Genetic characterization of viral blips in patients following suppressive HIV ART

Lindiwe Amanda Jhamba

Supervisors: Dr Kamini Gounder & Prof Thumbi Ndung'u

Submitted in fulfilment of the requirements for the degree of Master of Medical Science (Virology) in the School of Laboratory Medicine and Medical Science, HIV Pathogenesis Programme, Doris Duke Medical Research Institute, Nelson R. Mandela School of Medicine, University of KwaZulu-Natal.

2021

DECLARATION

I, Ms Lindiwe Amanda Jhamba, declare as follows:

(i) The research reported in this thesis, except where otherwise indicated, is my original research.

(ii) This thesis has not been submitted for any degree or examination at any other university.

(iii) This thesis does not contain other persons' data, pictures, graphs, or other information, unless specifically acknowledged as being sourced from other persons.

(iv) This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a) Their words have been re-written, but the general information attributed to them has been referenced,

b) Where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.

(v) Where I have reproduced a publication of which I am an author, co-author or editor, I have indicated in detail which part of the publication was actually written by myself alone and fully referenced such publications.

(vi) This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Student: LA Jhamba  Date: 31 October 2021

Supervisor: Dr K Gounder  Date: 31October 2021

Co-supervisor: Prof T Ndung'u  Date: 1 November 2021

DEDICATION

To GOD from Whom all blessings flow

To Tafara, Simphiwe, Mufaro, Nandi, thank you

ACKNOWLEGEMENTS

My sincere gratitude goes to Dr K Gounder for her academic guidance, dedication, and support.

Steven and Shiyiwe Jhamba, for emotional support and the motivation to press forward.

Lungile, Victor, Samukeliso, Ruth, thank you for walking with me.
Special thanks to my family and friends for your prayers and encouragement.

The staff and students of the HIV Pathogenesis Programme, for training and fostering a friendly research environment.

The participants of the Females Rising through Education, Support and Health (FRESH) acute infection cohort for their samples.

Prof T Ndung'u for the opportunity to join the Sub-Saharan African Network for TB/HIV Research Excellence (SANTHE) fellowship program.

The Poliomyelitis Research Foundation (PRF) for funding support.

This work was supported through the Sub-Saharan African Network for TB/HIV Research Excellence (SANTHE), a DELTAS Africa Initiative [grant # DEL-15-006]. The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust [grant # 107752/Z/15/Z] and the UK government. The views expressed in this publication are those of the author(s) and not necessarily those of AAS, NEPAD Agency, Wellcome Trust or the UK government.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

µL	Microlitre
AIDS	acquired immuno-deficiency syndrome
Amp	amplification
ART	Antiretroviral therapy
bnAb	broadly neutralizing antibodies
cDNA	complementary Deoxyribonucleic acid
CTL	Cytotoxic CD8+ T lymphocyte
D	aspartic acid
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
E	glutamic acid
EDTA	ethylenediaminetetraacetic acid
EFV	efavirenz
env	envelope
FDC	fixed-dose combination pill
FRESH	Females Rising through Education, Support and Health (FRESH) cohort
FTC	emtricitabine
g	relative centrifugal force
G	glycine
gag	group-specific antigen
HIV	human immuno-deficiency virus
HLA	human leukocyte antigen
LTR	Long terminal repeat
Min	minute
mL	millilitre
Mm	millimolar
N	asparagine
NaOAc	sodium acetate
NFLG	near full-length genome
NNRTI	non–nucleoside reverse transcriptase inhibitors

NRTI	nucleoside-analog reverse transcriptase inhibitors
pM	picomolar
pol	polymerase
PreART	pre-treatment initiation virus
R	arginine
RAL	raltegravir
RNA	Ribonucleic acid protease inhibitors
RNaseOUT	Recombinant RNase Inhibitor
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
s	seconds
S	serine
SGA	Single Genome Amplification
SIV	simian immune-deficiency virus
T	threonine
T/F	Transmitter/founder virus
TDF	tenofovir
vif	Viral infectivity factor
vpu	Viral protein

