells [211]. Th17 cells are preferentially depleted during pathogenic HIV and SIV infection of RMs resulting in a skewing away from Th17 to Th1 type memory subsets at gastrointestinal sites [209]. The extent of depletion of this subset correlates with levels of systemic immune activation and predicts the rate of disease progression [207]. Although the natural SIV hosts SMs and AGMs exhibit a depletion of mucosal memory CD4+ T cells early in infection, they maintain normal levels of Th17 cells, retaining gut integrity and therefore lack signs of microbial translocation [206]. It is thought that this unique characteristic of non-pathogenic SIV infection can explain the apparent lack of chronic immune activation in these hosts [107].

Recent evidence suggests that the loss of IL-21 producing CD4+ T cells is in turn correlated with the preferential depletion of Th17 CD4+ T cells [212]. IL-21 has several

important immune functions. IL-21 1) stimulates the maintenance of long term CD8+ T cells with cytotoxic functionality [213] 2) favors differentiation to and expansion of Th17 cells [214] 3) helps in the differentiation of B-cells towards memory and antibody-secreting plasma cells [215] and 4) regulates NK cell expansion and functionality [216]. Levels of IL-21 are decreased in pathogenic HIV/SIV infection but not in natural SIV infection of SMs or AGMs [212]. Experimental administration of IL-21 to SIV infected RMs results in reconstitution of Th17 cells without a negative impact on viral load and is associated with decreased immune activation and microbial translocation [217]. This highlights the important role of IL-21 producing CD4+ T cells in influencing Th17 cells and microbial translocation/immune activation, but it also demonstrates the precedence for pursuing therapeutics designed to restore IL-21 and Th17 mediated immunity in HIV infection.

Immune Exhaustion in HIV Pathogenesis:

During acute viral infections, viral replication is accompanied by the induction of antigen specific CD8+ T cells that function to control and eliminate viremia. These antigen-specific cells subsequently develop into differentiated memory cells. Chronic viral infection ensues when virally infected cells persist in the host over long periods of time either due to the inability of the host's immune response to clear the infection, to the presence of latently infected cells, or to a combination of both of these factors [218,219]. Under conditions of persistent antigen exposure, progressive loss of important CD8+ T cell functions, such as the ability to proliferate, secrete cytokines, and loss of cytotoxic activity, occurs leading to the exhaustion and anergy of these cells [219,220]. The functional exhaustion of virus specific T-cells has been extensively characterized in the murine mouse model of chronic lymphocytic choriomeningitis virus

(LCMV) infection [221], but has also more recently been described for a number of other chronic human infections such as hepatitis C virus [222], hepatitis B virus [223], and HIV [223-226] .

Despite the induction of a CD8+ T cell response during HIV infection, the majority of individuals are ultimately unable to control viral replication and progress to AIDS [227]. Numerous studies have described functional impairments of the CD8+ T cell response in HIV infection [220]. HIV infection is associated with a defect in the maturational status of CD8+ T cells as evidenced by a skewing of CD8+ T cells towards a more effector phenotype rather than effector memory [228]. Secondly, significant differences in the ability of CD8+ T cells to secrete cytokines and differences in cytotoxicity were observed when comparing CD8+ T cells from elite controllers and non-progressors to individuals that progress [229-231]. Furthermore, polyfunctional CD8+ T cells, meaning those that mediate a combination of functions (CD107a, IL-2, TNF-alpha, interferon, or MIP-1b), have been associated with more successful control of viremia and improved disease outcome [232]. The suppression of immune functions has been attributed to the upregulation of inhibitory molecules in the setting of chronic antigen stimulation [219]. Studies in murine infection of LCMV first described the role of the PD-1 (programmed death -1) /PDL1 pathway in dampening the functionality of virus specific CD8+ T cells during chronic infection [221]. PD-1 fits into the complex and finely tuned immune system by serving to regulate and turn off the immune response. PD-1 is a member of the B7:CD28 family members of co-stimulatory molecules that inhibit T-cell activation. PD-1 is upregulated after T-cell activation and functions to decrease TCR signaling, thus serves as an inhibitory feedback mechanism to prevent excessive T-cell activation [233,234]. An extensive number of studies have shown that the PD-1/PDL1 pathway leads to the cellular immune dysfunction and exhaustion of HIV-specific CD8+ T cells in pathogenic HIV/SIV infection [224,226,235,236]. Additionally, it has been

shown that blockade of PD-1 in RMs leads to the restoration of CD8+ T cell immune function and decreases viral load [237]. PD-1 has also been implicated in cell death. Studies have shown that PD-1 expression on virus-specific CTLs makes them more susceptible to apoptosis. Interestingly, this study also found that HIV-1 specific cells that expressed CD57 in the absence of PD-1 were more resistant to apoptosis and displayed higher cytotoxicity [238]. Traditionally CD57 has been utilized as a marker to indicate replicative senescence and terminal differentiation [130], however, accumulating evidence suggests that CD57 expression in HIV infection is not necessarily associated with replicative senescence as it is in aging and CMV infection [239,240]. Elucidating the underlying mechanisms of T-cell dysfunction in HIV infection and disease has important applications for future vaccine design as well as in therapeutic interventions to eradicate the latent viral reservoir.

The proviral burden in HIV infection

Retroviruses including HIV-1 are characterized by their ability to first reverse-transcribe their +ssRNA genome into DNA that is subsequently integrated into the host cell DNA. The integrated viral DNA genome (provirus) serves both as a template for viral RNA synthesis, but also is maintained and inherited as part of the host cell DNA [80]. HIV establishes a state of latency within particular cell subsets through integration of viral DNA into the host chromatin and the subsequent silencing of active viral transcription. This effectively hides the viral genome in specific cellular reservoirs that are unable to be recognized and cleared by host immune responses [241-243]. The HIV reservoir is established early during primary infection, an occurrence that poses unique obstacles for treating and curing HIV infected individuals [244]. Despite the high potency of ART to suppress viral replication, ART does not fully eradicate viral reservoirs

[103]. Although initiation of treatment early can reduce the size of the viral reservoir, a stable population of latently infected CD4+ T cells persists and is unaffected by ART treatment [245]. In addition, several practical considerations should be given to the strategy of treating the HIV-1 epidemic with ART alone. First, strict adherence to ART for the lifetime of the infected individual is absolutely necessary for the successful suppression of viral load. Additionally, despite virus control, ART regimens can themselves be associated with unwanted adverse effects. Therefore, the question of whether we will be able to eradicate HIV with ART becomes multifaceted and gives precedent towards developing novel therapeutics that target the latent viral reservoir and residual immune activation.

Most HIV proviral DNA can be found in CD4+ T cells that reside in lymphoid tissue. Within blood, HIV DNA is mostly found within central memory and transitional memory CD4+T cells, which likely maintain the viral reservoir, under fully suppressive ART treatment, due to their intrinsic ability to persist through homeostatic proliferation [147]. More recently, a subset of memory T cells with stem-cell like properties (TSCM) have been identified as potentially contributing to a long-lived viral reservoir [246]. In non-pathogenic SIV infection of SMs and AGMs, CD4+ T_{CM} [150] as well as TSCM [246] subsets are preferentially spared, presumably because of their reduced levels of CCR5. In pathogenic HIV infection, a higher viral burden in these compartments is observed, which in turn leads to a disruption in critical T-cell homeostasis. A bountiful supply of viral targets, namely CD4+ T_{EM} cells, cannot be regenerated from CD4+ T_{CM} precursors, which leads to the overt decline in total CD4+ T cells over time.

The cellular immune response to HIV-1 and CTL escape

Cytotoxic T lymphocytes (CTLs) are important for the control of many different types of viral infections. Once a virus infects a cell, a fraction of its cytosolic viral proteins are proteolytically cleaved, translocated into the ER, displayed on major histocompatibility complex class I (MHC-I) molecules, and then trafficked to the cell surface. The T-cell receptor (TCR) on a CD8+ T cell that is able to bind its cognate viral antigen (displayed on MHC-I) with high affinity will signal the release of cytolytic molecules that can lyse the infected target cell [218]. The MHC-I locus within humans is one the most polymorphic loci, possibly as a result of co-evolution with pathogenic organisms. The majority of the polymorphisms within a class I molecule result in differential binding to epitopes. Antigen presentation (to T cells) is therefore said to be restricted by MHC-I molecules, which are also referred to as HLA-I (human leukocyte antigen) alleles [247].

The role of cytotoxic T lymphocytes (CTLs) in suppressing HIV replication has been well documented. Reduction of high levels of viremia during the acute phase of infection is temporally associated with the development of HIV-specific CTLs [248,249]. However, the importance of CTLs in the control of HIV-1 replication is most directly supported through studies of SIV infected rhesus macaques that have been depleted of CD8+ T cells. CD8+ T cell depletion in these animals resulted in the loss of control of acute viremia and the animals progressed to AIDS-like symptoms [250]. It is now recognized that the control of acute viremia down to what is termed “set point viral load” is in part due to variation in HLA-I alleles. Several HLA-I alleles have been statistically associated with lower viral loads and/or delayed disease progression in humans including HLA-B*5701/03, B*5801, B*27, B*51, B*81, and A*74 among others [251-258]. Individuals with a broad Gag-biased CTL response have better clinical outcomes

[259,260]. Polyfunctional CD8+ T cells have also been associated with more favorable disease outcomes [113,232,261].

More recently, virus specific CD4+ T cells have been shown to be important for the suppression of HIV replication [262,263]. Soghoian et al. demonstrated that the expansion of HIV-specific CD4+ T cells that expressed the cytolytic molecule, granzyme A, during acute infection contributed to the control of viremia. Moreover, the HLA-class II allele, HLA-DRB1, was found to be associated with protective CD4+ T cell responses and reduced viral loads [264]. Interestingly, similar to what has been demonstrated for CD8+ T cell responses, Gag-specific CD4+T cell responses are more protective than Env-specific responses [265,266]

It has become increasingly apparent that infection with HIV cannot be completely eliminated with ART treatment or by naturally occurring immune responses alone, making the identification of an efficacious HIV-1 vaccine a high priority. A protective vaccine will likely need to include a combination of both humoral and cellular-mediated immune responses. Vaccine elicited humoral immune responses such as the development of broadly neutralizing antibodies could induce sterilizing immunity in which infection is completely blocked. Such strategies are, however, challenging due to the high levels of viral heterogeneity especially within the Env glycoprotein. On the other hand, vaccine elicited cellular immune responses, although unlikely to induce sterilizing immunity, could potentially be efficacious if they suppress early viral replication within the first few days after mucosal exposure, but prior to dissemination [267]. In support of this hypothesis, cellular based SIV vaccines that utilize rhesus CMV as a vector are able to elicit persistent and high frequency SIV-specific effector-memory CD8+ T cell responses at sites of SIV replication. These responses were associated with substantial control of viremia to undetectable levels in around half of the vaccinated animals. This

supports the notion that induction of a protective cellular immune able to control viral replication might be feasible [268].

Despite the initial reduction of acute viremia to set point by CTLs and the association of certain protective HLA-I alleles with long-term non-progression, the majority of infected individuals eventually progress to AIDS due to immune failure [227]. As HIV infection ensues, persistent levels of antigen leads to immune exhaustion and HIV-specific CD8+T cells become ineffective [224,226,235,236]. Adding to this complexity, HIV-1 adapts to immune pressure during the course of an infection in order to increase its fitness *in vivo*. Mutation of critical amino acid residues either directly within or adjacent to CTL epitopes can decrease antigen presentation on the cell surface, decrease recognition by the cognate TCR, or interfere with optimal antigen processing. The selection pressure that CTLs place on the virus is substancial; mutations in critical residues within the viral genome predominate in the face of associated replicative fitness costs, because the overall benefit to the virus is greater than remaining vulnerable to CTL recognition [269-273].

Clear examples of CTL escape in HIV-1 infection were first described in 1997 [269,273,274]. Subsequent studies in the SIV macaque model confirmed CTL escape and further demonstrated that escape can occur frequently [275]. The appearance of CTL escape mutations *in vivo* is commonly associated with loss of viral control and subsequent disease progression [274]. More recently, the evolution of CTL escape mutations has been demonstrated in the humanized BLT-mouse model of HIV infection indicating that these mice generate robust CD8+ T cell responses [276]. This opens up the utility of this animal model for vaccine studies. While CTL escape during chronic infection has been historically well documented, recent reports demonstrate that CTL escape can occur as early as one to two weeks post-infection. CTL escape is thought to dictate the timing and kinetics of the establishment of set point viral load during acute

infection [277]. The potency of the CTL pressure exerted on the virus can be directly related to the appearance of CTL-associated escape mutations. For instance, certain CTLs are unable to induce escape mutations, potentially because they exert minimal selective pressure; however, this is confounded by epitopes that fail to escape due to fitness constraints [278]. Investigating the kinetics and timing of CTL escape during natural infection will be integral for identifying CTL epitopes that would afford the best protection and would not escape with ease within the context of an HIV vaccine.

Just as antigen presentation is restricted by a host's HLA-I alleles, so is epitope escape. Therefore, patterns of CTL escape can be predicted based on a host's HLA-I allele [48]. Indeed, Moore et al. [279] first described the appearance of HLA footprints at the population level, suggesting that CTLs shape HIV-1 evolution, and can be utilized to identify novel CTL escape mutations. In order to correct for viral phylogeny, linkage disequilibrium, and covarying amino acids, a novel computational method initially utilized by Bhattacharya et al. [280] and further modified by Matthews [281] and Carlson [282] have identified HLA-I associated polymorphisms within both subtype B and C cohorts. The influence of HLA-I alleles on viral evolution at the population level was demonstrated in a global analysis of viral sequences from nine different cohorts. The prevalence of an HLA-I allele within the population was consistently correlated to the frequency of a viral polymorphism within a linked CTL epitope, even in individuals lacking that allele; this indicates that certain CTL escape mutations can become fixed and persist in a population if the restricting HLA-I allele is sufficiently prevalent within the population and the polymorphism does not rapidly revert [283].

Viral Replication Capacity of HIV-1

HIV-1 is able to escape from CTL mediated immune pressure; however, this may come at a cost to the virus if escape falls within a functionally constrained region of the genome, such as Gag [269,271-273,278,284]. Evidence suggests that part of the benefit observed from protective HLA-B*57, B*5801, and B*27 alleles is due to the CTL targeting of conserved regions in Gag [281,285-292]. Additionally, the number of CTL responses directed against fitness-constrained epitopes have been correlated to a lower viral load, highlighting the importance of fitness associated CTL escape mutations in defining disease progression [259,260]. A series of HLA-associated polymorphisms, mainly within p24 Gag and that are selected by protective alleles, have been extensively studied and found to reduce *in vitro* replication [281,285,286,288-291,293-297]. Studies have shown that some HLA-I associated polymorphisms will revert back to the consensus residue in the absence of the restricting HLA allele [272,281,298-300]. This also suggests that these mutations confer a fitness cost, and thus revert in order to maximize replicative fitness within the context of an HLA-disparate host. In support of this, HLA-B*57 and B*5801 associated polymorphisms within p24 revert *in vivo* and are statistically associated with a lower set-point plasma viral load in individuals that lack the restrictive HLA-I allele [290].

Certain CTL escape mutations that are associated with a fitness cost are able to persist in a population, and do not revert or revert slowly upon transmission to an HLA-mismatched individual. The mechanism of such persistence may be attributed to several complex factors including the evolution of compensatory mutations that restore fitness or the high frequency of the selecting HLA-I allele within the population [283,287,292]. In some instances, secondary mutations may play a role in compensating for a decrease in viral fitness by stabilizing protein structure, restoring critical protein-protein

interactions, or through facilitating further escape [286,291,292,294,297,301]. The mechanism of compensation might be direct, as is the case for two residues involved in a salt-bridge formation, or the compensation might act more indirectly through changing the global conformation of the protein [302].

To date, compensatory mutations have been described in detail for escape from the B*57/5*801 restricted TW10 epitope, as well as for the B*27 restricted KK10 epitope. These compensatory mutations have been statistically correlated to a higher viral load highlighting the clinical importance for understanding the pathways available for compensation within the context of a structurally constrained protein such as Gag [291,292,294,298,301]. Clearly, in response to CTL pressure, compensation occurs to restore fitness, however, the larger role that compensation plays in defining the evolution of viral fitness within a population has not been fully explored.

Despite significant progress in defining the *in vitro* fitness effects of individual CTL escape mutations and their associated compensatory mutations, naturally occurring HIV-1 strains contain unique and complex mutational patterns. These complex mutational patterns might represent multiple overlapping footprints of HLA-associated polymorphisms resulting from the HLA-mediated immune pressure of different immunogenetic backgrounds in the history of the virus. Alternatively, other selection pressures may drive viral sequence diversity within Gag or other viral proteins, and it is plausible to envision that such mutations also having functional consequences. Recent studies have demonstrated sequence polymorphisms that circumvent recognition by NK cells. KIR-associated amino acid polymorphisms were described for various HIV-1 proteins including Gag [303]. For example, sequence variation within HLA-C*03:34-presented HIV-1 epitopes in Gag was found to alter binding of KIR2DL2 binding to this HLA-allele [304]. Restriction factors such as TRIM (tripartite motif) 5 alpha [191,305-307] and Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like (APOBEC3)

might also contribute towards the accumulation of sequence diversity [308]. Moving forward, it is of great importance to understand how different combinations of mutations within Gag as well as other proteins work in concert to define the global *in vitro* fitness of viruses, and how this fitness in turn defines HIV-1 pathogenesis and disease progression.

One of the first indications that HIV-1 replicative capacity might impact clinical outcome came from the observation that the replicative capacity of *gag-pol* chimeric NL4-3 viruses derived from elite controllers was significantly lower in comparison with the viruses isolated from chronic progressors [289,309]. Subsequently, Brockman et al. demonstrated a link between the replicative capacity of *gag-pro* sequences isolated during chronic stage infection and viral load in both subtype C and B infections [310-313]. Although these studies demonstrated a clear relationship between replicative capacity and viral load during chronic infection, the experimental approach utilized in these studies was not adequate for studying viruses isolated during acute infection. The authors postulated that attenuation of viral replicative capacity through HLA-mediated immune escape drives the observed association with viral load. Alternatively, it is plausible to envision that polymorphisms within the virus that get transmitted to a newly infected individual could affect replicative capacity *in vitro,* and by extension, viral load and disease progression. It is therefore important to understand how early viral replication of the transmitted virus interplays with critical viral-host interactions to define HIV-1 pathogenesis and disease progression.

Through the study of HIV-1 subtype C infected heterosexual transmission pairs from Zambia, we are uniquely poised to investigate the how transmitted viral characteristics work in concert with host factors to define HIV-1 pathogenesis. While it has long been known that host factors such as gender and protective HLA-I alleles influence set point viral load, current estimates predict that together they only account

for ~ 22% of the variability [314]. Influential studies in heterosexual transmission pairs alluded to the possible role for viral characteristics in HIV-1 disease. In a series of studies, it was found that viral loads were correlated between chronically infected donors and their acutely infected linked recipients [315-319]. These findings suggest that viral loads are heritable outside of the influence of host factors, and perhaps this heritability can be attributed to a viral characteristic that is transmitted from donor to recipient. Yue et al. demonstrated that the correlation between donor and recipient viral load becomes stronger when host factors are taken into account in multivariable models, further highlighting the integral role that host factors and viral characteristics likely play in defining HIV-1 pathogenesis [319]. A key analysis by Goepfert et al. demonstrated that an accumulation of HLA-I associated polymorphisms within the transmitted Gag sequences derived from 88 acutely infected Zambians was associated with a lower point VL. This correlation was only observed for HLA-I associated polymorphisms within Gag, but not in the viral accessory protein, Nef [320]. The HIV-1 Gag polyprotein is composed of three major folded peptides, the capsid (CA), matrix (MA), and nucleocapsid (NC). These proteins serve to assemble the virus into distinct immature and, upon proteolytic cleavage, mature conical capsids. Multiple weak interactions between Gag molecules are necessary for the assembly of the immature and mature structures as well as for the associated conformational changes [85]. It is, therefore, plausible to envision HLA-I associated polymorphisms that, when present individually, may not adversely affect replicative capacity as measured by current *in vitro* assays, but in combination, could reduce viral fitness and limit the pathways that the virus has for reverting and compensating for this loss in fitness.

Taken together, this suggests that viral replicative capacity, specifically defined by Gag, might play a previously unrecognized role in defining the trajectory HIV-1 disease pathogenesis and might ultimately explain the variation observed in disease

severity between different individuals. Specifically, we hypothesize that the viral replication capacity of the transmitted Gag sequence will influence viral load as well as CD4+ T cell decline. It is also plausible to hypothesize that early viral replication might set in motion pathological events that cause irreversible damage to the immune system before the induction of an effective adaptive immune response.

Summary

In chapter 2, we investigated the role that replicative capacity of the transmitted Gag sequence plays in defining early HIV-1 infection. In order to accomplish this, we designed a novel cloning strategy that utilized restriction enzymes in order to clone the *gag* gene from acute timepoints from 150 recently infected linked recipients from Zambia into MJ4, a clade C proviral backbone derived from a primary isolate. *In vitro* replicative capacity was then assessed on a CEM-based T cell line that expresses CCR5, and viral particle production quantified via a radiolabeled reverse transcriptase assay. We demonstrate that viral replicative capacity (vRC) correlates with both donor viral load and set point viral load in newly infected individuals. These data highlights the role that transmitted viral characteristics play in influencing viral load during chronic infection, and then upon transmission, the set point viral load in acute infection. Moreover, we were able to elucidate HLA-associated polymorphisms that are statistically associated with changes in vRC, information that can be helpful for informed vaccine design. Finally, we go on to show that vRC predicts CD4+ T cell decline for a subset of individuals for whom longitudinal CD4+ T counts were available for three years post infection.

In chapter 3, we more fully explore the relationship between vRC of the transmitted Gag and CD4+ T cell decline. Expanding the dataset to include a total of 127 individuals with vRC and longitudinal CD4+ T cell counts out to five years allowed us to more definitively demonstrate the impact of vRC on disease progression. We demonstrate that vRC correlates with CD4+ T decline, even down to an endpoint of 200, which is the clinical definition of AIDS. Furthermore, we demonstrate that the impact of vRC on CD4+ T decline is independent of both set point viral load as well as host factors known to modulate set point viral load such as the protective alleles, HLA-B*57/B*5801. We further elucidated a mechanism by which vRC might be impacting CD4+ T decline early during infection before host factors enact to control acute viremia down to set point. We find that vRC was associated with levels of pro-inflammatory cytokines, cellular immune activation and exhaustion. Furthermore, we demonstrate that vRC is associated with increased cellular proliferation, and with the level of viral burden in critical CD4+ T cell compartments.

Chapter 2

Role of transmitted Gag CTL polymorphisms in defining replicative capacity and early HIV-1 pathogenesis

Published in *PLoS pathogens*. 2012;8(11)

Jessica L. Prince and Daniel Claiborne performed all RC experiments and data analysis for figures 1, 2, 4, 5, 6, 7, S1 and S2

Tianwei Yu performed the Cox-proportional hazard models (Table 2)

Jonathan Carlson and David Heckerman performed the statistical analysis in order to elucidate HLA-associated polymorphisms and residues associated with changes in fitness: Supplemental Table 1 and Figure 3

Jessica L. Prince, Daniel Claiborne and Eric Hunter wrote the manuscript

ABSTRACT

Initial studies of 88 transmission pairs in the Zambia Emory HIV Research Project cohort demonstrated that the number of transmitted HLA-B associated polymorphisms in Gag, but not Nef, was negatively correlated to set point viral load (VL) in the newly infected partners. These results suggested that accumulation of CTL escape mutations in Gag might attenuate viral replication and provide a clinical benefit during early stages of infection. Using a novel approach, we have cloned *gag* sequences isolated from the earliest seroconversion plasma sample from the acutely infected recipient of 149 epidemiologically linked Zambian transmission pairs into a primary isolate, subtype C proviral vector, MJ4. We determined the replicative capacity (RC) of these Gag-MJ4 chimeras by infecting the GXR25 cell line and quantifying virion production in supernatants via a radiolabeled reverse transcriptase assay. We observed a statistically significant positive correlation between RC conferred by the transmitted Gag sequence and set point VL in newly infected individuals ($p = 0.02$). Furthermore, the RC of Gag-MJ4 chimeras also correlated with the VL of chronically infected donors near the estimated date of infection ($p = 0.01$), demonstrating that virus replication contributes to VL in both acute and chronic infection. These studies also allowed for the elucidation of novel sites in Gag associated with changes in RC, where rare mutations had the greatest effect on fitness. Although we observed both advantageous and deleterious rare mutations, the latter could point to vulnerable targets in the HIV-1 genome. Importantly, RC correlated significantly ($p = 0.029$) with the rate of CD4+ T cell decline over the first 3 years of infection in a manner that is partially independent of VL, suggesting that the replication capacity of HIV-1 during the earliest stages of infection is a determinant of pathogenesis beyond what might be expected based on set point VL alone.

AUTHOR SUMMARY

In the majority of HIV-1 cases, a single virus establishes infection. However, mutations in the viral genome accumulate over time in order to avoid recognition by the host immune response. Certain mutations in the main structural protein, Gag, driven by cytotoxic T lymphocytes are detrimental to viral replication, and we showed previously that, upon transmission, viruses with higher numbers of escape mutations in Gag were associated with lower early set point viral loads. We hypothesized that this could be attributed to attenuation of the transmitted virus. Here, we have cloned the *gag* gene from 149 newly infected individuals from linked transmission pairs into a clade C proviral vector and determined the replicative capacity *in vitro*. We found that the replicative capacity conferred by the transmitted Gag correlated with set point viral loads in newly infected individuals, as well as with the viral load of the transmitting partner, and we identified previously unrecognized residues associated with increasing and decreasing replicative capacity. Importantly, we demonstrate that transmitted viruses with high replicative capacity cause more rapid CD4+ decline over the first three years, independent of viral load. This suggests that the trajectory of pathogenesis may be affected very early in infection, before adaptive immunity can respond.

INTRODUCTION

Despite a diverse HIV-1 quasispecies within chronically infected individuals, a single variant establishes infection in the majority of heterosexual transmission cases, resulting in a severe genetic bottleneck [1-3]. A more profound understanding of the interaction between host and viral characteristics and how they shape early pathogenesis and disease progression will be integral for understanding the trajectory and impact of early events after heterosexual transmission. While it is well established that host factors such as HLA-class I alleles can play a major role in determining clinical progression in those individuals recently infected with HIV-1 [4-10] the role of transmitted viral characteristics has been understudied due to the lack of suitable cohorts in which virus from both the donor and linked recipient are available. Accordingly, studies using epidemiologically linked heterosexual transmission pairs are essential for understanding how viral evolution, adaptation, and the characteristics of the transmitted variant influence HIV-1 pathogenesis.

Previous studies in both heterosexual and homosexual transmission pairs have demonstrated that viral loads (VL) in the newly infected partners are correlated with VL in the transmitting partner [11-13]. This finding is intriguing as the majority of the linked couples have disparate HLA-I alleles, and would therefore have varying immune responses to the incoming virus. Thus, the characteristics of the virus in the donor quasispecies that impact replication can similarly impact the replication of the newly infecting virus even in the context of what is frequently a distinct immunogenetic environment. Indeed, when host factors known to modulate VL such as the presence of protective or unfavorable HLA-I alleles, gender, age, and HLA-B sharing are taken into account in a multivariable analysis utilizing a general linearized model, the correlation

between donor and recipient VL becomes highly significant (Yue *et al.*, manuscript in submission).

It is clear that both humoral [14] and cellular immune responses can drive virus evolution over the course of infection. In the case of the latter, escape mutations arise that abrogate the ability of cytotoxic T lymphocytes (CTL) to kill virus-infected cells [15-22]. While the selection of CTL escape mutations provides an *in vivo* advantage for the virus, if a mutation occurs within a functionally constrained region of the genome such as Gag, it could reduce *in vitro* replicative fitness [17,23-30]. This phenomenon has been demonstrated for several CTL escape mutations associated with protective alleles such as HLA-B*57, B*5801, B*27, and B*81 [31-36]. The ability of protective alleles to target conserved regions of the genome that escape with difficulty, due to the fitness costs incurred by mutations at these epitopes, may partially explain the mechanism of enhanced protection from disease progression in individuals with these alleles [37-41]. While evasion from the CTL response may result in such deleterious mutations, the *in vivo* fitness benefit outweighs that of the replication cost [42], and the ongoing selection of additional mutations may allow the virus to compensate for these defects [17,29,32,33,35,43-49]. The functional consequence of escape and compensatory mutations upon transmission to an HLA-mismatched individual has not been fully explored.

Initial studies in the ZEHRP cohort of 88 transmission pairs demonstrated that the number of transmitted HLA-B associated polymorphisms in Gag, but not Nef, was negatively correlated to set point VL in the newly infected partners, suggesting that an accumulation of escape mutations might attenuate viral replication and provide a clinical benefit during early stages of infection [50]. In addition, in a smaller study of nine newly infected individuals infected by viruses with fitness reducing HLA-B*5703 associated epitope-escape mutations in p24, a lower set point VL was observed [24]. Previous

studies have also shown that elite controllers can harbor viruses with *gag-pro* sequences that confer reduced *in vitro* replicative capacities (RC) [51-53]. In a series of studies, Brockman and colleagues have shown that *in vitro* RC conferred by *gag-pro* variants isolated from both subtype B and C chronically infected individuals correlates to VL, demonstrating the role of intrinsic viral characteristics in defining this marker of pathogenesis [43,47,54,55]. However, in these studies the recombination of population-based PCR amplified sequences into the lab adapted NL4.3 provirus required the outgrowth of virus in a CEM-CCR5 based cell line, potentially skewing the nature of the virus recovered. In contrast, studies of HIV-1 fitness in acute infection did not yield a statistically significant correlation between RC and VL, potentially due to small sample sizes and the limitations of the methodologies used.

The identification of 149 heterosexual epidemiologically linked transmission pairs from a discordant couple cohort in Lusaka, Zambia, provides a unique opportunity to investigate the role that HLA-mediated adaptation of Gag within a chronically infected individual plays in modulating the RC of the transmitted variant. We hypothesize that HLA-mediated adaptation of HIV-1 resulting in Gag variants conferring varying levels of RCs will be a major viral characteristic linking donor and recipient VLs, and that the *in vitro* RC conferred by the transmitted Gag sequence defines early clinical parameters of HIV-1 pathogenesis.

To test this hypothesis and using a novel approach, we cloned *gag* sequences from the earliest seroconversion plasma sample from 149 newly infected recipients of epidemiologically linked Zambian transmission pairs into the MJ4 proviral backbone [56]. The RC of each Gag-MJ4 chimera was then determined and used to investigate how the RC conferred by the transmitted Gag sequence defines clinical parameters, such as early set point VL and CD4+ T cell decline in the newly infected individuals. These studies allowed us to identify novel residues in Gag that influence RC, and demonstrate a

strong correlation between RC and early set point VL, as well as between RC and CD4 decline during the first three years of infection, which was also found to be independent of VL. Thus, the RC of the transmitted virus as defined by its *gag* gene influences critical aspects of HIV-1 pathogenesis.

MATERIALS AND METHODS

Study subjects

All participants in the Zambia Emory HIV Research Project (ZEHRP) discordant couples cohort in Lusaka, Zambia were enrolled in human subjects protocols approved by both the University of Zambia Research Ethics Committee and the Emory University Institutional Review Board. Prior to enrollment, individuals received counseling and signed a written informed consent form agreeing to participate. The subjects selected from the cohort were initially HIV-1 serodiscordant partners in cohabiting heterosexual couples with subsequent intracouple (epidemiologically linked) HIV-1 transmission [57-59]. Epidemiological linkage was defined by phylogenetic analyses of HIV-1 *gp41* sequences from both partners [60]. Viral isolates from each partner in the transmission pair were closely related, with median and maximum nucleotide substitution rates of 1.5 and 4.0%, respectively. In contrast, median nucleotide substitution rate for unlinked HIV-1 C viruses from the Zambian cohort and elsewhere was 8.8% [60]. The algorithm used to determine the estimated date of infection (EDI) was previously described by Haaland *et al.* [2]. All patients in this cohort were antiretroviral therapy naïve. Zambian linked recipients were identified 45.5 days (median, IQR = 41.5 – 50.5) after the estimated date of infection, at which time plasma samples were obtained from both the transmitting partner (donor) and the seroconvertor (recipient). The vast majority (95%) of HIV-1 sequences derived from ZEHRP transmission pairs belonged to HIV-1 subtype C with subtypes A, D, G, and J being detected only occasionally [60]. All of the transmission pairs utilized in this study are infected with subtype C HIV-1.

Viral loads and CD4+ count measurements

Early set point VL for newly infected individuals was defined as the earliest stable nadir VL value measured between 3 and 9 months post infection and which did not show a significant increase in value within a 3-4 month window. HIV plasma VL was determined at the Emory Center for AIDS Research Virology Core Laboratory using the Amplicor HIV-1 Monitor Test (version 1.5; Roche). CD4+ T cell counts were based on T-cell immunophenotyping, with assays done using the FACScount System (Beckman Coulter Ltd., London, United Kingdom) in collaboration with the International AIDS Vaccine Initiative.

HLA-class I genotyping

Genomic DNA was extracted from whole blood or buffy coats (QIAamp blood kit; Qiagen). HLA class I genotyping relied on a combination of PCR-based techniques, involving sequence-specific primers (Invitrogen) and sequence-specific oligonucleotide probes (Innogenetics), as described previously [10]. Ambiguities were resolved by direct sequencing of three exons in each gene, using kits (Abbott Molecular, Inc.) designed for capillary electrophoresis and the ABI 3130*xl* DNA Analyzer (Applied Biosystems).

Amplification and sequencing of gag from donors and linked recipients

Viral RNA was extracted from 140 μL plasma samples using the Qiagen viral RNA extraction kit (Qiagen). *Gag-pol* population sequences were generated using nested gene specific primers. Combined RT-PCR and first round synthesis was performed using SuperScript III Platinum One Step RT-PCR (Invitrogen) and 5 μL viral RNA template.

RT-PCR and first round primers include GOF (forward) 5' ATTTGACTAGCGGAGGCTAGAA 3' and VifOR (RT-PCR and reverse) 5' TTCTACGGAGACTCCATGACCC 3'. Second round PCR was performed using Expand High Fidelity Enzyme (Roche) and 1 μL of the first round PCR product. Nested second round primers include GIF (forward) 5' TTTGACTAGCGGAGGCTAGAAGGA 3' and VifIR (reverse) 5' TCCTCTAATGGGATGTGTACTTCTGAAC 3'. Three positive amplicons per individual were pooled and purified via the Qiagen PCR purification kit (Qiagen). Purified products were sequenced by the University of Alabama at Birmingham DNA Sequencing Core. Sequence chromatograms were analyzed using Sequencher 5.0 (Gene Codes Corp.), and degenerate bases were denoted using the International Union of Pure and Applied Chemistry codes when minor peaks exceeded at least thirty percent of the major peak height.

The percent similarity between donor and recipient population *gag* sequences was determined by building a neighbor-joining tree using Geneious v5.5.7 (Biomatters Ltd.). The percent similarity between nucleotide and amino acid alignment was determined based on the output matrix from these neighbor-joining trees. In calculating the percent similarity between amino acid sequences, degenerate bases that resulted in non-synonymous changes and, thus, a mixture of amino acid residues, were translated as an "X". When one of the amino acids comprising a mixture in the donor was found in the recipient Gag sequence, this was counted as a mismatch, making the average percent similarity reported between donor and recipient *gag* sequences a maximal estimate of percent mismatch.

Generation of Gag-MJ4 chimeras

Viral RNA was extracted from linked recipients at the day of seroconversion time point using the Qiagen viral RNA extraction kit (Qiagen). First round PCR products were generated as was previously described for the *gag* sequencing of all 149 transmission pairs [50]. Second round PCR was performed to generate *gag* amplicons for Gag-MJ4 chimera generation using Phusion Hot Start II polymerase (Fisher) and 1 μL of the first round PCR product. Nested second round primers include GagInnerF1 (forward) 5' AGGCTAGAAGGAGAGAGATG 3' and BclIDegRev2 (reverse) 5' AGTATTTGATCATAYTGYYTYACTTTR 3', which generate a *gag* amplicon starting from the initiation codon of *gag* and extending 142 nucleotides after the *gag* stop codon and into *pro*. The 5' portion of the MJ4 long terminal repeat (LTR) was amplified using Phusion Hot Start II polymerase and the MJ4For1b (forward) 5' CGAAATCGGCAAAATCCC 3' and MJ4Rev (reverse) 5' CCCATCTCTCTCCTTCTAGC 3' primer set. In order to make the proper insert for cloning, the patient-specific *gag* and MJ4-LTR sequences were joined using a splice-overlap extension PCR utilizing the MJ4For1b (forward) and BclIRev (reverse) 5' TCTATAAGTATTTGATCATACTGTCTT 3' primer set. Joined splice-overlap-extension PCR products were gel purified using the Qiagen gel extraction kit (Qiagen). Purified Gag-LTR inserts and wild-type MJ4 vector (NIH AIDS Research and Reference Reagent Program, [56]) were digested with NgoMIV and BclI restriction enzymes (NEB) and ligated overnight at 4° C with T4 DNA ligase (Roche) at a 3:1 insert to vector ratio. JM109 competent cells were transformed with ligation products, plated onto LB/agar plates supplemented with 100 μg/ml ampicillin and grown at 30° C. Gag-MJ4 chimeric DNA was isolated from cultures using the Qiagen miniprep kit (Qiagen). Gag-MJ4 chimeras were sequenced to confirm *gag* insert fidelity as compared to previously amplified population sequences. Two identical independent

clones per linked recipient were chosen for replication assays in order to ensure backbone fidelity during the cloning process.

Generation and titration of viral stocks

Viral stocks were generated by transfecting 1.5 μg purified proviral plasmid DNA into 293T cells (American Type Culture Collection) using the Fugene HD transfection reagent (Roche) according the manufacturer's protocol. Viral stocks were collected 72 hrs post transfection, clarified by low-speed centrifugation, and frozen at -80° C. The titer of each viral stock was determined by infecting TZM-bl cells (NIH AIDS Research and Reference Reagent Program) with 5-fold serial dilutions of virus in a manner previously described [36,61].

Both 293T and TZM-bl cell lines were maintained in complete Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (HyClone Laboratories), 2mM L-glutamine, and 100 U/ml penicillin G sodium, and 100 μg/ml streptomycin sulfate (Gibco) at 37°C and 5% CO_2.

In vitro replication capacity (RC) assay

In order to assess the RC of Gag-MJ4 chimeras, 5 x 10^5 GXR25 cells [62] were infected at an MOI of 0.05, and 100 μl of viral supernatants were collected at 2 day intervals. Briefly, GXR25 cells and virus were incubated with 5 μg/ml polybrene at 37°C for 3 hours, washed 5 times with complete Roswell Park Memorial Institute (RPMI) medium (Gibco) and plated into 24-well plates. Cells were split 1:2 to maintain confluency, replaced with an equal volume of fresh media, and viral supernatants were taken at days 2, 4, 6, and 8 as previously described [36,61]. Virion production was

quantified using a ^{33}P-labeled reverse transcriptase assay. Based on values obtained for days 2-8, the optimal window for logarithmic growth for all viruses was determined to be between days 2 and 6, as by day 8 many high replicating viruses had exhausted target cells causing a flattening or decline of the replication curve. Therefore, log_{10}-transformed slopes were calculated based on days 2, 4, and 6 for all viruses. Replication scores were generated by dividing the log_{10}-transformed slope of the replication curve for each Gag-MJ4 chimera by the log_{10}-transformed slope of wild-type MJ4. Two independent Gag-MJ4 chimera clones per acutely infected linked recipient were run in duplicate to confirm cloning fidelity. After both independent clones were confirmed to have identical replicative capacities, one clone was subsequently run in triplicate in two independent experiments in order to generate consistent replicative capacity scores. GXR25 cells were maintained in complete RPMI medium supplemented with 10% fetal bovine serum (HyClone Laboratories), 100 U/ml penicillin G sodium, 100 μg/ml streptomycin sulfate (Gibco), and 10mM HEPES buffer at 37°C and 5% CO_2.

Quantification of HIV-1 reverse transcriptase

Aliquots of culture supernatants from infected cells were added to an RT-PCR master mix [63] and incubated at 37°C for 2 hours; then the RT-PCR product was blotted onto DE-81 paper, and allowed to dry. Blots were washed 5 times with 1X SSC (0.15 M NaCl, 0.015 M sodium citrate, pH 7.0) and 3 times with 90% ethanol, allowed to dry, and exposed to a phosphoscreen (Perkin Elmer) overnight. Counts were read using a Cyclone PhosphorImager [36,61].

HIV-1 Gag polymorphism scores

HLA-associated polymorphisms were defined as any non-consensus polymorphism that occurred at an amino acid position having known escape mutations adapted to specific HLA-class I alleles as defined by a list of associations generated in a manner similar to that described previously [44,64] from 1899 subtype C *gag* sequences from Zambia and South Africa (Carlson, *et al.*, manuscript in preparation). Polymorphisms that increased or decreased replicative capacity were defined based on amino acid associations with RC derived from an exploratory pair-wise analysis detailed in the experimental results. To generate a summed polymorphism score, the number of HLA-associated fitness-decreasing polymorphisms was subtracted from the number of HLA-associated fitness-increasing polymorphisms.

Statistical Analysis

The relationships between RC and set point VL, donor VL, and average CD4+ counts; RC and the number and quality of HLA-associated polymorphisms; and VL and the number and quality of HLA-associated polymorphisms were analyzed using the Spearman rank correlation. Linear regression analyses were utilized to generate trend lines to facilitate visualization of correlation graphs. Mann-Whitney tests were used to compare the differences in RC between rare and more common polymorphisms. Mann-Whitney tests were used to analyze the difference in median set point VLs between different RC groups (RC<1, RC=1-2, RC>2). All Spearman correlations, Mann-Whitney tests, and linear regression analyses were performed using Prism GraphPad v5.0 (GraphPad Software, Inc.). The Mann-Whitney U test was used to identify statistically significant differences in RC between two groups (e.g. sequences with or without a given

HIV polymorphism). Multiple tests were addressed using q-values [65], which estimates the expected proportion of significant tests that are false positives. To limit the number of tests, we considered only groups containing at least 3 individuals.

A subset of volunteers with longitudinal CD4+ T cell counts (n=63) was analyzed to characterize the relationship between RC and T cell decline. Survival analysis was used to estimate the association between replicative capacity (RC) and the drop of CD4+ cell count. The endpoint is defined as the time before CD4+ T cell counts drop below a threshold, e.g. 300 or 350 cells/mm^3. The set point VL is another factor being considered. Kaplan Meier curve and log-rank test were used to compare the survival between the groups with RC<1 and RC>2. Cox proportional hazard regression was used to assess risk associated with high RC with or without adjustment for the confounding factor of set point VL. The sample size (n=66) was inadequate for a more complete analysis with additional covariates.

RESULTS

Selection and characterization of transmitted gag sequences

Studying heterosexual transmission of HIV-1 within the context of discordant couples allows for full characterization of viruses from both donor and recipient. From a total of 294 epidemiologically linked transmission pairs identified from 1998-2010 in the Zambia Emory HIV Research Project (ZEHRP) discordant couple cohort [66], we selected 149 individuals that had been enrolled for at least 1 year and had at least 9 months of follow up post serconversion. The median time post estimated date of infection (EDI) for all 149 individuals was 45 days. Thus, this represents a unique early infection linked transmission pair cohort.

The median $\log_{10}$ VL for donors near the time of transmission and the median $\log_{10}$ set point VL for linked recipients was 5.02 (IQR = 4.51 – 5.45) and 4.39 (IQR = 3.91 – 4.99) respectively for all participants included in this study. Figure S1 depicts the phylogenetic clustering of population sequences for the *gag* gene of the 149 epidemiologically linked transmission pairs and highlights the high degree of sequence similarity between donor and recipient viruses.

The majority of population *gag* sequences isolated from acute/early time points of linked recipients were homogeneous and in most cases were identical to the donor population *gag* sequence isolated near the EDI. Overall, donor and linked recipient sequences differed in amino acid composition by only 2.7%. Mutations were counted when a mixture of nucleotides (amino acids) in the donor population sequence resolved to a single residue in the recipient and thus represent maximal values. Of the 149 pairs only 37 recipient sequences had evidence of potential early escape and a majority of these individuals (28/37) had only a single amino acid change. Therefore, we can

conclude that the majority of the sequence polymorphisms present at the seroconversion time point are derived from the chronically infected donor. Additional characteristics of the cohort including set point VL and CD4+ counts are listed in Table 1.

Construction of Gag-MJ4 chimeric viruses

Previous studies investigating the role of Gag viral fitness have employed a recombination approach in which sequences are PCR amplified as a bulk population, allowed to recombine into a *gag-deleted* NL4-3, and resulting viruses propagated in permissive cells [43,47,51,53-55]. This method has three distinct disadvantages: there is no control over the sites of recombination, it requires the outgrowth of virus which may select for the most fit virus in the population and could also select for sequence changes, and the introduction of subtype C sequences into a lab adapted subtype B proviral backbone may introduce biases due to the interaction of subtype B proteins with subtype C Gag. In order to avoid these limitations and because we are studying individuals recently infected with HIV, where the population is generally homogeneous, we have employed a direct cloning method that allows for the introduction of the entire *gag* gene into a replication competent, CCR5 tropic, clade C provirus, MJ4 [56].

A splice-overlap-extension PCR was employed to fuse the MJ4-LTR-U5 sequence with the transmitted *gag* sequence. This ensures that the cis-acting sequences upstream of Gag, which may influence expression levels, are constant throughout all constructs. The resulting chimeras include the entire transmitted *gag* sequence from the initiation codon to the end of Gag and extend into conserved region of protease by 47 amino acids. For each newly infected individual, at least two independent Gag-MJ4 chimeras were sequence confirmed and assayed for replicative capacity (RC). An analysis of variation of RC between the two independent clones derived from each newly infected linked

recipient was 8.5%. Testing two independent clones, therefore, ensures that the observed RC is not due to the confounding effect of backbone mutations that might have arisen during the cloning process and provides an estimate of experimentally induced variation.

Overall, a low amount of heterogeneity was detected in *gag* population sequences isolated from linked recipients, with only 28% having one or more mixed bases resulting in amino acid changes in the sequences from which the Gag-MJ4 chimeras were generated. When this was the case, the *gag* clone with the sequence closest to the donor *gag* sequence was chosen in order to avoid sampling *gags* containing *de novo* escape or reversion. In some cases in which multiple variants appeared to be transmitted, several *gag* variants were assayed for RC as described in the materials and methods section. In each case, these minor variants were found to have similar or identical RC values (data not shown).

Introduction of gag sequences from newly infected individuals into MJ4 drastically alters replicative capacity

In initial replication assays performed in order to test assay precision, wild-type MJ4 exhibited an intra-assay variability of 10.4% and an inter-assay variability of 8.7%. Figure 1 shows the results of a typical experiment for all 149 Gag-MJ4 chimeras, with wild-type MJ4 depicted in red. The normalized RC values of the chimeras ranged from less than 0.01 to greater than 3.5. Some viruses replicated more than 100 times more efficiently than MJ4, demonstrating that substitution of Gag can have a profound impact on the ability of the virus to replicate in cells.

Correlation between the replicative capacities conferred by transmitted *gag* sequences and viral loads in newly infected individuals and their transmitting partners

An examination of the RC of transmitted viruses allows us to determine the role of viral replication in defining set point VL in acutely infected individuals before significant viral adaptation to immune pressure of the host has taken place, which might confound the relationship of RC to VL. We observed a statistically significant positive correlation between the replicative capacities of Gag-MJ4 chimeras and set point VLs in newly infected individuals (Figure 2A; Spearman correlation $r = 0.17$, $p = 0.02$), a correlation that persists when conditioning on the presence of B*57 in, and the sex of, the linked recipient ($p = 0.009$). This finding indicates that the RC conferred by the transmitted Gag sequence clearly plays a role in defining early set point VL of newly infected Zambian linked recipients.

