

©Copyright 2013

Justine Sunshine

Contending with the Challenges of HIV Genetic Diversity:
The Role of Epitope Variant Recognition in Immune Control of HIV

Justine Sunshine

A dissertation submitted in
partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Nicole Frahm, Chair

James I. Mullins

Donald L. Sodora

Program Authorized to Offer Degree:

Program in Pathobiology

University of Washington

Abstract

Contending with the Challenges of HIV Genetic Diversity:

The Role of Epitope Variant Recognition in Immune Control of HIV

Justine Sunshine

Chair of the Supervisory Committee:

Nicole Frahm, Ph.D

Global Health

Despite over 30 years of intensive research, the development of an effective HIV vaccine has remained elusive. One of the most significant challenges in designing an HIV vaccine is contending with the extensive sequence variability found in circulating virus strains. An HIV vaccine will likely require the induction of both HIV-specific antibodies and cytotoxic T lymphocyte (CTL) responses. The efficacy of future HIV vaccines will likely depend on the ability of the induced immune effectors to cope with this widespread sequence diversity. In this thesis, I focused on investigating the capacity of HIV-specific T cells to handle genetic variation during natural HIV infection. In my first aim, I sought to determine the role of epitope variant recognition in immune control of HIV. To this end, I comprehensively assessed epitope variant recognition

in HIV-infected subjects with and without spontaneous control of viral replication by testing for T-cell responses to all Gag variants present in at least 5% of sequences found in the Los Alamos National Laboratory HIV Sequence Database using IFN γ /IL-2 FluoroSpot. Contrary to my expectations, I found that the observed levels of variant recognition were directly associated with viral load, and thus individuals with progressive infection demonstrated a marginally higher ability to recognize epitope variants. However, increased sequence coverage, defined as the overall proportion of HIV database sequences targeted through the Gag-specific repertoire, was found to be inversely associated with viral load. These results highlight that it is the ability to target the most frequently occurring variants, rather than simply a large number of variants, that is associated with control of viral replication. In my second aim, I studied dynamic CTL escape processes following primary HIV infection, with an emphasis on exploring the immune pressure mediated by early cross-reactive CTL responses. In this work, I screened for T-cell responses by IFN γ /IL-2 FluoroSpot using autologous peptide sets reflecting any variant present in at least 5% of Gag sequence reads from six linked HIV transmission pairs. Escape processes were found to be highly variable, with limited overlap in the patterns of escape between individuals. By testing for responses to variants that emerged in the viral population over time, I found that 70% of responses to evolving epitopes demonstrated some degree of epitope variant recognition. Interestingly, the variants that persisted in the viral population were recognized with a significantly lower functional avidity than responses to founder sequence, suggesting functional avidity as a surrogate for effective variant recognition. Several vaccine strategies designed to induce T-cell responses of enhanced variant recognition capacity are currently moving into phase I clinical trials. Taken together, these results provide evidence for sequence coverage and functional avidity to serve as critical measures to determine the protective potential of variant-inclusive vaccines.

Table of Contents

LIST OF FIGURES	VIII
LIST OF TABLES	X
ABBREVIATIONS	XI
CHAPTER I: INTRODUCTION AND LITERATURE REVIEW	1
<i>Overview.....</i>	<i>2</i>
<i>HIV Biology and Pathogenesis</i>	<i>3</i>
<i>Function of the Gag protein</i>	<i>7</i>
<i>Generation of HIV-specific CTL responses</i>	<i>10</i>
<i>Role of CTLs in immune control of HIV-1</i>	<i>12</i>
<i>HIV-1 genetic diversity.....</i>	<i>16</i>
<i>Challenges of HIV-1 diversity on CTL mediated immune control and vaccine development</i>	<i>19</i>
<i>Current vaccines designed to contend with HIV-1 diversity</i>	<i>21</i>
<i>Thesis aims</i>	<i>23</i>
CHAPTER II: INCREASED SEQUENCE COVERAGE THROUGH COMBINED TARGETING OF VARIANT AND CONSERVED EPITOPES CORRELATES WITH CONTROL OF HIV REPLICATION	25
ABSTRACT:	26
INTRODUCTION	27
MATERIALS AND METHODS	29
<i>Study Subjects.....</i>	<i>29</i>
<i>Peptide Set</i>	<i>30</i>
<i>IFN-γ/IL-2 FluoroSpot</i>	<i>30</i>
<i>Gag Sequencing.....</i>	<i>31</i>
PCR products were visualized using the QIAxcel DNA Fast Analysis Kit (Qiagen) and PCR positive samples were purified using NucleoSpin Gel and PCR Clean-Up Kit (Clontech) before being submitted directly for dideoxynucleotide chain termination (Sanger) sequencing. Chromatograms were assembled and edited using Geneious 5.6.5 (Biomatters, NZ).....	32
<i>Data Analysis.....</i>	<i>32</i>
<i>Statistical Analysis</i>	<i>34</i>
RESULTS	35
<i>Significant differences in the quality but not the quantity of Gag-specific responses between Controllers and Progressors.....</i>	<i>35</i>
<i>Controllers and Progressors target conserved and variable regions in Gag</i>	<i>38</i>
<i>Higher epitope variant recognition observed in Progressors.....</i>	<i>40</i>

<i>Epitope variant recognition is associated with viral load in Viremic Controllers and Progressors</i>	<i>41</i>
<i>Greater diversity in Gag epitopes found in Progressors</i>	<i>43</i>
<i>Coverage of frequently occurring Gag variants is associated with control of viral replication</i>	<i>45</i>
<i>Increased immune coverage is correlated with recognition of more founder viruses</i>	<i>46</i>
DISCUSSION	47
SUPPLEMENTARY MATERIALS:	52
CHAPTER III FITNESS-BALANCED ESCAPE DETERMINES RESOLUTION OF DYNAMIC CTL ESCAPE PROCESSES DURING PRIMARY HIV-1 INFECTION	58
ABSTRACT	59
INTRODUCTION	60
MATERIALS AND METHODS	62
<i>Study Subjects</i>	<i>62</i>
<i>Gag Sequencing</i>	<i>63</i>
<i>Peptide Set</i>	<i>64</i>
<i>IFN-γ/IL-2 FluoroSpot</i>	<i>65</i>
<i>Functional Avidity</i>	<i>66</i>
<i>Generation of mutant NL4-3 viral stocks</i>	<i>66</i>
<i>Pairwise Growth Competition Assay</i>	<i>68</i>
<i>Data Analysis</i>	<i>69</i>
<i>Statistical Analysis</i>	<i>69</i>
RESULTS	69
<i>Detection of primarily IFNγ-secreting Gag-specific responses in both Transmitting and Seroconverting Partners</i>	<i>69</i>
<i>CTL pressure is a major selective force shaping early Gag evolution</i>	<i>72</i>
<i>Evolution within TP targeted epitopes do not reflect classical reversion patterns</i>	<i>73</i>
<i>Entropy, functional avidity, and immunodominance in responses to evolving epitopes</i>	<i>76</i>
<i>Highly dynamic T-cell escape processes observed during early Gag evolution</i>	<i>78</i>
<i>Replacement variants are recognized with significantly lower avidity</i>	<i>87</i>
<i>Fixation of replacement variants reflects fitness-balanced escape processes</i>	<i>88</i>
DISCUSSION	91
CHAPTER IV DISCUSSION:	96
<i>Summary of results and caveats to consider</i>	<i>97</i>
<i>Sequence coverage as an evaluation tool of variant-inclusive vaccines</i>	<i>99</i>
<i>Immune correlates of effective epitope variant recognition</i>	<i>103</i>
<i>Variant-inclusive vaccines in therapeutic vaccination</i>	<i>106</i>
<i>Concluding remarks</i>	<i>109</i>

ACKNOWLEDGEMENTS	110
REFERENCES	112

List of Figures

	Page
Figure 2-1: Breadth and functionality of Gag-specific T-cell responses in 15 controllers and 15 progressors.....	37
Figure 2-2: Controllers and progressors target conserved and variable regions of Gag.....	39
Figure 2-3: Higher variant recognition in progressors.....	41
Figure 2-4: Viral load and variant recognition in viremic subjects.....	42
Figure 2-5: Viral diversity is higher in progressors than in VC.....	44
Figure 2-6: Increased sequence coverage is correlated with control of viral replication and recognition of founder viruses.....	46
Supplementary Figure 2-1: Controllers and progressors similarly targeted conserved epitopes.....	52
Supplementary Figure 2-2: Epitope variant recognition is extensive both within and between individuals.....	53
Supplementary Figure 2-3: Overall variant recognition is similar between individuals with and without B*27.....	54
Supplementary Figure 2-4: Functionality changes upon variant recognition.....	55
Supplementary Figure 2-5: Sequence coverage for all epitopic regions targeted by controllers and progressors.....	56
Supplementary Figure 2-6: Recognition of founder viruses by controllers and progressors.....	57
Figure 3-1: Breadth of Gag-specific responses in transmitting and seroconverting partners.....	72
Figure 3-2: Evolution of potentially reverting sites in SP viral populations.....	76
Figure 3-3: Shannon entropy, functional avidity, and immunodominance are associated with responses to evolving epitopes.	79
Figure 3-4: Cross-reactivity within targeted epitopes and selection of escape variants.....	81
Figure 3-5: Escape within PIC64236 targeted epitope QW9.....	82
Figure 3-6: Escape mutations promote responses to new epitopes in PIC64236.....	83
Figure 7-7: Dual selective pressures of reversion and escape within PIC44149	85

Figure 3-8: Cross-reactive responses detected in PIC68008 and PIC51861.....	87
Figure 3-9: Cross-reactive responses to epitopes GR11 and NL11 in PIC99203.	88
Figure 10: Functional avidity differences between responses to founder, replacement and transient variants.	89
Figure 3-11: HIVDB frequency differences between founder, transient, and replacement variants.	90
Figure 3-12: Pairwise fitness competitions with variants detected in PIC64236 RK9 and PIC68008 TW10 epitope evolution.....	92

List of Tables

	Page
Table 2-1: Example of data analysis	33
Table 2-2: Subject characteristics.....	35
Table 3-1: PIC partner pair characteristics	71
Table 3-2: Early Gag evolution in seroconverting partner	73
Table 3-3: SP Gag-specific responses within first year of infection.....	77

Abbreviations

AA= amino acid

AIDS= acquired immunodeficiency syndrome

APC= antigen presenting cell

ARV= antiretrovirals

ART= antiretroviral therapy

CA= Capsid

CE= Conserved element

CD= Cluster of differentiation

COT= Center-of-tree

CTL= Cytotoxic T lymphocyte

CRF= Circulating recombinant form

DPS= Days post onset of symptoms

EC50= Half maximal effective concentration

ELISPOT= Enzyme-linked immunosorbent spot

IFN= Interferon

IL= Interleukin

LTNP= Long-term non-progressor

LANL= Los Alamos National Laboratory

HIV= Human immunodeficiency virus

HIVDB= Los Alamos National Laboratory HIV sequence database

HLA= Human leukocyte antigen

MA= Matrix

MHC= Major histocompatibility complex

PMBC= Peripheral blood mononuclear cells

NC= Nucleocapsid

NHP= Non-human primate

SIV= Simian immunodeficiency virus

TCR= T cell receptor

VL= viral load

Chapter I:

Introduction and literature review

Overview

In 1983, the etiological agent behind Acquired Immunodeficiency Syndrome (AIDS) was discovered to be a retrovirus, Human Immunodeficiency Virus (HIV) (Barre-Sinoussi et al., 1983). By 2012, approximately 34 million individuals are living with HIV, 69% of whom reside in Sub-Saharan Africa (2012). Over 30 years of intensive research has provided a wealth of information regarding the origin, biology, and pathogenesis of this virus. However, despite rigorous efforts, we still do not have a commercially available HIV vaccine. With over 2.5 million new infections each year (2012), there is a pressing need for the development of a cost-effective and highly efficacious preventative vaccine.

One of the most significant challenges in the development of an HIV vaccine is the immense sequence diversity found in circulating HIV strains that must be covered by the induced immunity. With global HIV isolates capable of differing from one another by up to 35% in some proteins (Gaschen et al., 2002), efficacy of future HIV vaccines will likely depend on the capacity of the induced immunity effectors to cope with this extensive sequence diversity. One major goal in HIV vaccinology is to induce broadly neutralizing HIV-specific antibodies capable of recognizing a wide range of viral strains. However, no immunogen that has been tested to date has been successful in the induction of broadly neutralizing antibodies (Burton et al., 2012). Additionally, many of the broadly neutralizing antibodies isolated during natural HIV infection have developed after years of extensive somatic hypermutation and selective fine-tuning (Liao et al., 2013; Wu et al., 2011). Despite these difficulties, attempting to recreate these developmental steps via a short-term vaccination protocol remains a top priority in vaccine development.

Observational and correlative studies have strongly implicated a role for HIV-specific cytotoxic T lymphocyte (CTL) responses in control of viral replication (Kaslow et al., 1996; Kiepiela et al., 2007; Koup et al., 1994; Rolland et al., 2008; Schmitz et al., 1999). As such, the

induction of potent HIV-specific CTL responses has been a high priority in the development of an HIV vaccine. In order to contend with sequence variation found in circulating strains, two main vaccinology strategies have been proposed: induction of HIV-specific T-cell responses that (i) target conserved viral regions (Letourneau et al., 2007; Rolland et al., 2007) or (ii) are capable of recognizing a high number of epitope variants (Fischer et al., 2007; Nickle et al., 2007). Each of these strategies are currently moving into phase I human clinical trials. However, the ability of HIV-specific CTLs to cope with sequence variation during natural infection has not been well characterized. As such, the metrics with which to evaluate the immunogenicity and protection potential of these vaccines are not established. In this dissertation, several complimentary research projects are described that aimed to gain a better understanding of the ability of HIV-specific CTLs to tolerate sequence diversity during natural infection as well as identify immune correlates that will assist in vaccine evaluation.

HIV Biology and Pathogenesis

In the current HIV pandemic, HIV-1 is responsible for the majority of global infections whereas HIV-2 is primarily found in West Africa. These distinctive viral lineages are the result of separate cross-species transmission from Simian Immunodeficiency Virus (SIV) infected primates into human populations. Three separate transmission events from SIV infected chimpanzees (SIV_{cpz}) likely gave rise to the three HIV-1 groups: M (major, clades A-K), O (outlier) and N (non-M, non-O) (Montano, 2002; Zhu et al., 1998). The HIV-2 epidemic was likely the result of nine different transmission events from SIV infected sooty mangabeys (SIV_{smm}) (Christian Apetrei, personal communication). Given the results presented here are derived from clade B HIV-1 infections, the focus of this dissertation will be on HIV-1.

HIV-1 infects target cells by docking the surface envelope (Env) protein gp120 onto the CD4 receptor, a glycoprotein found on the surface of T lymphocytes, monocytes, macrophages, and dendritic cells (Montano, 2002). Binding to CD4 leads to a conformational change in gp120

that exposes the co-receptor binding site, and binding to the co-receptor (primarily CCR5 and in some cases, CXCR4 or other chemokine receptors). Further conformational changes within Env gp120 and the transmembrane Env protein gp41 promote fusion of the virus onto the target cell, where the viral capsid containing the replicative machinery is delivered (Montano, 2002).

Following uncoating of the capsid, the contained reverse transcriptase (RT) begins the process of converting viral RNA into proviral DNA (Montano, 2002). The viral enzyme integrase (IN) then facilitates incorporation of the synthesized proviral DNA (as part of a pre-integration complex) into the host chromosomes, thus permanently infecting the cell (Montano, 2002). Interestingly, one mechanism in which HIV-1 has evolved to take advantage of host machinery is to include a DNA binding site for transcription factor nuclear factor kappa B (NFκB) within the 5' long-terminal repeat (LTR). Upon activation of the infected cell (e.g. in response to viral infection), NFκB will enter the nucleus and promote transcription of the integrated viral DNA (Nabel and Baltimore, 1987). mRNA encoding for the accessory proteins Tat and Rev are first transcribed and later facilitate the transcription of alternative splice variants that allow for translation of all HIV-1 proteins (Montano, 2002). Polyproteins Gag (structural proteins) and Gag-Pol (enzymatic proteins), along with viral RNA and other viral and host factors, assemble with Env at the plasma membrane, eventually budding off the host cell to form an immature viral particle (Montano, 2002). After release from the host cell, HIV protease cleaves the Gag polyprotein and an infectious mature virus is produced (Montano, 2002). Remarkably, over 10^{10} viral particles are generated in an infected individual every day (Simon and Ho, 2003). Furthermore, some infected cells remain latent and do not actively produce virus, thus serving as a dormant reservoir within an infected individual (Donahue and Wainberg, 2013).

Sexual transmission is the most common mode of transmission in the current HIV epidemic (Simon et al., 2006). In the majority of these sexual transmissions, a single HIV-1 variant establishes infection following penetration of the mucosal barrier (Abrahams et al., 2009;

Herbeck et al., 2011; Keele et al., 2008; Shankarappa et al., 1999; Zhu et al., 1993). While CD4⁺ T-cells are the main source of target cells throughout HIV infection, the first target cell types at the site of infection remain controversial. Detailed investigations into the earliest dynamics of SIV infection has suggested that within the mucosa at the site of infection, multiple foci of CD4⁺/CCR5⁺ cells become infected and support local virus replication (Estes, 2013). However, often only one of these infected foci is the source for the systemic infection, although whether this phenomenon is the result of selective or stochastic forces remains unknown (Estes, 2013; McMichael et al., 2010). Exponential viral growth and dissemination begins with the transportation of HIV-1 from the site of infection into the draining lymph node, where large quantities of activated CD4⁺ T cells are available for infection (Estes, 2013; McMichael et al., 2010). As the virus predominately infiltrates gut-associated lymphoid tissues (GALT), a massive depletion of CD4⁺ T-cells in the GALT occurs within the first three weeks of infection as a result both productive infection and bystander activation and apoptosis (Klatt et al., 2013; Mattapallil et al., 2005). Unfortunately, this early depletion of CD4⁺ T cells in the gut is not reversed throughout infection even if an individual begins antiretroviral therapy (Mavigner et al., 2012).

Approximately 14 days post infection, the viremia within an individual hits a peak, which often clinically manifests with acute flu-like symptoms (McMichael et al., 2010; Schacker et al., 1996). Viremia begins to decline and later resolves into a viral load set point following 12-20 weeks of infection (McMichael et al., 2010; Rolland, 2013). Viral load set point is believed to be the result of a constant battle between ongoing viral replication versus innate and adaptive immune responses (McMichael et al., 2010; Rolland, 2013). Within days of infection, a cytokine storm of anti-viral compounds is released throughout the body (Borrow, 2011). Although these responses fail to contain the spread of infection, these earliest responses help develop the initial adaptive immune response, primarily mediated by CTLs but also including non-neutralizing binding antibodies that function in a variety of ways, including allowing for potent NK cell

responses (McMichael et al., 2010). Autologous neutralizing antibodies develop later in infection, generally occurring after three months of infection (Gray et al., 2007; Wei et al., 2003).

One of the most fascinating aspects of HIV pathogenesis is that individuals can remain asymptomatic for many years despite high levels of ongoing viral replication. Clinical definition of progression to AIDS is characterized by a CD4 count of <200 cells/ μ l and/or diagnosis of a AIDS-defining illness (i.e. Kaposi sarcoma, *Pneumocystis jiroveci*, toxoplasmosis of the brain, etc.) (Organization, 2007). In 1999, a seminal paper revealed that the levels of T-cell immune activation were more predictive of survival than virus burden, suggesting that it is generalized immune activation that promotes progression to AIDS (Giorgi et al., 1999). To support this, intensive investigations into characterizing non-progressive SIV infection of different natural non-human primate (NHP) hosts demonstrated that while these infections included modest depletion of mucosal CD4⁺ T cells upon SIV infection and high levels of plasma viremia, the levels of generalized immune activation were much lower compared to progressive infections (Gordon et al., 2007; Milush et al., 2011; Mir et al., 2011). Immune activation can manifest in many different ways including high levels of proinflammatory cytokines, loss of naïve and central memory T cells, hyperactivation and dysregulation of B cells, and alterations of different dendritic cell subsets (Klatt et al., 2013). Chronic stimulation of the immune system through viral replication is believed to be a direct cause of immune activation (Klatt et al., 2013). Additionally, as a consequence of the loss of gut integrity, microbial translocation of gastrointestinal products into systemic circulation has also been proposed to promote immune activation (Brenchley et al., 2006; Estes et al., 2010). The high levels of immune activation in progressive infection are likely the consequence of multiple complex viral and host mechanisms.

Amazing achievements in the development of antiretroviral treatments have drastically improved both the quality of life and survival time for HIV-infected individuals. However, even with treatment, persistent immune activation still occurs and has been linked to increased

morbidity and mortality compared to uninfected individuals (Kuller et al., 2008). Therefore, there have been great efforts in the field to develop cure therapies that can ablate all residual viremia or at least, induce a 'functional cure' that prevents disease progression. Additionally, with millions of new infections each year (2012), advancements in preventative methods are imperative to halt the epidemic. Therapeutic and prophylactic vaccination has the potential to make significant contributions towards eliminating the global burden of HIV. Identifying immune correlates of protection against disease progression is essential in order to better design future vaccine strategies.

Function of the Gag protein

HIV-1 possesses a *gag* gene that encodes for a polyprotein responsible for the structural components of the virion. The Gag protein is one of the most abundant proteins produced during HIV infection, with thousands of Gag proteins produced during replication cycles. The Gag precursor (p55) is comprised of four domains and two spacer proteins, all of which play critical roles in HIV-1 viral particle structure and virus replication.

Matrix (MA, p17)

The matrix (MA, p17) domain begins at the N-terminus of the Gag polyprotein and is comprised of 132 amino acids. After translation, the second position in MA (glycine) that is myristoylated and serves as a signal for the Gag polyprotein being trafficked and anchored to the plasma membrane (Bryant and Ratner, 1990). Within the first 30 amino acids lies a Highly Basic Region (HBR) comprised of multiple arginine and lysine residues that interact favorably with the negative charges of the phospholipid membrane in order to stabilize targeting (Chukkapalli and Ono, 2011). After virion assembly, protease cleaves MA from the polyprotein, allowing for MA to provide structure to the virion by associating with the lipid membrane derived from the host cell (Briggs and Krausslich, 2011). Following infection of a new target cell, MA has

also been proposed to assist in the transport of the pre-initiation complex into the host nucleus (Haffar et al., 2000).

Capsid (CA, p24)

Capsid (CA, p24) is the largest domain of the Gag polyprotein, comprised of 235 amino acids. Across all HIV-1 clades, CA is one of the most conserved components of the HIV genome. Numerous studies investigating mutational tolerance within CA have highlighted the fragility of this domain (Boutwell et al., 2009; Manochewa et al., 2013; Martinez-Picado et al., 2006; Rihn et al., 2013; Schneidewind et al., 2008; Troyer et al., 2009). In particular, one recent study using a library of single AA substitutions demonstrated that 70% of mutations in CA resulted in replication-defective viruses (Rihn et al., 2013). Given this genetic vulnerability, CA can be considered the 'Achilles heel' of the virus and has been widely studied in order to inform both vaccine and drug design.

CA consists of a N-terminal domain (NTD, 146 amino acids) and a C-terminal domain (CTD, 85 amino acids) with a linker region of four amino acids connecting the two components (Rihn et al., 2013). During assembly, the CTD promotes dimerization of Gag polyproteins, eventually resulting in higher order Gag oligomers (Montano, 2002; Morikawa et al., 1999). The major homology region (MHR), a feature conserved in all retroviruses, confers conformational stability of Gag oligomers (Montano, 2002). Following proteolytic cleavage, the mature capsid forms a conical core in the HIV virion that encapsulates the viral RNA/nucleocapsid complex as well as the replication machinery (Briggs and Krausslich, 2011). Approximately 1,100 cleaved capsid proteins arrange in a hexagonal lattice formed by 250 hexameric subunits and 12 pentamers (Briggs and Krausslich, 2011). NTD-NTD interactions mediate intra-hexamer contacts whereas CTD-CTD associations bring together adjacent hexamers (Rihn et al., 2013). CA also plays a critical role in infectivity and viral uncoating. CA NTD binds to host enzyme cyclophilin A (cypA) to determine nuclear importation, integration targeting, and replication

efficiency (Schaller et al., 2011). Additionally, CA is the target of intrinsic immune factor TRIM5 α , a host protein that restricts viral replication (Nakayama and Shioda, 2010). In order to infect human populations, HIV-1 CA had to evolve to overcome this TRIM5 α restriction (Nakayama and Shioda, 2010). The numerous requirements for CA-CA and CA-host interactions throughout the virus lifecycle highlight the structural and functional constraints that limit mutational tolerance.

Nucleocapsid (NC, p7)

Comprised of 55 amino acids, the nucleocapsid domain (NC, p7) plays a critical role in viral genome packaging. HIV-1 NC contains a zinc-binding site that functions to bind full-length HIV-1 viral RNA (Montano, 2002). Interestingly, swapping the NC domains between some retroviruses results in differential RNA packaging, indicating that each NC contains motifs that allow for preferential targeting of its specific genome (Lu et al., 2011). NC also helps facilitate dimerization of HIV-1 viral RNA (Lu et al., 2011). Although NC is believed to be critical for proper HIV-1 RNA incorporation into the genome, other viral (e.g., MA) and cellular (e.g., P-bodies) components may also assist in HIV-1 genome packaging (Lu et al., 2011; Reed et al., 2012). Within the mature virion, NC is in a complex with the viral RNA within the core (Montano, 2002). Following viral uncoating, NC assists in the synthesis of proviral DNA (Sleiman et al., 2012).

p6

The p6 domain lies at the C-terminal portion of the Gag polyprotein. Compared to other Gag domains, the functions mediated by p6 are the least characterized. The most well described role of p6 is its involvement in viral budding. The p6 protein contains two late-budding (L) domains, PTAP and YP(X)_nL. These L-domains permit hijacking of the endosomal-sorting complex required for transport (ESCRT) pathway, which then promotes budding of HIV virions

from the host cell (Bieniasz, 2009). The PTAP motif binds ESCRT protein Tsg101 (Garrus et al., 2001) whereas the YP(X)_nL interacts with the ALIX protein (Strack et al., 2003). Mutation of one L-domain severely affects budding efficacy of HIV-1, indicating that both domains are critical for proper virus release (Hahn et al., 2011). However, these deficiencies may be overcome by *in vitro* overexpression of the partner ESCRT protein of the domain that is not mutated (Hahn et al., 2011). Not surprisingly, these motifs are well conserved in HIV-1 sequences.

p1/p2 (spacer proteins)

p1 and p2 are short spacer domains between NC/p6 and CA/NC, respectively. These spacer domains contain sites where the HIV-1 protease cleaves the polyprotein to produce the separate Gag proteins during mature HIV-1 formation (Montano, 2002). An additional site lies between MA and CA but does not include a spacer region. Aside from the protease cleavage motifs, these spacers are highly variable among HIV-1 strains. The role that these regions may play during viral assembly outside of proteolytic cleavage remains poorly understood.

Generation of HIV-specific CTL responses

CTL responses directed against the virus(es) able to establish systemic infection (founder virus) develop within the first month of acute HIV-1 infection (Allen et al., 2005a; Herbeck et al., 2011; Liu et al., 2006; McMichael et al., 2010). The development of an HIV-specific CTL response first requires professional antigen presenting cells (APC), most often dendritic cells, to prime naïve CTLs (Janeway, 2005b). These processes usually begin at the site of infection, where viral replication at infected foci inevitably stimulates the activation of immature dendritic cells, which then uptake HIV antigens and migrate to the draining lymph node (Janeway, 2005b). Additionally, if infectious virus infiltrates lymph nodes, resident dendritic cells are also capable of antigen uptake. Activated mature dendritic cells will process ingested HIV antigens, referred to as peptides or epitopes, which are then presented on MHC class I

(Janeway, 2005a). Circulating naïve CTLs that express T-cell receptors (TCRs) capable of specifically recognizing the presented peptides by dendritic cells will then begin to differentiate into activated antigen-specific CTLs (Janeway, 2005b). Importantly, dendritic cells must also deliver the co-stimulatory signal CD80/86, which engages CD28 on the naïve CTL in order to permit activation and autocrine production of IL-2, a cytokine necessary for T-cell proliferation (Janeway, 2005b). Additionally, the APC delivers a third signal of pro-inflammatory cytokines and/or Type I interferons that promote differentiation into an effector phenotype that allows the CTL to kill infected cells and secrete inflammatory and antiviral compounds (Curtsinger and Mescher, 2010; Kaech et al., 2002). Following activation, there is a rapid expansion of the HIV-specific CTL and subsequent migration out of the lymph node to target infected cells in the periphery (Janeway, 2005b).

Infected CD4⁺ T-cells (or other CD4⁺ infected cell types) signal infection to circulating HIV-specific CTLs by presenting HIV epitopes on MHC class I at the surface of the cell (Janeway, 2005a). Epitope presentation begins following the degradation of HIV proteins by the immunoproteasome, specialized cellular machinery that is frequently generated in response to type I IFN signaling occurring as a result of localized viral infection (Janeway, 2005a). The immunoproteasome preferentially generates short peptide fragments that are then shuttled into the endoplasmic reticulum (ER) through the transporter associated with antigen processing (TAP) proteins (Janeway, 2005a). Following additional trimming, these peptides are loaded onto MHC class I and transported to the cell surface via the secretory pathway (Janeway, 2005a). The peptide binding cleft in MHC class I permits binding of epitopes generally between 8-11 amino acids in length (Janeway, 2005a). Each MHC class I allele has differential epitope binding motifs, which are determined by the class of amino acids in the second and C-terminal position of the peptide (Llano, 2009). For example, A*01 alleles preferentially bind peptides containing a negatively charged amino acid in the second position (p2) and a tyrosine in the C-

terminal position (C-term). Importantly, CTL TCRs recognize the MHC-epitope complex as a whole and as such each antigen-specific CTL is restricted by a specific HLA allele (Janeway, 2005b). The first, last, and number of amino acids in a peptide determine the epitope name (Llano, 2009). For example, epitope TSTLQEQIGW is referred to as TW10 and antigen-specific CTLs recognizing this peptide sequence are named TW10-specific CTLs.

Engagement of the HIV-specific CTL TCR to the MHC-epitope complex induces a polarization of the CTL that allows formation of an immunological synapse to permit effective delivery of cytotoxic compounds into the infected cells (Janeway, 2005b). CTL delivery of perforin produces holes within the target cell membrane, allowing for delivery of granzyme B into the cytoplasm, which initiates apoptotic pathways, ultimately resulting in the destruction of the cell (Janeway, 2005b). Additionally, CTLs can mount non-lytic responses in which upon peptide recognition, the CTL releases a variety of anti-viral cytokines including IFN γ and TNF α . These cytokines can function by signaling nearby cells to switch on antiviral defense mechanisms, including the cessation of intracellular protein synthesis (Janeway, 2005b).

Role of CTLs in immune control of HIV-1

In 1987, HIV-specific CTL responses were first identified in infected individuals (Walker et al., 1987). Later, it was discovered that the decline in peak viremia during acute HIV infection was temporally associated with the appearance of founder-specific CTL responses, indicating that these CTLs were capable of effectively eliminating infected CD4⁺ cells and thus may be important in immune control (Koup et al., 1994). To support this, in 1999 Schmitz *et al* demonstrated that depletion of CTLs in rhesus macaques during SIV infection caused a rapid increase in viremia that only subsided following reconstitution of the CTL population (Schmitz et al., 1999). Furthermore, there is an enrichment of specific HLA class I alleles (notably B*27 and B*57) in individuals with controlled viral replication and delayed disease progression (Kaslow et al., 1996). This rare and heterogeneous subset of individuals are collectively referred to as HIV-

controllers or in some cases, Long-Term Non-Progressors (LTNPs), when individuals have consistent control of viral replication for long periods of time (>11 years). Since these seminal investigations, there have been immense efforts in the field to further define the role of HIV-specific CTLs in control of viral replication and to identify immune correlates of protection from disease progression that can be translated into vaccine design.

Advances in the development of the IFN γ -ELISPOT, a high throughput assay that detects secretion of the frequently CTL produced cytokine in response to stimulation by HIV antigens, enabled widespread screening of HIV-specific CTL responses in infected populations. Surprisingly, two independent studies concluded that neither the breadth nor magnitude of total HIV-specific CTL responses was correlated with viral load (Addo et al., 2003; Cao et al., 2003), an observation that lead to more detailed investigations in the differences between HIV-specific CTL responses in individuals with and without immune control. In over ten years of intensive research efforts, there is still no consensus on the exact CTL parameters that determine immune control. However, the most consistent correlates of immune control include differential targeting of HIV proteins as well as qualitative differences in the CTL response.

In contrast to total HIV-specific CTL responses, preferential targeting of the Gag protein has been associated with lower viral loads in both clade B (Rolland et al., 2008; Zuniga et al., 2006) and C (Kiepiela et al., 2007; Masemola et al., 2004; Rolland et al., 2008) HIV-1 infections. The immunogenicity of Gag may be because there is an abundance of Gag molecules being produced in the cytosol during viral assembly and consequently, there is a high likelihood that a portion of these proteins will be degraded and shuttled into the MHC class I pathway immediately after infection. Gag-specific CTLs can recognize an infected cell as quickly as two hours post infection, prior to integration or new viral protein synthesis, suggesting that early Gag-specific responses may result in more efficient containment of viral spread (Sacha et al., 2007). Furthermore, compared to other HIV-1 proteins, Gag is relatively conserved due to the

various structural and functional constraints that limit tolerance to mutation (Rihn et al., 2013). Control mediated by Gag-specific responses may be the result of effective targeting against the most vulnerable regions within the Gag protein. However, the identification of which Gag regions are most vulnerable to attack has been unclear. Several studies have suggested that relative dominance of conserved p24-specific CTL responses plays a role in control of viral replication (Dahirel et al., 2011; Streeck et al., 2007; Zuniga et al., 2006). Many of the immunodominant epitopes restricted by HLA alleles associated with immune control are located within the p24 protein and are highly conserved among HIV-1 isolates. Furthermore, escape mutations within p24 epitopes can result in replicative fitness costs, highlighting the vulnerability of this region to immune attack (Miura et al., 2009; Schneidewind et al., 2008; Troyer et al., 2009). Alternatively, Gag-specific CTL responses that have an increased ability to maintain recognition upon epitope variation have also been proposed as a factor contributing to the superiority of Gag-specific targeting (Chen et al., 2012; Gillespie et al., 2002; Kosmrlj et al., 2010; Mothe et al., 2012; Turnbull et al., 2006). With wide epitope variant recognition, these responses have the potential to prevent the outgrowth of escape mutations, thus leading to immune control (Ladell et al., 2013). Theoretically, a highly cross-reactive Gag-specific CTL repertoire extends the vulnerability of the Gag protein beyond conserved regions as long as variation arising in the more variable regions is effectively cross-recognized. However, individuals with progressive disease can also mount Gag-specific CTL responses, including targeting of p24 (Liu et al., 2011; Mothe et al., 2012) and cross-reactive responses (Henn et al., 2012; Hoof et al., 2010; Liu et al., 2011; Malhotra et al., 2009; Zembe et al., 2011). Therefore, there must be additional immunological factors responsible for the *in vivo* effectiveness of some Gag-specific responses.

One consistent factor associated with immune-mediated viral control is CTL polyfunctionality. Comprehensive interrogations into the functionality of Gag-specific CTL

responses require the use of intracellular cytokine staining (ICS), a technique that allows for simultaneous detection of many different immunological parameters. Compared to progressors, HIV-controllers have been shown to consistently maintain more polyfunctional CTL responses, where functionality is usually measured by the production of: perforin, granzyme B, CD107a (a marker for degranulation following release of perforin and granzyme B), IFN- γ , TNF- α , MIP1- β (a cytokine that binds to the CCR5 receptor), and IL-2 (Akinsiku et al., 2011; Betts et al., 2006; Julg et al., 2010; Migueles et al., 2002; Migueles et al., 2008). Enhanced polyfunctionality has been associated with more effective CTL responses. Superior up-regulation of lytic granules (e.g. perforin, granzyme B) was shown to be most effective at the elimination of infected CD4⁺ T cells (Migueles et al., 2008). Additionally, Julg *et al* demonstrated that greater polyfunctionality of Gag-specific CTLs (defined as at least 2 or 3 functions) translated to more effective inhibition of HIV-1 replication, as measured by an *in vitro* viral inhibition assay (VIA) (Julg et al., 2010). Furthermore, proliferative capacity of the responding CTL may also be an essential feature of a protective CTL response. Production of IL-2 as well as actively evading suppression from T-regulatory responses has been associated with greater proliferative capacity (Elahi et al., 2011; Migueles et al., 2002; Zimmerli et al., 2005).

Although polyfunctionality has been consistently associated with responses of HIV-controllers, it is difficult to assess whether this greater functionality is the cause or a consequence of control of viral replication. Constant antigenic stimulation, as mediated by high viral loads, has been shown to induce T-cell exhaustion and subsequent loss of polyfunctionality (Shin et al., 2007). The limited viremia in HIV-controllers may not induce exhaustion and thus functionality is maintained. Furthermore, HIV-specific CTLs in controllers may have greater proliferative capacity due to IL-2 production by CD4⁺ T cells, which are not destroyed to the same extent as seen in progressive HIV infection. However, given the connection between

greater polyfunctionality and superior inhibition of viral replication, the induction of polyfunctional Gag-specific CTLs remains a priority in vaccine development.

HIV-1 genetic diversity

One of the most distinguishing characteristics of the HIV epidemic is the enormous amount of genetic diversity found both within and between individuals. There are several factors inherent in the virus replication lifecycle that contribute to the generation of such incredible diversity. First, during the synthesis of viral DNA, the viral RT is highly error prone due to an absence of exonuclease activity that would otherwise function to remove erroneously incorporated nucleotides (Preston et al., 1988; Renjifo, 2002). This lack of a proofreading mechanism results in an error rate of 10^{-4} per nucleotide that roughly translates into one incorporated error per replicated genome (Renjifo, 2002). Coupled with a relatively short generation time, each day an infected individual produces a vast collection of mutations that can ultimately be selected upon. The likelihood that a mutation will be retained in the viral population is dependent on the effects of that mutation on replicative fitness as well as whether it provides any benefit to the viral population (e.g. evasion of immune responses or resistance to antiretroviral treatment). Furthermore, in cases in which two viruses infect the same cell, recombination between these parental genomes can occur, resulting in further viral diversification (Renjifo, 2002). Due to these mechanisms allowing for sustained levels of high intra-population variance, the viral population within an individual is referred to as a quasispecies.