ABSTRACT

Antiretroviral therapy (ART) has resulted in the decline of HIV-related mortality worldwide. On initiating treatment, most patients can suppress plasma viral RNA to undetectable levels (<20 copies/mL). Patients on ART frequently experience intermittent viremia (viral blips), however the genetic nature and source of these rebounding viruses while on suppressive ART remains unclear. The study of these genetic characteristics would be essential in the development of targeted vaccines and adjunct treatments. We identified four HIV-1 subtype C infected women from the Females Rising through Education, Support and Health (FRESH) acute infection cohort who experienced viral blips following suppressive ART (median 584 days). Two participants initiated treatment during the chronic infection phase (~625 days post detection) and the other two initiated treatment immediately upon first HIV-1 RNA detection. RNA was extracted from stored plasma samples of participants' transmitter/founder (T/F) virus (~3 days post detection), pre-treatment initiation and during viral blips (>2,000 copies/mL). *Gag* and *env* genes were amplified by single genome amplification followed by sequencing. The protease and reverse transcriptase region of the *pol* gene were also amplified and bulk sequenced. Phylogenetic relatedness and genetic differences were visualized using Maximum-likelihood trees and Highlighter plots respectively (Los Alamos HIV-1 database). The *gag* and *env* blip sequences of the acute-treated participants were similar to those of the T/F, while those of the chronic-treated participants were genetically distinct from the T/F but similar to the pre-treatment initiation virus (PreART). In the acute-treated participants, all transmitted HLA-associated *gag* cytotoxic CD8+ T lymphocyte (CTL) escape identified was retained in the blip-derived sequences, however the chronic-treated participants experienced an increase of ~0.8% CTL escape at the blip time point. This increase coupled with development of a reduced replication capacity mutation (HLA-B*57:01/58:01 T242N), may be indicative of immune pressure prior to ART initiation. Mutations associated with bnAb escape in the CD4 binding, gp120/gp41 and V1V2 sites were identified only in the PreART and blip sequences of the chronic-treated participants, whereas the acute-treated retained the same amino acid residues at T/F and blip. With the exception of one chronic-treated participant who developed mutations associated with resistance to efavirenz, the viral blips were not associated with mutations linked to drug resistance. This data suggests that those who initiate treatment late are less likely to benefit from an immune response-inducing vaccine or bnAb therapy.

CHAPTER 1: INTRODUCTION

1.1 HIV origin

The human immunodeficiency virus (HIV) and the acquired immunodeficiency syndrome (AIDS) which it causes, have been a global health issue for several decades since its discovery in the 1980s, with the virus being believed to have crossed species from other primates to humans several decades earlier (1). The transmission of the simian immunodeficiency virus (SIV) to humans is widely accepted to have been due to contact with blood during the slaughter and or consumption of infected chimpanzees, within the central African communities of Congo (1). Advancements in global trade and migration contributed to the spread of the virus across the world, leading up to the first documented cases in the USA in the early 1980s (1, 2). HIV was distinguished into two major types in 1983 and 1986, namely HIV-1 and HIV-2, with the latter having crossed over from sooty mangabey monkeys and mostly associated with west Africa and the former having crossed over from chimpanzees and responsible for the majority of the global infection (3, 4). Since then, the virus has continued to evolve and the HIV/AIDS pandemic has claimed an estimated 36.3 million lives according to a 2021 UNAIDS report (5).

1.2 HIV epidemiology

According to the UNAIDS 2021 report, of the 37.3 million living with HIV globally, 20.6 million are from eastern and southern Africa (5). South Africa bears the greatest burden of disease, and it has the largest number of people enrolled for treatment in the world (6). A report published by Statistics South Africa in 2021, estimated that 8.2 million South Africans were living with HIV (6). Approximately a quarter of all South African women in their reproductive ages (15-49 years) are HIV positive (6). The KwaZulu-Natal and Gauteng provinces have the highest prevalence rates in South Africa (7). Of those newly infected, women aged 15 years and older were twice as likely to be infected than men of the same age (5). A recent study identified young women aged 15-24, from developing countries such as South Africa, as being at high risk of acquiring infection due to, cultural, socioeconomic, biological, and structural factors (8, 9). Prevention and support strategies, remain crucial in reducing incidence rates in Africa (8, 9).