In several cohorts VL in the transmitting partner and that in the linked seroconvertor have been shown to be correlated [11-13]. In order to more fully explore the possible contribution of RC in explaining this phenomenon, we compared VLs of the transmitting partner at the time of transmission to the RC defined by the transmitted Gag sequence. Despite both a higher maximum and wider range of VLs within transmitting partners, we observed a statistically significant positive correlation between RC and the set point VL of the donors, similar to that of their newly infected partners (Figure 2B; Spearman correlation $r = 0.18$, $p = 0.01$). This supports the concept that RC, defined by Gag, is a viral characteristic contributing to the positive correlation between donor and recipient VLs that has been previously reported [11-13].

Several amino acids in Gag significantly correlate to changes in replicative capacity. Uncovering sites of vulnerability in HIV-1 is a high priority for the informed design of an effective HIV vaccine [21]. Therefore, we examined all 149 Gag sequences and their RC using an exploratory pairwise analysis described previously [43,54], to uncover residues that significantly affect the virus' ability to replicate *in vitro*. We found 49 residues at 31 unique positions that had a statistically significant effect on RC at $p <$

0.05 ($q < 0.51$) and 4 residues at 3 unique positions that were significant at $p<0.002$ ($q<0.2$) (Table S1). In what follows, we will use $q<0.2$ as the significance threshold when individual sites of significance are considered, and $q<0.51$ ($p<0.05$) as the significance threshold when we are testing broad trends, in which we are willing to increase our expected false positive rate as a tradeoff to substantially reduce our expected false negative rate.

The locations of all statistically significant polymorphisms ($p < 0.05$), along with their effects on RC as compared to the median RC of all viruses, are plotted linearly on a graphical representation of the Gag protein (Figure 3A.). Residues that dramatically modulate RC were enriched in p17 and p2 (Fisher's exact test, $p<0.0001$). In addition, roughly two-thirds of the non-consensus residues with $p<0.05$ increase fitness relative to the median RC for the entire population.

An expanded data set of 1899 subtype C *gag* sequences from Zambia and South Africa (Carlson, Schaefer et. al., manuscript in preparation) was utilized to identify residues that affected RC and were also HLA-associated. Within this data set, HLA-associated polymorphisms are classified as being either adapted or non-adapted. An adapted residue is one that is escaped relative to the HLA-allele in question. In contrast, a residue that is non-adapted is the susceptible form and may render the virus vulnerable to immunological targeting by the HLA-allele in question. This new dataset has identified a total of 199 HLA-linked polymorphisms ($q<0.2$, $p<0.0007$) vs. 59 associations utilized previously from a smaller subset of *gag* sequences [27,50]. Within the 49 residues associated with changes in RC, 7 polymorphisms were found to be adapted to specific HLA class I alleles, clearly demonstrating the impact of the cellular immune response in affecting viral fitness (Figure 3B, * denotes $q< 0.2$). Six polymorphisms were found to be non-adapted to specific HLA class I alleles (Figure 3C, * denotes $q < 0.2$). Since these are non-consensus polymorphisms, it is possible that

consensus at these residues is escaped relative to these HLAs, potentially explaining why an adapted consensus residue at this position is the less fit variant. Indeed, 5 consensus residues (62K, 451S, 488S, 85L, and 309A) with $p<0.05$ were found to be adapted to HLA-I alleles, demonstrating that the cellular immune response can drive selection for consensus residues.

Rare polymorphisms have the greatest effect on replicative capacity

During our analysis of amino acid polymorphisms linked to changes in RC ($p < 0.05$), we observed a negative correlation between the frequency of polymorphisms and the magnitude of their effect on RC (Spearman correlation, $r = -0.89$, $p < 0.0001$). Indeed, rare polymorphisms, those occurring in less than 10 of the 149 individuals studied, had significantly greater impact (both negative and positive) on RC than polymorphisms that occurred more frequently (Figure 4A and 4B). This finding is especially intriguing in the case of rare deleterious mutations, as these residues may highlight epitopes at which HIV escapes or compensates for fitness defects with great difficulty, similar to those described for elite controllers [52], and may, therefore, be attractive targets for a cellular-based vaccine.

The cumulative and qualitative effect of HLA associated polymorphisms in Gag on replicative capacity

In order to investigate whether there is a cumulative effect of viral escape from cellular immune pressure in Gag on RC, the expanded dataset of HLA-associated polymorphisms generated from an analysis of 1899 *gag* sequences from Zambia and South Africa (Carlson, Schaefer *et al.*, manuscript in preparation), described above, was

employed. The number of non-consensus polymorphisms located at HLA-associated positions was determined for each MJ4 chimera and then correlated with the RC defined by those Gag sequences. Surprisingly, we found a positive association between the number of HLA-associated polymorphisms and RC (Fig 5A; $r = 0.14$, $p = 0.05$). Although counterintuitive, this is consistent with the fact that not all HLA-associated polymorphisms within a particular Gag sequence will necessarily reduce fitness. We have shown in the previous sections that several non-adapted (or "susceptible" to HLA pressure) HLA-associated polymorphisms increase fitness relative to the median of all sequences. Indeed, we observe a highly statistically significant positive correlation between the number of non-adapted HLA-associated polymorphisms and RC (Spearman correlation, $r = 0.23$, $p = 0.003$; data not shown). Thus, the inclusion of both adapted (or escaped with respect to specific HLA alleles) and non-adapted polymorphisms within this expanded HLA-associated dataset may explain the observed positive association between numbers of HLA-associated polymorphisms and RC. Therefore, we hypothesize that it is the balance and interaction of both fitness increasing and fitness decreasing polymorphisms within a particular sequence that ultimately determines the RC of the virus.

In order to more accurately determine how the number and quality of HLA-associated polymorphisms affects RC and to correct for the opposing influence of both increasing and decreasing polymorphisms within a particular sequence, a summed polymorphism score was calculated by assigning fitness increasing polymorphisms a score of +1, fitness decreasing polymorphisms a score of -1, and neutral polymorphisms a score of 0. HLA-associated polymorphisms were defined as being positive, negative, or neutral based on the previously described univariate analysis that correlated specific residues within our 149 sequences with changes in RC. In this modified analysis, we observed a highly statistically significant correlation between the summed

polymorphism score and RC (Figure 5B: Spearman rank correlation; r = 0.6, p = <0.0001), confirming that the sequence features are approximately independent of each other and suggesting that the offsetting influence of fitness decreasing and increasing polymorphisms is a strong contributor to RC. This finding may explain the observation that, in general, the most-fit viruses are less like the consensus subtype C Gag sequence, consistent with a majority of polymorphisms increasing fitness (Figure S2, [55]).

The cumulative and qualitative effect of HLA-associated polymorphisms in Gag on set point viral load

In a previous report using 88 Zambian linked seroconverters, we reported that increasing numbers of transmitted HLA-B associated polymorphisms within or adjacent to well defined epitopes were associated with lower set point VLs [50]. When we expand this analysis to include all 149 Zambian linked recipients and use the same dataset of HLA-linked polymorphisms used by Goepfert *et al.* [50] we observe the same correlation (r = -0.15, p = 0.03, Figure 6A). However, when we use the expanded HLA-associated data set (199 associations) to define HLA-associated polymorphisms, we no longer observe a statistically significant negative association between the number of transmitted HLA-associated polymorphisms in Gag and set point VL (Figure 6B).

We therefore hypothesized that, as with RC, this correlation between the total number of transmitted HLA-associated polymorphisms in Gag and set point VL in newly infected individuals may be confounded by not taking into account whether polymorphisms increase or decrease fitness. Indeed, using the summed polymorphism score, we observe a highly significant correlation between the summed score of HLA-associated polymorphisms and set point VL (Figure 6C: r = 0.21, p = 0.006). This demonstrates that it is not merely the quantity of HLA-associated polymorphisms

present in the transmitted Gag sequence that ultimately defines set point VL, but it is the influence of both fitness increasing and decreasing polymorphisms that contributes to RC and in turn set point VL in newly infected individuals.

Transmission of viruses with low replicative capacities provides recipients with a longer-term clinical benefit

Though set point VL has been shown to be a relevant marker for disease progression [67,68], CD4+ T cell counts are traditionally used to define those individuals that have progressed to AIDS and are at a higher risk for opportunistic infections [69,70]. Therefore, we analyzed a subset of individuals (n=66) for whom longitudinal CD4+ T cells counts for at least one-year post-infection are available. We observed a statistically significant correlation between the average CD4+ T cell counts and the replicative capacities of Gag-MJ4 chimeras (Figure 7A, Spearman correlation, $r = -0.24$, $p = 0.02$). This demonstrates that infection with attenuated viruses may impart some survival benefit to newly infected individuals, at least within the first year of infection. Subsequently, we determined whether individuals infected with poorly replicating viruses exhibit differential pathogenesis over the first three years of infection or whether this early benefit is transient and quickly lost. To answer this question, we studied a subset of the linked recipients (n=63) for whom CD4+ T cell counts were available at regular three month intervals for greater than one year post-infection.

In a Kaplan-Meier survival analysis, in which we defined the endpoint as having a CD4+ T cell count >350 cells/mm^3 (WHO recommendation for initiation of anti-retroviral therapy [71]), we observed a statistically significant difference in the number of individuals that maintain CD4+ counts >350 cells/mm^3 between those infected with viruses that replicate very poorly (RC<1) and those infected with highly replicating

viruses (RC>2), within the first 3 years of infection (Figure 7B, Mantel-Cox test p = 0.029). This disparity in disease progression was even more pronounced when the endpoint was defined as having CD4+ T cell counts >300 cells/mm^3, demonstrating a median difference of 896 days before falling below the CD4+ count cut off between individuals infected with low and high replicating viruses (Figure 7C, Mantel-Cox test p = 0.014). Using a Cox proportional hazard model, we demonstrate a significantly increased risk of CD4+ T cell counts falling below 350 (Hazard Ratio (HR) 2.36; p = 0.034) or 300 (HR 3.80; p = 0.021) over the first three years of infection for individuals whose Gag conferred an RC>2 vs. RC<1.

Interestingly, the benefit conferred by low replicating viruses could not be wholly explained by differences in set point VL within this smaller data set. Although there was a trend towards higher VLs between the two most disparate groups, with a 2.5 fold difference in median VLs (Figure 7D), we observed no statistically significant differences in median set point VLs between individuals infected with low (RC<1), medium (RC= 1-2), and highly (RC>2) replicating viruses. Further, in Cox proportional hazard models that take into account VL, the HR remained high (2.18 and 3.12 respectively) and p values continued to trend or remain borderline significant (p = 0.093 and 0.051) (Table 2), indicating that both VL and RC can independently affect CD4 loss. Moreover, the HR associated with log10 increases in set point VL alone was lower than that for RC alone (HR = 1.75 versus 2.62, and 2.09 versus 3.80; CD4 <350 and 300 respectively; Table 2). These results suggest that infection with a low replicating virus confers clinical benefit outside of the effect of RC on set point VL, and that the kinetics of viral replication early in infection can ultimately dictate long-term pathogenesis.

DISCUSSION

In this study of 149 linked Zambian heterosexual transmission pairs from the ZEHRP discordant couple cohort we have more fully characterized the role that HIV-1 viral characteristics, in particular the replicative capacity (RC) conferred by the transmitted *gag* sequence, plays in defining parameters of early HIV-1 pathogenesis. Identification and comparison of both donor and recipient *gag* sequences for all 149 transmission pairs revealed a high degree of similarity (97.6%) within each linked pair, allowing us to conclude that the majority of polymorphisms in Gag present at the seroconversion time point were transmitted from the chronically infected donor.

Transmitted gag sequences from newly infected individuals engineered into MJ4 drastically alters replicative capacity

Since all of the transmission pairs in this study were infected with subtype C viruses, our approach of precisely cloning *gag* genes from acutely-infected recipients into a primary isolate (MJ4) provirus has many important advantages over previously employed methods. MJ4 is a CCR5 tropic infectious molecular clone derived from a subtype C clinical isolate from Botswana, providing greater homology to viruses circulating within the Zambian population than other previously used subtype B lab-adapted strains [33,47,54,55]. Additionally, this cloning method for generating Gag-MJ4 chimeric viruses does not rely on recombination based technologies that require the outgrowth of viral quasispecies, which may select for the most fit virus, and in some cases, amino acid changes in the viral stocks that are not present in the individual from which they were derived [32,47,53-55]. The use of a common BclI restriction site located 137 nucleotides after the *gag* stop codon in MJ4 does result in a chimeric protease,

however, this region is 96.5% conserved in this cohort and we did not observe a high prevalence of dead or inactive Gag-MJ4 chimeras.

The impact of engineering foreign *gag* sequences into MJ4 on virus replication was highly significant, with many of the chimeras exhibiting RC values greater than a hundred-fold higher than wild-type MJ4, which in this assay is one of the poorer replicators. This indicates that substitution of Gag can drastically alter the *in vitro* RC of the virus when all other viral components are constant. Multiple intra-molecular contacts as well as host protein interactions in Gag are necessary for effective intracellular Gag trafficking [72,73], particle formation [74], budding [75,76], maturation [77] and disassembly [78]. Therefore, immune mediated adaptation of this functionally constrained protein could have clear consequences for viral replication through disruption of these many interactions.

Correlation between the replicative capacities conferred by transmitted *gag* sequences and viral loads in newly infected individuals and their transmitting partners

It has been well established that the set point VL in those recently infected with HIV-1 is correlated to disease progression and clinical outcome [67,68]. Previous data from our group demonstrated that transmission of sequences with increasing numbers of CTL escape mutations in Gag resulted in lower set point VLs in newly infected individuals, a finding that suggested that transmitted HLA-associated polymorphisms in Gag might negatively affect viral replication [50]. We have confirmed this association in the current study after increasing the number of transmission pairs analyzed from 88 to 149. This result is consistent with studies by Brockman *et al.*, which have demonstrated a statistically significant link between the RC conferred by *gag-pro* genes in subtype B and C chronically infected individuals to VL [43,47,54,55]. However, a statistically significant correlation between Gag RC derived from acutely infected individuals and set point VL has not previously been definitively reported in a subtype C cohort.

In contrast, in this large group of very early ZEHRP seroconvertors (with samples drawn a median of 45 days post-EDI) we observed a clear statistically significant correlation ($p=0.02$) between the RC conferred by the transmitted *gag* sequence and the early set point VL in newly infected individuals. This result implies that RC plays a role in defining the overall level of virus replication during the first year of infection. Moreover, in multivariable analyses that take into account the early viral control imposed by the B*57 allele and by gender, the impact of RC on set point VL was found to be independent of these two host factors ($p=0.009$). Other factors such as NK cells and restriction factors such as TRIM or APOBEC may potentially affect RC and VL, however little is currently known regarding these potential effects, and future efforts should evaluate the role of such factors. While we observed a statistically significant positive correlation between RC and set point VL, outliers in the data exist that do not fit the trend, and in some cases can be explained by the presence of protective HLA-alleles or by a large number of escape mutations present in the transmitted sequence that are relevant to the HLA background of the newly infected individual. Set point VL is clearly determined by a combination of both host factors, including HLA-alleles, and viral factors such as RC, and this may explain the differences in the absolute correlation for each individual.
The RC of Gag-MJ4 chimeras also correlated with VLs near the estimated date of infection in chronically infected donors, consistent with the previously reported observation that donor and recipient VLs are correlated within epidemiologically linked transmission pairs [11-13]. The data presented here would suggest that the RC conferred by the transmitted Gag sequence is a contributing viral characteristic of that donor virus responsible for influencing early set point VL in the newly infected partner.

Several amino acids in Gag significantly correlate to changes in replicative capacity

In a pair-wise analysis, a large number of residues were associated with changes in fitness ($p<0.05$, $q<0.51$), with 4 residues at 3 unique positions at $q<0.2$. These residues included the polymorphisms 30R and 31I in p17 (MA), and 309S in p24 (CA). However, in an exploratory analysis of those residues associated with changes in RC with a p value <0.05, it was clear that associated polymorphisms were noticeably enriched on a per residue basis in p17 and p2 (Fisher's exact test, $p<0.001$). The former plays critical roles in intracellular trafficking, and membrane association of Gag [73,79,80], while the latter is an important structural element involved in formation of the immature protein shell [81,82] and a target of the novel drug Bevirimat during maturational cleavage of the Gag precursor [83,84]. Surprisingly, only one third of the associated mutations negatively affected virus replication, while nearly two-thirds of the associations increased fitness. Some of these fitness-increasing mutations represent adapted polymorphisms (i.e. selected as immune escape) and in terms of vaccine design it may be important to avoid the inclusion of such epitopes.

Polymorphisms positively or negatively affecting replication in the p24 region of Gag were limited to just six residues (4 positive, 2 negative), in accordance with the conserved nature of this protein. Surprisingly, none of the canonical B*57/B*5801 associated escape mutations within p24, whose fitness defects have been well documented [24,31-34], were found to be significantly associated with decreases in RC in our present study. This may be due to the high prevalence of B*57/B*5801 positive individuals within this cohort (25%), which could promote viral adaptation to these alleles over time through compensatory mutations [22]. It is also possible that some fitness defects such as those associated with T242N within the TW10 epitope might be

missed in the current study, as a previous study has shown that it is cell-type dependent [32].

The most deleterious HLA-associated mutation that we observed was K12E, which reduced RC by almost 10-fold relative to the median RC of the cohort. This polymorphism is found quite rarely in the population (3 out of 149), and is statistically associated with HLA-A*74, an allele found to be highly protective in both this Zambian subtype C cohort as well as others [9,85]. The protective effect of A*74 has recently been demonstrated to be independent of HLA-B*57 [86]. The negative *in vitro* impact of mutations at residue 12 on replication is supported by a longitudinal study of a subset of this seroconvertor cohort (n=81), in whom polymorphisms at residue 12 were found to revert at a high rate (25%/yr), over the first two years of infection (Schaefer *et al.*, manuscript in preparation). Furthermore, in this same study, escape at position 12 occurred only once and at 24 months post-infection in a total of ten A*74 positive individuals, confirming the high fitness cost associated with CTL-induced escape mutations at this position. We hypothesize that the targeting of this putative epitope, KR9 [86], may account for part of the protective effect conferred by A*74 and indicates that protective immune responses can target regions of Gag outside of p24. While the nature of the replication defect in viruses encoding K12E remains to be determined, this residue does lie in the highly basic region at the N-terminus of p17 (MA), which is involved in membrane targeting and membrane association of Gag [80,87,88].

Rare polymorphisms have the greatest effect on replicative capacity

Rare mutations, such as K12E, which occur in a small subset of the population studied here (less than 10 individuals of the 149), affected fitness to a statistically greater degree than more common polymorphisms. Rare fitness decreasing mutations are likely

unique to specific circumstances such as those where a considerable decrease in RC is warranted in the face of a very effective cellular immune response that is largely abrogated upon mutation. Such mutations have been found to subsequently revert after transmission to individuals lacking the selecting HLA-allele [17,42,89] and in whom they now confer a fitness deficit for the virus. These sites of rare fitness reducing polymorphisms may emphasize vulnerable epitopes at which HIV-1 escapes from immune pressure with great difficulty. Alternatively, it is possible that, when escape occurs, it is consistently associated with a decrease in RC that cannot be completely compensated.

A similar observation was made for rare mutations that greatly increase RC. Global compensatory mutations do exist that can compensate multiple deleterious mutations, such as those within the cyclophilin binding loop [42,46]. Some of the rare fitness increasing mutations may be of this type, although those reported previously have generally been quite common in the population. Compensatory mutations can also be secondary site-suppressors of deleterious mutations [90]. Frequently, such mutations are only conditionally beneficial and can be deleterious in a different context, which could explain why some fitness increasing mutations are rare. It is also possible that these mutations do carry some unrecognized *in vivo* fitness cost that cannot be captured in the *in vitro* replication system used here. Due to the fact that these mutations are rare, they are difficult to statistically link to HLA alleles or to link to other residues with which they may covary, making the potential fitness defects that these mutations mitigate difficult to elucidate.

The cumulative and qualitative effect of HLA associated polymorphisms in Gag on replicative capacity and VL

A key goal of this study was to understand how the cellular immune response might select for mutations in Gag that reduce viral RC, and while identification of specific amino acid polymorphisms that either increase or decrease fitness can be informative, it is equally important to elucidate how the accumulation of specific HLA-associated polymorphisms in Gag affects both RC of the virus and VL in the newly infected person. Previous efforts to correlate the total number of HLA-associated polymorphisms in Gag to RC have yielded inconclusive results [32,54], perhaps because the quality of the polymorphisms in question was not considered. Using an expanded list of HLA-associated polymorphisms (Carlson, Schaefer *et al.*, manuscript in preparation) we report a weak positive correlation between the total number of HLA-associated polymorphisms in Gag and RC. The fact that this correlation was positive is consistent with our observation that a large fraction of the non-consensus HLA-associated polymorphisms increased RC. In particular, in the expanded data set of HLA-associated polymorphisms, we observed that non-adapted residues, which would be predicted to render the virus susceptible to the linked HLA allele, were statistically associated with increased fitness.

These findings suggest that CTL escape mutations, which decrease the overall RC of the virus, are being driven to consensus as a result of population level immune pressure. In the absence of immune pressure, the non-escaped (non-adapted) residues would be expected to predominate, but if they render the virus susceptible to a large portion of the population, then the consensus residue will be escaped rather than susceptible, despite reducing *in vitro* fitness. This is consistent with the findings of Kawashima *et al.* [22] that the frequency of certain HLA-class I alleles within a particular

population can influence the fixation of escape mutations in the overall population. Moreover, Wright *et al.* [55] showed that Gag-NL43 recombinant viruses encoding *gag-pro* sequences most disparate from the subtype C consensus *gag-pro* sequence had statistically higher replicative capacities than their more consensus-like counterparts, and this finding has been recapitulated in this current study. Taken together, these data suggest that overall, HLA-mediated adaptation is driving the fixation of consensus residues that are less fit than their susceptible counter-parts.

When we account for this ability of HLA-associated polymorphisms to either increase or decrease fitness by assigning a summed polymorphism score, which subtracts the number of fitness decreasing polymorphisms from the number of fitness increasing polymorphisms in a particular sequence, we find a highly statistically significant correlation between RC and the summed polymorphism score ($p<0.0001$). Although this p-value should be interpreted cautiously, since it reflects the summation of features previously identified to be correlated with RC, the data do suggest that the effect of polymorphisms is cumulative, and that as a Gag sequence accumulates an excess of fitness-reducing polymorphisms, the RC decreases proportionally. Similarly, utilization of a summed polymorphism score improved previously reported correlations between the total number of HLA-associated polymorphisms in Gag and set point VL in newly infected individuals [50]. We observed a highly statistically significantly correlation ($p = 0.006$) between the summed polymorphism score and set point VL in newly infected individuals. Just as this balance of fitness increasing and decreasing polymorphisms impacts RC, it simultaneously influences the set point VL of the newly infected individual.

Transmission of viruses with low RCs provides recipients with a longer-term clinical benefit

While VL has been demonstrated to influence the rate of disease progression in HIV-1 infected individuals [68,69], it is possible that, during the very earliest stages of infection and before host immune control, the replication rate of the virus may affect the rate of future damage to the immune system. Indeed, we observed a statistically significant negative correlation between RC and average CD4 counts for the first year post infection, suggesting a role for RC in defining this important parameter of pathogenesis at early stages after infection. However, it is possible that this early benefit could be quickly lost due to further adaptation of the virus to the new host's immunogenetic background and further compensation for *de novo* escape mutations. Consequently, we analyzed individuals with longitudinal CD4 counts out to three-years post infection in order to determine if the observed early benefit was sustained in early chronic stages of infection. Using Kaplan-Meier survival analyses to examine the relative time for individuals infected with viruses encoding *gag* genes conferring RC values of <1 and >2 to reach CD4 T cell counts of 350 after 3 years of infection, we observed a clear and statistically significant difference. This was even more striking when CD4 counts less than 300 were used as the endpoint. Moreover a Cox proportional hazard model demonstrated a significantly increased risk of CD4 counts falling below both 350 (HR 2.36) or 300 (HR 3.80) over the first three years of infection for individuals whose *gag* gene conferred an RC>2 vs. RC<1. These findings indicate that the RC conferred by the transmitted Gag sequence may have profound and prolonged effects on HIV-1 pathogenesis from acute to early chronic stages of infection.

While RC and VL are correlated in the full data set, set point VL does not fully explain the effect of RC on CD4, because we did not observe any statistically significant

differences in set point VL between the two groups (RC<1 and RC>2) for the subset of individuals with CD4+T cell counts (n=63). Moreover, in Cox proportional hazard models which take into account VL, the HR remained high (2.17 and 3.11 respectively) and p values continued to trend or remain borderline significant (p=0.093 and 0.051). This suggests that both VL and RC have independent effects on CD4 decline, however, because this analysis was conducted on a subset of less than half of our initial cohort, additional work is underway to further confirm and extend these results.

It seems possible, therefore, that the RC of the transmitted variant may initiate crucial events, early in infection and dissemination, that dictate both acute and later stage pathogenesis regardless of the ability of the immune system to control viral replication down to set point. Infection with highly replicating variants could lead to a more complete depletion of central memory CD4+ T cell pools at this early time that could predispose an individual to more rapid CD4+ T cell loss, irrespective of adequate control of viral replication. This is evidenced in a few individuals infected with highly replicating Gag variants (RC>2), who go on to control VL to a low set point, but whose CD4+ T cells counts rapidly drop below 300 (data not shown). Additionally, a high level of peak viremia or initial high antigen loads could establish an inflammatory environment that leads to sustained immune activation, which has been implicated as a more reliable marker for disease progression [91]. These possibilities are the focus of ongoing work.

In summary, using an *in vitro* approach to define the impact of polymorphisms in Gag on transmitted virus RC has clearly shown that this property of the virus is a significant contributor to early set point VL in a newly infected individual. More importantly, however, these studies suggest a critical role for RC in defining the trajectory of immune depletion and pathogenesis, beyond simply its impact on VL, and highlight the importance of the very earliest events in virus-host interactions. It also

raises the possibility that a vaccine that can attenuate early virus replication would have a positive impact both on vaccinated individuals, as well as non-vaccinated individuals by weakening the transmitted/founder virus and increasing the likelihood of transmission of low replicating variants.

AUTHOR'S CONTRIBUTION

The study was conceived and designed by: DTC, JLP, SAV, PG, JMC and EH. All experiments were performed by: DTC, JLP, MS. Data analysis was performed by EH, DTC, JLP, JMC, DH, TY, JT, RAK, MP, HAP, PG, and SAA. LY, PF, RAK, JT, SL, JM, WK, JG, and SAA contributed reagents, materials and analysis tools. DTC, JLP, EH, CD, and JMC wrote and edited the manuscript. SL, JM, WK, JG, SAA were critically involved in sample collection, site maintenance, participant recruitment and follow-up. All authors read and approved the manuscript.

ACKNOWLEDGEMENTS

The investigators thank all the volunteers in Zambia who participated in this study and all the staff at the Zambia Emory HIV Research Project in Lusaka who made this study possible. The investigators would like to thank Jon Allen, Smita Chavan, and Mackenzie Hurlston for technical assistance and sample management. We would also like to thank Dr. Mark Brockman for his discussions and generous donation of the GXR25 cells.

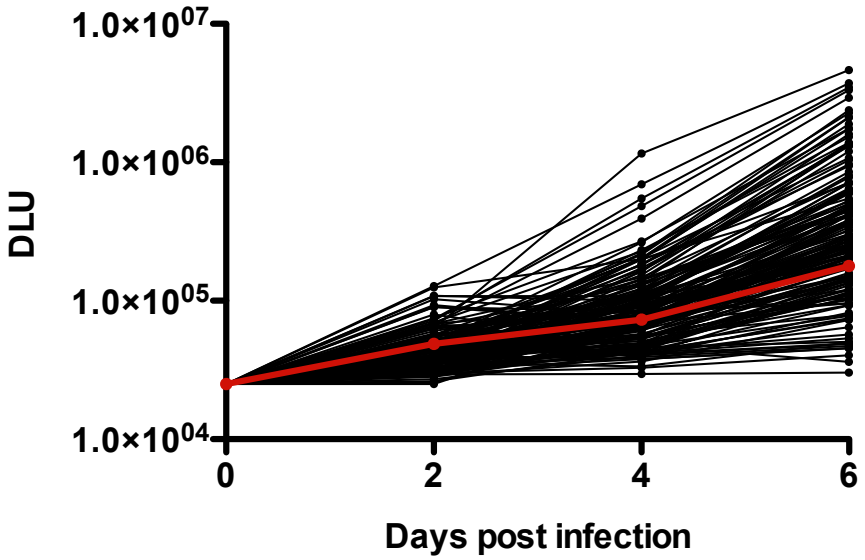


Figure 1. Insertion of the *gag* gene from newly infected individuals dramatically alters the replicative capacity of MJ4

In order to generate replication curves, 5 x 10^5 GXR25 cells were infected with Gag-MJ4 chimeras at an MOI of 0.05. Viral supernatants (100 μL) were collected on days 2, 4, and 6, and the amount of virus in supernatants was quantified using a radiolabeled reverse transcriptase assay. Replicative capacity scores were generated by dividing the $\log_{10}$-transformed slope of replication of each Gag-MJ4 chimera by the $\log_{10}$-transformed slope of MJ4. It is clear that the insertion of the *gag* gene alone can dramatically alter the *in vitro* RC, generating viruses with up to 1000-fold different replication curves. DLU, digital light units.

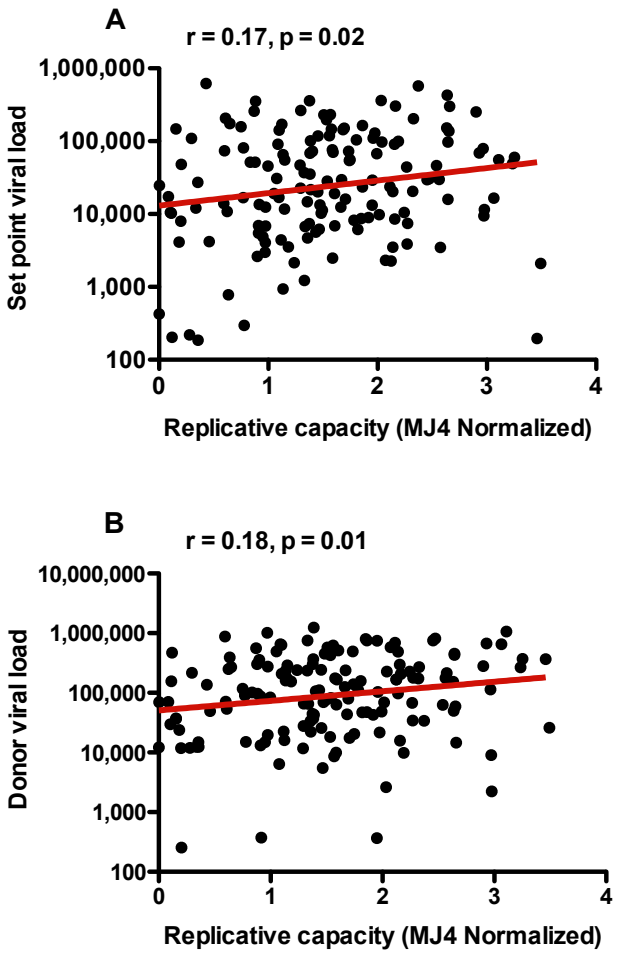
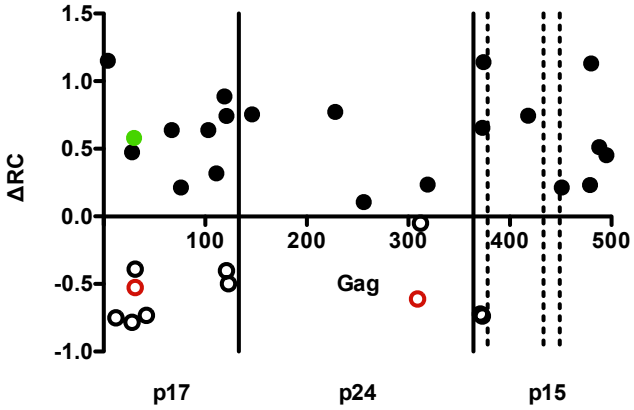
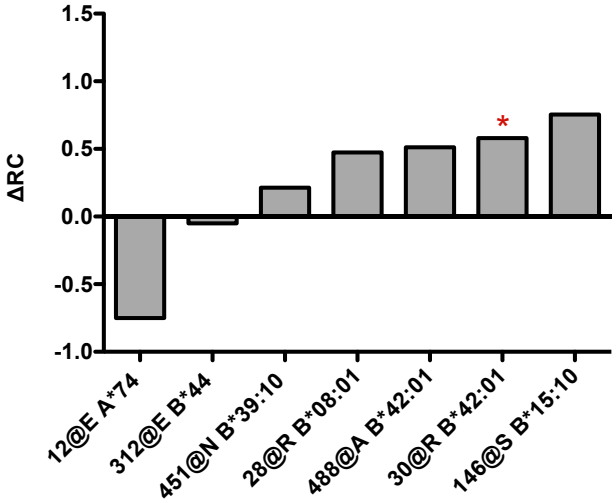


Figure 2. Replicative capacity is correlated to viral load in recipients and donors (A) The RC of Gag-MJ4 chimeras generated from *gag* sequences isolated from epidemiologically linked recipients at acute time points correlates to early set point VL in the same acutely infected recipients (Spearman correlation, r = 0.17, p = 0.02). Replicative capacity scores were generated by normalizing the $\log_{10}$-transformed slopes of replication curves from days 2 through 6 for each Gag-MJ4 chimeric virus to the $\log_{10}$-transformed slope of wild-type MJ4. (B) The RC of Gag-MJ4 chimeric viruses also correlates to the VL near the estimated date of infection in chronically infected donors (Spearman correlation, r = 0.18, p = 0.01). Trend lines were generated using linear regression analysis, and are shown in order to facilitate visualization of correlations.


1.5
1.0
0.5
0.0
-0.5
-1.0
ΔRC
100
200
300
400
500
Gag
p17
p24
p15

B


1.5
1.0
0.5
0.0
-0.5
-1.0
ΔRC
*
12@E A*74
312@E B*44
451@N B*39:10
28@R B*08:01
488@A B*42:01
30@R B*42:01
146@S B*15:10

C

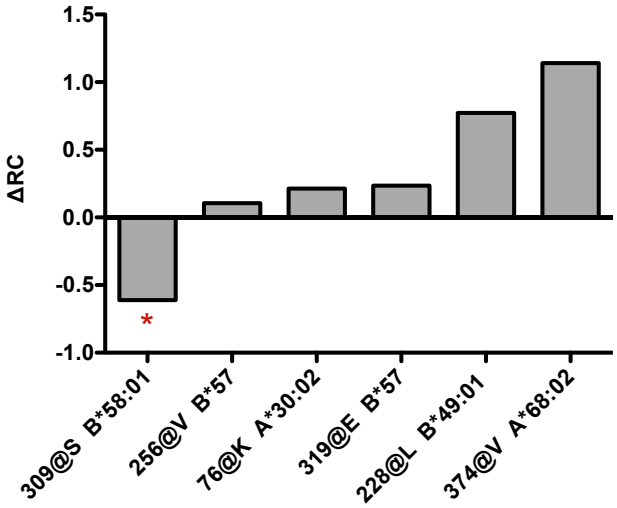

1.5
1.0
0.5
0.0
-0.5
-1.0
ΔRC
*
309@S B*58:01
256@V B*57
76@K A*30:02
319@E B*57
228@L B*49:01
374@V A*68:02

Figure 3. Identification of polymorphisms in Gag that significantly affect RC, several of which can be linked to HLA-class I alleles (A) In a pair-wise analysis of all amino acids represented in the 149 *gag* sequences tested, we observed 49 amino acids at 31 unique positions that were statistically associated ($p<0.05$) with changes in RC (open circles RC decreasing, filled circles RC increasing), 3 of which were significant at ($q<0.2$) when adjusted for multiple comparisons (red for RC reducing and green for RC increasing). The x-axis shows the polymorphism position in the primary Gag sequence (HXB2 numbering), and the y-axis depicts the impact of the polymorphism on RC relative to the median RC of all viruses (~1.5). (B) In a separate study analyzing 1899 subtype C *gag* sequences from Zambia and South Africa, 199 residues were linked to HLA-I alleles ($q< 0.2$, Carlson, Schaefer *et al.,* manuscript in preparation). From this, a total of 7 polymorphisms associated with changes in RC ($p< 0.05$) were found to be adapted to specific HLA-I alleles, adapted (amino acid is present when HLA-I allele is also present.) (* denotes $q<0.2$). (C) Six polymorphisms associated with changes in RC ($p<0.05$) were found to be non-adapted (amino acid is present only when HLA-I allele is absent) to specific HLA-I alleles (* denotes $q<0.2$).

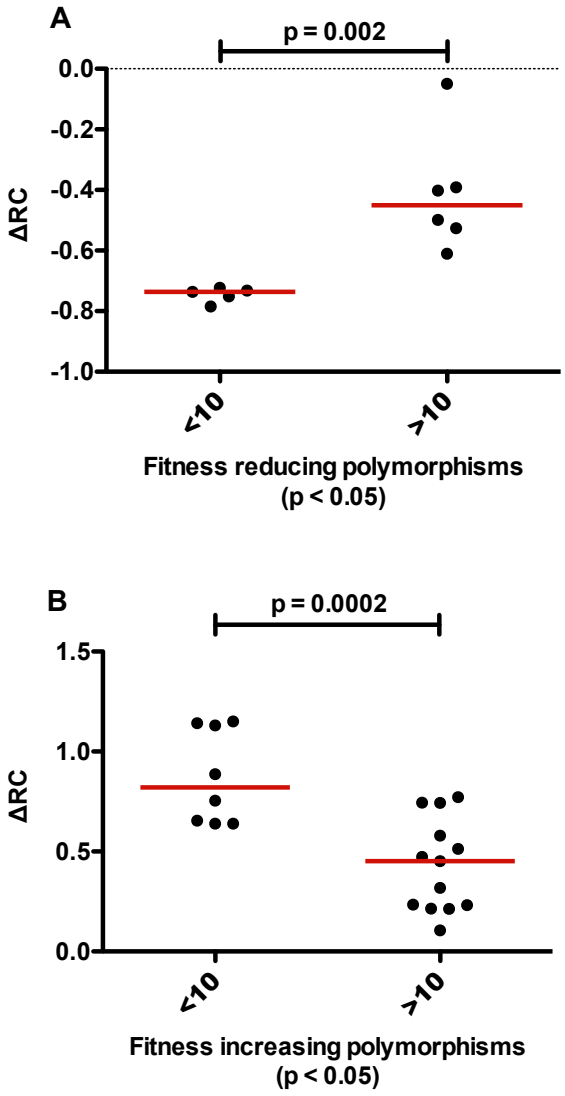


Figure 4. Rare polymorphisms have a significantly greater impact on RC (A) In an exploratory analysis, rare fitness reducing polymorphisms, occurring in less than 10 out of 149 of the sequences tested, were found to decrease fitness to a significantly greater degree than more common fitness decreasing mutations (Mann Whitney U test, p = 0.002). The y-axis depicts the impact of the polymorphism on RC relative to the median RC of all viruses (~1.5) (B) Rare fitness increasing polymorphisms, occurring in less than 10 out of 149 of the sequences tested, were found to increase fitness to a

significantly greater degree than more common ones (Mann Whitney U test, p = 0.0002).

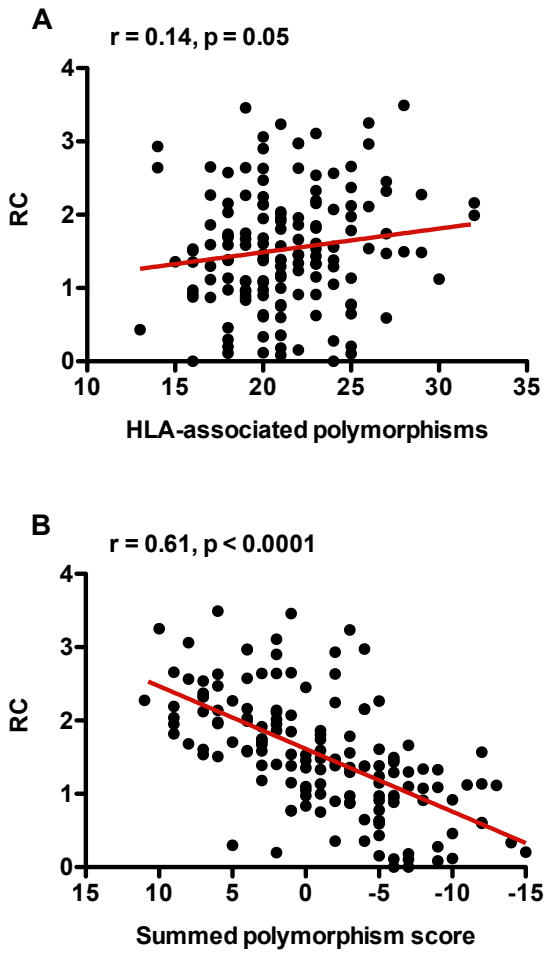


Figure 5. The balance of fitness increasing and decreasing HLA-associated polymorphisms strongly correlates with RC (A) The total number of HLA-associated polymorphisms positively correlates with RC. HLA-associated polymorphisms were defined as non-consensus residues at any amino acid position known to have polymorphisms statistically linked to HLA-I alleles. (B) RC of Gag-MJ4 chimeras is highly correlated to the summed polymorphism score in Gag. Each amino acid polymorphism was given a score of 1 (fitness increasing) or -1 (fitness decreasing). Summed polymorphism scores were generated by summing these scores for each Gag protein. Trend lines were generated using linear regression analysis, and are shown in order to facilitate visualization of correlations.

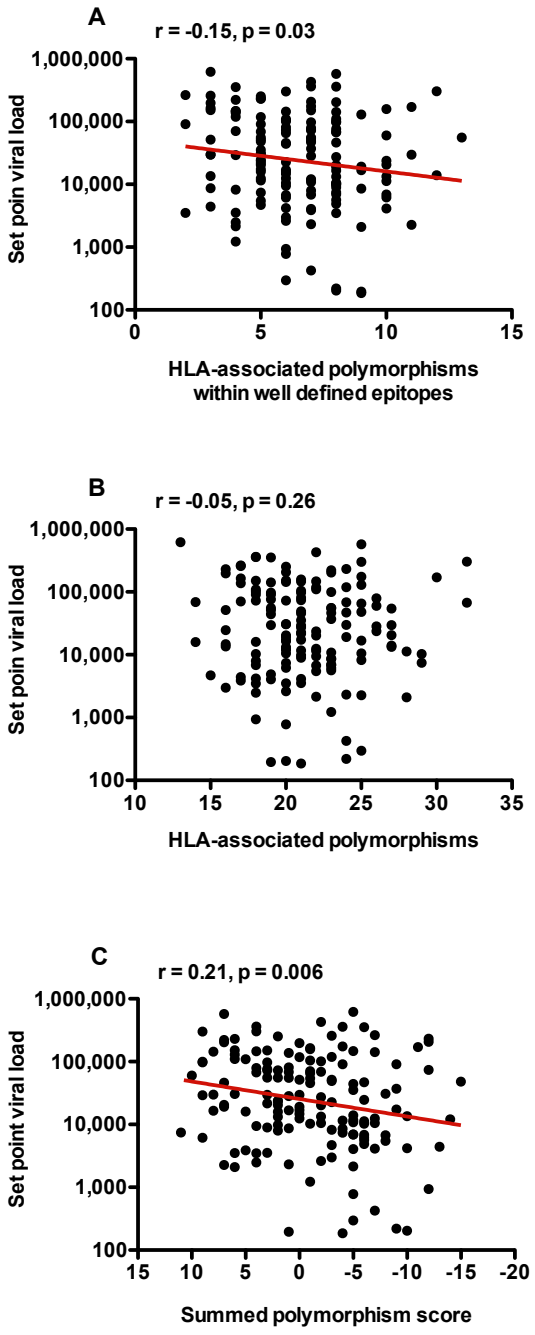


Figure 6. The balance of HLA-associated fitness increasing and decreasing mutations strongly correlates with set point viral load in newly infected individuals (A) The number of HLA-associated amino acid polymorphisms within well-

defined epitopes for each Gag protein negatively correlates with set point VL. (B) The total number of HLA-associated polymorphisms (including those outside well-defined CTL epitopes) in Gag does not correlate to set point VL in newly infected individuals. (C) When the quality of HLA-associated polymorphisms is considered, a strong correlation between the summed polymorphism score (as defined in Figure 5) and set point VL is observed. Trend lines were generated using linear regression analysis, and are shown in order to facilitate visualization of correlations.

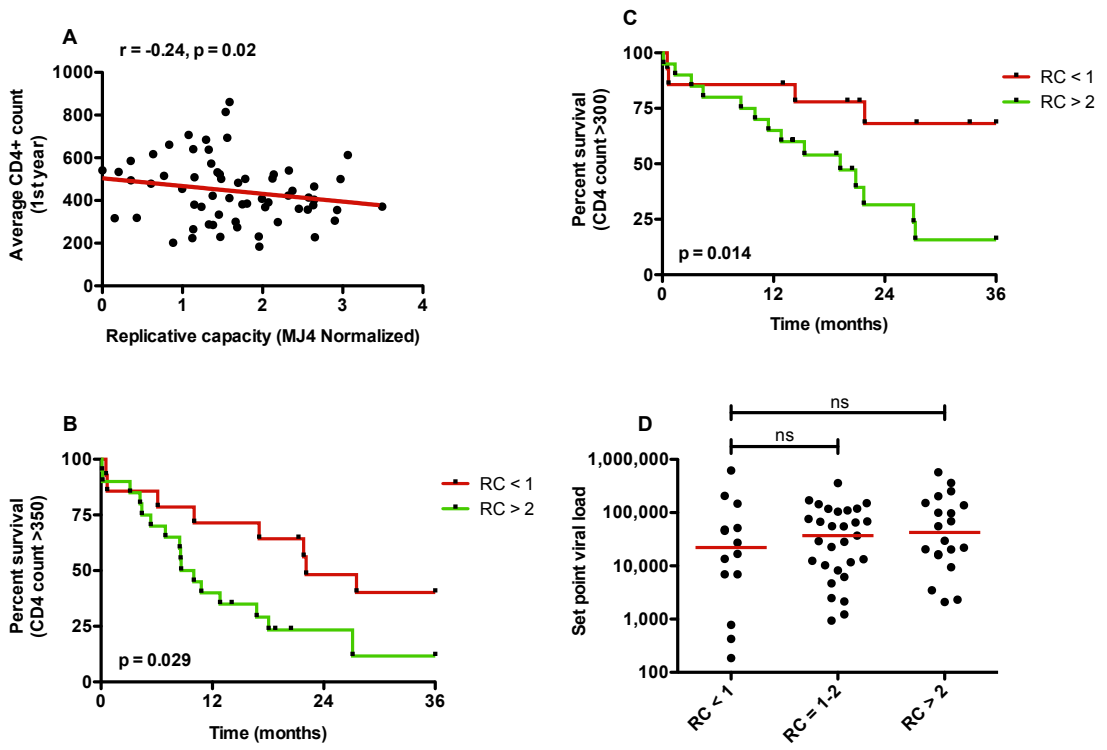


Figure 7. RC affects the rate of CD4 decline in a manner that may be independent of viral load (A) In a subset of individuals for which longitudinal CD4+ counts were available for at least 1 year post infection (n=63), RC of Gag-MJ4 chimeras negatively correlated with the average CD4+ counts of linked recipients within the first year (Spearman rank correlation, r = -0.24, p = 0.02). (B) Kaplan-Meier plots in which the endpoint was defined as the first CD4+ count below 350. Interval cut-off for endpoint was set to 36 months. The difference in median time to endpoint between those receiving viruses with RC<1 (n = 14) and RC>2 (n = 20) was 384 days (Log-rank test, p = 0.029). (C) Kaplan-Meier plots in which the endpoint was defined as the first CD4+ count below 300. Interval cut-off for endpoint was set to 36 months. The difference in median time to endpoint between those receiving viruses with RC<1 (n = 14) and RC>2 (n = 20) was >800 days (Log-rank test, p = 0.014). (D) The difference in median VL between RC groups RC<1, RC 1-2, and RC >2 were not significantly different (Mann

Whitney U test), consistent with the effect of RC on CD4 decline being independent of VL.