Typically, a single variant, which is referred to as the founder virus, establishes HIV-1 infection (Abrahams et al., 2009; Herbeck et al., 2011; Keele et al., 2008; Shankarappa et al., 1999). In the early weeks of infection, prior to peak viremia, the viral population is relatively homogenous with little divergence from the founder virus sequence (Herbeck et al., 2011; Keele et al., 2008; Zhu et al., 1993). The minor levels of diversity observed during this time are

believed to be primarily due to stochastic evolution (Herbeck et al., 2011). Approximately one month following transmission, signatures of positive selection become apparent resulting in an increasingly diverse viral population (McMichael et al., 2010). The founder-specific CTL responses are believed to be a major selective force responsible for divergence and diversification during primary infection (Allen et al., 2005a; Herbeck et al., 2011; Liu et al., 2006; Liu et al., 2011). Interestingly, the effects of a transmitting partner's previous CTL response may also contribute to early diversification of the viral population through a process called CTL reversion. In these cases, if the founder virus carries CTL mutations previously selected to confer escape in the transmitting partner, these mutations may revert to the more sensitive form upon transmission to a new host of a different HLA background (Allen et al., 2004; Friedrich et al., 2004; Li et al., 2007).

Throughout the course of HIV-1 infection, Shankarappa *et al* proposed that changes in divergence (mean genetic distance of all sequences from founder virus) and diversity (mean genetic distance between all contemporaneous viral sequences) occur in three phases. In the first phase, both divergence and diversity increase linearly, occurring at an estimated rate of 1% each year in the C2-V5 region of *env* (Shankarappa et al., 1999). These changes in genetic diversity are believed to be, in part, due to the consequence of viral-host interactions, in which the virus constantly has to adapt to escape ongoing immune pressure (McMichael et al., 2010; Rolland, 2013). In the second phase, the rate of diversity begins to shrink while divergence from founder virus continues at the same rate (Shankarappa et al., 1999). During this phase, viruses able to bind the co-receptor CXCR4 (X4 viruses), either exclusively or in addition to CCR5 binding, emerge in the population (Jensen et al., 2003; Shankarappa et al., 1999). This co-receptor shift enables infection of a new array of target cells. Finally, in the third phase, diversity begins to decline and the rate of divergence hits a plateau (Shankarappa et al., 1999). These diversity/divergence characteristics generally occur right before the onset of AIDS. Potentially,

the generalized immune dysfunction may disrupt the immune response, thus removing the selective pressure to keep the viral population diverse (Bunnik et al., 2008; Curlin et al., 2010).

While the level of diversity within an individual is remarkable, the amount of global diversity found across all viral isolates is astonishing. In the early 1990s, it was noticed that viral strains originated from African populations had greater genetic variation than strains isolated in United States and European populations (Renjifo, 2002). With advancements in viral sequencing methodology, it was soon discovered that the global HIV-1 pandemic (group M) contained many different genetically distinct viral clades (referred to as subtypes A-K) (Renjifo, 2002; Robertson et al., 2000). The most accurate way of distinguishing viral clades is to use phylogenetic methods to determine the relatedness between viral sequences (Renjifo, 2002). Virus strains belonging to a particular clade will cluster together on a phylogenetic tree and are distinct from the viruses comprising other clades. Most often, sequenced short fragments of the genome, such as *gag* or *env*, are used in determining clade distinction. Different clade isolates can be over 35% genetically distinct from one other in some HIV proteins (Gaschen et al., 2002).

The North American and European HIV-1 epidemics are caused almost exclusively by subtype B (Renjifo, 2002). In contrast, many different circulating subtypes are found in African countries, the most prevalent of which is subtype C, which is mostly found in sub-Saharan African nations (Renjifo, 2002). In some countries, such as the Democratic Republic of Congo, most of all known subtypes can be found in the infected populations, which is likely a reflection of that fact that this was the region where HIV-1 first emerged to infect human populations (Vidal et al., 2000). In some cases, when different clades are frequently circulating within a population, it allows for dual infections to occur and recombinants to be formed between different clades (Renjifo, 2002; Robertson et al., 2000). In some cases, these recombinants can become prevalent within a population and thus form what is referred to as a circulating recombinant form (CRF). CRFs can become the predominant source of infection within a region. For example, in

Thailand, most of the infections are caused by CRF01_AE (recombinant form between clades A and E); and in West Africa, CRF02_AG (clades A and G) is highly prevalent (Renjifo, 2002). In contrast to inter-clade distinctions, detailed classification of CRFs requires full-length genome sequencing to identify which clades compose the new strain (Renjifo, 2002).

Challenges of HIV-1 diversity on CTL mediated immune control and vaccine development

CTL recognition of an infected cell requires contact between the TCR and the MHC-epitope complex. Disruption of any of these contacts can result in abolishing the CTL response. The high propensity of HIV-1 to develop mutations provides a steady pool of potential escape mutations that could be rapidly selected for. There are three classes of escape mutations that will disrupt CTL recognition (Borrow and Shaw, 1998). First, mutations that occur closely nearby targeted epitopes have the potential to alter peptide processing, and as a result these substitutions prevent the generation of the epitope. As previously mentioned, each HLA allele has a binding motif that permits preferential binding of particular epitopes. Mutations occurring in the p2 or C-term position of the targeted epitope may disrupt HLA binding, thus preventing the presentation of these epitopes on the infected cell surface. Finally, as the TCR recognizes the epitope as a whole, changes occurring within the epitope may abolish T-cell recognition despite maintaining binding to the HLA molecule.

While founder-specific CTL responses develop a month into infection, CTL escape mutations are often selected shortly thereafter, rendering the original response ineffective (McMichael et al., 2010). As these escape processes repeat throughout the course of infection, the vast majority of infected individuals are unable to control viral replication through CTL responses (Ammaranond et al., 2011; Feeney et al., 2004; Goulder et al., 1997). In controllers, escape mutations can have both positive and negative effects on disease progression, as illustrated by CTL escape processes in controllers with HLA alleles associated with immune

control. B*57 HIV-controllers frequently target an epitope within p24 (TW10) early during infection (Altfeld et al., 2003a). The most common CTL escape pathway within this immunodominant epitope involves a T242N mutation at the third position of the epitope (Brockman et al., 2007). As this mutation disrupts interactions with host protein cypA during viral uncoating, this mutation carries a fitness cost (Brockman et al., 2007; Martinez-Picado et al., 2006). Although this fitness cost may be compensated by additional mutations 25-30 amino acids upstream of the epitope (Brockman et al., 2007), this disabling of the virus early in infection is believed to contribute to immune control achieved by these individuals. In contrast, B*27 controllers have been shown to lose control of viral replication once escape mutations within the immunodominant KK10 epitope and compensatory mutations develop (Goulder et al., 1997). As this network of mutations is difficult to arise given that each single mutation has high fitness costs, it often takes years until escape is achieved, resulting in late disease progression in these individuals (Schneidewind et al., 2008).

Global HIV-1 diversity has been a significant challenge in the development of a universal prophylactic vaccine. Individuals have to be re-vaccinated each year in order to contend with the diversity found within the influenza virus, which diverges at a rate of 1-2% per year (Korber et al., 2001). Considering that HIV-1 can differ by over 35% in some proteins between isolates (Gaschen et al., 2002), it is daunting to think of the vaccination strategies that will be necessary to overcome the challenges of HIV sequence diversity. Induction of HIV-specific T cells via vaccination has been an important goal in the design of HIV immunogens. Including a CTL component in an HIV vaccine is supported by the wealth of literature supporting a role for CTLs in immune control during natural infection as well as several NHP studies that demonstrate vaccine induced SIV-specific CTL responses to be protective from disease progression (Barouch et al., 2012; Hansen et al., 2011; Kaslow et al., 1996; Kiepiela et al., 2007; Koup et al., 1994; Mudd et al., 2012; Patel et al., 2013; Rolland et al., 2008; Schmitz et al.,

1999; Wilson et al., 2009). However, in human clinical trials, all T-cell based vaccines test to date have shown no positive clinical efficacy (Buchbinder et al., 2008; Gray et al., 2011; Health, 2013). The best studied of these is the Step Study, a phase IIb clinical trial that was shown to be ineffective at both prevention of HIV-1 infection and reducing viral load post-infection (Buchbinder et al., 2008). The immunogen used in the Step Study included a single viral sequence each of HIV-1 proteins Gag, Pol, and Nef (McElrath et al., 2008). Remarkably, despite the absence of positive efficacy, sequencing of HIV-1 strains in trial participants revealed that the vaccine-induced CTLs did have an impact on breakthrough viral populations in that the viruses infecting vaccinees were of a greater genetic distance to the sequence used in the immunogen than the viruses infecting placebos (Rolland et al., 2011). Therefore, one of the explanations for why this vaccine failed is that the limited breadth of responses provided insufficient coverage of infecting strains. Since the results of the Step Study, there has been considerable work in the field in developing concepts for T-cell based vaccines that are better equipped to handle the challenges posed by the extensive diversity found in HIV populations.

Current vaccines designed to contend with HIV-1 diversity

As the Step Study suggested, the choice of immunogen may be critical in inducing responses that can contend with sequence variation. One strategy to improve HIV immunogens is design the antigen to be reflective of a centralized sequence that is equally genetically related to all circulating sequences (Korber et al., 2009). There are several different ways to calculate a centralized sequence including consensus, ancestral, and center-of-tree (COT) methodology (Korber et al., 2009). Importantly, these sequences can be generated for specific subtypes or for all group M sequences, depending on the vaccination strategy goals. To build upon the sequence coverage of potentially infecting strains achieved through centralized sequences, several groups generated polyvalent vaccines that were designed to incorporate additional variation found in circulating sequences (Fischer et al., 2007; Nickle et al., 2007). In mosaic

immunogens, sequences are computationally optimized to maximize variation included in potential T-cell epitopes while still mimicking an intact protein (Fischer et al., 2007). Alternatively, computational strategy in the design of COT+ immunogens involves building the variation up from COT centralized sequence by maximizing variation within the number of 9-mer (9 amino acids long) potential epitopes (Nickle et al., 2007). In both mosaics and COT+, the immunogen design remarkably increases the coverage of sequence diversity compared to consensus or COT immunogens alone. Both concepts hope to induce CTL responses with high breadth (number of HIV-specific responses) and wide depth (the number of epitope variants targeted). Increasing the valence (i.e. number of components) of the immunogens increases the level of coverage that can be achieved. In theory, these responses may be capable of blocking infection of a broad range of viral variants. Furthermore, if breakthrough infections occur, these responses of wide epitope variant recognition may be able to block common escape pathways and thus select for rare variants that may carry high fitness costs, allowing for post-acquisition control of viral replication.

In contrast to incorporating variation into immunogen design, an alternative strategy to overcome the challenges of sequence diversity is to focus induced responses only on highly conserved regions of the virus (Letourneau et al., 2007; Rolland et al., 2007). Concepts of conserved vaccines are rooted in the principles that (a) targeting conserved regions will increase the sequence coverage of potentially infecting strains, (b) that avoiding responses to variable regions (i.e. regions prone to escape) will improve the quality of the adaptive immune response, and (c) escape mutations occurring within these regions will severely comprise the viability of the virus. In vaccines designed to induce CTL responses, conserved regions are selected based on known immunogenicity and therefore, these regions contain previously described CTL epitopes (Letourneau et al., 2007; Rolland et al., 2007). However, even conserved regions are not entirely invariant in HIV-1 populations. To address this, one vaccine

strategy, the conserved elements (CE) vaccine, incorporates polymorphisms at the few sites that can tolerate variation within the conserved regions, thus constructing a polyvalent immunogen (Kulkarni et al., 2013; Niu et al., 2011; Rolland et al., 2007). As a result, the vaccine will ideally induce cross-reactive responses against the limited toggle sites within the conserved elements and preserve targeting. Using COT centralized sequence and incorporating toggle sites, the segments included in the CE vaccine are reflective of over 99% of group M sequences, and thus provides immense coverage of global HIV-1 strains (Rolland et al., 2007).

Thesis aims

HIV-1 diversity is a significant challenge in the development of an efficacious vaccine. Inducing CTL responses of wide epitope variant recognition or targeting primarily conserved regions have each been proposed as a tactic to overcome this challenge. Interestingly, despite vast differences in the concepts of mosaic and CE immunogen design, the success of both vaccination strategies is dependent on the ability for HIV-specific CTLs to recognize epitope variants (Fischer et al., 2007; Rolland et al., 2007). With human phase I trials currently being planned to measure the immunogenicity of both of these vaccine strategies, it is imperative to further investigate the capacity of HIV-specific T cells to cope with epitope variation. While the ability to recognize epitope variants has previously been reported during both acute and chronic HIV-1 infection (Gillespie et al., 2002; Henn et al., 2012; Liu et al., 2011; Malhotra et al., 2009; Mothe et al., 2012; Turnbull et al., 2006; Zembe et al., 2011), these studies were limited to investigating a small number of epitopes and testing recognition of a narrow number of epitope variants. As a result, the consequences of epitope variant recognition during natural HIV-1 infection remain largely unknown. Here I describe results that contribute to a greater understanding of the role of epitope variant recognition in control of viral replication as well as the selective pressure exerted by early cross-reactive CTL responses. Two areas of research are described:

1. The role of epitope variant recognition in immune control of HIV-1.
2. Investigation of CTL escape processes following acute HIV-1 infection and characterization of the impact of cross-reactive CTL responses on selection of escape mutations.

These data provide critical insight into the complexity of epitope variant recognition during HIV infection and thus will contribute significantly to the future evaluation of both mosaic and CE vaccine responses.

Chapter II:

**Increased sequence coverage through
combined targeting of variant and
conserved epitopes correlates with control
of HIV replication**

Abstract:

A major challenge in the development of an effective HIV vaccine is contending with the extensive sequence variability found in circulating virus strains. Induction of HIV-specific T-cell responses targeting conserved regions or recognizing a high number of epitope variants have each been proposed as a strategy to overcome this challenge. We addressed the ability of cytotoxic T-lymphocytes from untreated HIV-infected subjects with and without control of virus replication to recognize epitope variants by assessing responses by IFN γ /IL-2 FluoroSpot to all circulating clade B Gag sequence variants encoded by at least 5% of the sequences in the Los Alamos National Laboratory HIV database (1300 peptides). While targeting of conserved regions was similar in both groups ($p=0.47$), we found that subjects with spontaneous control of virus replication demonstrate marginally lower recognition of Gag epitope variants compared to subjects with normal progression ($p=0.05$). In viremic controllers and progressors, we found variant recognition to be associated with viral load ($r=0.62$, $p=0.001$). Interestingly, we show that increased overall sequence coverage, a metric combining variant recognition and targeting of conserved epitopes that reflects the proportion of HIV database sequences targeted, is inversely associated with viral load ($r=-0.38$, $p=0.03$). Furthermore, we found that sequence coverage, but not epitope variant recognition, correlated with increased recognition of a panel of clade B HIV founder viruses ($r=0.50$, $p=0.004$). Our results suggest that the ability to target the most commonly occurring Gag variants is associated with control of viral replication. We therefore propose sequence coverage by HIV Gag-specific immune responses as a possible correlate of protection that may contribute to control of virus replication. Additionally, sequence coverage serves as a valuable measure in which to evaluate the protective potential of future vaccination strategies.

Introduction

Observational and correlative studies strongly suggest a role for virus-specific cytotoxic T lymphocytes (CTLs) in the control of HIV and SIV infection (Hansen et al., 2011; Kaslow et al., 1996; Kiepiela et al., 2007; Koup et al., 1994; Mudd et al., 2012; Rolland et al., 2008; Schmitz et al., 1999). Translating these observations into the development of a successful HIV vaccine has been incredibly challenging, in part due to the immense sequence diversity of HIV that must be covered by the induced immunity, and to the difficulty in defining accurate correlates of protective immunity from disease progression. Several lines of evidence have implicated a superior role for Gag-specific CTL responses in control of viral replication (Julg et al., 2010; Kiepiela et al., 2007; Rolland et al., 2008). However, as individuals with progressive disease can also mount Gag-specific CTL responses, factors other than simple recognition must be involved that contribute to the *in vivo* effectiveness of some anti-viral Gag-specific CTL responses.

One factor that has been proposed as an immune correlate of HIV control is an enhanced ability of Gag-specific CTLs to cross-react with epitope variants (Gillespie et al., 2002; Kosmrlj et al., 2010; Mothe et al., 2012; Turnbull et al., 2006). In the majority of individuals, HIV infection leads to the induction of a great number of CTLs that are able to recognize and kill infected cells, but the extraordinary tolerance for sequence variability in the viral population can rapidly lead to the accumulation of escape mutations (Allen et al., 2005a; Herbeck et al., 2011; Liu et al., 2006). In many cases, this immune adaptation renders the CTL responses ineffective and can lead to increasing viremia during the course of infection (Ammaranond et al., 2011; Feeney et al., 2004; Goulder et al., 1997). The induction and maintenance of a broadly cross-reactive CTL repertoire may facilitate control of viral replication by suppressing the outgrowth of escape variants. Furthermore, broad epitope variant recognition has the potential to be beneficial in a vaccine setting, as recognition of epitope variants may afford targeting of a

greater proportion of potentially infecting strains. Indeed, this rationale has already been translated into the design of innovative immunogens designed to incorporate as much sequence variability as feasible (Fischer et al., 2007; Nickle et al., 2007). One approach currently moving into human phase I clinical trials include polyvalent mosaic vaccines, which have been shown to improve epitope variant recognition in non-human primate (NHP) models compared to consensus or natural immunogens (Barouch et al., 2010; Santra et al., 2010; Santra et al., 2012). At the same time, the opposite approach focusing variant responses on only highly conserved regions of HIV to provide recognition of more infecting strains is also being moved into phase I clinical trials (Kulkarni et al., 2013; Letourneau et al., 2007; Mothe et al., 2012; Niu et al., 2011; Rolland et al., 2007).

It is important to consider that epitope variant recognition has been widely reported in the context of progressive infection (Hoof et al., 2010; Liu et al., 2011; Malhotra et al., 2009; Zembe et al., 2011). To date, variant recognition assessments have often focused on a limited number of epitopes and/or epitope variants, largely due to limited sample availability and the prohibitive cost of generating proteome-wide variant peptide sets. Most often, the effects of intra-clade variation on T-cell recognition has been assessed in the setting of immune escape, and therefore the evaluation of variant-specific responses are limited to only those variants that arose in the autologous virus over time (Altfeld et al., 2003b; Henn et al., 2012; Liu et al., 2006; Liu et al., 2011). While greatly informative in our understanding of the dynamics of host-viral interactions, these types of approaches inevitably only incrementally extend understanding of the potential for HIV-specific CTLs to recognize naturally occurring epitope variants. As a result, the relative contribution that broad epitope variant recognition may play in the control of HIV replication remains unclear.

Here we report a comprehensive evaluation of the cross-reactivity of Gag-specific T-cell responses to frequently found clade B variants in the Los Alamos National Laboratory HIV

Sequence Database (HIVDB; <http://www.hiv.lanl.gov/>) in individuals with and without spontaneous control of viral replication. Contrary to our expectations, we observed extensive epitope variant recognition in progressors, which was associated with viral load. However, we found an inverse correlation between viral load and sequence coverage of frequently occurring variants, suggesting that it is the ability to target the most commonly occurring variants, rather than simply a large number of variants, that contribute to control of viral replication. These findings provide a greater understanding of epitope variant recognition during natural infection and offer important insight for informing the evaluation of variant-inclusive vaccines.

Materials and Methods

Study Subjects

Thirty HIV-infected participants in the Seattle Long-Term Non-Progressor (Horton et al., 2006), Natural Progression, Primary Infection (Schacker et al., 1996; Schacker et al., 1998; Stekler et al., 2008), and NIAID/NIH (Migueles et al., 2000) cohorts were selected for this study. HIV controllers (n=15) were defined by viral loads <2,000 RNA copies/ml and CD4⁺ T cell counts >500. Elite Controllers (EC) were further defined as having undetectable viral load (<50 RNA copies/ml), whereas viremic controllers (VC) had detectable viral loads under <2,000 RNA copies/ml. Progressors (n=15) were defined as having median viral loads >10,000 RNA copies/ml in the last year. Three of the progressors were identified from the Seattle Primary Infection cohort and have previously assigned publication identification numbers (P11=55097, P12=59530, P15=75688). All subjects were studied in chronic infection and were antiretroviral therapy naïve. The relevant institutional review boards approved all human subject protocols, and all subjects provided written informed consent before enrollment.

Peptide Set

A Gag peptide set of 11mer peptides overlapping by 10 amino acids (AA) was synthesized for use in this study (Sigma-Aldrich, St. Louis, MO). All 11mers were generated to reflect any Gag variants present in at least 5% of clade B sequences (n=1621) in the HIVDB in 2010, resulting in a total of 1300 peptides covering all 500 AA positions in Gag. Peptides were pooled based on each peptide's starting position in the alignment, resulting in a total of 489 pools.

IFN- γ /IL-2 FluoroSpot

IFN- γ /IL-2 FluoroSpot assay was used for detection of Gag-specific T-cell responses. Cryopreserved peripheral blood mononuclear cells (PBMC) were thawed and incubated in R10 media (RPMI 1640 [GibcoBRL, Carlsbad, CA], 10% FBS [Gemini Bioproducts, West Sacramento, CA], 2 mM L-glutamine [Gibco], 100 μ g/ml streptomycin sulfate [Gibco], 100 U/ml penicillin G [Gibco]) overnight before stimulation. Between 70,000-100,000 PBMC/well were plated in 96-well IPFL plate (Millipore, Bedford, MA), which had been pre-coated overnight with 1 μ g/ml anti-IFN- γ monoclonal antibody 1-D1K (Mabtech, Stockholm, Sweden) and 1 μ g/ml anti-IL-2 antibody 2516KZ (BD Biosciences, New Jersey, USA). Peptide pools or individual 11mers were added at a final concentration of 10 μ g/ml. Each plate contained cells stimulated with 1.8 μ g/ml PHA as a positive control and six negative control wells of cells incubated with media alone. Plates were incubated overnight at 37°C in 5% CO₂. Plates were developed using anti-IFN- γ 7-B6-1-FS-FITC antibody diluted 1:400 (Mabtech) and 0.1 μ g/ml anti-IL-2 antibody 2311KZ (BD) for 2h, followed by anti-FITC-green diluted 1:400 (Mabtech) and streptavidin-coupled Cy3 diluted 1:1000 (Biolegend, San Diego, CA) for another hour. Plates were washed six times with PBS between all steps. Spots were counted using AID iSpot FluoroSpot Reader System.

All study subjects were screened for responses to the 489 peptide pools. For each positive pool, all contained 11mers (i.e., the different variants of the 11mer sequence) were tested individually in a confirmatory IFN- γ /IL-2 FluoroSpot. Positivity for both pool and single variant responses was defined as i) >55 Spot Forming Cells (SFC)/million PBMC, ii) >4 times the average of at least six negative control wells from the same experiment, iii) >3 standard deviations above average negative control wells, and iv) at least 5 spots per well.

Gag Sequencing

10-25 *gag* sequences were obtained from plasma RNA from five VC and five progressors. For 8/10 study subjects, sequencing was performed at the same visit date used in the immunological assays. Because of plasma availability, PCR products from VCs C12 and C15 were sequenced from plasma at the next available visit date (2 and 3 months later, respectively).

HIV-1 RNA was purified from plasma using the QIAamp Viral RNA Mini Kit (Qiagen, Valencia, CA). HIV-1-specific PCR was performed to amplify HIV-1 *gag* products with use of end point-limiting dilution to obtain single templates. The HIV-1 *gag* gene (1590 nt fragment; the complete coding region of Gag and 103 nt of PR) was amplified by nested-PCR with the following primer sets. For the first round PCR, nt 683-705 of HIV-1_{HXB2} (F683, 5'-CTC TCG ACG CAG GAC TCG GCT TG-3'); and nt 2484-2511 of HIV-1_{HXB2} (StepGR_1.0_2511, 5'-TTC CAA TTA TGT TGA CAG GTG TAG GTC C-3') were used. For the second round, nt 762-786 of HIV-1_{HXB2} (F762, 5'-TTG ACT AGC GGA GGC TAG AAG GAG A-3'); and nt 2377-2403 of HIV-1_{HXB2} (StepGR_2.0_2403, 5'-CAA TTC CCC CTA TCA TTT TTG GTT TCC-3') were used. The first round of PCR used the Advantage 2 polymerase (Clontech, Mountain View, CA) and the following conditions: 95°C for 2 min, 35 cycles of 95°C for 30s, 58°C for 30s, 68°C for 2 mins and a final extension step at 68°C for 10 min. The second round of PCR used Biolase Taq

(Bioline, Taunton, MA) and the following conditions: 95°C for 2 min, 35 cycles of 95°C for 30s, 62°C for 30s, 68°C for 2 min and a final extension step at 68°C for 10 min.

PCR products were visualized using the QIAxcel DNA Fast Analysis Kit (Qiagen) and PCR positive samples were purified using NucleoSpin Gel and PCR Clean-Up Kit (Clontech) before being submitted directly for dideoxynucleotide chain termination (Sanger) sequencing. Chromatograms were assembled and edited using Geneious 5.6.5 (Biomatters, NZ).

Data Analysis

The breadth of Gag-specific responses was calculated conservatively based on epitopic regions, which means that responses for up to 4 subsequent 11mer peptides (i.e., potential shared 8-mer) were counted as one epitopic region. Dual functional responses were determined by an IFN- γ plus IL-2 response to at least one 11mer peptide within a targeted epitopic region. Variant recognition was calculated as the proportion of variants eliciting a response among all the 11mer variants tested in the targeted epitopic region (**Table 2-1**). Responses in which only a single sequence was tested (i.e., highly conserved, with no other variant present in $\geq 5\%$ of sequences) were excluded from the variant recognition analysis as no variants were tested. Immune coverage of frequently occurring Gag variants was calculated per epitopic region as the sum [frequency of targeted 11mers] over the sum [frequency of tested 11mers], where frequency refers to the proportion of the 11mer sequence in the clade B sequence alignment used to design the peptide set (**Table 2-1**). We chose to normalize immune coverage for each epitopic region by sum [frequency of tested 11mers] to account for the fact many epitopic regions had overlapping 11mers. Without normalization, sum [frequency of targeted 11mers] would often be over 100%. The median sum [frequency of tested 11mers] across all targeted epitopic regions was 81%, highlighting that although our peptide set does not cover all possible variation in clade B Gag sequences, the peptide design ensured that we were able to test for responses to a high proportion of circulating Gag variants.

Table 2-1: Example of Data Analysis						
Epitope ID	IFN- γ SFC/M	IL-2 SFC/M	DUAL SFC/M	11mer peptide		11mer Frequency
145a	770	---	---	QAISPRTLNAW		53.10%
145b	---	---	---	QPISPRTLNAW		17.49%
145c	1050	---	---	QALSPRTLNAW		14.02%
145d	---	---	---	QPLSPRTLNAW		5.96%
		# Variants Recognized		2	Sum Frequency of Recognized Variants	67.12%
		# Variants Tested		4	Sum Frequency of Tested Variants	90.57%
		% Variant Recognition		50%	% Coverage of frequently occurring variants	74.16%
SFC/M=Spot forming Cells/Million PBMC						

Autologous *gag* sequences (whole protein and viral portions corresponding to targeted epitopic region) were aligned using the Muscle algorithm within Seaview v4.4.0 (Edgar, 2004). Le and Gascuel (LG) substitution model (Le et al., 2012) corrected pairwise diversity and divergence estimates from clade B consensus were calculated using DIVEIN (Deng et al., 2010). Analyses were restricted to only those epitopic regions in which variant recognition was assessed. To generate the panel of founder viruses, we obtained previously generated sequences isolated from the first timepoint in early acute infection (median 7 days post onset of symptoms) from eight individuals (Herbeck et al., 2011). For each individual, a consensus sequence was generated from the first time point sequences using Seaview v4.4.0 and was used to represent the founder virus.

Statistical Analysis

For each study subject, the median variant recognition or median epitope coverage was calculated across all epitopes and used in comparisons between groups using a non-parametric Mann-Whitney test. Differences between pairwise distances between groups were assessed using a non-parametric Mann-Whitney test. Spearman rank was used to assess all correlations. GraphPad Prism X was used for all statistical analyses.

Results

Significant differences in the quality but not the quantity of Gag-specific responses between Controllers and Progressors

To assess whether broad recognition of Gag epitope variants was associated with control of viral replication, we screened for Gag-specific T-cell responses in 15 HIV controllers and 15 HIV progressors. The controller cohort was comprised of seven Elite Controllers (EC, viral load undetectable with a cutoff of 50 copies/ml) and eight Viremic Controllers (VC, viral load <2000 copies/ml) (**Table 2-2**). Approximately half of the controllers possessed HLA class I alleles previously described to be associated with control of HIV replication (B*57 and/or B*27), however, these alleles did not segregate with EC over VC (**Table 2-2**). To reduce HLA bias, B*57-expressing progressors were included in this study (**Table 2-2**). All study subjects were antiretroviral therapy naïve.

Table 2-2: Subject Characteristics				
Study Subject	Status^a	CD4 count^b (cells/μl)	Viral Load^c (RNA copies/ml)	Protective Allele
Controller 1	EC	809	<50	B*27/B*57
Controller 2	EC	1211	<50	
Controller 3	EC	540	<50	B*57
Controller 4	EC	603	<50	B*27
Controller 5	EC	643	<50	
Controller 6	EC	1389	<50	B*57
Controller 7	EC	1916	<50	
Controller 8	VC	523	246	B*57
Controller 9	VC	784	125	B*57
Controller 10	VC	580	432	B*27/B*57
Controller 11	VC	747	250	
Controller 12	VC	570	386	B*57
Controller 13	VC	702	163	
Controller 14	VC	790	191	B*57
Controller 15	VC	1181	430	

Controller Median		747	125	
Progressor 1	TP	349	22,730	
Progressor 2	TP	550	286,000	
Progressor 3	DP	441	78,400	B*57
Progressor 4	TP	522	39,550	
Progressor 5	TP	510	60,000	
Progressor 6	TP	175	115,240	
Progressor 7	TP	440	622,800	
Progressor 8	TP	371	217,200	
Progressor 9	TP	407	3,455	
Progressor 10	TP	340	23,750	
Progressor 11	TP	317	8,606	
Progressor 12	TP	448	153,286	
Progressor 13	TP	575	24,224	B*57
Progressor 14	TP	416	23,279	B*57
Progressor 15	TP	500	9,333	
Progressor Median		440	39,550	

^aVC= Viremic Controller, EC= Elite Controller, TP= Typical Progressor, DP= Delayed Progressor. DP refers to an individual who previously was a controller and later progressed. At the time of study, this individual had VL>10,000 copies/ml over the prior year.

^bCD4 counts within 4 weeks of sample collection

^cViral loads on the day of sample collection for assays (median viral load over last year >10,000 for progressors).

Study subjects were screened for cytokine-secreting T cells by IFN- γ /IL-2 FluoroSpot using a peptide set of 11mers overlapping by 10 amino acids (AA) spanning all 500 AA of Gag. The FluoroSpot assay uses unique fluorophore-conjugated antibodies to allow for simultaneous detection within the same well of IFN- γ , IL-2 and dual IFN- γ /IL-2 responses following peptide stimulation. Given the tight overlap of the peptide set, we took a conservative approach to calculating breadth in that responses of up to four sequential 11mers were counted as a response to a single epitopic region. We detected responses to a total of 347 epitopic regions. The breadth of Gag-specific responses was comparable between controllers and progressors, with controllers targeting a median of 12 epitopic regions and progressors targeting a median of 9 ($p=0.24$, **Figure 2-1A**). The peptide set used was very sensitive for the detection of HIV-specific T-cell responses, detecting more than double the number of responses previously described using other peptide sets (Frahm et al., 2004; Kiepiela et al., 2007). This increase in

sensitivity was likely due to the combination of using shorter peptides (11 rather than 15 AA) that more closely resemble the optimal HLA class I epitope and increased variant coverage relative to other peptide sets (Frahm et al., 2008; Mothe et al., 2012).

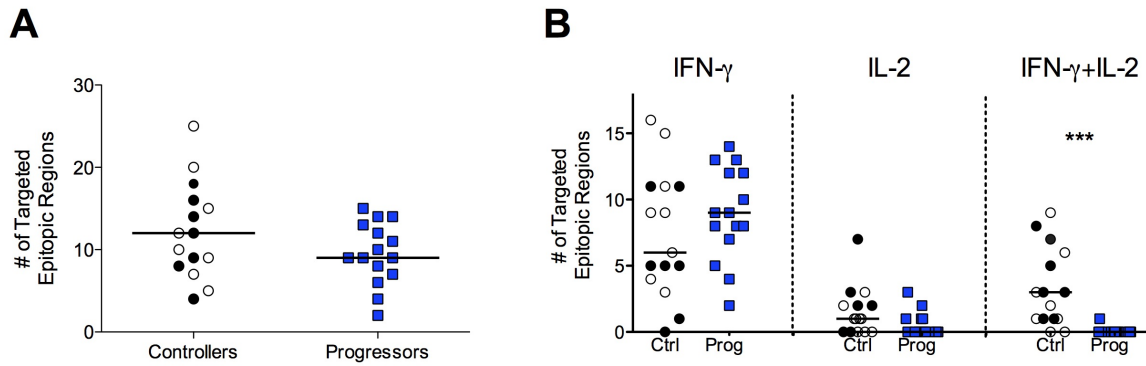


Figure 2-1: Breadth and functionality of Gag-specific T-cell responses in 15 controllers and 15 progressors. Controllers include seven Elite Controllers (viral load <50 RNA copies/ml, black circles) and eight viremic controllers (viral load detectable but <2,000 RNA copies/ml, open circles). Progressors had median viral loads >10,000 RNA copies/ml (blue squares). A) Breadth was defined as the number of targeted epitopic regions, with responses to up to 4 consecutive 11mer peptides counted as a single epitopic region. Breadth for calculated for all study subjects and compared between groups ($p=0.24$). B) Breadth of Gag-specific responses stratified by functionality and compared between groups: monofunctional IFN- γ (left, $p=0.35$) or IL-2 (middle, $p=0.06$) and dual functional IFN- γ /IL-2 responses (right column, ***: $p<0.0001$). Responses were considered dual functional if any peptide in the epitopic region elicited a dual response.

There was no difference in overall magnitude of responses between the two groups (data not shown). However, we observed a significant qualitative difference between the two groups in functionality of the detected responses: while progressors made predominantly single IFN- γ responses ($p=0.35$ compared to controllers), controllers mounted a marginally greater number of IL-2 responses ($p=0.06$) and a significantly higher amount of dual IFN- γ /IL-2 responses ($p<0.0001$, **Figure 2-1B**). CTL epitopes presented by HLA class I are typically 8-12 AA long and therefore our peptide set will preferentially detect CD8⁺ T-cell responses; nonetheless, as the FluoroSpot uses bulk PBMC to detect responses, we cannot rule out the possibility we were also detecting CD4⁺ T-cell responses.

Controllers and Progressors target conserved and variable regions in Gag

The overlapping design of our peptide set coupled with the extensive sequence representation of frequently occurring Gag variants allowed us to comprehensively detect Gag-specific responses without using autologous peptide sets. Overall, we found the targeting of Gag by T cells to be remarkably similar between controllers and progressors (**Figure 2-2**). We found no evidence for differential targeting of variable regions and no significant differences overall in the number of responses towards conserved epitopic regions (**Figure 2-2**, **Supplementary Figure 2-1**). The only significant difference we observed between groups was the preferential targeting of residues #362-367 (₃₆₂VLA₃₆₇EAM₃₆₇) in controllers ($p < 0.05$ for all residues, Fisher's exact test, **Figure 2-2**). These conserved residues lie at the proteolytic cleavage site between p24 and p2, and have recently been reported to be part of a highly constrained co-evolving group of Gag residues with a low tolerance for mutation (Dahirel et al., 2011). We did not find evidence for an HLA bias in preferential targeting; however, study of larger cohorts would be needed to define a connection between these residues and a potential contribution towards immune control. Given that half of our controllers possessed B*57 and/or B*27 alleles, we did observe more responders among controllers to residues #162-172 ($p = 0.46$) and #263-272 ($p = 0.65$), corresponding to the immunodominant B*57 KF11 and B*27 KK10 epitopes, respectively. However, these differences were not significant.