1.3 Antiretroviral therapy

The HIV/AIDS pandemic is a global health issue for which there is neither a cure nor vaccine to date (10). There is however, antiretroviral therapy (ART), the introduction of which has seen a marked decrease in HIV-related mortality worldwide and an improved quality of life for those infected,

effectively turning a fatal condition into a manageable chronic illness (11). In South Africa the government bears the economic burden of providing this lifelong treatment to over 4.7 million people for free (5). Treatment is available in the form of six drug types: nucleoside-analog reverse transcriptase inhibitors (NRTIs), non–nucleoside reverse transcriptase inhibitors (NNRTIs), integrase inhibitors, protease inhibitors (PIs), fusion inhibitors, and coreceptor antagonists (12). ART regimens often include 3 drugs from the first 4 drug classes (12). When adhered to, ART suppresses the HIV RNA in most patients' blood plasma, to undetectable levels, of less than 50 copies/mL within a few months (13).

While ART is effective, there are several factors which may impact patient adherence, and these include stigma, access to health care and treatment, and adverse side effects to treatment. Some of these side effects include nausea, fatigue, diarrhoea, vomiting, high cholesterol, hepatotoxicity, and lipodystrophy (13). Notwithstanding its benefits, ART is not curative, as evidenced by reports of viral mRNA in peripheral blood mononuclear cells (14, 15), viral sequence evolution (16) and the presence of detectable, low level, intermittent viral blips in the plasma (17-19).The virus persists during treatment in these anatomical sites: gastrointestinal tract, spleen, brain, lungs, bone marrow and lymph node (13). These reservoir sites harbour latent infection, are impenetrable to treatment and so remain one of the greatest hurdles in the development of cure strategies (13). If treatment is interrupted, the virus rebounds within a few weeks and becomes detectable in the plasma again (20). Intermittent treatment interruption, subtherapeutic dosage due to patient non-compliance and drug-drug interactions due to undisclosed patient self-medication are just some of the known causes of ART drug resistance in patients (21, 22).

1.4 Viral rebound

Determining the source and cause of the rebound virus is key, to developing targeted eradication strategies. The source of these rebound viruses, and the mechanism by which they are maintained within the viral reservoirs remains debatable. While the exact source of the rebound virus is unknown, some studies suggest that it may be varied cross several sites (23). Two different mechanisms have been proposed for the maintenance of viral reservoirs during suppressive ART: the long-lived nature of infected cells and their proliferation or ongoing viral replication.

The question of ongoing HIV replication is a controversial one, and some recent studies, have questioned whether it occurs during suppressive ART (24, 25). In studies conducted by Deng et al. (2015) and Kearney *et al.*, (2014), no evidence of viral evolution in the rebound virus was observed (26, 27), however only plasma samples were used. A study by Lorenzo-Redondo et al, (2016) suggested that the observation of ongoing viral replication in the lymph nodes of patients who had been virally

suppressed for 5 years, was responsible for the maintenance of viral reservoirs (24). These findings were disputed by Kearney et al. (2017), who argued that the data set used was limited, had a limited number of unique viral sequences, and showed no comparative change in the level of sequence diversity prior to treatment versus during ART (27-29).

On the other hand, several studies argue that viral reservoirs are likely maintained by the clonal expansion of infected cells and their persistence in the long term (28, 30-34). In addition to the controversy, is the debated clinical significance of intermittent viral blips, observed in the plasma of patients on suppressive ART (20). While a study by Crowell *et al.* (2020), reported that early treatment initiation was associated with reduced viral blip frequency (35), little is known about the source, cause, and genetic characteristics of viral blips in patients who initiate treatment in the acute phase of infection and those starting treatment in the chronic phase of infection (35).