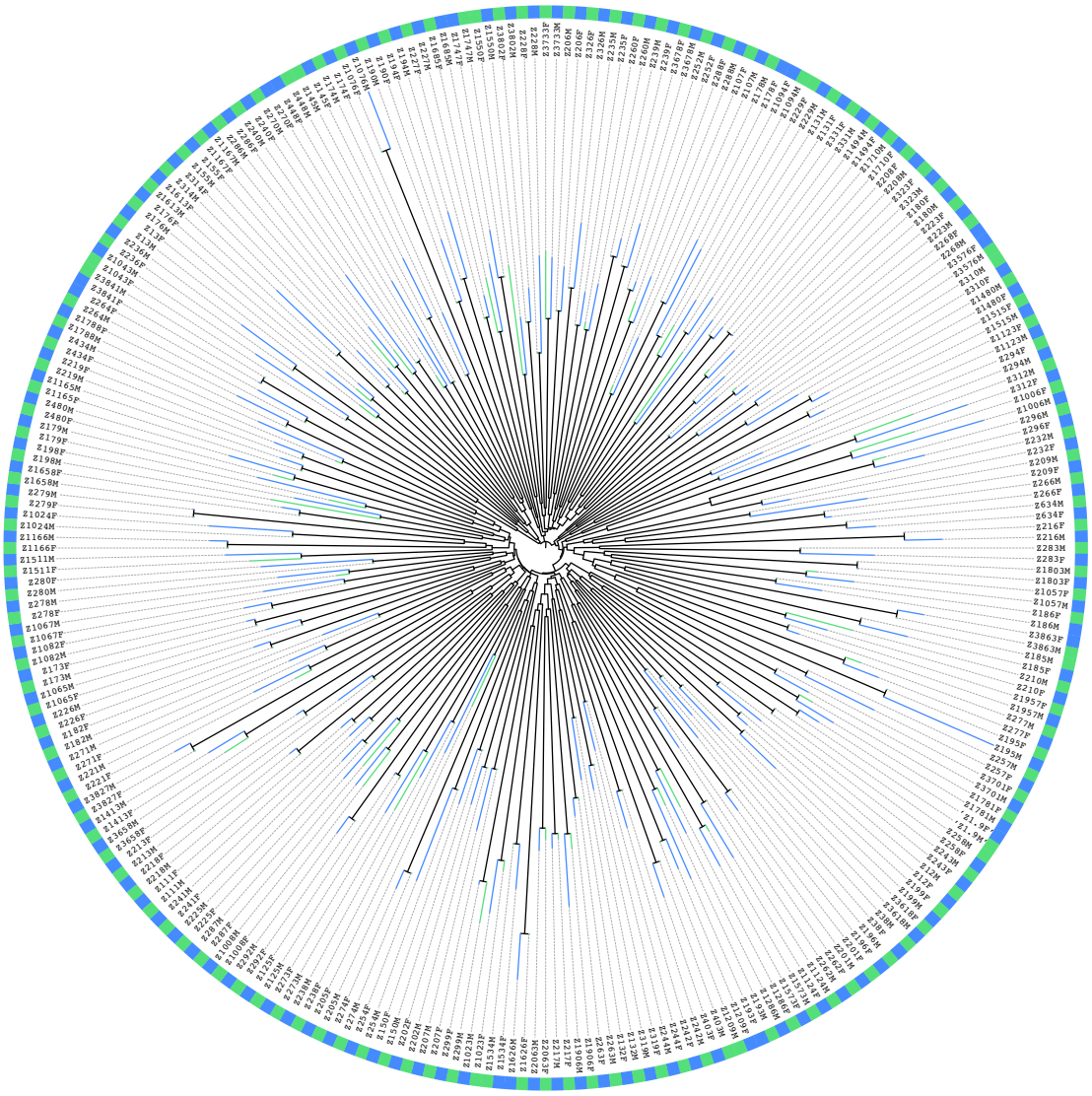


Figure S1. Donor and recipient population *gag* sequences cluster with one another Gag population sequences from donors and linked recipients were amplified and sequenced as described in the methods section. Nucleotide *gag* sequences were aligned using the Gene Cutter tool accessible on the Los Alamos Nation Lab HIV Sequence Database (http://www.hiv.lanl.gov/content/sequence/GENE_CUTTER/cutter.html) and a neighbor-joining tree was generated using the Geneious sequence analysis software

v5.5.7 (Biomatters Ltd.). The radial tree was annotated using the Interactive Tree of Life (iTOL) online tool for phylogenetic tree display and annotation (Letunic and Bork, Bioinformatics, 2006). Blue denotes transmitting partners (donors) and green denotes linked recipients. This tree demonstrates the high degree of similarity between donor and linked recipient *gag* population sequences within epidemiologically linked transmission pairs.

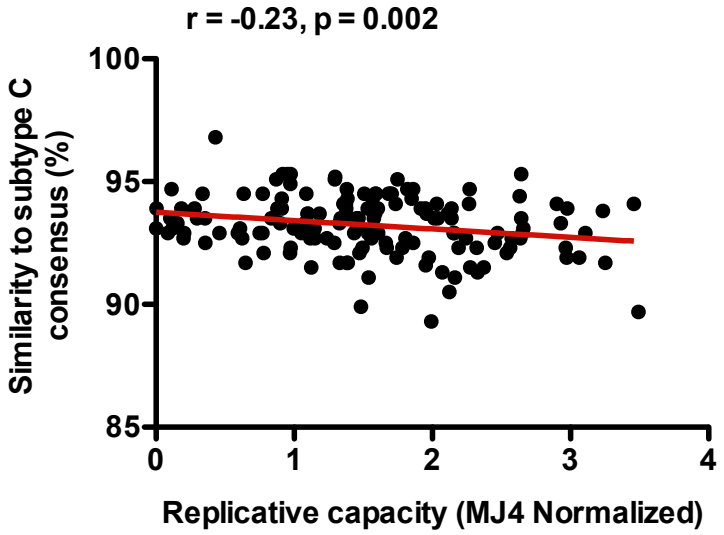


Figure S2. Gag sequences that are less like the Gag subtype C consensus sequence replicate more efficiently *in vitro* The Gag amino acid sequences of all Gag-MJ4 chimeras were compared to a Zambian subtype C consensus Gag sequence (generated using the LANL Consensus Maker tool; http://www.hiv.lanl.gov/content/sequence/CONSENSUS/consensus.html) by building a neighbor-joining tree using the Geneious sequence analysis software v5.5.7 (Biomatters Ltd.) and the percent similarity to consensus was determined for each sequence. Notably, viruses encoding Gag sequences most disparate from the subtype C consensus Gag sequence replicated to higher levels, and we observed a statistically significant negative correlation between RC and the percent similarity of Gag to consensus (Pearson correlation, $p = 0.002$, $r = -0.23$).

Table 1. Cohort statistics generated from the 149 transmission pairs selected from the ZEHRP cohort.

Parameters	Linked Seroconverter (SC)
Number of individuals	149
Number of Females (%)	77 (52%)
Age of SC at time of seroconversion	28 (24 – 35)[a]
Estimated days post infection	45.5 (41.5 – 50.5)[a]
Log_{10} Set-point VL of SC	4.39 (3.91 – 4.99)[a]
Log_{10} VL of Donor Partner[b]	5.02 (4.51 – 5.45)[a]
CD4 cell count 1yr post SC[c]	383 (308 – 503)[a]

[a] Median (Inter-Quartile Range)

[b] Measured at or within one month of seroconversion of LR

[c] LR for whom we have longitudinal CD4+ T cell counts (n=63)

Table 2. Cox proportional hazard models demonstrate the independent effects of RC and VL on CD4 decline

Risk of CD4 below 350 in first 3 years	**p**	**Hazard Ratio[c]**	**lower .95 [d]**	**upper .95 [d]**
RC>2[a]	0.034	2.62	1.077	6.374
Log10(SetPoint VL)[b]	0.04	1.75	1.014	3.004
RC>2 + Log10(SetPointVL)	0.093	2.18	0.878	5.395
Risk of CD4 below 300 in first 3 years	**p**	**Hazard Ratio[c]**	**lower .95 [d]**	**upper .95 [d]**
RC>2[a]	0.021	3.80	1.226	11.801
Log10(SetPoint VL)[b]	0.023	2.09	1.105	3.94
RC>2 + log10(SetPointVL)	0.051	3.12	0.995	9.751

[a] as compared to RC<1

[b] Model in which $\log_{10}$ set point viral load in the linked recipient is taken into account

[c] The ratio of hazard rates between the variables RC>2 and RC<1

[d] The upper and lower bounds of confidence

Codon	Amino Acid	Consensus Amino Acid	Sequences w/o residue	Sequences with residue	Median RC w/o residue	Median RC with residue	p value	q value	ΔRC[a]	Within Epitope[b]	Epitope HLA Restriction[c]
31	L	L	49	103	1.051613	1.694342	0.00001	0.00545	0.19326	YES	A*2402; B*0801
31	I	L	119	33	1.606183	0.974471	0.00016	0.04360	-0.52661	YES	A*2402; B*0801
30	R	M	130	22	1.442259	2.080627	0.00107	0.18394	0.57955	YES	A*2402; B*0801
309	S	A	126	26	1.572069	0.890472	0.00135	0.18394	-0.61061	YES	B*4402; B*5301; B*5701; Cw5
28	R	H	125	27	1.433357	1.974499	0.00326	0.32077	0.47342	YES	A*0301; A*2402; B*0801
309	A	A	29	123	0.910234	1.567691	0.00397	0.32077	0.06661	YES	B*4402; B*5301; B*5701; Cw5
370	A	A	48	104	1.797895	1.380252	0.00412	0.32077	-0.12083	YES	B*4501
4	S	R	149	3	1.484227	2.651962	0.00965	0.40790	1.15088	NO	
4	R	R	3	149	2.651962	1.484227	0.01005	0.40790	-0.01685	NO	
451	N	S	96	56	1.417936	1.713877	0.01179	0.40790	0.21280	NO	
111	S	S	61	90	1.74016	1.383489	0.01199	0.40790	-0.11759	NO	
418	K	K	22	130	2.0575	1.468355	0.01217	0.40790	-0.03272	NO	
373	A	N	125	4	1.534145	0.764576	0.01258	0.40790	-0.73650	YES	B*4501
146	S	A	142	10	1.468355	2.255673	0.01380	0.40790	0.75460	YES	A*2501; B*1510; B*5701
119	A	E	145	4	1.472093	2.387669	0.01388	0.40790	0.88659	NO	
121	A	D	137	15	1.472093	2.244058	0.01418	0.40790	0.74298	NO	
42	D	E	145	7	1.534145	0.769036	0.01449	0.40790	-0.73204	YES	A*30; B*3501
451	S	S	63	89	1.68293	1.386771	0.01499	0.40790	-0.11431	NO	
69	Q	Q	26	126	1.822397	1.442259	0.01663	0.40790	-0.05882	NO	
12	E	K	149	3	1.532977	0.750097	0.01901	0.40790	-0.75098	YES	B*4002
418	R	K	136	16	1.47816	2.245238	0.01981	0.40790	0.74416	NO	
319	E	D	125	27	1.464616	1.735435	0.02014	0.40790	0.23436	NO	
319	D	D	27	125	1.735435	1.464616	0.02014	0.40790	-0.03646	NO	
85	L	L	4	148	2.364864	1.47816	0.02077	0.40790	-0.02292	YES	A*0201; A*1101; A*2902; A*3002; B*4403; B*58; B*63; Cw14
111	C	S	110	41	1.417936	1.819739	0.02115	0.40790	0.31866	NO	
480	G	D	145	3	1.49407	2.632309	0.02118	0.40790	1.13123	NO	
62	K	K	61	91	1.68293	1.386563	0.02145	0.40790	-0.11451	NO	
312	E	D	79	73	1.588149	1.451162	0.02178	0.40790	-0.04992	YES	B*4402; B*5301; B*5701; Cw5
312	D	D	73	79	1.451162	1.588149	0.02178	0.40790	0.08707	YES	B*4402; B*5301; B*5701; Cw5
123	E	G	133	13	1.535189	1.002543	0.02266	0.40790	-0.49853	NO	
495	N	S	117	35	1.451162	1.952761	0.02395	0.40790	0.45168	NO	
495	S	S	35	117	1.952761	1.451162	0.02395	0.40790	-0.04992	NO	
228	L	M	141	11	1.472093	2.273388	0.02650	0.43765	0.77231	YES	B*13
103	K	K	10	142	2.086015	1.474422	0.03202	0.48755	-0.02666	NO	
371	S	N	148	3	1.520532	0.778439	0.03300	0.48755	-0.72264	YES	B*4501
67	S	A	145	7	1.472093	2.139988	0.03310	0.48755	0.63891	NO	
67	A	A	7	145	2.139988	1.472093	0.03310	0.48755	-0.02898	NO	
103	R	K	143	9	1.484227	2.139988	0.03564	0.50140	0.63891	NO	
373	Q	N	124	5	1.47816	2.15603	0.03588	0.50140	0.65495	YES	B*4501
42	E	E	10	142	0.903087	1.534667	0.03691	0.50290	0.03359	YES	A*30; B*3501
121	G	D	135	17	1.538821	1.099119	0.03823	0.50552	-0.40196	NO	
76	K	R	110	42	1.394643	1.714888	0.04012	0.50552	0.21381	YES	A*3002; B*0801; B*4403; B*58; B*63
28	T	H	148	4	1.533561	0.716807	0.04064	0.50552	-0.78427	YES	A*0301; A*2402; B*0801
488	A	S	141	11	1.472093	2.013763	0.04174	0.50552	0.51269	YES	B*4001
488	S	S	11	141	2.013763	1.472093	0.04174	0.50552	-0.02898	YES	B*4001
479	R	K	128	22	1.459385	1.733306	0.04370	0.51775	0.23223	NO	
31	M	L	136	16	1.534083	1.109909	0.04710	0.54616	-0.39117	YES	A*2402; B*0801
256	V	I	99	53	1.454154	1.606183	0.04825	0.54784	0.10511	YES	B*3501
374	V	T	149	3	1.484227	2.642477	0.04981	0.55401	1.14140	YES	B*4501

Table S1. Amino acids in Gag associated with changes in replicative capacity

This table lists all amino acids associated with changes in RC. Residues that remain significantly associated with changes in RC after correction for multiple comparisons

(q<0.2) are depicted in green. A total of 152 sequences and RC values were available for association analysis, with 149 of these with sufficient clinical follow-up for inclusion in the broader study.

[a] ΔRC is defined as median RC of all viruses tested (~1.5) subtracted from the median RC of all viruses with the particular polymorphism.

[b] The location of epitopes was defined by the compendium of "A-list" epitopes available in the LANL Immunology Database (*HIV Molecular Immunology 2009)*.

[c] *HLA class I alleles restricting epitopes harboring these polymorphisms that affect RC were also defined base on the LANL Immunology Database compilation of "A-list" epitopes.*

REFERENCES CITED:

1. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, et al. (2008) Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc Natl Acad Sci U S A 105: 7552-7557.

2. Haaland RE, Hawkins PA, Salazar-Gonzalez J, Johnson A, Tichacek A, et al. (2009) Inflammatory genital infections mitigate a severe genetic bottleneck in heterosexual transmission of subtype A and C HIV-1. PLoS Pathog 5: e1000274.

3. Derdeyn CA, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, et al. (2004) Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. Science 303: 2019-2022.

4. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, et al. (2007) A whole-genome association study of major determinants for host control of HIV-1. Science 317: 944-947.

5. Kaslow RA, Carrington M, Apple R, Park L, Munoz A, et al. (1996) Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. Nat Med 2: 405-411.

6. Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C, et al. (2004) Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature 432: 769-775.

7. Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, et al. (2010) The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science 330: 1551-1557.

8. Tang J, Kaslow RA (2003) The impact of host genetics on HIV infection and disease progression in the era of highly active antiretroviral therapy. AIDS 17 Suppl 4: S51-60.

9. Tang J, Malhotra R, Song W, Brill I, Hu L, et al. (2010) Human leukocyte antigens and HIV type 1 viral load in early and chronic infection: predominance of evolving relationships. PLoS One 5: e9629.

10. Tang J, Tang S, Lobashevsky E, Myracle AD, Fideli U, et al. (2002) Favorable and unfavorable HLA class I alleles and haplotypes in Zambians predominantly infected with clade C human immunodeficiency virus type 1. J Virol 76: 8276-8284.

11. Tang J, Tang S, Lobashevsky E, Zulu I, Aldrovandi G, et al. (2004) HLA allele sharing and HIV type 1 viremia in seroconverting Zambians with known transmitting partners. AIDS Res Hum Retroviruses 20: 19-25.

12. Hollingsworth TD, Laeyendecker O, Shirreff G, Donnelly CA, Serwadda D, et al. (2010) HIV-1 transmitting couples have similar viral load set-points in Rakai, Uganda. PLoS Pathog 6: e1000876.

13. Hecht FM, Hartogensis W, Bragg L, Bacchetti P, Atchison R, et al. (2010) HIV RNA level in early infection is predicted by viral load in the transmission source. AIDS 24: 941-945.

14. Rong R, Li B, Lynch RM, Haaland RE, Murphy MK, et al. (2009) Escape from autologous neutralizing antibodies in acute/early subtype C HIV-1 infection requires multiple pathways. PLoS Pathog 5: e1000594.

15. Borrow P, Lewicki H, Wei X, Horwitz MS, Peffer N, et al. (1997) Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. Nat Med 3: 205-211.

16. Goulder PJ, Phillips RE, Colbert RA, McAdam S, Ogg G, et al. (1997) Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. Nat Med 3: 212-217.

17. Leslie AJ, Pfafferott KJ, Chetty P, Draenert R, Addo MM, et al. (2004) HIV evolution: CTL escape mutation and reversion after transmission. Nat Med 10: 282-289.

18. Phillips RE, Rowland-Jones S, Nixon DF, Gotch FM, Edwards JP, et al. (1991) Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. Nature 354: 453-459.

19. Price DA, Goulder PJ, Klenerman P, Sewell AK, Easterbrook PJ, et al. (1997) Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. Proc Natl Acad Sci U S A 94: 1890-1895.

20. Jones NA, Wei X, Flower DR, Wong M, Michor F, et al. (2004) Determinants of human immunodeficiency virus type 1 escape from the primary CD8+ cytotoxic T lymphocyte response. J Exp Med 200: 1243-1256.

21. Goulder PJ, Watkins DI (2004) HIV and SIV CTL escape: implications for vaccine design. Nat Rev Immunol 4: 630-640.

22. Kawashima Y, Pfafferott K, Frater J, Matthews P, Payne R, et al. (2009) Adaptation of HIV-1 to human leukocyte antigen class I. Nature 458: 641-645.

23. Allen TM, Altfeld M, Yu XG, O'Sullivan KM, Lichterfeld M, et al. (2004) Selection, transmission, and reversion of an antigen-processing cytotoxic T-lymphocyte escape mutation in human immunodeficiency virus type 1 infection. J Virol 78: 7069-7078.

24. Chopera DR, Woodman Z, Mlisana K, Mlotshwa M, Martin DP, et al. (2008) Transmission of HIV-1 CTL escape variants provides HLA-mismatched recipients with a survival advantage. PLoS Pathog 4: e1000033.

25. Friedrich TC, Dodds EJ, Yant LJ, Vojnov L, Rudersdorf R, et al. (2004) Reversion of CTL escape-variant immunodeficiency viruses in vivo. Nat Med 10: 275-281.

26. Li B, Gladden AD, Altfeld M, Kaldor JM, Cooper DA, et al. (2007) Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. J Virol 81: 193-201.

27. Matthews PC, Prendergast A, Leslie A, Crawford H, Payne R, et al. (2008) Central role of reverting mutations in HLA associations with human immunodeficiency virus set point. J Virol 82: 8548-8559.

28. Rousseau CM, Daniels MG, Carlson JM, Kadie C, Crawford H, et al. (2008) HLA class I-driven evolution of human immunodeficiency virus type 1 subtype c proteome: immune escape and viral load. J Virol 82: 6434-6446.

29. Schneidewind A, Brockman MA, Sidney J, Wang YE, Chen H, et al. (2008) Structural and functional constraints limit options for cytotoxic T-lymphocyte escape in the immunodominant HLA-B27-restricted epitope in human immunodeficiency virus type 1 capsid. J Virol 82: 5594-5605.

30. Dahirel V, Shekhar K, Pereyra F, Miura T, Artyomov M, et al. (2011) Coordinate linkage of HIV evolution reveals regions of immunological vulnerability. Proc Natl Acad Sci U S A 108: 11530-11535.

31. Boutwell CL, Rowley CF, Essex M (2009) Reduced viral replication capacity of human immunodeficiency virus type 1 subtype C caused by cytotoxic-T-lymphocyte escape mutations in HLA-B57 epitopes of capsid protein. J Virol 83: 2460-2468.

32. Brockman MA, Schneidewind A, Lahaie M, Schmidt A, Miura T, et al. (2007) Escape and compensation from early HLA-B57-mediated cytotoxic T-lymphocyte pressure on human immunodeficiency virus type 1 Gag alter capsid interactions with cyclophilin A. J Virol 81: 12608-12618.

33. Crawford H, Prado JG, Leslie A, Hue S, Honeyborne I, et al. (2007) 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. J Virol 81: 8346-8351.

34. Martinez-Picado J, Prado JG, Fry EE, Pfafferott K, Leslie A, et al. (2006) Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. J Virol 80: 3617-3623.

35. Schneidewind A, Brockman MA, Yang R, Adam RI, Li B, et al. (2007) 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. J Virol 81: 12382-12393.

36. Wright JK, Naidoo VL, Brumme ZL, Prince JL, Claiborne DT, et al. (2012) Impact of HLA-B*81-associated mutations in HIV-1 Gag on viral replication capacity. J Virol 86: 3193-3199.

37. Mothe B, Llano A, Ibarrondo J, Daniels M, Miranda C, et al. (2011) Definition of the viral targets of protective HIV-1-specific T cell responses. J Transl Med 9: 208.

38. Rolland M, Heckerman D, Deng W, Rousseau CM, Coovadia H, et al. (2008) Broad and Gag-biased HIV-1 epitope repertoires are associated with lower viral loads. PLoS One 3: e1424.

39. Wang YE, Li B, Carlson JM, Streeck H, Gladden AD, et al. (2009) Protective HLA class I alleles that restrict acute-phase CD8+ T-cell responses are associated with viral escape mutations located in highly conserved regions of human immunodeficiency virus type 1. J Virol 83: 1845-1855.

40. Wagner R, Leschonsky B, Harrer E, Paulus C, Weber C, et al. (1999) Molecular and functional analysis of a conserved CTL epitope in HIV-1 p24 recognized from a

long-term nonprogressor: constraints on immune escape associated with targeting a sequence essential for viral replication. J Immunol 162: 3727-3734.

41. Peyerl FW, Barouch DH, Letvin NL (2004) Structural constraints on viral escape from HIV- and SIV-specific cytotoxic T-lymphocytes. Viral Immunol 17: 144-151.

42. Crawford H, Lumm W, Leslie A, Schaefer M, Boeras D, et al. (2009) Evolution of HLA-B*5703 HIV-1 escape mutations in HLA-B*5703-positive individuals and their transmission recipients. J Exp Med 206: 909-921.

43. Brockman MA, Brumme ZL, Brumme CJ, Miura T, Sela J, et al. (2010) Early selection in Gag by protective HLA alleles contributes to reduced HIV-1 replication capacity that may be largely compensated for in chronic infection. J Virol 84: 11937-11949.

44. Carlson JM, Brumme ZL, Rousseau CM, Brumme CJ, Matthews P, et al. (2008) Phylogenetic dependency networks: inferring patterns of CTL escape and codon covariation in HIV-1 Gag. PLoS Comput Biol 4: e1000225.

45. Chopera DR, Mlotshwa M, Woodman Z, Mlisana K, de Assis Rosa D, et al. (2011) Virological and immunological factors associated with HIV-1 differential disease progression in HLA-B 58:01-positive individuals. J Virol 85: 7070-7080.

46. Crawford H, Matthews PC, Schaefer M, Carlson JM, Leslie A, et al. (2011) The hypervariable HIV-1 capsid protein residues comprise HLA-driven CD8+ T-cell escape mutations and covarying HLA-independent polymorphisms. J Virol 85: 1384-1390.

47. Huang KH, Goedhals D, Carlson JM, Brockman MA, Mishra S, et al. (2011) Progression to AIDS in South Africa is associated with both reverting and compensatory viral mutations. PLoS One 6: e19018.

48. Rolland M, Carlson JM, Manocheewa S, Swain JV, Lanxon-Cookson E, et al. (2010) Amino-acid co-variation in HIV-1 Gag subtype C: HLA-mediated selection pressure and compensatory dynamics. PLoS One 5.

49. Schneidewind A, Brumme ZL, Brumme CJ, Power KA, Reyor LL, et al. (2009) Transmission and long-term stability of compensated CD8 escape mutations. J Virol 83: 3993-3997.

50. Goepfert PA, Lumm W, Farmer P, Matthews P, Prendergast A, et al. (2008) Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. J Exp Med 205: 1009-1017.

51. Miura T, Brumme ZL, Brockman MA, Rosato P, Sela J, et al. (2010) Impaired replication capacity of acute/early viruses in persons who become HIV controllers. J Virol 84: 7581-7591.

52. Miura T, Brockman MA, Schneidewind A, Lobritz M, Pereyra F, et al. (2009) 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. J Virol 83: 2743-2755.

53. Miura T, Brockman MA, Brumme ZL, Brumme CJ, Pereyra F, et al. (2009) HLA-associated alterations in replication capacity of chimeric NL4-3 viruses carrying gag-protease from elite controllers of human immunodeficiency virus type 1. J Virol 83: 140-149.

54. Wright JK, Novitsky V, Brockman MA, Brumme ZL, Brumme CJ, et al. (2011) Influence of Gag-protease-mediated replication capacity on disease progression in individuals recently infected with HIV-1 subtype C. J Virol 85: 3996-4006.

55. Wright JK, Brumme ZL, Carlson JM, Heckerman D, Kadie CM, et al. (2010) Gag-protease-mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. J Virol 84: 10820-10831.

56. Ndung'u T, Renjifo B, Essex M (2001) Construction and analysis of an infectious human Immunodeficiency virus type 1 subtype C molecular clone. J Virol 75: 4964-4972.

57. Allen S, Karita E, Chomba E, Roth DL, Telfair J, et al. (2007) Promotion of couples' voluntary counselling and testing for HIV through influential networks in two African capital cities. BMC Public Health 7: 349.

58. Kempf MC, Allen S, Zulu I, Kancheya N, Stephenson R, et al. (2008) Enrollment and retention of HIV discordant couples in Lusaka, Zambia. J Acquir Immune Defic Syndr 47: 116-125.

59. McKenna SL, Muyinda GK, Roth D, Mwali M, Ng'andu N, et al. (1997) Rapid HIV testing and counseling for voluntary testing centers in Africa. AIDS 11 Suppl 1: S103-110.

60. Trask SA, Derdeyn CA, Fideli U, Chen Y, Meleth S, et al. (2002) Molecular epidemiology of human immunodeficiency virus type 1 transmission in a heterosexual cohort of discordant couples in Zambia. J Virol 76: 397-405.

61. Bhakta SJ, Shang L, Prince JL, Claiborne DT, Hunter E (2011) Mutagenesis of tyrosine and di-leucine motifs in the HIV-1 envelope cytoplasmic domain results in a loss of Env-mediated fusion and infectivity. Retrovirology 8: 37.

62. Brockman MA, Tanzi GO, Walker BD, Allen TM (2006) Use of a novel GFP reporter cell line to examine replication capacity of CXCR4- and CCR5-tropic HIV-1 by flow cytometry. J Virol Methods 131: 134-142.

63. Ostrowski MA, Chun TW, Cheseboro B, Stanley SK, Tremblay M (2006) Detection assays for HIV proteins. Curr Protoc Immunol Chapter 12: Unit 12 15.

64. Carlson JM, Listgarten J, Pfeifer N, Tan V, Kadie C, et al. (2012) Widespread impact of HLA restriction on immune control and escape pathways of HIV-1. J Virol 86: 5230-5243.

65. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 100: 9440-9445.

66. Fideli US, Allen SA, Musonda R, Trask S, Hahn BH, et al. (2001) Virologic and immunologic determinants of heterosexual transmission of human immunodeficiency virus type 1 in Africa. AIDS Res Hum Retroviruses 17: 901-910.

67. Saag MS, Holodniy M, Kuritzkes DR, O'Brien WA, Coombs R, et al. (1996) HIV viral load markers in clinical practice. Nat Med 2: 625-629.

68. Mellors JW, Rinaldo CR, Jr., Gupta P, White RM, Todd JA, et al. (1996) Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science 272: 1167-1170.

69. Rodriguez B, Sethi AK, Cheruvu VK, Mackay W, Bosch RJ, et al. (2006) Predictive value of plasma HIV RNA level on rate of CD4 T-cell decline in untreated HIV infection. JAMA 296: 1498-1506.

70. Fahey JL, Taylor JM, Detels R, Hofmann B, Melmed R, et al. (1990) The prognostic value of cellular and serologic markers in infection with human immunodeficiency virus type 1. N Engl J Med 322: 166-172.

71. Crowley S, Rollins N, Shaffer N, Guerma T, Vitoria M, et al. (2010) New WHO HIV treatment and prevention guidelines. Lancet 375: 874-875.

72. Freed EO, Orenstein JM, Buckler-White AJ, Martin MA (1994) Single amino acid changes in the human immunodeficiency virus type 1 matrix protein block virus particle production. J Virol 68: 5311-5320.

73. Spearman P, Horton R, Ratner L, Kuli-Zade I (1997) Membrane binding of human immunodeficiency virus type 1 matrix protein in vivo supports a conformational myristyl switch mechanism. J Virol 71: 6582-6592.

74. Ganser-Pornillos BK, von Schwedler UK, Stray KM, Aiken C, Sundquist WI (2004) Assembly properties of the human immunodeficiency virus type 1 CA protein. J Virol 78: 2545-2552.

75. Weiss ER, Gottlinger H (2011) The role of cellular factors in promoting HIV budding. J Mol Biol 410: 525-533.

76. Gottlinger HG, Dorfman T, Sodroski JG, Haseltine WA (1991) Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. Proc Natl Acad Sci U S A 88: 3195-3199.

77. von Schwedler UK, Stemmler TL, Klishko VY, Li S, Albertine KH, et al. (1998) Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. EMBO J 17: 1555-1568.

78. Forshey BM, von Schwedler U, Sundquist WI, Aiken C (2002) Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. J Virol 76: 5667-5677.

79. Dorfman T, Mammano F, Haseltine WA, Gottlinger HG (1994) Role of the matrix protein in the virion association of the human immunodeficiency virus type 1 envelope glycoprotein. J Virol 68: 1689-1696.

80. Freed EO, Martin MA (1996) Domains of the human immunodeficiency virus type 1 matrix and gp41 cytoplasmic tail required for envelope incorporation into virions. J Virol 70: 341-351.

81. Krausslich HG, Facke M, Heuser AM, Konvalinka J, Zentgraf H (1995) The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. J Virol 69: 3407-3419.

82. Wright ER, Schooler JB, Ding HJ, Kieffer C, Fillmore C, et al. (2007) Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. EMBO J 26: 2218-2226.

83. Li F, Goila-Gaur R, Salzwedel K, Kilgore NR, Reddick M, et al. (2003) PA-457: a potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. Proc Natl Acad Sci U S A 100: 13555-13560.

84. Lu W, Salzwedel K, Wang D, Chakravarty S, Freed EO, et al. (2011) A single polymorphism in HIV-1 subtype C SP1 is sufficient to confer natural resistance to the maturation inhibitor bevirimat. Antimicrob Agents Chemother 55: 3324-3329.

85. Lazaryan A, Song W, Lobashevsky E, Tang J, Shrestha S, et al. (2011) The influence of human leukocyte antigen class I alleles and their population frequencies on human immunodeficiency virus type 1 control among African Americans. Hum Immunol 72: 312-318.

86. Matthews PC, Adland E, Listgarten J, Leslie A, Mkhwanazi N, et al. (2011) HLA-A*7401-mediated control of HIV viremia is independent of its linkage disequilibrium with HLA-B*5703. J Immunol 186: 5675-5686.

87. Fledderman EL, Fujii K, Ghanam RH, Waki K, Prevelige PE, et al. (2010) Myristate exposure in the human immunodeficiency virus type 1 matrix protein is modulated by pH. Biochemistry 49: 9551-9562.

88. Zhou W, Parent LJ, Wills JW, Resh MD (1994) Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. J Virol 68: 2556-2569.

89. Navis M, Matas DE, Rachinger A, Koning FA, van Swieten P, et al. (2008) Molecular evolution of human immunodeficiency virus type 1 upon transmission between human leukocyte antigen disparate donor-recipient pairs. PLoS One 3: e2422.

90. Poon A, Chao L (2005) The rate of compensatory mutation in the DNA bacteriophage phiX174. Genetics 170: 989-999.

91. Hunt PW, Martin JN, Sinclair E, Bredt B, Hagos E, et al. (2003) T cell activation is associated with lower CD4+ T cell gains in human immunodeficiency virus-infected patients with sustained viral suppression during antiretroviral therapy. J Infect Dis 187: 1534-1543.

Chapter 3

Replicative capacity of HIV-1 drives early inflammation, T cell activation, and infection of central memory CD4+ T cells

Manuscript in preparation

Jessica L. Prince and Daniel Claiborne performed all *in vitro* replication assays, CD4+ T cell FACS sorting and staining panels (Figure 4-5).

Eileen Scully and Marcus Altfeld performed ELISA and Luminex assays for measuring plasma cytokines (Figure 2).

Daniel Claiborne performed Cox-proportional hazard models, PCA and Network analysis (Supplementary Table 1, Figure 2A and B).

Jessica Prince analyzed all FACS analysis in Flow Jo, and performed data analysis for Figures 3-4.

Gladys Macharia performed multi-color flow cytometric staining of CD8+T cells and cytotoxicity (Fig. 3).

Benton Lawson performed the qPCR for analysis of viral burden in memory T-cell subsets (Figure 5A).

Jessica L. Prince, Daniel Claiborne, and Eric Hunter wrote the manuscript

Abstract

The degree to which transmitted viral characteristics influence early and late stage pathogenesis in HIV-1 infection in the context of a vigorous cellular immune response, and the mechanisms by which this may occur, remain poorly defined. Previously, we showed that viral replicative capacity (vRC), as defined by the *gag* gene, influences set point viral load (SPVL) and CD4+ T cell decline in seroconvertors. Here, we show that the association of vRC with CD4+ T cell decline is independent of SPVL, and that vRC drives an early inflammatory cytokine response that is associated with increased T cell activation, exhaustion, and proliferation, all strong predictors of HIV-1 disease progression in this cohort. Furthermore, vRC is positively linked to the magnitude of viral burden in naïve and central memory CD4+ T cells during early infection. Taken together, these results support an unprecedented role for vRC, a transmitted viral characteristic, in driving many facets of HIV-1 immunopathology.

Introduction

From the start of the epidemic, HIV-1 infection has been characterized by a steady decline in CD4+ T cells resulting in a state of overt immunodeficiency marked by an increased susceptibility to opportunistic infections and malignancies [1,2]. Yet, despite the fact that a majority of HIV-1 infected individuals eventually progress to AIDS, they do so at varying rates [3]. To date, research has focused on identifying host factors that contribute to viral load and disease progression [4]. Although the results of such studies have highlighted the importance of several HLA-class I alleles, it has been estimated that host genetic factors explain merely 22% of the variability in viral load [5].

Viral characteristics, which have been relatively understudied, might further explain variations in disease severity. In support of this, viral loads in epidemiologically linked transmission pairs were found to be correlated between chronically infected donors and acutely infected recipients, implying that transmitted viral characteristics can impact viral control in individuals with disparate immunogenetic backgrounds [6-8]. This is supported by our observation that the number of transmitted cytotoxic T lymphocyte (CTL) induced escape mutations in the structural protein precursor, Gag, but not the accessory protein, Nef, correlated with a reduction in early set point viral load (SPVL), consistent with escape mutations within Gag attenuating viral replication and thereby impacting viral control [9].

In an initial study, we showed that in individuals recently infected with HIV-1 subtype C, low viral replicative capacity (vRC), defined *in vitro* by the transmitted Gag sequence in chimeric viruses, was associated with a delayed loss of CD4+ T cells. Because this effect appeared to be partially independent of SPVL, we hypothesized that high levels of replication of the transmitted virus might initiate irreversible pathogenic events early in infection that dictate the kinetics of subsequent disease progression, before host

factors can significantly control virus replication [10]. Here, in a study of 127 acutely HIV-1 infected individuals, we test this hypothesis and show that infection by high vRC HIV-1, as defined by Gag, is associated with increased levels of inflammatory cytokines, greater cellular immune activation and exhaustion, and higher levels of proviral burden in memory T cell subsets at early times after infection. Thus, the replicative capacity conferred by the *gag* gene of transmitted HIV-1 is a critical factor in defining the initial inflammatory state, the preservation or loss of CD4+ T cell homeostasis, and the trajectory of disease progression. Interventions including early antiretroviral therapy or vaccine-induced immunity that impact these early events may have a significant effect on the development of clinical disease.

Results

Viral characteristics determine HIV-1 pathogenesis

Previous work by our group and others has shown that transmitted viral characteristics significantly correlate with early SPVL as well as CD4+ T cell decline up to three years post infection (10). Here, we sought to determine if the vRC of HIV-1, conferred by the *gag* gene alone, durably affects the trajectory of CD4+ T cell decline even in the chronic stages of infection, and if this effect might be independent of vRC's established association with SPVL.

To assess the replicative capacity conferred by the transmitted *gag* sequence, we amplified the *gag* gene from plasma virus during the acute infection time point (median 45 days post estimated date of infection), generated replication competent virus by cloning the *gag* gene into a common proviral backbone, and assessed vRC in an *in vitro* cell culture assay as described previously [10]. In this cohort of 127 acutely infected individuals from Zambia, low viral RC significantly delayed the time to CD4+ T cell counts <300 for up to 5 years post infection **(Figure 1A)**. The clearest benefit is observed with the lowest vRC tercile compared to the middle and highest tercile. A clear benefit remained even down to CD4 counts of <200, the clinical definition of AIDS, when individuals infected with intermediate and highly replicating viruses were combined into one group **(Figure 1B)**.

We have previously shown an association between early SPVL and vRC [10]. Thus, we sought to more definitively determine if the replication capacity defined by the *gag* gene affected CD4 decline in a manner linked to, or independent of, the well-documented effect of early SPVL on subsequent disease progression. In this cohort, we found SPVLs >10^5 RNA copies/ml to be highly hazardous **(Supplemental Figure 1A)**;

however, vRC significantly dichotomized the trajectory of CD4 decline ($p<0.0001$) in individuals with SPVLs $<10^5$ **(Figure 1C)**, a majority of the cohort. This suggested independent, but additive effects, of both vRC and early SPVL on HIV disease progression. This is confirmed in a multivariable Cox proportional hazards model assessing the relative risk of vRC in the context of other well-established predictors of HIV disease progression. We find that low vRC, early SPVL, and canonical protective HLA class I alleles (B*57, B*5801) were each independent predictors of CD4 decline **(Supplemental Table 1)**. Moreover, vRC significantly dichotomizes disease trajectories of those with (and without) these protective HLA alleles **(Figure 1D)**. Taken together, these data firmly establish vRC as a distinct contributor to HIV disease progression, and suggests that it may modulate events very early after infection that set the trajectory of disease. In order to test this hypothesis, we aimed to determine early levels of circulating inflammatory cytokines, immune activation and exhaustion in T cell compartments, as well as viral burden in memory CD4+ T cells.

Viral RC alters early inflammatory cytokine profiles

Acute HIV infection is characterized by a rapid and robust expression of type I interferons (IFN-I), IFN-I stimulated genes, and inflammatory cytokines [11]. Gut disruption and subsequent microbial translocation has also been shown to contribute significantly to this inflammatory state, possibly through a positive feedback loop [12]. This inflammatory response, particularly in chronic infection, contributes to disease progression [13-16]. Therefore, we analyzed the levels of 16 inflammatory cytokines, chemokines, and markers of gut damage and microbial translocation at or before seroconversion to assess the effect of vRC on the early inflammatory milieu (n = 33; previously dichotomized into low and high RC phenotypes (10)). We found that vRC

was positively correlated with a number of inflammatory cytokines **(Supplemental Table 2)**, most notably IL-6 and IL-1B, two inflammatory cytokines previously implicated in driving aberrant CD4+ T cell turnover and impairing homeostatic proliferation [17].

In an effort to define distinct inflammatory "profiles", we employed an unsupervised data reduction tool, Principal Component Analysis (PCA), which groups linear variables into combinations, principal components (PCs) **(Figure 2A)**. Strikingly, principal component 1 (PC1), which describes the greatest variation in the data set, significantly correlates with viral RC, in that individuals with positive loadings for PC1 (elevated levels of inflammation) tend to have higher vRC, while those with negative loadings are significantly enriched for poorly replicating viruses **(Supplemental Figure 2)**. Moreover, principal component 2 (PC2), which by definition is unrelated to PC1, describes the second greatest variation in the data and is significantly correlated with SPVL **(Supplemental Figure 2)**. This further highlights the independence of initial viral replication and subsequent adaptive immune control of VL set point.

Though this analysis is unsupervised, the first six principal components significantly describe biological phenotypes (such as gender, protective HLA alleles, and age) known to influence HIV pathogenesis (data not shown). The inflammatory cytokines that contribute to latent variables 1 and 2 are depicted in **Figure 2B**. Thus viruses with high vRC are correlated with a distinct inflammatory cytokine profile characterized by a heightened interferon response and elevated levels of key inflammatory cytokines such as IL-6 and IL-1B.

Replicative capacity is associated with levels of CD8+T cell activation and exhaustion

Chronic immune activation is a hallmark of HIV-1 infection; it often persists

following ART and is a more reliable predictor of disease progression than viral load [18-22]. Therefore, we assessed the impact of vRC on levels of cellular immune activation by measuring the co-expression of CD38 and HLA-DR on CD8+ T cells isolated within three months post-infection for a subset of individuals under study (n=36). We found that vRC is positively correlated with the expression of CD38 and HLA-DR on CD8+ T cells (**Figure 3A**). Consistent with previous studies [18,20,22], we observe that higher CD8+ T cell activation in this cohort is associated with faster CD4+ T cell decline and that this increased risk is independent of SPVL (**Supplementary Figure 3A**), thus positioning T cell activation as a likely mechanistic link between vRC and subsequent CD4+ T cell decline.

In addition to cellular immune activation, CD8+ T cell exhaustion is characteristic of pathogenic HIV/SIV infection. Exhaustion of CD8+ T cells is marked by the increased expression of the inhibitory receptor Programmed Death 1 (PD-1), and levels of PD-1 expression predict the rate of disease progression [23,24]. PD-1^{hi} CD8+ T cells are typically $CD57^{low}$ [23,24] while, in contrast, Petrovas et al. showed that PD-$1^{low}CD57^{high}$ CD8+T cells are more resistant to apoptosis [25]. We find that a greater percentage of CD8+ T cells isolated from individuals infected with low RC viruses displayed high levels of CD57 while maintaining low levels of PD-1 (**Figure 3B**) relative to individuals infected by high RC viruses. In CD8+ T cells, markers of exhaustion are often associated with impaired cytotoxic function [26]. Indeed, dual expression of granzyme B and perforin was positively correlated with the frequency of PD-$1^{-}CD57^{+}$ CD8+ T cells (**Figure 3C**) and was inversely correlated with vRC (**Figure 3D**). Consequently, CD57 expression is protective in terms of CD4+ T cell decline while PD-1 expression is associated with faster disease progression in this cohort (**Supplementary Figure 3B and 3C**). Taken together, these data suggest that individuals infected with attenuated viruses mount a more functional and less exhausted CTL response early in

infection that may be more effective in controlling viral replication, thus, providing an extended clinical benefit for the individual.

Replicative capacity predicts levels of CD4+ T cell activation, proliferation, and exhaustion

Immune activation in HIV-1 infection is associated with global immunological dysfunction, characterized by increased cellular-turnover, and ultimately the disruption of critical CD4+ T cell homeostasis, indicative of progressive disease [27-29]. The percentage of CD4+ T cells expressing CD38/HLA-DR and Ki67 were measured in PBMCs isolated less than 3-months post infection for a subset of individuals (n=19; see methods for sample selection). Expression of these markers was assessed on total CD4+ T cells as well as in the context of different CD4+ T cell subsets: naïve (T_N, CD27+, CD45RO-, CCR7+), central memory (T_{CM}, CD27+/-, CD45RO+, CCR7+), and effector memory (T_{EM}, CD27 +/-, CD45RO+, CCR7-).

High vRC was associated with increased expression of CD38 and HLA-DR on total (**Supplementary Figure 4A,** P = 0.02) and T_{EM} (**Figure 4A,** P = 0.02) CD4+ T cells, however this association was most pronounced in the T_{CM} (**Figure 4B,** P = 0.006) compartment. Consistent with previous reports [18,20], high levels of CD4+ T cell activation (CD38+/HLA-DR+) are also associated with faster CD4+ T cell decline (**Figure 4C**). Furthermore, individuals infected with high RC viruses exhibited significantly greater levels of CD4+ T cell proliferation as measured by Ki67 expression. This association was most striking in the effector memory T cell subset (**Figure 4D,** P = 0.003), but was also significant for total (**Supplementary figure 4B,** P = 0.006) and central memory CD4 T cells (**Figure 4E,** P = 0.008). Increased CD4+ T cell turnover, demonstrated by elevated expression of Ki-67 on CD4+ T cell subsets was also associated

with faster CD4+T cell decline, as has been reported previously [28,30]. **(Figure 4F)**. Furthermore, individuals infected with poorly replicating viruses displayed expression levels of CD38/HLA-DR and Ki67 on their CD4+ T cells, which, although somewhat elevated, for the most part did not differ significantly when compared to HIV-uninfected Zambians **(Figure 4 A-D),** suggesting that shortly after transmission, individuals infected with viruses exhibiting low vRC preserve an immune system akin to that of a healthy individual.

Early inflammatory cytokine profiles associated with vRC are linked to activated T cell phenotype

To determine whether changes in T cell activation are associated with the distinct inflammatory cytokine profiles associated with vRC, we compared individuals with positive (increased inflammatory cytokine levels) and negative loadings for Principal Component 1 (PC1; **shown in Figure 2**). Individuals with positive loadings presented with significantly increased CD8+ T cell activation (**Figure 5A**) and exhaustion (**Figure 5B and 5C**). Positive loadings were further associated with increased frequency of $PD1^{+}CD4+$ T_{CM} cells, as well as higher levels of CD4+ T_{EM} cell proliferation, as measured by Ki67 expression (**Figure 5D**). Thus, higher levels of inflammatory cytokines are closely linked to the activated, exhausted T cell phenotype observed in individuals infected with high vRC viruses.