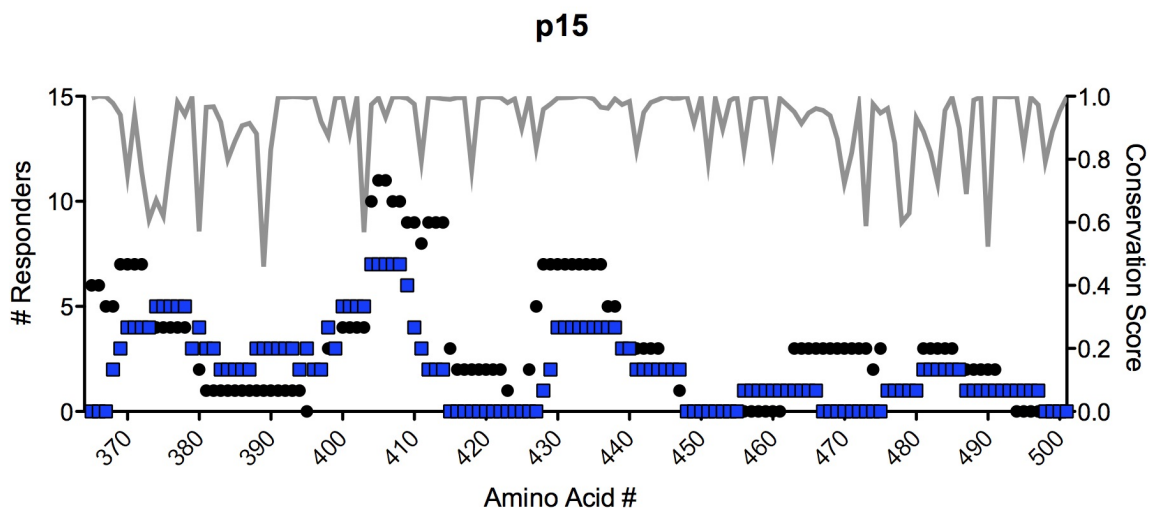
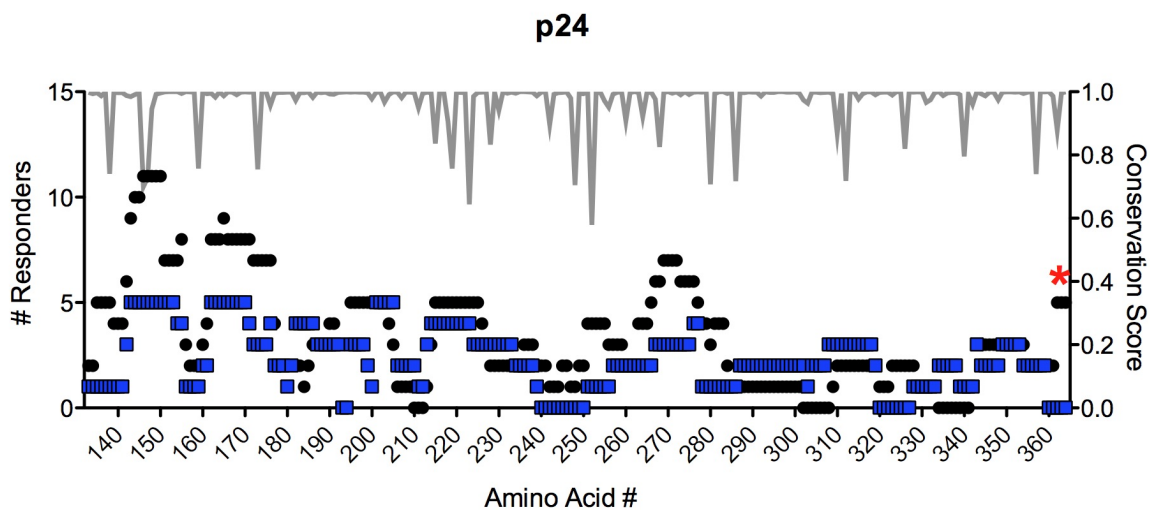
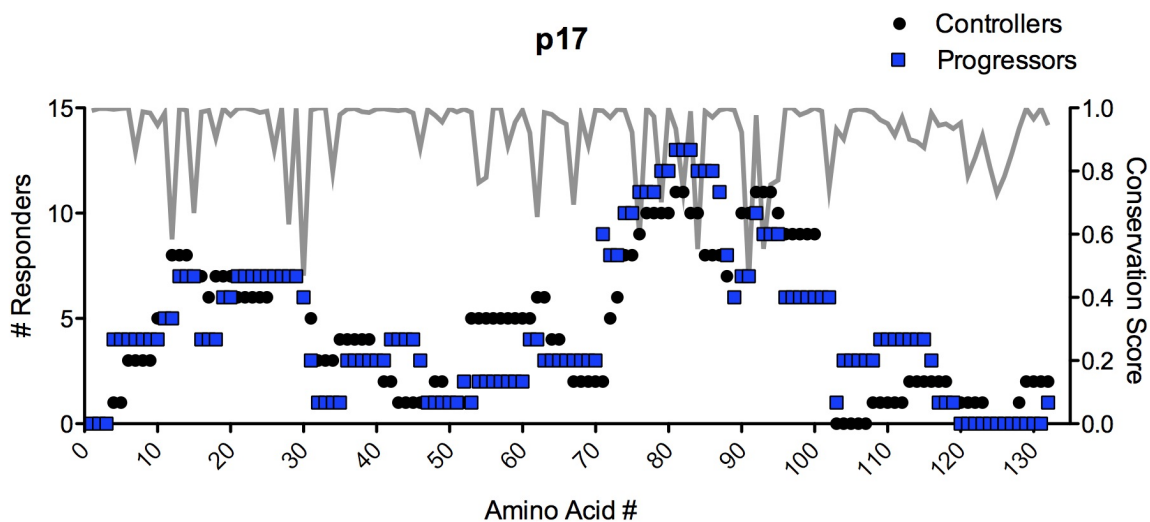


Figure 2-2: Controllers and progressors target conserved and variable regions of Gag.

Shown is the number of controllers (black circles) and progressors (blue squares) that target each amino acid residue in Gag p17 (top), p24 (middle) and p15 (bottom panel), as well as the conservation score for each residue (grey line). *: $p < 0.05$.

Higher epitope variant recognition observed in Progressors

As breadth and the targeting of Gag residues was comparable, we next determined if there were differences in the number of distinct sequence variances recognized (“variant recognition”) by Gag-specific T-cell responses between the controller and progressor groups. The peptide set used in this study was designed to include any variant present in at least 5% of clade B Gag sequences from the HIVDB, allowing us to comprehensively identify responses to frequently occurring Gag variants. For each targeted epitope, we calculated variant recognition as the number of 11mers recognized out of the number of 11mers tested per epitopic region (**Table 2-1**). Both within and between individuals, the level of variant recognition per epitopic region varied considerably (**Supplementary Figure 2-2A and B**). Previous reports have suggested that B*57 and B*27 restricted T cells exhibit enhanced cross-reactivity capability (Chen et al., 2012; Gillespie et al., 2002; Kosmrlj et al., 2010). However, we did not detect a significant difference in overall variant recognition between individuals with ($n=10$) or without ($n=20$) B*57 and/or B*27 alleles ($p=0.61$, **Supplementary Figure 2-3**). Interestingly, functionality of the responding T cells varied depending on the epitope variant. In both controllers and progressors, the majority of responses to epitope variants were detected primarily by IFN- γ secretion. Only in a few targeted epitopic regions did all recognized variants elicit an IFN- γ plus IL-2 response (e.g., epitopes 74 and 262 in C5, **Supplementary Figure 2-4**). We observed marginally higher median epitope variant recognition in progressors ($p=0.05$, **Figure 2-3**).

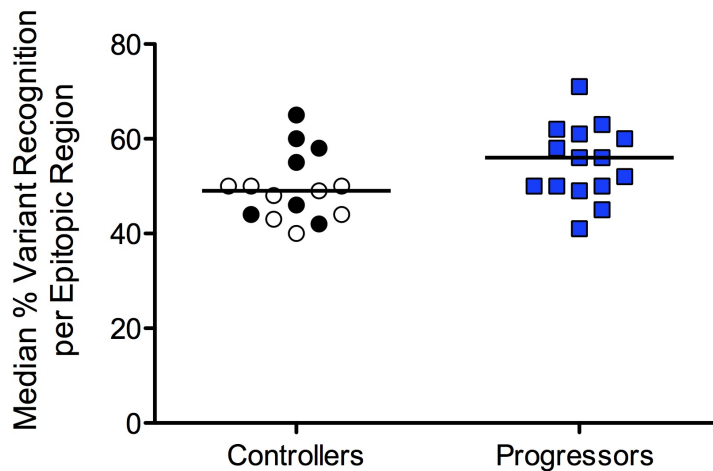


Figure 2-3: Higher variant recognition in progressors.

Epitope variant recognition is measured as the number of 11mers recognized per number of 11mers tested in a targeted epitopic region. For each study subject, the median percent variant recognition was calculated across all targeted epitopic regions and compared between controllers and progressors ($p=0.05$). EC (black circles), VC (open circles) and progressors (blue squares).

Epitope variant recognition is associated with viral load in Viremic Controllers and Progressors

The extensive epitope variant recognition observed in progressors may be a consequence of persistently higher viral loads, which may sustain greater intra-population variance. When we examined all study subjects, we observed a trend towards a correlation between epitope variant recognition and viral load ($r=0.33$, $p=0.07$; **Figure 2-4A**). Exclusion of individuals with undetectable viral load from the analysis revealed a significant correlation between epitope variant recognition and viral load ($r=0.62$, $p=0.001$; **Figure 2-4B**).

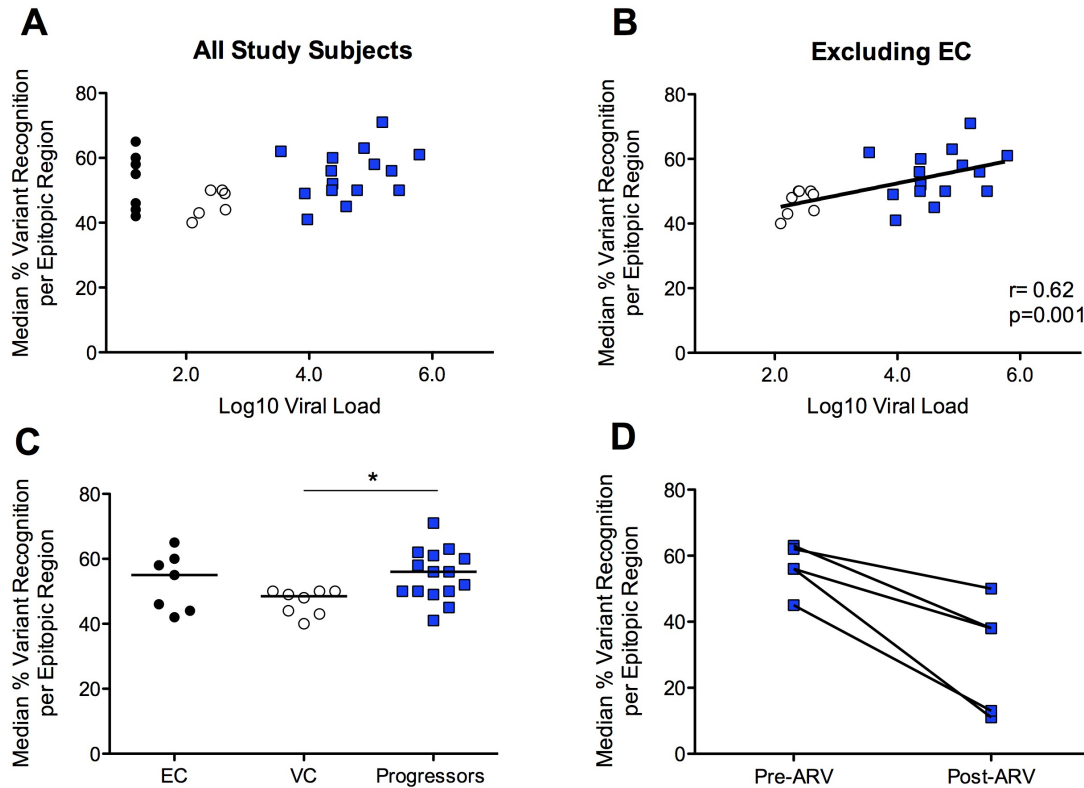


Figure 2-4: Viral load and variant recognition in viremic subjects. Median variant recognition per subject for EC (black circles), VC (open circles) and progressors (blue squares) is shown in relation to $\log_{10}$ viral load for A) all subjects or B) excluding EC. C) Variant recognition stratified by elite and viremic control compared to progressors (KW $p=0.047$, *: Dunn's post test $p<0.05$ for VC vs Progressor). D) Variant recognition in five progressors before (Pre-ARV) and after 6 months of treatment (Post-ARV, $p=0.06$).

Interestingly, when we segregated our controller cohort into elite controllers and viremic controllers, we found that progressors recognized significantly more variants only when compared to viremic controllers (KW $p=0.047$, Dunn's post test $p<0.05$ for VC vs progressors, **Figure 2-4C**). To further explore the relationship between viral load and variant recognition, we repeated the variant recognition assessment in five progressors who started antiretroviral therapy during the study period. Following six months of treatment, we observed a trend towards reduced epitope variant recognition in all five progressors ($p=0.06$, **Figure 2-4D**). Collectively, these data suggest that while viral load can explain the differential variant recognition observed between viremic controllers and progressors, there are likely alternative

mechanisms responsible for driving and maintaining the high levels of variant recognition observed in elite controllers.

Greater diversity in Gag epitopes found in Progressors

Persistence of viral antigen has been shown to promote proliferation and maintain antigen-specific CTL populations during chronic viral infection (Shin et al., 2007). Therefore, the high levels of variant-specific responses observed in progressive infection may be maintained due to the presence of highly diverse Gag variants in the viral population. To address the association between viral diversity and epitope variant recognition, we directly sequenced 10-25 individual Gag genes from plasma RNA from five viremic controllers and five progressors. We observed a significant correlation between viral load and mean pairwise distance ($r=0.75$, $p=0.02$, **Figure 2-5A**). Diversity of the entire Gag protein was marginally higher in progressors than controllers ($p=0.05$, **Figure 2-5B**), even though controllers were on average infected for significantly longer (median 20 years) than progressors (4 years, $p=0.008$, data not shown), and diversity usually increases with duration of infection (Shankarappa et al., 1999). This effect was more pronounced when examining pairwise distance only in viral regions corresponding to targeted epitopes ($p=0.0005$, **Figure 2-5C**).

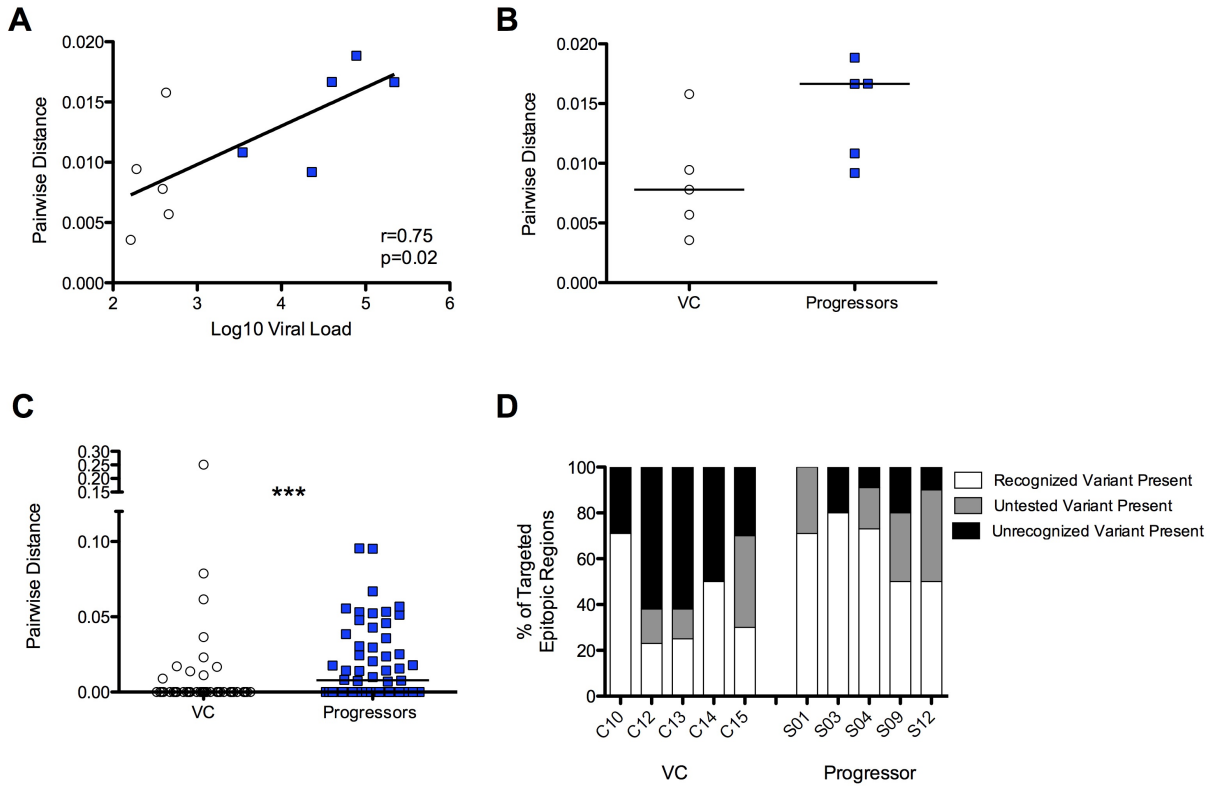


Figure 2-5: Viral diversity is higher in progressors than in VC. A) Pairwise distance is correlated with viral load in this cohort. Pairwise distance between Gag sequences in plasma viral RNA populations isolated from progressors and controllers and compared between groups B) within the whole protein ($p=0.05$) and C) within targeted epitopic regions (***: $p=0.0005$). D) Proportion of targeted epitopic regions for which recognized variants were present in the viral population (white); variants present in the viral population that were rare and not represented in the peptide set (grey); and variants present in the viral population that were tested but not recognized (black).

We next determined the relationship between the epitope variants that were recognized and the viral variants present in the autologous population. If the observed variant-specific T cells were effective at eliminating infected cells, we would not expect to see recognized epitope variants present in the plasma viral population. Within targeted epitopic regions, viremic controllers had more viral variants present in the plasma that were tested but not recognized than progressors ($p=0.008$, **Figure 2-5D**), suggesting greater evidence for escape in viremic controllers. In both controllers and progressors, we found variants present within epitopic regions that were not tested and therefore represent uncommon variants present in less than 5% of clade B sequences in the HIVDB ($p=0.34$, **Figure 2-5D**). Intriguingly, viral populations in

progressors tended to more often match the recognized epitope variants compared to viremic controllers ($p=0.07$, **Figure 2-5D**).

Coverage of frequently occurring Gag variants is associated with control of viral replication

An alternative approach to investigating the capacity of CTLs to cope with HIV sequence diversity is to evaluate recognized variants relative to the frequency in which they occur in circulating clade B sequences. For example, if an individual recognizes 3/5 tested variants (60%) but each of these variants is only present in 5% of circulating sequences, then the actual sequence coverage afforded through that variant recognition is quite poor (i.e., 5% each or 15% total). To assess the level of coverage in a given epitopic region, we summed the frequency of recognized 11mers versus the summed frequency of all tested 11mers in that region (**Table 1**). Although our peptide set did not test every possible clade B Gag variant in the database, its design ensured that we were evaluating coverage of all frequently occurring Gag variants. For each targeted epitopic region, including both conserved and variable responses, we determined the level of sequence coverage through the recognized variants (**Supplementary Figure 2-5A and 2-B**). Controllers and progressors had comparable overall immune coverage through the Gag-specific repertoire ($p=0.29$); however, elite controllers trended towards greater sequence coverage than both viremic controllers and progressors (KW $p=0.09$, **Figure 2-6A**). Interestingly, although viremic controllers recognized significantly fewer variants than progressors (**Figure 2-4C**), coverage was equivalent in the two groups, indicating that the number of variants recognized does not necessarily translate into broad sequence coverage. We found a significant inverse correlation between sequence coverage and viral load ($r=-0.38$, $p=0.03$, **Figure 2-6B**).

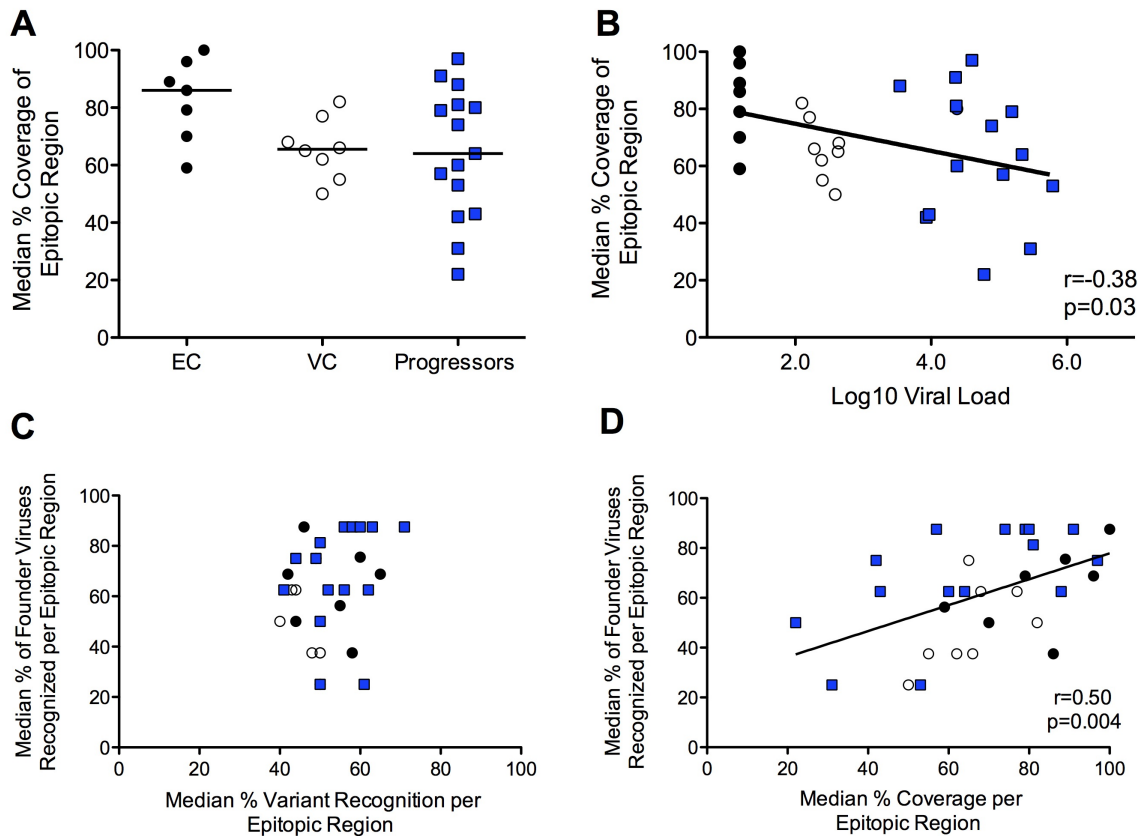


Figure 2-6: Increased sequence coverage is correlated with control of viral replication and recognition of founder viruses. A) For each study subject, the median coverage across targeted epitopic regions was calculated and compared between groups (KW $p=0.09$). B) Coverage is inversely correlated with viral load. Median recognition of founder viruses (y-axis) is C) not correlated with median variant recognition, but D) does correlate with median coverage. EC (black circles), VC (open circles), and progressors (blue squares).

Increased immune coverage is correlated with recognition of more founder viruses

Several vaccination strategies have designed multivalent immunogens with the hope of inducing T-cell responses that provide broad immune coverage (Fischer et al., 2007; Nickle et al., 2007; Rolland et al., 2007). Intuitively, increased immune coverage should increase the number of potentially infecting sequences that would be recognized. To explore how the immune coverage measured in this study can inform future vaccine evaluation, we analyzed all 347 responses relative to how often they matched a panel of eight founder viruses. This panel is a collection of clade B viruses isolated from recently infected individuals (median 7 days post onset of symptoms) in the same demographic population as the majority of participants in this

study (Herbeck et al., 2011). For each targeted epitopic region, we determined what proportion of the founder viruses would be recognized given the epitope specificity of that individual's T-cell response (**Supplementary Figure 2-6A and B**). While we did not see a correlation between recognition of founder viruses and median epitope variant recognition ($r=0.22$, $p=0.23$, **Figure 2-6C**), we did find a significant correlation between the number of founder viruses that were recognized and overall immune coverage ($r=0.50$, $p=0.004$ **Figure 2-6D**), suggesting that for vaccine-induced responses, a greater protective potential may be achieved by targeting the most commonly occurring variants rather than simply a large number of variants.

Discussion

Our study provides the first investigation into the ability to recognize all frequently occurring Gag variants in subjects with and without spontaneous control of HIV replication. In agreement with a recent study by Mothe *et al*, the shorter peptide length, tight overlap and vast sequence representation in our peptide set enabled us to detect a larger breadth of Gag-specific responses, and that targeting of the Gag protein is remarkably similar between controllers and progressors (Mothe et al., 2012). We found that recognition of epitope variants was widespread in both controlled and progressive HIV infection. Interestingly, while responses to conserved epitopes and high numbers of HIV variants have each previously been associated with immune control of HIV (Gillespie et al., 2002; Mothe et al., 2012; Turnbull et al., 2006; Zuniga et al., 2006), our study identified overall immune coverage provided through the Gag-specific repertoire as a superior correlate of immune responses associated with virologic control.

Cross-reactivity has been proposed as an essential feature of T-cell receptors (TCR) (Mason, 1998; Sewell, 2012). Recognition of naturally occurring Gag variants can result either from a T cell expressing a TCR that is exceptionally flexible, or through several different clonotypes, each recognizing one or a range of epitope variants. As we identified responses by

cytokine secretion using bulk PBMC, our assay was unable to distinguish between these two possibilities. In viremic individuals, we found that contemporaneous viremia was associated with the extent of epitope variant recognition. Investigations into viral-host dynamics during acute HIV infection have shown that *de novo* responses to viral variants develop throughout infection (Allen et al., 2005b; Henn et al., 2012; Liu et al., 2006; Liu et al., 2011). Therefore, as different TCRs have differential abilities to tolerate variation within epitopes, sequential *de novo* responses associated with high viremia have the potential to increase the breadth and depth of responses, resulting in the extensive epitope variant recognition observed in progressive infection. Interestingly, we observed greater evidence for escape (i.e., fixation of a viral variant that was tested but not recognized) in epitopes targeted by viremic controllers compared to progressors. While these responses appear to be effective at driving escape, the lower variant recognition in viremic controllers may reflect that these individuals make fewer *de novo* responses to viral variants, possibly due to the weak potential of low viremia to prime new responses, or selection of escape mutations that abrogate epitope-HLA binding (Carlson et al., 2012).

Constant antigenic stimulation has been shown to be necessary for long-term persistence of virus-specific CTLs during chronic viral infection (Liu et al., 2011; Shin et al., 2007). The lack of IL-2 production in the majority of responses to epitope variants, coupled with the contraction of variant recognition upon ARV-mediated suppression of viral load, supports a model in which the maintenance of variant-specific responses is antigen-dependent. Increasingly diverse viral quasispecies can also provide more opportunities for different TCRs to engage with a recognized antigen to stimulate T-cell proliferation. However, persistent antigenic stimulation can also promote T-cell exhaustion and functional impairment (Appay et al., 2000; Day et al., 2006; Janbazian et al., 2012; Streeck et al., 2008; Wherry et al., 2003). Exhausted T cells characteristically exhibit weak cytotoxic potential and loss of polyfunctionality, but maintain

the ability to secrete IFN- γ (Wherry et al., 2003). In the viral population of progressors, and to a lesser extent in viremic controllers, we found evidence for ineffective variant-specific responses, as many viral variants were present in the population despite being recognized. As this is a cross-sectional study, we cannot rule out that the viral populations were in the process of escaping (in the cases in which a recognized variant was a minor variant) or eventually will escape in response to immune pressure (in the cases in which a recognized variant was a major variant). Given that the majority of responses to epitope variants detected in this analysis were monofunctional IFN- γ -positive, the ineffectiveness of these responses may be a consequence of exhaustion. Our results are consistent with previous reports that have shown that the ability to recognize peptide variants as measured by IFN- γ secretion does not directly translate into an ability to inhibit replication of viruses harboring recognized variants (Bennett et al., 2008; Valentine et al., 2008). Taken together, our data suggest a perpetual cycle exists in viremic individuals, by which the combination of high viral loads and diversity can drive the development and persistence of variant-specific responses while simultaneously rendering these responses ineffective at fully suppressing viral replication, ultimately permitting the emergence of new viral variants.

Elite controllers maintain levels of epitope variant recognition in the absence of antigenic stimulation comparable to those in progressors, suggesting that there may be alternative immunological mechanisms responsible for preserving the ability to recognize epitope variants. T-cell functional avidity has been associated with mediating epitope variant recognition (Bennett et al., 2010; Kosmrly et al., 2010; Mothe et al., 2012). Given that functional avidity increases in the absence of antigen (Janbazian et al., 2012), the observed high levels of variant recognition in elite controllers may be due to the maintenance of a high avidity Gag repertoire. Additionally, it has been reported that compared to typical progressors, B*27 and B*57 elite controllers possess distinct clonotypes that have greater cross-reactivity to epitope variants (Chen et al.,

2012). Although we did not detect a significant difference in the overall variant recognition between individuals with and without B*57 and B*27 alleles, on a per T cell basis individual highly cross-reactive clonotypes may be contributing to the observed levels of variant recognition found in elite controllers.

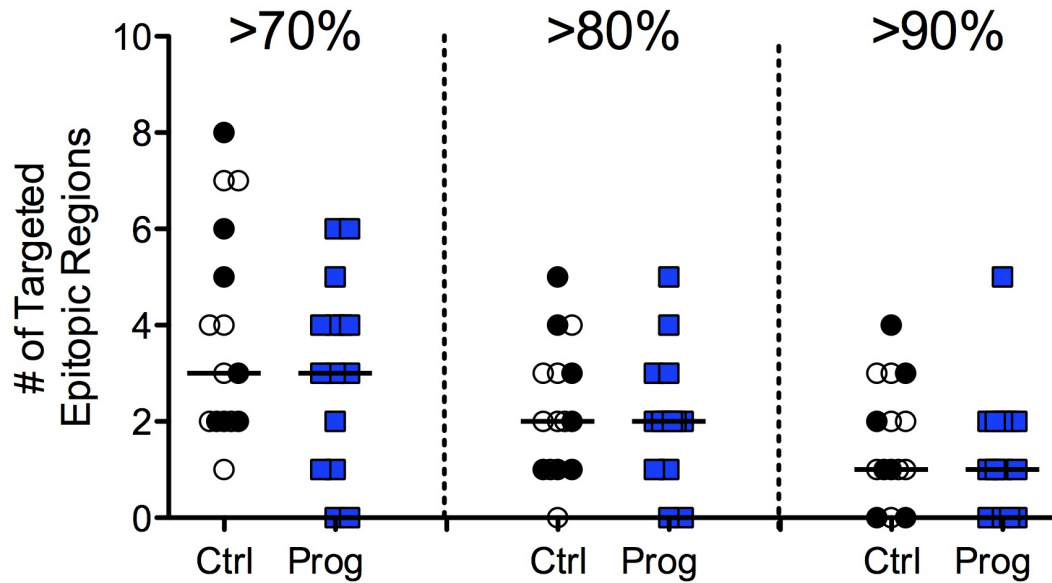
We found that increased sequence coverage by Gag-specific T-cell responses is associated with lower viral loads in infected subjects. The combination of all variant-specific and conserved responses contributes to overall immune coverage, highlighting the importance of the collective Gag-specific repertoire in control of viral replication. On a single epitope level, coverage of frequently occurring variants can lead to the selection of rare mutations to permit escape. Given the weak relationship between sequence conservation and viral fitness costs (Boutwell et al., 2013; Manochewa et al., 2013; Rihn et al., 2013; Rolland M, 2013) even high immune coverage of a single epitope may not translate directly to control. However, recent analyses have revealed extensive co-variation and networks of interdependent amino acids within the Gag protein (Carlson et al., 2008; Dahirel et al., 2011; Rolland et al., 2010). Structural and functional constraints limit the tolerance to multiple simultaneous mutations within co-evolving Gag sites (Dahirel et al., 2011). Additionally, accumulation of mutations within Gag has recently been shown to decrease HIV replicative capacity as well as contribute to lower viral load set point if transmitted to a new recipient (Goepfert et al., 2008; Prince et al., 2012). Therefore, broad immune coverage afforded through a functionally effective Gag-specific repertoire has the potential to select for multiple rare mutations across the protein. Increasing the overall level of coverage increases the likelihood that the consequences of escape will compromise the viability of the virus and allow for immune control.

Increased immune coverage through an individual's Gag-specific repertoire correlated with higher recognition of different founder viruses. Although our study was limited to clade B infections, these results are encouraging for the protective potential of several vaccination

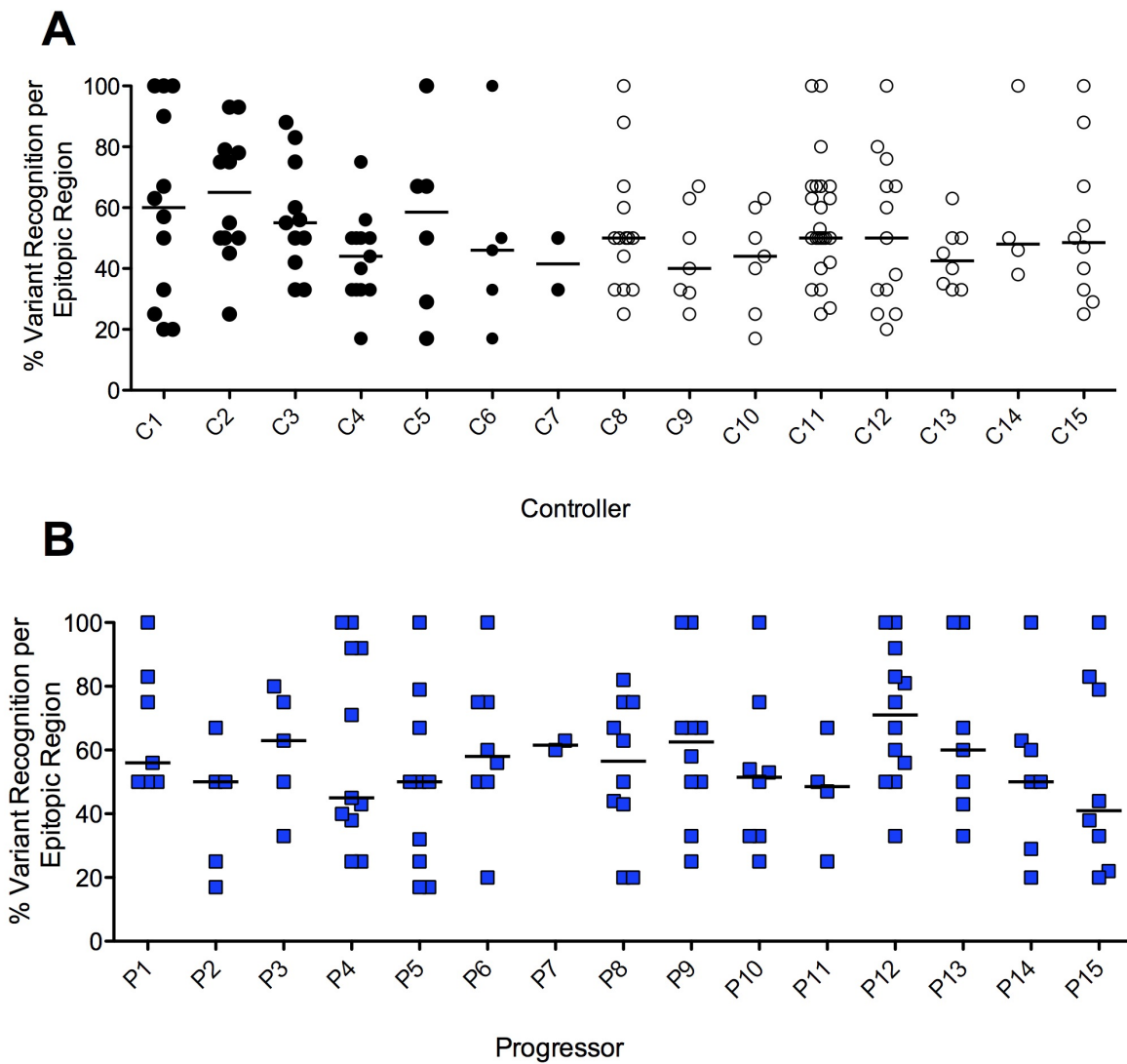
strategies designed to elicit responses of broad coverage to global HIV sequences. Polyvalent mosaic and Conserved Element vaccines both utilize multi-valent immunogens to induce variant-specific responses that will ideally provide broad coverage either against whole proteins (Fischer et al., 2007) or primarily against conserved regions (Rolland et al., 2007). Our results suggest that emphasis in the evaluation of these vaccines should be on the combined sequence coverage that can be achieved through all induced responses. However, our results were defined based on only Gag-specific responses, and it remains to be determined whether the protective potential of sequence coverage is applicable to other HIV proteins. Additionally, immunogenicity evaluation of these vaccines may depend on the identification of T-cell functional qualities that correlate with effective variant recognition. Given that dual IFN- γ /IL-2 responses to epitope variants were uncommon in our study, we were unable to evaluate whether the IFN- γ /IL-2 FluoroSpot better predicts effective variant recognition than the traditional IFN- γ ELISpot.

Taken together, our study provides a new potential correlate, “sequence coverage”, of protection from progression in HIV infection. Sequence coverage combines previously suggested markers for viral control, variant recognition and targeting of conserved epitopes, into a single measure that is more informative in predicting viral control. The inverse relationship between sequence coverage and viral load, as well as the positive correlation with the recognition of founder viruses, make sequence coverage a candidate for assessment in novel vaccine strategies that aim at increasing the breadth and depth of HIV-specific immune responses, as well as those designed to focus the immune response on conserved regions.