1.5 HIV immune response evasion

1.5.1 CD8+ T lymphocyte escape

Early treatment initiation has been associated with the development of smaller viral reservoirs and low viral diversity, making it an ideal model for cure research (29, 36-38). An understanding of the genetic characteristics of rebound viruses and their interaction with the host's immune response, is crucial in the development of targeted eradication strategies. The human leukocyte antigen (HLA) alleles are surface proteins which are responsible for antigen presentation during an immune response (39). HLA alleles have been shown to be associated with both positive and negative disease outcomes (39, 40). Administering therapies which induce an immune response, may be one such potential strategy against viral blips. Shan et al. (2012) demonstrated that CD8+ T cells (CTLs) from patients on treatment, were not able to kill previously latently infected CD4+ T cells, following their reactivation (41). However, when the CTLs were first stimulated using *gag* peptides, the reactivated infected cells were successfully killed. On conducting a similar study with patients who initiated treatment in the chronic phase of infection, Deng et al. (2015) demonstrated that the use of random *gag* epitopes did not stimulate CTLs to kill the reactivated autologous infected CD4+ T cells (26). This showed that the latent reservoirs of the chronic-treated patients predominantly consisted of CTL escape variants (26). It would therefore be necessary to investigate the level of *gag* CTL escape in the viral blip, in order to determine, which *gag* epitopes remain conserved, and would be ideal candidates, for therapeutic vaccine design.

1.5.2 Broadly neutralizing antibody escape

The HIV envelope (*env*) is critical in the establishment of the initial infection, as its gp120 subunit is the first site of virus-host attachment pre-infection, which is then followed by membrane fusion and viral invasion (42, 43). It is estimated that 1 in every 4 newly infected people, develops strain-specific neutralizing antibodies within the first 2-3 years of infection (43-45). These antibodies target specific *env* sites: V1V2, V3, gp120-gp41 interface, MPER and CD4-binding site, thus inhibiting infection (43, 44, 46, 47). Notwithstanding the potential for the development of neutralizing antibody-based vaccines, viral escape remains a considerable obstacle (43, 44).

Previous studies have shown that while seroconversion antibodies are highly strain specific to the autologous viremia, the initial homologous transmitter/founder (T/F) virus population diversifies over time, resulting in the antibody-based selection of resistant *env* variants (43, 44). The serial coevolution of antibodies and antibody-resistant HIV-1 variants in response to viral escape, yields broadly neutralizing antibodies (bnAbs) (46, 48). While bnAbs are incapable of neutralizing HIV infection in the individual that produces them, they may be potentially used by others in conjunction with ART in order to address the occurrence of intermittent viral blips while on ART (46, 48). These bnAbs may be potentially used as vaccines, to elicit an immune response against rebound viruses (42).Therefore, it is necessary to determine whether rebound viruses are susceptible to broadly neutralizing antibodies.

The polyvalent HIV-1 mosaic antigen strategy is a recent example of vaccination design, in which viral proteins with optimized epitopes were used to elicit a broad, potent, immune response. There was however, no notable reduction in acquisition or viral load at infection, owing to some evidence of pre-seroconversion resistance and the variability of *env* (49, 50). A study by Mendoza *et al.* (2018) showed how patients remained virally suppressed, when bnAbs were administered following a fortnight analytic treatment interruption. This study illustrated the potential of inducing CTL responses and using the vaccine effect of bnAbs to neutralize the rebound virus. It is therefore important to characterize HIV-1 *env*, in order to determine the relationship between genetic variations in antibody-binding epitopes and sensitivity to bnAbs.

1.6 Problem statement

Individuals on HIV ART frequently experience intermittent viremia (viral blips), however the genetic nature and source of these rebounding viruses while on suppressive ART remains unclear. The study of these genetic characteristics would be essential in the development of targeted vaccines and adjunct treatments.

1.7 Hypothesis

In our study it was hypothesized that the viral blip would be genetically distinct from the T/F virus that establishes the infection as well as the PreART virus.

1.8 Aim

We aimed to determine the genetic characteristics of the viral blips which occurred in patients following a period of suppressive HIV ART.

1.9 Objectives

- To determine whether the blip virus was genetically distinct from the T/F and PreART viruses
- To investigate whether the blip virus develops CTL escape mutations
- To determine the development of HIV drug resistance at the time of infection (T/F), before initiation of HIV treatment (PreART) and at the time when a blip
- To determine whether the blip virus develops any mutations associated with bnAb escape

1.10 Thesis structure

This thesis is structured such that chapter 1 is an overview of the literature, research questions and objectives of the study. The methods used to achieve these objectives are detailed in chapter 2, followed by a description of the findings of the study in chapter 3. Chapter 4 contains the discussion and conclusion of the results as they relate to the literature.