High vRC is associated with increased viral burden in CD4+ T cell subsets

Establishment of the latent HIV reservoir occurs early during acute infection and sets the stage for viral persistence [31-34]. Thus, even in the context of suppressive

antiretroviral therapy, HIV-1 cannot be fully eradicated. To determine if the replicative capacity of the transmitted virus influences the proviral DNA burden during early infection, levels of cell-associated viral DNA were measured in naïve, central memory, and effector memory CD4+ T cell subsets sorted from PBMCs isolated 3 months post-infection. We observed that high vRC was associated with a significant increase in the amount of cell-associated viral DNA in both T_{CM} ($P = 0.01$) and T_N ($P = 0.001$) CD4+ subsets (**Figure 6A**). Because high vRC was associated with increased levels of immune activation and proliferation, we determined if there was a direct association between these markers and the magnitude of viral burden in T_{CM}, a key population in CD4+ T cell homeostasis. We find that HIV-1 DNA in T_{CM} positively correlates with expression of CD38+/HLA-DR+ on T_{EM} (**Figure 6B**, $P<0.0001$) and, to a lesser extent, on T_{CM} ($R^2 = 0.3$, $P = 0.01$, data not shown). Similarly, HIV-1 DNA in T_{CM} was positively correlated with Ki67 expression on CD4+ T_{EM} cells (**Figure 6C**), indicating that greater infection of CD4+ T_{CM} may lead to higher activation and proliferation of the CD4+ T_{EM} population. Moreover, higher levels of cell associated HIV DNA in central memory CD4+ T cells was associated with an accelerated loss of CD4+ T cell counts <300 (**Figure 6D**).

Discussion

Untreated HIV-1 infection is characterized by a progressive depletion of CD4+ T cells and development of a fatal state of immunodeficiency, a disease course with kinetics that vary among individuals and between different HIV-1 subtypes [3,35]. Although numerous studies have implicated host immunogenetic factors in governing the tempo of disease progression [36], viral characteristics have received less scrutiny. Recent studies have implicated transmitted viral characteristics in explaining the heritability of HIV-1 pathogenesis [5]. In support of this, we previously showed that the replicative capacity of the transmitted virus, as conferred by the transmitted *gag* sequence, is correlated both with chronic donor viral loads near the time of transmission and with SPVL in recently infected recipients, establishing vRC as a heritable viral trait impacting HIV-1 disease [10].

Here we have highlighted an integral role for replicative capacity of transmitted virus, as defined by Gag, in driving multiple facets of HIV-1 immunopathology. In a cohort of Zambian seroconvertors acutely infected with HIV-1 subtype C we demonstrate that vRC is (I) an independent contributor to CD4+ T cell decline outside of host immunogenetic factors that dictate SPVL, (II) linked to an inflammatory state early in infection that is characterized by elevated levels of key inflammatory cytokines known to drive pathogenesis, (III) linked to aberrant CD8+ and CD4+ T cell phenotypes characterized by increased levels of cellular activation, exhaustion, and proliferation, all of which correlated with CD4+ T cell decline within this cohort and (IV) associated with increased viral burden in naïve CD4+ T cells and CD4+ T_{CM}, the latter being integral in establishing the latent viral reservoir [32] and preferentially spared in non-pathogenic SIV infection of natural hosts [37]. Taken together, these results support an unprecedented role for transmitted viral characteristics, specifically viral replicative

capacity, in determining the early inflammatory state of HIV-1 infected individuals and their subsequent disease trajectory.

Previous studies show that the level of immune activation is established very early after infection and varies greatly between individuals, but remains quite stable over time [18]. Furthermore, immune activation predicts disease progression even when viral loads are suppressed either immunologically or by ART [18,20-22]. For the first time, we demonstrate that levels of immune activation established during early infection can be significantly attributed to the replicative capacity of the transmitted virus, even though we are only measuring the contribution of the *gag* gene to this process.

We furthermore demonstrated that the amount of HIV DNA harbored by both T_{CM} and CD27+/CD45RO- CD4+ T cells is highly associated with viral replicative capacity. It is possible that the increased levels of HIV-1 DNA found in these "naïve" CD4+ T cells could be due in part to the presence of T memory stem cells (T_{SCM}), that would be found within our sorted naïve T cell populations [38,39]. These long-lived memory T cells with stem cell-like properties are susceptible to HIV-1 or SIV in pathogenic infections but spared in non-pathogenic SIV disease [40].

T_{CM} CD4+ cells, which are characterized as being more stable and longer-lived, are preferentially spared in non-pathogenic SIV infection and infection of this subset has been linked to the immunopathology of HIV-1 infection [28,37,41]. The results of the current study provide further evidence that HIV-1 infection of this subset occurs early, is linked to altered CD4 T cell homeostasis, and significantly predicts disease progression.

T_{CM} CD4+ cell have also been highlighted as an integral population for the maintenance of latency and viral persistence [32]. Our findings suggest that the extent to which HIV infects T_{CM} CD4+ T cells and establishes a latent reservoir might be influenced by characteristics of the transmitted founder virus. This has important

therapeutic implications, and suggests that individuals infected with attenuated (low vRC) viruses may be better candidates for proof-of-concept cure strategies.

The data presented here suggests that viral characteristics can provide independent information about an individual's risk for disease progression and point to a novel target for interventions to reduce immune activation and viral burden. Immune responses and interventions that effectively attenuate vRC in the earliest stages of infection could not only have a dramatic impact on viral control and disease progression, but, based on our recent findings (Carlson et al – Science, in press), could also impact the efficiency of subsequent transmission to other partners. Moreover, reducing the size of the viral burden in key T cell populations prior to ARV treatment could significantly augment cure strategies aimed at eliminating the latent reservoir.

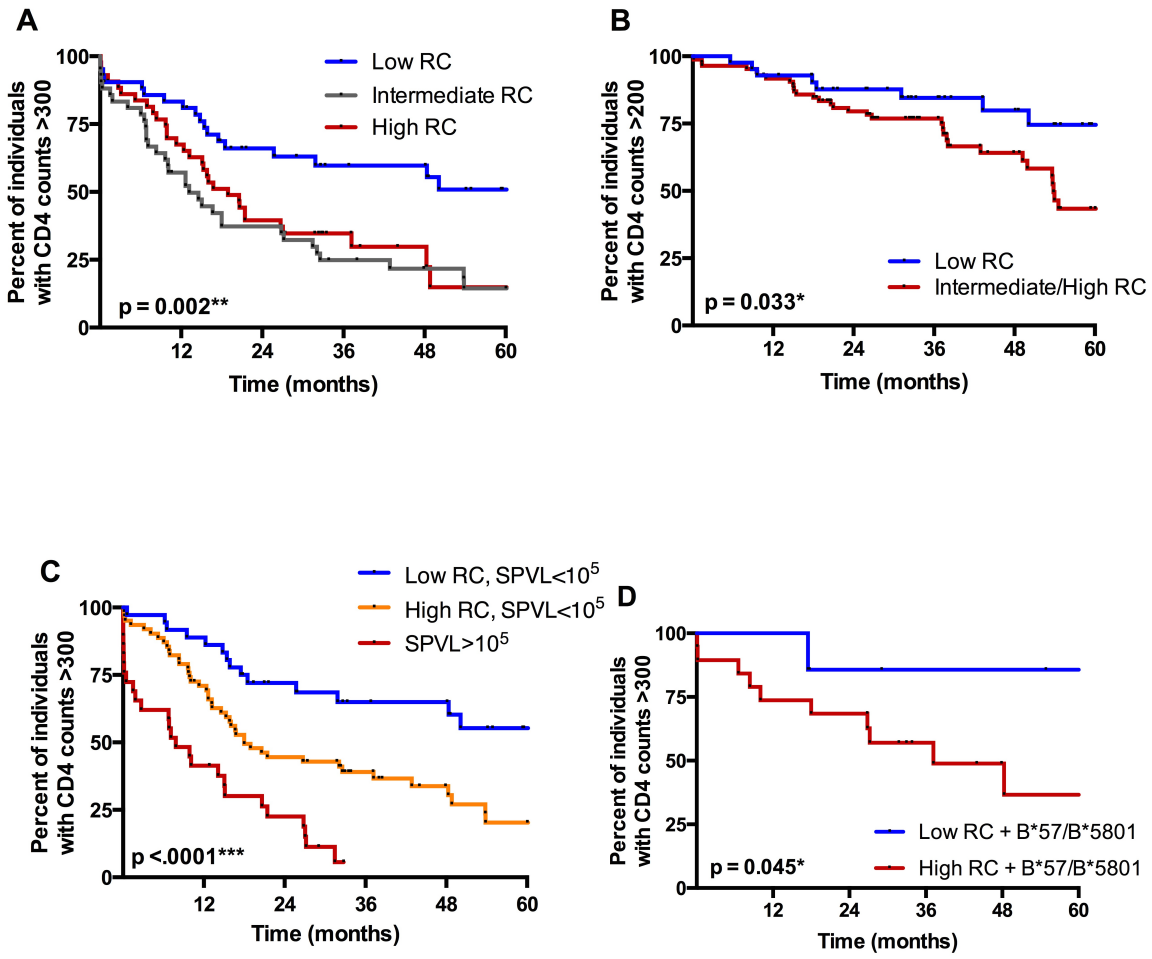
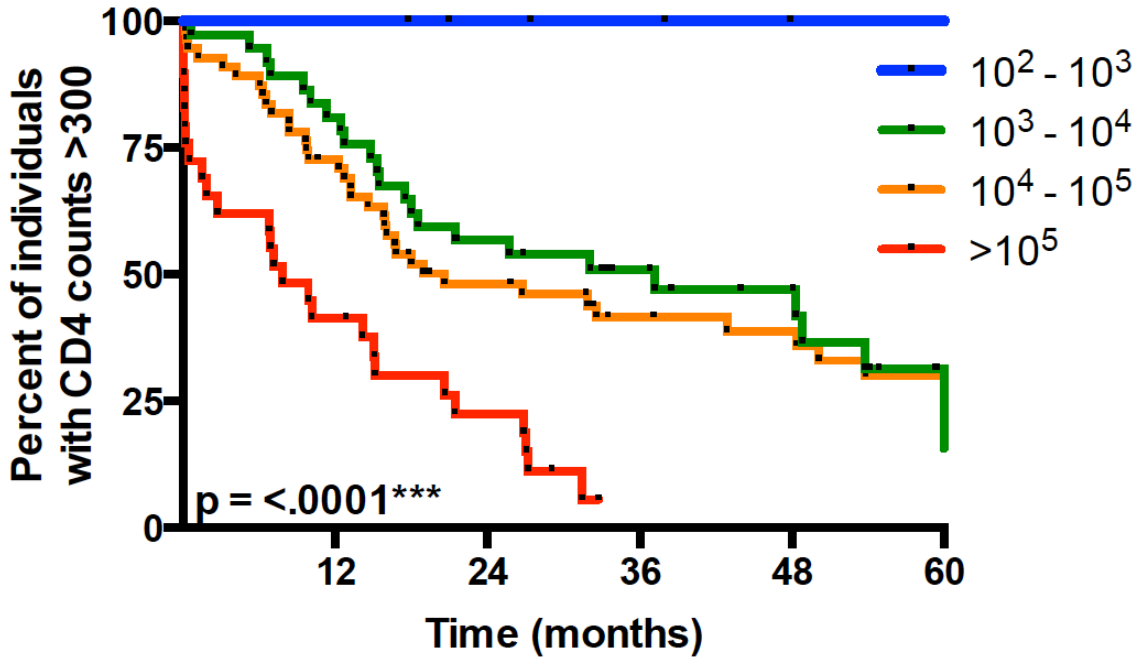


Figure 1 | HIV-1 replicative capacity, when defined by the transmitted *gag* sequence, predicts CD4 T cell decline in ART-naïve, HIV-1 infected individuals. A. In a Kaplan-Meier survival analysis with an endpoint defined as CD4+ T cells counts falling below 300, low vRC, defined as the lowest tercile of vRC values in 127 acutely infected individuals, provides a significant protective benefit from CD4+ T cell decline. **B.** When the middle and upper terciles of vRC groups are combined, low vRC provides significant protection from CD4+ T cell counts falling below 200. **C.** vRC has an independent, but additive, effect with early set point VL in determining CD4+ decline. High vRC is defined as the upper 67% of vRC scores, and low vRC is defined as

the lowest 33% of vRC scores. **D.** In individuals with protective HLA class I alleles (n = 26), vRC significantly dichotomizes CD4 decline trajectory.



Supplementary Figure 1 | The effect of $\log_{10}$-increases in early set point VL on longitudinal CD4 T cell decline post seroconversion. In a Kaplan-Meier survival analysis with an endpoint defined as reaching CD4+ T cell counts <300, individuals with set point VLs greater than 10^5 were at the highest risk of rapid CD4+ T cell decline (n = 127). Groups were defined as the $\log_{10}$-transformed set point VL measurement rounded down to the nearest power.

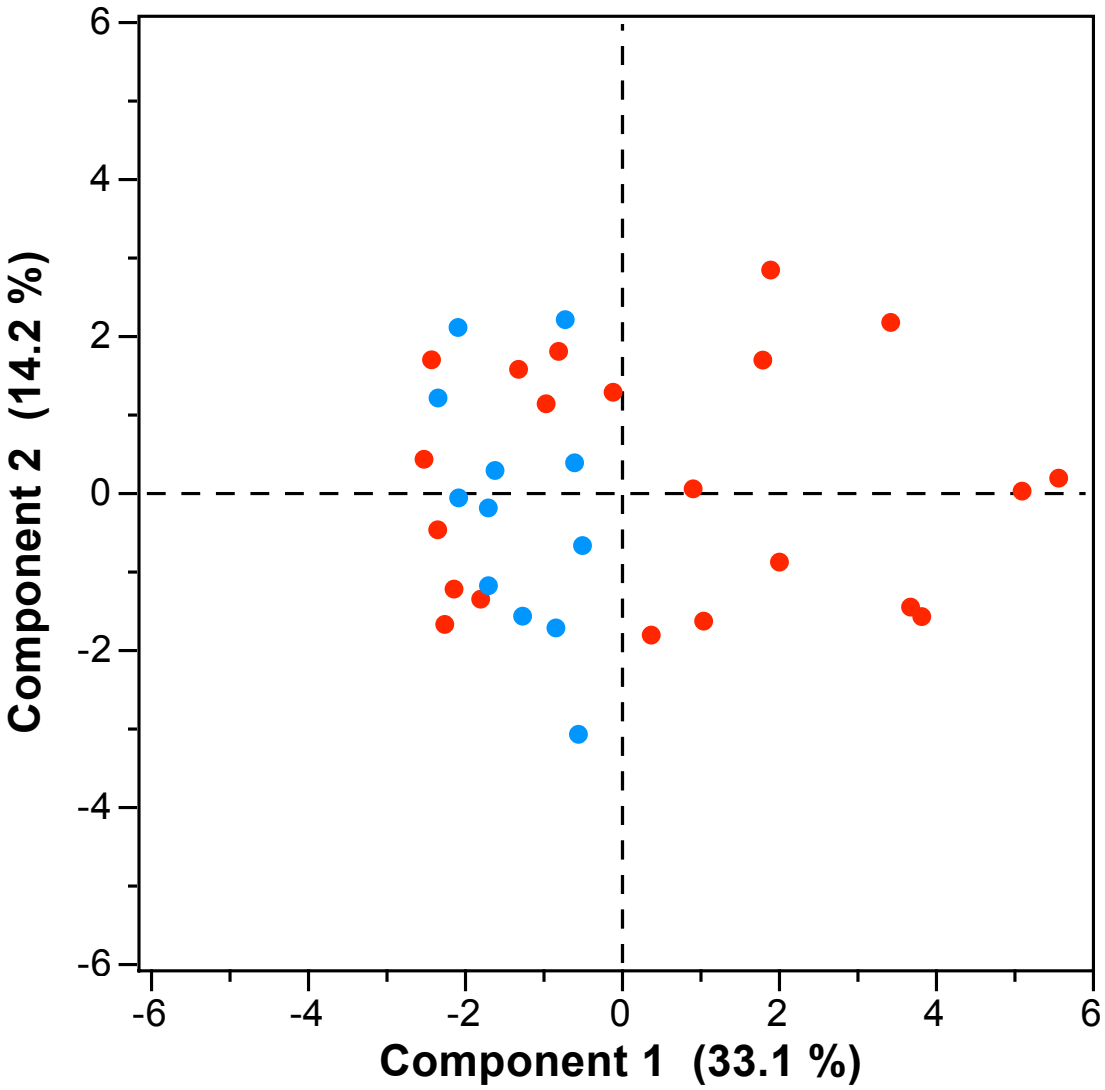
	Cox Proportional Hazards Model (Time to CD4<300)		
Factors Tested	**HR**	**95% CI**	***P*-value**
Female	1.10	0.67 – 1.70	0.78
Low vRC (lowest tercile)	0.48	0.28 – .80	0.004
B*57/5801	0.45	0.23 – .81	0.006
Set point VL	10.00	2.86 – 44.14	0.0004

Supplementary Table 1 | Host and viral characteristics independently predict CD4+ T cell decline. In a multivariable Cox proportional hazards model with an endpoint defined as CD4+ T cell counts <300, low vRC, carriage of HLA-B*57 or B*5801, and set point VL were independent predictors of CD4 decline (n = 127). Transmission of a virus exhibiting low vRC (defined as the lowest tercile of vRC scores) and carriage of protective HLA-I alleles were found to be independently and similarly protective, while high set point VLs were hazardous.

Cytokine/Analyte	Mean (Low RC)	Mean (High RC)	Two-tailed p-value	One-tailed p-value
IL-10 (0M)	5.5	10.73	0.0076	0.0038
IL-6 (0M)	1.88	3.94	0.0086	0.0043
IL-1B (0M)	0.21	0.57	0.0158	0.0079
IFNg (0M)	4.38	10.26	0.0286	0.0143
IP-10 (0M)	639.56	1108.43	0.0352	0.0176
TNFa (0M)	10.76	13.87	0.0566	0.0283
IL-7 (0M)	1.81	2.65	0.0922	0.0461
IFNa2 (0M)	21.09	31.48	0.0968	0.0484

Supplementary Table 2 | High vRC significantly increases early inflammatory cytokine levels. In an analysis of 16 inflammatory cytokines, chemokines, and markers of gut integrity and microbial translocation in plasma at the time of seroconversion (median 46 days post estimated date of infection), 8 analytes were found to be significantly increased in individuals infected with highly replicating viruses (n = 33; statistical comparisons were made using the Student's t-test). When correcting for multiple comparisons using the Storey method, IL-10, IL-6, IL-1B, and IFNg remained significantly elevated in individuals infected with highly replicating viruses.

A



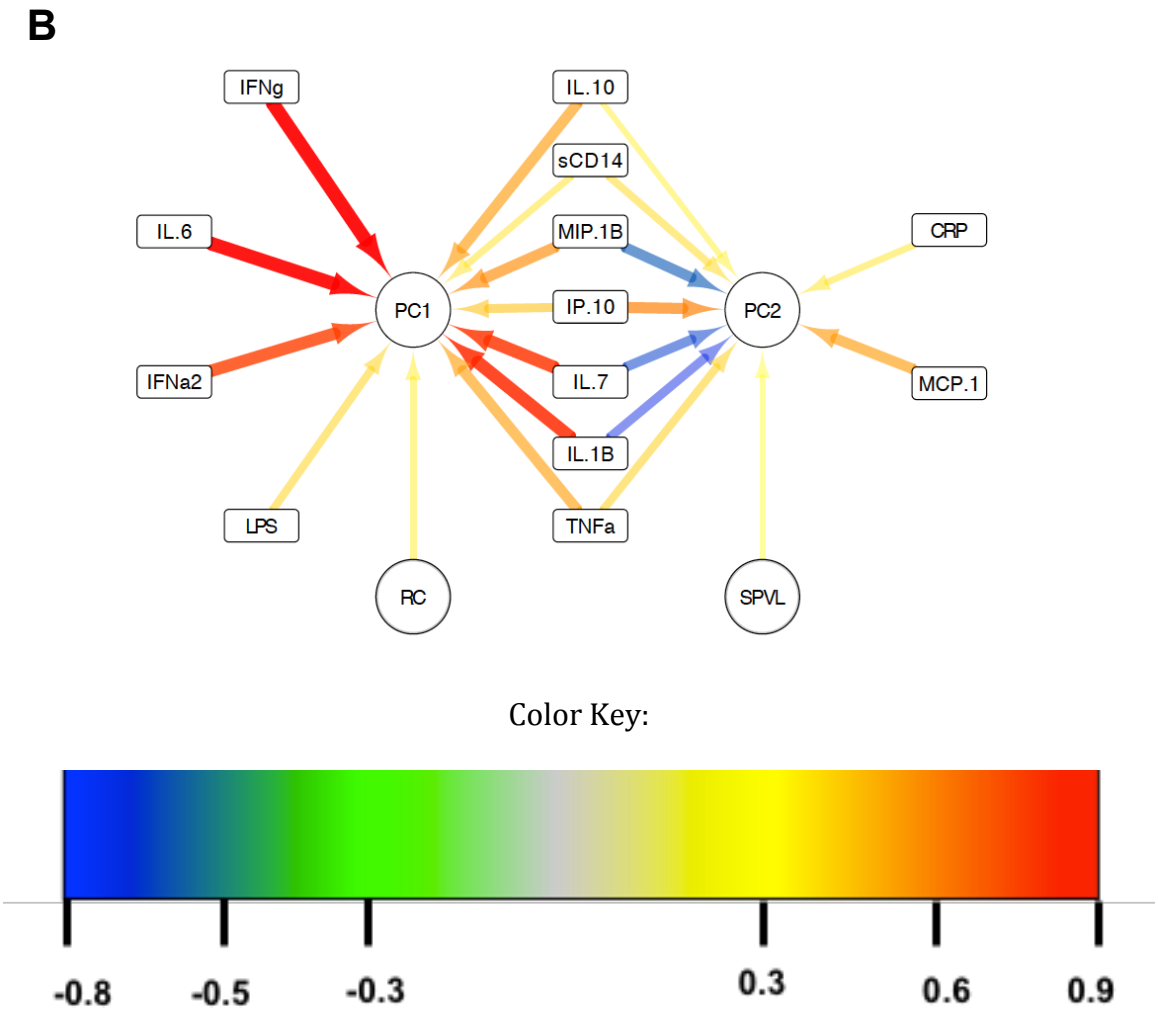
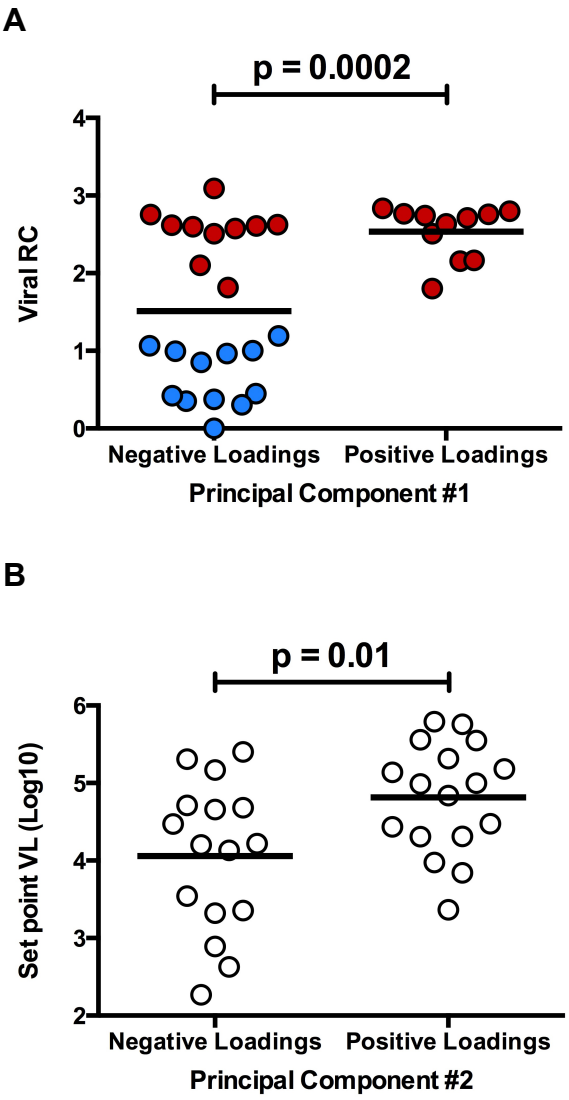


Figure 2 | Low vRC is associated with a distinct cytokine profile early in infection, characterized by muted inflammatory cytokine levels. A. Principal Component Analysis (PCA) was used to reduce dimensionality and extract latent variables comprised of linear combinations of 16 different analytes measured at the seroconversion time point (n = 33). The loadings of the first two principal components are depicted in a 2-dimensional scatter plot. The first principal component, which describes the greatest variation in the data set, has positive loadings associated with high vRC. **B.** A relational network depicts the analytes most positively (red) and negatively (blue) associated with the first two principal components, which themselves are also

correlated with two distinct biological phenotypes, vRC (PC1) and set point VL (PC2). The color key describes how the degree of linear dependence, based on the Pearson correlation coefficient (r), corresponds to the color of connections (the threshold for visible connections was set to a Pearson correlation coefficient of $r > 0.4$ or < -0.4).



Supplementary Figure 2 | The first two principal components are significantly correlated with vRC and set point VL, respectively. **A.** Positive loadings for PC#1, which are associated with high levels of inflammatory cytokines, are significantly enriched for individuals with high vRC. **B.** Positive loadings for PC#2, which are associated with high levels of IP-10 and MCP-1, are significantly enriched for individuals with higher set point VLs (n = 33; statistical comparisons were made using the Student's t-test).

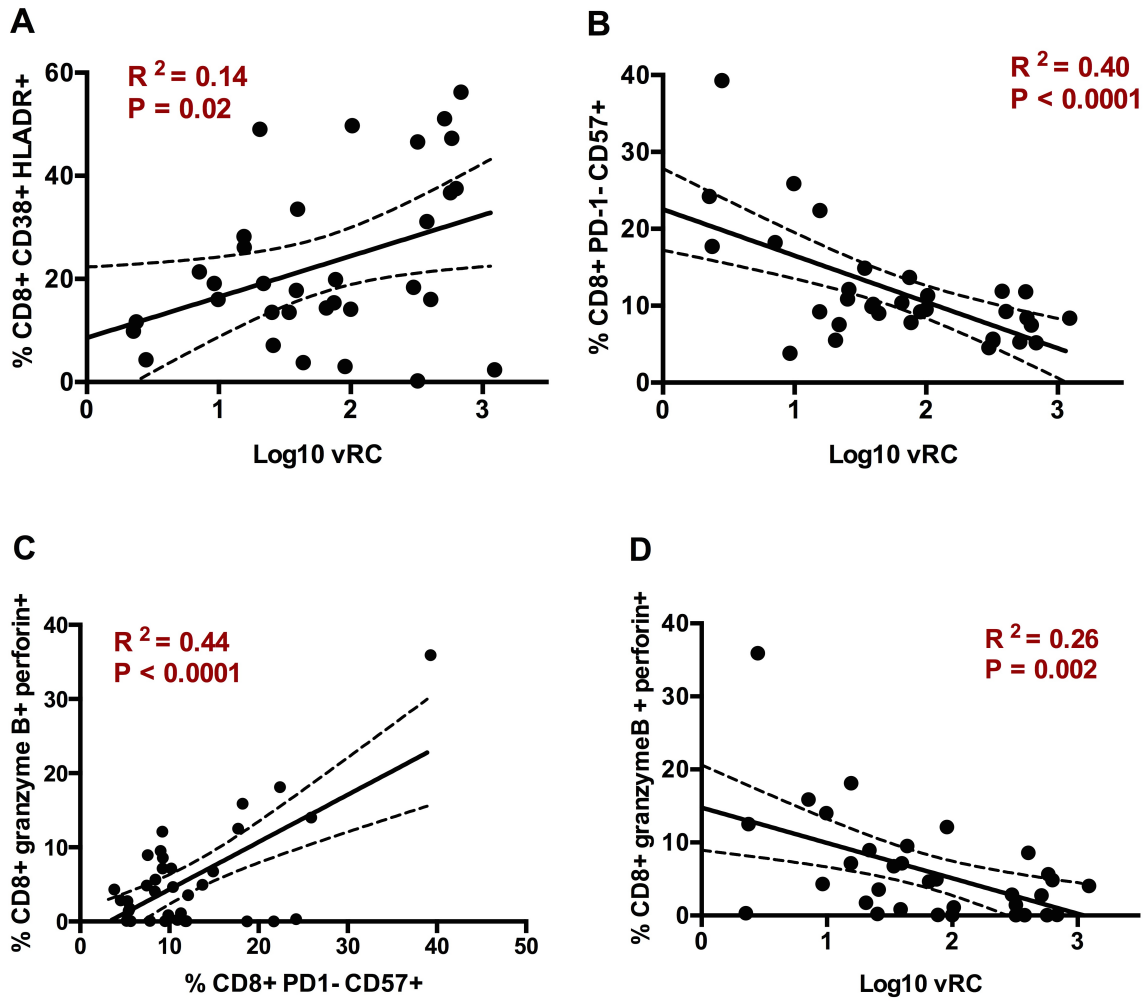
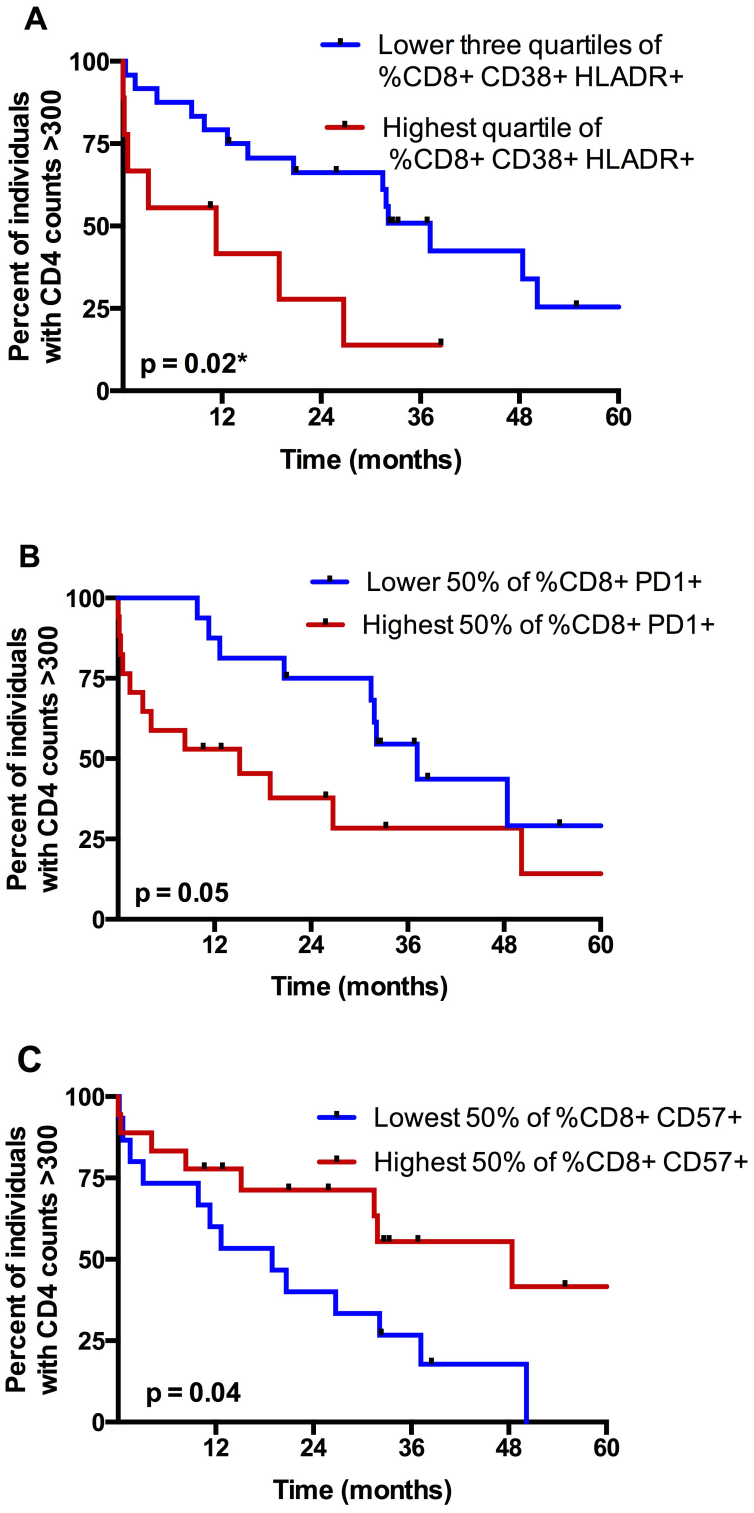


Figure 3 | High vRC is associated with increased CD8 T cell activation and lower cytotoxic potential. A-D. Cryopreserved PBMCs were collected less than 3-months post estimated date of infection (n = 35) and stained with two multi-color flow cytometry panels evaluating either memory subsets and markers of cellular activation and exhaustion (a, b) or markers of T cell cytotoxicity (c, d). **A.** Correlation between vRC and the percentage of total CD8+ T cells co-expressing CD38 and HLA-DR. **B.** Correlation between vRC and the percentage of CD8 T cells that are $CD57^{+}$ but remain $PD\text{-}1^{-}$. **C.** Individuals with a greater percentage of $CD57^{+}/PD\text{-}1^{-}$ CD8+ T cells also have CD8+ T cells with greater cytotoxic potential, as measured by the percentage of cells expressing both granzyme B and perforin. **D.** Correlation between cytotoxic potential in

CD8+ T cells, as measured by co-expression of granzyme B and perforin, and vRC. Correlation statistics were generated using linear regression. Solid lines indicate trend lines, and dashed lines represent 95% confidence bands.



Supplementary Figure 3 | CD8+ T cell activation phenotypes early after infection are associated with CD4 T cell decline. **A-C.** Kaplan-Meier survival analysis with an

endpoint defined as CD4+ T cells counts <300 and statistics generated from the log-rank test (*n*=35). **A.** The highest quartile of CD38 and HLA-DR co-expression on CD8+ T cells is significantly correlated with accelerated CD4+ T cell decline. **B.** High levels of PD-1 expression on CD8+ T cells show a trend for accelerated CD4+ T cell decline. **C.** Individuals with a high percentage of $CD57^+$ CD8+ T cells exhibit delayed CD4+ T cell decline.

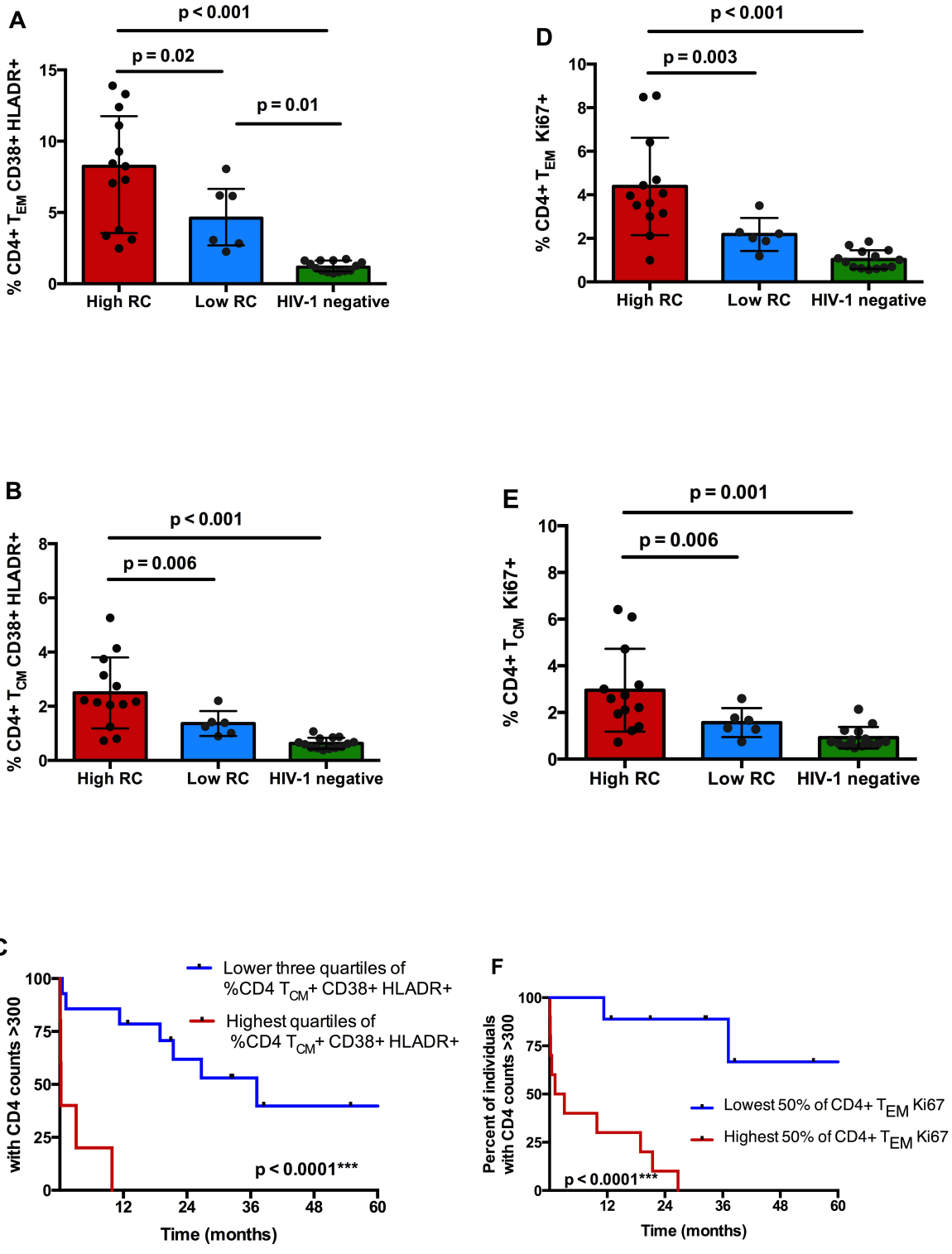
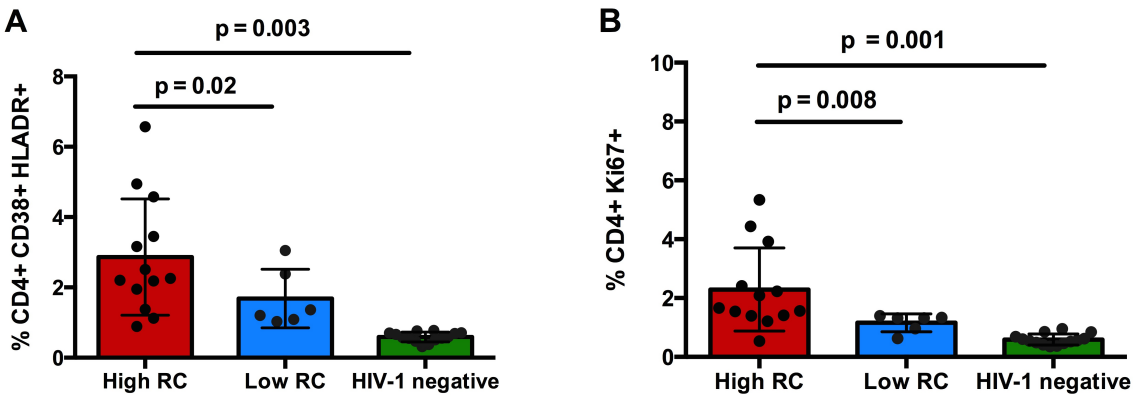

A
p < 0.001
p = 0.02
p = 0.01
% CD4+ T_EM CD38+ HLADR+
High RC
Low RC
HIV-1 negative
D
p < 0.001
p = 0.003
% CD4+ T_EM Ki67+
B
p < 0.001
p = 0.006
% CD4+ T_CM CD38+ HLADR+
E
p = 0.001
p = 0.006
% CD4+ T_CM Ki67+
C
Lower three quartiles of %CD4 T_CM+ CD38+ HLADR+
Highest quartiles of %CD4 T_CM+ CD38+ HLADR+
Percent of individuals with CD4 counts >300
p < 0.0001***
Time (months)
F
Lowest 50% of CD4+ T_EM Ki67
Highest 50% of CD4+ T_EM Ki67
p < 0.0001***

Figure 4 | vRC is associated with increased cellular activation and proliferation in CD4 T cell memory subsets. A-D. Cryopreserved PBMCs from individuals with high vRC (n = 13), low vRC (n = 6), and from HIV-negative Zambian individuals (n = 14) isolated at 3-months post estimated date of infection were stained for markers of activation, cellular turnover, and delineation of memory T cell subsets. **A-B.** Co-expression of the activation markers CD38 and HLA-DR on CD4+ T_{EM} and T_{CM} cells was significantly increased in individuals infected with highly replicating viruses as compared to individuals infected with attenuated viruses or HIV-negative individuals. The percentage of CD4+ T_{CM} cells co-expressing CD38 and HLA-DR did not differ significantly between individuals infected with poorly replicating viruses and those that were HIV-negative. **D-E.** CD4+ T cell proliferation, as measured by Ki67 expression, was significantly higher in T_{CM} and T_{EM} of individuals infected with highly replicating viruses. There was no statistically significant difference in the percentage of T_{CM} and T_{EM} expressing Ki67 between individuals infected with attenuated viruses and HIV-negative individuals (statistical comparisons were made using the Student's t-test). **C,F** In a Kaplan-Meier survival analysis with an endpoint defined as CD4 T cell counts falling below 300, high percentages of CD4+ T_{CM} cells co-expressing CD38 and HLA-DR or CD4+ T_{EM} expressing Ki67 was highly deleterious and lead to rapid disease progression.



Supplementary Figure 4 | High vRC is associated with an increased level of activation and turnover in CD4+ T cells that is highly deleterious. **A.** High vRC is associated with an increased percentage of total CD4+ T cells that co-express CD38 and HLA-DR as compared to individuals infected with attenuated viruses or to HIV-negative individuals. Individuals infected with attenuated viruses exhibit CD4 activation phenotypes that more closely resemble that of uninfected individuals. **B.** High vRC is associated with increased cellular turnover of total CD4+ T cells as evidenced by intracellular Ki67 expression. Individuals infected with attenuated viruses more closely resemble uninfected individuals in terms of the percentage of total CD4+ T cells expressing Ki67.

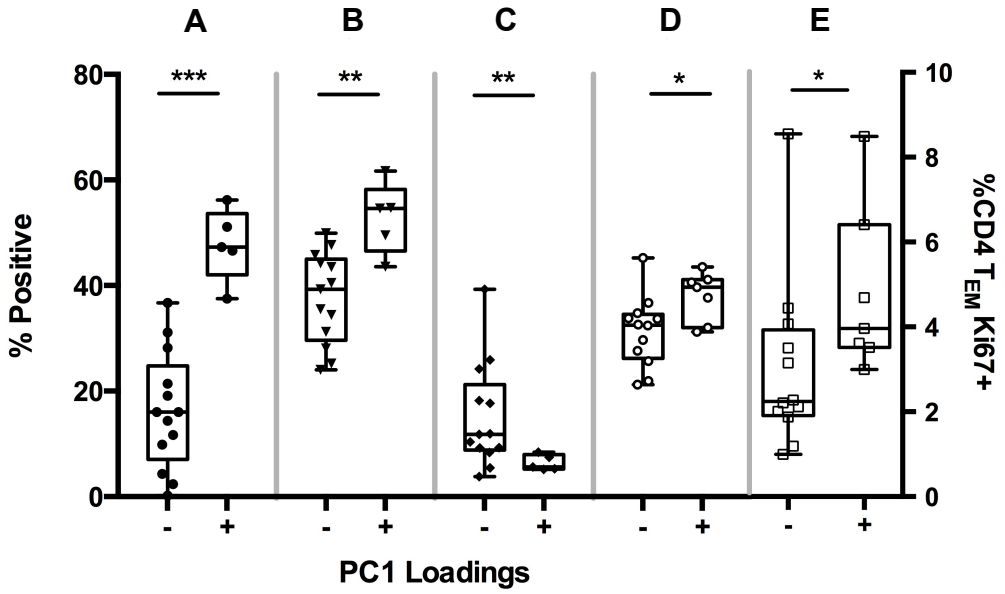


Figure 5 | Inflammatory cytokine profiles associated with vRC correlate with T cell activation. Inflammatory cytokines measured at seroconversion in 33 acutely infected individuals were used to define distinct cytokine profiles via Principal Component Analysis. Of this group, a subset of 18 and 19 individuals had also been immunophenotyped for CD8+ and CD4+ T cell activation, respectively. Positive loadings were defined as a PC1 score >0 and negative loadings were defined as a PC1 score <0. **A-C.** Positive loadings for PC1 are associated with increased CD8+ T cell activation (CD38+/HLA-DR+) and exhaustion (PD-1+, CD57-). **D-E.** Positive loadings for PC1 were also associated with increased levels of PD-1 on CD4+ T_{CM} and increased CD4+ T_{EM} proliferation (Ki67+). The left hand y-axis displays percent positive cells for panels A-D; the right hand y-axis displays percent Ki67+ cells for panel E. (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; statistical comparisons were made using the Student's t-test).