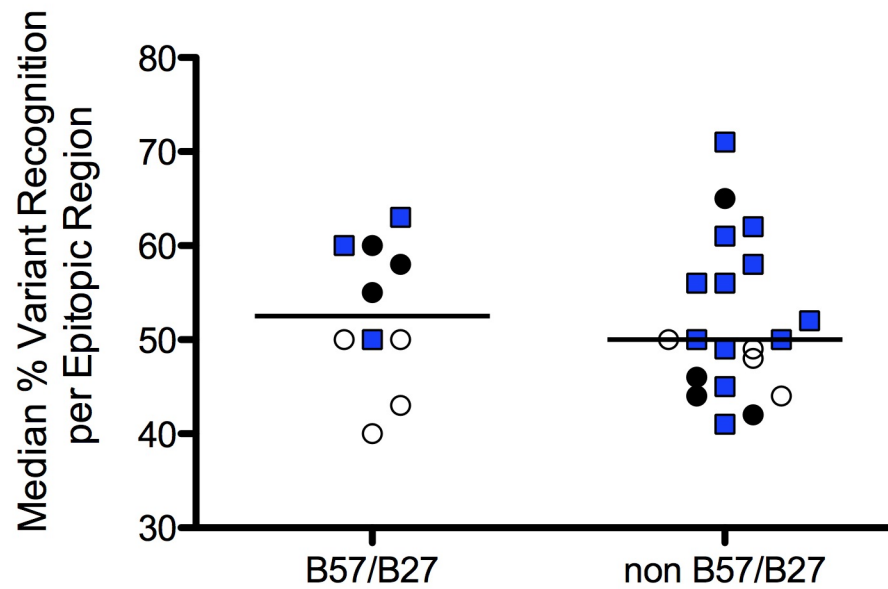
Supplementary Materials:



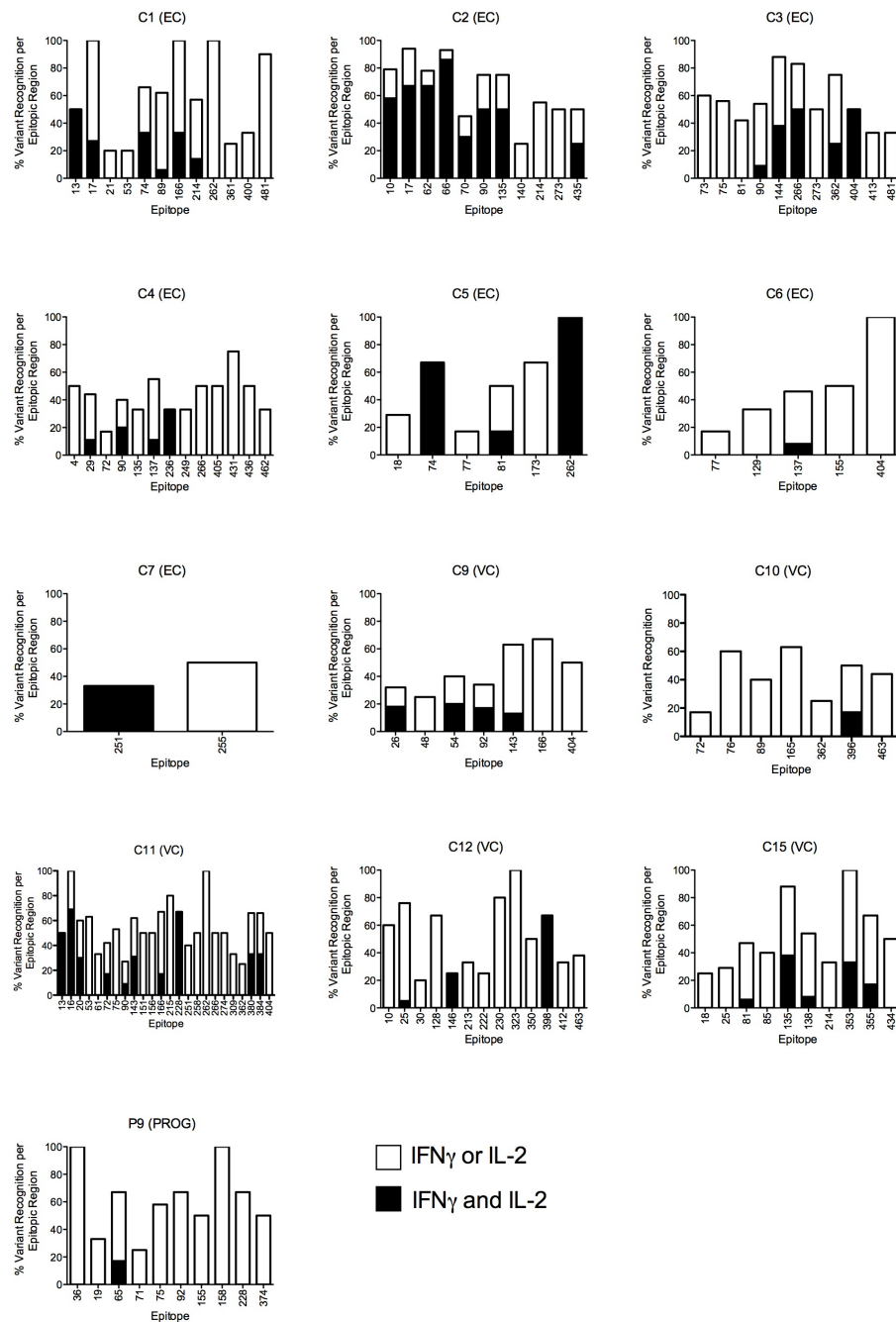
Supplementary Figure 2-1: Controllers and progressors similarly targeted conserved epitopes. Conservation was assessed based on the frequency of an 11mer sequence in the clade B HIVDB sequence alignment used to design the peptide set. A response to a conserved epitopic region reflects the targeting of at least one conserved 11mer peptide within the targeted region. We found no significant differences in the targeting of conserved epitopic regions, regardless of whether the threshold for conservation was set at >70% ($p=0.47$), >80% ($p=0.74$) or >90% ($p=0.60$) of sequences in the alignment. EC (black circles), VC (open circles), Progressors (blue squares)



Supplementary Figure 2-2: Epitope variant recognition is extensive both within and between individuals. We determined the percent variant recognition by calculating the number of 11mer variants recognized out of the number of variants tested in each epitopic region for A) EC (black circles) and VC (open circles) and B) Progressors (blue squares). Median variant recognition across all targeted epitopic regions is shown.

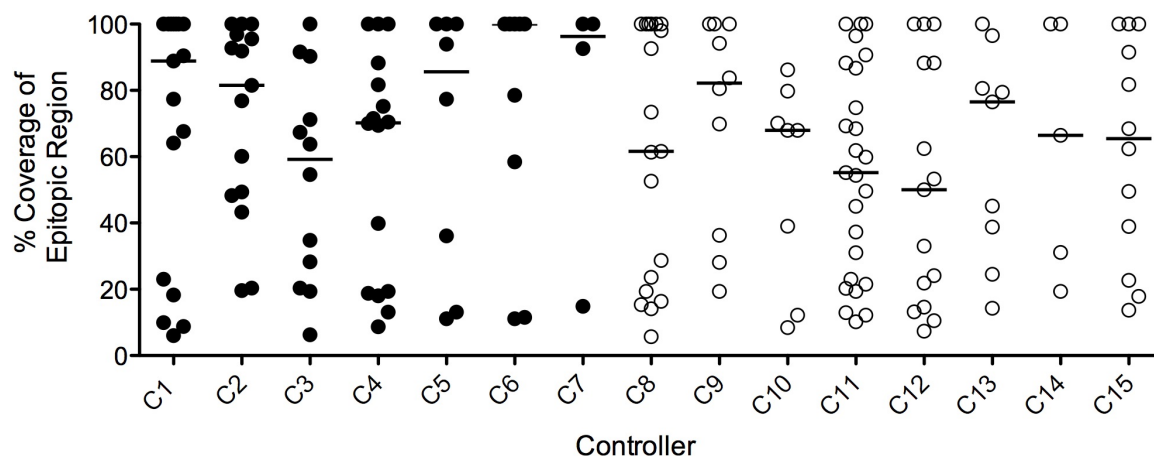


Supplementary Figure 2-3: Overall variant recognition is similar between individuals with and without B*27 and B*57 alleles. We observed no difference in median percent variant recognition between study subjects (EC black circles, VC open circles, and progressors blue squares) with and without B*27 and B*57 alleles ($p=0.61$). There is no difference between groups even if analysis is restricted to only controllers ($p=0.45$).

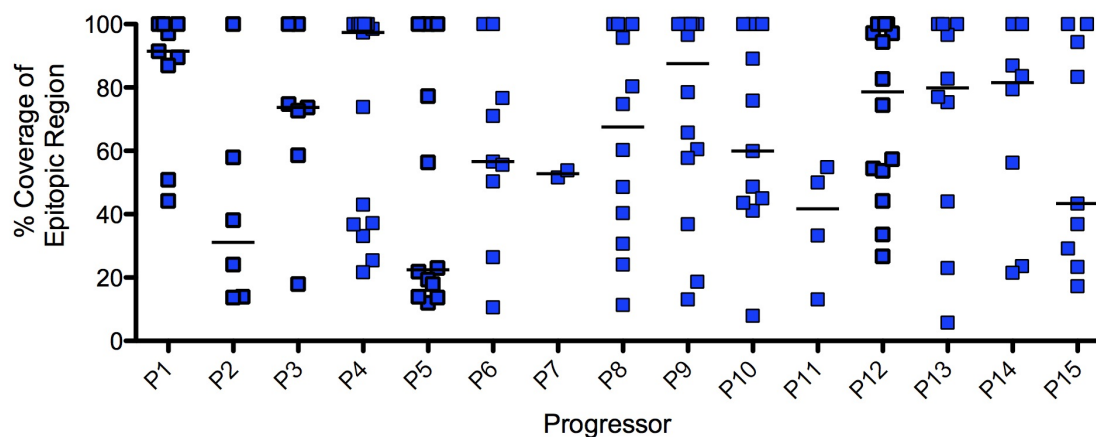


Supplementary Figure 2-4: Functionality changes upon variant recognition. In all study subjects who elicited a dual IFN- γ /IL-2 response in an epitopic region, we analyzed how functionality changes upon epitope variation. In each graph, the height of the bars reflect the overall variant recognition in a given targeted epitopic region. The proportion of variants that were recognized via a monofunctional (white, IFN- γ or IL-2) or a polyfunctional (black, IFN- γ and IL-2) response are indicated. Polyfunctional responses to all variants were uncommon.

A

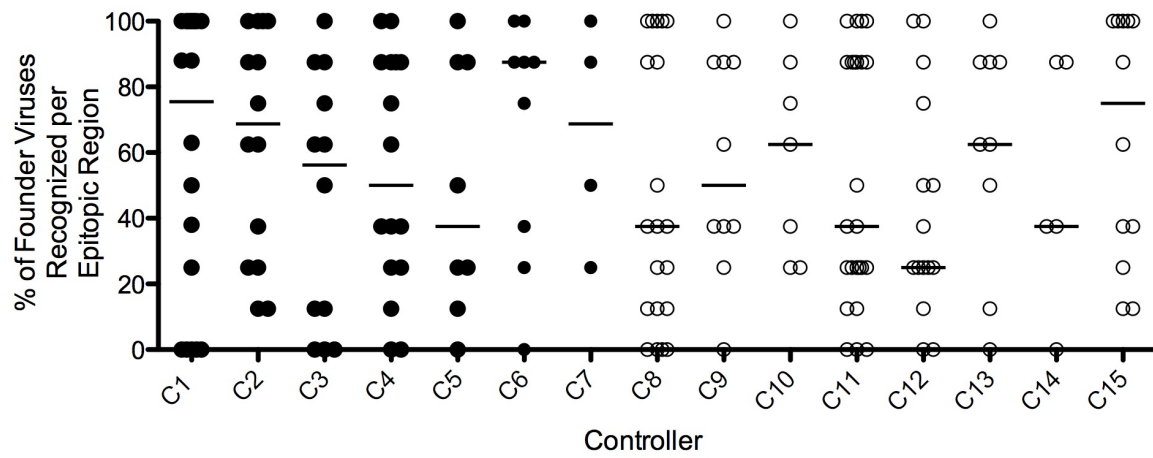


B

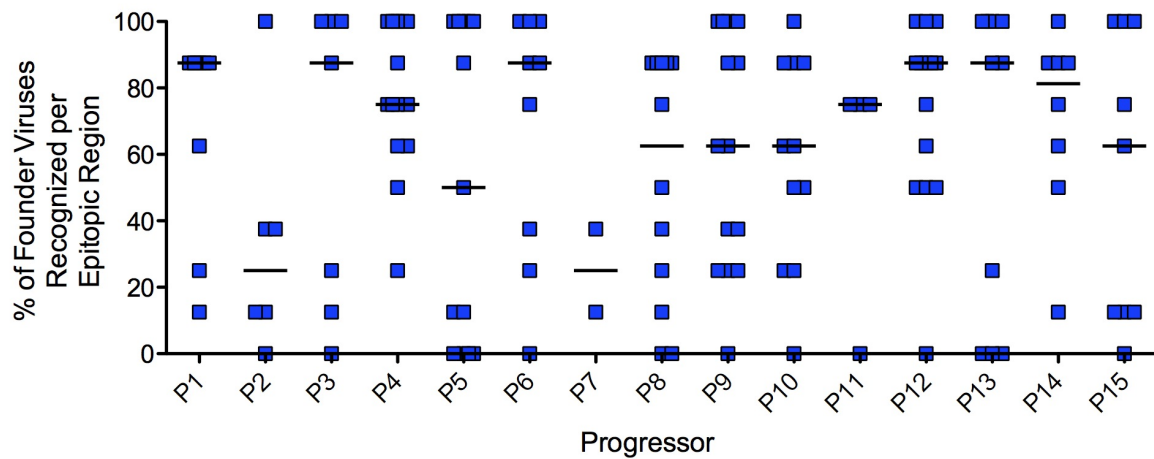


Supplementary Figure 2-5: Sequence coverage for all epitopic regions targeted by controllers and progressors. We determined the sequence coverage for each targeted epitopic region by calculating the sum [frequency targeted 11mers] out of the sum [frequency of tested 11mers] where frequency refers to the proportion of the 11mer sequence in the clade B sequence alignment. Sequence coverage was calculated for A) EC (black circles) and VC (open circles) and B) Progressors (blue squares). Lines represent median coverage across all targeted epitopic regions.

A



B



Supplementary Figure 2-6: Recognition of founder viruses by controllers and progressors. We analyzed whether the recognized variants in each targeted epitopic region would match the sequences found in the founder virus panel (n=8). For example, if a person's recognized variants matched 6/8 founder viruses, then the overall recognition would be 75%. The proportion of founder viruses that would be recognized per epitopic region was calculated for A) EC (black circles) and VC (open circles) and B) Progressors (blue squares). Lines represent median percent of founder viruses recognized across all targeted epitopic regions.

Chapter III:

**Fitness-balanced escape determines
resolution of dynamic CTL escape
processes during primary HIV-1 infection**

Abstract

Investigations into host-virus interactions during primary HIV-1 infection are critical for the design of future vaccine immunogens. To understand the dynamic interplay between host cytotoxic t-lymphocyte (CTL) responses and the mechanisms by which HIV-1 evades those responses, we studied viral evolutionary patterns and host CTL responses in six linked transmission pairs. Viral *gag* sequences were generated by 454 pyrosequencing in all pairs and seroconverting partners were followed extensively for a median of 847 days post infection. T-cell responses were screened by IFN γ /IL-2 FluoroSpot using autologous peptide sets reflecting any variant present in at least 5% of sequence reads. We detected a total of 12 responses in the transmitting partners and 16 responses in the seroconverting partners, ten of which evolved over the study period. Interestingly, cross-reactivity to epitope variants was found in 70% of evolving epitopes. We observed little evidence for CTL reversions and found escape processes during primary infection to be highly variable, with limited overlap in the patterns of escape between individuals. In cases in which multiple escape mutations developed within a targeted epitope, we found that the variant that ultimately became fixed in the viral population was selected due to achieving optimal fitness-balanced escape, a process that may be predicted by the differential database frequencies of the multiple variants. These results further our understanding of the impact of CTL escape and reversion during acute HIV infection as well as contribute to the identification of Gag regions most vulnerable to attack.

Introduction

Exploring the dynamic interplay between human immune responses and the mechanisms by which HIV-1 evades the immune system can help inform which types of responses may be most effective to induce by vaccination. In approximately 75% of HIV sexual transmissions, a single variant establishes infection followed by exponential viral growth and rapid expansion throughout the body (Abrahams et al., 2009; Herbeck et al., 2011; Keele et al., 2008; Shankarappa et al., 1999; Zhu et al., 1993). At peak viral loads, the viral population is relatively homogenous (Herbeck et al., 2011; Keele et al., 2008; Zhu et al., 1993). The appearance of the first HIV-specific CTLs responses is temporally associated with the initial decline in peak viremia, likely due to ability of these early CTLs to effectively eliminate HIV-infected cells and inhibit viral replication (Koup et al., 1994). Therefore, the founder-specific CTL responses incur early selective pressures on the viral population. Consequently, the virus rapidly selects for mutations both within and flanking targeted epitopes in order to escape the immune response (Allen et al., 2005a; Liu et al., 2006). Furthermore, if the founder virus carries CTL mutations previously selected to confer escape in the transmitting partner, these mutations may revert to the more sensitive form upon transmission to a new host of a different HLA background (Allen et al., 2004; Friedrich et al., 2004; Li et al., 2007; Liu et al., 2006). Taken together, CTL pressure in the forms of selective escape and reversion has been suggested to be a major driving force of early HIV evolution.

Preferential CTL targeting of the Gag protein has been associated with lower viral loads during chronic infection and thus is considered to be the primary target for CTL based vaccines (Kiepiela et al., 2007; Rolland et al., 2008; Zuniga et al., 2006). Escape and potential reversion mutations within Gag epitopes have been widely reported during both acute and chronic HIV infection. One factor that is believed to influence the rate of escape and reversion is the impact of the mutation on the replicative fitness of the virus (Ganusov et al., 2011; Liu et al., 2007;

Martinez-Picado et al., 2006; Troyer et al., 2009). Given the differential structural and functional constraints of the Gag protein, some amino acid sites exhibit high replicative fitness costs upon mutation whereas others can readily tolerate multiple different amino acids substitutions (Boutwell et al., 2013; Manochewa et al., 2013; Rihn et al., 2013; Rolland M, 2013). Following common escape and reversion CTL pathways during acute infection can help identify vulnerable viral regions that exhibit the greatest fitness consequences if changed. The most well defined CTL escape and reversion processes are from responses targeted by alleles associated with control of HIV replication (B*27 and B*57). These responses consistently promote dominant escape and reversion pathways in which only a select few mutations are required (Brockman et al., 2007; Miura et al., 2009; Schneidewind et al., 2008). However, in responses restricted by more common alleles, clear evidence of reversions is inconsistent and escape processes appear to be highly dynamic with often multiple simultaneous or rapidly sequential escape mutations arising within a targeted epitope (Goonetilleke et al., 2009; Herbeck et al., 2011; Liu et al., 2011). In order to better design Gag immunogens that are widely immunogenic and encompass the most vulnerable regions of the protein, fitness-balanced escape processes in complex escape pathways must be better understood.

Characterizing the immune pressure necessary to promote escape pathways has been a long-standing goal in investigations of acute HIV-1 infection. Several immune correlates have been proposed to predict ensuing epitopic escape, including immunodominance of the epitope-specific response and the epitope entropy (Ferrari et al., 2011; Liu et al., 2013). These investigations have been focused on CTL responses directed against a single sequence reflecting the founder virus. However, the ability of founder-specific CTLs to recognize epitope variants has been reported in both acute HIV and SIV infection (Cale et al., 2011; Henn et al., 2012; Liu et al., 2011; Malhotra et al., 2009). Interestingly, there has been several reported examples of variants becoming fixed in the viral population despite being recognized early in

infection, suggesting that some cross-reactive CTL responses recognition fail to contain the emergence of epitope variants (Cale et al., 2011; Liu et al., 2011). Therefore, it is important to investigate the role, if any, that epitope variant recognition may play in the selection of escape variants, particularly in cases of dynamic escape processes when multiple simultaneous substitutions are circulating in the population.

To gain critical insight into early CTL responses, escape and reversion processes, and subsequent fitness consequences, we studied the founder-specific immune responses and ensuing viral evolution in six phylogenetically and epidemiologically linked HIV transmission pairs. Interestingly, we found that many of the detected responses in both seroconverting and transmitter partners displayed some degree of cross-reactivity. While there was limited evidence for CTL reversion occurring during acute infection, we identified numerous patterns of escape processes. Our results provide further understanding into the complicated viral-host interactions occurring during acute HIV-1 infection.

Materials and Methods

Study Subjects

All 12 subjects were recruited through the University of Washington Primary Infection Clinic (PIC) (Schacker et al., 1996; Schacker et al., 1998; Stekler et al., 2008). The relevant institutional review boards approved all human subject protocols, and all subjects provided written informed consent before enrollment. Seroconverting partners (SP) (n=6) were recruited during primary HIV infection (Fiebig stages I-V) and followed extensively. Transmitting partners (TP) (n=6) were recruited for a single visit around the time of transmission. Duration of infection was determined by 'days post onset of symptoms' (dps), which refers to the onset of physical symptoms of primary infection and is a median of 12 days behind transmission (Schacker et al.,

1996). All SPs were antiretroviral therapy naïve during the study period and all TP's were not on therapy at the timepoint analyzed.

Gag Sequencing

All sequencing was performed in collaboration with the Mullins lab. HIV-1 RNA was extracted from the plasma of infected individuals using the QIAamp Viral RNA Mini Kit (Qiagen, Valencia, CA) according to the manufacturer's protocol. A total of 560µL plasma was extracted and eluted in 80µL elution buffer. cDNA was synthesized using Takara BluePrint First Strand Synthesis Kit (Clontech 6115A) according to the manufacturer's protocol with 10 µL RNA. cDNA was synthesized with the gag specific primer, R3337-1 (5'-TTTCCYACTAAYTTYTGATATTCATTGAC-3').

To sequence a 2.6 kb fragment of the HIV-1 *gag - pol* gene region, first- and second-round PCRs were performed using Kapa HiFi HS (Kapa Biosystems, Boston). In brief, first-round PCR reactions were conducted in two separate 25 µL reaction volumes with 1-6 µL of cDNA as template, corresponding to 200 to 500 input template copies based on limiting endpoint dilution PCR using the template estimator program Quality (<http://indra.mullins.microbiol.washington.edu/quality/>). Second-round PCR reactions were conducted in 25 µL, using 2 µL of the first round reaction as template. A total of five multiplex PCR and one monoplex PCR were performed to generate 11 overlapping PCR amplicons ranging from 384-603 bp in size. Second-round PCR reaction products were pooled and purified using AMPure beads according to the manufacturer's protocol (Agencourt, Beverly, MA). Briefly, 180 µL of PCR product was mixed with 144 µL of AMPure beads. The solution was then vortexed, incubated, washed, and purified. Finally, the beads were resuspended in 60 µL of 10mM Tris-Cl. A Nanodrop instrument (Thermo Scientific; Waltham, MA) was used to determine DNA concentration and purity. The purified product was then diluted to 100ng/µL with 10mM Tris-HCl. All purified products were stored at -20°C prior to sequencing. The second-round

primers also contained universal adaptors A and B, as well as one of seven different MIDs described by Roche: ACGAGTGC GT, ACGCTCGACA, AGACGCACTC, AGCACTGTAG, ATCAGACACG, ATATCGCGAG, and CGTGTCTCTA (Roche; Basel, Switzerland). Second-round PCR products were visualized on a QIAxcel (Qiagen). Confirmed positive reactions were pooled and purified using AMPure beads as outlined above.

Briefly, each pooled and purified PCR product was quantified using the Quan-it PicoGreen dsDNA assay (Invitrogen; Carlsbad, CA). Each product was then diluted to a working stock of 10^7 molecules/ μ l in TE buffer. All 11 amplicons were then pooled together in equimolar ratios into a single tube for pyrosequencing. Emulsion PCR and sequencing were performed according to the manufacturer's GS FLX Titanium protocols (Roche). PCR products were added to the emulsion PCR at a ratio of 1-2 molecules per bead. The picotiter sequencing plate was prepared with a single gasket, creating two regions. Four million enriched beads were distributed equally across the two regions.

To process output from 454 sequencing, raw sequences were screened to remove any reads that were less than 100 bp or contained the ambiguous base 'N'. Sequences generated for each sample were then aligned to a participant-specific consensus using BLAST with the following parameters: match: 1; mismatch: -1; gap existence: 1; gap extension: 2. We corrected the reads using a series of scripts developed in the Mullins lab (Deng et al., 2013). For each pairwise alignment between read and the reference, the script counts the number of aligned nucleotides in the reference sequence, as well as the numbers of insertions, deletions and substitutions in the read compared to the reference.

Peptide Set

For each subject, an autologous peptide set of 15mer peptides overlapping by 11 amino acids (AA) was synthesized for use in this study (Sigma-Aldrich, St. Louis, MO). Peptide sets were designed to reflect all Gag p17 and p24 variants present in >5% of corrected sequence

reads from each individual's viral population. Peptides from all subjects were pooled based on 15mer windows, resulting in a total of 85 peptide pools.

IFN- γ /IL-2 FluoroSpot

This study employed the IFN- γ /IL-2 FluoroSpot assay for detection of Gag-specific T-cell responses. Cryopreserved peripheral blood mononuclear cells (PBMC) were thawed and incubated in R10 media (RPMI 1640 [GibcoBRL, Carlsbad, CA], 10% FBS [Gemini Bioproducts, West Sacramento, CA], 2 mM L-glutamine [Gibco], 100 μ g/ml streptomycin sulfate [Gibco], 100 U/ml penicillin G [Gibco]) overnight before stimulation. Between 70,000-100,000 PBMC/well were plated in 96-well IPFL plates (Millipore, Bedford, MA), which had been pre-coated overnight with 1 μ g/ml anti-IFN- γ monoclonal antibody 1-D1K (Mabtech, Stockholm, Sweden) and 1 μ g/ml anti-IL-2 antibody 2516KZ (BD Biosciences, New Jersey, USA). Peptide pools or individual 11mers were added at a final concentration of 2 μ g/ml. Each plate contained cells stimulated with 1.8 μ g/ml PHA as a positive control and six negative control wells of cells incubated with media alone. Plates were incubated overnight at 37°C in 5% CO₂. Plates were developed using anti-IFN- γ 7-B6-1-FS-FITC antibody diluted 1:400 (Mabtech) and 0.1 μ g/ml anti-IL-2 antibody 2311KZ (BD) for 2h, followed by anti-FITC-green diluted 1:400 (Mabtech) and streptavidin-coupled Cy3 diluted 1:1000 (Biolegend, San Diego, CA) for another hour. Plates were washed six times with PBS between all steps. Spots were counted using AID iSpot FluoroSpot Reader System.

All TPs and SPs with sufficient PBMC samples (n=10) were first screened for responses to the 85-peptide pools. SPs were screened at 1-2 timepoints during the first year of infection and TPs were screened at one time point following transmission. For each positive pool, the partner pair autologous peptides (i.e., variants from both the TP and SP) were tested in duplicate in a confirmatory IFN- γ /IL-2 FluoroSpot. Positivity for both pool and single variant responses was defined as i) >55 Spot Forming Cells (SFC)/million PBMC, ii) >4 times the

average of at least six negative control wells, iii) >3 standard deviations above average negative control wells, and iv) at least 5 spots per well.

Functional Avidity

In SPs, we determined the optimal epitope (the shortest and most immunogenic sequence) for all positive responses detected to the 15mer peptides. Optimal epitopes were first predicted based on well-known epitopes in the HIVDB as well as on HLA binding motifs (Llano, 2009). Single AA truncations and extensions of the predicted optimal epitope were also designed as previously described (Frahm et al., 2005). All potential optimal epitopes were synthesized (Sigma-Aldrich, St. Louis, MO) and tested over 8 log₁₀ dilutions using IFN- γ /IL-2 FluoroSpot. The peptide concentration needed to induce half-maximal response rate (half maximal effective concentration, EC₅₀) was determined for each peptide. The sequence with the lowest EC₅₀ value (i.e. highest avidity) was designated as the optimal epitope. If an individual recognized variants in a particular region as determined by the 15mer peptide analysis, then additional peptides were generated to reflect the optimal epitope containing these variants. Variant optimals were also titrated over 8 log₁₀ dilutions using IFN- γ /IL-2 FluoroSpot and the EC₅₀ value was calculated. The majority of detected responses were IFN- γ secreting and therefore all EC₅₀ values were calculated based on IFN- γ .

Generation of mutant NL4-3 viral stocks

Transient and replacement mutants identified during PIC64236 RK9 (K28T, K28E, K28R, K28Q) and PIC68008 TW10 (G248E, T242N) epitope evolution were introduced into the NL4-3 viral strain. Site-directed mutagenesis was performed using restriction-free cloning of synthesized gblocks (synthetic gene products up to 500 base pairs) containing the mutation of interest (Integrated DNA Technologies; IDT, Coralville, IA). We found this method to be more efficient than traditional methods for site directed mutagenesis. The gblock containing the

mutated site of interest was amplified in a first round PCR using primers PIC64236_RK9 [RK9_F AGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGG, RK9_R CTTCTGATGTCTCTAAAAGGCCAGGATTAAGTGCGAATCGTTCTAGCTCC] and PIC68008_TW10 [TW10_F ACAAGATTTAAATACCATGCTAAACACAGTGGGGGGACATCAAGCAGCCAT, TW10_R TGTCATCCAATTTTTTACCTCTTGTGAAGCTTGCTCGGCTCTTAGAGTTTTATAGA]. The first round of PCR used Phusion Hot Start Flex polymerase (New England BioLabs, Ipswich, MA) and the following conditions: 98°C for 30s, 25 cycles of 98°C for 8s, 55°C for 20s, 72°C for 15s and a final extension step at 72°C for 5 min. Amplified gblock PCR fragments were PCR purified using a PCR clean-up kit (Clontech). To insert the gblock into NL4-3 backbone, 100ng of purified amplified gblock was used as the primer in the second round PCR. The second round PCR used the same conditions except primer annealing was done at 60°C instead of 55°C.

Second round PCR products were digested with Dpn1 and transformed into One Shot TOP10 chemically competent *E. coli* (Invitrogen, New York, USA). Single colonies were picked and grown overnight followed by plasmid DNA extraction using QIAprep Spin Miniprep Kit (Qiagen) and sequenced in *gag* to confirm the desired mutagenesis.

Plasmid preps were prepared using the QIAprep HiSpeed Plasmid Midi Kit (Qiagen) with an additional endotoxin removal step for use in mammalian cell transfection experiments (Lanxon-Cookson et al., 2013). HEK 293T cells were transfected with 2µg of plasmid DNA using the XtremeGENE 9 DNA transfection reagent (Roche, San Francisco, CA). Following 48 hours at 37°C, cell-free supernatants were collected and filtered by double centrifugation. Viral stock aliquots were stored at -80°C until use. The viral titer for each stock was calculated by determining endpoints in a serial dilution assay (Reed, 1938) on PBMC from a single uninfected donor, where positive wells were quantified by p24 ELISA (McClure et al., 2007).

Pairwise Growth Competition Assay

PBMC from the same uninfected donor used for virus titration were infected with NL4-3 and variant virus 72 hours after PHA stimulation (Lanxon-Cookson et al., 2013). Cells were cultured in Iscove's Modified Dulbecco's Medium (IMDM) supplemented with 20U/ml of human interleukin 2 (hIL-2; Roche), 10% fetal bovine serum (Sigma-Aldrich) and 1% penicillin/streptomycin. Viruses were added at multiplicity of infection (MOI) of 0.005 in 3×10^5 PBMC/well and were washed 24 hours post infection (Lanxon-Cookson et al., 2013). All dual infections were done in triplicate. Cell-free supernatant (400 μ l) was collected from each replicate on day 0, 3, 5, 7, and 9 and stored at -80°C in 200 μ l aliquots.

Viral RNA was extracted from 200 μ l supernatant aliquots with a QIAextractor robot (Qiagen) as described in the manufacturer's protocol. cDNA was synthesized using SuperScript III (Invitrogen) with primer RT2 with the following conditions: 50°C for 1.5 hours followed by 70°C for 15 min. Depending on the competition, the PIC64236_RK9 or PIC68008_TW10 regions were PCR amplified using Taq polymerase and the following conditions: 94°C for 60s, 34 cycles of 94°C for 15s, 58°C for 30s, 72°C for 60s and a final extension step at 72°C for 5 min. Purified PCR products containing gene of interest were submitted directly for dideoxynucleotide chain termination (Sanger) sequencing. Chromatograms were assembled and edited using Geneious 5.6.5 (Biomatters, NZ).

The relative proportion of the viruses in culture over time was determined by measuring the area under the curve on chromatograms at the mutated nucleotide of interest using an in-house web tool (<http://indra.mullins.microbiol.washington.edu/cgi-bin/chromatquant.cgi>) (Lanxon-Cookson et al., 2013). The relative fitness (1+s) was estimated using a web-based linear regression algorithm (<http://bis.urmc.rochester.edu/vFitness/FitnessMulti.aspx>) that allows for the assessment of the rate of change in mutant frequency over multiple data points, where data points here refer to days of culture (Lanxon-Cookson et al., 2013).

Data Analysis

Breadth of responses in SPs was calculated as the number of optimal epitopes defined during the first year of infection. In TPs, breadth was calculated based on the number of targeted 15mer peptides. Responses to overlapping 15mers were counted as a single epitope. HIVDB frequency of a particular AA site was calculated using the QuickAlign web-tool (http://www.hiv.lanl.gov/content/sequence/QUICK_ALIGN/QuickAlign.html). Shannon Entropy for each AA in Gag was calculated using LANL entropy web-tool (<http://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy.html>) using an alignment of 410 clade B Gag sequences extracted from the HIVDB. Escape mutations were defined as variant AA that did not elicit a response in the 15mer peptide screen.

Statistical Analysis

Comparisons between evolving and persistent epitopes were assessed using a non-parametric Mann-Whitney test. Non-linear regressions were performed to determine the half maximal effective concentration (EC_{50}) value in the functional avidity analyses. Functional avidity differences between founder and variant epitopes were assessed using a non-parametric Wilcoxon signed rank test. Relative fitness differences between viruses were determined using a student's T-test. GraphPad Prism X was used for all statistical analyses.

Results

Detection of primarily IFN γ -secreting Gag-specific responses in both Transmitting and Seroconverting Partners

Six epidemiologically and phylogenetically linked HIV-1 transmission pairs whose viral populations had been previously sequenced were included in this study (**Table 3-1**).

Seroconverting Partners (SPs) were first sequenced by a median of 25 days post onset of

symptoms (dps) and then followed up longitudinally for a median of 827 dps, whereas Transmitting Partners (TP) were sequenced at one timepoint near the time of transmission (**Table 3-1**). Viral sequences in Gag p17 and p24 were obtained by 454 pyrosequencing or by generating 15-20 Sanger sequences when material was limiting. Phylogenetic analysis supported a single viral variant establishing infection in 5/6 SPs whereas PIC51861 was infected with two variants.

For each partner pair, we generated autologous 15mer peptides overlapping by 11 amino acids (AA) corresponding to Gag p17 and p24 sequences. Peptides were designed to reflect any AA variant found in $\geq 5\%$ of sequence reads obtained from both the TP and SP. Each subject was screened for responses by IFN- γ /IL-2 FluoroSpot using their partner pair's autologous peptide set. TPs with available PBMC samples (n=5) were screened for responses at the same timepoint used for sequencing. SPs were screened for responses within the first year of HIV-1 infection at 1-2 timepoints depending on sample availability (**Table 3-1**). One SP could not be screened due to insufficient PBMC samples. We also fine mapped the optimal epitopes for responses detected in the SPs. All SPs were antiretroviral therapy naïve and no TPs were on treatment at the timepoint used for immunological assays.

Table 3-1: PIC Partner Pair Characteristics

Partner Pair	Sequenced Timepoint (DPS) ^a	Viral Load	Partner Pair	Sequenced Timepoint (DPS)	Viral Load
PIC37628 (TP)	NT^b	11934	PIC39522 (TP)	NT	5483
PIC64236 (SP)	29	26713	PIC99203 (SP)	27	N/A
	49	4283		48	1938
	91	7123		63	768
	146	4202		111	2803
	257	12293		174	2308
	428	12631		212	603
	896	8796		370	3636
	1265	11330			
PIC67505 (TP)	NT	14009	PIC11473 (TP)	NT	102720
PIC68008 (SP)	24	11736	PIC90770 (SP)	3	469830
	32	7503		11	N/A
	60	2922		16	42810
	207	515		43	41710

	366	905		71	12220
	872	7065		99	14170
	1250	1219		129	18210
				197	70780
				330	81970
				576	66063
				751	12137
PIC68415 (TP)	NT	22243	PIC15332 (TP)	NT	5483
PIC44149 (SP)	22	3755	PIC51861 (SP)	26	11535
	63	4100		47	4233
	128	394		84	388
	161	<50		112	4375
	259	1408		194	990
	350	1563		250	N/A
				532	<50
				903	190

^aDays post onset of symptoms

^bNT: Near transmission

Timepoint used in immunological assay

We identified Gag-specific responses to 12 epitopes in the TPs and 16 epitopes in the SPs within the first year of infection (**Figure 3-1A and B**). We detected primarily IFN- γ responses to the targeted epitopes in both the TPs and the SPs (**Figure 3-1A and B**). Previous studies have shown that in subjects with progressive HIV infection, 95% of responses were found to be CTL mediated when detected by IFN γ ELISpot, a technically similar assay (Altfeld et al., 2003b). However, given the use of bulk PBMC to detect responses, we cannot rule out the possibility we are also detecting CD4⁺ T-cell responses.

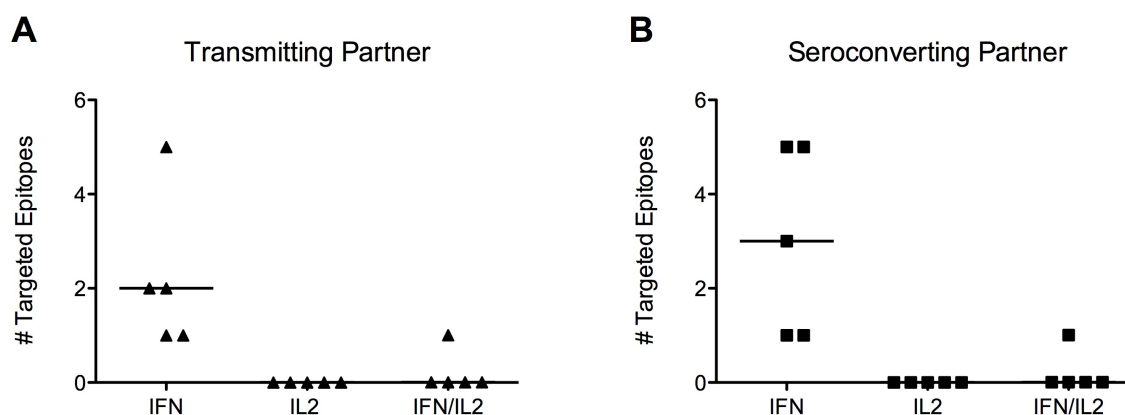


Figure 3-1: Breadth of Gag-specific responses in transmitting and seroconverting partners. Cohort includes six linked HIV-transmission pairs, sufficient PBMC samples for analysis were available for 10 subjects. Breadth of

Gag-specific responses are stratified by functionality: monofunctional IFN- γ (left) or IL-2 (middle) and dual functional IFN- γ /IL-2 responses (right column). **A)** Breadth in Transmitting Partners was calculated as the number of recognized 15mer peptides. Responses to overlapping 15-mer peptides were calculated as one targeted epitope. **B)** Breadth in Seroconverting Partners was calculated as the number of optimal epitopes identified during the first year of infection.

CTL pressure is a major selective force shaping early Gag evolution

CTL escape and reversion have been repeatedly reported to be major selective forces driving early HIV-1 evolution (Allen et al., 2005a; Herbeck et al., 2011; Li et al., 2007; Liu et al., 2006). To determine the role of CTL pressure in shaping early Gag evolution in our seroconverting partners, we determined how many evolving AA sites identified through sequence analysis were located within or nearby detected epitopes targeted by both the SP (to identify potential escape mutations) and TP (to identify putative reversion mutations). We found mutations occurring in a median of 9.5 AA p17/p24 sites in the SP viral populations during the observed study period (**Table 3-2**). Most of the mutations identified were changes to an AA of lower database frequency (**Table 3-2**). A median of 33% of these mutations were found to be occurring in AA sites within or nearby epitopes targeted by the SP, suggesting that the mutations developed in order to facilitate escape from host CTL pressure by disruption of the processing, recognition, or presentation of targeted epitopes (**Table 3-2**). Although the majority of these potential escape mutations reflected a decrease in database frequency, mutations within targeted SP epitopes did not account for all observed changes to lower database frequency across Gag. Our results are slightly lower than the previously described ~50% of mutations being associated with escape from CTL responses (Allen et al., 2005a; Liu et al., 2006). However, our analysis is restricted to only Gag p17 and p24 compared to previously reported whole proteome studies.

Transmitted escape mutations may contain associated fitness costs and thus reverting these AA sites to more consensus-like AA during primary HIV infection may improve the overall fitness of the viral population (Friedrich et al., 2004; Liu et al., 2007). Because most escape mutations are often of lower database frequency, potential reversion mutations are frequently

calculated as substitutions reflecting an increase in database frequency (Herbeck et al., 2011). A median of 18% of evolving sites involved a mutation from a lower to higher database frequency (**Table 3-2**). A median of only 11% of mutating sites in the SP were located within or nearby epitopes targeted by the TP (**Table 3-2**) and only two of these potentially reverting sites involved a mutation to a higher database frequency.

Table 3-2: Early Gag Evolution in Seroconverting Partner

Partner Pair (TP/SP)	#Evolving AA sites in SP	#Evolving AA sites to lower DB frequency ^a	#Evolving AA sites to higher DB frequency ^a	# Evolving AA sites within/nearby SP epitopes	# Evolving AA sites within/nearby TP epitopes
37628/64236	19	14 (74%)	2 (11%)	8 (23%)	2 (11%)
67505/68008	13	10 (77%)	1 (8%)	3 (54%)	1 (8%)
68415/44149	5	4 (80%)	1 (20%)	4 (80%)	1 (20%)
15332/51861	10	6 (67%)	2 (22%)	2 (20%)	1 (10%)
39522/99203	6	4 (67%)	1 (17%)	2 (20%)	ND ^b
11473/90770	10	8 (80%)	2 (20%)	ND ^b	2 (20%)
Median	9.5	7 (75%)	1.5 (18%)	3 (33%)	1 (11%)

^a Amino acid changes to equal DB frequency are not included

^bNot determined: no PBMC precluded analysis of targeted epitopes

Evolution within TP targeted epitopes do not reflect classical reversion patterns

The classical reversion pattern involves the transmission of a lower database frequency AA that confers escape from a TP CTL response followed by a reversion to the higher database frequency sensitive AA (i.e., pre-escape AA) in the new host. Interestingly, we found no evidence for classical reversion patterns in the five TP epitopes that were evolving over time in the SP viral population.

We found two similar reversion patterns in TPs PIC37628 and PIC67505. While we only detected V variants in the contemporaneous viral population, both TPs recognized V and I variants in the sequences shown in **Figure 3-2A and B**. In both cases, the V variant emerged in

their respective SP and was maintained in the viral population until eventual replacement of the I variant. Both of these mutations reflected a decrease in database frequency (AA#159: V 73% I 23%, AA#191 V 98% I 1%). As both variants were recognized in these examples, it is unclear which, if any, of these mutations initially developed to confer escape in the TP. The ongoing evolution to a lower database frequency AA in the SP viral population suggests that these changes may be due to an alternative selection pressure in the SP rather than selection for an AA that could improve replicative fitness following transmission of an escape mutant.

In TP PIC15332, we detected responses to 15mer peptides containing the E11 variant but not to peptides representing the A11 variant, suggesting that E11A was an escape mutation at the time of transmission (**Figure 3-2C**). PIC15332 transmitted both E11 and A11 variants to SP PIC51861. The E11 variant became fixed in the SP by 532dps (**Figure 3-2C**). Although we found evidence for A11 variant to be a transmitted escape mutation, the eventual fixation of E11 was due to outcompeting A11 in the SP viral population and not due to reversion.

TP PIC67505 only targeted peptides containing Y79 and not F79, indicating that Y79F is an escape mutation (**Figure 3-2D**). Interestingly, the sensitive variant (Y79) was the major variant in the TP viral population at the time of transmission and was transmitted to SP PIC68008 (**Figure 3-2D**). After 366dps, the Y79 variant declines in the SP population and F79 becomes the major viral variant. This targeted region contains the A*01 restricted GY9 epitope (₇₁GSEELKSLY₇₉) and both TP and SP express HLA A*01. We did not detect a response to GY9 during the first year of infection but this response may have developed later resulting in the selection of F79 by 872dps. Therefore, it is possible that because the sensitive form was transmitted, this allowed for PIC68008 to target this region and promote the same escape pathway as seen in the TP. Finally, in the last evolving epitope previously targeted by a TP, we found evidence for simultaneous reversion and escape processes in TP PIC68415 and SP PIC44149 (presented below in Figure 3-7).

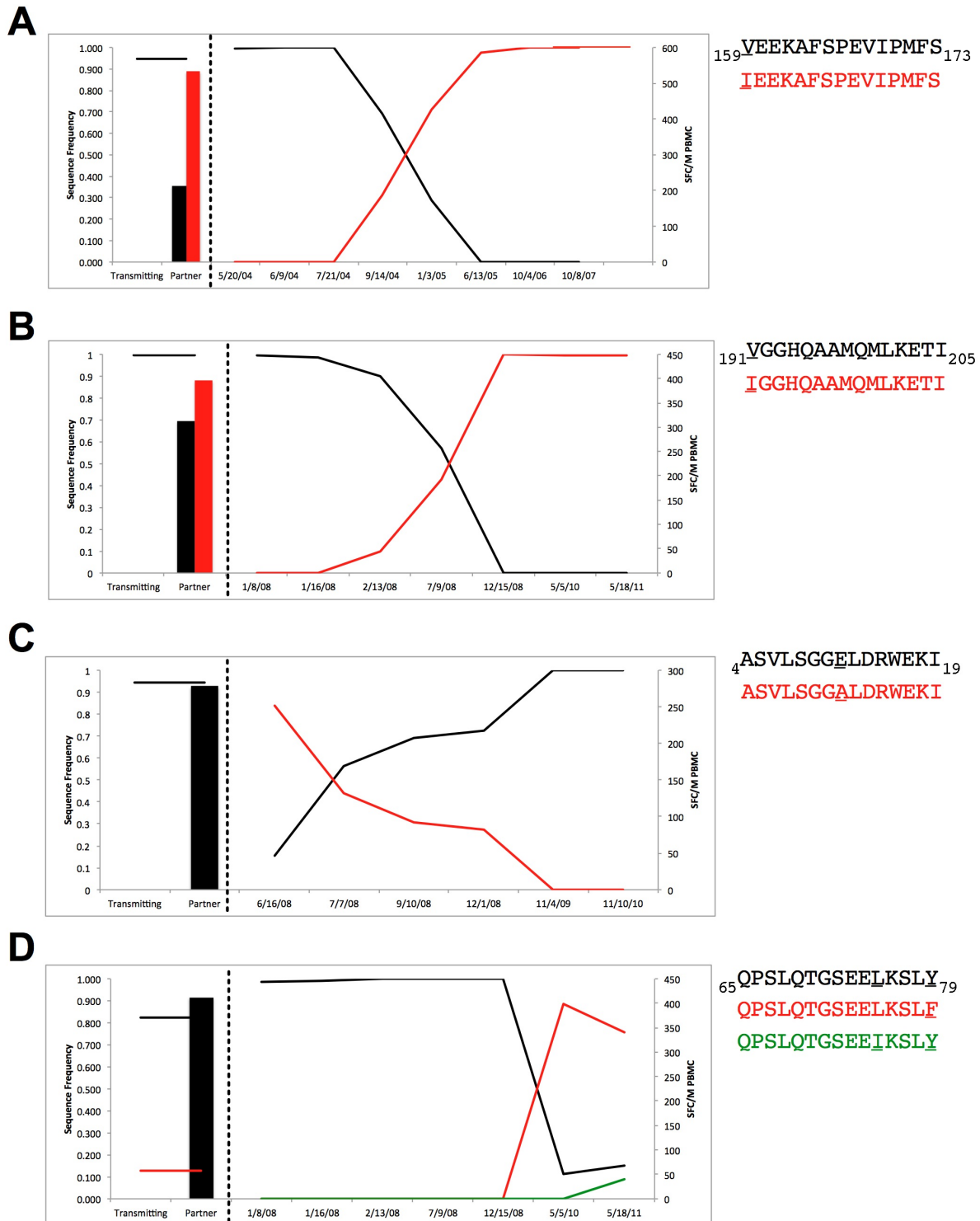


Figure 3-2: Evolution of potentially reverting sites in SP viral populations. In all graphs: sequence frequency of variant over time (line graph, left y-axis), magnitude of IFN γ response to 15mer peptide (bar graph, right y-axis) and visit date (x-axis). Please note, x-axis is not on a linear scale. Sequence frequency and response magnitude left of the dotted lines refers to viral population and responses in the transmitting partner. All data right of dotted line represents sequence evolution within the SP viral population. All graphs are color-coordinated to the specific variant

that was tested. All founder variants are in black. Evolution and TP immune responses directed at site **A)** 159 in TP37628 and SP64236 **B)** 191 in TP67505 and SP68008 **C)** 11 in TP15332 and SP51861 and **D)** 75/79 in TP37628 and SP64236.

Entropy, functional avidity, and immunodominance in responses to evolving epitopes

Of the 16 epitopes identified in the SPs, ten epitopes evolved over the course of the study period whereas six targeted epitopes remained unchanged (**Table 3-3**). Due to differential structural and functional constraints across the Gag protein, some viral regions are more tolerant of accumulating mutations. One measure that can predict this toleration is Shannon entropy, as entropy is a measure of population diversity at a particular site or group of sites (Liu et al., 2013). We calculated the Shannon entropy for each targeted epitope utilizing sequences deposited in the HIVDB and found that epitopes that evolved over time were of higher entropy than those that persisted ($p=0.004$, **Figure 3-3A**). These results are in agreement with several recent reports, highlighting the importance of the influence of epitope entropy on escape (Ferrari et al., 2011; Liu et al., 2013).

Table 3-3: SP Gag-specific responses within first year of infection

SP	Epitope /HLA ^A	Optimal	p17/p24	Evolving or Persistent
1691	RK9 (A30)	RLRPGGKKK	p17	Evolving
1691	LY12(?)	LKSLYNTVAVLY	p17	Evolving
1691	LI10(C06)	LYYVHQRIEI	p17	Evolving
1691	QW9(B53)	QASQDVKNW	p24	Evolving
1802	TW10(B57)	TSTLQEQVGW	p24	Evolving
1853	TV8(B08)	TVAVLLCV	p17	Evolving
1853	DL9(B08)	DIKDTKEAL	p17	Evolving
1853	EI8(B08)	EIYKRWII	p24	Persistent
1853	DV9(B08)	DYVDRFYKV	p24	Persistent
1853	NL8(B08)	NPDCKTIL	p24	Persistent
1810	KY12(C07)	KIRLRPGGKKKY	p17	Evolving
1839	GR11(A68)	GQMVHQALSPR	p24	Evolving
1839	NL11(A02)	NLQGQMVHQAL	p24	Evolving
1839	KW11(?)	KETINEEAAEW	p24	Persistent
1839	LY10(C03)	LGLNKIVRMV	p24	Persistent
1839	EL15	EKIRLRPGGKKKYRL	p17	Persistent

^Apredicted HLA restriction. ? refers to unknown HLA restriction

^BNot determined: limited PBMC precluded determination of optimal

Functional avidity, defined as the antigen concentration needed to elicit the half-maximal response of antigen-specific T-cells (EC_{50}), has been reported to be predictive of the selective pressure being exerted by the responding T cell, as high avidity T cells have been shown to be superior in the elimination of virally-infected target cells (Alexander-Miller et al., 1996; Derby et al., 2001). To determine the functional avidity of the founder-specific responses detected within the first year of infection, we titrated peptides representing the optimal epitope over seven $\log_{10}$ dilutions in each SP using IFN- γ /IL-2 FluoroSpot. Additionally, we evaluated the functional avidity of responses to any epitope variant that was recognized in the original 15mer screen. A high functional avidity is reflected as low EC_{50} (i.e., a low peptide concentration is needed to induce the response). We found that within the first year of infection, founder-specific responses to epitopes that subsequently evolve are of significantly higher functional avidity than founder-specific responses to epitopes that remain persistent ($p=0.03$, **Figure 3-3B**). This is in line with the rapid escape observed in high avidity epitopes in acute SIV infection (O'Connor et al., 2002).

In addition to functional avidity, T-cell immunodominance has also been reported as a parameter in driving escape (Liu et al., 2013). Immunodominant epitopes are those eliciting the highest magnitude within an individual (intra-individual immunodominance), or those most frequently targeted across individuals (inter-individual immunodominance). While there was no difference in the overall magnitude of responses against evolving or persistent epitopes ($p=0.84$, data not shown), we found a significant difference in the dominance between evolving and persistent epitopes ($p=0.03$, **Figure 3-3C**). Taken together, our results suggest that the contribution of both the strength of immune pressure as measured through avidity and immunodominance together with the ability to tolerate mutations determine whether the targeted epitopes will evolve or persist over time.

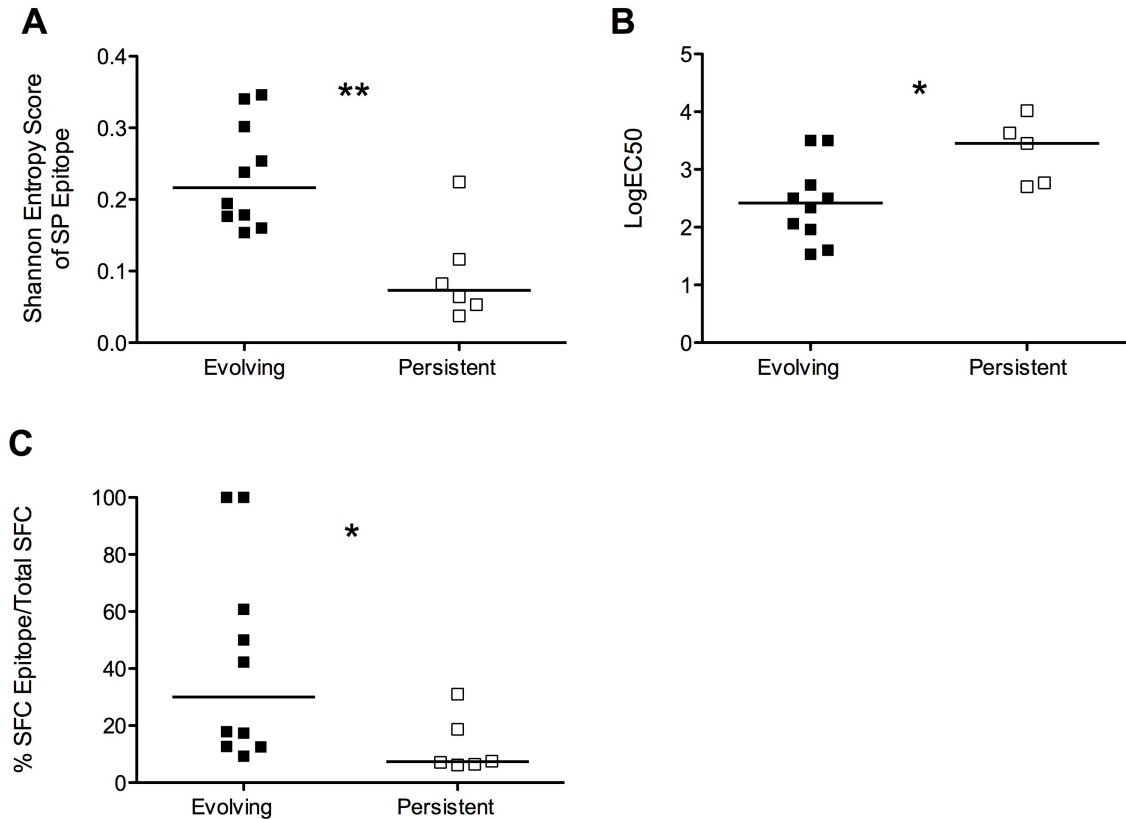


Figure 3-3: Shannon entropy, functional avidity, and immunodominance are associated with responses to evolving epitopes. A) Shannon entropy was calculated for each AA within the Gag protein. The average entropy across all AA within an epitope is shown for each evolving and persistent epitope and compared between groups (**: $p=0.004$). **B)** Functional avidity (EC_{50}) differences between responses to evolving and persistent epitopes (*: $p=0.03$.) **C)** Immunodominance is calculated as the magnitude of the response (SFC/M PBMC) to a particular epitope over the total magnitude of responses detected within the first year of infection. For each evolving or persistent epitope, the relative immunodominance is shown and compared between groups (*: $p=0.03$).

Highly dynamic T-cell escape processes observed during early Gag evolution

In the ten epitopes in which we identified subsequent viral evolution, we found limited overlap in the patterns of CTL escape, highlighting the remarkable plasticity in which CTL escape processes can occur. We observed five distinct patterns of CTL escape:

Cross-reactivity and eventual escape

In PIC64236, the founder sequence in the RK9 epitope ($_{20}$ RLRPGGKKK $_{28}$) began to decline at 257dps until it was undetectable in the viral population by 896dps (**Figure 3-4A**). During this time, multiple independent variants arose in the viral population after 146dps (K28T,

K28E, P23Q, K28R) and one variant (K28Q) became fixed by the end of the study period (**Figure 3-4A**). At 41dps, both the founder sequence and the K28R variant were recognized with similar avidity (**Figure 3-4B**). Peptides reflecting all other variants (K28T, K28E, P23Q) did not elicit a response, indicating that those substitutions conferred escape (**Figure 3-4B**).

In the PIC44149 viral population, the founder sequence in the DL9 epitope (₉₃DIKDTKEAL₁₀₁) began to decrease in frequency after 161dps but remained detectable until the end of follow up (350dps) (**Figure 3-4C**). Three additional variants (I94V, E99E, and D93A) were present in the population by 259dps. We detected responses of comparable avidity to the founder sequence and variants I94V and E99E but not to variant D93A (**Figure 3-4D**).

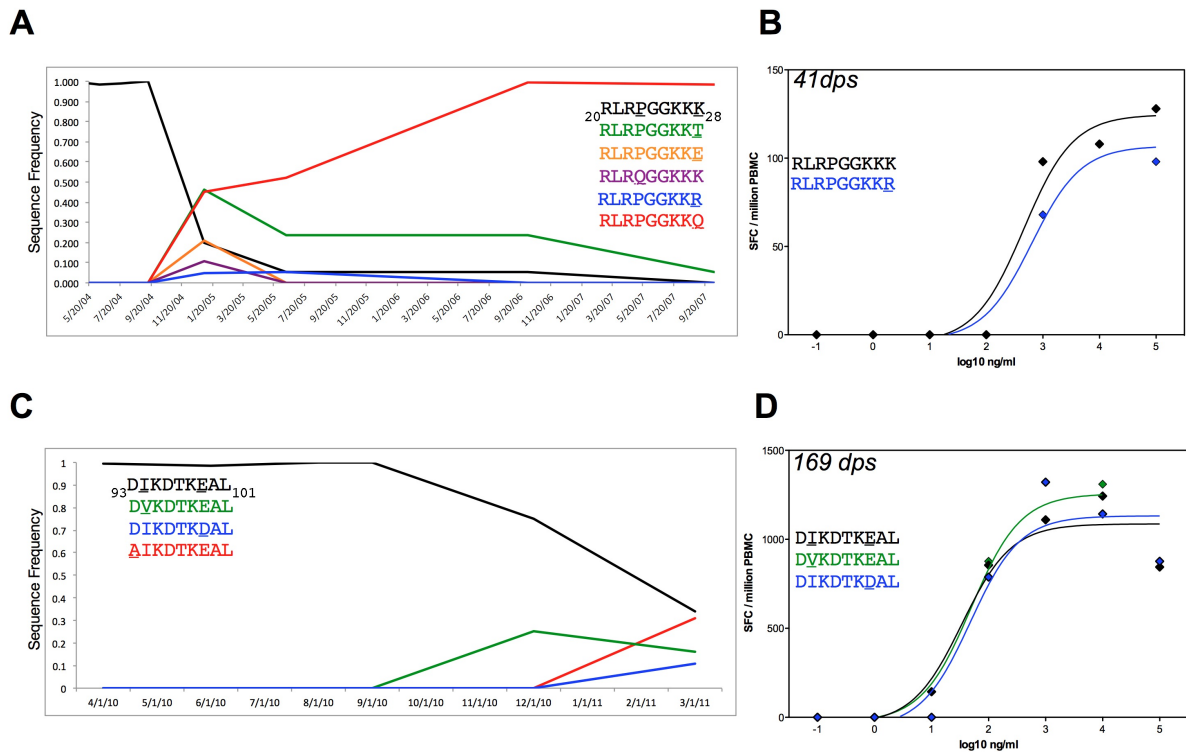


Figure 3-4: Cross-reactivity within targeted epitopes and selection of escape variants. A) Sequence frequency (y-axis) of variants at site 23 and 28 within the RK9 epitope over 1265dps of follow-up (x-axis) in PIC64236. **B)** Functional avidity titration curves at 41dps to K28 (EC₅₀ 2.06 ng/ml, black) and R28 (3.07 ng/ml, blue) in PIC64236. All variants in RK9 are colored coordinated between graphs A and B. **C)** Sequence frequency (y-axis) of variants at sites 94 and 99 within the DL9 epitope over 350dps of follow-up (x-axis) in PIC44149. **D)** Functional avidity titration curves at 169dps to I94 (1.53 ng/ml, black), V94 (1.69 ng/ml, green), and D99 (1.68 ng/ml, blue) in PIC44149. All

variants in DL9 are colored coordinated between graphs C and D. Underlined residues correspond to sites evolving in the SP over time.

Escape with epitope 'drift'

We observed a distinctive escape pattern in the evolution of the PIC64236 targeted QW9 epitope. The epitope founder sequence ($_{308}\text{QASQDVKNW}_{316}$) began to decline after 257dps and persisted as a minor variant until the end of follow up at 1265dps (**Figure 3-5A**). Between 257 and 428dps, a variant emerged with a single D312E mutation. After 428dps, we also detected D312E variants each with an additional unique mutation: S310T or N315S. By 1265dps, we also detected a minor variant with only a S310T mutation within the epitope (**Figure 3-5A**). We only identified a response targeting the founder sequence, indicating the epitope mutations observed all conferred escape (**Figure 3-5B**). It is currently unclear why after the rise of the initial D312E escape variant additional D312E-containing variants emerged in the viral population. Although the D312E marks a mutation to a higher database frequency AA, in the context of PIC64236 viral population, this mutation could infer a fitness cost in which the additional mutations were selected for as compensatory mutations.

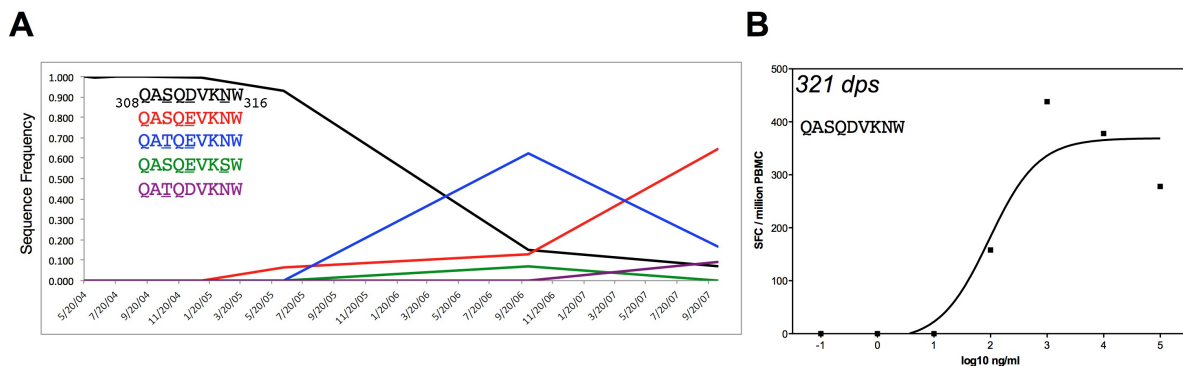


Figure 3-5: Escape within PIC64236 targeted epitope QW9. A) Sequence frequency (y-axis) of variants at sites 310, 312, and 315 within the QW9 epitope over 1265dps of follow-up (x-axis) in PIC64236. **B)** Functional avidity titration curve at 321dps to founder sequence (EC_{50} 1.96 ng/ml, black) in PIC64236. All variants in QW9 are colored coordinated between graphs A and B. Underlined residues correspond to sites evolving in the SP over time.

Sequential responses and escapes

De novo responses against mutations originally selected to confer escape from founder-specific viruses have been reported during acute infection. In PIC64236, we identified a unique development of *de novo* responses that offers insight into the broadening of Gag-specific responses over time. In the viral region corresponding to AA sites 75-99, the founder sequence (C87, K95) began to decline following 91dps until it was undetectable in the population by 896dps (**Figure 3-6A**). At 321dps, we detected a response to the LY12 epitope (₇₅LKSLYNTVAVLY₈₆) contained within this region (**Figure 3-6B**). Following 91dps, we observed the emergence of a transient variant (C87Y) that persisted until 896dps. As this mutation occurred one AA away from the targeted epitope, this mutation may have been selected as a processing escape mutation. Interestingly, the emergence of the C87Y variant resulted in a response to a distinct epitope, LI10 (₈₅LYYVHQRIE₉₃), which was also detected by 321dps (**Figure 3-6B**). Variant K95R emerged and became fixed in the population between 428 and 896dps, replacing both the founder and the C87Y variant. Again, the close proximity of the K95R mutation may disrupt peptide processing, potentially permitting escape from the LI10-specific response. However, the viral sequence containing the K95R variant retains the LY12 founder sequence and thus it is currently unclear if this epitope is again susceptible to the LY12-specific response. Potentially, the LY12-specific response may have waned over time, alleviating the selection pressure necessary to maintain the C87Y escape mutation.

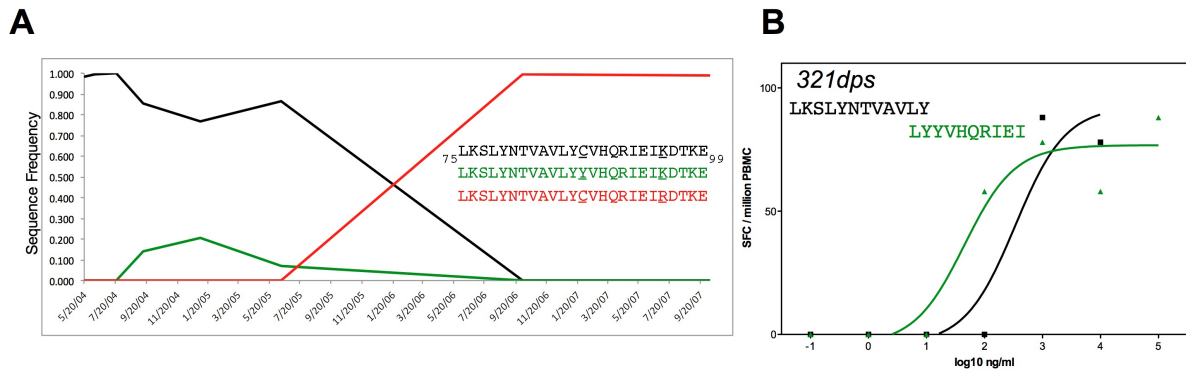


Figure 3-6: Escape mutations may promote responses to new epitopes in PIC64236. A) Sequence frequency (y-axis) of variants at sites 75 and 99 within the LY12 and LI10 epitopes over 1265dps of follow-up (x-axis) in PIC64236. B) Functional avidity titration curve at 321dps to LY12 (EC₅₀ 2.5 ng/ml, black) and LI10 (1.6 ng/ml, green) in PIC64236. All variants are colored coordinated between graphs A and B.

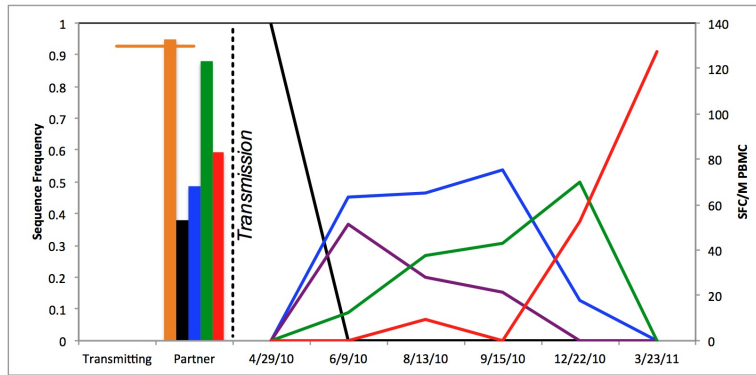
Simultaneous escape and reversion

In all transmission pairs, we identified only one epitope in which we found evidence for simultaneous escape and reversion processes. Around the time of transmission, TP PIC68415 elicited a response to 15mer peptides representing the viral region between sites 77 and 91, which contains the immunodominant A*02 SL9 epitope (77SLFNTIAVL85). Although we did not fine map the response due to limited sample availability, given that PIC68415 expresses HLA-A*02, it is highly likely that this epitope is being targeted. Variant I82/F86 is the major variant present in the PIC68415 population around the time of transmission (**Figure 7-7A**). While I82 is present in approximately 24% of sequences from the LANL DB, F86 is extremely rare and is found in <3% of deposited sequences. All tested 15mer variants, including those that arose following transmission, are recognized by PIC68415, indicating that both the I82 and V82 variants are recognized and that the variation occurring at site 86 does not disrupt recognition, potentially because it is occurring outside of the SL9 epitope (**Figure 7-7A**).

Interestingly, in SP PIC44149, the founder sequence that was inferred at 22dps (V82/F86) was not the major variant observed in the TP viral population (**Figure 7-7A**). It is unclear if the I82/F86 variant was transmitted and reverted quickly to include the more

commonly occurring V82 (present in 75% of database sequences) or if a minor variant containing V82/F86 was directly transmitted. Over the study period, we observed multiple mutations arising at position 86 (F86L, F86S, F86C) and eventual fixation of F86Y. Notably, Y86 is the most common AA in this position (present in 97% of HIVDB sequences) whereas all other observed mutations are also extremely rare. At 169dps, we detected a response by PIC44149 to the TV8 epitope (₈₁TVALLCV₈₇) in this region (**Figure 7-7B**). Although most variants were cross recognized, the highest avidity response was to the L86 variant, which arose early in the viral population before 63dps, suggesting that it was actually the L86 variant that primed the TV8-specific response (**Figure 7-7B**). Collectively, these results suggest that following a rapid selection for mutations to replace the F86 variant transmitted by the TP, one of these variants (L86) became the target of the SP host immune response. These dual selective pressures thus promoted a complicated mutually exclusive mutational process that eventually resulted in the fixation of the Y86 variant, which reflected the most frequently occurring AA at this site (reversion) and was also poorly recognized by the TV8-specific response (escape).

A



77SLFNTIAVLF₇₇CVHQR₉₁
 SLFNTVAVLF₇₇CVHQR₉₁
 SLFNTVAVLSCVHQR
 SLFNTVAVLL₇₇CVHQR
 SLFNTVAVLCCVHQR
 SLFNTVAVLY₇₇CVHQR

B

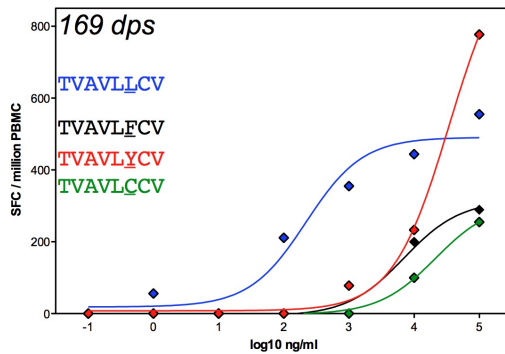


Figure 3-7: Dual selective pressures of reversion and escape within PIC44149. A) Sequence frequency of variant over time (line graph, left y-axis) with the magnitude of IFN γ response to 15mer peptides (bar graph, right y-axis) and days post onset of symptoms (x-axis). Sequence frequency and response magnitude left of the dotted lines refers to viral population and responses in TP PIC68415. All data right of dotted line is sequence evolution within the SP PIC44149 viral population. All graphs are color-coordinated to the specific variant that was tested. Founder sequence is in black and replacement variant (Y86) in red. **B)** Functional avidity titration curve at 169dps to L86 (EC₅₀ 2.34 ng/ml, blue), F86 (3.84 ng/ml, black), Y86 (4.53 ng/ml, red), and C86 (4.32 ng/ml, green) in PIC44149.

Cross-reactive responses

In three SPs, we detected cross-reactive responses to all viral variants that arose during the study period. In PIC68008, between 32dps and 207dps, the founder sequence to the TW10 epitope (₂₄₀TSTLQEQVGW₂₄₉) was first replaced by a G248E variant before eventual fixation of a T242N variant by 366dps (**Figure 3-8A**). We also detected an additional variant T242N/V247I present at low levels in the viral population at both 366dps and 1250dps (**Figure 3-8A**). The V247I mutation reflected an increase in database frequency from <2.5% to 97.5%. Although the TP PIC67505 did not target this region at the time of transmission, this low level variant may be

the beginning of a potential reversion at site 247. At 47dps, SP PIC68008 recognized all TW10 epitope variants (**Figure 3-8B**). While the founder sequence and the G248E variant were recognized with similar avidity, we detected a decrease in avidity upon recognition of peptides representing the T242N and T242N/V247I variants (**Figure 3-8B**).

In the PIC51861 targeted epitope KY12, we observed a replacement of founder sequence (₁₈KIRLRPGGKKKY₂₉) with a K28R variant by 532dps (**Figure 3-8C**). We also detected two transient variants each containing a R20W mutation (R20W and R20W/K28R). At 54dps, we detected the highest avidity response to the founder sequence and both the K28R and the R20W variants were recognized with diminished avidity (**Figure 3-8D**). No response was detected to the R20W/K28R variant.

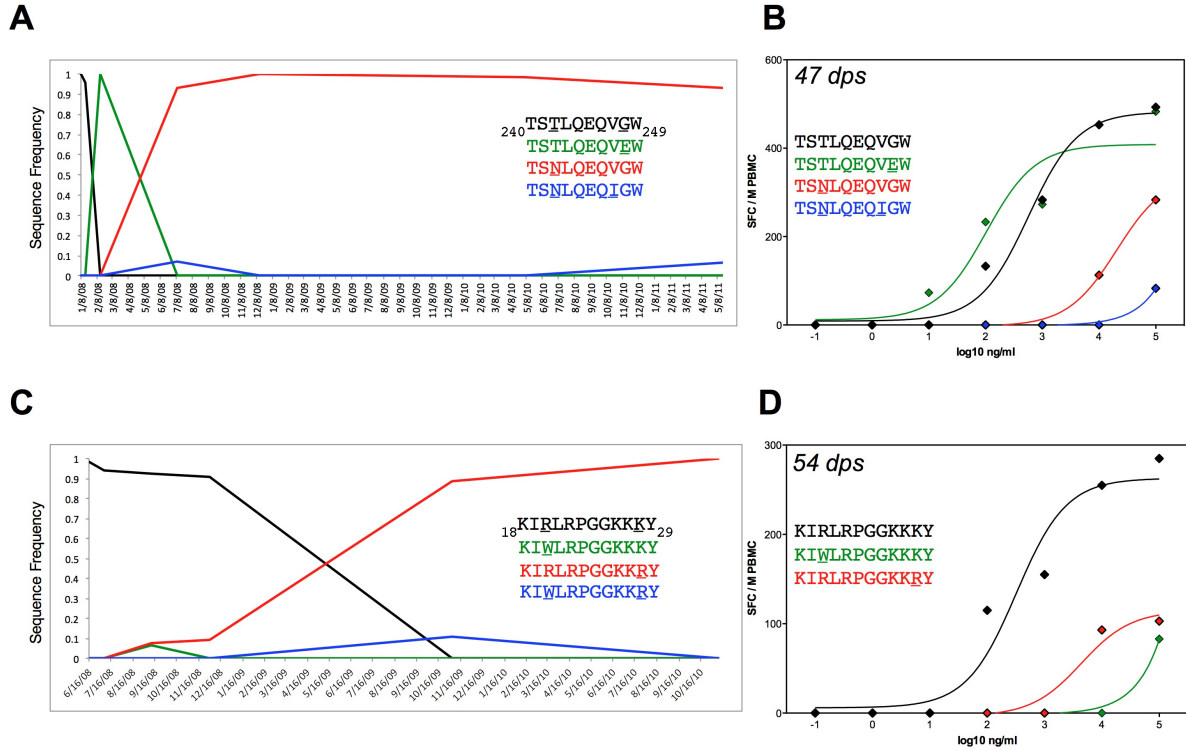


Figure 3-8: Cross-reactive responses detected in PIC68008 and PIC51861. A) Sequence frequency (y-axis) of variants at sites 242, 247, and 248 within the TW10 epitope over 1250dps of follow-up (x-axis) in PIC68008. **B)** Functional avidity titration curve at 47dps to founder sequence (EC_{50} 2.7 ng/ml, black), E248 (2.0 ng/ml, green), N242 (4.3 ng/ml, red) and N424/I247 (7.4 ng/ml, blue) in PIC68008. All variants in TW10 are colored coordinated between graphs A and B. **C)** Sequence frequency (y-axis) of variants at sites 20 and 28 within the KY12 epitope over 903dps of follow-up (x-axis) in PIC51861. **D)** Functional avidity titration curve at 54dps to founder sequence (2.5 ng/ml, black), W20 (7.4 ng/ml, green), and R28 (3.64 ng/ml, red) in PIC51861. All variants in KY12 are colored coordinated between graphs C and D.