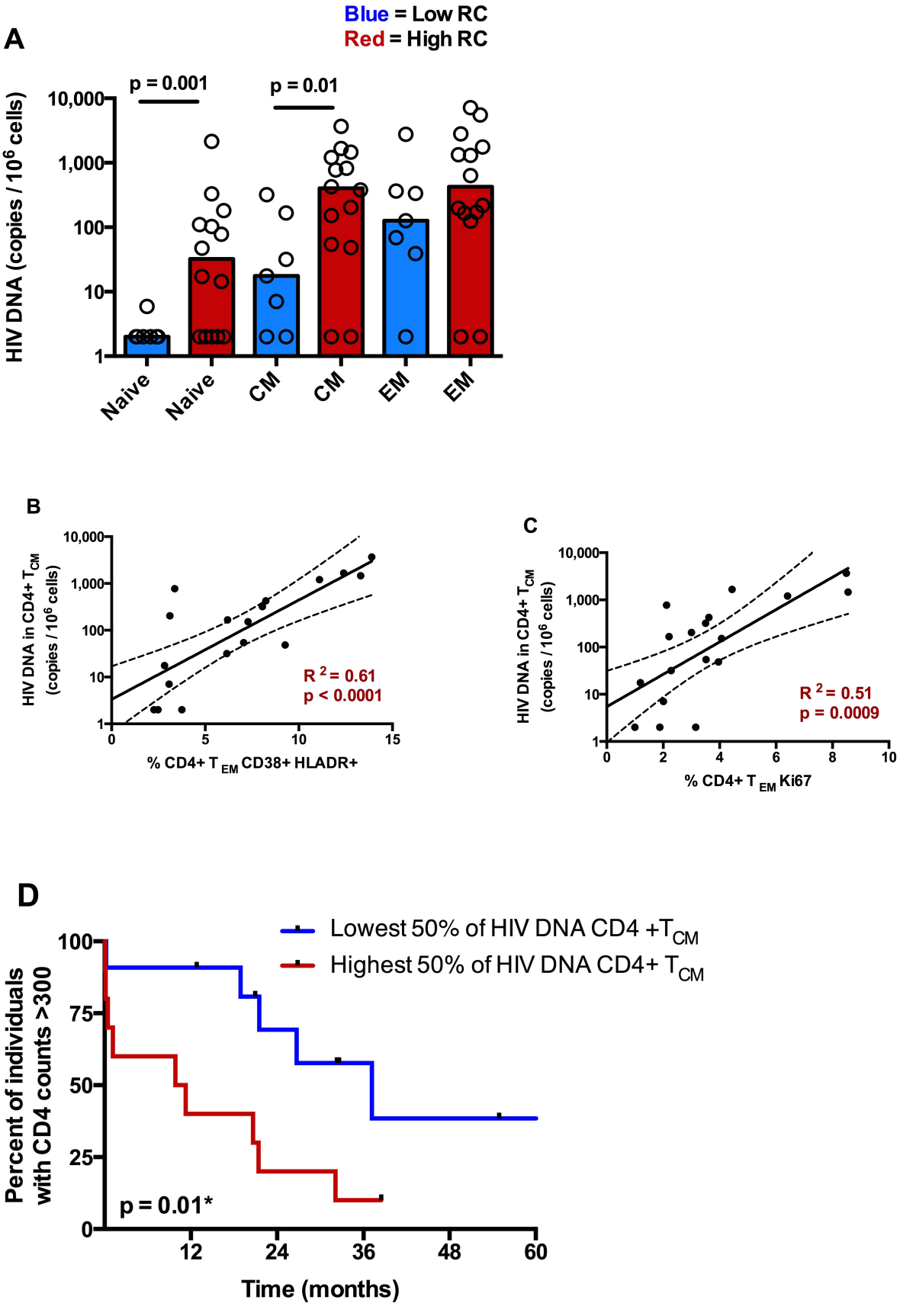

A
Blue = Low RC
Red = High RC
p = 0.001
p = 0.01
HIV DNA (copies / 10^6 cells)
10,000
1,000
100
10
1
Naive
Naive
CM
CM
EM
EM
B
HIV DNA in CD4+ T_CM (copies / 10^6 cells)
R^2 = 0.61
p < 0.0001
% CD4+ T_EM CD38+ HLADR+
C
HIV DNA in CD4+ T_CM (copies / 10^6 cells)
R^2 = 0.51
p = 0.0009
% CD4+ T_EM Ki67
D
Lowest 50% of HIV DNA CD4 +T_CM
Highest 50% of HIV DNA CD4+ T_CM
Percent of individuals with CD4 counts >300
p = 0.01*
Time (months)

Figure 6| vRC correlates with the burden of HIV-1 viral DNA in CD4+ T_{CM} and T_N. A. Cryopreserved PBMCs isolated 3-months post infection (*n* = 21) were stained and CD4+ T cells were sorted into 3 distinct populations, T_N ($CD27^+$, $CD45RO^-$, $CCR7^+$), T_{EM} ($CD27^{+/-}$, $CD45RO^+$, $CCR7^-$), and T_{CM} ($CD27^{+/-}$, $CD45RO^+$, $CCR7^+$), and cell-associated HIV-1 DNA was quantified by real-time PCR amplification of HIV-1 subtype C *integrase* gene. Cell copy number was normalized based on real-time PCR amplification of the human *albumin* gene. Low vRC was associated with significantly fewer CD4+ T_N and T_{CM} cells harboring HIV-1 DNA. **B-C.** In a majority of the same individuals (*n* = 18), a fraction of unsorted PBMCs was stained for markers of cellular activation and proliferation. The viral burden in CD4+ T_{CM} was positively correlated with levels of proliferation (C) and activation (B) in CD4+ T_{EM} cells. Correlation statistics were generated using linear regression and $\log_{10}$-transformed viral DNA copies/10^6 cells. **D.** In a Kaplan-Meier survival analysis with an endpoint defined as CD4 T cell counts falling below 300, high viral burden in CD4+ T_{CM} was associated with an increased risk of CD4 T cell decline.

Literature Cited

1. Masur H, Ognibene FP, Yarchoan R, Shelhamer JH, Baird BF, et al. (1989) CD4 counts as predictors of opportunistic pneumonias in human immunodeficiency virus (HIV) infection. Ann Intern Med 111: 223-231.

2. Small CB, Klein RS, Friedland GH, Moll B, Emeson EE, et al. (1983) Community-acquired opportunistic infections and defective cellular immunity in heterosexual drug abusers and homosexual men. Am J Med 74: 433-441.

3. Munoz A, Wang MC, Bass S, Taylor JM, Kingsley LA, et al. (1989) Acquired immunodeficiency syndrome (AIDS)-free time after human immunodeficiency virus type 1 (HIV-1) seroconversion in homosexual men. Multicenter AIDS Cohort Study Group. Am J Epidemiol 130: 530-539.

4. Martin MP, Carrington M (2013) Immunogenetics of HIV disease. Immunol Rev 254: 245-264.

5. Fellay J, Ge D, Shianna KV, Colombo S, Ledergerber B, et al. (2009) Common genetic variation and the control of HIV-1 in humans. PLoS Genet 5: e1000791.

6. Hecht FM, Hartogensis W, Bragg L, Bacchetti P, Atchison R, et al. (2010) HIV RNA level in early infection is predicted by viral load in the transmission source. AIDS 24: 941-945.

7. Hollingsworth TD, Laeyendecker O, Shirreff G, Donnelly CA, Serwadda D, et al. (2010) HIV-1 transmitting couples have similar viral load set-points in Rakai, Uganda. PLoS Pathog 6: e1000876.

8. Yue L, Prentice HA, Farmer P, Song W, He D, et al. (2013) Cumulative impact of host and viral factors on HIV-1 viral-load control during early infection. J Virol 87: 708-715.

9. Goepfert PA, Lumm W, Farmer P, Matthews P, Prendergast A, et al. (2008) Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. J Exp Med 205: 1009-1017.

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

11. Stacey AR, Norris PJ, Qin L, Haygreen EA, Taylor E, et al. (2009) Induction of a striking systemic cytokine cascade prior to peak viremia in acute human immunodeficiency virus type 1 infection, in contrast to more modest and delayed responses in acute hepatitis B and C virus infections. J Virol 83: 3719-3733.

12. Brenchley JM, Price DA, Schacker TW, Asher TE, Silvestri G, et al. (2006) Microbial translocation is a cause of systemic immune activation in chronic HIV infection. Nat Med 12: 1365-1371.

13. Keating SM, Jacobs ES, Norris PJ (2012) Soluble mediators of inflammation in HIV and their implications for therapeutics and vaccine development. Cytokine Growth Factor Rev 23: 193-206.

14. Kuller LH, Tracy R, Belloso W, De Wit S, Drummond F, et al. (2008) Inflammatory and coagulation biomarkers and mortality in patients with HIV infection. PLoS Med 5: e203.

15. Sandler NG, Wand H, Roque A, Law M, Nason MC, et al. (2011) Plasma levels of soluble CD14 independently predict mortality in HIV infection. J Infect Dis 203: 780-790.

16. Vaidya SA, Korner C, Sirignano MN, Amero M, Bazner S, et al. (2014) Tumor Necrosis Factor alpha Is Associated With Viral Control and Early Disease Progression in Patients With HIV Type 1 Infection. J Infect Dis.

17. Shive CL, Mudd JC, Funderburg NT, Sieg SF, Kyi B, et al. (2014) Inflammatory Cytokines Drive CD4+ T-Cell Cycling and Impaired Responsiveness to Interleukin 7: Implications for Immune Failure in HIV Disease. J Infect Dis.

18. Deeks SG, Kitchen CM, Liu L, Guo H, Gascon R, et al. (2004) Immune activation set point during early HIV infection predicts subsequent CD4+ T-cell changes independent of viral load. Blood 104: 942-947.

19. Giorgi JV, Ho HN, Hirji K, Chou CC, Hultin LE, et al. (1994) CD8+ lymphocyte activation at human immunodeficiency virus type 1 seroconversion: development of HLA-DR+ CD38- CD8+ cells is associated with subsequent stable CD4+ cell levels. The Multicenter AIDS Cohort Study Group. J Infect Dis 170: 775-781.

20. Hunt PW, Brenchley J, Sinclair E, McCune JM, Roland M, et al. (2008) Relationship between T cell activation and CD4+ T cell count in HIV-seropositive individuals with undetectable plasma HIV RNA levels in the absence of therapy. J Infect Dis 197: 126-133.

21. Hunt PW, Martin JN, Sinclair E, Bredt B, Hagos E, et al. (2003) T cell activation is associated with lower CD4+ T cell gains in human immunodeficiency virus-

infected patients with sustained viral suppression during antiretroviral therapy. J Infect Dis 187: 1534-1543.

22. Liu Z, Cumberland WG, Hultin LE, Prince HE, Detels R, et al. (1997) Elevated CD38 antigen expression on CD8+ T cells is a stronger marker for the risk of chronic HIV disease progression to AIDS and death in the Multicenter AIDS Cohort Study than CD4+ cell count, soluble immune activation markers, or combinations of HLA-DR and CD38 expression. J Acquir Immune Defic Syndr Hum Retrovirol 16: 83-92.

23. Day CL, Kaufmann DE, Kiepiela P, Brown JA, Moodley ES, et al. (2006) PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. Nature 443: 350-354.

24. Trautmann L, Janbazian L, Chomont N, Said EA, Gimmig S, et al. (2006) Upregulation of PD-1 expression on HIV-specific CD8+ T cells leads to reversible immune dysfunction. Nat Med 12: 1198-1202.

25. Petrovas C, Chaon B, Ambrozak DR, Price DA, Melenhorst JJ, et al. (2009) Differential association of programmed death-1 and CD57 with ex vivo survival of CD8+ T cells in HIV infection. J Immunol 183: 1120-1132.

26. Zhang D, Shankar P, Xu Z, Harnisch B, Chen G, et al. (2003) Most antiviral CD8 T cells during chronic viral infection do not express high levels of perforin and are not directly cytotoxic. Blood 101: 226-235.

27. Douek DC (2003) Disrupting T-cell homeostasis: how HIV-1 infection causes disease. AIDS Rev 5: 172-177.

28. Okoye A, Meier-Schellersheim M, Brenchley JM, Hagen SI, Walker JM, et al. (2007) Progressive CD4+ central memory T cell decline results in CD4+ effector memory insufficiency and overt disease in chronic SIV infection. J Exp Med 204: 2171-2185.

29. Okoye AA, Picker LJ (2013) CD4(+) T-cell depletion in HIV infection: mechanisms of immunological failure. Immunol Rev 254: 54-64.

30. Orendi JM, Bloem AC, Borleffs JC, Wijnholds FJ, de Vos NM, et al. (1998) Activation and cell cycle antigens in CD4+ and CD8+ T cells correlate with plasma human immunodeficiency virus (HIV-1) RNA level in HIV-1 infection. J Infect Dis 178: 1279-1287.

31. Ananworanich J, Schuetz A, Vandergeeten C, Sereti I, de Souza M, et al. (2012) Impact of multi-targeted antiretroviral treatment on gut T cell depletion and HIV reservoir seeding during acute HIV infection. PLoS One 7: e33948.

32. Chomont N, El-Far M, Ancuta P, Trautmann L, Procopio FA, et al. (2009) HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. Nat Med 15: 893-900.

33. Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, et al. (1997) Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. Science 278: 1295-1300.

34. Strain MC, Little SJ, Daar ES, Havlir DV, Gunthard HF, et al. (2005) Effect of treatment, during primary infection, on establishment and clearance of cellular reservoirs of HIV-1. J Infect Dis 191: 1410-1418.

35. Ng OT, Laeyendecker O, Redd AD, Munshaw S, Grabowski MK, et al. (2014) HIV type 1 polymerase gene polymorphisms are associated with phenotypic differences in replication capacity and disease progression. J Infect Dis 209: 66-73.

36. Carrington M, Walker BD (2012) Immunogenetics of spontaneous control of HIV. Annu Rev Med 63: 131-145.

37. Paiardini M, Cervasi B, Reyes-Aviles E, Micci L, Ortiz AM, et al. (2011) Low levels of SIV infection in sooty mangabey central memory CD(4)(+) T cells are associated with limited CCR5 expression. Nat Med 17: 830-836.

38. Gattinoni L, Lugli E, Ji Y, Pos Z, Paulos CM, et al. (2011) A human memory T cell subset with stem cell-like properties. Nat Med 17: 1290-1297.

39. Lugli E, Dominguez MH, Gattinoni L, Chattopadhyay PK, Bolton DL, et al. (2013) Superior T memory stem cell persistence supports long-lived T cell memory. J Clin Invest 123: 594-599.

40. Cartwright EK, McGary CS, Cervasi B, Micci L, Lawson B, et al. (2014) Divergent CD4+ T Memory Stem Cell Dynamics in Pathogenic and Nonpathogenic Simian Immunodeficiency Virus Infections. J Immunol 192: 4666-4673.

41. Letvin NL, Mascola JR, Sun Y, Gorgone DA, Buzby AP, et al. (2006) Preserved CD4+ central memory T cells and survival in vaccinated SIV-challenged monkeys. Science 312: 1530-1533.

Discussion

A unifying theme in HIV/AIDS research is the pursuit of a cure capable of ending the global pandemic. Decades of research have proven that a multidisciplinary approach will be crucial for combating the biological and social complexities of this disease. Although the rate of newly acquired infections has decreased over time, and the advent of ART has prolonged life expectancy for those living with HIV/AIDS, the epidemic ensues unabated. The complete or near complete eradication of infectious diseases such as small pox and polio have required the development of an efficacious vaccine and HIV/AIDS will be no exception.

To date, a preventative vaccine for HIV-1 has proven difficult and elusive to achieve. Vaccines that have made it to Phase III clinical trials have either failed to show protection (the Merck Step trial [372]) or have only shown modest protection (the RV144 Thailand trial [373]). Vaccine research, although making small steps forward, is often misguided, because very little is known about what constitutes a protective immune response against HIV-1 infection. The extensive genetic diversity of HIV-1 further complicates vaccine endeavors; this fact perhaps implies that a multifaceted approach that induces multiple arms of the immune system will be key [12].

Although a vaccine is considered the holy grail of HIV-1 prevention research, its development is likely years in the making. Thus in the interim, we rely on alternative prevention strategies to curb the epidemic. Thanks to the development of ART, growing numbers of individuals living with AIDS experience a healthier longer life, while ART-suppression of viral load has been proven to reduce transmission efficiency. Despite this, AIDS-related co-morbidities still exist in treated individuals, and ART has not successfully halted the spread of the disease. Thus, research looking beyond anti-retroviral treatment alone is important for future efforts to eliminate this disease.

The HIV/AIDS field stands to gain novel insights into HIV-1 pathogenesis through investigating the natural course of HIV-1 infection. Such insights could reveal novel vulnerabilities that could be targeted in the next-generation of prevention and therapeutic strategies. To date, most research on HIV has focused on the role of the host genetic factors and immune responses in defining pathogenesis [369]. This research has, without a doubt, been instrumental in characterizing the role of host genetic factors such as HLA-I alleles and gender, as well as identified numerous immunological abnormalities in HIV-1 infection. Unfortunately, this research fails to recognize or take into account the virus genotype and how this might affect pathogenesis [374]. The results presented in this body of work strongly suggest a dominant role for viral factors in defining multiple facets of HIV-1 pathogenesis. Furthermore, this work suggests that an integrative approach that seeks to understand the interplay between the host and virus will ultimately yield more meaningful insights into pathogenesis.

Through a long-standing collaboration with one of the largest discordant couples cohort in the world, the Zambia Emory HIV Research Project, we have been afforded the unique opportunity to study the host and viral factors important in defining HIV-1 pathogenesis. Preliminary studies with this cohort supported a role for transmitted viral characteristics in defining clinical parameters during early infection. First, it was observed that viral loads appeared to be “heritable”, i.e. high viral loads in the chronically infected partner translated into a higher set point viral load within the newly infected linked recipient [318,319]. This “heritable” factor was hypothesized to be attributable to an unknown viral characteristic. Subsequently, it was shown that transmitted CTL escape mutations in Gag, but not other proteins, correlated with lower set point viral loads, implicating that the transmission of an attenuated Gag might provide a clinical benefit to the newly infected individual [320]. We show here that

replicative capacity of the transmitted Gag contributes to the heritability of viral loads between chronically infected donors and their acutely infected linked recipients.

The work presented here has revolutionized the field's ability to assess important viral characteristics in the context of disease progression. We developed a restriction enzyme based cloning strategy in which *gag* sequences are inserted into a common a common clade C proviral backbone, MJ4. Previous methods to assess *in vitro* replicative capacity of HIV isolates are less than ideal. The major limitations of past approaches revolved around the technical difficulties in cloning retroviruses, which have extremely high genetic variability and their unstable replication in bacteria. While the cloning of *env* genes for the analysis of antibody and co-receptor utilization is well established, cloning strategies for other genes are underdeveloped. Previous methods circumvented the difficulty in cloning by propagating the virus from plasma, or by recombining viral DNA with NL4-3 in T-cells. These methods risk significantly skewing the genetic nature of the recovered virus. Additionally, researchers in the field tend to use clade B laboratory-adapted strains such as NL4-3 when creating chimeric viruses [310-313,334,336,375,376]. Therefore, the enzyme-based cloning method utilized here provides the field with a superior alternative to assessing *in vitro* replicative capacity. A more in-depth discussion of the methods can be found in the Appendix. Although initially designed for the cloning of *gag*, these methods can be extended to other viral gene products and have opened the door for future research aimed at identifying additional viral determinants of HIV-1 pathogenicity.

Perhaps one of the most intriguing as well as frustrating aspects of HIV-1 is the rapid rate of evolution, which has led to a seemingly endless amount of global HIV-1 diversity [26,41]. Evolution of viral diversity occurs largely in response to strong selection forces such as immune responses enacting within the host. Viral escape from host immune responses increases the *in vivo* fitness of the virus, but such escape might

impact the *in vitro* replicative capacity of the virus [287]. HIV-1 viral evolution follows a pattern that is dictated by viral fitness, but the term "fitness" can have several meanings, and the meaning might have different connotations within different contexts. *In vivo* fitness is influenced not only by *in vitro* replicative capacity, but also by all of the necessary elements for *in vivo* replication and survival such as immune evasion.

Within an epidemic, viruses are transmitted from host to host, and in the case of HIV-1, a severe genetic bottleneck occurs suggesting selective transmission of certain viral genetic characteristics [68,69]. Thus, transmission likely imposes additional levels of selection and influences HIV-1 genetic variability within a population. Adding to this complexity, geographically distinct populations of people are often genetically different and display different distributions of alleles such as HLA-I [283]. A cross-protective vaccine would need to protect against many diverse strains of HIV-1, but would have to be designed such that the epitopes targeted could be presented by a majority of HLA-I alleles worldwide. Therefore, studying the genetic diversity of HIV-1 in the context of different populations is highly important for understanding pathogenesis and extracting out the most meaningful correlates of protection that might be conserved across multiple epidemics.

The results presented here encompass a large-scale study of over 200 subtype C acutely infected individuals from Zambia in which the *in vitro* replicative capacity of the transmitted virus as conferred by Gag, was measured. Although Gag is one of the more conserved HIV-1 encoded proteins due to structural and functional constraints [85,377], significant sequence variation in Gag was observed between different individuals. Substitution of the transmitted *gag* into MJ4 resulted in drastic differences of *in vitro* replicative capacity. *In vitro* replicative capacity of the transmitted Gag, in turn, was found to correlate with disease progression independent of set-point viral load and host factors such as gender and HLA-I alleles.

These studies have highlighted the contribution of Gag in defining *in vitro* replicative capacity of the virus. Although these data suggest that Gag plays a major role in defining the kinetics of the viral lifecycle, certainly other viral proteins might additionally influence this. A recent study has demonstrated that *in vitro* RC of Pol was found to correlate with disease progression in both subtype A and D [368]. Furthermore, Mann et al. demonstrated that sequence variation in Nef resulted in differential downregulation of CD4 and HLA-I [378,379]. Novel cloning strategies that allow for feasible cloning of a large number of full-length infectious molecular clones have been recently developed by our research group (Deymier et al. manuscript in preparation). Investigation of the *in vitro* replication kinetics of full-length infectious molecular clones will further our understanding of how viral characteristics interplay with the host to define transmission efficiency as well as HIV-1 pathogenesis.

Based on the surmounting evidence that CD8+ T cells are key players in the immune response to HIV-1, research efforts have turned towards defining what type of CD8+ T cell responses would likely protect in a vaccine. A successful cellular based vaccine will likely need to induce responses that recognize vulnerable regions within the virus that escape with difficulty, i.e. “protective epitopes”. Another feature of ideal “protective epitopes” would be more ubiquitous presentation by a wide diversity of HLA-I alleles globally. The field has focused an exhaustive amount of research on studying the immunological mechanisms by which protective HLA-I alleles such as B*27 and B*57 confer control of viral replication. These alleles have been shown to target structurally conserved regions in Gag, which incur fitness costs upon escape [251,253-257,357,369,380]. A major caveat to such studies is the fact that most people do not have these protective alleles. Additionally, the observed protective effect of HLA-I alleles as well as *in vitro* fitness costs can be cohort specific and attributed to genetic differences in circulating HIV-1 strains [381-383]. The challenge then becomes to discover novel

epitopes both provide protection but that can also be targeted by multiple different HLA-I alleles.

A recent comprehensive statistical analysis (Carlson and Shaefer et al, Science, in press), which encompasses 1899 subtype C sequences from Zambia and South Africa, has identified 199 HLA-associated polymorphisms within Gag, Pol, and Nef. Additionally, longitudinal sequence analysis of 80 recently infected linked transmission pairs from this cohort has shed light on the kinetics of CTL epitope escape as well as inferred mutations likely to carry fitness costs based on the rate of reversion. Through studying the *in vitro* replicative capacity of naturally occurring Gag sequences, vulnerable regions of Gag can be more thoroughly identified within the context of this subtype C infected cohort.

In chapter 2, we identify several previously unrecognized HLA-associated polymorphisms that were statistically associated with differences in vRC. An overabundance of polymorphisms that affected RC was found in matrix and p2 of Gag. To date, the majority of protective HLA-I epitopes have been described in P24, thus this work further characterizes the functional constraints of additional regions in Gag, which might better inform rational vaccine design. The MA domain of HIV-1 plays key roles in the targeting of Gag to the plasma membrane and in incorporation of Env during virus assembly [85,87,89,359]. Reflecting on its diverse roles in the retroviral lifecycle, it is therefore not surprising that we observed mutations within MA that were statistically associated with changes in RC. We found that the HLA-associated polymorphism, K12E, was associated with the greatest fitness defect across all of the sequences. Moreover, 12E is associated with HLA-A*7401, an allele previously described as protective in multiple cohorts [257,281,384]. Position 12 is also predicted to be the critical anchor residue within the A*74 epitope Gag-KR9 [258]. It is likely that polymorphisms at this position would interfere with binding, although further studies are required to demonstrate that

this epitope is indeed presented *in vivo* and that polymorphisms confer functional escape.

Position 12 lies in a highly basic region of the N-terminus of matrix important for membrane targeting [359]. Mutational analysis by Freed et al. demonstrated that the mutation L13E was associated with decreased infectivity and a disruption in Env incorporation [88]. Interestingly, 13L is 100% conserved among all HIV-1 subtype C MA sequences within Zambia. It is plausible, therefore, that changes at an adjacent residue might, to some degreem impact *in vitro* replicative capacity. Based on longitudinal sequencing of 80 recently infected Zambians from the ZHERP cohort, polymorphisms at position 12 reverted at a rate of 25% per year. This was comparable to the kinetics of the well-documented B*57 escape mutation T242N, which reverts at a rate of 27.5% per year. Furthermore, in a total of ten recently infected individuals with A*7401, escape only occurred in one individual and occurred late at 24-months post infection.

Several different polymorphisms were observed in addition to glutamate at position 12. The basic amino acid, lysine (K), is the consensus residue within Zambian clade C sequences. The polar, but non-charged glutamine (Q) is the second most prevalent residue followed by the negatively charged glutamate (E), and lastly by asparagine (N), which is the least prevalent. 12Q has also been statistically associated with HLA-A*7401 (Carlson and Shaefer et al, Science, in press). While 12N might also be driven by HLA-A*7401, it is likely too rare to demonstrate a statistical association with the allele. In order to investigate the *in vitro* fitness impact of A*7401 associated CTL escape mutations, I engineered all naturally occurring polymorphisms at residue 12 into MJ4, and assessed the *in vitro* replicative capacity of each virus. Interestingly, MJ4's Gag sequence naturally contains an arginine at position 12. Preliminary results demonstrate that, in the context of MJ4, N12E resulted in a half-log reduction in *in vitro* replicative capacity. (Appendix, Figure 1A and 1B). MJ4's naturally occurring asparagine was

associated with the highest *in vitro* replication followed by N12K and N12Q, both equal in their growth kinetics. From this, it appears as if the only polymorphism associated with a drastic replication defect is 12E. If the presence of an N, Q, or K results in minimal *in vitro* replication defects, it seems puzzling why 12E would be selected for at all. Even though 12E is rare, occurring in only 5% of all sequences, it is selected for under certain circumstances; otherwise it would not appear in naturally occurring sequences. One hypothesis might be that not all variants confer the same phenotype in terms of escape from CTL recognition; 12E might inhibit binding to a greater degree than the other residues, thus, leading to more complete escape.

An alternative hypothesis would be that *in vitro* fitness defects might be context dependent. Therefore, 12E might be deleterious in some sequence contexts, but not others, potentially due to the presence of compensatory mutations. In order to address this, I engineered 12E into six additional Gag-MJ4 chimeric viruses all with drastically different Gag sequences, but that all harbored the consensus residue (K) at 12. Interestingly, a wide range of *in vitro* fitness defects were observed for the same mutations, K12E, in six different Gag sequences (Appendix, Figure 2). For some sequences, K12E did not have an *in vitro* fitness defect, while in others, *in vitro* replication was drastically reduced to almost background levels. While these studies are admittedly preliminary, they suggest that CTL escape mutations might impact *in vitro* fitness to varying extents depending on what other viral polymorphisms are present. Ongoing studies are investigating the potential role of co-varying amino acids in rescuing the *in vitro* replicative defects of K12E.

The varying *in vitro* fitness defects of K12E might explain the reversion kinetics, which is around 25% per year. It is possible that minimal defects in viral replication of particular sequence contexts can explain the instances in which reversion does not occur. To date, research identifying the *in vitro* fitness defects associated with CTL escape has

often engineered desired mutations into laboratory-adapted backbones such as NL4-3. This data argues that such approaches might not accurately predict how certain polymorphisms would impact the *in vitro* fitness within the context of certain populations of HIV-1 infected individuals.

Taken together, this suggests that perhaps the protective effect of HLA-A*7401 within this cohort might be attributed, at least in part, to an effective CTL response directed against the Gag-KR9 epitope in MA. It is possible that a prolonged CTL response due to a lack of escape might suppress viremia in A*7401 individuals and thus would delay disease progression. Immunological studies aimed at characterizing the effect of differential epitope escape in the context of varying replication kinetics will be key for further elucidating the mechanism of protection for HLA-A*7401 in this cohort.

In addition to helping elucidate potential vulnerabilities in virus structure and function that can be harnessed to develop a more effective cellular based vaccine, the studies presented here have demonstrated a role of transmitted viral characteristics in defining HIV-1 disease progression and pathogenesis. RC of the transmitted virus was associated with set point VL in newly infected linked recipients. This was the first definitive evidence that RC of the virus played a role in early acute subtype C infection. Moreover, we demonstrated that RC correlated with viral loads in chronically infected donors. Thus, the observed “heritability” of viral load between partners within a transmission pair can be explained, at least in part, by replicative capacity of the transmitted virus. By assessing the *in vitro* replicative capacity of Gag-MJ4 chimeric viruses, we demonstrate that this effect on HIV-1 viral load can be attributable to Gag.

Moving forward, it will be important to understand how events during transmission might impose selective pressures for viral replicative capacity. Genetic heterogeneity within the quasispecies of a chronically infected individual is likely to result in different *in vitro* replicative capacities of between different viruses present in

the quasispecies. A recent study by Carlson and Shafer et al. (Science, in press) demonstrated a selection bias towards the transmission of consensus residues in the viral genome. This was the case even when mixtures of consensus and non-consensus residues were present within the quasispecies of the chronically infected donor. These results are in congruence with previous observations that ancestral sequences, or those that are genetically similar to the most recent common ancestor, are more often transmitted [77,385]. These results suggest that through each transmission event, the viral genotype is set backwards in evolutionary time towards consensus. Therefore, immune-mediated evolution might occur slower than expected on population level. A recent study comparing ancestral subtype B sequences (1979-1989) to modern sequences from (2000-2011) isolated from four cities in North America demonstrated slow adaptation in Gag and Nef over time. Furthermore, replicative capacity of Gag did not differ between ancestral and modern sequences while they observed an increase of Nef-mediated CD4 down-regulation over time [386]. It will be important to extend these studies to other subtypes and epidemics in order to elucidate how viral evolution might affect pathogenesis of the epidemic over time, both in terms of protein function as well as transmission of pre-adapted CTL epitopes.

Certain residues within the virus are likely absolutely critical for the viral lifecycle and these functions must be preserved in order to ensure the virus' survival on a population level. This might explain the selection bias towards transmission of consensus residues; however, we also demonstrate in Chapter 2 that a majority of polymorphisms present within Gag were in fact associated with increases in RC. Moreover, we observed a statistically significant negative correlation between sequence similarity to the Zambian consensus and RC. This finding has been independently demonstrated in separate studies [313] and suggests that while certain residues might be driven to consensus due to functional constraints, consensus-like Gag sequences confer a

lower *in vitro* fitness than those with an accumulation of additional polymorphisms, which are likely compensatory in nature.

In chapter 3, we demonstrate that high vRC was associated with a distinct cytokine profile during early acute infection characterized by higher levels of interferon as well as higher levels of the pro-inflammatory cytokines IL-6 and IL-1beta. Early viral replication and/or recognition of Gag antigen might induce an early innate immune response resulting in a cytokine storm and cellular activation. It is possible that the consequences of rapid viral replication might be selected against during transmission. Rapid replication might induce a potent innate antiviral response that would suppress viral replication before the virus could replicate to sufficient levels at the site of infection to allow for dissemination. Potent antiviral responses are often characterized by the induction of interferon [387]. It has recently been shown that HIV-1 transmitted founder viruses are more resistant to interferon than viruses derived from chronically infected individuals [79,388] highlighting the strong selective pressure imposed by innate antiviral responses such as interferon on the virus during transmission.

Therefore, although the selection of consensus residues during transmission likely results in the preservation of critical virus structure and function, it is also plausible that viruses with very high levels of replication might be selected against during the transmission event. An optimal RC might allow the virus to slowly replicate unnoticed by the innate immune system, thus, avoiding a robust innate antiviral response, but such viruses would also retain necessary functions for the viral lifecycle.

The data presented in chapter 3 describes the mechanisms by which viral replicative capacity influences CD4+ T cell decline. We demonstrate that in individuals recently infected with HIV-1 subtype C, low vRC was associated with a delay in CD4+ T cell decline. A clear and durable benefit was observed even down to CD4+ T cell counts below 200, the clinical definition of AIDS. This suggests that even in the face of continual

viral adaptation and evolution of RC throughout infection, RC of the initial infecting viral strain initiates crucial events early during pathogenesis that dictate the long-term trajectory of disease progression. In support of this, we further demonstrate that vRC predicts CD4+ T cell decline independent of set point viral load. Acute infection is characterized by a burst of viral load in plasma that peaks around 2-3 weeks post infection and is quickly resolved down to a more steady state set point [389]. The establishment of set point is influenced by host factors such as HLA-I alleles and gender [319]. Individuals in this cohort were not identified early enough during acute infection to be able to accurately define peak viral load. Nevertheless, we hypothesize that peak viral load might be more reflective of vRC than set point viral load, which is defined by the balance between viral replication and host factors that enact to control viral replication.

We additionally demonstrate that vRC predicts CD4+ T cell decline independent of the canonical protective HLA-class I alleles B*57 and B*5801. Not all individuals with protective alleles become non-progressors or elite controllers [390]. We find that vRC significantly dichotomizes the trajectories of CD4+ T cell decline in individuals that harbor B*57/B*5801. Therefore, infection with a highly pathogenic virus might mask the protective benefits of these HLA-I alleles, and would explain why some HLA-B*57 positive individuals progress to AIDS. This collectively suggests that the transmitted Gag plays a crucial role in defining the early assault on the immune system before host factors can control acute viremia down to set point viral load.

We go on to show in chapter 3 that vRC influences many facets of HIV-1 immunopathology that have been shown to be important factors associated with CD4+T cell decline. In acute HIV-1 infection, a sharp rise in plasma viremia is associated with an intense cytokine storm characterized by high levels of pro-inflammatory cytokines such as IFN, TNF-alpha, IL-6, IL-8, and a late peak in levels of the immunoregulatory

cytokine IL-10 [363]. This extensive cytokine storm is unique to HIV/SIV infection. Limited levels of cytokines characterize other chronic infections such as HCV and HBV, which is generally cleared [363]. This suggests that the early inflammatory environment in the case of HIV-1 translates into highly deleterious immunopathology. In fact, a hallmark of pathogenic HIV-1 infection is chronic immune activation, which is characterized by an increase in circulating pro-inflammatory cytokines as well as an increase in T-cell activation [152,154,156,160,164,170,171,173,202]. Moreover, the translocation of microbial products from the gut into systemic circulation has been implicated in exacerbating immune activation [197].

We used Luminex technology to measure the levels of 16 different cytokines, chemokines and markers of microbial translocation at seroconversion. Furthermore, we performed principal component analysis in order to characterize a unique inflammatory profile associated with differences in vRC. We find that the first principal component, which describes the highest variation within the data, is associated with vRC. Conversely, subsequent principal components, which explain less variation in the data, are correlated with viral load, gender, and HLA-I alleles. This suggests that during early infection, viral replicative capacity induces a unique inflammatory milieu that is less influenced by host factors. Individuals infected with attenuated viruses exhibited dampened levels of pro-inflammatory cytokines and markers of microbial translocation.

Most notably, vRC was associated with the pro-inflammatory cytokines IFN-gamma, IFN-a, IL-6, IL-1beta, and to a lesser extent, sCD14 and LPS, which are markers of microbial translocation. Interestingly, IL-6 and IL-1B are two inflammatory cytokines that have been implicated in driving CD4+ T cell turnover and impairing homeostatic proliferation, key characteristics in pathogenic HIV infection [367]. Consequently, vRC was also correlated with levels of CD4+ T cell proliferation, which in turn predicted disease progression.

Residual immune activation persists in some individuals on ART despite suppressed viral loads, and activation levels predict non-AIDS related co-morbidities such as adverse cardiovascular events, neurological impairments, and aging [104,105,152,154,156,202]. Based on the data presented here, it is plausible to hypothesize that individuals initially infected with a highly replicating virus might present with residual immune activation despite suppressive ART therapy. Furthermore, the fact that vRC correlated with early establishment of the inflammatory milieu further supports the notion that early ART treatment would more effectively limit immune activation, and thus, would provide additional clinical benefit. Finally, non-pathogenic SIV infection is characterized by a robust cytokine storm that is quickly dampened [391]. It is possible that low vRC induces a similar phenomenon in humans in which the levels of pro-inflammatory cytokines decrease over time. In-depth longitudinal studies as well as measurement of immunoregulatory cytokines such as TGF-beta will help elucidate if this is the case.

We found it interesting that vRC was more highly associated with differences in pro-inflammatory cytokines rather than microbial translocation, although this evidence is not sufficient to definitively conclude that vRC does not influence gut damage and microbial translocation. The causes of microbial translocation have not been fully elucidated, and it is not yet unclear if early inflammation is a driver of gut damage. Higher levels of viral replication and antigen load might result in higher levels of innate recognition of viral PAMPS. The three-dimensional Gag lattice structure is also recognized by TRIM5 alpha, which can act as an innate immune sensor to induce the IFN-pathway [191]. In addition, higher levels of early viral replication might deplete critical CD4+ T subsets such as Th17 cells to a more severe extent. This might then cause immunological dysfunction at mucosal sites, leading to more severe microbial translocation, which itself in turn leads to exacerbated immune activation. Therefore,

investigating how vRC of the virus, specifically in the context of Gag, interacts with the host to induce immune activation during the earliest stages of infection will require more in-depth studies. These studies might include viral stimulation and characterization of APCs, explant models, measuring additional markers such as IL-17, as well as analogous *in vivo* studies using a humanized mouse model.

Chronic viral infections such as HIV-1 induce immune activation, but are also associated with a functional exhaustion of the immune system [219,220,231,236]. We find that vRC correlated with early levels of CD8+ T cell activation. Furthermore, levels of CD8+ T cell activation correlated with CD4+ T cell decline, thus providing a likely mechanistic link between vRC and CD4+ T cell decline. It is well accepted that immune activation predicts CD4+ T cell decline better than viral load, however potential mechanisms for this were undefined until now. We demonstrate, for the first time, that this phenomenon can be, at least in part, attributed to the RC of the transmitted virus, which we show is driven by Gag.

In addition to lower levels of activation, low vRC was associated with CD8+ T cells that were PD-$1^{low}CD57^{high}$. Although CD57 has traditionally been used as a marker of senescence and differentiation [130], CD57 expression in HIV-1 infection has been shown to have different functional consequences than CD57 expression in CMV infection or in aging [239]. Furthermore, CD57 was associated with delayed disease progression, and expression levels were significantly restored upon ART highlighting the protective role of CD57 in HIV-1 infection [240]. Furthermore, Petrovas et al. showed that CD8+T cells that maintain high levels of CD57 but low levels of PD-1 are more resistant to apoptosis [238]. Therefore, low vRC, and thus by extension, lower early antigen load might limit CD8+ T cells from becoming exhaustion and protect cells from apoptosis. The association between reduced exhaustion and low vRC translated into increased CD8+ T cell cytotoxicity. Moreover, the levels of CD8+ T cell activation and exhaustion

predicted CD4+ T cell decline within this cohort. Therefore, it seems that infection with an attenuated virus affords protection from disease progression in part by inducing a more functional CTL response in which the CD8+ T cells do not become exhausted and survive long enough to suppress viral replication. Understanding the innate immune pathways that lead to this type of response will enhance vaccine efforts aimed at inducing cellular immune responses.

HIV-1 pathogenesis is also characterized by increased activation and cellular turnover within the CD4+ T cell compartment, which results in a disruption of critical homeostasis necessary for maintaining healthy levels of peripheral CD4+ T cells [140,150,175]. vRC was associated with levels of CD38/HLA-DR and Ki67 on CD4+ T cells, which in turn translated into differential disease progression within this cohort. Interestingly, individuals infected with attenuated viruses maintained low levels of activation and proliferation that were akin to that of HIV-uninfected individuals. This suggests that attenuated viral replication might lead to a muted inflammatory response that limits cellular turn over thus preserving the homeostatic signals necessary for the replenishment of susceptible CD4+ T_{EM} populations without the depletion of CD4+ T_{CM} cells. Low vRC was also associated with a lower viral burden in CD4+ T_{CM}, which is a population characterized as being longer-lived and more stable [132,140,150]. Direct infection of this subset might further increase immune activation, proliferation, and disrupt critical homeostasis.

Although ART treatment has revolutionized the care of HIV-1 infected individuals, it is a life-long treatment that has side effects, financial costs and is not available to many who need it. These data have turned research efforts towards discovering a cure. One of the main barriers to a sterilizing cure is the presence of a latent viral reservoir.

Properly followed ART regimens suppress viral loads in HIV-infected patients; however, upon treatment interruption, viral loads rebound [102,103,146,147,241]. Thus, ART treatment does not effectively eradicate all infected cells, and replication-competent virus may still be present in a latent form in certain tissue compartments. Additionally, several studies have investigated the effects of treating HIV-1 early during acute infection. One study that has received a lot of media attention is that of the "Mississippi baby". This baby was born to an HIV+ mother, treated with aggressive treatment almost immediately after birth, and then stopped receiving treatment at around 15-18 months. Tests revealed that the child had, in fact, been infected during birth, but showed no signs of HIV RNA or DNA after stopping treatment [392]. Although anecdotal, this case has certainly raised our hopes about the feasibility of eradicating the reservoir early in infection. Larger scale studies in adults have shown that a small percentage of individuals who initiated treatment during early acute infection maintained undetectable levels of viral load after treatment interruption [244,245]. Understanding the mechanisms by which this occurs will provide insight into describing clinical parameters that define when virus has been eradicated and when it might be safe to take individuals off therapy.

The population highlighted to be integral for the maintenance of latency has been shown to be mainly CD4+ T_{CM}, which have the capacity to survive for extended periods of time, and constitute a long-lived reservoir [147]. We provide further evidence that infection of this subset occurs early during infection, and show that viral burden in CD4+ T_{CM} ultimately dictates disease progression within this cohort. Unlike previous studies, we demonstrate that vRC of the transmitted virus impacts HIV-1 viral burden in this compartment, and might actually impact the size of the viral reservoir, although studies measuring viral burden within individuals on suppressive ART would be warranted in order to definitively conclude that vRC impacts the reservoir size. Therefore, it is

plausible to hypothesize that individuals infected with attenuated viruses might be more easily cured due to due to the establishment of a smaller reservoir.

IL-7 has been shown to promote CD4+ T cell survival and proliferation. When administered to HIV-infected individuals on ART, IL-7 led to the rapid proliferation of CD4+ memory T cells and substantially increased the number of CD4+ T cells that harbored HIV DNA [393]. We demonstrate that infection with viruses with high replicative capacity was associated with higher levels of IL-7, which might promote maintenance of the latent viral reservoir in these individuals. Therefore, infection with an attenuated virus might limit the viral reservoir and make it more vulnerable to elimination.

In summary, we highlight a previously unrecognized role for viral characteristics in determining HIV-1 disease progression and pathogenesis. Viral replicative capacity as defined by Gag was correlated with long-term disease progression independent of host factors such as protective HLA-I alleles. We show that a potential mechanism for this is the induction of differential states of immune activation based on the replicative capacity of the transmitted virus. In fact, individuals infected with attenuated viruses displayed immune phenotypes more similar to healthy individuals and reminiscent of non-pathogenic SIV infection of natural hosts. These results are intriguing and perhaps suggest a role for viral attenuation in viral-host adaptation over the course of evolution. Understanding the mechanisms that underlie HIV-1 pathogenesis will give us insight into what pathways might need to be targeted in order to reverse the immune dysfunction associated with HIV-1 infection.

Literature cited in the Introduction and Discussion

1. Centers for Disease C (1981) Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. MMWR Morb Mortal Wkly Rep 30: 305-308.

2. Masur H, Michelis MA, Greene JB, Onorato I, Stouwe RA, et al. (1981) An outbreak of community-acquired Pneumocystis carinii pneumonia: initial manifestation of cellular immune dysfunction. N Engl J Med 305: 1431-1438.

3. Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, et al. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science 220: 868-871.

4. Gallo RC, Salahuddin SZ, Popovic M, Shearer GM, Kaplan M, et al. (1984) Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science 224: 500-503.

5. <20121120_unaids_global_report_2012_with_annexes_en.pdf>.

6. Trask SA, Derdeyn CA, Fideli U, Chen Y, Meleth S, et al. (2002) Molecular epidemiology of human immunodeficiency virus type 1 transmission in a heterosexual cohort of discordant couples in Zambia. J Virol 76: 397-405.

7. Hemelaar J, Gouws E, Ghys PD, Osmanov S (2006) Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. AIDS 20: W13-23.

8. Osmanov S, Pattou C, Walker N, Schwardlander B, Esparza J, et al. (2002) Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. J Acquir Immune Defic Syndr 29: 184-190.

9. Palella FJ, Jr., Delaney KM, Moorman AC, Loveless MO, Fuhrer J, et al. (1998) Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N Engl J Med 338: 853-860.

10. Hammer SM, Eron JJ, Jr., Reiss P, Schooley RT, Thompson MA, et al. (2008) Antiretroviral treatment of adult HIV infection: 2008 recommendations of the International AIDS Society-USA panel. JAMA 300: 555-570.

11. Jaffar S, Govender T, Garrib A, Welz T, Grosskurth H, et al. (2005) Antiretroviral treatment in resource-poor settings: public health research priorities. Trop Med Int Health 10: 295-299.

12. Barouch DH (2008) Challenges in the development of an HIV-1 vaccine. Nature 455: 613-619.

13. Clavel F, Guetard D, Brun-Vezinet F, Chamaret S, Rey MA, et al. (1986) Isolation of a new human retrovirus from West African patients with AIDS. Science 233: 343-346.

14. Chakrabarti L, Guyader M, Alizon M, Daniel MD, Desrosiers RC, et al. (1987) Sequence of simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses. Nature 328: 543-547.

15. Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR (1989) An African primate lentivirus (SIVsm) closely related to HIV-2. Nature 339: 389-392.

16. Worobey M, Telfer P, Souquiere S, Hunter M, Coleman CA, et al. (2010) Island biogeography reveals the deep history of SIV. Science 329: 1487.

17. VandeWoude S, Apetrei C (2006) Going wild: lessons from naturally occurring T-lymphotropic lentiviruses. Clin Microbiol Rev 19: 728-762.

18. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, et al. (1999) Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature 397: 436-441.

19. Hahn BH, Shaw GM, De Cock KM, Sharp PM (2000) AIDS as a zoonosis: scientific and public health implications. Science 287: 607-614.

20. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, et al. (2006) Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. Science 313: 523-526.

21. Sharp PM, Hahn BH (2011) Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med 1: a006841.

22. Vidal N, Peeters M, Mulanga-Kabeya C, Nzilambi N, Robertson D, et al. (2000) Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. J Virol 74: 10498-10507.

23. Korber B, Muldoon M, Theiler J, Gao F, Gupta R, et al. (2000) Timing the ancestor of the HIV-1 pandemic strains. Science 288: 1789-1796.

24. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, et al. (2008) Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature 455: 661-664.

25. Zhu T, Korber BT, Nahmias AJ, Hooper E, Sharp PM, et al. (1998) An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. Nature 391: 594-597.

26. McCutchan FE (2000) Understanding the genetic diversity of HIV-1. AIDS 14 Suppl 3: S31-44.

27. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, et al. (2000) HIV-1 nomenclature proposal. Science 288: 55-56.

28. Gilbert MT, Rambaut A, Wlasiuk G, Spira TJ, Pitchenik AE, et al. (2007) The emergence of HIV/AIDS in the Americas and beyond. Proc Natl Acad Sci U S A 104: 18566-18570.

29. Peeters M, Gueye A, Mboup S, Bibollet-Ruche F, Ekaza E, et al. (1997) Geographical distribution of HIV-1 group O viruses in Africa. AIDS 11: 493-498.

30. Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, et al. (2009) A new human immunodeficiency virus derived from gorillas. Nat Med 15: 871-872.

31. Vallari A, Bodelle P, Ngansop C, Makamche F, Ndembi N, et al. (2010) Four new HIV-1 group N isolates from Cameroon: Prevalence continues to be low. AIDS Res Hum Retroviruses 26: 109-115.

32. Keele BF, Jones JH, Terio KA, Estes JD, Rudicell RS, et al. (2009) Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. Nature 460: 515-519.

33. Gao F, Yue L, White AT, Pappas PG, Barchue J, et al. (1992) Human infection by genetically diverse SIVSM-related HIV-2 in west Africa. Nature 358: 495-499.

34. Damond F, Descamps D, Farfara I, Telles JN, Puyeo S, et al. (2001) Quantification of proviral load of human immunodeficiency virus type 2 subtypes A and B using real-time PCR. J Clin Microbiol 39: 4264-4268.

35. Pieniazek D, Ellenberger D, Janini LM, Ramos AC, Nkengasong J, et al. (1999) Predominance of human immunodeficiency virus type 2 subtype B in Abidjan, Ivory Coast. AIDS Res Hum Retroviruses 15: 603-608.

36. de Silva TI, Cotten M, Rowland-Jones SL (2008) HIV-2: the forgotten AIDS virus. Trends Microbiol 16: 588-595.

37. Popper SJ, Sarr AD, Gueye-Ndiaye A, Mboup S, Essex ME, et al. (2000) Low plasma human immunodeficiency virus type 2 viral load is independent of proviral load: low virus production in vivo. J Virol 74: 1554-1557.

38. Rowland-Jones SL, Whittle HC (2007) Out of Africa: what can we learn from HIV-2 about protective immunity to HIV-1? Nat Immunol 8: 329-331.

39. Roberts JD, Bebenek K, Kunkel TA (1988) The accuracy of reverse transcriptase from HIV-1. Science 242: 1171-1173.

40. Onafuwa-Nuga A, Telesnitsky A (2009) The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. Microbiol Mol Biol Rev 73: 451-480, Table of Contents.