In the viral population of PIC99203, at AA sites 133-135, we observed the sequence frequency of the founder (A146/L147) change over time until it was no longer detectable by 370dps and replaced with two variants (A146P and L147I) (**Figure 3-9A**). Upon fine mapping of the targeted epitopes, we identified responses to two distinct overlapping epitopes, GR11 (₁₄₀GQMVHQALSPR₁₅₀) and NL11 (₁₃₇NLQGQMVHQAL₁₄₇) (**Figure 3-9B**). In both epitopes, all variants were recognized with comparable avidity (**Figure 3-9C and D**). This variant recognition and replacement pattern is in contrast to what was observed in the cross-reactive responses in PIC68008 and PIC51861, in which the eventual replacement variant was recognized with lower avidity than the founder sequence.

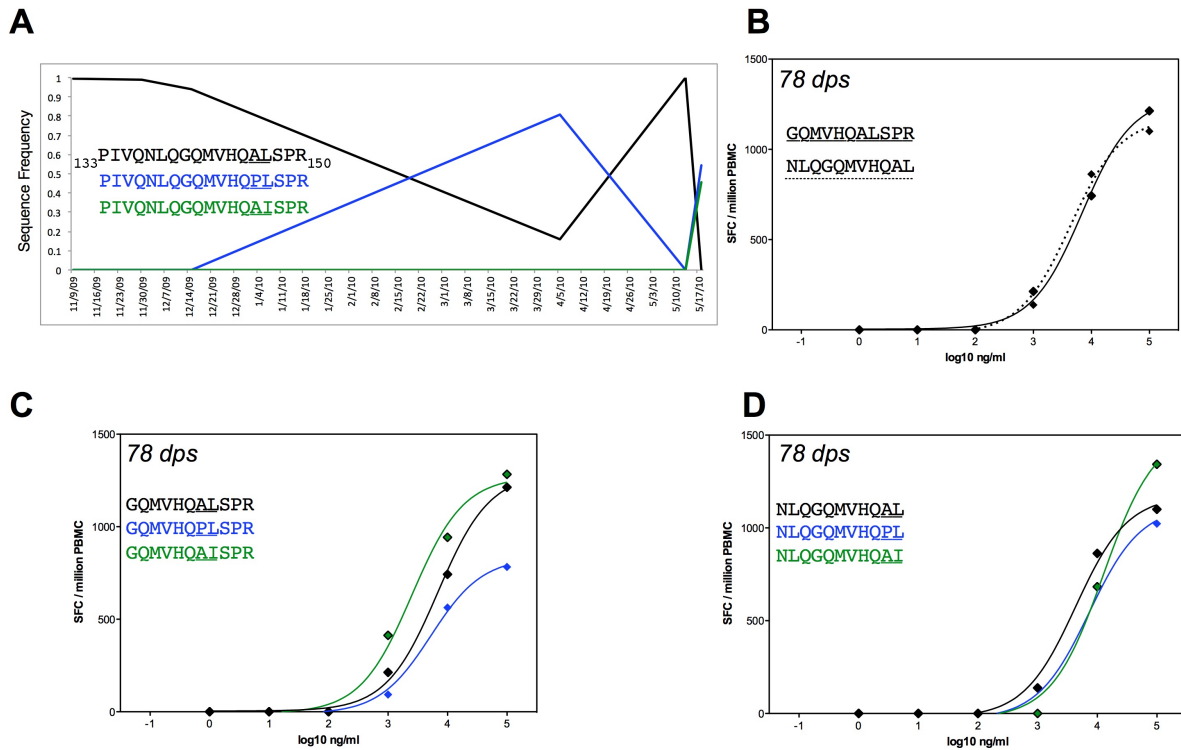


Figure 3-9: Cross-reactive responses to epitopes GR11 and NL11 in PIC99203. A) Sequence frequency (y-axis) of variants at sites 145 and 146 within viral region spanning AA sites 133-150 over 370dps of follow-up (x-axis) in PIC99203. **B)** Two optimal epitopes were identified within this area, GR11 and NL11. Functional avidity titration curve at 47dps to GR11 (EC_{50} 3.5 ng/ml, black solid) and NL11 (3.5 ng/ml, black dashed). **C)** Variant recognition within GR11. Functional avidity titration curve at 47dps to P145 (3.72 ng/ml, blue) and I146 (3.4 ng/ml, green). **D)** Variant recognition within NL11. Functional avidity titration curve at 47dps to P145 (3.86 ng/ml, blue) and I146 (4.1 ng/ml, green).

Replacement variants are often recognized with lower avidity

We identified some degree of cross-reactivity in seven of the ten evolving epitopes identified during the first year of HIV infection. Given the relationship between functional avidity and target cell elimination (Alexander-Miller et al., 1996; Derby et al., 2001), we set out to determine how functional avidity changes upon epitope variant recognition and if avidity alterations can predict the persistence of a particular variant. In the epitopes in which we detected a response to both the founder and the replacement variant, we found that responses to the replacement variant generally were of significantly lower avidity than that to the founder sequence ($p=0.03$, **Figure 3-10A**). In the epitopes in which there was a response to variants

that only existed transiently during the study period, we did not detect a significant difference in the functional avidity between the founder and transient variants ($p=0.25$, **Figure 3-10B**).

Notably, most of the responses to the transient variants were of similar decreased avidity to that seen with the replacement variants, suggesting that these transient variants did effectively permit immune escape and yet were not selected to eventual fixation.

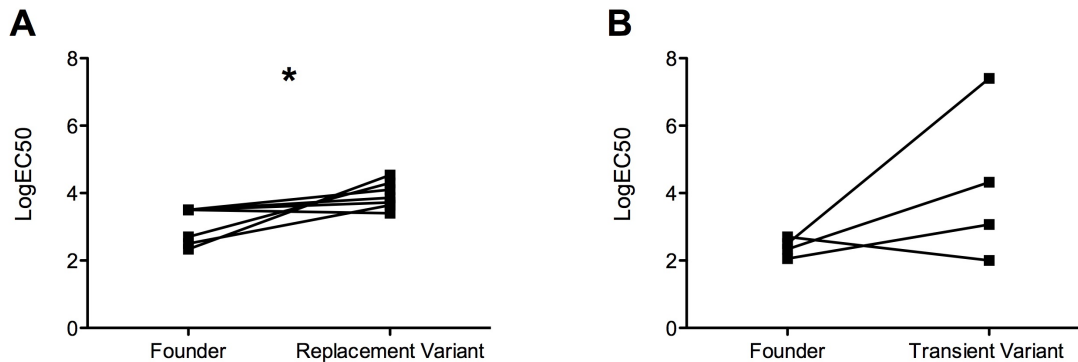


Figure 10: Functional avidity differences between responses to founder, replacement and transient variants. Responses in which there was cross-reactivity to either replacement variants (defined as variant(s) replacing the founder sequence by end of follow-up) or transient variants (defined as variants arising but then become undetectable by end of study period) were included in this analysis. A) Responses to replacement variants are of significantly lower avidity than responses to founder sequence (*: $p=0.03$). B) No significant difference of avidity between founder and transient variant-specific responses ($p=0.25$).

Fixation of replacement variants reflects fitness-balanced escape processes

Selection of a CTL escape mutation is believed to be the optimal result of a balance between the replicative fitness costs of a particular mutation with its ability to confer immune escape (Liu et al., 2007; Troyer et al., 2009). We hypothesized that while transient mutations observed in our study could permit escape, either through abolishing recognition or decreasing avidity, the fitness costs of these mutations prevented their fixation. Several studies have reported fitness costs associated with mutations that are infrequently found in the HIVDB (Prince et al., 2012; Rihn et al., 2013; Troyer et al., 2009). Therefore, in all the epitopes that we observed both transient sequence(s), and eventual replacement sequence(s) (RK9, TW10, TV8, and KY12), we compared the HIVDB sequence frequency of the founder sequence to the

HIVDB frequency of the variants. In this analysis, we included HIVDB frequencies of all transient and replacement variants regardless of whether or not the variant was recognized. Remarkably, we observed a consistent pattern across epitopes in which the transient variant(s) was typically of lower database frequency (median 0.09%) compared to the founder sequence (median 57.46%), and the eventual replacement variant reflected an increase in database frequency compared to the transient variant (median 11.94%) (**Figure 3-11**). The only case in which the replacement sequence was of higher database frequency than the founder sequence was in epitope TV8.

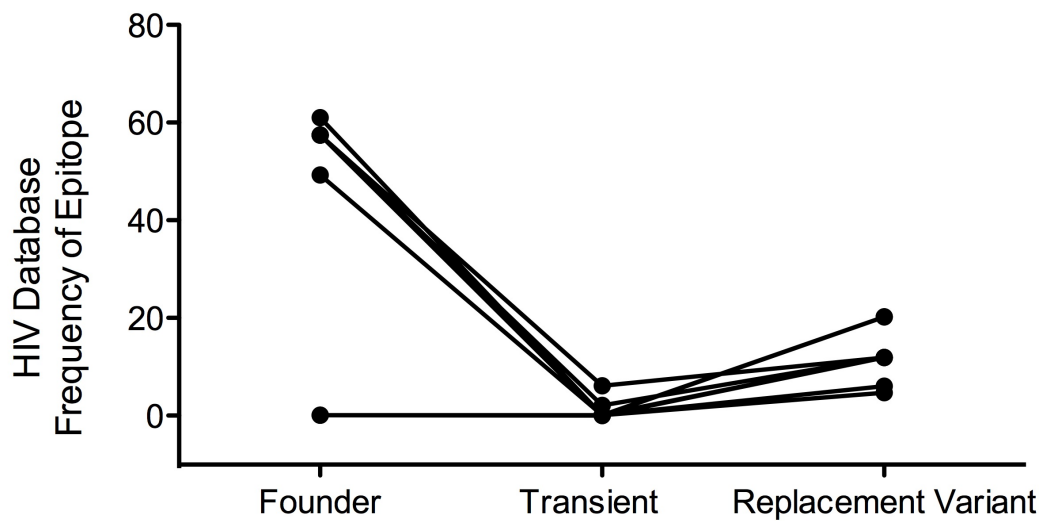


Figure 3-11: HIVDB frequency differences between founder, transient, and replacement variants. All epitopes in which we detected both transient and replacement variants were included in this analysis (RK9, TW10, TV8, and KY12). Each detected transient variant and their relationship to founder and replacement variant database frequency are represented in this graph.

To explore the interplay between immune escape and viral fitness, we performed *in vitro* pairwise fitness competitions between parental NL4-3 virus and NL4-3 viral variants with mutated sites corresponding to the detected variants observed in PIC64236 RK9 and PIC68008 TW10 epitope evolution (**Figure 3-4A and 3-8A**). These epitopes displayed interesting mutually

exclusive substitution patterns of transient variant(s) before eventual fixation of a replacement variant and therefore were prime candidates for further study. Pairwise fitness competitions were performed in PBMC cultures from a healthy donor and the relative fitness was calculated by measuring the proportion of parental NL4-3 versus mutant virus at different timepoints following *in vitro* infection (Lanxon-Cookson et al., 2013).

Despite vast differences in the frequencies in the database, we found little to no fitness costs in the different variants that arose during PIC64236 RK9 epitope evolution (**Figure 3-12A**). K28R and K28Q did not exact a fitness cost on the virus (**Figure 3-12A**). Compared to parental NL4-3, the relative fitness value was 0.95 for K28T and 0.97 for K28E. Although these were slight reductions, we did observe a trend towards these transient mutations having a higher fitness cost compared to eventual replacement variant K28Q (28T $p=0.06$, 28E $p=0.09$, students T-test). The only transient variants for which we did not observe a decrease in fitness (K28R in RK9) did not permit immune escape, as it was equally recognized by the founder-specific response in PIC64236 (**Figure 3-4A**). In fitness competitions between variants observed during PIC68008 TW10 evolution, we determined a relative fitness value for G248E to be 0.87 whereas T242N was 0.92 (**Figure 3-12B**). These results are consistent with previous reports describing the fitness costs associated with the T242N mutation (Brockman et al., 2007; Lanxon-Cookson et al., 2013). Again, we observed a marginal decrease in fitness of G248E compared to the replacement variant T242N ($p=0.05$). Interestingly, when we plotted the relative fitness values for each variant within an epitope, we observed a dip in fitness with the transient variants in 3 of 4 cases and a restoration of relative fitness with the replacement variant (**Figure 3-12C**). These results mirror the patterns, albeit less dramatically, to the observations made in the database frequencies differences.

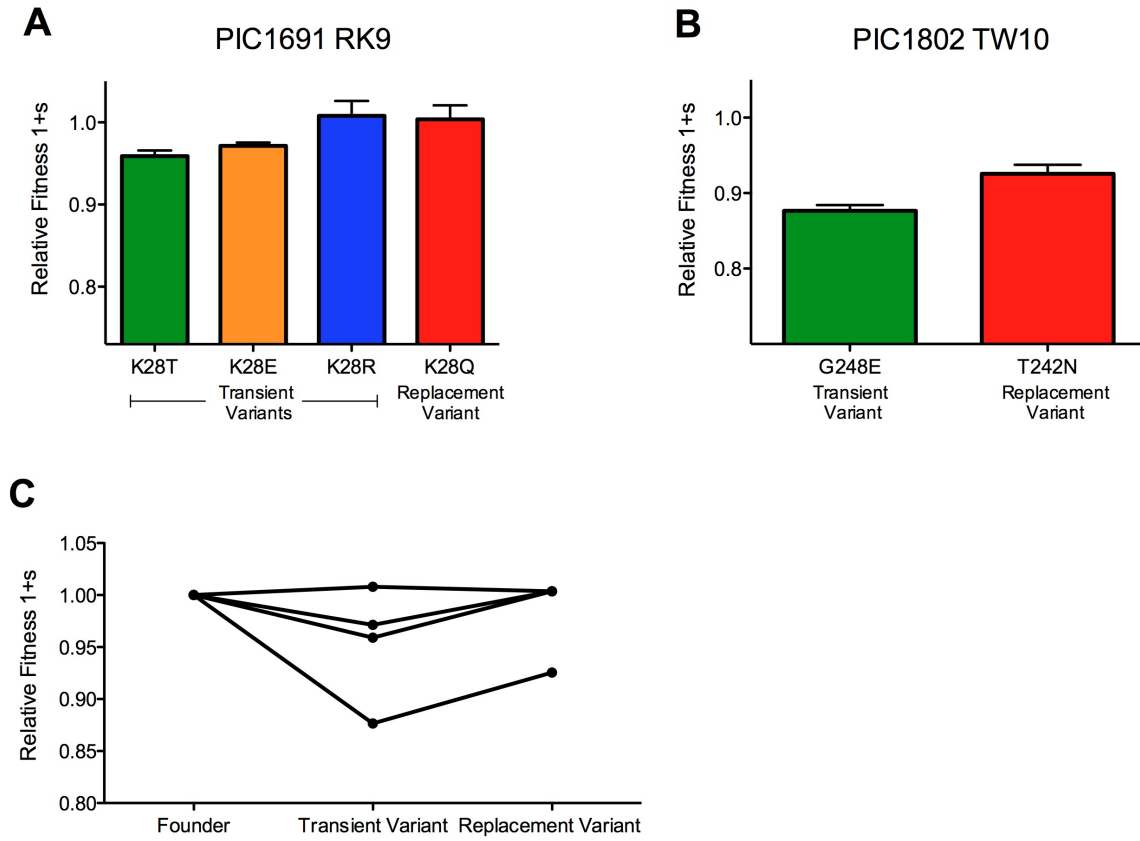


Figure 3-12: Pairwise fitness competitions with variants detected in PIC64236 RK9 and PIC68008 TW10 epitope evolution. Relative fitness of **A**) PIC64236 RK9 and **B**) PIC68008 TW10 evolution variants. All fitness competitions were performed in healthy PBMCs. Values shown are from three replicates. Error bars represent standard deviation. **C**) Relative fitness values stratified by founder, transient, and replacement variants from PIC64236 RK9 and PIC68008 TW10 evolution.

Discussion

In this study, we combined extensive deep sequencing of viral populations with detection of T-cell responses by autologous peptide screening in six linked HIV-transmission pairs in order to provide a unique perspective into CTL escape and reversion processes during acute HIV infection.

Reports have been variable on the extent of CTL reversions during primary HIV infection (Goonetilleke et al., 2009; Herbeck et al., 2011; Li et al., 2007; Liu et al., 2006). As our study included transmission pairs, we were able to screen TPs for T-cell responses at the time of

transmission, allowing us to determine potential reverting sites based on evolution in currently targeted TP epitopes within SP viral populations. We found a low frequency of potentially reverting sites, and the ongoing evolution at these sites did not reflect classical reversion patterns. Although we tested for responses in the TP with peptides representing the contemporaneous viral population at the time of transmission, it remains possible that we were unable to detect all the original Gag-specific responses, as responses can wane following fixation of escape mutations (Liu et al., 2011). Therefore, we also treated mutations to higher database frequencies as potential reversion mutations as this is the most commonly used methodology in determining reversions. With this analysis, the frequency of potential reversions remained relatively low compared to mutations towards a decrease in database frequency. Taken together, we found only slight evidence for reversions occurring in our six SPs during primary infection. However, several of the SPs had relatively low viral loads during early infection. As transmitted polymorphisms has been associated with lower viral loads (Goepfert et al., 2008; Matthews et al., 2008), the lack of reversions in these subjects may be contributing towards the maintenance of low viremia. Additional investigations in larger transmission pair cohorts are necessary to understand the importance of reversions in acute infection and whether this phenomenon can be harnessed in vaccine design.

CTL escape processes were often found to be highly dynamic with multiple potentially simultaneous mutations developing within targeted epitopes. We found that replacement variants often conferred escape from founder-specific T-cell responses by either abolishing recognition or by causing a decrease in functional avidity. This decrease in avidity may be a surrogate for a decrease in cytotoxic potential upon recognition, and as such these variants may persist in the viral population due to a weakened CTL-mediated selection pressure (Leslie et al., 2006). We observed that many of the transient variants detected during the study period also permitted escape but were not selected as the eventual replacement variant likely due to the

fitness costs associated with the mutations. Some of these transient mutations conferred only minor fitness costs as detected in our assay. However, we may be underestimating these costs due to the mutations being engineered into the lab strain NL4-3 backbone rather than within the subjects' autologous virus. Nonetheless, our data is suggestive that while some mutations may permit escape, even minor fitness defects can prevent the fixation of a particular variant. Collectively, these results indicate that eventual selection of a replacement variant and resolution of dynamic escape processes is the result of an optimal balance between immune escape and replicative fitness costs. Furthermore, the time to resolution of dynamic escape processes may be predictive of the fitness costs of the initial escape mutations. In PIC64236 RK9 evolution, it took 896dps for eventual fixation of a replacement variant, consistent with the fitness costs of the initial escape mutations (K28T, K28E) being low ($1+s=0.95$, 0.97 respectively). Conversely, in PIC68008 TW10 evolution, the higher fitness costs of G248E ($1+s=0.87$) may have resulted in more rapid fixation (366dps) of the replacement variant, T242N. Presumably, if multiple escape mutations arise within a targeted epitope and all have no impact on replicative fitness, these variants may be maintained in the virus population indefinitely.

The relationship between HIVDB frequency and replicative fitness costs of mutated HIV amino acids has been inconsistent. Although some conserved sites exhibit high fitness costs upon mutation, there appears to be no strong correlation between replicative fitness as measured *in vitro* and sequence conservation (Boutwell et al., 2013; Manochewa et al., 2013; Rolland M, 2013). In agreement, our study found that the measured changes in viral fitness between founder, transient, and replacement variants in dynamic escape processes were not proportional to the dramatic changes observed in the database frequency differences. However, we found that the trending directions of database frequency differences mirrored the changes found in the viral fitness assays. Therefore, although fitness costs and sequence frequency may not be tightly correlated, differential database frequencies between variants during dynamic

escape processes may predict fitness consequences of the mutations and help inform why one variant was selected over another.

Identifying regions within Gag that are especially vulnerable to a targeted CTL attack is critical for the design of new immunogens. In this study, we found strong immune pressure, as measured by functional avidity and immunodominance, on epitopes of high entropy during primary HIV infection, suggesting that these individuals were focusing early responses on viral regions with a high tolerance for variation and thus more pathways to achieve escape. While these regions are variable, they also appear to be highly immunogenic and thus it may be easier to induce responses to these regions via vaccination. Therefore, in order to make these immunodominant regions more susceptible to CTL-mediated control, cross-reactive responses of comparable functional avidity should be induced towards frequently occurring variants within the region (i.e., potential founder and replacement variants). Effective immune responses encompassing these variants may push the virus to select for rare variants, similar to what was observed with the transient variants, and the fitness costs associated with these mutations may contribute to control of viral replication.

Furthermore, we observed a significantly weaker immune pressure applied to epitopes that persisted throughout the study period. Interestingly, 5/6 of the persistent epitopes were located with the capsid domain of Gag (p24), which has been consistently reported to be the least tolerant of accumulating mutations compared to other Gag domains (Martinez-Picado et al., 2006; Schneidewind et al., 2008; Shin et al., 2007; Troyer et al., 2009). A recent elegant study exploring the toleration of mutation within p24 using a library of single AA substitutions demonstrated that 70% of mutations resulted in replication-defective viruses (Rihn et al., 2013). The failure to focus immune responses towards the more genetically fragile and vulnerable regions in p24 during early infection may have contributed to the higher viral loads in some of our study subjects (e.g., PIC64236). For vaccination purposes, in addition to inducing cross-

reactive responses in immunodominant regions, it may also be beneficial to induce multiple strong immune responses directed across the p24 domain. In cases of breakthrough infection post vaccination, having these potent p24-specific responses rapidly launched during early infection may exert a substantial selective pressure on the virus to select mutations that would cripple the virus replicative capacity (Janes et al., 2013).

In summary, we have shown that the principles of fitness-balanced immune escape guide the selection of mutations even in complicated dynamic escape processes. These results can be informative in the design of immunogens aimed to block escape pathways within immunogenic regions of the Gag protein.

Chapter IV:

Discussion

Summary of results and caveats to consider

Our investigations provided important insight into epitope variant recognition during natural HIV-1 infection. In chronically infected viremic individuals, we found that the median level of variant recognition across all targeted epitopic regions was correlated with viral load. While individuals with progressive infection maintained extensive epitope variant recognition, viruses harboring these recognized variants were frequently found in viral populations, suggesting that these variant-specific responses were ineffective in viral suppression. In elite controllers, we detected similar levels of epitope variant recognition to what was observed in progressors. As elite controllers have no detectable viremia, it is currently unclear what mechanisms are responsible for the high levels of variant recognition. Importantly, we found that increased overall sequence coverage, defined as the overall proportion of HIVDB sequences targeted through the Gag-specific repertoire, was inversely associated with viral load. Our results highlight that it is the ability to target the most frequently occurring variants, rather than simply a large number of variants, that is associated with control of viral replication. Furthermore, we show that sequence coverage, but not overall variant recognition, was associated with recognition of a greater proportion of founder viruses.

In chapter three, in addition to observing highly dynamic escape processes from founder-specific responses, we found evidence for epitope cross-reactivity in 70% of responses to evolving epitopes. Interestingly, recognized variants that persisted in the viral population were recognized with significantly lower functional avidity than responses to the founder sequence. These data suggest functional avidity as an immune parameter that can predict the effectiveness of the variant-specific response. Taken together, our results provide evidence for sequence coverage and functional avidity to serve as critical measures to estimate and then assess the protective potential of variant-inclusive vaccines currently entering phase I human clinical trials.

There are several important caveats to consider in the interpretation of the presented results. First, in both studies, our cohort size was relatively small and therefore it is possible that the conclusions reached in these studies may not directly translate to the wider HIV-1 infected population. However, given the prohibitive costs of generating variant-inclusive or autologous peptide sets and the massive PBMC quantities required for these types of assays, it is difficult to perform these experiments in large cohorts. Additionally, we restricted our studies to only testing for CTL responses directed against the Gag protein. Preferential targeting of the Gag protein has been associated with control of viral replication (Kiepiela et al., 2007; Rolland et al., 2008; Zuniga et al., 2006) and thus our results are informative for future vaccine design and evaluation. However, we may have over or under-estimated the effects of variant recognition and sequence coverage during natural HIV-1 infection due to our limited focus.

Another important caveat is that in our investigations presented in chapter 2, we calculated breadth, variant recognition, and sequence coverage based on targeted epitopic regions. Responses up to four sequential 11mer peptides were counted as a single epitopic region. In one scenario, a shared 8mer may actually be recognized within the epitopic region and thus our variant recognition and coverage calculations may be incorrect if the variation within the 8mer alone is different than the variation in the overall epitopic region. Conversely, if there are multiple epitopes being recognized within the epitopic region, then our calculation of breadth is underestimated. Given these conflicting considerations, we feel that using epitopic regions in our study is a conservative compromise. However, in order to be most accurate in the characterization of Gag-specific responses, the optimal epitope would have to be mapped for all responses. While we were unable to do this in this current study, future investigations may be able to fine-map a subset of individuals and determine how variant recognition and/or coverage may change between analysis of epitopic regions versus mapped epitopes.

Despite these caveats, these investigations have provided a unique insight into epitope variant recognition during acute and chronic HIV-1 infection. This discussion will focus on the implications of these results as they relate to current variant-inclusive vaccine strategies.

Sequence coverage as an evaluation tool of variant-inclusive vaccines

Several vaccination strategies designed to induce some level of epitope variant recognition are entering phase I clinical trials, where these vaccines will be tested for safety and immunogenicity for the first time in human subjects (Fischer et al., 2007; Rolland et al., 2007). Clinical testing of mosaic vaccines will include a trivalent HIV immunogen resembling intact HIV proteins (Santra et al., 2012). Presumably, these constructs will be degraded and processed similarly as in natural infection and thus induce T-cell and antibody responses. As it is a trivalent immunogen, there is a maximum of three amino acid variants at each position within a potential epitope (all nine amino acid fragments are considered potential epitopes) (Fischer et al., 2007; Santra et al., 2012). In the conserved element (CE) vaccine, the immunogen is designed to direct immune responses against conserved regions ranging from 8-20 amino acids in length, which are connected via amino acid linkers chosen to optimize natural peptide processing (Kulkarni et al., 2013; Niu et al., 2011; Rolland et al., 2007). The CE vaccine is a bivalent immunogen, in which the two constructs are identical except for the variation included at toggle sites. There is a maximum of two amino acids represented at each of these toggle sites (Rolland et al., 2007). Our results can help inform the evaluation and provide important considerations for both of these vaccines.

Epitope variant recognition through Gag-specific responses contributes to achieving a high level of sequence coverage. By design, variant-inclusive vaccines aim to induce T-cell responses capable of recognizing variation within epitopes. However, even a high level of variant recognition to a particular epitope will become obsolete if mutations are selected to disrupt epitope processing or HLA binding. One strategy to overcome this limitation is to induce

a high breadth of overlapping Gag-specific responses. Ideally, mutations that may disrupt epitope presentation in one epitope may still be targeted by a separate T-cell response encompassing an overlapping region. Although variant-inclusive vaccines have been shown to improve the breadth of T-cell responses compared to consensus or lab-strain immunogens in NHP studies (Barouch et al., 2010; Santra et al., 2010), it is unknown if these immunogens are capable of inducing responses to overlapping epitopes in human subjects. Identification of additional vaccine strategies that result in increased breadth of responses may be critical in improving the protective potential of variant-inclusive vaccines. Interestingly, several studies have demonstrated that artificially manipulating Gag to decrease protein stability results in increased shuttling of Gag into the MHC class I presentation pathway (Goldwich et al., 2008; Hahn et al., 2011). Recent work has shown that one of the Gag p6 late budding domain sequence motifs (PTAP) additionally functions to negatively regulate entry into the MHC class I pathway (Hahn et al., 2011). Mutation of PTAP results in increased Gag ubiquitination, leading to an increased rate of protein degradation and enhanced MHC class I presentation of Gag epitopes (Hahn et al., 2011). It is possible that increasing the processing of Gag protein can generate an abundance of diverse, overlapping Gag peptides that when presented by MHC class I, can lead to the induction of multiple CTL responses broadly directed at the Gag protein. As mosaic immunogens resemble natural proteins, mutation of the PTAP motif within the immunogen is feasible. Additionally, in immunogens with fragmented Gag regions, such as in the CE vaccine, fusing ubiquitin to the fragment itself may result in increased proteasome shuffling and increased breadth of responses (Benlahrech et al., 2012). Future experiments involving NHP vaccination studies of variant-inclusive immunogens should consider testing the impact of PTAP mutations and/or ubiquitination on vaccine-induced responses.

Our results suggest sequence coverage to be important for vaccine responses, if supported by larger studies, it will be of interest to determine what level of coverage is

necessary to induce in order to provide protection from acquisition or post-infection control. Currently, mosaic vaccines will present a maximum of three variants per potential epitope. Individual coverage of an epitope may be high or low, depending on the variability of the region. For example, targeting the three most common variants of the immunodominant epitope SL9 (SLYNTVATL) in variable p17 will provide only 54% coverage against all group M HIV-1 strains, whereas targeting of the relatively conserved p24 epitope variants in KK10 (KRWIILGLNK) will provide 92% coverage (http://www.hiv.lanl.gov/content/sequence/QUICK_ALIGN/QuickAlign.html). By design, CE vaccination can provide extremely high sequence coverage (>99%) of group M sequences (Rolland et al., 2007). Our results suggest that it is the collective sequence coverage achieved through all Gag-specific responses that is associated with control of viral replication. Both of these vaccination strategies have the potential to induce immune responses with considerable coverage. However, if mosaic vaccination preferentially induces responses against only the variable regions within Gag, or if the CE vaccine fails to direct durable immunodominant responses against the conserved elements, then the level of sequence coverage achieved may not be sufficient to provide protection.

Identification of the minimal threshold of sequence coverage that results in effective vaccine-induced responses will be valuable in the evaluation of variant-inclusive vaccines. Theoretically, this threshold may be different for goals of blocking establishment of infection or controlling viral replication post-acquisition. In our work, we found higher coverage of frequently occurring variants per epitopic region (80-100%) correlated with a higher proportion (80-90%) of founder viruses that were recognized, suggesting that fairly high coverage per epitope may be needed in order to block infection. In cases of breakthrough viral infections, it is possible that high coverage achieved by a Gag-specific CTL response may result in the selection of rare escape mutations within the targeted epitope. In our fitness investigations presented in chapter

three, we found that even rare mutations in global HIV sequences may not result in high replicative fitness costs. For example, in PIC64236, a K28T mutation arose in the p17 RK9 epitope (RLRPGGKKK) to confer escape. While this mutation is only found in 1.8% of group M HIV-1 sequences, it caused a minimal reduction in fitness cost compared to wild type virus (Figure 3-12) and persisted in the PIC64236 viral population for almost three years. Thus, in order to control breakthrough infections, protective responses may only be those in which high coverage can block off all possible escape pathways such that the virus cannot maintain adequate replicative capacity. As our chapter three studies highlight the extreme plasticity of the virus in response to CTL pressure, this may be a difficult goal to achieve.

To date, the use of the NHP model in determining a coverage threshold for protection with variant-inclusive vaccines has been limited. In published NHP studies, all of the mosaic constructs have been tested for immunogenicity rather than protection because mosaics are reflective of HIV sequence diversity and no currently used NHP model can be challenged with HIV strains. As breadth and depth of vaccine-induced responses have been previously mapped in these studies (Barouch et al., 2010; Santra et al., 2010; Santra et al., 2012), it would be simple to determine the level of sequence coverage (as defined in our work) that was achieved for each induced epitopic response via the vaccination strategy. However, in order to correlate the levels of induced coverage to vaccine efficacy in the NHP model, variant-inclusive vaccines reflective of SIV diversity would be needed. Fortunately, SIV mosaic constructs have been designed and are currently being tested in SIV challenge trials (Fischer et al., 2012). If the current strategies are found to result in insufficient coverage for protection or post-acquisition control, there may be several adjustments that can help improve induced coverage. First, as previously mentioned, immunogens can include mutations in PTAP or ubiquitination that may result in the breadth of responses (Benlahrech et al., 2012; Hahn et al., 2011), which may improve overall sequence coverage by expanding the Gag-specific repertoire. Additionally,

variant-inclusive vaccine strategies may require prime-boost regimens in which there are different variants represented in each component. The exposure to increased variation may synergize to achieve high levels of sequence coverage. Since the evaluation of variant-inclusive vaccines are just beginning, intensive further investigations of phase I clinical trials and future SIV/NHP challenge studies are necessary in order to gauge the optimal level of sequence coverage that may be necessary for protection.

Immune correlates of effective epitope variant recognition

Importantly, the functionality of the responding T cell upon epitope variation will also influence the ability to control through coverage. Our results suggest that ineffective T-cell responses result in recognized variants circulating in the viral population. These observations join a growing body of literature indicating peptide recognition of an epitope variant as measured by IFN γ secretion does not directly translate into the ability to inhibit replication of viruses harboring the recognized variant (Bennett et al., 2008; Valentine et al., 2008). Identification of an immune correlate of effective variant recognition will be useful in the evaluation of induced responses by variant-inclusive vaccination.

The functional avidity of the responding CTL has been correlated with increased variant recognition in both Hepatitis C and HIV-1 infections (Bennett et al., 2010; Mothe et al., 2012; Yerly et al., 2008). The mechanisms that connect functional avidity with enhanced variant recognition are currently unknown. One potential explanation is that high avidity T cells require strong interactions with a limited number of contact residues within the peptide rather than a collection of weaker interactions with the entire peptide-MHC class I complex (Kosmrlj et al., 2010). Because of the few contact residues (1-2 amino acids), the TCR can tolerate a greater number of epitope variations without disrupting recognition (Kosmrlj et al., 2010). However, even though correlations exist between avidity and variant recognition, often the functional avidity of the T cell will alter upon epitope variation (Goulder et al., 2001; Leslie et al., 2004;

Streeck et al., 2008; Turnbull et al., 2011). In our acute infection investigations, we found a number of cross-reactive responses to epitope variants as measured by IFN γ secretion. Upon titrating peptides representing the optimal epitope and epitope variants, we found that epitope variation from the founder sequence most often caused a decrease in CTL functional avidity. Low avidity T cells have been shown to demonstrate weaker cytotoxic potential compared to their high-avidity counterparts (Alexander-Miller et al., 1996). Interestingly, in early cross-reactive responses, we found that the variants that persisted in the viral population were recognized with a significantly lower avidity than responses to founder sequence.

Our results suggest that ineffective variant recognition may be a consequence of these variants being recognized with a lower avidity IFN γ response. As such, we propose variant recognition through only high avidity CTLs as a potential immune correlate of effective variant recognition. The criteria for defining a threshold for a “high avidity” response have not been established. Bennett *et al* proposed that responses with EC₅₀ values of 2 μ g/ml or less were associated with greater cytotoxic potential (Bennett et al., 2007). However, these conclusions were derived from investigating only A*02 restricted SL9 responses (Bennett et al., 2007). It is unknown how avidity thresholds may change from epitope to epitope or MHC class I allele to allele. In our studies, we observed that variants that were recognized with a significantly lower avidity were persistent in the viral population. Therefore, significant differences in functional avidity upon epitope variation may be more informative to determine response efficacy rather than identifying an avidity threshold applicable to all possible CTL responses.

Determining the functional avidity for vaccine-induced responses may be difficult in large-scale clinical trials. These assays first require determination of the optimal epitope (see chapter 3 methods) followed by titration of epitope variants over a wide range of peptide concentrations. In addition to the costs associated with generating a wealth of different peptides, testing responses to three variants in a single epitope requires up to six million PMBC. Epitope

mapping through the detection of IFN γ responses is widely performed in current HIV clinical trials and therefore completing these assays is feasible in terms of available equipment and technician expertise. However, in order to properly investigate differences in functional avidity, protocol design of the sample collections in variant-inclusive vaccine trials will need to ensure that enough PMBC samples are collected in order to accurately determine the avidity of the responses induced. Furthermore, since these assays are both cell and labor intensive, these analyses may need to be restricted to a subset of trial participants. Despite these challenges, we believe that understanding how the avidity of vaccine-induced responses changes upon epitope variation will be greatly informative in evaluating the protective potential of these variant-inclusive vaccine induced responses.

Although functional avidity is a promising immune correlate of effective variant recognition, other CTL functional qualities may also prove useful in predicting the efficacy of variant-specific responses. Polyfunctionality and proliferative capacity have both been associated with mediating inhibition of viral replication (Akinsiku et al., 2011; Julg et al., 2010; Migueles et al., 2002; Migueles et al., 2008). Our assay was limited to the detection of IFN γ (an antiviral cytokine) and IL-2 (cytokine associated with T-cell proliferation). Although few responses maintained IL-2 expression upon epitope variation, it is unclear if this directly translates into an inability of the T cell to proliferate upon variant recognition with our current data. Additional investigations are required in order to explore how polyfunctionality and proliferative capacity may alter upon variant recognition. One potentially informative experiment would be integrating intracellular cytokine staining (measurement of polyfunctionality and/or proliferative capacity) with an *in vitro* viral inhibition assay. In this experiment, one would first measure how different immunological parameters do or do not change upon stimulation of different epitope variants. These variant-specific CTLs could be sorted by epitope specificity and used as effectors in viral inhibition assay, in which a lab viral strain has previously infected

autologous CD4⁺ T cells. Importantly, viral strains would have to be engineered to reflect the different epitope variants that are recognized. By measuring the levels of virus in the presence of the different variant-specific effectors, one could extrapolate which functions need to be maintained in order to result in effective viral inhibition. While these experiments may be informative, they are also more cell and labor intensive than functional avidity assays and as such, may not be useful in evaluation of vaccine-induced responses.