41. Taylor BS, Hammer SM (2008) The challenge of HIV-1 subtype diversity. N Engl J Med 359: 1965-1966.

42. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, et al. (1995) Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature 373: 123-126.

43. Wei X, Ghosh SK, Taylor ME, Johnson VA, Emini EA, et al. (1995) Viral dynamics in human immunodeficiency virus type 1 infection. Nature 373: 117-122.

44. Korber B, Gaschen B, Yusim K, Thakallapally R, Kesmir C, et al. (2001) Evolutionary and immunological implications of contemporary HIV-1 variation. Br Med Bull 58: 19-42.

45. Kuiken CL, Lukashov VV, Baan E, Dekker J, Leunissen JA, et al. (1996) Evidence for limited within-person evolution of the V3 domain of the HIV-1 envelope in the amsterdam population. AIDS 10: 31-37.

46. Bernardin F, Kong D, Peddada L, Baxter-Lowe LA, Delwart E (2005) Human immunodeficiency virus mutations during the first month of infection are preferentially found in known cytotoxic T-lymphocyte epitopes. J Virol 79: 11523-11528.

47. Fischer W, Ganusov VV, Giorgi EE, Hraber PT, Keele BF, et al. (2010) Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. PLoS One 5: e12303.

48. Goulder PJ, Watkins DI (2004) HIV and SIV CTL escape: implications for vaccine design. Nat Rev Immunol 4: 630-640.

49. McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF (2010) The immune response during acute HIV-1 infection: clues for vaccine development. Nat Rev Immunol 10: 11-23.

50. Richman DD, Wrin T, Little SJ, Petropoulos CJ (2003) Rapid evolution of the neutralizing antibody response to HIV type 1 infection. Proc Natl Acad Sci U S A 100: 4144-4149.

51. Rong R, Li B, Lynch RM, Haaland RE, Murphy MK, et al. (2009) Escape from autologous neutralizing antibodies in acute/early subtype C HIV-1 infection requires multiple pathways. PLoS Pathog 5: e1000594.

52. Salazar-Gonzalez JF, Salazar MG, Keele BF, Learn GH, Giorgi EE, et al. (2009) Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. J Exp Med 206: 1273-1289.

53. Hladik F, McElrath MJ (2008) Setting the stage: host invasion by HIV. Nat Rev Immunol 8: 447-457.

54. McElrath MJ, De Rosa SC, Moodie Z, Dubey S, Kierstead L, et al. (2008) HIV-1 vaccine-induced immunity in the test-of-concept Step Study: a case-cohort analysis. Lancet 372: 1894-1905.

55. Dunkle KL, Stephenson R, Karita E, Chomba E, Kayitenkore K, et al. (2008) New heterosexually transmitted HIV infections in married or cohabiting couples in urban Zambia and Rwanda: an analysis of survey and clinical data. Lancet 371: 2183-2191.

56. Chomba E, Allen S, Kanweka W, Tichacek A, Cox G, et al. (2008) Evolution of couples' voluntary counseling and testing for HIV in Lusaka, Zambia. J Acquir Immune Defic Syndr 47: 108-115.

57. Boily MC, Baggaley RF, Wang L, Masse B, White RG, et al. (2009) Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. Lancet Infect Dis 9: 118-129.

58. Shaw GM, Hunter E (2012) HIV transmission. Cold Spring Harb Perspect Med 2.

59. Galvin SR, Cohen MS (2004) The role of sexually transmitted diseases in HIV transmission. Nat Rev Microbiol 2: 33-42.

60. Fideli US, Allen SA, Musonda R, Trask S, Hahn BH, et al. (2001) Virologic and immunologic determinants of heterosexual transmission of human immunodeficiency virus type 1 in Africa. AIDS Res Hum Retroviruses 17: 901-910.

61. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, et al. (2000) Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. N Engl J Med 342: 921-929.

62. Miller WC, Rosenberg NE, Rutstein SE, Powers KA (2010) Role of acute and early HIV infection in the sexual transmission of HIV. Curr Opin HIV AIDS 5: 277-282.

63. El-Sadr WM, Mayer KH, Hodder SL (2010) AIDS in America--forgotten but not gone. N Engl J Med 362: 967-970.

64. Dorak MT, Tang J, Penman-Aguilar A, Westfall AO, Zulu I, et al. (2004) Transmission of HIV-1 and HLA-B allele-sharing within serodiscordant heterosexual Zambian couples. Lancet 363: 2137-2139.

65. Tang J, Shao W, Yoo YJ, Brill I, Mulenga J, et al. (2008) Human leukocyte antigen class I genotypes in relation to heterosexual HIV type 1 transmission within discordant couples. J Immunol 181: 2626-2635.

66. Quinn TC (2007) Circumcision and HIV transmission. Curr Opin Infect Dis 20: 33-38.

67. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, et al. (2011) Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 365: 493-505.

68. Derdeyn CA, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, et al. (2004) Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. Science 303: 2019-2022.

69. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, et al. (2008) Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc Natl Acad Sci U S A 105: 7552-7557.

70. Keele BF, Li H, Learn GH, Hraber P, Giorgi EE, et al. (2009) Low-dose rectal inoculation of rhesus macaques by SIVsmE660 or SIVmac251 recapitulates human mucosal infection by HIV-1. J Exp Med 206: 1117-1134.

71. Derdeyn CA, Hunter E (2008) Viral characteristics of transmitted HIV. Curr Opin HIV AIDS 3: 16-21.

72. Keele BF (2010) Identifying and characterizing recently transmitted viruses. Curr Opin HIV AIDS 5: 327-334.

73. Boeras DI, Hraber PT, Hurlston M, Evans-Strickfaden T, Bhattacharya T, et al. (2011) Role of donor genital tract HIV-1 diversity in the transmission bottleneck. Proc Natl Acad Sci U S A 108: E1156-1163.

74. Chohan B, Lang D, Sagar M, Korber B, Lavreys L, et al. (2005) Selection for human immunodeficiency virus type 1 envelope glycosylation variants with shorter V1-V2 loop sequences occurs during transmission of certain genetic subtypes and may impact viral RNA levels. J Virol 79: 6528-6531.

75. Haaland RE, Hawkins PA, Salazar-Gonzalez J, Johnson A, Tichacek A, et al. (2009) Inflammatory genital infections mitigate a severe genetic bottleneck in heterosexual transmission of subtype A and C HIV-1. PLoS Pathog 5: e1000274.

76. Ping LH, Joseph SB, Anderson JA, Abrahams MR, Salazar-Gonzalez JF, et al. (2013) Comparison of viral Env proteins from acute and chronic infections with subtype C human immunodeficiency virus type 1 identifies differences in glycosylation

and CCR5 utilization and suggests a new strategy for immunogen design. J Virol 87: 7218-7233.

77. Sagar M, Laeyendecker O, Lee S, Gamiel J, Wawer MJ, et al. (2009) Selection of HIV variants with signature genotypic characteristics during heterosexual transmission. J Infect Dis 199: 580-589.

78. Isaacman-Beck J, Hermann EA, Yi Y, Ratcliffe SJ, Mulenga J, et al. (2009) Heterosexual transmission of human immunodeficiency virus type 1 subtype C: Macrophage tropism, alternative coreceptor use, and the molecular anatomy of CCR5 utilization. J Virol 83: 8208-8220.

79. Parrish NF, Gao F, Li H, Giorgi EE, Barbian HJ, et al. (2013) Phenotypic properties of transmitted founder HIV-1. Proc Natl Acad Sci U S A 110: 6626-6633.

80. Liu L (2014) Book Review: Fields Virology, 6th Edition. Clin Infect Dis.

81. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, et al. (1996) Identification of a major co-receptor for primary isolates of HIV-1. Nature 381: 661-666.

82. Feng Y, Broder CC, Kennedy PE, Berger EA (1996) HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science 272: 872-877.

83. Bosch ML, Earl PL, Fargnoli K, Picciafuoco S, Giombini F, et al. (1989) Identification of the fusion peptide of primate immunodeficiency viruses. Science 244: 694-697.

84. Wyatt R, Sodroski J (1998) The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. Science 280: 1884-1888.

85. Sundquist WI, Krausslich HG (2012) HIV-1 assembly, budding, and maturation. Cold Spring Harb Perspect Med 2: a006924.

86. Resh MD (2005) Intracellular trafficking of HIV-1 Gag: how Gag interacts with cell membranes and makes viral particles. AIDS Rev 7: 84-91.

87. Dorfman T, Mammano F, Haseltine WA, Gottlinger HG (1994) Role of the matrix protein in the virion association of the human immunodeficiency virus type 1 envelope glycoprotein. J Virol 68: 1689-1696.

88. Freed EO, Martin MA (1995) Virion incorporation of envelope glycoproteins with long but not short cytoplasmic tails is blocked by specific, single amino acid substitutions in the human immunodeficiency virus type 1 matrix. J Virol 69: 1984-1989.

89. Spearman P, Horton R, Ratner L, Kuli-Zade I (1997) Membrane binding of human immunodeficiency virus type 1 matrix protein in vivo supports a conformational myristyl switch mechanism. J Virol 71: 6582-6592.

90. Freed EO (1998) HIV-1 gag proteins: diverse functions in the virus life cycle. Virology 251: 1-15.

91. Ganser-Pornillos BK, von Schwedler UK, Stray KM, Aiken C, Sundquist WI (2004) Assembly properties of the human immunodeficiency virus type 1 CA protein. J Virol 78: 2545-2552.

92. Darlix JL, Godet J, Ivanyi-Nagy R, Fosse P, Mauffret O, et al. (2011) Flexible nature and specific functions of the HIV-1 nucleocapsid protein. J Mol Biol 410: 565-581.

93. Saadatmand J, Kleiman L (2012) Aspects of HIV-1 assembly that promote primer tRNA(Lys3) annealing to viral RNA. Virus Res 169: 340-348.

94. Wright ER, Schooler JB, Ding HJ, Kieffer C, Fillmore C, et al. (2007) Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. EMBO J 26: 2218-2226.

95. Gottlinger HG, Dorfman T, Sodroski JG, Haseltine WA (1991) Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. Proc Natl Acad Sci U S A 88: 3195-3199.

96. Weiss ER, Gottlinger H (2011) The role of cellular factors in promoting HIV budding. J Mol Biol 410: 525-533.

97. Klatzmann D, Barre-Sinoussi F, Nugeyre MT, Danquet C, Vilmer E, et al. (1984) Selective tropism of lymphadenopathy associated virus (LAV) for helper-inducer T lymphocytes. Science 225: 59-63.

98. Masur H, Ognibene FP, Yarchoan R, Shelhamer JH, Baird BF, et al. (1989) CD4 counts as predictors of opportunistic pneumonias in human immunodeficiency virus (HIV) infection. Ann Intern Med 111: 223-231.

99. Munoz A, Wang MC, Bass S, Taylor JM, Kingsley LA, et al. (1989) Acquired immunodeficiency syndrome (AIDS)-free time after human immunodeficiency virus type 1 (HIV-1) seroconversion in homosexual men. Multicenter AIDS Cohort Study Group. Am J Epidemiol 130: 530-539.

100. Small CB, Klein RS, Friedland GH, Moll B, Emeson EE, et al. (1983) Community-acquired opportunistic infections and defective cellular immunity in heterosexual drug abusers and homosexual men. Am J Med 74: 433-441.

101. Chun TW, Finzi D, Margolick J, Chadwick K, Schwartz D, et al. (1995) In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. Nat Med 1: 1284-1290.

102. Chun TW, Justement JS, Murray D, Hallahan CW, Maenza J, et al. (2010) Rebound of plasma viremia following cessation of antiretroviral therapy despite profoundly low levels of HIV reservoir: implications for eradication. AIDS 24: 2803-2808.

103. Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, et al. (1997) Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. Science 278: 1295-1300.

104. Baker JV, Peng G, Rapkin J, Krason D, Reilly C, et al. (2008) Poor initial CD4+ recovery with antiretroviral therapy prolongs immune depletion and increases risk for AIDS and non-AIDS diseases. J Acquir Immune Defic Syndr 48: 541-546.

105. Battegay M, Nuesch R, Hirschel B, Kaufmann GR (2006) Immunological recovery and antiretroviral therapy in HIV-1 infection. Lancet Infect Dis 6: 280-287.

106. Thompson MA, Aberg JA, Cahn P, Montaner JS, Rizzardini G, et al. (2010) Antiretroviral treatment of adult HIV infection: 2010 recommendations of the International AIDS Society-USA panel. JAMA 304: 321-333.

107. Chahroudi A, Bosinger SE, Vanderford TH, Paiardini M, Silvestri G (2012) Natural SIV hosts: showing AIDS the door. Science 335: 1188-1193.

108. Silvestri G (2013) Embracing the complexity of HIV immunology. Immunol Rev 254: 5-9.

109. Brenchley JM, Douek DC (2008) HIV infection and the gastrointestinal immune system. Mucosal Immunol 1: 23-30.

110. Nishimura Y, Brown CR, Mattapallil JJ, Igarashi T, Buckler-White A, et al. (2005) Resting naive CD4+ T cells are massively infected and eliminated by X4-tropic

simian-human immunodeficiency viruses in macaques. Proc Natl Acad Sci U S A 102: 8000-8005.

111. Reimann KA, Watson A, Dailey PJ, Lin W, Lord CI, et al. (1999) Viral burden and disease progression in rhesus monkeys infected with chimeric simian-human immunodeficiency viruses. Virology 256: 15-21.

112. Gordon SN, Klatt NR, Bosinger SE, Brenchley JM, Milush JM, et al. (2007) Severe depletion of mucosal CD4+ T cells in AIDS-free simian immunodeficiency virus-infected sooty mangabeys. J Immunol 179: 3026-3034.

113. Meythaler M, Wang Z, Martinot A, Pryputniewicz S, Kasheta M, et al. (2011) Early induction of polyfunctional simian immunodeficiency virus (SIV)-specific T lymphocytes and rapid disappearance of SIV from lymph nodes of sooty mangabeys during primary infection. J Immunol 186: 5151-5161.

114. Pandrea IV, Gautam R, Ribeiro RM, Brenchley JM, Butler IF, et al. (2007) Acute loss of intestinal CD4+ T cells is not predictive of simian immunodeficiency virus virulence. J Immunol 179: 3035-3046.

115. Douek DC, Picker LJ, Koup RA (2003) T cell dynamics in HIV-1 infection. Annu Rev Immunol 21: 265-304.

116. Doms RW (2001) Chemokine receptors and HIV entry. AIDS 15 Suppl 1: S34-35.

117. Grivel JC, Penn ML, Eckstein DA, Schramm B, Speck RF, et al. (2000) Human immunodeficiency virus type 1 coreceptor preferences determine target T-cell depletion and cellular tropism in human lymphoid tissue. J Virol 74: 5347-5351.

118. Nishimura Y, Igarashi T, Donau OK, Buckler-White A, Buckler C, et al. (2004) Highly pathogenic SHIVs and SIVs target different CD4+ T cell subsets in rhesus monkeys, explaining their divergent clinical courses. Proc Natl Acad Sci U S A 101: 12324-12329.

119. Qin S, Rottman JB, Myers P, Kassam N, Weinblatt M, et al. (1998) The chemokine receptors CXCR3 and CCR5 mark subsets of T cells associated with certain inflammatory reactions. J Clin Invest 101: 746-754.

120. Biancotto A, Iglehart SJ, Vanpouille C, Condack CE, Lisco A, et al. (2008) HIV-1 induced activation of CD4+ T cells creates new targets for HIV-1 infection in human lymphoid tissue ex vivo. Blood 111: 699-704.

121. Maier R, Bartolome-Rodriguez MM, Moulon C, Weltzien HU, Meyerhans A (2000) Kinetics of CXCR4 and CCR5 up-regulation and human immunodeficiency virus expansion after antigenic stimulation of primary CD4(+) T lymphocytes. Blood 96: 1853-1856.

122. Zaunders JJ, Munier ML, Kaufmann DE, Ip S, Grey P, et al. (2005) Early proliferation of CCR5(+) CD38(+++) antigen-specific CD4(+) Th1 effector cells during primary HIV-1 infection. Blood 106: 1660-1667.

123. Geldmacher C, Ngwenyama N, Schuetz A, Petrovas C, Reither K, et al. (2010) Preferential infection and depletion of Mycobacterium tuberculosis-specific CD4 T cells after HIV-1 infection. J Exp Med 207: 2869-2881.

124. Bronke C, Palmer NM, Jansen CA, Westerlaken GH, Polstra AM, et al. (2005) Dynamics of cytomegalovirus (CMV)-specific T cells in HIV-1-infected individuals progressing to AIDS with CMV end-organ disease. J Infect Dis 191: 873-880.

125. Casazza JP, Brenchley JM, Hill BJ, Ayana R, Ambrozak D, et al. (2009) Autocrine production of beta-chemokines protects CMV-Specific CD4 T cells from HIV infection. PLoS Pathog 5: e1000646.

126. Geldmacher C, Koup RA (2012) Pathogen-specific T cell depletion and reactivation of opportunistic pathogens in HIV infection. Trends Immunol 33: 207-214.

127. Saharia KK, Koup RA (2013) T cell susceptibility to HIV influences outcome of opportunistic infections. Cell 155: 505-514.

128. Brenchley JM, Ruff LE, Casazza JP, Koup RA, Price DA, et al. (2006) Preferential infection shortens the life span of human immunodeficiency virus-specific CD4+ T cells in vivo. J Virol 80: 6801-6809.

129. Chattopadhyay PK, Roederer M (2005) Immunophenotyping of T cell subpopulations in HIV disease. Curr Protoc Immunol Chapter 12: Unit 12 12.

130. Chattopadhyay PK, Roederer M (2010) Good cell, bad cell: flow cytometry reveals T-cell subsets important in HIV disease. Cytometry A 77: 614-622.

131. Lanzavecchia A, Sallusto F (2005) Understanding the generation and function of memory T cell subsets. Curr Opin Immunol 17: 326-332.

132. Sallusto F, Geginat J, Lanzavecchia A (2004) Central memory and effector memory T cell subsets: function, generation, and maintenance. Annu Rev Immunol 22: 745-763.

133. Picker LJ, Reed-Inderbitzin EF, Hagen SI, Edgar JB, Hansen SG, et al. (2006) IL-15 induces CD4 effector memory T cell production and tissue emigration in nonhuman primates. J Clin Invest 116: 1514-1524.

134. Veazey RS, Mansfield KG, Tham IC, Carville AC, Shvetz DE, et al. (2000) Dynamics of CCR5 expression by CD4(+) T cells in lymphoid tissues during simian immunodeficiency virus infection. J Virol 74: 11001-11007.

135. Scarlatti G, Tresoldi E, Bjorndal A, Fredriksson R, Colognesi C, et al. (1997) In vivo evolution of HIV-1 co-receptor usage and sensitivity to chemokine-mediated suppression. Nat Med 3: 1259-1265.

136. Schuitemaker H, Koot M, Kootstra NA, Dercksen MW, de Goede RE, et al. (1992) Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus population. J Virol 66: 1354-1360.

137. Mehandru S, Poles MA, Tenner-Racz K, Horowitz A, Hurley A, et al. (2004) Primary HIV-1 infection is associated with preferential depletion of CD4+ T lymphocytes from effector sites in the gastrointestinal tract. J Exp Med 200: 761-770.

138. Veazey RS, DeMaria M, Chalifoux LV, Shvetz DE, Pauley DR, et al. (1998) Gastrointestinal tract as a major site of CD4+ T cell depletion and viral replication in SIV infection. Science 280: 427-431.

139. Picker LJ, Hagen SI, Lum R, Reed-Inderbitzin EF, Daly LM, et al. (2004) Insufficient production and tissue delivery of CD4+ memory T cells in rapidly progressive simian immunodeficiency virus infection. J Exp Med 200: 1299-1314.

140. Okoye A, Meier-Schellersheim M, Brenchley JM, Hagen SI, Walker JM, et al. (2007) Progressive CD4+ central memory T cell decline results in CD4+ effector memory insufficiency and overt disease in chronic SIV infection. J Exp Med 204: 2171-2185.

141. Cao J, Park IW, Cooper A, Sodroski J (1996) Molecular determinants of acute single-cell lysis by human immunodeficiency virus type 1. J Virol 70: 1340-1354.

142. Douek DC (2003) Disrupting T-cell homeostasis: how HIV-1 infection causes disease. AIDS Rev 5: 172-177.

143. Lenardo MJ, Angleman SB, Bounkeua V, Dimas J, Duvall MG, et al. (2002) Cytopathic killing of peripheral blood CD4(+) T lymphocytes by human

immunodeficiency virus type 1 appears necrotic rather than apoptotic and does not require env. J Virol 76: 5082-5093.

144. Picker LJ (2006) Immunopathogenesis of acute AIDS virus infection. Curr Opin Immunol 18: 399-405.

145. Zeng M, Smith AJ, Wietgrefe SW, Southern PJ, Schacker TW, et al. (2011) Cumulative mechanisms of lymphoid tissue fibrosis and T cell depletion in HIV-1 and SIV infections. J Clin Invest 121: 998-1008.

146. Boulassel MR, Chomont N, Pai NP, Gilmore N, Sekaly RP, et al. (2012) CD4 T cell nadir independently predicts the magnitude of the HIV reservoir after prolonged suppressive antiretroviral therapy. J Clin Virol 53: 29-32.

147. Chomont N, El-Far M, Ancuta P, Trautmann L, Procopio FA, et al. (2009) HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. Nat Med 15: 893-900.

148. Letvin NL, Mascola JR, Sun Y, Gorgone DA, Buzby AP, et al. (2006) Preserved CD4+ central memory T cells and survival in vaccinated SIV-challenged monkeys. Science 312: 1530-1533.

149. Mattapallil JJ, Douek DC, Buckler-White A, Montefiori D, Letvin NL, et al. (2006) Vaccination preserves CD4 memory T cells during acute simian immunodeficiency virus challenge. J Exp Med 203: 1533-1541.

150. Paiardini M, Cervasi B, Reyes-Aviles E, Micci L, Ortiz AM, et al. (2011) Low levels of SIV infection in sooty mangabey central memory CD(4)(+) T cells are associated with limited CCR5 expression. Nat Med 17: 830-836.

151. Pandrea I, Apetrei C, Gordon S, Barbercheck J, Dufour J, et al. (2007) Paucity of CD4+CCR5+ T cells is a typical feature of natural SIV hosts. Blood 109: 1069-1076.

152. Deeks SG, Kitchen CM, Liu L, Guo H, Gascon R, et al. (2004) Immune activation set point during early HIV infection predicts subsequent CD4+ T-cell changes independent of viral load. Blood 104: 942-947.

153. Giorgi JV, Hultin LE, McKeating JA, Johnson TD, Owens B, et al. (1999) 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. J Infect Dis 179: 859-870.

154. Hunt PW, Brenchley J, Sinclair E, McCune JM, Roland M, et al. (2008) Relationship between T cell activation and CD4+ T cell count in HIV-seropositive individuals with undetectable plasma HIV RNA levels in the absence of therapy. J Infect Dis 197: 126-133.

155. Douek DC, Roederer M, Koup RA (2009) Emerging concepts in the immunopathogenesis of AIDS. Annu Rev Med 60: 471-484.

156. Hunt PW, Martin JN, Sinclair E, Bredt B, Hagos E, et al. (2003) T cell activation is associated with lower CD4+ T cell gains in human immunodeficiency virus-infected patients with sustained viral suppression during antiretroviral therapy. J Infect Dis 187: 1534-1543.

157. Marchetti G, Gori A, Casabianca A, Magnani M, Franzetti F, et al. (2006) Comparative analysis of T-cell turnover and homeostatic parameters in HIV-infected patients with discordant immune-virological responses to HAART. AIDS 20: 1727-1736.

158. Broussard SR, Staprans SI, White R, Whitehead EM, Feinberg MB, et al. (2001) Simian immunodeficiency virus replicates to high levels in naturally infected African green monkeys without inducing immunologic or neurologic disease. J Virol 75: 2262-2275.

159. Silvestri G, Sodora DL, Koup RA, Paiardini M, O'Neil SP, et al. (2003) Nonpathogenic SIV infection of sooty mangabeys is characterized by limited bystander immunopathology despite chronic high-level viremia. Immunity 18: 441-452.

160. Lane HC, Masur H, Edgar LC, Whalen G, Rook AH, et al. (1983) Abnormalities of B-cell activation and immunoregulation in patients with the acquired immunodeficiency syndrome. N Engl J Med 309: 453-458.

161. Hazenberg MD, Stuart JW, Otto SA, Borleffs JC, Boucher CA, et al. (2000) T-cell division in human immunodeficiency virus (HIV)-1 infection is mainly due to

immune activation: a longitudinal analysis in patients before and during highly active antiretroviral therapy (HAART). Blood 95: 249-255.

162. Hellerstein M, Hanley MB, Cesar D, Siler S, Papageorgopoulos C, et al. (1999) Directly measured kinetics of circulating T lymphocytes in normal and HIV-1-infected humans. Nat Med 5: 83-89.

163. Valdez H, Lederman MM (1997) Cytokines and cytokine therapies in HIV infection. AIDS Clin Rev: 187-228.

164. Connolly NC, Riddler SA, Rinaldo CR (2005) Proinflammatory cytokines in HIV disease-a review and rationale for new therapeutic approaches. AIDS Rev 7: 168-180.

165. Klatt NR, Estes JD, Sun X, Ortiz AM, Barber JS, et al. (2012) Loss of mucosal CD103+ DCs and IL-17+ and IL-22+ lymphocytes is associated with mucosal damage in SIV infection. Mucosal Immunol 5: 646-657.

166. Kwa S, Kannanganat S, Nigam P, Siddiqui M, Shetty RD, et al. (2011) Plasmacytoid dendritic cells are recruited to the colorectum and contribute to immune activation during pathogenic SIV infection in rhesus macaques. Blood 118: 2763-2773.

167. Reeves RK, Evans TI, Gillis J, Wong FE, Kang G, et al. (2012) SIV infection induces accumulation of plasmacytoid dendritic cells in the gut mucosa. J Infect Dis 206: 1462-1468.

168. Wallet MA, Rodriguez CA, Yin L, Saporta S, Chinratanapisit S, et al. (2010) Microbial translocation induces persistent macrophage activation unrelated to HIV-1 levels or T-cell activation following therapy. AIDS 24: 1281-1290.

169. Reeves RK, Gillis J, Wong FE, Yu Y, Connole M, et al. (2010) CD16- natural killer cells: enrichment in mucosal and secondary lymphoid tissues and altered function during chronic SIV infection. Blood 115: 4439-4446.

170. Giorgi JV, Detels R (1989) T-cell subset alterations in HIV-infected homosexual men: NIAID Multicenter AIDS cohort study. Clin Immunol Immunopathol 52: 10-18.

171. Giorgi JV, Ho HN, Hirji K, Chou CC, Hultin LE, et al. (1994) CD8+ lymphocyte activation at human immunodeficiency virus type 1 seroconversion: development of HLA-DR+ CD38- CD8+ cells is associated with subsequent stable CD4+ cell levels. The Multicenter AIDS Cohort Study Group. J Infect Dis 170: 775-781.

172. Dianzani U, Funaro A, DiFranco D, Garbarino G, Bragardo M, et al. (1994) Interaction between endothelium and CD4+CD45RA+ lymphocytes. Role of the human CD38 molecule. J Immunol 153: 952-959.

173. Liu Z, Cumberland WG, Hultin LE, Prince HE, Detels R, et al. (1997) Elevated CD38 antigen expression on CD8+ T cells is a stronger marker for the risk of chronic HIV disease progression to AIDS and death in the Multicenter AIDS Cohort Study than CD4+ cell count, soluble immune activation markers, or combinations

of HLA-DR and CD38 expression. J Acquir Immune Defic Syndr Hum Retrovirol 16: 83-92.

174. Liu Z, Hultin LE, Cumberland WG, Hultin P, Schmid I, et al. (1996) Elevated relative fluorescence intensity of CD38 antigen expression on CD8+ T cells is a marker of poor prognosis in HIV infection: results of 6 years of follow-up. Cytometry 26: 1-7.

175. Orendi JM, Bloem AC, Borleffs JC, Wijnholds FJ, de Vos NM, et al. (1998) Activation and cell cycle antigens in CD4+ and CD8+ T cells correlate with plasma human immunodeficiency virus (HIV-1) RNA level in HIV-1 infection. J Infect Dis 178: 1279-1287.

176. Klatt NR, Funderburg NT, Brenchley JM (2013) Microbial translocation, immune activation, and HIV disease. Trends Microbiol 21: 6-13.

177. Ancuta P, Kamat A, Kunstman KJ, Kim EY, Autissier P, et al. (2008) Microbial translocation is associated with increased monocyte activation and dementia in AIDS patients. PLoS One 3: e2516.

178. Deeks SG (2011) HIV infection, inflammation, immunosenescence, and aging. Annu Rev Med 62: 141-155.

179. Pandrea I, Cornell E, Wilson C, Ribeiro RM, Ma D, et al. (2012) Coagulation biomarkers predict disease progression in SIV-infected nonhuman primates. Blood 120: 1357-1366.

180. Herbeuval JP, Hardy AW, Boasso A, Anderson SA, Dolan MJ, et al. (2005) Regulation of TNF-related apoptosis-inducing ligand on primary CD4+ T cells by HIV-1: role of type I IFN-producing plasmacytoid dendritic cells. Proc Natl Acad Sci U S A 102: 13974-13979.

181. Markle TJ, Philip M, Brockman MA (2013) HIV-1 Nef and T-cell activation: a history of contradictions. Future Virol 8.

182. Schindler M, Munch J, Kutsch O, Li H, Santiago ML, et al. (2006) Nef-mediated suppression of T cell activation was lost in a lentiviral lineage that gave rise to HIV-1. Cell 125: 1055-1067.

183. Lee B (2013) HIV provides ample PAMPs for innate immune sensing. Proc Natl Acad Sci U S A 110: 19183-19184.

184. Meier A, Alter G, Frahm N, Sidhu H, Li B, et al. (2007) MyD88-dependent immune activation mediated by human immunodeficiency virus type 1-encoded Toll-like receptor ligands. J Virol 81: 8180-8191.

185. Simmons RP, Scully EP, Groden EE, Arnold KB, Chang JJ, et al. (2013) HIV-1 infection induces strong production of IP-10 through TLR7/9-dependent pathways. AIDS 27: 2505-2517.

186. Dixit E, Kagan JC (2013) Intracellular pathogen detection by RIG-I-like receptors. Adv Immunol 117: 99-125.

187. Stetson DB, Medzhitov R (2006) Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. Immunity 24: 93-103.

188. Galao RP, Le Tortorec A, Pickering S, Kueck T, Neil SJ (2012) Innate sensing of HIV-1 assembly by Tetherin induces NFkappaB-dependent proinflammatory responses. Cell Host Microbe 12: 633-644.

189. Luban J (2012) Innate immune sensing of HIV-1 by dendritic cells. Cell Host Microbe 12: 408-418.

190. Tokarev A, Suarez M, Kwan W, Fitzpatrick K, Singh R, et al. (2013) Stimulation of NF-kappaB activity by the HIV restriction factor BST2. J Virol 87: 2046-2057.

191. Uchil PD, Hinz A, Siegel S, Coenen-Stass A, Pertel T, et al. (2013) TRIM protein-mediated regulation of inflammatory and innate immune signaling and its association with antiretroviral activity. J Virol 87: 257-272.

192. Finkel TH, Tudor-Williams G, Banda NK, Cotton MF, Curiel T, et al. (1995) Apoptosis occurs predominantly in bystander cells and not in productively infected cells of HIV- and SIV-infected lymph nodes. Nat Med 1: 129-134.

193. Gougeon ML, Lecoeur H, Dulioust A, Enouf MG, Crouvoiser M, et al. (1996) Programmed cell death in peripheral lymphocytes from HIV-infected persons: increased susceptibility to apoptosis of CD4 and CD8 T cells correlates with lymphocyte activation and with disease progression. J Immunol 156: 3509-3520.

194. Doitsh G, Galloway NL, Geng X, Yang Z, Monroe KM, et al. (2014) Cell death by pyroptosis drives CD4 T-cell depletion in HIV-1 infection. Nature 505: 509-514.

195. Fink SL, Cookson BT (2005) Apoptosis, pyroptosis, and necrosis: mechanistic description of dead and dying eukaryotic cells. Infect Immun 73: 1907-1916.

196. Okoye AA, Picker LJ (2013) CD4(+) T-cell depletion in HIV infection: mechanisms of immunological failure. Immunol Rev 254: 54-64.

197. Brenchley JM, Price DA, Schacker TW, Asher TE, Silvestri G, et al. (2006) Microbial translocation is a cause of systemic immune activation in chronic HIV infection. Nat Med 12: 1365-1371.

198. Brenchley JM, Douek DC (2012) Microbial translocation across the GI tract. Annu Rev Immunol 30: 149-173.

199. Wolochow H, Hildebrand GJ, Lamanna C (1966) Translocation of microorganisms across the intestinal wall of the rat: effect of microbial size and concentration. J Infect Dis 116: 523-528.

200. Bukh AR, Melchjorsen J, Offersen R, Jensen JM, Toft L, et al. (2011) Endotoxemia is associated with altered innate and adaptive immune responses in untreated HIV-1 infected individuals. PLoS One 6: e21275.

201. Marchetti G, Tincati C, Silvestri G (2013) Microbial translocation in the pathogenesis of HIV infection and AIDS. Clin Microbiol Rev 26: 2-18.

202. Sandler NG, Wand H, Roque A, Law M, Nason MC, et al. (2011) Plasma levels of soluble CD14 independently predict mortality in HIV infection. J Infect Dis 203: 780-790.

203. Kelesidis T, Kendall MA, Yang OO, Hodis HN, Currier JS (2012) Biomarkers of microbial translocation and macrophage activation: association with progression of subclinical atherosclerosis in HIV-1 infection. J Infect Dis 206: 1558-1567.

204. Wurfel MM, Hailman E, Wright SD (1995) Soluble CD14 acts as a shuttle in the neutralization of lipopolysaccharide (LPS) by LPS-binding protein and reconstituted high density lipoprotein. J Exp Med 181: 1743-1754.

205. Pelsers MM, Namiot Z, Kisielewski W, Namiot A, Januszkiewicz M, et al. (2003) Intestinal-type and liver-type fatty acid-binding protein in the intestine. Tissue distribution and clinical utility. Clin Biochem 36: 529-535.

206. Brenchley JM, Paiardini M, Knox KS, Asher AI, Cervasi B, et al. (2008) Differential Th17 CD4 T-cell depletion in pathogenic and nonpathogenic lentiviral infections. Blood 112: 2826-2835.

207. Favre D, Lederer S, Kanwar B, Ma ZM, Proll S, et al. (2009) Critical loss of the balance between Th17 and T regulatory cell populations in pathogenic SIV infection. PLoS Pathog 5: e1000295.

208. Pandrea I, Gaufin T, Brenchley JM, Gautam R, Monjure C, et al. (2008) Cutting edge: Experimentally induced immune activation in natural hosts of simian

immunodeficiency virus induces significant increases in viral replication and CD4+ T cell depletion. J Immunol 181: 6687-6691.

209. Cecchinato V, Trindade CJ, Laurence A, Heraud JM, Brenchley JM, et al. (2008) Altered balance between Th17 and Th1 cells at mucosal sites predicts AIDS progression in simian immunodeficiency virus-infected macaques. Mucosal Immunol 1: 279-288.

210. Estes JD, Harris LD, Klatt NR, Tabb B, Pittaluga S, et al. (2010) Damaged intestinal epithelial integrity linked to microbial translocation in pathogenic simian immunodeficiency virus infections. PLoS Pathog 6: e1001052.

211. Huber S, Gagliani N, Flavell RA (2012) Life, death, and miracles: Th17 cells in the intestine. Eur J Immunol 42: 2238-2245.

212. Micci L, Cervasi B, Ende ZS, Iriele RI, Reyes-Aviles E, et al. (2012) Paucity of IL-21-producing CD4(+) T cells is associated with Th17 cell depletion in SIV infection of rhesus macaques. Blood 120: 3925-3935.

213. Yi JS, Ingram JT, Zajac AJ (2010) IL-21 deficiency influences CD8 T cell quality and recall responses following an acute viral infection. J Immunol 185: 4835-4845.

214. Korn T, Bettelli E, Gao W, Awasthi A, Jager A, et al. (2007) IL-21 initiates an alternative pathway to induce proinflammatory T(H)17 cells. Nature 448: 484-487.

215. Ozaki K, Spolski R, Feng CG, Qi CF, Cheng J, et al. (2002) A critical role for IL-21 in regulating immunoglobulin production. Science 298: 1630-1634.

216. Parrish-Novak J, Dillon SR, Nelson A, Hammond A, Sprecher C, et al. (2000) Interleukin 21 and its receptor are involved in NK cell expansion and regulation of lymphocyte function. Nature 408: 57-63.

217. Pallikkuth S, Micci L, Ende ZS, Iriele RI, Cervasi B, et al. (2013) Maintenance of intestinal Th17 cells and reduced microbial translocation in SIV-infected rhesus macaques treated with interleukin (IL)-21. PLoS Pathog 9: e1003471.

218. Masopust D, Vezys V, Wherry EJ, Ahmed R (2007) A brief history of CD8 T cells. Eur J Immunol 37 Suppl 1: S103-110.

219. Virgin HW, Wherry EJ, Ahmed R (2009) Redefining chronic viral infection. Cell 138: 30-50.

220. Shin H, Wherry EJ (2007) CD8 T cell dysfunction during chronic viral infection. Curr Opin Immunol 19: 408-415.

221. Barber DL, Wherry EJ, Masopust D, Zhu B, Allison JP, et al. (2006) Restoring function in exhausted CD8 T cells during chronic viral infection. Nature 439: 682-687.

222. Maier H, Isogawa M, Freeman GJ, Chisari FV (2007) PD-1:PD-L1 interactions contribute to the functional suppression of virus-specific CD8+ T lymphocytes in the liver. J Immunol 178: 2714-2720.

223. Zhang Z, Zhang JY, Wherry EJ, Jin B, Xu B, et al. (2008) Dynamic programmed death 1 expression by virus-specific CD8 T cells correlates with the outcome of acute hepatitis B. Gastroenterology 134: 1938-1949, 1949 e1931-1933.

224. Day CL, Kaufmann DE, Kiepiela P, Brown JA, Moodley ES, et al. (2006) PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. Nature 443: 350-354.

225. Streeck H, Brumme ZL, Anastario M, Cohen KW, Jolin JS, et al. (2008) Antigen load and viral sequence diversification determine the functional profile of HIV-1-specific CD8+ T cells. PLoS Med 5: e100.

226. Trautmann L, Janbazian L, Chomont N, Said EA, Gimmig S, et al. (2006) Upregulation of PD-1 expression on HIV-specific CD8+ T cells leads to reversible immune dysfunction. Nat Med 12: 1198-1202.

227. Draenert R, Verrill CL, Tang Y, Allen TM, Wurcel AG, et al. (2004) Persistent recognition of autologous virus by high-avidity CD8 T cells in chronic, progressive human immunodeficiency virus type 1 infection. J Virol 78: 630-641.

228. Champagne P, Ogg GS, King AS, Knabenhans C, Ellefsen K, et al. (2001) Skewed maturation of memory HIV-specific CD8 T lymphocytes. Nature 410: 106-111.

229. Migueles SA, Laborico AC, Shupert WL, Sabbaghian MS, Rabin R, et al. (2002) HIV-specific CD8+ T cell proliferation is coupled to perforin expression and is maintained in nonprogressors. Nat Immunol 3: 1061-1068.

230. Shankar P, Russo M, Harnisch B, Patterson M, Skolnik P, et al. (2000) Impaired function of circulating HIV-specific CD8(+) T cells in chronic human immunodeficiency virus infection. Blood 96: 3094-3101.

231. Zhang D, Shankar P, Xu Z, Harnisch B, Chen G, et al. (2003) Most antiviral CD8 T cells during chronic viral infection do not express high levels of perforin and are not directly cytotoxic. Blood 101: 226-235.

232. Betts MR, Nason MC, West SM, De Rosa SC, Migueles SA, et al. (2006) HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells. Blood 107: 4781-4789.

233. Chen L (2004) Co-inhibitory molecules of the B7-CD28 family in the control of T-cell immunity. Nat Rev Immunol 4: 336-347.

234. Porichis F, Kaufmann DE (2012) Role of PD-1 in HIV pathogenesis and as target for therapy. Curr HIV/AIDS Rep 9: 81-90.

235. Petrovas C, Price DA, Mattapallil J, Ambrozak DR, Geldmacher C, et al. (2007) SIV-specific CD8+ T cells express high levels of PD1 and cytokines but have impaired proliferative capacity in acute and chronic SIVmac251 infection. Blood 110: 928-936.

236. Velu V, Kannanganat S, Ibegbu C, Chennareddi L, Villinger F, et al. (2007) Elevated expression levels of inhibitory receptor programmed death 1 on simian immunodeficiency virus-specific CD8 T cells during chronic infection but not after vaccination. J Virol 81: 5819-5828.

237. Velu V, Titanji K, Zhu B, Husain S, Pladevega A, et al. (2009) Enhancing SIV-specific immunity in vivo by PD-1 blockade. Nature 458: 206-210.

238. Petrovas C, Chaon B, Ambrozak DR, Price DA, Melenhorst JJ, et al. (2009) Differential association of programmed death-1 and CD57 with ex vivo survival of CD8+ T cells in HIV infection. J Immunol 183: 1120-1132.

239. Lee SA, Sinclair E, Hatano H, Hsue PY, Epling L, et al. (2014) Impact of HIV on CD8+ T cell CD57 expression is distinct from that of CMV and aging. PLoS One 9: e89444.

240. Lee SA, Sinclair E, Jain V, Huang Y, Epling L, et al. (2014) Low Proportions of CD28- CD8+ T cells Expressing CD57 Can Be Reversed by Early ART Initiation and Predict Mortality in Treated HIV Infection. J Infect Dis.

241. Chun TW, Carruth L, Finzi D, Shen X, DiGiuseppe JA, et al. (1997) Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. Nature 387: 183-188.

242. Finzi D, Blankson J, Siliciano JD, Margolick JB, Chadwick K, et al. (1999) Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. Nat Med 5: 512-517.

243. Hermankova M, Siliciano JD, Zhou Y, Monie D, Chadwick K, et al. (2003) Analysis of human immunodeficiency virus type 1 gene expression in latently infected resting CD4+ T lymphocytes in vivo. J Virol 77: 7383-7392.

244. Ananworanich J, Schuetz A, Vandergeeten C, Sereti I, de Souza M, et al. (2012) Impact of multi-targeted antiretroviral treatment on gut T cell depletion and HIV reservoir seeding during acute HIV infection. PLoS One 7: e33948.

245. Strain MC, Little SJ, Daar ES, Havlir DV, Gunthard HF, et al. (2005) Effect of treatment, during primary infection, on establishment and clearance of cellular reservoirs of HIV-1. J Infect Dis 191: 1410-1418.

246. Cartwright EK, McGary CS, Cervasi B, Micci L, Lawson B, et al. (2014) Divergent CD4+ T Memory Stem Cell Dynamics in Pathogenic and Nonpathogenic Simian Immunodeficiency Virus Infections. J Immunol 192: 4666-4673.

247. Shastri N, Cardinaud S, Schwab SR, Serwold T, Kunisawa J (2005) All the peptides that fit: the beginning, the middle, and the end of the MHC class I antigen-processing pathway. Immunol Rev 207: 31-41.

248. Borrow P, Lewicki H, Hahn BH, Shaw GM, Oldstone MB (1994) Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. J Virol 68: 6103-6110.

249. Koup RA, Safrit JT, Cao Y, Andrews CA, McLeod G, et al. (1994) Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. J Virol 68: 4650-4655.

250. Jin X, Bauer DE, Tuttleton SE, Lewin S, Gettie A, et al. (1999) Dramatic rise in plasma viremia after CD8(+) T cell depletion in simian immunodeficiency virus-infected macaques. J Exp Med 189: 991-998.

251. Altfeld M, Addo MM, Rosenberg ES, Hecht FM, Lee PK, et al. (2003) Influence of HLA-B57 on clinical presentation and viral control during acute HIV-1 infection. AIDS 17: 2581-2591.

252. Costello C, Tang J, Rivers C, Karita E, Meizen-Derr J, et al. (1999) HLA-B*5703 independently associated with slower HIV-1 disease progression in Rwandan women. AIDS 13: 1990-1991.

253. Kaslow RA, Carrington M, Apple R, Park L, Munoz A, et al. (1996) Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. Nat Med 2: 405-411.

254. Migueles SA, Sabbaghian MS, Shupert WL, Bettinotti MP, Marincola FM, et al. (2000) HLA B*5701 is highly associated with restriction of virus replication in a

subgroup of HIV-infected long term nonprogressors. Proc Natl Acad Sci U S A 97: 2709-2714.

255. Tang J, Kaslow RA (2003) The impact of host genetics on HIV infection and disease progression in the era of highly active antiretroviral therapy. AIDS 17 Suppl 4: S51-60.

256. Tang J, Malhotra R, Song W, Brill I, Hu L, et al. (2010) Human leukocyte antigens and HIV type 1 viral load in early and chronic infection: predominance of evolving relationships. PLoS One 5: e9629.

257. Tang J, Tang S, Lobashevsky E, Myracle AD, Fideli U, et al. (2002) Favorable and unfavorable HLA class I alleles and haplotypes in Zambians predominantly infected with clade C human immunodeficiency virus type 1. J Virol 76: 8276-8284.

258. Matthews PC, Adland E, Listgarten J, Leslie A, Mkhwanazi N, et al. (2011) HLA-A*7401-mediated control of HIV viremia is independent of its linkage disequilibrium with HLA-B*5703. J Immunol 186: 5675-5686.

259. Kiepiela P, Ngumbela K, Thobakgale C, Ramduth D, Honeyborne I, et al. (2007) CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. Nat Med 13: 46-53.

260. Rolland M, Heckerman D, Deng W, Rousseau CM, Coovadia H, et al. (2008) Broad and Gag-biased HIV-1 epitope repertoires are associated with lower viral loads. PLoS One 3: e1424.

261. Brenchley JM, Knox KS, Asher AI, Price DA, Kohli LM, et al. (2008) High frequencies of polyfunctional HIV-specific T cells are associated with preservation of mucosal CD4 T cells in bronchoalveolar lavage. Mucosal Immunol 1: 49-58.

262. Gloster SE, Newton P, Cornforth D, Lifson JD, Williams I, et al. (2004) Association of strong virus-specific CD4 T cell responses with efficient natural control of primary HIV-1 infection. AIDS 18: 749-755.

263. Soghoian DZ, Jessen H, Flanders M, Sierra-Davidson K, Cutler S, et al. (2012) HIV-specific cytolytic CD4 T cell responses during acute HIV infection predict disease outcome. Sci Transl Med 4: 123ra125.

264. Ranasinghe S, Cutler S, Davis I, Lu R, Soghoian DZ, et al. (2013) Association of HLA-DRB1-restricted CD4(+) T cell responses with HIV immune control. Nat Med 19: 930-933.

265. Ranasinghe S, Flanders M, Cutler S, Soghoian DZ, Ghebremichael M, et al. (2012) HIV-specific CD4 T cell responses to different viral proteins have discordant associations with viral load and clinical outcome. J Virol 86: 277-283.

266. Schieffer M, Jessen H, Oster A, Pissani F, Soghoian DZ, et al. (2014) Induction of Gag-specific CD4 T cell responses during acute HIV infection is associated with improved viral control. J Virol.

267. Stephenson KE, Barouch DH (2013) A global approach to HIV-1 vaccine development. Immunol Rev 254: 295-304.

268. Hansen SG, Piatak M, Jr., Ventura AB, Hughes CM, Gilbride RM, et al. (2013) Immune clearance of highly pathogenic SIV infection. Nature 502: 100-104.