Variant-inclusive vaccines in therapeutic vaccination

Since its inception in the mid 1990s, combined antiretroviral therapy (cART) has drastically improved the morbidity and mortality rates among HIV-infected individuals. However, these drugs are extremely expensive, require strict adherence requirements, and have varying degrees of side effects. Furthermore, it is difficult to ensure that each infected individual is directed to and maintained in care. Because of these challenges, there is a pressing need for alternative therapeutic strategies that control or eliminate viral replication. One proposed strategy is to use therapeutic vaccination in infected individuals in order to induce a “functional cure,” which is defined as control of viral replication without cART, similar to what is observed in elite controllers (Garcia et al., 2012). Ideally, therapeutic vaccines administered during cART would function to transform an individual’s HIV-specific immune response such that when treatment is ceased, the autologous immune system effectively controls viral replication (Garcia et al., 2012). As controllers are hypothesized to control viremia through HIV-specific CTL responses, many therapeutic vaccination strategies strive to induce highly effective CTLs (Garcia et al., 2012). Although an interesting concept, it is currently unclear what quality of CTL responses would be most effective to induce, a similar challenge in the development of prophylactic vaccines. To date, there have been several therapeutic vaccine clinical trials. While some products were successful in altering an individual’s HIV-specific CTL responses (i.e.,

inducing new CTL epitopes, increasing specificity to Gag), these studies either showed only minor reductions in viral load or no clinical benefit (Garcia et al., 2012).

To our knowledge, variant-inclusive vaccines have not been tested as therapeutic vaccination strategies. Interestingly, in our studies, we found that epitope variant recognition was reduced following initiation of cART. For example, in one individual we may detect responses against four out of five tested variants in an epitopic region (80% variant recognition) whereas following treatment, we are only able to detect responses to two out of five variants (40%). The FluoroSpot assay detects responses by quantifying the number of 'spots' that are formed upon antigen simulation, where one spot equals one spot-forming cell (SFC). Thus, the FluoroSpot measures the quantity of an antigen-specific response. Following suppression of viremia through cART, there is a contraction in the overall HIV-specific CTL repertoire (Gray et al., 1999; Ogg et al., 1999). This observation makes sense in light that many of the circulating HIV-specific CTL responses in an individual with high viremia are effector-like and thus need constant antigenic stimulation in order to proliferate and be maintained in the population (Liu et al., 2011; Shin et al., 2007). The removal of this stimulation can result in the demise of many effector CTLs and potentially the only remaining responses are from a pool of effector-memory or central-memory CTLs.

It is tempting to speculate that the majority of variant-specific responses that are associated with viremia during progressive infection are effector-like responses and that only a small proportion of variants are actually recognized through long-lived memory responses. This could have interesting implications for therapeutic vaccination in that variant-inclusive vaccines may be able to alter the variant recognition repertoire resulting in the majority of variant-specific responses having an effector-memory or central-memory phenotype. In prophylactic SIV vaccine studies, induced effector-memory (Barouch et al., 2012; Hansen et al., 2011; Yamamoto et al., 2012) and central memory (Barouch et al., 2012) CTL responses have been

associated with post-acquisition control of viral replication. Therefore, shifting variant-specific responses to effector-memory and/or central-memory CTLs via therapeutic vaccination may result in more effective inhibition of viruses harboring recognized variants compared what we observed with pre-ARV responses in progressive infection (Figure 2-5D).

While our current data is interesting in regard to the effects of viral suppression on variant recognition, more detailed investigations are necessary before variant-inclusive vaccines can be assessed for their potential in therapeutic vaccination. First, it is important to characterize the CTL phenotype pre and post ARV treatment via flow cytometry. This will allow for the determination if effector responses predominate during high viremia and then perish upon removal of antigenic stimulation, as well as examine if memory responses persist post-ARV. Additionally, it could be interesting to follow the clonotypic composition of variant-specific T cells within an epitope pre and post ARV. It remains to be determined if the observed variant recognition is the result of a single highly cross-reactive CTL, a collection of distinct clonotypes all recognizing a single sequence, or some combination of the two. Potentially, these different mechanisms of variant recognition may align with the phenotype of the responding CTL. For example, variant recognition through a polyclonal response may primarily be through effector responses whereas high cross-reactivity may be a feature of the memory response. Although pure speculation, it is interesting to note that elite controllers maintained levels of variant recognition comparable to that of progressors in the absence of high viremia and antigenic stimulation. It is possible that variant recognition in elite controllers may be a consequence of the Gag-specific repertoires being predominately cross-reactive memory responses. These investigations, coupled with ongoing prophylactic vaccine studies, can provide important insight into the role that variant-inclusive vaccines may play in future therapeutic vaccination strategies.

Concluding remarks

The road to the development of an efficacious HIV vaccine has been paved with inspiring successes, disappointing failures, and innovative forward thinking. Although a long journey remains ahead, there is optimism in the field that this goal will one day be achieved. In addition to the encouraging results from the RV144 clinical trial (Rerks-Ngarm et al., 2009), there are several new vaccine strategies that have proven to be very promising in the NHP model (Barouch et al., 2012; Hansen et al., 2011). Additionally, as highlighted throughout this dissertation, variant-inclusive vaccines have enormous potential to succeed where previous T-cell vaccine strategies have failed. The results presented here will be valuable in evaluating the immunogenicity of these vaccines as well as potentially informing vaccine alterations that may improve the induced immunity. It is an exciting time in the HIV vaccine field as vaccine-inclusive vaccines, along with other promising vaccine candidates, move through the development pipeline and may one day prove to be successful vaccine strategies. If the field continues to cultivate innovation and constantly strive for improvements, a preventative HIV vaccine will become a reality.

Acknowledgements

First and foremost, I would like to thank all study participants who donating their valuable time to make this research possible. Similarly, I would like to thank all clinic staff for their critical efforts and dedication, especially Julie Czartoski, Janine Maenza, and Anne Collier for their incredible work in maintaining unique Seattle area cohorts.

For the research presented in chapter two, I would like to thank Moon Kim for all sequencing efforts, David Heckerman for designing the Gag peptide set, Jonathan Carlson for thoughtful insight regarding sequence coverage analyses, and Stephen A. Migueles for providing valuable samples. I thank Paula Walsh and David Friedrich for technical support in setting up the FluoroSpot assay. I am most grateful to Julie McElrath, whose support allowed for these investigations to be conducted.

In my studies presented in chapter three, I would especially like to thank Brandon Maust, Brendan Larson and Ellie Casey for a wonderful collaboration on host-viral interactions during acute infection. I would like to thank Moon Kim, Lennie Chen, Dylan Westfall, Suvankar Ghorai, Hong Zhao, Kim Wong, and Wenjie Deng for their incredible work in generating all the 454 sequences for the PIC transmission pairs. Additionally, I would like to thank Vic Swain and Erinn Lanxon-Cookson for training in viral fitness competition assays.

I would like to thank my committee members, James I. Mullins, Michael Emerman, Donald L. Sodora, and Daniel Stetson for all the valuable insight and direction provided over my graduate career. I would like to thank Sheila Lukehart and everyone behind the STD/AIDS Research Training Fellowship Program for providing me with my pre-doctoral fellowship. Additionally, I want to especially thank James I. Mullins for serving as co-mentor on my training grant fellowship and providing me with expertise and direction in my virology investigations.

Finally, I would like to thank my mentor, Nicole Frahm. Over the past four years of graduate school, the vast support and amazing mentorship that I have received being a member in her lab has been truly incredible. I am very grateful to have had the amazing opportunities that have been provided for me under her mentorship.

References

- Abrahams, M.R., J.A. Anderson, E.E. Giorgi, C. Seoighe, K. Mlisana, L.H. Ping, G.S. Athreya, F.K. Treurnicht, B.F. Keele, N. Wood, J.F. Salazar-Gonzalez, T. Bhattacharya, H. Chu, I. Hoffman, S. Galvin, C. Mapanje, P. Kazembe, R. Thebus, S. Fiscus, W. Hide, M.S. Cohen, S.A. Karim, B.F. Haynes, G.M. Shaw, B.H. Hahn, B.T. Korber, R. Swanstrom, and C. Williamson. 2009. Quantitating the multiplicity of infection with human immunodeficiency virus type 1 subtype C reveals a non-poisson distribution of transmitted variants. *J Virol* 83:3556-3567.
- Addo, M.M., X.G. Yu, A. Rathod, D. Cohen, R.L. Eldridge, D. Strick, M.N. Johnston, C. Corcoran, A.G. Wurcel, C.A. Fitzpatrick, M.E. Feeney, W.R. Rodriguez, N. Basgoz, R. Draenert, D.R. Stone, C. Brander, P.J.R. Goulder, E.S. Rosenberg, M. Altfeld, and B.D. Walker. 2003. Comprehensive epitope analysis of HIV-1-specific T cell responses directed against the entire expressed HIV-1 genome demonstrate broadly directed responses, but no correlation to viral load. *J. Virol.* 77:2081-2092.
- Akinsiku, O.T., A. Bansal, S. Sabbaj, S.L. Heath, and P.A. Goepfert. 2011. Interleukin-2 production by polyfunctional HIV-1-specific CD8 T cells is associated with enhanced viral suppression. *J Acquir Immune Defic Syndr* 58:132-140.
- Alexander-Miller, M.A., G.R. Leggatt, and J.A. Berzofsky. 1996. Selective expansion of high- or low-avidity cytotoxic T lymphocytes and efficacy for adoptive immunotherapy. *Proceedings of the National Academy of Sciences of the United States of America* 93:4102-4107.
- Allen, T.M., M. Altfeld, S.C. Geer, E.T. Kalife, C. Moore, M. O'Sullivan K, I. Desouza, M.E. Feeney, R.L. Eldridge, E.L. Maier, D.E. Kaufmann, M.P. Lahaie, L. Reyor, G. Tanzi, M.N. Johnston, C. Brander, R. Draenert, J.K. Rockstroh, H. Jessen, E.S. Rosenberg, S.A. Mallal, and B.D. Walker. 2005a. Selective escape from CD8+ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution. *J Virol* 79:13239-13249.
- Allen, T.M., M. Altfeld, X.G. Yu, K.M. O'Sullivan, M. Lichterfeld, S. Le Gall, M. John, B.R. Mothe, P.K. Lee, E.T. Kalife, D.E. Cohen, K.A. Freedberg, D.A. Strick, M.N. Johnston, A. Sette, E.S. Rosenberg, S.A. Mallal, P.J. Goulder, C. Brander, and B.D. Walker. 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.
- Allen, T.M., X.G. Yu, E.T. Kalife, L.L. Reyor, M. Lichterfeld, M. John, M. Cheng, R.L. Allgaier, S. Mui, N. Frahm, G. Alter, N.V. Brown, M.N. Johnston, E.S. Rosenberg, S.A. Mallal, C. Brander, B.D. Walker, and M. Altfeld. 2005b. De novo generation of escape variant-specific CD8+ T-cell responses following cytotoxic T-lymphocyte escape in chronic human immunodeficiency virus type 1 infection. *J Virol* 79:12952-12960.
- Altfeld, M., M.M. Addo, E.S. Rosenberg, F.M. Hecht, P.K. Lee, M. Vogel, X.G. Yu, R. Draenert, M.N. Johnston, D. Strick, T.M. Allen, M.E. Feeney, J.O. Kahn, R.P. Sekaly, J.A. Levy, J.K. Rockstroh, P.J. Goulder, and B.D. Walker. 2003a.

- Influence of HLA-B57 on clinical presentation and viral control during acute HIV-1 infection. *AIDS* 17:2581-2591.
- Altfeld, M., M.M. Addo, R. Shankarappa, P.K. Lee, T.M. Allen, X.G. Yu, A. Rathod, J. Harlow, K. O'Sullivan, M.N. Johnston, P.J. Goulder, J.I. Mullins, E.S. Rosenberg, C. Brander, B. Korber, and B.D. Walker. 2003b. Enhanced detection of human immunodeficiency virus type 1-specific T-cell responses to highly variable regions by using peptides based on autologous virus sequences. *J. Virol.* 77:7330-7340.
- Ammaranond, P., D.J. van Bockel, K. Petoumenos, M. McMurchie, R. Finlayson, M.G. Middleton, M.P. Davenport, V. Venturi, K. Suzuki, L. Gelgor, J.M. Kaldor, D.A. Cooper, and A.D. Kelleher. 2011. HIV immune escape at an immunodominant epitope in HLA-B*27-positive individuals predicts viral load outcome. *J Immunol* 186:479-488.
- Appay, V., D.F. Nixon, S.M. Donahoe, G.M. Gillespie, T. Dong, A. King, G.S. Ogg, H.M. Spiegel, C. Conlon, C.A. Spina, D.V. Havlir, D.D. Richman, A. Waters, P. Easterbrook, A.J. McMichael, and S.L. Rowland-Jones. 2000. HIV-specific CD8(+) T cells produce antiviral cytokines but are impaired in cytolytic function. *The Journal of experimental medicine* 192:63-75.
- Barouch, D.H., J. Liu, H. Li, L.F. Maxfield, P. Abbink, D.M. Lynch, M.J. Iampietro, A. SanMiguel, M.S. Seaman, G. Ferrari, D.N. Forthal, I. Ourmanov, V.M. Hirsch, A. Carville, K.G. Mansfield, D. Stablein, M.G. Pau, H. Schuitemaker, J.C. Sadoff, E.A. Billings, M. Rao, M.L. Robb, J.H. Kim, M.A. Marovich, J. Goudsmit, and N.L. Michael. 2012. Vaccine protection against acquisition of neutralization-resistant SIV challenges in rhesus monkeys. *Nature* 482:89-93.
- Barouch, D.H., K.L. O'Brien, N.L. Simmons, S.L. King, P. Abbink, L.F. Maxfield, Y.H. Sun, A. La Porte, A.M. Riggs, D.M. Lynch, S.L. Clark, K. Backus, J.R. Perry, M.S. Seaman, A. Carville, K.G. Mansfield, J.J. Szinger, W. Fischer, M. Muldoon, and B. Korber. 2010. Mosaic HIV-1 vaccines expand the breadth and depth of cellular immune responses in rhesus monkeys. *Nat Med* 16:319-323.
- Barre-Sinoussi, F., J.C. Chermann, F. Rey, M.T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, and L. Montagnier. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220:868-871.
- Benlahrech, A., A. Meiser, S. Herath, T. Papagatsias, T. Athanasopoulos, F. Li, S. Self, V. Bachy, C. Hervouet, K. Logan, L. Klavinskis, G. Dickson, and S. Patterson. 2012. Fragmentation of SIV-gag vaccine induces broader T cell responses. *PloS one* 7:e48038.
- Bennett, M.S., A. Joseph, H.L. Ng, H. Goldstein, and O.O. Yang. 2010. Fine-tuning of T-cell receptor avidity to increase HIV epitope variant recognition by cytotoxic T lymphocytes. *AIDS* 24:2619-2628.
- Bennett, M.S., H.L. Ng, A. Ali, and O.O. Yang. 2008. Cross-clade detection of HIV-1-specific cytotoxic T lymphocytes does not reflect cross-clade antiviral activity. *The Journal of infectious diseases* 197:390-397.
- Bennett, M.S., H.L. Ng, M. Dagarag, A. Ali, and O.O. Yang. 2007. Epitope-dependent avidity thresholds for cytotoxic T-lymphocyte clearance of virus-infected cells. *J Virol* 81:4973-4980.

- Betts, M.R., M.C. Nason, S.M. West, S.C. De Rosa, S.A. Migueles, J. Abraham, M.M. Lederman, J.M. Benito, P.A. Goepfert, M. Connors, M. Roederer, and R.A. Koup. 2006. HIV nonprogressors preferentially maintain highly functional HIV-specific CD8⁺ T cells. *Blood* 107:4781-4789.
- Bieniasz, P.D. 2009. The cell biology of HIV-1 virion genesis. *Cell host & microbe* 5:550-558.
- Borrow, P. 2011. Innate immunity in acute HIV-1 infection. *Current opinion in HIV and AIDS* 6:353-363.
- Borrow, P., and G.M. Shaw. 1998. Cytotoxic T-lymphocyte escape viral variants: how important are they in viral evasion of immune clearance in vivo? *Immunological reviews* 164:37-51.
- Boutwell, C.L., J.M. Carlson, T.H. Lin, A. Seese, K.A. Power, J. Peng, Y. Tang, Z.L. Brumme, D. Heckerman, A. Schneidewind, and T.M. Allen. 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.
- Boutwell, C.L., C.F. Rowley, and M. Essex. 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.
- Brenchley, J.M., D.A. Price, T.W. Schacker, T.E. Asher, G. Silvestri, S. Rao, Z. Kazzaz, E. Bornstein, O. Lambotte, D. Altmann, B.R. Blazar, B. Rodriguez, L. Teixeira-Johnson, A. Landay, J.N. Martin, F.M. Hecht, L.J. Picker, M.M. Lederman, S.G. Deeks, and D.C. Douek. 2006. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat Med* 12:1365-1371.
- Briggs, J.A., and H.G. Krausslich. 2011. The molecular architecture of HIV. *Journal of molecular biology* 410:491-500.
- Brockman, M.A., A. Schneidewind, M. Lahaie, A. Schmidt, T. Miura, I. Desouza, F. Ryvkin, C.A. Derdeyn, S. Allen, E. Hunter, J. Mulenga, P.A. Goepfert, B.D. Walker, and T.M. Allen. 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.
- Bryant, M., and L. Ratner. 1990. Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proceedings of the National Academy of Sciences of the United States of America* 87:523-527.
- Buchbinder, S.P., D.V. Mehrotra, A. Duerr, D.W. Fitzgerald, R. Mogg, D. Li, P.B. Gilbert, J.R. Lama, M. Marmor, C. Del Rio, M.J. McElrath, D.R. Casimiro, K.M. Gottesdiener, J.A. Chodakewitz, L. Corey, and M.N. Robertson. 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*
- Bunnik, E.M., L. Pisas, A.C. van Nuenen, and H. Schuitemaker. 2008. Autologous neutralizing humoral immunity and evolution of the viral envelope in the course of subtype B human immunodeficiency virus type 1 infection. *J Virol* 82:7932-7941.
- Burton, D.R., R. Ahmed, D.H. Barouch, S.T. Butera, S. Crotty, A. Godzik, D.E. Kaufmann, M.J. McElrath, M.C. Nussenzweig, B. Pulendran, C.N. Scanlan, W.R. Schief, G. Silvestri, H. Streeck, B.D. Walker, L.M. Walker, A.B. Ward, I.A. Wilson,

- and R. Wyatt. 2012. A Blueprint for HIV Vaccine Discovery. *Cell host & microbe* 12:396-407.
- Cale, E.M., P. Hraber, E.E. Giorgi, W. Fischer, T. Bhattacharya, T. Leitner, W.W. Yeh, C. Gleasner, L.D. Green, C.S. Han, B. Korber, and N.L. Letvin. 2011. Epitope-specific CD8⁺ T lymphocytes cross-recognize mutant simian immunodeficiency virus (SIV) sequences but fail to contain very early evolution and eventual fixation of epitope escape mutations during SIV infection. *J Virol* 85:3746-3757.
- Cao, J., J. McNevin, S. Holte, L. Fink, L. Corey, and M.J. McElrath. 2003. Comprehensive analysis of human immunodeficiency virus type 1 (HIV-1)-specific gamma interferon-secreting CD8⁺ T cells in primary HIV-1 infection. *J Virol* 77:6867-6878.
- Carlson, J.M., C.J. Brumme, E. Martin, J. Listgarten, M.A. Brockman, A.Q. Le, C.K. Chui, L.A. Cotton, D.J. Knapp, S.A. Riddler, R. Haubrich, G. Nelson, N. Pfeifer, C.E. Deziel, D. Heckerman, R. Apps, M. Carrington, S. Mallal, P.R. Harrigan, M. John, and Z.L. Brumme. 2012. Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. *J Virol* 86:13202-13216.
- Carlson, J.M., Z.L. Brumme, C.M. Rousseau, C.J. Brumme, P. Matthews, C. Kadie, J.I. Mullins, B.D. Walker, P.R. Harrigan, P.J. Goulder, and D. Heckerman. 2008. Phylogenetic dependency networks: inferring patterns of CTL escape and codon covariation in HIV-1 Gag. *PLoS computational biology* 4:e1000225.
- Chen, H., Z.M. Ndhlovu, D. Liu, L.C. Porter, J.W. Fang, S. Darko, M.A. Brockman, T. Miura, Z.L. Brumme, A. Schneidewind, A. Piechocka-Trocha, K.T. Cesa, J. Sela, T.D. Cung, I. Toth, F. Pereyra, X.G. Yu, D.C. Douek, D.E. Kaufmann, T.M. Allen, and B.D. Walker. 2012. TCR clonotypes modulate the protective effect of HLA class I molecules in HIV-1 infection. *Nature immunology* 13:691-700.
- Chukkapalli, V., and A. Ono. 2011. Molecular determinants that regulate plasma membrane association of HIV-1 Gag. *Journal of molecular biology* 410:512-524.
- Curlin, M.E., R. Zioni, S.E. Hawes, Y. Liu, W. Deng, G.S. Gottlieb, T. Zhu, and J.I. Mullins. 2010. HIV-1 envelope subregion length variation during disease progression. *PLoS pathogens* 6:e1001228.
- Curtsinger, J.M., and M.F. Mescher. 2010. Inflammatory cytokines as a third signal for T cell activation. *Current opinion in immunology* 22:333-340.
- Dahirel, V., K. Shekhar, F. Pereyra, T. Miura, M. Artyomov, S. Talsania, T.M. Allen, M. Altfeld, M. Carrington, D.J. Irvine, B.D. Walker, and A.K. Chakraborty. 2011. Coordinate linkage of HIV evolution reveals regions of immunological vulnerability. *Proceedings of the National Academy of Sciences of the United States of America* 108:11530-11535.
- Day, C.L., D.E. Kaufmann, P. Kiepiela, J.A. Brown, E.S. Moodley, S. Reddy, E.W. Mackey, J.D. Miller, A.J. Leslie, C. DePierres, Z. Mncube, J. Duraiswamy, B. Zhu, Q. Eichbaum, M. Altfeld, E.J. Wherry, H.M. Coovadia, P.J. Goulder, P. Klenerman, R. Ahmed, G.J. Freeman, and B.D. Walker. 2006. PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. *Nature* 443:350-354.

- Deng, W., B.S. Maust, D.C. Nickle, G.H. Learn, Y. Liu, L. Heath, S.L. Kosakovsky Pond, and J.I. Mullins. 2010. DIVEIN: a web server to analyze phylogenies, sequence divergence, diversity, and informative sites. *BioTechniques* 48:405-408.
- Deng, W., B.S. Maust, D.H. Westfall, L. Chen, H. Zhao, B. Larsen, S. Iyer, Y. Liu, and J.I. Mullins. 2013. Indel and Carryforward Correction (ICC): A new analysis approach for processing 454 pyrosequencing data. *Bioinformatics*
- Derby, M., M. Alexander-Miller, R. Tse, and J. Berzofsky. 2001. High-avidity CTL exploit two complementary mechanisms to provide better protection against viral infection than low-avidity CTL. *J Immunol* 166:1690-1697.
- Donahue, D.A., and M.A. Wainberg. 2013. Cellular and molecular mechanisms involved in the establishment of HIV-1 latency. *Retrovirology* 10:11.
- Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* 32:1792-1797.
- Elahi, S., W.L. Dinges, N. Lejarcegui, K.J. Laing, A.C. Collier, D.M. Koelle, M.J. McElrath, and H. Horton. 2011. Protective HIV-specific CD8+ T cells evade Treg cell suppression. *Nat Med* 17:989-995.
- Estes, J. 2013. Early events following SIV mucosal challenge. In HIV Vaccines Keystone Symposia. Keystone, CO
- Estes, J.D., L.D. Harris, N.R. Klatt, B. Tabb, S. Pittaluga, M. Paiardini, G.R. Barclay, J. Smedley, R. Pung, K.M. Oliveira, V.M. Hirsch, G. Silvestri, D.C. Douek, C.J. Miller, A.T. Haase, J. Lifson, and J.M. Brenchley. 2010. Damaged intestinal epithelial integrity linked to microbial translocation in pathogenic simian immunodeficiency virus infections. *PLoS pathogens* 6:e1001052.
- Feeney, M.E., Y. Tang, K.A. Roosevelt, A.J. Leslie, K. McIntosh, N. Karthas, B.D. Walker, and P.J. Goulder. 2004. Immune escape precedes breakthrough human immunodeficiency virus type 1 viremia and broadening of the cytotoxic T-lymphocyte response in an HLA-B27-positive long-term-nonprogressing child. *J Virol* 78:8927-8930.
- Ferrari, G., B. Korber, N. Goonetilleke, M.K. Liu, E.L. Turnbull, J.F. Salazar-Gonzalez, N. Hawkins, S. Self, S. Watson, M.R. Betts, C. Gay, K. McGhee, P. Pellegrino, I. Williams, G.D. Tomaras, B.F. Haynes, C.M. Gray, P. Borrow, M. Roederer, A.J. McMichael, and K.J. Weinhold. 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 pathogens* 7:e1001273.
- Fischer, W., C. Apetrei, M.L. Santiago, Y. Li, R. Gautam, I. Pandrea, G.M. Shaw, B.H. Hahn, N.L. Letvin, G.J. Nabel, and B.T. Korber. 2012. Distinct evolutionary pressures underlie diversity in simian immunodeficiency virus and human immunodeficiency virus lineages. *J Virol* 86:13217-13231.
- Fischer, W., S. Perkins, J. Theiler, T. Bhattacharya, K. Yusim, R. Funkhouser, C. Kuiken, B. Haynes, N.L. Letvin, B.D. Walker, B.H. Hahn, and B.T. Korber. 2007. Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants. *Nat Med* 13:100-106.
- Frahm, N., S. Adams, P. Kiepiela, C.H. Linde, H.S. Hewitt, M. Lichterfeld, K. Sango, N.V. Brown, E. Pae, A.G. Wurcel, M. Altfeld, M.E. Feeney, T.M. Allen, T. Roach, M.A. St John, E.S. Daar, E. Rosenberg, B. Korber, F. Marincola, B.D. Walker, P.J. Goulder, and C. Brander. 2005. HLA-B63 presents HLA-B57/B58-restricted

- cytotoxic T-lymphocyte epitopes and is associated with low human immunodeficiency virus load. *J Virol* 79:10218-10225.
- Frahm, N., B.T. Korber, C.M. Adams, J.J. Szinger, R. Draenert, M.M. Addo, M.E. Feeney, K. Yusim, K. Sango, N.V. Brown, D. SenGupta, A. Piechocka-Trocha, T. Simonis, F.M. Marincola, A.G. Wurcel, D.R. Stone, C.J. Russell, P. Adolf, D. Cohen, T. Roach, A. StJohn, A. Khatiri, K. Davis, J. Mullins, P.J. Goulder, B.D. Walker, and C. Brander. 2004. Consistent cytotoxic-T-lymphocyte targeting of immunodominant regions in human immunodeficiency virus across multiple ethnicities. *J Virol* 78:2187-2200.
- Frahm, N., D.C. Nickle, C.H. Linde, D.E. Cohen, R. Zuniga, A. Lucchetti, T. Roach, B.D. Walker, T.M. Allen, B.T. Korber, J.I. Mullins, and C. Brander. 2008. Increased detection of HIV-specific T cell responses by combination of central sequences with comparable immunogenicity. *AIDS* 22:447-456.
- Friedrich, T.C., E.J. Dodds, L.J. Yant, L. Vojnov, R. Rudersdorf, C. Cullen, D.T. Evans, R.C. Desrosiers, B.R. Mothe, J. Sidney, A. Sette, K. Kunstman, S. Wolinsky, M. Piatak, J. Lifson, A.L. Hughes, N. Wilson, D.H. O'Connor, and D.I. Watkins. 2004. Reversion of CTL escape-variant immunodeficiency viruses in vivo. *Nat Med* 10:275-281.
- Ganusov, V.V., N. Goonetilleke, M.K. Liu, G. Ferrari, G.M. Shaw, A.J. McMichael, P. Borrow, B.T. Korber, and A.S. Perelson. 2011. Fitness costs and diversity of the cytotoxic T lymphocyte (CTL) response determine the rate of CTL escape during acute and chronic phases of HIV infection. *J Virol* 85:10518-10528.
- Garcia, F., A. Leon, J.M. Gatell, M. Plana, and T. Gallart. 2012. Therapeutic vaccines against HIV infection. *Human vaccines & immunotherapeutics* 8:569-581.
- Garrus, J.E., U.K. von Schwedler, O.W. Pornillos, S.G. Morham, K.H. Zavitz, H.E. Wang, D.A. Wettstein, K.M. Stray, M. Cote, R.L. Rich, D.G. Myszka, and W.I. Sundquist. 2001. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* 107:55-65.
- Gaschen, B., J. Taylor, K. Yusim, B. Foley, F. Gao, D. Lang, V. Novitsky, B. Haynes, B.H. Hahn, T. Bhattacharya, and B. Korber. 2002. Diversity considerations in HIV-1 vaccine selection. *Science* 296:2354-2360.
- Gillespie, G.M., R. Kaul, T. Dong, H.B. Yang, T. Rostron, J.J. Bwayo, P. Kiama, T. Peto, F.A. Plummer, A.J. McMichael, and S.L. Rowland-Jones. 2002. Cross-reactive cytotoxic T lymphocytes against a HIV-1 p24 epitope in slow progressors with B*57. *AIDS* 16:961-972.
- Giorgi, J.V., L.E. Hultin, J.A. McKeating, T.D. Johnson, B. Owens, L.P. Jacobson, R. Shih, J. Lewis, D.J. Wiley, J.P. Phair, S.M. Wolinsky, and R. Detels. 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. *The Journal of infectious diseases* 179:859-870.
- Goepfert, P.A., W. Lumm, P. Farmer, P. Matthews, A. Prendergast, J.M. Carlson, C.A. Derdeyn, J. Tang, R.A. Kaslow, A. Bansal, K. Yusim, D. Heckerman, J. Mulenga, S. Allen, P.J. Goulder, and E. Hunter. 2008. 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.

- Goldwich, A., S.S. Hahn, S. Schreiber, S. Meier, E. Kampgen, R. Wagner, M.B. Lutz, and U. Schubert. 2008. Targeting HIV-1 Gag into the defective ribosomal product pathway enhances MHC class I antigen presentation and CD8⁺ T cell activation. *J Immunol* 180:372-382.
- Goonetilleke, N., M.K. Liu, J.F. Salazar-Gonzalez, G. Ferrari, E. Giorgi, V.V. Ganusov, B.F. Keele, G.H. Learn, E.L. Turnbull, M.G. Salazar, K.J. Weinhold, S. Moore, N. Letvin, B.F. Haynes, M.S. Cohen, P. Hraber, T. Bhattacharya, P. Borrow, A.S. Perelson, B.H. Hahn, G.M. Shaw, B.T. Korber, and A.J. McMichael. 2009. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. *The Journal of experimental medicine* 206:1253-1272.
- Gordon, S.N., N.R. Klatt, S.E. Bosinger, J.M. Brenchley, J.M. Milush, J.C. Engram, R.M. Dunham, M. Paiardini, S. Klucking, A. Danesh, E.A. Strobert, C. Apetrei, I.V. Pandrea, D. Kelvin, D.C. Douek, S.I. Staprans, D.L. Sodora, and G. Silvestri. 2007. Severe depletion of mucosal CD4⁺ T cells in AIDS-free simian immunodeficiency virus-infected sooty mangabeys. *J Immunol* 179:3026-3034.
- Goulder, P.J., C. Brander, Y. Tang, C. Tremblay, R.A. Colbert, M.M. Addo, E.S. Rosenberg, T. Nguyen, R. Allen, A. Trocha, M. Altfeld, S. He, M. Bunce, R. Funkhouser, S.I. Pelton, S.K. Burchett, K. McIntosh, B.T. Korber, and B.D. Walker. 2001. Evolution and transmission of stable CTL escape mutations in HIV infection. *Nature* 412:334-338.
- Goulder, P.J., R.E. Phillips, R.A. Colbert, S. McAdam, G. Ogg, M.A. Nowak, P. Giangrande, G. Luzzi, B. Morgan, A. Edwards, A.J. McMichael, and S. Rowland-Jones. 1997. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nat Med* 3:212-217.
- Gray, C.M., J. Lawrence, J.M. Schapiro, J.D. Altman, M.A. Winters, M. Crompton, M. Loi, S.K. Kundu, M.M. Davis, and T.C. Merigan. 1999. Frequency of class I HLA-restricted anti-HIV CD8⁺ T cells in individuals receiving highly active antiretroviral therapy (HAART). *J Immunol* 162:1780-1788.
- Gray, E.S., P.L. Moore, I.A. Choge, J.M. Decker, F. Bibollet-Ruche, H. Li, N. Leseka, F. Treurnicht, K. Mlisana, G.M. Shaw, S.S. Karim, C. Williamson, and L. Morris. 2007. Neutralizing antibody responses in acute human immunodeficiency virus type 1 subtype C infection. *J Virol* 81:6187-6196.
- Gray, G.E., M. Allen, Z. Moodie, G. Churchyard, L.G. Bekker, M. Nchabeleng, K. Mlisana, B. Metch, G. de Bruyn, M.H. Latka, S. Roux, M. Mathebula, N. Naicker, C. Ducar, D.K. Carter, A. Puren, N. Eaton, M.J. McElrath, M. Robertson, L. Corey, and J.G. Kublin. 2011. Safety and efficacy of the HVTN 503/Phambili study of a clade-B-based HIV-1 vaccine in South Africa: a double-blind, randomised, placebo-controlled test-of-concept phase 2b study. *The Lancet infectious diseases* 11:507-515.
- Haffar, O.K., S. Popov, L. Dubrovsky, I. Agostini, H. Tang, T. Pushkarsky, S.G. Nadler, and M. Bukrinsky. 2000. Two nuclear localization signals in the HIV-1 matrix protein regulate nuclear import of the HIV-1 pre-integration complex. *Journal of molecular biology* 299:359-368.

- Hahn, S., C. Setz, J. Wild, and U. Schubert. 2011. The PTAP sequence within the p6 domain of human immunodeficiency virus type 1 Gag regulates its ubiquitination and MHC class I antigen presentation. *J Immunol* 186:5706-5718.
- Hansen, S.G., J.C. Ford, M.S. Lewis, A.B. Ventura, C.M. Hughes, L. Coyne-Johnson, N. Whizin, K. Oswald, R. Shoemaker, T. Swanson, A.W. Legasse, M.J. Chiuchiolo, C.L. Parks, M.K. Axthelm, J.A. Nelson, M.A. Jarvis, M. Piatak, Jr., J.D. Lifson, and L.J. Picker. 2011. Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine. *Nature* 473:523-527.
- Health, N.I.o. 2013. NIH Discontinues Immunizations in HIV Vaccine Study. *NIH News*
- Henn, M.R., C.L. Boutwell, P. Charlebois, N.J. Lennon, K.A. Power, A.R. Macalalad, A.M. Berlin, C.M. Malboeuf, E.M. Ryan, S. Gnerre, M.C. Zody, R.L. Erlich, L.M. Green, A. Berical, Y. Wang, M. Casali, H. Streeck, A.K. Bloom, T. Dudek, D. Tully, R. Newman, K.L. Axten, A.D. Gladden, L. Battis, M. Kemper, Q. Zeng, T.P. Shea, S. Gujja, C. Zedlack, O. Gasser, C. Brander, C. Hess, H.F. Gunthard, Z.L. Brumme, C.J. Brumme, S. Bazner, J. Rychert, J.P. Tinsley, K.H. Mayer, E. Rosenberg, F. Pereyra, J.Z. Levin, S.K. Young, H. Jessen, M. Altfeld, B.W. Birren, B.D. Walker, and T.M. Allen. 2012. Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. *PLoS pathogens* 8:e1002529.
- Herbeck, J.T., M. Rolland, Y. Liu, S. McLaughlin, J. McNevin, H. Zhao, K. Wong, J.N. Stoddard, D. Raugi, S. Sorensen, I. Genowati, B. Birditt, A. McKay, K. Diem, B.S. Maust, W. Deng, A.C. Collier, J.D. Stekler, M.J. McElrath, and J.I. Mullins. 2011. Demographic processes affect HIV-1 evolution in primary infection before the onset of selective processes. *J Virol* 85:7523-7534.
- Hoof, I., C.L. Perez, M. Buggert, R.K. Gustafsson, M. Nielsen, O. Lund, and A.C. Karlsson. 2010. Interdisciplinary analysis of HIV-specific CD8+ T cell responses against variant epitopes reveals restricted TCR promiscuity. *J Immunol* 184:5383-5391.
- Horton, H., I. Frank, R. Baydo, E. Jalbert, J. Penn, S. Wilson, J.P. McNevin, M.D. McSweyn, D. Lee, Y. Huang, S.C. De Rosa, and M.J. McElrath. 2006. Preservation of T cell proliferation restricted by protective HLA alleles is critical for immune control of HIV-1 infection. *J Immunol* 177:7406-7415.
- Janbazian, L., D.A. Price, G. Canderan, A. Filali-Mouhim, T.E. Asher, D.R. Ambrozak, P. Scheinberg, M.R. Boulassel, J.P. Routy, R.A. Koup, D.C. Douek, R.P. Sekaly, and L. Trautmann. 2012. Clonotype and repertoire changes drive the functional improvement of HIV-specific CD8 T cell populations under conditions of limited antigenic stimulation. *J Immunol* 188:1156-1167.
- Janes, H., D.P. Friedrich, A. Krambrink, R.J. Smith, E.G. Kallas, H. Horton, D.R. Casimiro, M. Carrington, D.E. Geraghty, P.B. Gilbert, M.J. McElrath, and N. Frahm. 2013. Vaccine-Induced Gag-Specific T Cells Are Associated With Reduced Viremia After HIV-1 Infection. *The Journal of infectious diseases*
- Janeway, C.T., P. Walport, M. Shlomchik, M. 2005a. Antigen Presentation to T Lymphocytes. In Immunobiology. Garland Science, New York, New York.
- Janeway, C.T., P. Walport, M. Shlomchik, M. 2005b. T-Cell Mediated Immunity. In Immunobiology. Garland Science, New York, New York.