269. Borrow P, Lewicki H, Wei X, Horwitz MS, Peffer N, et al. (1997) Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. Nat Med 3: 205-211.

270. Cardinaud S, Consiglieri G, Bouziat R, Urrutia A, Graff-Dubois S, et al. (2011) CTL escape mediated by proteasomal destruction of an HIV-1 cryptic epitope. PLoS Pathog 7: e1002049.

271. Jones NA, Wei X, Flower DR, Wong M, Michor F, et al. (2004) Determinants of human immunodeficiency virus type 1 escape from the primary CD8+ cytotoxic T lymphocyte response. J Exp Med 200: 1243-1256.

272. Leslie AJ, Pfafferott KJ, Chetty P, Draenert R, Addo MM, et al. (2004) HIV evolution: CTL escape mutation and reversion after transmission. Nat Med 10: 282-289.

273. Price DA, Goulder PJ, Klenerman P, Sewell AK, Easterbrook PJ, et al. (1997) Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. Proc Natl Acad Sci U S A 94: 1890-1895.

274. Goulder PJ, Phillips RE, Colbert RA, McAdam S, Ogg G, et al. (1997) Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. Nat Med 3: 212-217.

275. Evans DT, O'Connor DH, Jing P, Dzuris JL, Sidney J, et al. (1999) Virus-specific cytotoxic T-lymphocyte responses select for amino-acid variation in simian immunodeficiency virus Env and Nef. Nat Med 5: 1270-1276.

276. Dudek TE, No DC, Seung E, Vrbanac VD, Fadda L, et al. (2012) Rapid evolution of HIV-1 to functional CD8(+) T cell responses in humanized BLT mice. Sci Transl Med 4: 143ra198.

277. Ferrari G, Korber B, Goonetilleke N, Liu MK, Turnbull EL, et al. (2011) 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. PLoS Pathog 7: e1001273.

278. Brumme ZL, Brumme CJ, Carlson J, Streeck H, John M, et al. (2008) 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. J Virol 82: 9216-9227.

279. Moore CB, John M, James IR, Christiansen FT, Witt CS, et al. (2002) Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. Science 296: 1439-1443.

280. Bhattacharya T, Daniels M, Heckerman D, Foley B, Frahm N, et al. (2007) Founder effects in the assessment of HIV polymorphisms and HLA allele associations. Science 315: 1583-1586.

281. Matthews PC, Prendergast A, Leslie A, Crawford H, Payne R, et al. (2008) Central role of reverting mutations in HLA associations with human immunodeficiency virus set point. J Virol 82: 8548-8559.

282. Carlson JM, Listgarten J, Pfeifer N, Tan V, Kadie C, et al. (2012) Widespread impact of HLA restriction on immune control and escape pathways of HIV-1. J Virol 86: 5230-5243.

283. Kawashima Y, Pfafferott K, Frater J, Matthews P, Payne R, et al. (2009) Adaptation of HIV-1 to human leukocyte antigen class I. Nature 458: 641-645.

284. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, et al. (2007) A whole-genome association study of major determinants for host control of HIV-1. Science 317: 944-947.

285. Boutwell CL, Rowley CF, Essex M (2009) Reduced viral replication capacity of human immunodeficiency virus type 1 subtype C caused by cytotoxic-T-

lymphocyte escape mutations in HLA-B57 epitopes of capsid protein. J Virol 83: 2460-2468.

286. Brockman MA, Schneidewind A, Lahaie M, Schmidt A, Miura T, et al. (2007) Escape and compensation from early HLA-B57-mediated cytotoxic T-lymphocyte pressure on human immunodeficiency virus type 1 Gag alter capsid interactions with cyclophilin A. J Virol 81: 12608-12618.

287. Crawford H, Lumm W, Leslie A, Schaefer M, Boeras D, et al. (2009) Evolution of HLA-B*5703 HIV-1 escape mutations in HLA-B*5703-positive individuals and their transmission recipients. J Exp Med 206: 909-921.

288. Martinez-Picado J, Prado JG, Fry EE, Pfafferott K, Leslie A, et al. (2006) Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. J Virol 80: 3617-3623.

289. Miura T, Brockman MA, Schneidewind A, Lobritz M, Pereyra F, et al. (2009) 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. J Virol 83: 2743-2755.

290. Rousseau CM, Daniels MG, Carlson JM, Kadie C, Crawford H, et al. (2008) HLA class I-driven evolution of human immunodeficiency virus type 1 subtype c proteome: immune escape and viral load. J Virol 82: 6434-6446.

291. Schneidewind A, Brockman MA, Yang R, Adam RI, Li B, et al. (2007) 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. J Virol 81: 12382-12393.

292. Schneidewind A, Brumme ZL, Brumme CJ, Power KA, Reyor LL, et al. (2009) Transmission and long-term stability of compensated CD8 escape mutations. J Virol 83: 3993-3997.

293. Boutwell CL, Carlson JM, Lin TH, Seese A, Power KA, et al. (2013) Frequent and variable cytotoxic-T-lymphocyte escape-associated fitness costs in the human immunodeficiency virus type 1 subtype B Gag proteins. J Virol 87: 3952-3965.

294. Crawford H, Prado JG, Leslie A, Hue S, Honeyborne I, et al. (2007) 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. J Virol 81: 8346-8351.

295. Rolland M, Manocheewa S, Swain JV, Lanxon-Cookson EC, Kim M, et al. (2013) HIV-1 conserved-element vaccines: relationship between sequence conservation and replicative capacity. J Virol 87: 5461-5467.

296. Schneidewind A, Brockman MA, Sidney J, Wang YE, Chen H, et al. (2008) Structural and functional constraints limit options for cytotoxic T-lymphocyte escape in the immunodominant HLA-B27-restricted epitope in human immunodeficiency virus type 1 capsid. J Virol 82: 5594-5605.

297. Wright JK, Naidoo VL, Brumme ZL, Prince JL, Claiborne DT, et al. (2012) Impact of HLA-B*81-associated mutations in HIV-1 Gag on viral replication capacity. J Virol 86: 3193-3199.

298. Friedrich TC, Dodds EJ, Yant LJ, Vojnov L, Rudersdorf R, et al. (2004) Reversion of CTL escape-variant immunodeficiency viruses in vivo. Nat Med 10: 275-281.

299. Huang KH, Goedhals D, Carlson JM, Brockman MA, Mishra S, et al. (2011) Progression to AIDS in South Africa is associated with both reverting and compensatory viral mutations. PLoS One 6: e19018.

300. Li B, Gladden AD, Altfeld M, Kaldor JM, Cooper DA, et al. (2007) Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. J Virol 81: 193-201.

301. Crawford H, Matthews PC, Schaefer M, Carlson JM, Leslie A, et al. (2011) The hypervariable HIV-1 capsid protein residues comprise HLA-driven CD8+ T-cell escape mutations and covarying HLA-independent polymorphisms. J Virol 85: 1384-1390.

302. Poon A, Chao L (2005) The rate of compensatory mutation in the DNA bacteriophage phiX174. Genetics 170: 989-999.

303. Alter G, Heckerman D, Schneidewind A, Fadda L, Kadie CM, et al. (2011) HIV-1 adaptation to NK-cell-mediated immune pressure. Nature 476: 96-100.

304. Van Teijlingen NH, Holzemer A, Korner C, Garcia-Beltran WF, Schafer JL, et al. (2014) Sequence variations in HIV-1 p24 Gag-derived epitopes can alter binding of KIR2DL2 to HLA-C*03:04 and modulate primary natural killer cell function. AIDS.

305. Battivelli E, Migraine J, Lecossier D, Yeni P, Clavel F, et al. (2011) Gag cytotoxic T lymphocyte escape mutations can increase sensitivity of HIV-1 to human TRIM5alpha, linking intrinsic and acquired immunity. J Virol 85: 11846-11854.

306. Kuroishi A, Bozek K, Shioda T, Nakayama EE (2010) A single amino acid substitution of the human immunodeficiency virus type 1 capsid protein affects viral sensitivity to TRIM5 alpha. Retrovirology 7: 58.

307. Maillard PV, Zoete V, Michielin O, Trono D (2011) Homology-based identification of capsid determinants that protect HIV1 from human TRIM5alpha restriction. J Biol Chem 286: 8128-8140.

308. Harris RS, Hultquist JF, Evans DT (2012) The restriction factors of human immunodeficiency virus. J Biol Chem 287: 40875-40883.

309. Miura T, Brockman MA, Brumme ZL, Brumme CJ, Pereyra F, et al. (2009) HLA-associated alterations in replication capacity of chimeric NL4-3 viruses carrying gag-protease from elite controllers of human immunodeficiency virus type 1. J Virol 83: 140-149.

310. Brockman MA, Brumme ZL, Brumme CJ, Miura T, Sela J, et al. (2010) Early selection in Gag by protective HLA alleles contributes to reduced HIV-1 replication capacity that may be largely compensated for in chronic infection. J Virol 84: 11937-11949.

311. Brumme ZL, Li C, Miura T, Sela J, Rosato PC, et al. (2011) Reduced replication capacity of NL4-3 recombinant viruses encoding reverse transcriptase-integrase sequences from HIV-1 elite controllers. J Acquir Immune Defic Syndr 56: 100-108.

312. Wright JK, Brumme ZL, Carlson JM, Heckerman D, Kadie CM, et al. (2010) Gag-protease-mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. J Virol 84: 10820-10831.

313. Wright JK, Novitsky V, Brockman MA, Brumme ZL, Brumme CJ, et al. (2011) Influence of Gag-protease-mediated replication capacity on disease progression in individuals recently infected with HIV-1 subtype C. J Virol 85: 3996-4006.

314. Fellay J, Ge D, Shianna KV, Colombo S, Ledergerber B, et al. (2009) Common genetic variation and the control of HIV-1 in humans. PLoS Genet 5: e1000791.

315. Hecht FM, Hartogensis W, Bragg L, Bacchetti P, Atchison R, et al. (2010) HIV RNA level in early infection is predicted by viral load in the transmission source. AIDS 24: 941-945.

316. Hollingsworth TD, Laeyendecker O, Shirreff G, Donnelly CA, Serwadda D, et al. (2010) HIV-1 transmitting couples have similar viral load set-points in Rakai, Uganda. PLoS Pathog 6: e1000876.

317. Lingappa JR, Thomas KK, Hughes JP, Baeten JM, Wald A, et al. (2013) Partner characteristics predicting HIV-1 set point in sexually acquired HIV-1 among African seroconverters. AIDS Res Hum Retroviruses 29: 164-171.

318. Tang J, Tang S, Lobashevsky E, Zulu I, Aldrovandi G, et al. (2004) HLA allele sharing and HIV type 1 viremia in seroconverting Zambians with known transmitting partners. AIDS Res Hum Retroviruses 20: 19-25.

319. Yue L, Prentice HA, Farmer P, Song W, He D, et al. (2013) Cumulative impact of host and viral factors on HIV-1 viral-load control during early infection. J Virol 87: 708-715.

320. Goepfert PA, Lumm W, Farmer P, Matthews P, Prendergast A, et al. (2008) Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. J Exp Med 205: 1009-1017.

321. Buchbinder SP, Mehrotra DV, Duerr A, Fitzgerald DW, Mogg R, et al. (2008) Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. Lancet 372: 1881-1893.

322. Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, et al. (2009) Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N Engl J Med 361: 2209-2220.

323. Carrington M, Walker BD (2012) Immunogenetics of spontaneous control of HIV. Annu Rev Med 63: 131-145.

324. Fraser C, Lythgoe K, Leventhal GE, Shirreff G, Hollingsworth TD, et al. (2014) Virulence and pathogenesis of HIV-1 infection: an evolutionary perspective. Science 343: 1243727.

325. Dalmau J, Rotger M, Erkizia I, Rauch A, Reche P, et al. (2014) Highly pathogenic adapted HIV-1 strains limit host immunity and dictate rapid disease progression. AIDS.

326. Huang KH, Goedhals D, Carlson JM, Brockman MA, Mishra S, et al. (2011) Progression to AIDS in South Africa is associated with both reverting and compensatory viral mutations. PLoS One 6: e19018.

327. Miura T, Brumme ZL, Brockman MA, Rosato P, Sela J, et al. (2010) Impaired replication capacity of acute/early viruses in persons who become HIV controllers. J Virol 84: 7581-7591.

328. Prado JG, Prendergast A, Thobakgale C, Molina C, Tudor-Williams G, et al. (2010) Replicative capacity of human immunodeficiency virus type 1 transmitted from

mother to child is associated with pediatric disease progression rate. J Virol 84: 492-502.

329. Rihn SJ, Wilson SJ, Loman NJ, Alim M, Bakker SE, et al. (2013) Extreme genetic fragility of the HIV-1 capsid. PLoS Pathog 9: e1003461.

330. Ng OT, Laeyendecker O, Redd AD, Munshaw S, Grabowski MK, et al. (2014) HIV type 1 polymerase gene polymorphisms are associated with phenotypic differences in replication capacity and disease progression. J Infect Dis 209: 66-73.

331. Mann JK, Byakwaga H, Kuang XT, Le AQ, Brumme CJ, et al. (2013) Ability of HIV-1 Nef to downregulate CD4 and HLA class I differs among viral subtypes. Retrovirology 10: 100.

332. Mwimanzi P, Markle TJ, Martin E, Ogata Y, Kuang XT, et al. (2013) Attenuation of multiple Nef functions in HIV-1 elite controllers. Retrovirology 10: 1.

333. Bailey JR, Williams TM, Siliciano RF, Blankson JN (2006) Maintenance of viral suppression in HIV-1-infected HLA-B*57+ elite suppressors despite CTL escape mutations. J Exp Med 203: 1357-1369.

334. Lazaryan A, Song W, Lobashevsky E, Tang J, Shrestha S, et al. (2011) The influence of human leukocyte antigen class I alleles and their population frequencies on human immunodeficiency virus type 1 control among African Americans. Hum Immunol 72: 312-318.

335. Chopera DR, Cotton LA, Zawaira A, Mann JK, Ngandu NK, et al. (2012) Intersubtype differences in the effect of a rare p24 gag mutation on HIV-1 replicative fitness. J Virol 86: 13423-13433.

336. Matthews PC, Koyanagi M, Kloverpris HN, Harndahl M, Stryhn A, et al. (2012) Differential clade-specific HLA-B*3501 association with HIV-1 disease outcome is linked to immunogenicity of a single Gag epitope. J Virol 86: 12643-12654.

337. Payne RP, Branch S, Kloverpris H, Matthews PC, Koofhethile CK, et al. (2014) Differential escape patterns within the dominant HLA-B*57:03-restricted HIV Gag epitope reflect distinct clade-specific functional constraints. J Virol 88: 4668-4678.

338. Zhou W, Parent LJ, Wills JW, Resh MD (1994) Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. J Virol 68: 2556-2569.

339. Koehler RN, Walsh AM, Saathoff E, Tovanabutra S, Arroyo MA, et al. (2010) Class I HLA-A*7401 is associated with protection from HIV-1 acquisition and disease progression in Mbeya, Tanzania. J Infect Dis 202: 1562-1566.

340. Herbeck JT, Nickle DC, Learn GH, Gottlieb GS, Curlin ME, et al. (2006) Human immunodeficiency virus type 1 env evolves toward ancestral states upon transmission to a new host. J Virol 80: 1637-1644.

341. Cotton LA, Kuang XT, Le AQ, Carlson JM, Chan B, et al. (2014) Genotypic and functional impact of HIV-1 adaptation to its host population during the North American epidemic. PLoS Genet 10: e1004295.

342. Schoggins JW (2014) Interferon-stimulated genes: roles in viral pathogenesis. Curr Opin Virol 6C: 40-46.

343. Fenton-May AE, Dibben O, Emmerich T, Ding H, Pfafferott K, et al. (2013) Relative resistance of HIV-1 founder viruses to control by interferon-alpha. Retrovirology 10: 146.

344. Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, et al. (2003) Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS 17: 1871-1879.

345. Baker BM, Block BL, Rothchild AC, Walker BD (2009) Elite control of HIV infection: implications for vaccine design. Expert Opin Biol Ther 9: 55-69.

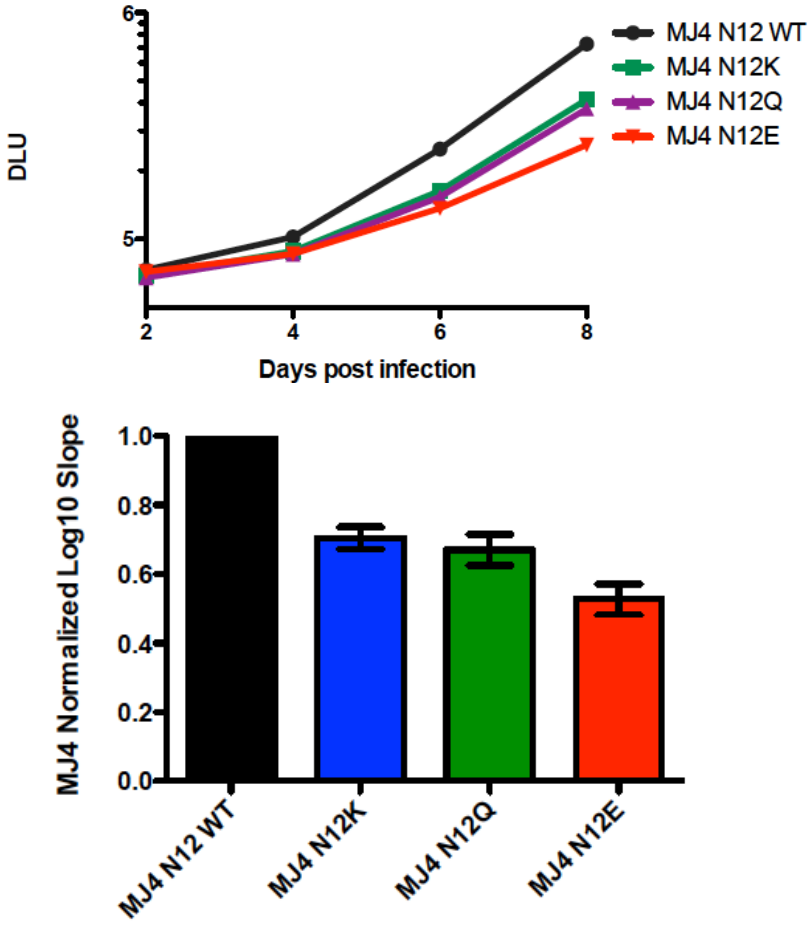
346. Stacey AR, Norris PJ, Qin L, Haygreen EA, Taylor E, et al. (2009) Induction of a striking systemic cytokine cascade prior to peak viremia in acute human immunodeficiency virus type 1 infection, in contrast to more modest and delayed responses in acute hepatitis B and C virus infections. J Virol 83: 3719-3733.

347. Shive CL, Mudd JC, Funderburg NT, Sieg SF, Kyi B, et al. (2014) Inflammatory Cytokines Drive CD4+ T-Cell Cycling and Impaired Responsiveness to Interleukin 7: Implications for Immune Failure in HIV Disease. J Infect Dis.

348. Jacquelin B, Mayau V, Targat B, Liovat AS, Kunkel D, et al. (2009) Nonpathogenic SIV infection of African green monkeys induces a strong but rapidly controlled type I IFN response. J Clin Invest 119: 3544-3555.

349. Persaud D, Gay H, Ziemniak C, Chen YH, Piatak M, Jr., et al. (2013) Absence of detectable HIV-1 viremia after treatment cessation in an infant. N Engl J Med 369: 1828-1835.

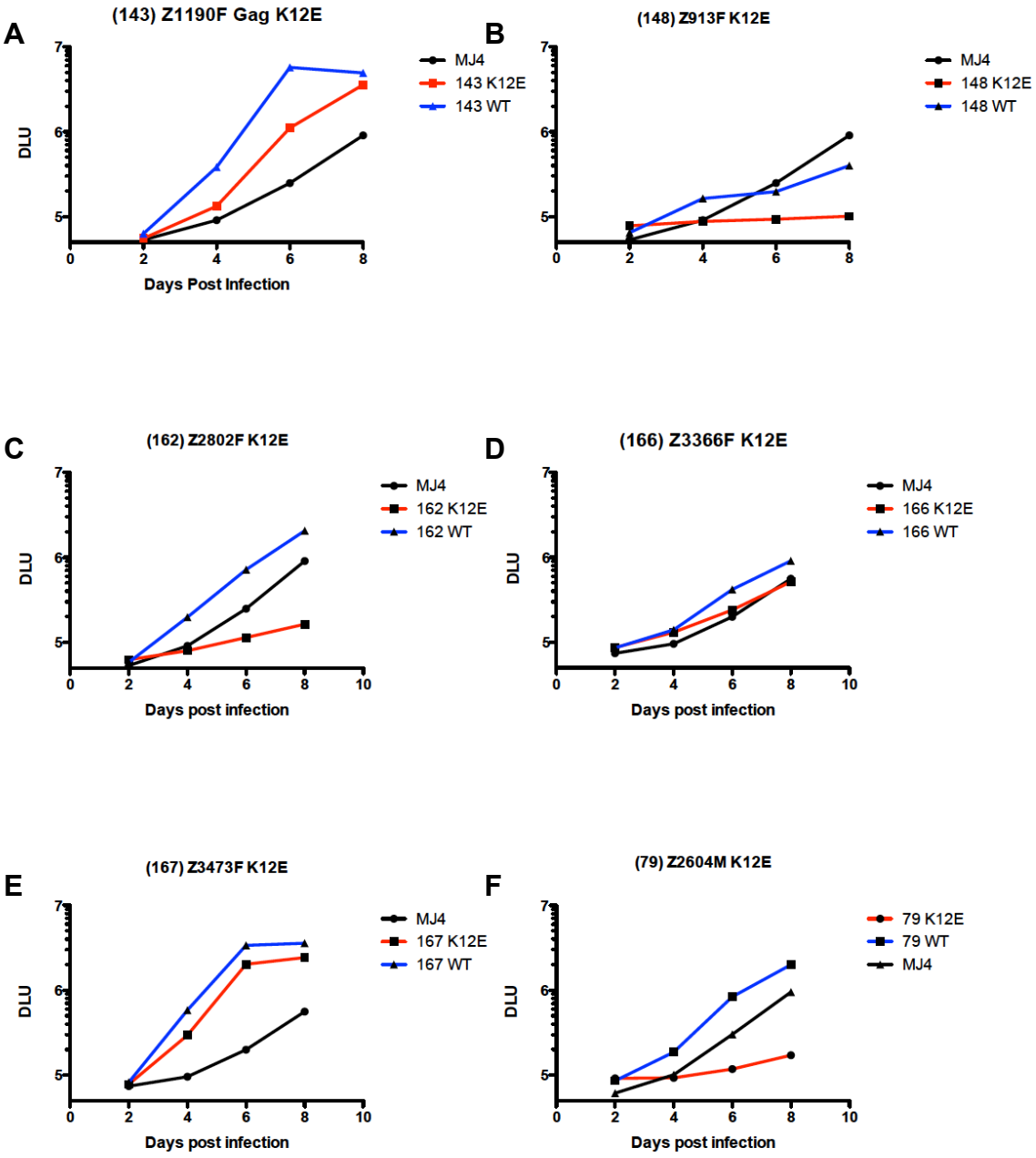
350. Vandergeeten C, Fromentin R, DaFonseca S, Lawani MB, Sereti I, et al. (2013) Interleukin-7 promotes HIV persistence during antiretroviral therapy. Blood 121: 4321-4329.

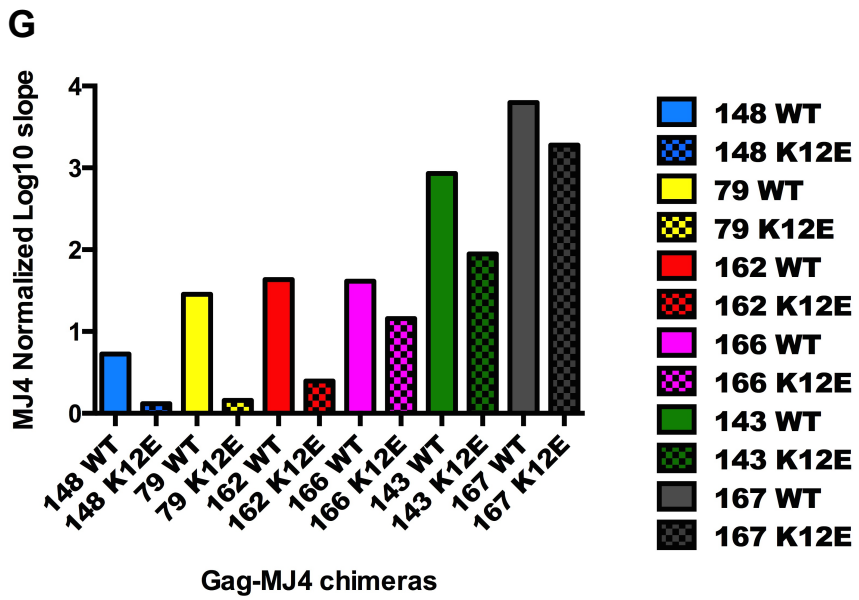


Appendix Figure 1: *In vitro* replication of all naturally occurring residues at the A*74-associated position 12 in Matrix engineered into WT MJ4.

(A) Site-directed mutagenesis was performed in order to engineer the naturally occurring residues at postion 12 of matrix - 12K, 12Q, and 12E - into MJ4, which has an N at that position. $5x10^5$ GXR25 cells were infected with wild-type MJ4 (black), MJ4 N12Q (green), MJ4 N12K (blue), and MJ4 N12E (red) an MOI of 0.05 in triplicate and with two independent clones. Supernatants were collected at two day intervals post infection, and virion production quantified by a radiolabeled RT assay. The digital light unit (DLU) values were averaged from the three replicates and two independent clones

to create replication curves. **G.** Replication capacity scores were generated by calculating the log10 slope from days 2-8 for each virus and then normalizing that value by dividing it by the log10 slope of WT MJ4. Normalized RC scores were averaged for each virus based on the three replicates for two independent clones.


A
(143) Z1190F Gag K12E
DLU
Days Post Infection
MJ4
143 K12E
143 WT
B
(148) Z913F K12E
MJ4
148 K12E
148 WT
C
(162) Z2802F K12E
Days post infection
MJ4
162 K12E
162 WT
D
(166) Z3366F K12E
Days post infection
MJ4
166 K12E
166 WT
E
(167) Z3473F K12E
Days post infection
MJ4
167 K12E
167 WT
F
(79) Z2604M K12E
Days post infection
79 K12E
79 WT
MJ4



Appendix Figure 2: *In vitro* replication defects of the A*74-associated polymorphism K12E engineered into various Gag-MJ4 sequences

(A-F) Site-directed mutagenesis was performed in order to engineer the A*74 associated polymorphism, K12E, into six different Gag-MJ4 chimeric viruses, in which the *gag* gene was derived from an acutely infected Zambian. 5 x 10^5 GXR25 cells were infected with wild-type MJ4 (black), the WT Gag-MJ4 chimeras (blue), as well as the K12E mutation for each Gag-MJ4 chimera (red) at an MOI of 0.05. Supernatants were collected at two day intervals post infection, and virion production quantified by a radiolabeled RT assay, and digital light unit values (DLU) plotted for each time point. **G.** Replication capacity scores were generated by calculating the $\log_{10}$ slope from days 2-8 for each virus and then normalizing that value by dividing it by the $\log_{10}$ slope of WT MJ4. The normalized RC values for the WT Gag-MJ4 chimeras (which all naturally have a 12K residue) are shown in solid colors, while Gag-MJ4 chimeras with the K12E mutation are shown in a hatched pattern. Insertion of K12E into various subtype C derived *gag* genes can have varying replication defects. K12E was almost lethal in some sequence contexts, while in others, it did not have a major impact.

A restriction enzyme based cloning method to assess the *in vitro* replication capacity of HIV-1 subtype C Gag-MJ4 chimeric viruses

Manuscript in press; Journal of Visualized Experiments, 2014

Short Abstract:

HIV-1 pathogenesis is defined by both viral characteristics and host genetic factors. Here we describe a robust method that allows for reproducible measurements to assess the impact of the *gag* gene sequence variation on the *in vitro* replication capacity of the virus.

Abstract:

The protective effect of many HLA class I alleles on HIV-1 pathogenesis and disease progression is, in part, attributed to their ability to target conserved portions of the HIV-1 genome that escape with difficulty. Sequence changes attributed to cellular immune pressure arise across the genome during infection, and if found within conserved regions of the genome such as Gag, can affect the ability of the virus to replicate *in vitro*. Transmission of HLA-linked polymorphisms in Gag to HLA-mismatched recipients has been associated with reduced set point viral loads. We hypothesized this may be due to a reduced replication capacity of the virus. Here we present a novel method for assessing the *in vitro* replication of HIV-1 as influenced by the *gag* gene isolated from acute time points from subtype C infected Zambians. This method uses restriction enzyme based cloning to insert the *gag* gene into a common subtype C HIV-1 proviral backbone, MJ4. This makes it more appropriate to the study of subtype C sequences than previous recombination based methods that have assessed the *in vitro* replication of chronically derived *gag-pro* sequences. Nevertheless, the protocol could be readily modified for studies of viruses from other subtypes. Moreover, this protocol details a robust and reproducible method for assessing the replication capacity of the Gag-MJ4 chimeric viruses on a CEM-based T cell line. This method was utilized for the study of Gag-MJ4

chimeric viruses derived from 149 subtype C acutely infected Zambians, and has allowed for the identification of residues in Gag that affect replication. More importantly, the implementation of this technique has facilitated a deeper understanding of how viral replication defines parameters of early HIV-1 pathogenesis such as set point viral load and longitudinal CD4+ T cell decline.

Introduction:

Determining both the host and viral characteristics that influence HIV-1 pathogenesis and disease progression is paramount for rational vaccine design. The cellular immune response is a key component of the human immune response to HIV-1 infection. Cytotoxic T lymphocytes (CTL) are necessary for the initial control of acute viremia, and allow the host to establish a steady state (set point) viral load [248,249]. Experimental depletion of these effector cells results in loss of viral control [250,394]. Despite this, escape mutations arise within the viral genome that subvert CTL recognition of virally infected cells [272-274,278,324].

Certain HLA alleles have been associated with lower viral loads and slower disease progression including B*57, B*27 and B*81 [251,253,254,257,395,396]. Part of the protective benefit of HLA class I alleles can be attributed to the fact that they target functionally constrained regions of the genome such as Gag and select for escape mutations that decrease the ability of the virus to replicate *in vitro* [259,260,328-331]. Although escape from the cellular immune system is beneficial to the virus in the context of the selecting HLA class I allele, the effect of these mutations may have differential consequences for the host upon transmission to an HLA-mismatched individual [287,326]. Therefore, understanding the effects of transmitted HLA-associated escape mutations on viral replication capacity will be important to further our understanding of early HIV-1 pathogenesis.

While much progress has been made to identify and characterize the fitness defects of individual escape mutations associated with specific HLA class I alleles

[286,288,291,293,294,297], naturally occurring HIV-1 isolates have unique and complex footprints of HLA-associated polymorphisms, likely arising from the HLA-mediated immune pressure of different immunogenetic backgrounds [283]. In a previous analysis, Goepfert et. al. showed that an accumulation of HLA-associated mutations in the transmitted Gag sequences derived from 88 acutely infected Zambians was associated with a reduction in set point viral load [320]. This suggested that the transmission of deleterious escape mutations, specifically in Gag, to HLA-mismatched recipients provides a clinical benefit, and may be due to attenuated viral replication. Moving forward, it is imperative to study how complex combinations of Gag polymorphisms within naturally occurring isolates work in concert to define characteristics of the transmitted virus such as replication capacity, and how early replication might in turn affect HIV-1 clinical parameters and late-stage pathogenesis.

Brockman et. al. first demonstrated a link between the replication capacity of *gag-pro* sequences isolated during chronic stage infection and viral load in both subtype C and B infections [310,312,313,334]. The experimental approach presented in these studies, although appropriate for examining the *in vitro* replication capacity of sequences derived from chronically infected individuals, has several technical caveats and limitations that make studying HIV-1 replicative capacity in subtype C acutely infected individuals difficult. This method relies on the recombination of population based PCR-amplified sequences into the subtype B NL4-3 provirus, which was derived in part from LAV, a laboratory adapted virus stock [397]. Virus generation was accomplished by co-transfection of a CEM-based T-cell line [342] with PCR amplicons and digested delta-*gag-pro* NL4-3 DNA. This method requires the outgrowth of virus over a period of weeks to months, potentially skewing the nature of the recovered virus stock in relation to the viral quasispecies *in vivo*, and therefore altering the measurement of replication

capacity *in vitro*. This method is more appropriate for studying chronically infected individuals, where it effectively selects for virus with the highest replicative capacity, and where cloning numerous different viral variants from a large number of chronically infected individuals is quite labor intensive and therefore not feasible. However, within an acutely infected individual, there are generally one to two variants present, and thus eliminating the risk of skewing the nature of the recovered virus stock, through *in vitro* selection pressures, allows for a more accurate assessment of *in vitro* replication capacity. Secondly, this method requires recombining subtype C *gag-pro* sequences into a subtype B derived backbone, and could introduce backbone incompatibility biases into the analysis. Due to these limitations, large numbers of sequences must be analyzed in order to overcome any potential biases introduced.

Here we describe an alternative experimental approach appropriate for studying sequences derived from subtype C acutely infected individuals. We use a restriction enzyme based cloning strategy to introduce the *gag* gene derived from acute infection time points of HIV-1 subtype C infected individuals into the subtype C proviral backbone, MJ4. The use of MJ4 as a common backbone in which to clone *gag* genes is crucial for the analysis of subtype C derived sequences. MJ4 is derived from a primary isolate [337], and thus would be less likely to introduce bias due to subtype incompatibility between the backbone and *gag* gene. In addition, the approach of using enzyme based restriction cloning allows for the proviral constructs to be transfected directly into 293T cells, and for the recovery of a clonal virus stock identical to the cloned *gag* sequence.

The method presented below is a high throughput method for assessing the replication capacity of subtype C derived Gag-MJ4 chimeric viruses. Transfection into 293T cells is

straightforward and recovery of virus takes only three days. *In vitro* replication capacity is assayed on the same CEM-CCR5 based T-cell line created by Brockman et. al. [342], but using important protocol modifications necessary for the successful replication of subtype C MJ4 chimeric viruses. The use of an appropriate T-cell line rather than PBMCs allows for large numbers of subtype C MJ4-chimeric viruses to be tested with high assay reproducibility. Finally, using a radiolabeled reverse transcriptase assay for quantification of virus in the supernatant is more cost effective than using commercially available p24 ELISA kits. It also gives a higher dynamic range, which was important for detecting both poorly and highly replicating viruses within the same assay and for detecting subtle differences in replication between isolates.

In conclusion, the method presented here has allowed for the in-depth study of the replication capacity of *gag* sequences derived from HIV-1 subtype C acutely infected individuals from Zambia, and as written, could also be expanded to study other subtype C infected populations. A high degree of variation in replication capacities between different Gag isolates was observed. In addition, we were able to show a statistical association between the replication capacity of the transmitted Gag and set point viral load as well as with CD4+ decline over a three-year period [362]. Such results highlight the importance of studying how transmitted viral characteristics, such as replication capacity, interact with the host immune system to influence pathogenesis during early infection and will be integral for developing effective vaccine interventions as well as treatment.

Protocol Text:

1. Amplification of the HIV-1 *gag* gene from infected, frozen plasma

1.1 Extract viral RNA from 140µL thawed HIV-1 infected plasma using an extraction kit.

1.1.1 When feasible, immediately proceed to cDNA synthesis after RNA extraction as unfrozen viral RNA yields the best amplification results. If possible, set up PCR master mix for first-round DNA amplification and store at 4°C prior to viral RNA extraction.

1.2 Reverse-transcribe cDNA from RNA and amplify first-round DNA products using reverse-transcriptase and a thermostable DNA polymerase in a one-step RT-PCR. Table 1A and 1B††

1.2.1 Take RNA out of -80°C freezer (if freezing was necessary) and briefly thaw at room temperature then place in cold block. Transfer 5µL of RNA to each PCR tube containing 45µL of master mix to give a final volume of 50µL. Immediately place remaining RNA samples into -80°C freezer.

1.2.2 After enzyme is added to the first-round PCR master mix (Table 1A), aliquot 45µL into each thin-walled PCR amplification tube for the number of reactions desired. Make sure to use PCR tubes with individual caps and not strip caps to ensure minimal cross-contamination between samples. Place PCR tubes in a cold block to protect the temperature sensitive RT enzyme and RNA template. Note: samples should be run in triplicate in order to adequately sample the viral quasispecies. It is also important to utilize a product with a high fidelity DNA polymerase enzyme in order to minimize PCR-

introduced misincorporation. Please refer to the reagents list for the recommended product.

1.2.3 After RNA template has been added to all reaction tubes, transfer PCR tubes to a thermocycler for amplification using the cycler program described in Table 1B. Note: the product of this amplification will be used in a second-round nested PCR.

1.3 Perform a nested second-round PCR amplification, using 1µL of the first round PCR amplification (1.2) as the DNA template. Table 2A, 2B††

1.3.1 Aliquot 49µL of second-round PCR master mix (Table 2A) to each thin-walled PCR tube for the number of reactions desired. Transfer 1µL from the first-round PCR amplification to each reaction tube to give a final volume of 50µL. Note: this will serve as template DNA for the second-round amplification. It is also important to utilize a DNA polymerase with very high fidelity and proofreading capabilities in order to minimize PCR-introduced misincorporation. Please refer to the reagents list for the recommended product.

1.3.2 Transfer second-round PCR reaction mix to thermocycler and run program described in Table 2B. Note: after completion of the program, the products of this reaction will be run on an agarose gel to confirm the production of product.

1.4 Add 3µL of 5x loading dye to 5µL of the 50µL second-round reaction volume and run at 120V on a 1% agarose-TAE gel containing a UV-fluorescent DNA stain until bands are resolved. Visualize 1.6kb bands on a blue light illuminator. †† Figure 1A

1.5 Run the remaining 45µL reaction volume on a 1% agarose-TAE gel containing a DNA stain, and excise the appropriate bands with a clean razor blade.

1.5.1 Extract DNA from the gel slice using a gel purification kit, elute in nuclease-free water, combine three positive reactions per individual, and freeze product at -20°C for subsequent use.

2. Preparing *gag* amplicons for cloning by introduction of the necessary restriction sites

2.1 Amplify the Long Terminal Repeat (LTR)/5' UTR portion of the MJ4 plasmid. †† Table 3

2.2 Visualize 1.3 kb PCR products on illuminator, excise positive amplicons, gel purify and freeze as previously described (1.4,1.5).†† Figure 1B

2.3 Create fused MJ4 LTR-*gag* amplicons via "splice-overlap-extension" PCR. †† Table 4

2.4 Visualize, excise, purify and freeze the 3.2kb-long amplicons as in 1.4,1.5. †† Figure 1C

3. Cloning amplified *gag* genes into the MJ4, subtype C, infectious molecular clone

3.1 Digest 1.5µg of MJ4 plasmid and 1.5µg of purified LTR-*gag* PCR product with the BclI (methylation sensitive) restriction endonuclease for 1.5 hours at 50°C. †† Table 5A

3.2 Add 1µL of NgoMIV to the digestion reaction and incubate at 37°C for 1 hour.

3.3 Add 5x loading dye to restriction digest reactions and slowly electrophorese the total volume on a 1% agarose-TAE gel containing a DNA gel stain for 1-2 hours at 100 V. Visualize, excise, and purify indicated bands for cloning as previously described. †† Figure 2

3.4 Prepare ligation reactions using purified LTR-*gag* insert and MJ4 vector DNA at a 3:1 insert to vector ratio. Incubate ligation reactions overnight at 4°C. †† Table 5B

3.5 Transform JM109 chemically competent bacteria with ligation products and spread on LB agar plates with 100µg/mL ampicillin and grow at 30°C for 22+ hours.

3.5.1 Thaw JM109 competent cells on ice for 15 minutes. Label 1.5mL microcentrifuge tubes and chill on ice.

3.5.2 Aliquot 50 – 100µL of JM109 cells to pre-chilled 1.5mL microcentrifuge tubes. Add 2.5 – 5µL of ligation reaction to JM109 cells, flicking tube lightly to mix and immediately returning to ice. Incubate competent cells with ligation product on ice for 30 minutes.

3.5.3 Heat shock 1.5mL microcentrifuge tubes containing JM109 competent cells and ligation reaction in a 42°C heat block for 45 seconds and return to ice for at least 3 minutes.

3.5.4 Add 50µL of SOC media to each microcentrifuge tube and plate the entire transformation reaction onto room temperature LB-agar plates supplemented with 100ug/mL of ampicillin.

3.5.5 Transfer plates to a 30°C incubator and leave for 20 hours, or until colonies are clearly formed.

3.6 Pick isolated colonies and grow in 4mL of LB with 100µg/mL ampicillin at 30°C for 22+ hours.

3.7 Spin down cultures at 3,200 x g for 15 min, pour off broth, and extract plasmid DNA.

3.8 To confirm cloning fidelity, cut miniprep DNA with NgoMIV and HpaI restriction enzymes in a double digest at 37°C for 2 hours. Analyze on 1% agarose-TAE gel. †† Table 5C.

3.9 Sequence the LTR-gag insert region of each plasmid using the following primers: GagInnerF1, GagF2, Rev1, Rev3, and GagR6 (Table 6) to confirm sequence identity.

3.10 Compare the cloned sequences obtained with the population sequences derived from the initial purified amplicon from 1.5.1. Note: it is important to ultimately assess the replication capacity of two independent clones in order to ensure that any *in vitro* replication phenotypes are not due to backbone errors introduced during the cloning process.

4. Generation and titering of replication competent Gag-MJ4 chimeric viruses

4.1 Generate replication competent virus by transfecting 1.5µg of the chimeric MJ4 plasmid miniprep DNA into 293T cells using a 4:1 ratio of Fugene HD as described by Prince et. al [362].

4.2 Titer harvested virus stocks on TZM-bl cells as described below and in Prince et. al [362].

4.2.1 Plate TZM-bl cells 24 hours before infection with virus in order to have a 30-40% confluent cell monolayer on the following day. This can usually be achieved by adding 5 x 10^4 cells in a total volume of 800µL per well in a 24-well plate.

4.2.2 After adding cells to the well, gently move the 24-well plate forward and back, then side to side in order to efficiently distribute cells. Never swirl – this will result in cells accumulating in the center of the plate.

4.2.3 The following day, prepare 1% FBS in DMEM; this will be used to dilute DEAE-Dextran and to dilute virus stocks. Stock DEAE-Dextran is 10mg/mL or 125X, and a final concentration of 80µg/mL or 1X is desired.

4.2.4 Take out virus stock to be tested from -80°C freezer and place on shaker to thaw at room temperature.

4.2.5 Using a multichannel pipette, serially dilute virus stocks in a round-bottomed tissue culture treated 96-well plate with lid. This is a 3-fold dilution protocol. See table 7 for the dilution scheme.

4.2.7 Remove media from seeded TZM-bl 24-well plates using a vacuum aspirator making sure not to disturb cell monolayer. Only remove the media from one 24-well plate at a time so that cells do not dry out.

4.2.8 Add 150µL of the 1X DEAE-Dextran + 1% FBS in DMEM mixture to each well using a 1000µL pipette. Do not use repeat pipette as this disrupts the monolayer.

4.2.9 Add 150µL of each virus dilution to the appropriate well. Add virus to the center of the well and swirl gently to evenly distribute virus across the cell monolayer. Incubate infected cells for 2 hours in a tissue culture incubator at 37°C and 5% CO_2.

4.2.10 After the 2 hour incubation, add 0.5mL 10% FBS in DMEM to each well and return plates to incubator for an additional 48 hours.

4.2.11 After 48 hours, cells should be 100% confluent. Because MJ4 is a replication competent virus, there will be some cell death, but this is to be expected.

4.2.12 Remove media from each well and add 400µL/well of fixing solution (see Table 8A for recipe). Fix only one plate at a time. Add fixing solution using a 1000µL pipette, and let sit for 5 minutes at room temperature.

4.2.13 Remove fixing solution and wash 3x with PBS with Ca^{2+}/Mg^{2+}. Use a squirt bottle to gently add PBS to the side of the well to avoid disrupting the monolayer. After the third wash, blot plate dry on absorbent paper.

4.2.14 Add 400µL/well of staining solution (see Table 8B for recipe) to each well of the 24-well plate, and incubate at 37°C for at least 2 hours. Longer incubation times are acceptable, but incubation times should be kept consistent between titering experiments.

4.2.15 Wash 2x with PBS with Ca^{2+}/Mg^{2+} and score for infection, or add PBS and store at 4°C for scoring later. Plates can be kept at 4°C for up to four days, as long as the monolayer is kept hydrated.

4.2.16 To score for infection, use a permanent marker to divide wells of the 24-well plate into quadrants. Count all blue cells within a field of view, using a total magnification of 200X, once in each of the four quadrants of a well.

4.2.17 Compute Infectious Units per µL as follows: [(# blue cells/4) x 67]/(µL virus added) = IU/µL. Refer to Table 8C for volume in µL of virus added to each well. Average several wells together; the numbers should be relatively similar for an accurate titration.

5. Preparation of culture media and propagation of GXR25 cells for *in vitro* replication assay

5.1 Prepare complete RPMI (cRPMI) for propagation of GXR25 cells. Note: It is extremely important to culture GXR25 cells at least four months prior to planned replication experiments.†† Table 9

5.2 Split cells 1:10 twice per week and maintain cell density between $1x10^5$ – $2x10^6$ cells/mL.

6. *In vitro* replication of Gag-MJ4 chimeras in GXR25 (CEM-CCR5-GFP) cells

6.1 The day before the infection, split cells to a concentration between 2 and $3x10^5$ cells/mL so that they are in logarithmic growth on the day of the infection.

6.2 On the day of infection, remove virus stocks from -80°C freezer and thaw. Calculate the volume of virus needed from each stock based on the IU/µL titer from TZM-bl cell assay in order to infect $5x10^5$ GXR25 cells at a multiplicity of infection of 0.05, using the formula: Volume of virus for infection (µL) = 1µL/X IU x 0.05 x 500,000. Note: we have chosen an MOI of 0.05 as this results in a logarithmic growth phase for all of the viruses that we have tested. It is important to conduct initial experiments in order to determine the optimal MOI to capture logarithmic growth for the virus to be tested.

6.3 Dilute virus in cRPMI to a volume of 100µL (in order to achieve a 0.05 MOI) and add into a well of a V-Bottom tissue culture treated 96-well plate. Note: Include both a positive (MJ4 WT infected) and a negative (mock infected cells) control in each set of replication experiments. Set up the plate such that every other column is blank to limit cross-contamination between wells. The positive control is imperative, as it will be used later as a standard to normalize the replication slopes, which are a measure of the replication rate of experimental replicates.

6.4 Count GXR25 cells using an automated cell counter. Calculate the number of cells needed in total for all infections ($5x10^5$ x # infections). Aliquot the required volume into a 50mL conical tube and centrifuge to pellet cells. Always calculate for 25% more infections than needed.

6.5 Aspirate media carefully and resuspend in cRPMI at a concentration of $5x10^5$ cells/100µL.

6.6 Pipette cells into a sterile trough and mix thoroughly. Using a multi-channel pipette, add 100µL of cells into each well of the 96 well plate containing the diluted virus. Mix thoroughly.

6.7 Add 2µL of 5mg/mL (100x) solution of polybrene to each well and mix cells thoroughly.

6.8 Incubate at 37°C in tissue culture incubator with 5% CO_2 for 3 hours.

6.9 In order to wash infected cells, centrifuge 96-well V-bottom plate to pellet cells. Then carefully remove 150µL of the medium and replace with fresh 150µL of cRPMI without disturbing the cell pellet. Repeat 2 more times (including centrifugation) to sufficiently wash cells.

6.10 After the last centrifugation and addition of 150µL cRPMI, resuspend the cell pellet with a multichannel pipette and add the entire cell/virus mixture (approximately 200µL) from each well independently to 800µL of cRPMI in a well of a 24-well tissue culture plate.

6.11 Place plate in 5% CO_2 tissue culture incubator at 37°C.

6.12 Every two days, remove 100µL of supernatant from surface of culture well and transfer to a 96-well U-bottom plate and store frozen at -80°C until running virus quantitation assay. Note: Careful arrangement of supernatants in 96-well plates will allow for multi-channel pipette transfer of culture supernatants during virion quantification via a radiolabeled reverse transcriptase (RT) assay. Every plate should contain samples from an infection standard in order to normalize inter-plate variation in the RT assay readout.

6.13 After removing each 100µL sample, split cells 1:2 by thoroughly resuspending cells and removing half of the remaining volume (450µL). Restore original volume (1mL) by adding 550µL of fresh cRPMI to each well.

7. Analysis of reverse transcriptase (RT) in cell culture supernatants. Protocol adapted from Ostrowski et. al. [343]

7.1 Set up biological safety hood in a BSL-3 facility for use with radioactive materials.

7.1.1 Place absorbent paper in hood and replace aspirator for aspirator designated for radioactive use. Note: all radioactive waste must be properly disposed of with accordance to safety regulations.

7.2 Add 1-2µL of 10mCi/mL of [∝-P33] dTTP and 4µL of 1M dithiothreitol to each 1mL aliquot of RT master mix. †† Table 10 and Ostrowski et. al.[343]

7.3 Carefully mix each 1.5 mL microcentrifuge tube of RT master mix, DTT, and [∝-P33] dTTP with a 1000µL pipette and transfer to a sterile trough.

7.4 Dispense 25µL of labeled RT mix into each well of 96 well thin-walled PCR plate. Note: Include space for a positive control (MJ4 WT infection) and negative control (Mock infection).

7.5 Add 5µL of each supernatant to the PCR plate containing the RT master mix.

7.6 Seal PCR plate with adhesive foil cover and incubate for 2 hours at 37°C in a thermocycler. For safety reasons, rinse pipette tips with Amphyl and dispose of tips in small container containing Amphyl waste, which will later be disposed of in radioactive waste.

7.7 After incubation, make small holes in the foil cover using a 200µL multi-channel pipette. Note: If pipette tips touch liquid in well, replace tips before moving to the next column or row.

7.8 Mix samples 5 times and transfer 5µL of each sample to the DE-81 paper. Air dry 10 min.

7.9 Wash blots 5x with 1X SSC (sodium chloride, sodium citrate), and then 2x with 90% ethanol at 5 minutes per wash. Allow to air dry.

7.9.1 Place each blot in a separate washing container, such as a sandwich storage box, add enough wash buffer (1X SSC or 90% EtOH) to cover blot, and shake for 5 min.

7.9.2 Pour off wash buffer into separate container and repeat. Note: The first 3 washes are considered radioactive and should be disposed of properly. The last 2 washes can be disposed of regularly.

7.10 Once dry, carefully wrap blot in saran wrap and expose to a phosphoscreen in a tightly sealed cassette overnight at room temperature.

7.11 Analyze phosphoscreens with a phosphorimager and quantify radioactive transcripts.

7.12 Draw a circle around each radioactive signal using OptiQuant software. Graph the Digital Light Unit (DLU) values to generate replication curves.

7.13 In order to generate replication capacity scores, DLU values were $\log_{10}$-transformed and slopes were calculated using the day 2, 4, and 6 time points. Replication slopes are then divided by the slope of WT MJ4 in order generate replication capacity scores. It is important to normalize based on the MJ4 WT values obtained from the same RT plate to reduce intra-assay variability.

Representative Results:

In order to properly execute this protocol, which creates a proviral plasmid capable of assembling fully functional, infectious Gag-MJ4 chimeras, great care must be taken to generate the appropriate PCR amplicons. Determining whether the PCR has generated the appropriately sized *gag* amplicon is crucial. Products should be within 100 base pairs (bp) of the approximately 1,700 bp amplicon depicted in Figure 1A. The exact length of this fragment will vary depending on the *gag* gene under study. Next, the 5' long terminal repeat (LTR) portion of the MJ4 molecular clone must be amplified and spliced to the *gag* amplicon in order to make it suitable for subsequent cloning. The MJ4 LTR amplicon should be 1,474 bp in length. Figure 1B shows a representative gel image for which correct band sizes are indicated. After the splice-overlap-extension PCR [398], combined LTR-*gag* products should be approximately 3,200 bp in length, as depicted in Figure 1C.

Once the *gag* gene has been made suitable for cloning by fusion to the 5' LTR from MJ4, which contains the necessary NgoMIV restriction site, both vector and *gag* insert must be digested with NgoMIV and BclI restriction enzymes and excised from an agarose gel after electrophoretic separation. It is imperative to excise the appropriate vector and insert bands. A representative gel is shown in Figure 2. The vector portion of the MJ4 plasmid should be approximately 10,000 bp in length, while the LTR-*gag* insert should remain at approximately 3,000 bp in length, as the restriction sites are located at the extreme ends of the amplicon. Any significant decrease in size will indicate an additional cut-site within the *gag* gene under study.

Following ligation of the two fragments, bacterial transformation, and isolation of plasmid DNA, the Gag-MJ4 chimeras must be checked for appropriate size by

performing a double digest with NgoMIV and HpaI restriction enzymes. Full-length Gag-MJ4 chimeras that have not incurred any deletion events during bacterial replication should have a restriction pattern similar to that depicted in Figure 3, with two bands of approximately 8,700 and 4,300 bp.

An important distinction of this protocol compared to previous approaches, is the use of the HIV-1 subtype C infectious molecular clone, MJ4, rather than the more common laboratory-adapted NL4-3 virus. However, the approaches described in the previous section could be modified for cloning of subtype B *gag* sequences into pNL4-3.

An optimization of the multiplicity of infection (MOI) for use in subsequent experiments was performed in order to select the ideal MOI that showed logarithmic growth of a majority of the viruses tested. Figure 4 depicts representative replication curves from three different MOIs (0.01, 0.05, and 0.25) for MJ4 (Figure 4A) as well as for NL4.3 (Figure 4B). MJ4 replicates much less efficiently in GXR25 cells than NL4-3, which is important to take into account, as an MOI appropriate for NL4-3 replication would likely be too low to detect efficient MJ4 replication. As seen in Figure 4A, an MOI of 0.05 as opposed to 0.01 or 0.25, was the ideal choice, because logarithmic growth was observed between days 2-6 for MJ4. For the lower MOI, 0.01, day 6 DLU values are barely detectable, and we anticipated the generation of Gag-MJ4 chimeric viruses that replicated lower than MJ4. Therefore this MOI would not capture the growth of the more attenuated Gag-MJ4 chimeras, which may also be the most biologically critical. Additionally, an MOI of 0.25 was not ideal, because the rapid kinetics of viral replication killed a substantial amount of cells even by day 4. This causes the replication curve to plateau, and calculating a slope based on a curve such as this would underestimate the replication capacity. Based on curves generated for NL4-3, even at an MOI of 0.01,

available cell targets have been noticeably exhausted by day 6 post infection. In conclusion, an MOI of 0.05 was found to be optimal for a large panel of Gag-MJ4 chimeric viruses, all of which had diverse Gag sequences and varying degrees of replication.

A total of 149 Gag-MJ4 chimeric viruses derived from acute subtype C Gag sequences have been tested for *in vitro* replication using this assay. The normalized RC values ranged from 0.01 to over 3.5 with some viruses replicating more than 100 times more efficiently than wild-type MJ4. Figure 5 shows the replication curves from nine representative Gag-MJ4 chimeric viruses, with wild-type MJ4 depicted in red, and demonstrates the wide range of replication capacities observed. Thus, the sequence diversity within the *gag* gene alone can drastically impact the ability of the virus to replicate *in vitro*. While this is representative of Gag-MJ4 chimeric viruses derived from acutely infected Zambians, other subtype C sequences have not been extensively tested, and may exhibit different replication kinetics. Therefore, great care must be taken to optimize the MOI to suit the specific replication of the viruses of a particular study, because there can be a wide range in the levels of replication between different HIV-1 backbones and Gag isolates.

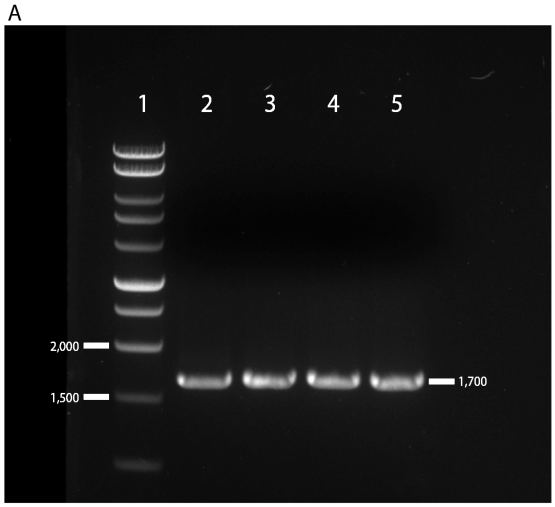
One of the advantages of using a T-cell line such as the GXR25 cell line, that supports the replication of MJ4, is the level of reproducibility observed relative to replication experiments using stimulated peripheral blood mononuclear cells as targets. In initial optimization experiments, MJ4 wild type exhibited an intra-assay variability of 8.7%, and different clones of the same Gag-MJ4 chimeric virus exhibited variability in replication of 8.5%. Because different master mixes and phosphoscreen exposures may give DLU values that differ slightly in magnitude, intra-assay variability can further be

controlled by running the same virus standard (in our case wild-type MJ4) in each RT assay plate. Figure 6 graphs the DLU values derived from the same MJ4 infection, quantified in eight different RT plates. Normalizing to a virus that is common to all RT assays can help to mitigate potential error induced by these slight changes in signal magnitude between assays.

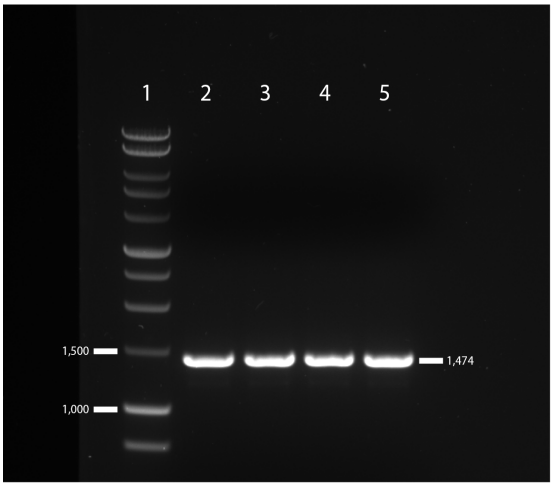
Inter-assay variably was also tested and replicates repeated on different days were highly correlated. Figure 7 plots the normalized RC score values from two independent experiments performed approximately one year apart. A high degree of correlation with the absence of major outliers (R^2 = 0.873) was observed between the two independent replicates.

Although highly correlated, there is some variability in the overall magnitude of replication kinetics between the two independent experiments. This can be attributed in part to the difference in passage numbers between the GXR25 cells stocks used in each experiment. In general, GXR25 cells stocks that have been passaged for a period of time greater than 6 months tend to support more efficient replication of Gag-MJ4 chimeras. Therefore, it is advisable to assess replication capacity among groups of chimeras within a one-month time frame. When the previous steps are followed closely, this assay is capable of producing highly robust and reproducible results, which are applicable to a wide range of studies.

A


1
2
3
4
5
2,000
1,500
1,700

B


1
2
3
4
5
1,500
1,000
1,474

C

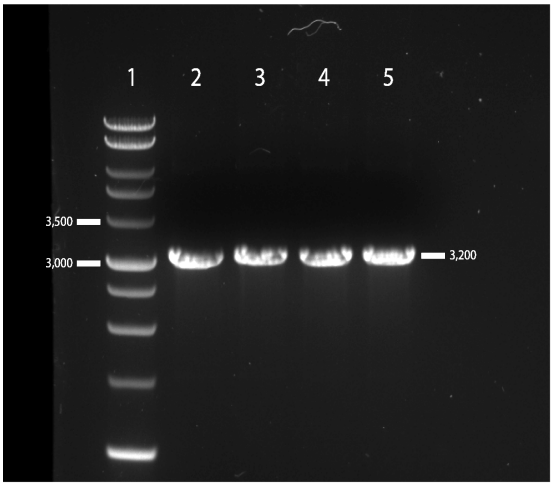

1
2
3
4
5
3,500
3,000
3,200

Figure 1. Representative gel images depicting electrophoretic separation of PCR products. For all PCR products, 5µL of each 50µL reaction was mixed with 3µL of 5x loading dye, loaded into a 1% agarose-TAE gel supplemented with 1X SYBR-safe DNA gel stain, and separated by electrophoresis at 120V for 45 minutes. The Promega 1kb DNA ladder (Lane 1) was used to approximate amplicon sizes **(A)** The *gag* gene was amplified from viral RNA using a nested PCR approach. Due to insertions and deletions, the *gag* amplicon may vary from 1600 – 1700 bp in length and appears slightly above the 1,500 bp DNA ladder marker. Lanes 2 – 5 depict successful *gag* gene amplification. **(B)** The 5' LTR of MJ4 was amplified from the wild-type MJ4 plasmid and visualized via electrophoretic separation. Lanes 2 – 5 depict successful amplification of the 1,474 bp LTR product, which appears slightly below the 1,500 bp DNA ladder marker. **(C)** The 5' LTR derived from wild-type MJ4 and the *gag* gene amplified from patient plasma are fused together via splice-overlap-extension PCR and visualized via electrophoretic separation. Lanes 2 – 5 depict successfully fused amplicons that are approximately 3,200 bp in size, and which appear slightly above the 3,000 bp DNA ladder marker.

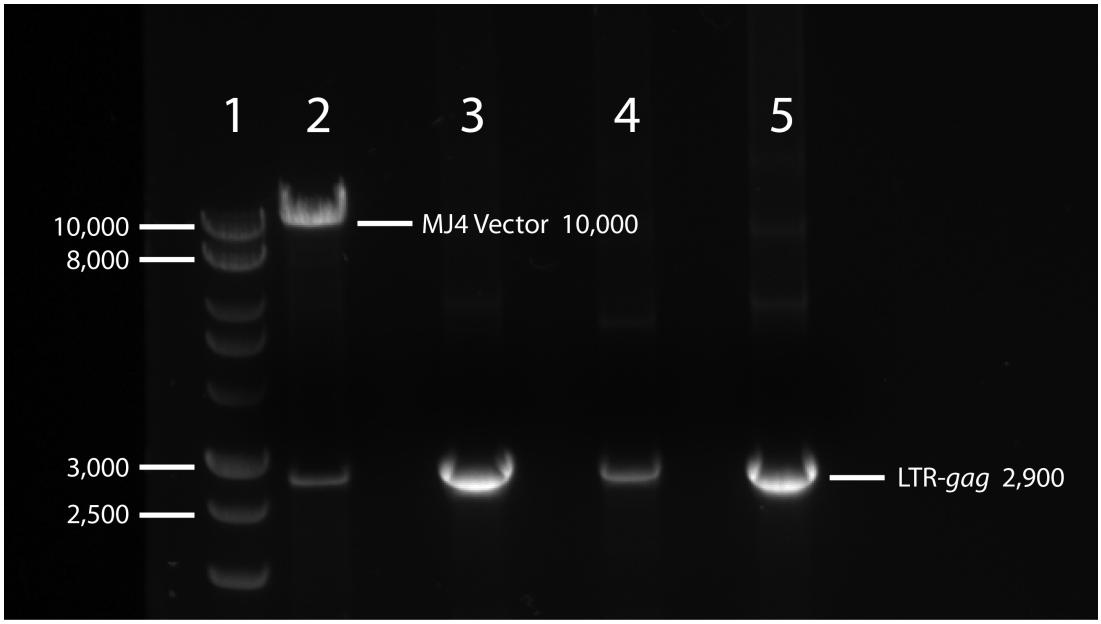


Figure 2. Representative gel image depicting electrophoretic separation of restriction digests for cloning patient *gag* genes into MJ4. Wild-type MJ4 plasmid and LTR-*gag* fusion products were digested with BclI for 1.5 hours at 50°C and NgoMIV for 1 hour at 37°C. Vector and insert fragments were visualized via electrophoretic separation on a 1% agarose-TAE gel supplemented with 1X SYBR-safe DNA gel stain at 100V for 2 hours and using a blue light illuminator in order to reduce UV-induced DNA damage. The vector and insert fragments suitable for subsequent cloning steps appear at approximately 10,000 bp and 2,900 bp respectively.

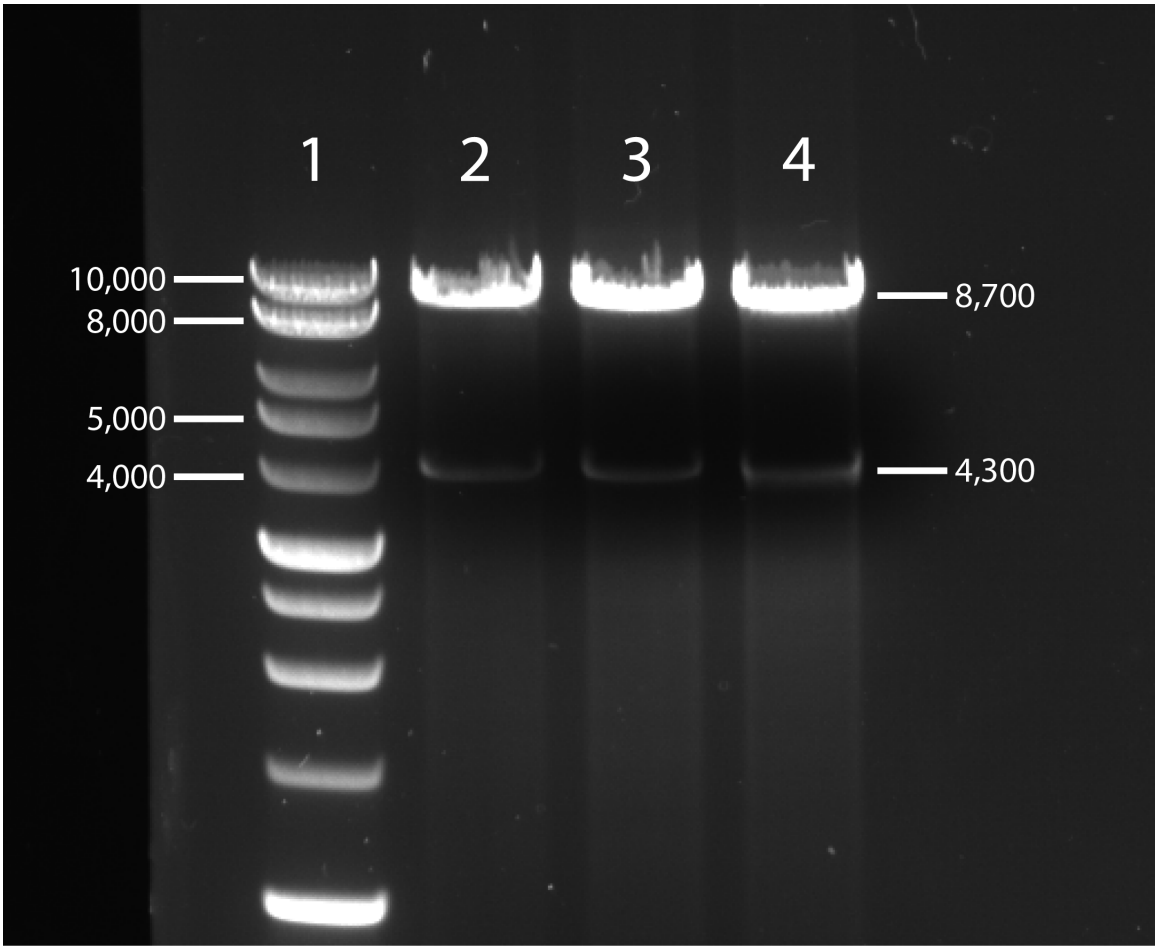


Figure 3. Representative gel image depicting electrophoretic separation of restriction digests of purified Gag-MJ4 chimera plasmid DNA. Purified Gag-MJ4 chimera plasmid DNA was double-digested with NgoMIV and HpaI restriction enzymes for 2 hours at 37°C. Restriction digests were visualized via electrophoretic separation on a 1% agarose-TAE gel supplemented with 1X SYBR-safe DNA gel stain at 120V for 45 minutes. Plasmids without large deletions will resolve to two distinct bands at approximately 8,700 and 4,300 bp.

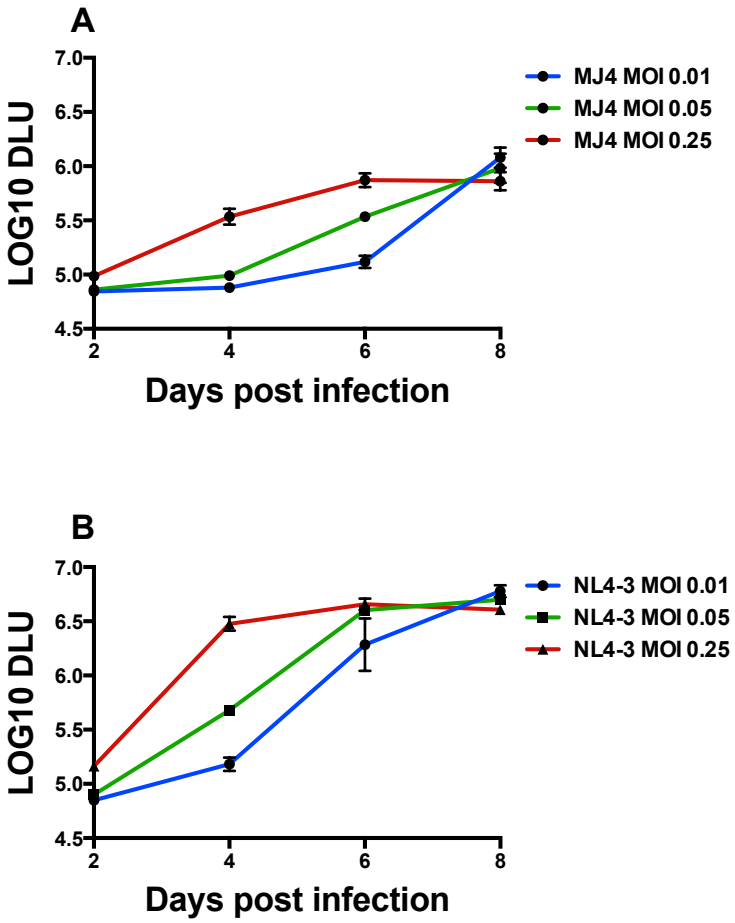


Figure 4. Replication of MJ4 and NL4-3 isolates of HIV-1 in the GXR25 cell line at different multiplicities of infection (MOI). $5x10^5$ GXR25 cells were infected as described in the method protocol with 5-fold increasing MOI of each virus stock. Supernatants were collected on days 2, 4, 6, and 8 post infection and virion production was quantified via a radiolabeled reverse transcriptase assay. Infections were run in triplicate and error bars denote the standard deviation for the three replicates. **(A)** MJ4 **(B)** NL4-3.

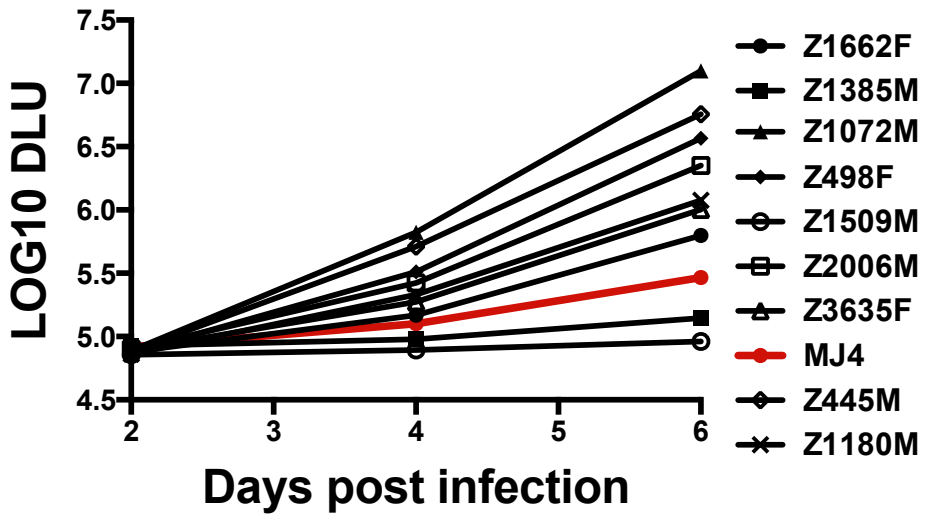


Figure 5. Representative range of replication for different Gag-MJ4 chimeras.

As described in the method protocol, $5x10^5$ GXR25 cells were infected with wild-type MJ4 or Gag-MJ4 chimeras at an MOI of 0.05, supernatants were collected at two day intervals post infection, and virion production quantified by a radiolabeled RT assay. Insertion of various subtype C derived *gag* genes can have a dramatic impact on the replication capacity of MJ4. Wild-type MJ4 replication is denoted in red.

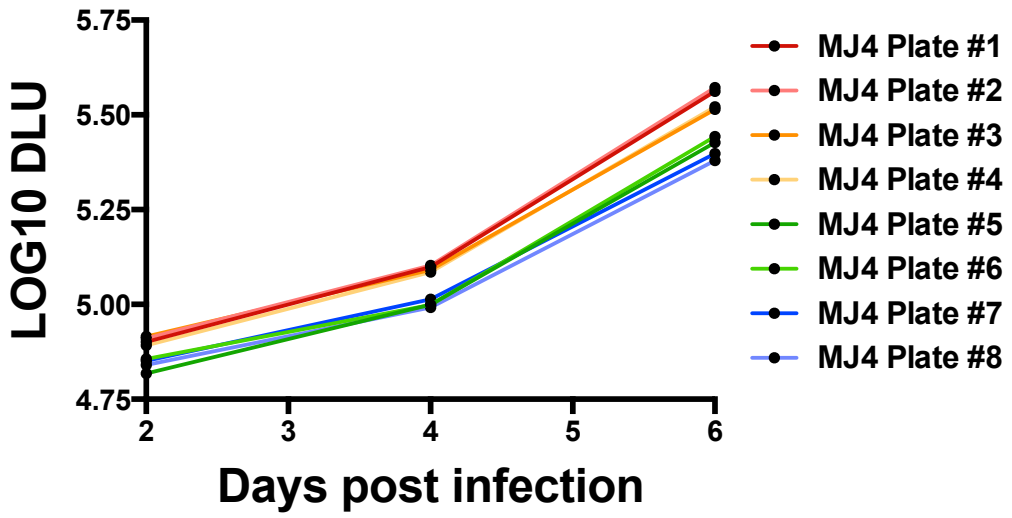


Figure 6. Comparison of intra-assay variation in the radiolabeled reverse transcriptase (RT) quantification assay. The graph depicts inherent intra-assay variability by plotting the DLU values of the same supernatants from a single wild-type MJ4 infection in eight different RT assay plates. The variation in curves reflects the slight changes in signal magnitude between plates, which can be corrected for by running a standard on each RT plate, which can be subsequently used to normalize the slopes of Gag-MJ4 chimeras assayed on the same plate.

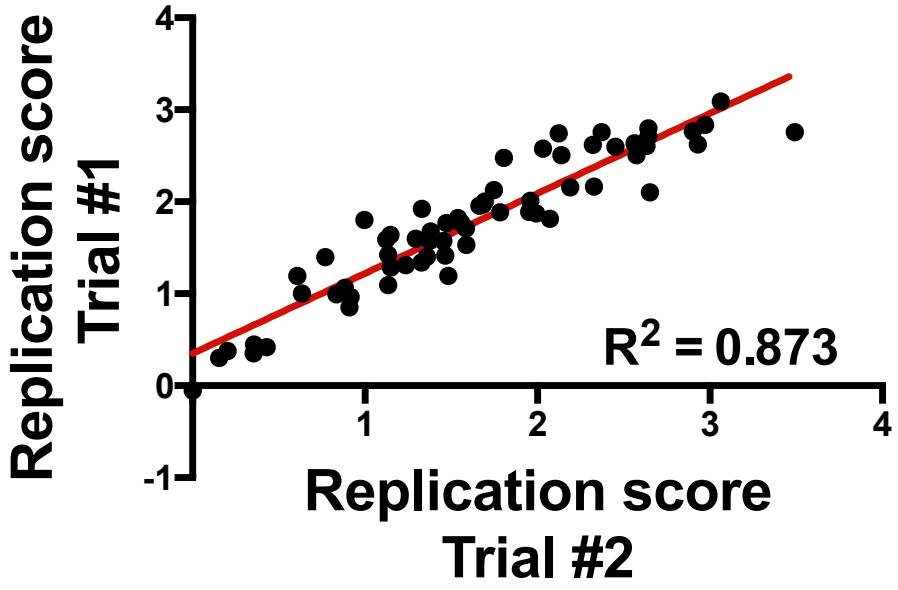


Figure 7. Reproducibility of the replication assay over time in the GXR25 cell line. The same Gag-MJ4 chimeric viruses were used to infect GXR25 cells in two independent experiments performed approximately one year apart. Replication scores were generated by calculating the slope of log-transformed DLU values and normalizing that slope to wild-type MJ4. Gag-MJ4 chimeras that replicate more efficiently than wild-type MJ4 have replication scores greater than 1, and those that replicate less efficiently than wild-type MJ4 have replication scores less than 1. The two independent measurements are strongly correlated (R^2 = 0.87, linear regression) and highlight the reproducibility of assays performed at different times and with cells at different passages.

Discussion:

Due to the length and technical nature of this protocol, there are several steps that are critical for both the successful construction of chimeric Gag-MJ4 plasmids as well as for quantification of viral replication capacity. Although the restriction enzyme based cloning strategy for the introduction of foreign *gag* genes into MJ4 outlined in this protocol has numerous advantages over previously used recombination based methods, the protocol can be technically challenging if critical steps are not followed precisely.

First, it is absolutely essential to use MJ4 plasmid DNA that has been generated in a competent bacterial strain lacking the dcm and dam DNA methylases. This is necessary as the enzymatic activity of the BclI restriction endonuclease, which is used to clone *gag* genes into the MJ4 backbone, is dam/dcm methylation sensitive. The JM110 and SCS110 (an endA negative JM110 derivative) *E. coli* strains are suitable for generating unmethylated MJ4 plasmid DNA. Additionally, for the excision of vector and insert bands, it is highly recommended that a blue light illuminator be used for visualization. This will reduce UV wavelength-dependent DNA damage and drastically increase cloning efficiency. If a blue light illuminator is unavailable, cloning efficiency can be maximized by visualizing DNA with SYBR Safe DNA gel stain instead of ethidium bromide and minimizing UV exposure time.

Finally, molecular cloning with large (>10kb) and/or retroviral plasmids such as MJ4 has traditionally been difficult for a variety of reasons. Large plasmids reduce transformation efficiency of competent bacterial strains[399] while retroviral inserts, which contain long terminal repeat (LTR) sequences, reduce stability of the plasmid and compromise replication fidelity of the plasmid DNA within the bacterial host leading to deletions of the retroviral genome [400]. Bacteria transformed with MJ4 or Gag-MJ4

plasmid products must be grown at 30°C rather than the traditional 37°C for most protocols. Recovery steps after heat shock transformation, growth of transformed bacteria on agar plates, and growth of bacterial colonies in liquid culture should all be performed at 30°C. This lower temperature reduces the growth rate of the bacteria and thus helps to ensure replication fidelity of the MJ4 plasmid. Additionally, replication of the MJ4 plasmid is more stable when using the JM109 *E. coli* strain over the DH5a strain, in our hands. Due to unstable nature of the plasmid, purified plasmid products should always be checked for correct plasmid size by restriction enzyme digestion; here, a double digest with the NgoMIV and HpaI restriction endonucleases at 37°C for 2 hours.

Once successful generation of chimeric Gag-MJ4 plasmids has been accomplished, virus is generated via transfection of 293T cells, titered on an indicator cell line, TZM-bl cells, and replication capacity is measured using a CEM-based T cell line. The CEM-based GXR25 cell line used for these replication studies is one of the few established T cell lines able to support entry and replication of CCR5-tropic strains of HIV-1. This has been achieved by retroviral transduction to allow stable expression of human CCR5 [342]. This cell line naturally expresses CXCR4 and will support replication of CXCR4-tropic HIV-1, such as the laboratory-adapted strain NL4-3. However, in order to support efficient replication of CCR5-tropic strains, such as MJ4, the GXR25 cells must be propagated for no less than 4 months prior to infection. Properly passaged cultures can support replication even after passaging for up to 1 year. Careful monitoring of CCR5-tropic replication throughout passaging is essential for successful experiments.

As with any technique, there are limitations to the protocol that must be considered. Due to the location of restriction sites in the MJ4 plasmid as well as the availability of

conserved restriction sites in naturally occurring HIV-1 isolates, the 3' distal restriction site, BclI, is located 137 nucleotides from the *gag* stop codon. Although this generates a chimeric *protease* gene, this region is 96.5% conserved in this cohort, and we did not observe an abundance of dead or defective chimeric viruses.

One of the advantages of using the MJ4 subtype C infectious molecular clone with subtype C derived sequences is that it reduces the risk of suboptimal gene pairing between *gag* genes and backbone vectors of different subtypes. However, a certain amount of within-clade diversity exists as evidenced by the clustering of HIV-1 sequences by country or region even when found within the same subtype [320]. This could contribute to suboptimal pairing between the *gag* genes derived from acutely infected Zambians and the MJ4 infectious molecular clone backbone, which was derived from a chronically infected individual from Botswana [337]. However, a majority of the analyzed constructs produced infectious progeny virus. As different HIV-1 subtype C infectious molecular clones become more widely available, it will be important to further validate this system by cloning in these HIV-1 clade C *gag* genes into other backbones in order to ensure that there is a minimal bias introduced due to backbone incompatibilities.

The GXR25 cell line is a unique cell line, specifically due to its ability to support both CXCR4 and CCR5-tropic strains and its HIV-1 inducible GFP reporter [342]. However, some limitations exist and should be carefully considered before using this cell line for experiments adapted from this protocol. The GXR25 cell line does not appear to support entry of a majority of subtype C or A, CCR5-tropic, primary isolates we have tested. Additionally, the parent CEM cell line from which the GXR25 cell line was derived exhibits high levels of cyclophilin A, up to 2 to 4-fold higher expression than the Jurkat

cell line [401]. Due to high levels of cyclophilin A, the replication defect normally associated with the canonical HLA-B*57 associated escape mutation, T242N, which is attributed to a decreased ability of capsid to bind cyclophilin A, cannot be easily detected in this particular cell line [286]. Thus the CEM-based GXR25 cell line is not ideal for studying replication defects associated with mutations in the HIV-1 capsid cyclophilin-binding loop.

Finally, while this protocol is ideal for analyzing the replication capacity of *gag* sequences derived from acute time points, modifications must be made in order to study the replication capacity of *gag* genes derived from chronically infected individuals. This protocol involves the amplification of population sequences from acute time points (median 45 days post estimated date of infection) when viral diversity is limited. The *gag* gene from each chimera is then sequenced and compared to the initial population PCR amplicon to ensure cloning fidelity. Due to limited sequence diversity at acute time points, cloning from population PCR products is possible. However, sequence diversity exists within the viral quasispecies of a chronically infected individual; therefore, in order to accurately assess the replication capacity of the chronic quasispecies, single genome amplification must be employed to capture several representative variants. Each of these variants must then be assayed for *in vitro* replication.

This technique has several broader applications, which stem from its advantages over existing methods. Since this process results in a clonal replication competent plasmid, it is simple to use constructs for additional mutagenesis studies, which can help to elucidate the contributions of specific residues to viral replication. Furthermore, by cloning in *gag* genes from longitudinal time points, one can assess the evolution of viral

replication capacity over time and how these changes may affect pathogenesis in an HIV-1 infected individual.

This technique can be modified in order to expand its utility for different applications. Additional HIV-1 viral proteins can be engineered into the MJ4 plasmid in order to assess their effects on viral replication or their interactions with host proteins. This can be accomplished by engineering additional restrictions sites at desired regions in the genome through the introduction of silent nucleotide changes. However, special consideration must be taken when engineering novel restriction sites into the 3' half of the HIV-1 genome as many accessory proteins are encoded in alternate reading frames, and silent changes in one protein could lead to amino acid substitutions in another. In this instance, restriction site independent cloning methods such as those discussed by Dudley et al.[402] may overcome this limitation. Viral sequences derived from different populations may have varying ranges of replication capacities, and the MOI can be adjusted accordingly in order to capture the majority of viral isolates within their logarithmic phase of growth. Finally, the GXR25 cell line has been stably transfected with GFP under an LTR-driven, Tat-inducible promoter, and viral spread can be measured as a function of GFP positive cells via flow cytometry [342] as an alternative to the RT assay described here or a traditional p24 ELISA.

In conclusion, this protocol provides an efficient and powerful technique for assessing HIV-1 viral replication as conferred by the *gag* gene, which encodes a conserved structural protein necessary for proper virion formation, budding, maturation, and disassembly [91,95,352]. Furthermore, this technique generates a replication competent clonal plasmid, which is ideal for mutagenesis studies and, thus, provides a method for elucidating the specific amino acid determinants of viral fitness. Studies such as these

are imperative to enhance the understanding of how immune-driven viral evolution affects pathogenesis and disease progression in HIV-1 infected individuals.

Acknowledgements:

The investigators thank all the volunteers in Zambia who participated in this study and all the staff at the Zambia Emory HIV Research Project in Lusaka who made this study possible. The investigators would like to thank Jon Allen, Smita Chavan, and Mackenzie Hurlston for technical assistance and sample management. We would also like to thank Dr. Mark Brockman for his discussions and generous donation of the GXR25 cells.

This study was funded by R01 AI64060 and R37 AI51231 (EH) and the International AIDS Vaccine Initiative. This work was made possible in part by the generous support of the American people through the United States Agency for International Development (USAID). The contents are the responsibility of the study authors and do not necessarily reflect the views of USAID or the United States Government. This work also was supported, in part, by the Virology Core at the Emory Center for AIDS Research (Grant P30 AI050409). DC and JP were supported in part by Action Cycling Fellowships. This work was supported in part by the Yerkes National Primate Research Center base grant (2P51RR000165-51). This project was also funded in part by the National Center for Research Resources P51RR165 and is currently supported by the Office of Research Infrastructure Programs/OD P51OD11132.

Disclosures: The authors have nothing to disclose.

References:

1 Borrow, P., Lewicki, H., Hahn, B. H., Shaw, G. M. & Oldstone, M. B. Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. *Journal of virology* **68**, 6103-6110 (1994).

2 Koup, R. A. *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* **68**, 4650-4655 (1994).

3 Jin, X. *et al.* Dramatic rise in plasma viremia after CD8(+) T cell depletion in simian immunodeficiency virus-infected macaques. *The Journal of experimental medicine* **189**, 991-998 (1999).

4 Schmitz, J. E. *et al.* Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. *Science* **283**, 857-860 (1999).

5 Brumme, Z. L. *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* **82**, 9216-9227, doi:JVI.01041-08 [pii]10.1128/JVI.01041-08 (2008).

6 Leslie, A. J. *et al.* HIV evolution: CTL escape mutation and reversion after transmission. *Nature medicine* **10**, 282-289, doi:10.1038/nm992nm992 [pii] (2004).

7 Phillips, R. E. *et al.* Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature* **354**, 453-459, doi:10.1038/354453a0 (1991).

8 Price, D. A. *et al.* Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 1890-1895 (1997).

9 Goulder, P. J. *et al.* Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nature medicine* **3**, 212-217 (1997).

10 Tang, J. *et al.* Favorable and unfavorable HLA class I alleles and haplotypes in Zambians predominantly infected with clade C human immunodeficiency virus type 1. *Journal of virology* **76**, 8276-8284 (2002).

11 Prentice, H. A. *et al.* HLA-B*57 versus HLA-B*81 in HIV-1 infection: slow and steady wins the race? *Journal of virology* **87**, 4043-4051, doi:10.1128/JVI.03302-12 (2013).

12 Migueles, S. A. *et al.* HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 2709-2714, doi:10.1073/pnas.050567397050567397 [pii] (2000).

13 Leslie, A. *et al.* Additive contribution of HLA class I alleles in the immune control of HIV-1 infection. *Journal of virology* **84**, 9879-9888, doi:10.1128/JVI.00320-10 (2010).

14 Kaslow, R. A. *et al.* Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. *Nature medicine* **2**, 405-411 (1996).

15 Altfeld, M. *et al.* Influence of HLA-B57 on clinical presentation and viral control during acute HIV-1 infection. *AIDS* **17**, 2581-2591, doi:10.1097/01.aids.0000096870.36052.b6 (2003).

16 Kiepiela, P. *et al.* CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nature medicine* **13**, 46-53, doi:nm1520 [pii]10.1038/nm1520 (2007).

17 Mothe, B. *et al.* Definition of the viral targets of protective HIV-1-specific T cell responses. *Journal of translational medicine* **9**, 208, doi:10.1186/1479-5876-9-208 (2011).

18 Peyerl, F. W., Barouch, D. H. & Letvin, N. L. Structural constraints on viral escape from HIV- and SIV-specific cytotoxic T-lymphocytes. *Viral immunology* **17**, 144-151, doi:10.1089/0882824041310658 (2004).

19 Rolland, M. *et al.* Broad and Gag-biased HIV-1 epitope repertoires are associated with lower viral loads. *PloS one* **3**, e1424, doi:10.1371/journal.pone.0001424 (2008).

20 Wagner, R. *et al.* Molecular and functional analysis of a conserved CTL epitope in HIV-1 p24 recognized from a long-term nonprogressor: constraints on immune escape associated with targeting a sequence essential for viral replication. *J Immunol* **162**, 3727-3734 (1999).

21 Wang, Y. E. *et al.* Protective HLA class I alleles that restrict acute-phase CD8+ T-cell responses are associated with viral escape mutations located in highly conserved regions of human immunodeficiency virus type 1. *Journal of virology* **83**, 1845-1855, doi:10.1128/JVI.01061-08 (2009).

22 Chopera, D. R. *et al.* Transmission of HIV-1 CTL escape variants provides HLA-mismatched recipients with a survival advantage. *PLoS pathogens* **4**, e1000033, doi:10.1371/journal.ppat.1000033 (2008).

23 Crawford, H. *et al.* Evolution of HLA-B*5703 HIV-1 escape mutations in HLA-B*5703-positive individuals and their transmission recipients. *The Journal of experimental medicine* **206**, 909-921, doi:10.1084/jem.20081984 (2009).

24 Boutwell, C. L. *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* **87**, 3952-3965, doi:10.1128/JVI.03233-12 (2013).

25 Brockman, M. A. *et al.* Escape and compensation from early HLA-B57-mediated cytotoxic T-lymphocyte pressure on human immunodeficiency virus type 1 Gag alter capsid interactions with cyclophilin A. *Journal of virology* **81**, 12608-12618, doi:10.1128/JVI.01369-07 (2007).

26 Crawford, H. *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* **81**, 8346-8351, doi:JVI.00465-07 [pii]10.1128/JVI.00465-07 (2007).

27 Martinez-Picado, J. *et al.* Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. *Journal of virology* **80**, 3617-3623, doi:80/7/3617 [pii]10.1128/JVI.80.7.3617-3623.2006 (2006).

28 Schneidewind, A. *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* **81**, 12382-12393, doi:JVI.01543-07 [pii]10.1128/JVI.01543-07 (2007).

29 Wright, J. K. *et al.* Impact of HLA-B*81-associated mutations in HIV-1 Gag on viral replication capacity. *Journal of virology* **86**, 3193-3199, doi:10.1128/JVI.06682-11 (2012).

30 Kawashima, Y. *et al.* Adaptation of HIV-1 to human leukocyte antigen class I. *Nature* **458**, 641-645, doi:nature07746 [pii]10.1038/nature07746 (2009).

31 Goepfert, P. A. *et al.* Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. *The Journal of experimental medicine* **205**, 1009-1017, doi:jem.20072457 [pii]10.1084/jem.20072457 (2008).

32 Brockman, M. A. *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* **84**, 11937-11949, doi:10.1128/JVI.01086-10 (2010).

33 Huang, K. H. *et al.* Progression to AIDS in South Africa is associated with both reverting and compensatory viral mutations. *PloS one* **6**, e19018, doi:10.1371/journal.pone.0019018 (2011).

34 Wright, J. K. *et al.* Gag-protease-mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. *Journal of virology* **84**, 10820-10831, doi:10.1128/JVI.01084-10 (2010).

35 Wright, J. K. *et al.* Influence of Gag-protease-mediated replication capacity on disease progression in individuals recently infected with HIV-1 subtype C. *Journal of virology* **85**, 3996-4006, doi:10.1128/JVI.02520-10 (2011).

36 Adachi, A. *et al.* Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *Journal of virology* **59**, 284-291 (1986).

37 Brockman, M. A., Tanzi, G. O., Walker, B. D. & Allen, T. M. Use of a novel GFP reporter cell line to examine replication capacity of CXCR4- and CCR5-tropic HIV-1 by flow cytometry. *Journal of virological methods* **131**, 134-142, doi:10.1016/j.jviromet.2005.08.003 (2006).

38 Ndung'u, T., Renjifo, B. & Essex, M. Construction and analysis of an infectious human Immunodeficiency virus type 1 subtype C molecular clone. *Journal of virology* **75**, 4964-4972, doi:10.1128/JVI.75.11.4964-4972.2001 (2001).

39 Prince, J. L. *et al.* Role of transmitted Gag CTL polymorphisms in defining replicative capacity and early HIV-1 pathogenesis. *PLoS pathogens* **8**, e1003041, doi:10.1371/journal.ppat.1003041 (2012).

40 Ostrowski, M. A., Chun, T. W., Cheseboro, B., Stanley, S. K. & Tremblay, M. Detection assays for HIV proteins. *Current protocols in immunology / edited by John E. Coligan ... [et al.]* **Chapter 12**, Unit 12 15, doi:10.1002/0471142735.im1205s70 (2006).

41 Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. & Pease, L. R. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61-68 (1989).

42 Inoue, H., Nojima, H. & Okayama, H. High efficiency transformation of Escherichia coli with plasmids. *Gene* **96**, 23-28 (1990).

43 Bichara, M., Pinet, I., Schumacher, S. & Fuchs, R. P. Mechanisms of dinucleotide repeat instability in Escherichia coli. *Genetics* **154**, 533-542 (2000).

44 Ackerson, B., Rey, O., Canon, J. & Krogstad, P. Cells with high cyclophilin A content support replication of human immunodeficiency virus type 1 Gag mutants with decreased ability to incorporate cyclophilin A. *Journal of virology* **72**, 303-308 (1998).

45 Dudley, D. M. *et al.* A novel yeast-based recombination method to clone and propagate diverse HIV-1 isolates. *BioTechniques* **46**, 458-467, doi:10.2144/000113119 (2009).

46 Ganser-Pornillos, B. K., von Schwedler, U. K., Stray, K. M., Aiken, C. & Sundquist, W. I. Assembly properties of the human immunodeficiency virus type 1 CA protein. *Journal of virology* **78**, 2545-2552 (2004).

47 Forshey, B. M., von Schwedler, U., Sundquist, W. I. & Aiken, C. Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *Journal of virology* **76**, 5667-5677 (2002).

48 Gottlinger, H. G., Dorfman, T., Sodroski, J. G. & Haseltine, W. A. Effect of mutations affecting the p6 gag protein on human immunodeficiency virus

particle release. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 3195-3199 (1991).