- Jensen, M.A., F.S. Li, A.B. van 't Wout, D.C. Nickle, D. Shriner, H.X. He, S. McLaughlin, R. Shankarappa, J.B. Margolick, and J.I. Mullins. 2003. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences. *J Virol* 77:13376-13388.
- Julg, B., K.L. Williams, S. Reddy, K. Bishop, Y. Qi, M. Carrington, P.J. Goulder, T. Ndung'u, and B.D. Walker. 2010. Enhanced anti-HIV functional activity associated with Gag-specific CD8 T-cell responses. *J Virol* 84:5540-5549.
- Kaech, S.M., E.J. Wherry, and R. Ahmed. 2002. Effector and memory T-cell differentiation: implications for vaccine development. *Nature reviews. Immunology* 2:251-262.
- Kaslow, R.A., M. Carrington, R. Apple, L. Park, A. Munoz, A.J. Saah, J.J. Goedert, C. Winkler, S.J. O'Brien, C. Rinaldo, R. Detels, W. Blattner, J. Phair, H. Erlich, and D.L. Mann. 1996. Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. *Nat Med* 2:405-411.
- Keele, B.F., E.E. Giorgi, J.F. Salazar-Gonzalez, J.M. Decker, K.T. Pham, M.G. Salazar, C. Sun, T. Grayson, S. Wang, H. Li, X. Wei, C. Jiang, J.L. Kirchherr, F. Gao, J.A. Anderson, L.H. Ping, R. Swanstrom, G.D. Tomaras, W.A. Blattner, P.A. Goepfert, J.M. Kilby, M.S. Saag, E.L. Delwart, M.P. Busch, M.S. Cohen, D.C. Montefiori, B.F. Haynes, B. Gaschen, G.S. Athreya, H.Y. Lee, N. Wood, C. Seoighe, A.S. Perelson, T. Bhattacharya, B.T. Korber, B.H. Hahn, and G.M. Shaw. 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proceedings of the National Academy of Sciences of the United States of America* 105:7552-7557.
- Kiepiela, P., K. Ngumbela, C. Thobakgale, D. Ramduth, I. Honeyborne, E. Moodley, S. Reddy, C. de Pierres, Z. Mncube, N. Mkhwanazi, K. Bishop, M. van der Stok, K. Nair, N. Khan, H. Crawford, R. Payne, A. Leslie, J. Prado, A. Prendergast, J. Frater, N. McCarthy, C. Brander, G.H. Learn, D. Nickle, C. Rousseau, H. Coovadia, J.I. Mullins, D. Heckerman, B.D. Walker, and P. Goulder. 2007. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nat Med* 13:46-53.
- Klatt, N.R., N.T. Funderburg, and J.M. Brenchley. 2013. Microbial translocation, immune activation, and HIV disease. *Trends in microbiology* 21:6-13.
- Korber, B., B. Gaschen, K. Yusim, R. Thakallapally, C. Kesmir, and V. Detours. 2001. Evolutionary and immunological implications of contemporary HIV-1 variation. *British medical bulletin* 58:19-42.
- Korber, B.T., N.L. Letvin, and B.F. Haynes. 2009. T-cell vaccine strategies for human immunodeficiency virus, the virus with a thousand faces. *J Virol* 83:8300-8314.
- Kosmrlj, A., E.L. Read, Y. Qi, T.M. Allen, M. Altfeld, S.G. Deeks, F. Pereyra, M. Carrington, B.D. Walker, and A.K. Chakraborty. 2010. Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. *Nature* 465:350-354.
- Koup, R.A., J.T. Safrit, Y. Cao, C.A. Andrews, G. McLeod, W. Borkowsky, C. Farthing, and D.D. Ho. 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.

- Kulkarni, V., M. Rosati, A. Valentin, B. Ganneru, A.K. Singh, J. Yan, M. Rolland, C. Alicea, R.K. Beach, G.M. Zhang, S. Le Gall, K.E. Broderick, N.Y. Sardesai, D. Heckerman, B. Mothe, C. Brander, D.B. Weiner, J.I. Mullins, G.N. Pavlakis, and B.K. Felber. 2013. HIV-1 p24(gag) derived conserved element DNA vaccine increases the breadth of immune response in mice. *PloS one* 8:e60245.
- Kuller, L.H., R. Tracy, W. Belloso, S. De Wit, F. Drummond, H.C. Lane, B. Ledergerber, J. Lundgren, J. Neuhaus, D. Nixon, N.I. Paton, and J.D. Neaton. 2008. Inflammatory and coagulation biomarkers and mortality in patients with HIV infection. *PLoS medicine* 5:e203.
- Ladell, K., M. Hashimoto, M.C. Iglesias, P.G. Wilmann, J.E. McLaren, S. Gras, T. Chikata, N. Kuse, S. Fastenackels, E. Gostick, J.S. Bridgeman, V. Venturi, Z.A. Arkoub, H. Agut, D.J. van Bockel, J.R. Almeida, D.C. Douek, L. Meyer, A. Venet, M. Takiguchi, J. Rossjohn, D.A. Price, and V. Appay. 2013. A molecular basis for the control of preimmune escape variants by HIV-specific CD8+ T cells. *Immunity* 38:425-436.
- Lanxon-Cookson, E.C., J.V. Swain, S. Manochewa, R.A. Smith, B. Maust, M. Kim, D. Westfall, M. Rolland, and J.I. Mullins. 2013. Factors Affecting Relative Fitness Measurements in Pairwise Competition Assays of Human Immunodeficiency Viruses. *Journal of virological methods*
- Le, S.Q., C.C. Dang, and O. Gascuel. 2012. Modeling protein evolution with several amino acid replacement matrices depending on site rates. *Molecular biology and evolution* 29:2921-2936.
- Leslie, A., D.A. Price, P. Mkhize, K. Bishop, A. Rathod, C. Day, H. Crawford, I. Honeyborne, T.E. Asher, G. Luzzi, A. Edwards, C.M. Rousseau, J.I. Mullins, G. Tudor-Williams, V. Novelli, C. Brander, D.C. Douek, P. Kiepiela, B.D. Walker, and P.J. Goulder. 2006. Differential selection pressure exerted on HIV by CTL targeting identical epitopes but restricted by distinct HLA alleles from the same HLA supertype. *J Immunol* 177:4699-4708.
- Leslie, A.J., K.J. Pfafferott, P. Chetty, R. Draenert, M.M. Addo, M. Feeney, Y. Tang, E.C. Holmes, T. Allen, J.G. Prado, M. Altfeld, C. Brander, C. Dixon, D. Ramduth, P. Jeena, S.A. Thomas, A. St John, T.A. Roach, B. Kupfer, G. Luzzi, A. Edwards, G. Taylor, H. Lyall, G. Tudor-Williams, V. Novelli, J. Martinez-Picado, P. Kiepiela, B.D. Walker, and P.J. Goulder. 2004. HIV evolution: CTL escape mutation and reversion after transmission. *Nat Med* 10:282-289.
- Letourneau, S., E.J. Im, T. Mashishi, C. Brereton, A. Bridgeman, H. Yang, L. Dorrell, T. Dong, B. Korber, A.J. McMichael, and T. Hanke. 2007. Design and pre-clinical evaluation of a universal HIV-1 vaccine. *PloS one* 2:e984.
- Li, B., A.D. Gladden, M. Altfeld, J.M. Kaldor, D.A. Cooper, A.D. Kelleher, and T.M. Allen. 2007. Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. *J Virol* 81:193-201.
- Liao, H.X., R. Lynch, T. Zhou, F. Gao, S.M. Alam, S.D. Boyd, A.Z. Fire, K.M. Roskin, C.A. Schramm, Z. Zhang, J. Zhu, L. Shapiro, J.C. Mullikin, S. Gnanakaran, P. Hraber, K. Wiehe, G. Kelsoe, G. Yang, S.M. Xia, D.C. Montefiori, R. Parks, K.E. Lloyd, R.M. Searce, K.A. Soderberg, M. Cohen, G. Kamanga, M.K. Louder, L.M. Tran, Y. Chen, F. Cai, S. Chen, S. Moquin, X. Du, M.G. Joyce, S. Srivatsan, B. Zhang, A. Zheng, G.M. Shaw, B.H. Hahn, T.B. Kepler, B.T. Korber, P.D. Kwong,

- J.R. Mascola, and B.F. Haynes. 2013. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* 496:469-476.
- Liu, M.K., N. Hawkins, A.J. Ritchie, V.V. Ganusov, V. Whale, S. Brackenridge, H. Li, J.W. Pavlicek, F. Cai, M. Rose-Abrahams, F. Treurnicht, P. Hrabec, C. Riou, C. Gray, G. Ferrari, R. Tanner, L.H. Ping, J.A. Anderson, R. Swanstrom, M. Cohen, S.S. Karim, B. Haynes, P. Borrow, A.S. Perelson, G.M. Shaw, B.H. Hahn, C. Williamson, B.T. Korber, F. Gao, S. Self, A. McMichael, and N. Goonetilleke. 2013. Vertical T cell immunodominance and epitope entropy determine HIV-1 escape. *The Journal of clinical investigation* 123:380-393.
- Liu, Y., J. McNevin, J. Cao, H. Zhao, I. Genowati, K. Wong, S. McLaughlin, M.D. McSweyn, K. Diem, C.E. Stevens, J. Maenza, H. He, D.C. Nickle, D. Shriner, S.E. Holte, A.C. Collier, L. Corey, M.J. McElrath, and J.I. Mullins. 2006. Selection on the human immunodeficiency virus type 1 proteome following primary infection. *J Virol* 80:9519-9529.
- Liu, Y., J. McNevin, H. Zhao, D.M. Tebit, R.M. Troyer, M. McSweyn, A.K. Ghosh, D. Shriner, E.J. Arts, M.J. McElrath, and J.I. Mullins. 2007. Evolution of human immunodeficiency virus type 1 cytotoxic T-lymphocyte epitopes: fitness-balanced escape. *J Virol* 81:12179-12188.
- Liu, Y., J.P. McNevin, S. Holte, M.J. McElrath, and J.I. Mullins. 2011. Dynamics of viral evolution and CTL responses in HIV-1 infection. *PloS one* 6:e15639.
- Llano, A.F., N. Brander, C. . 2009. How to Optimally Define Optimal Cytotoxic T Lymphocyte Epitopes in HIV Infection? In HIV Molecular Immunology. B.T.M.K. Karina Yusim, Christian Brander, Barton F. Haynes, Richard Koup, John P., and B.D.W. Moore, and David I. Watkins., editors. Los Alamos National Laboratory, Theoretical Biology and Biophysics Los Alamos, New Mexico. 3-24.
- Lu, K., X. Heng, and M.F. Summers. 2011. Structural determinants and mechanism of HIV-1 genome packaging. *Journal of molecular biology* 410:609-633.
- Malhotra, U., J. Nolin, H. Horton, F. Li, L. Corey, J.I. Mullins, and M.J. McElrath. 2009. Functional properties and epitope characteristics of T-cells recognizing natural HIV-1 variants. *Vaccine* 27:6678-6687.
- Manocheewa, S., J.V. Swain, E. Lanxon-Cookson, M. Rolland, and J.I. Mullins. 2013. Fitness Costs of Mutations at the HIV-1 Capsid Hexamerization Interface. *PloS one* 8:e66065.
- Martinez-Picado, J., J.G. Prado, E.E. Fry, K. Pfafferott, A. Leslie, S. Chetty, C. Thobakgale, I. Honeyborne, H. Crawford, P. Matthews, T. Pillay, C. Rousseau, J.I. Mullins, C. Brander, B.D. Walker, D.I. Stuart, P. Kiepiela, and P. Goulder. 2006. Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. *J Virol* 80:3617-3623.
- Masemola, A., T. Mashishi, G. Khoury, P. Mohube, P. Mokgotho, E. Vardas, M. Colvin, L. Zijenah, D. Katzenstein, R. Musonda, S. Allen, N. Kumwenda, T. Taha, G. Gray, J. McIntyre, S.A. Karim, H.W. Sheppard, and C.M. Gray. 2004. Hierarchical targeting of subtype C human immunodeficiency virus type 1 proteins by CD8+ T cells: correlation with viral load. *J Virol* 78:3233-3243.
- Mason, D. 1998. A very high level of crossreactivity is an essential feature of the T-cell receptor. *Immunology today* 19:395-404.

- Mattapallil, J.J., D.C. Douek, B. Hill, Y. Nishimura, M. Martin, and M. Roederer. 2005. Massive infection and loss of memory CD4⁺ T cells in multiple tissues during acute SIV infection. *Nature* 434:1093-1097.
- Matthews, P.C., A. Prendergast, A. Leslie, H. Crawford, R. Payne, C. Rousseau, M. Rolland, I. Honeyborne, J. Carlson, C. Kadie, C. Brander, K. Bishop, N. Mlotshwa, J.I. Mullins, H. Coovadia, T. Ndung'u, B.D. Walker, D. Heckerman, and P.J. Goulder. 2008. Central role of reverting mutations in HLA associations with human immunodeficiency virus set point. *J Virol* 82:8548-8559.
- Mavigner, M., M. Cazabat, M. Dubois, F.E. L'Faqihi, M. Requena, C. Pasquier, P. Klopp, J. Amar, L. Alric, K. Barange, J.P. Vinel, B. Marchou, P. Massip, J. Izopet, and P. Delobel. 2012. Altered CD4⁺ T cell homing to the gut impairs mucosal immune reconstitution in treated HIV-infected individuals. *The Journal of clinical investigation* 122:62-69.
- McClure, J., A.B. van't Wout, T. Tran, and J.E. Mittler. 2007. Granulocyte-monocyte colony-stimulating factor upregulates HIV-1 replication in monocyte-derived macrophages cultured at low density. *J Acquir Immune Defic Syndr* 44:254-261.
- McElrath, M.J., S.C. De Rosa, Z. Moodie, S. Dubey, L. Kierstead, H. Janes, O.D. Defawe, D.K. Carter, J. Hural, R. Akondy, S.P. Buchbinder, M.N. Robertson, D.V. Mehrotra, S.G. Self, L. Corey, J.W. Shiver, and D.R. Casimiro. 2008. HIV-1 vaccine-induced immunity in the test-of-concept Step Study: a case-cohort analysis. *Lancet* 372:1894-1905.
- McMichael, A.J., P. Borrow, G.D. Tomaras, N. Goonetilleke, and B.F. Haynes. 2010. The immune response during acute HIV-1 infection: clues for vaccine development. *Nature reviews. Immunology* 10:11-23.
- Migueles, S.A., A.C. Laborico, W.L. Shupert, M.S. Sabbaghian, R. Rabin, C.W. Hallahan, D. Van Baarle, S. Kostense, F. Miedema, M. McLaughlin, L. Ehler, J. Metcalf, S. Liu, and M. Connors. 2002. HIV-specific CD8⁺ T cell proliferation is coupled to perforin expression and is maintained in nonprogressors. *Nature immunology* 3:1061-1068.
- Migueles, S.A., C.M. Osborne, C. Royce, A.A. Compton, R.P. Joshi, K.A. Weeks, J.E. Rood, A.M. Berkley, J.B. Sacha, N.A. Coglianor-Shutta, M. Lloyd, G. Roby, R. Kwan, M. McLaughlin, S. Stallings, C. Rehm, M.A. O'Shea, J. Mican, B.Z. Packard, A. Komoriya, S. Palmer, A.P. Wiegand, F. Maldarelli, J.M. Coffin, J.W. Mellors, C.W. Hallahan, D.A. Follman, and M. Connors. 2008. Lytic granule loading of CD8⁺ T cells is required for HIV-infected cell elimination associated with immune control. *Immunity* 29:1009-1021.
- Migueles, S.A., M.S. Sabbaghian, W.L. Shupert, M.P. Bettinotti, F.M. Marincola, L. Martino, C.W. Hallahan, S.M. Selig, D. Schwartz, J. Sullivan, and M. Connors. 2000. 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.
- Milush, J.M., K.D. Mir, V. Sundaravaradan, S.N. Gordon, J. Engram, C.A. Cano, J.D. Reeves, E. Anton, E. O'Neill, E. Butler, K. Hancock, K.S. Cole, J.M. Brenchley, J.G. Else, G. Silvestri, and D.L. Sodora. 2011. Lack of clinical AIDS in SIV-infected sooty mangabeys with significant CD4⁺ T cell loss is associated with double-negative T cells. *The Journal of clinical investigation* 121:1102-1110.

- Mir, K.D., M.A. Gasper, V. Sundaravaradan, and D.L. Sodora. 2011. SIV infection in natural hosts: resolution of immune activation during the acute-to-chronic transition phase. *Microbes and infection / Institut Pasteur* 13:14-24.
- Miura, T., M.A. Brockman, A. Schneidewind, M. Lobritz, F. Pereyra, A. Rathod, B.L. Block, Z.L. Brumme, C.J. Brumme, B. Baker, A.C. Rothchild, B. Li, A. Trocha, E. Cutrell, N. Frahm, C. Brander, I. Toth, E.J. Arts, T.M. Allen, and B.D. Walker. 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.
- Montano, M.W., C. 2002. The Molecular Virology of HIV-1. In *AIDS in Africa*. M. Essex, editor Kluwer Academic/Plenum Publishers, New York, New York. 11-34.
- Morikawa, Y., T. Goto, and K. Sano. 1999. In vitro assembly of human immunodeficiency virus type 1 Gag protein. *The Journal of biological chemistry* 274:27997-28002.
- Mothe, B., A. Llano, J. Ibarrondo, J. Zamarreno, M. Schiaulini, C. Miranda, M. Ruiz-Riol, C.T. Berger, M.J. Herrero, E. Palou, M. Plana, M. Rolland, A. Khatri, D. Heckerman, F. Pereyra, B.D. Walker, D. Weiner, R. Paredes, B. Clotet, B.K. Felber, G.N. Pavlakis, J.I. Mullins, and C. Brander. 2012. CTL responses of high functional avidity and broad variant cross-reactivity are associated with HIV control. *PloS one* 7:e29717.
- Mudd, P.A., M.A. Martins, A.J. Ericson, D.C. Tully, K.A. Power, A.T. Bean, S.M. Piaskowski, L. Duan, A. Seese, A.D. Gladden, K.L. Weisgrau, J.R. Furlott, Y.I. Kim, M.G. Veloso de Santana, E. Rakasz, S. Capuano, 3rd, N.A. Wilson, M.C. Bonaldo, R. Galler, D.B. Allison, M. Piatak, Jr., A.T. Haase, J.D. Lifson, T.M. Allen, and D.I. Watkins. 2012. Vaccine-induced CD8+ T cells control AIDS virus replication. *Nature* 491:129-133.
- Nabel, G., and D. Baltimore. 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* 326:711-713.
- Nakayama, E.E., and T. Shioda. 2010. Anti-retroviral activity of TRIM5 alpha. *Reviews in medical virology* 20:77-92.
- Nickle, D.C., M. Rolland, M.A. Jensen, S.L. Pond, W. Deng, M. Seligman, D. Heckerman, J.I. Mullins, and N. Jojic. 2007. Coping with viral diversity in HIV vaccine design. *PLoS computational biology* 3:e75.
- Niu, L., J.M. Termini, S.K. Kanagavelu, S. Gupta, M.M. Rolland, V. Kulkarni, G.N. Pavlakis, B.K. Felber, J.I. Mullins, M.A. Fischl, and G.W. Stone. 2011. Preclinical evaluation of HIV-1 therapeutic ex vivo dendritic cell vaccines expressing consensus Gag antigens and conserved Gag epitopes. *Vaccine* 29:2110-2119.
- O'Connor, D.H., T.M. Allen, T.U. Vogel, P. Jing, I.P. DeSouza, E. Dodds, E.J. Dunphy, C. Melsaether, B. Mothe, H. Yamamoto, H. Horton, N. Wilson, A.L. Hughes, and D.I. Watkins. 2002. Acute phase cytotoxic T lymphocyte escape is a hallmark of simian immunodeficiency virus infection. *Nat Med* 8:493-499.
- Ogg, G.S., X. Jin, S. Bonhoeffer, P. Moss, M.A. Nowak, S. Monard, J.P. Segal, Y. Cao, S.L. Rowland-Jones, A. Hurley, M. Markowitz, D.D. Ho, A.J. McMichael, and D.F. Nixon. 1999. Decay kinetics of human immunodeficiency virus-specific effector cytotoxic T lymphocytes after combination antiretroviral therapy. *J Virol* 73:797-800.

Organization, W.H. 2007. WHO CASE DEFINITIONS OF HIV FOR SURVEILLANCE AND REVISED CLINICAL

STAGING AND IMMUNOLOGICAL CLASSIFICATION OF HIV-RELATED DISEASE IN ADULTS AND CHILDREN.

- Patel, V., R. Jalah, V. Kulkarni, A. Valentin, M. Rosati, C. Alicea, A. von Gegerfelt, W. Huang, Y. Guan, B.F. Keele, J.W. Bess, Jr., M. Piatak, Jr., J.D. Lifson, W.T. Williams, X. Shen, G.D. Tomaras, R.R. Amara, H.L. Robinson, W. Johnson, K.E. Broderick, N.Y. Sardesai, D.J. Venzon, V.M. Hirsch, B.K. Felber, and G.N. Pavlakis. 2013. DNA and virus particle vaccination protects against acquisition and confers control of viremia upon heterologous simian immunodeficiency virus challenge. *Proceedings of the National Academy of Sciences of the United States of America* 110:2975-2980.
- Preston, B.D., B.J. Poiesz, and L.A. Loeb. 1988. Fidelity of HIV-1 reverse transcriptase. *Science* 242:1168-1171.
- Prince, J.L., D.T. Claiborne, J.M. Carlson, M. Schaefer, T. Yu, S. Lahki, H.A. Prentice, L. Yue, S.A. Vishwanathan, W. Kilembe, P. Goepfert, M.A. Price, J. Gilmour, J. Mulenga, P. Farmer, C.A. Derdeyn, J. Tang, D. Heckerman, R.A. Kaslow, S.A. Allen, and E. Hunter. 2012. Role of transmitted Gag CTL polymorphisms in defining replicative capacity and early HIV-1 pathogenesis. *PLoS pathogens* 8:e1003041.
- Reed, J.C., B. Molter, C.D. Geary, J. McNevin, J. McElrath, S. Giri, K.C. Klein, and J.R. Lingappa. 2012. HIV-1 Gag co-opts a cellular complex containing DDX6, a helicase that facilitates capsid assembly. *The Journal of cell biology* 198:439-456.
- Reed, L.J.M., H. 1938. A simple method of estimating fifty percent endpoints. *The American Journal of Hygiene* 27:493-497.
- Renjifo, B.E., M. 2002. HIV-1 Subtypes and Recombinants. In AIDS in Africa. M. Essex, editor Kluwer Academic/Plenum Publishers, New York, New York. 263-281.
- Rerks-Ngarm, S., P. Pitisuttithum, S. Nitayaphan, J. Kaewkungwal, J. Chiu, R. Paris, N. Premisri, C. Namwat, M. de Souza, E. Adams, M. Benenson, S. Gurunathan, J. Tartaglia, J.G. McNeil, D.P. Francis, D. Stablein, D.L. Birx, S. Chunsuttiwat, C. Khamboonruang, P. Thongcharoen, M.L. Robb, N.L. Michael, P. Kunasol, and J.H. Kim. 2009. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *The New England journal of medicine* 361:2209-2220.
- Rihn, S.J., S.J. Wilson, N.J. Loman, M. Alim, S.E. Bakker, D. Bhella, R.J. Gifford, F.J. Rixon, and P.D. Bieniasz. 2013. Extreme Genetic Fragility of the HIV-1 Capsid. *PLoS pathogens* 9:e1003461.
- Robertson, D.L., J.P. Anderson, J.A. Bradac, J.K. Carr, B. Foley, R.K. Funkhouser, F. Gao, B.H. Hahn, M.L. Kalish, C. Kuiken, G.H. Learn, T. Leitner, F. McCutchan, S. Osmanov, M. Peeters, D. Pieniazek, M. Salminen, P.M. Sharp, S. Wolinsky, and B. Korber. 2000. HIV-1 nomenclature proposal. *Science* 288:55-56.
- Rolland, M. 2013. Early HIV-1 adaptation in the context of natural infection or following vaccination. In VIDD Faculty Seminar Series. Seattle, WA.
- Rolland, M., J.M. Carlson, S. Manochewa, J.V. Swain, E. Lanxon-Cookson, W. Deng, C.M. Rousseau, D.N. Raugi, G.H. Learn, B.S. Maust, H. Coovadia, T. Ndung'u, P.J. Goulder, B.D. Walker, C. Brander, D.E. Heckerman, and J.I. Mullins. 2010.

- Amino-acid co-variation in HIV-1 Gag subtype C: HLA-mediated selection pressure and compensatory dynamics. *PloS one* 5:
- Rolland, M., D. Heckerman, W. Deng, C.M. Rousseau, H. Coovadia, K. Bishop, P.J. Goulder, B.D. Walker, C. Brander, and J.I. Mullins. 2008. Broad and Gag-biased HIV-1 epitope repertoires are associated with lower viral loads. *PloS one* 3:e1424.
- Rolland M, M.S., Swain JV, Lanxon-Cookson EC, Kim M, Westfall DH, Larsen BB, Gilbert PB, Mullins JI. 2013. HIV-1 Conserved-Element Vaccines: Relationship between Sequence Conservation and Replicative Capacity. *Journal of Virology* 87:5461-5467.
- Rolland, M., D.C. Nickle, and J.I. Mullins. 2007. HIV-1 group M conserved elements vaccine. *PLoS pathogens* 3:e157.
- Rolland, M., S. Tovanabutra, A.C. deCamp, N. Frahm, P.B. Gilbert, E. Sanders-Buell, L. Heath, C.A. Magaret, M. Bose, A. Bradfield, A. O'Sullivan, J. Crossler, T. Jones, M. Nau, K. Wong, H. Zhao, D.N. Raugi, S. Sorensen, J.N. Stoddard, B.S. Maust, W. Deng, J. Hural, S. Dubey, N.L. Michael, J. Shiver, L. Corey, F. Li, S.G. Self, J. Kim, S. Buchbinder, D.R. Casimiro, M.N. Robertson, A. Duerr, M.J. McElrath, F.E. McCutchan, and J.I. Mullins. 2011. Genetic impact of vaccination on breakthrough HIV-1 sequences from the STEP trial. *Nat Med* 17:366-371.
- Sacha, J.B., C. Chung, E.G. Rakasz, S.P. Spencer, A.K. Jonas, A.T. Bean, W. Lee, B.J. Burwitz, J.J. Stephany, J.T. Loffredo, D.B. Allison, S. Adnan, A. Hoji, N.A. Wilson, T.C. Friedrich, J.D. Lifson, O.O. Yang, and D.I. Watkins. 2007. Gag-specific CD8+ T lymphocytes recognize infected cells before AIDS-virus integration and viral protein expression. *J Immunol* 178:2746-2754.
- Santra, S., H.X. Liao, R. Zhang, M. Muldoon, S. Watson, W. Fischer, J. Theiler, J. Szinger, H. Balachandran, A. Buzby, D. Quinn, R.J. Parks, C.Y. Tsao, A. Carville, K.G. Mansfield, G.N. Pavlakis, B.K. Felber, B.F. Haynes, B.T. Korber, and N.L. Letvin. 2010. Mosaic vaccines elicit CD8+ T lymphocyte responses that confer enhanced immune coverage of diverse HIV strains in monkeys. *Nat Med* 16:324-328.
- Santra, S., M. Muldoon, S. Watson, A. Buzby, H. Balachandran, K.R. Carlson, L. Mach, W.P. Kong, K. McKee, Z.Y. Yang, S.S. Rao, J.R. Mascola, G.J. Nabel, B.T. Korber, and N.L. Letvin. 2012. Breadth of cellular and humoral immune responses elicited in rhesus monkeys by multi-valent mosaic and consensus immunogens. *Virology* 428:121-127.
- Schacker, T., A.C. Collier, J. Hughes, T. Shea, and L. Corey. 1996. Clinical and epidemiologic features of primary HIV infection. *Annals of internal medicine* 125:257-264.
- Schacker, T.W., J.P. Hughes, T. Shea, R.W. Coombs, and L. Corey. 1998. Biological and virologic characteristics of primary HIV infection. *Annals of internal medicine* 128:613-620.
- Schaller, T., K.E. Ocwieja, J. Rasaiyaah, A.J. Price, T.L. Brady, S.L. Roth, S. Hue, A.J. Fletcher, K. Lee, V.N. KewalRamani, M. Noursadeghi, R.G. Jenner, L.C. James, F.D. Bushman, and G.J. Towers. 2011. HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS pathogens* 7:e1002439.

- Schmitz, J., M. Kuroda, S. Santra, V. Sasseville, M. Simon, M. Lifton, P. Racz, K. Tenner-Racz, M. Dalesandro, B. Scallon, J. Ghayeb, M. Forman, D. Montefiori, E. Rieber, N. Letvin, and K. Reimann. 1999. Control of viremia in simian immunodeficiency virus infection by CD8⁺ lymphocytes. *Science* 277:333-338.
- Schneidewind, A., M.A. Brockman, J. Sidney, Y.E. Wang, H. Chen, T.J. Suscovich, B. Li, R.I. Adam, R.L. Allgaier, B.R. Mothe, T. Kuntzen, C. Oniangue-Ndza, A. Trocha, X.G. Yu, C. Brander, A. Sette, B.D. Walker, and T.M. Allen. 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.
- Sewell, A.K. 2012. Why must T cells be cross-reactive? *Nature reviews. Immunology* 12:669-677.
- Shankarappa, R., J.B. Margolick, S.J. Gange, A.G. Rodrigo, D. Upchurch, H. Farzadegan, P. Gupta, C.R. Rinaldo, G.H. Learn, X. He, X.L. Huang, and J.I. Mullins. 1999. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J Virol* 73:10489-10502.
- Shin, H., S.D. Blackburn, J.N. Blattman, and E.J. Wherry. 2007. Viral antigen and extensive division maintain virus-specific CD8 T cells during chronic infection. *The Journal of experimental medicine* 204:941-949.
- Simon, V., and D.D. Ho. 2003. HIV-1 dynamics in vivo: implications for therapy. *Nature reviews. Microbiology* 1:181-190.
- Simon, V., D.D. Ho, and Q. Abdool Karim. 2006. HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. *Lancet* 368:489-504.
- Sleiman, D., V. Goldschmidt, P. Barraud, R. Marquet, J.C. Paillart, and C. Tisne. 2012. Initiation of HIV-1 reverse transcription and functional role of nucleocapsid-mediated tRNA/viral genome interactions. *Virus research* 169:324-339.
- Stekler, J., B.J. Sycks, S. Holte, J. Maenza, C.E. Stevens, J. Dragavon, A.C. Collier, and R.W. Coombs. 2008. HIV dynamics in seminal plasma during primary HIV infection. *AIDS research and human retroviruses* 24:1269-1274.
- Strack, B., A. Calistri, S. Craig, E. Popova, and H.G. Gottlinger. 2003. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. *Cell* 114:689-699.
- Streeck, H., Z.L. Brumme, M. Anastario, K.W. Cohen, J.S. Jolin, A. Meier, C.J. Brumme, E.S. Rosenberg, G. Alter, T.M. Allen, B.D. Walker, and M. Altfeld. 2008. Antigen load and viral sequence diversification determine the functional profile of HIV-1-specific CD8⁺ T cells. *PLoS medicine* 5:e100.
- Streeck, H., M. Lichterfeld, G. Alter, A. Meier, N. Teigen, B. Yassine-Diab, H.K. Sidhu, S. Little, A. Kelleher, J.P. Routy, E.S. Rosenberg, R.P. Sekaly, B.D. Walker, and M. Altfeld. 2007. Recognition of a defined region within p24 gag by CD8⁺ T cells during primary human immunodeficiency virus type 1 infection in individuals expressing protective HLA class I alleles. *J Virol* 81:7725-7731.
- Troyer, R.M., J. McNevin, Y. Liu, S.C. Zhang, R.W. Krizan, A. Abraha, D.M. Tebit, H. Zhao, S. Avila, M.A. Lobritz, M.J. McElrath, S. Le Gall, J.I. Mullins, and E.J. Arts. 2009. Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. *PLoS pathogens* 5:e1000365.

- Turnbull, E.L., J. Baalwa, K.E. Conrod, S. Wang, X. Wei, M. Wong, J. Turner, P. Pellegrino, I. Williams, G.M. Shaw, and P. Borrow. 2011. Escape is a more common mechanism than avidity reduction for evasion of CD8+ T cell responses in primary human immunodeficiency virus type 1 infection. *Retrovirology* 8:41.
- Turnbull, E.L., A.R. Lopes, N.A. Jones, D. Cornforth, P. Newton, D. Aldam, P. Pellegrino, J. Turner, I. Williams, C.M. Wilson, P.A. Goepfert, M.K. Maini, and P. Borrow. 2006. HIV-1 epitope-specific CD8+ T cell responses strongly associated with delayed disease progression cross-recognize epitope variants efficiently. *J Immunol* 176:6130-6146.
- UNAIDS. 2012. UNAIDS World AIDS Day Report
- Valentine, L.E., S.M. Piaskowski, E.G. Rakasz, N.L. Henry, N.A. Wilson, and D.I. Watkins. 2008. Recognition of escape variants in ELISPOT does not always predict CD8+ T-cell recognition of simian immunodeficiency virus-infected cells expressing the same variant sequences. *J Virol* 82:575-581.
- Vidal, N., M. Peeters, C. Mulanga-Kabeya, N. Nzilambi, D. Robertson, W. Ilunga, H. Sema, K. Tshimanga, B. Bongo, and E. Delaporte. 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.
- Walker, B.D., S. Chakrabarti, B. Moss, T.J. Paradis, T. Flynn, A.G. Durno, R.S. Blumberg, J.C. Kaplan, M.S. Hirsch, and R.T. Schooley. 1987. HIV-specific cytotoxic T lymphocytes in seropositive individuals. *Nature* 328:345-348.
- Wei, X., J.M. Decker, S. Wang, H. Hui, J.C. Kappes, X. Wu, J.F. Salazar-Gonzalez, M.G. Salazar, J.M. Kilby, M.S. Saag, N.L. Komarova, M.A. Nowak, B.H. Hahn, P.D. Kwong, and G.M. Shaw. 2003. Antibody neutralization and escape by HIV-1. *Nature* 422:307-312.
- Wherry, E.J., J.N. Blattman, K. Murali-Krishna, R. van der Most, and R. Ahmed. 2003. Viral persistence alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment. *J Virol* 77:4911-4927.
- Wilson, N.A., B.F. Keele, J.S. Reed, S.M. Piaskowski, C.E. MacNair, A.J. Bett, X. Liang, F. Wang, E. Thoryk, G.J. Heidecker, M.P. Citron, L. Huang, J. Lin, S. Vitelli, C.D. Ahn, M. Kaizu, N.J. Maness, M.R. Reynolds, T.C. Friedrich, J.T. Loffredo, E.G. Rakasz, S. Erickson, D.B. Allison, M. Piatak, Jr., J.D. Lifson, J.W. Shiver, D.R. Casimiro, G.M. Shaw, B.H. Hahn, and D.I. Watkins. 2009. Vaccine-induced cellular responses control simian immunodeficiency virus replication after heterologous challenge. *J Virol* 83:6508-6521.
- Wu, X., T. Zhou, J. Zhu, B. Zhang, I. Georgiev, C. Wang, X. Chen, N.S. Longo, M. Louder, K. McKee, S. O'Dell, S. Perfetto, S.D. Schmidt, W. Shi, L. Wu, Y. Yang, Z.Y. Yang, Z. Yang, Z. Zhang, M. Bonsignori, J.A. Crump, S.H. Kapiga, N.E. Sam, B.F. Haynes, M. Simek, D.R. Burton, W.C. Koff, N.A. Doria-Rose, M. Connors, J.C. Mullikin, G.J. Nabel, M. Roederer, L. Shapiro, P.D. Kwong, and J.R. Mascola. 2011. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science* 333:1593-1602.
- Yamamoto, T., M.J. Johnson, D.A. Price, D.I. Wolinsky, J.R. Almeida, C. Petrovas, M. Nason, W.W. Yeh, L. Shen, M. Roederer, S.S. Rao, A.B. McDermott, F. Lefebvre, G.J. Nabel, E.K. Haddad, N.L. Letvin, D.C. Douek, and R.A. Koup. 2012. Virus

- inhibition activity of effector memory CD8(+) T cells determines simian immunodeficiency virus load in vaccinated monkeys after vaccine breakthrough infection. *J Virol* 86:5877-5884.
- Yerly, D., D. Heckerman, T.M. Allen, J.V. Chisholm, 3rd, K. Faircloth, C.H. Linde, N. Frahm, J. Timm, W.J. Pichler, A. Cerny, and C. Brander. 2008. Increased cytotoxic T-lymphocyte epitope variant cross-recognition and functional avidity are associated with hepatitis C virus clearance. *J Virol* 82:3147-3153.
- Zembe, L., W.A. Burgers, H.B. Jaspan, L.G. Bekker, H. Bredell, G. Stevens, J. Gilmour, J.H. Cox, P. Fast, P. Hayes, E. Vardas, C. Williamson, and C.M. Gray. 2011. Intra- and Inter-clade Cross-reactivity by HIV-1 Gag Specific T-Cells Reveals Exclusive and Commonly Targeted Regions: Implications for Current Vaccine Trials. *PloS one* 6:e26096.
- Zhu, T., B.T. Korber, A.J. Nahmias, E. Hooper, P.M. Sharp, and D.D. Ho. 1998. An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. *Nature* 391:594-597.
- Zhu, T., H. Mo, N. Wang, D.S. Nam, Y. Cao, R.A. Koup, and D.D. Ho. 1993. Genotypic and phenotypic characterization of HIV-1 patients with primary infection. *Science* 261:1179-1181.
- Zimmerli, S.C., A. Harari, C. Cellerai, F. Vallelian, P.A. Bart, and G. Pantaleo. 2005. HIV-1-specific IFN-gamma/IL-2-secreting CD8 T cells support CD4-independent proliferation of HIV-1-specific CD8 T cells. *Proceedings of the National Academy of Sciences of the United States of America* 102:7239-7244.
- Zuniga, R., A. Lucchetti, P. Galvan, S. Sanchez, C. Sanchez, A. Hernandez, H. Sanchez, N. Frahm, C.H. Linde, H.S. Hewitt, W. Hildebrand, M. Altfeld, T.M. Allen, B.D. Walker, B.T. Korber, T. Leitner, J. Sanchez, and C. Brander. 2006. Relative dominance of Gag p24-specific cytotoxic T lymphocytes is associated with human immunodeficiency virus control. *J Virol* 80:3122-3125.