CHAPTER 2: METHODOLOGY

2.1 Study participants

The stored blood plasma samples of HIV-1 subtype C acutely infected donors for this investigation, were obtained from the Females Rising Through Education, Support and Health (FRESH) cohort study (2012 to present), which enrolled high risk females aged between 18 and 23 years, from the Durban township of Umlazi, South Africa, from 2012 to present (51). Participants received HIV-1 RNA testing twice weekly during a 48–96-week life-skills empowerment programme and blood samples were collected prior to infection and post infection. Those who acquired infection were recruited and followed longitudinally. The blood plasma samples of four participants who experienced viral rebound following suppressive ART were used in this study. Viral blips were defined as viral loads of 2,000 copies/mL or more, following a period of more than 150 days of viral suppression on ART. The viral load criterion for a blip was selected based on the WHO virological failure guidelines and the ease of viral DNA amplification. All the protocols used were reviewed by the University of KwaZulu-Natal (South Africa) Biomedical Research Ethics Committee (UKZN BREC), reference #BREC/00001445/2020.

2.2 Viral RNA extraction

Viral RNA was extracted from plasma using the QIAamp viral RNA Minikit Spin protocol (Qiagen, Valencia, CA) as per the manufacturer's instructions. Briefly, 500 µL plasma was concentrated by centrifuging at 14 000 x g at 4°C for 120 min, discarding 360 µL supernatant and resuspending the pellet. The concentrated plasma (140 µL) was then added to a microcentrifuge tube containing 560 µL Buffer AVE-AVL-carrier RNA, pulse vortexed and then incubated at room temperature for 10 min. Absolute ethanol (560 µL) was added to the tube followed by pulse vortexing and brief centrifuging. The solution was then added to the QIAamp Mini column in 630 µL aliquots and centrifuged at 6 000 x g at 4°C for 1 min. The collection tube was subsequently replaced following each centrifugation. Buffer AW1 (500 µL) was added to the column and centrifuged at 6 000 x g at 4°C for 1 min. Buffer AW2 (500 µL) was then added to the column and centrifuged at 20 000 x g at 4°C for 3 min. The Mini column was then further centrifuged at 6 000 x g at 4°C for 1 min and transferred to a new 2 mL collection tube. RNA was eluted by adding 70 µL of buffer AVE and centrifuged at 6 000 x g at 4°C for 1 min.

2.3 Amplification of HIV-1 by Single Genome Amplification (SGA)

Single genome amplification is a method used to amplify the virus from a single template of DNA. This is accomplished by performing a limiting dilution of cDNA prior to amplification. Based on the Poisson distribution, there is a high probability that the dilution which yields 30% PCR positivity was amplified from a single cDNA template (52). We first attempted to obtain the near full-length genome (NFLG) of HIV-1 (~9kb) by the amplification of two overlapping fragments (5´ and 3´ HIV genome fragments) SGAs were initially generated from all 4 participants, however due to challenges with amplication, the gene-specific approach was used to generate SGAs from all the participants.

2.3.1 Amplification of the NFLG HIV-1 genome in two overlapping fragments

This method was adapted from literature (53). The first fragment (5´ fragment) was amplified using a forward primer (0682F) (Table 2.1) which binds to the 5´ LTR region and a reverse primer (6352R) which binds to the *vpu* region producing a 5.5kb amplicon. The second overlapping fragment (3´ fragment) had primers that bind and amplify fragments from the *vif* region (5550F) to the 3´ LTR region (9555R) producing a 3.3kb amplicon.

2.3.1.1 Amplification of 5´ fragment of HIV-1 genome

cDNA was synthesized using the SuperScript III kit (Invitrogen, Life Technologies Corporation Carlsbad, USA) as per manufacturers' instructions. The protocol for cDNA synthesis was adapted from literature (53). Briefly, 10 µL of RNA template was added to the first cDNA master mix which contained deionized water, 10mM of deoxynucleoside triphosphate mix and 10 µM of gene-specific primer (6352R) and incubated at 65°C for 5 min then 4°C for 1 min. Thereafter, 7µL of the second master mix containing 5X Strand Buffer, 100mM DTT, 40U/µL RNaseOUT (Recombinant RNase Inhibitor) and 200U/µL Superscript III Reverse Transcriptase was added to the reaction, and incubated as follows: 50°C for 60 min, 55°C for 60 min, followed by 70°C for 15 min. Reverse transcription reaction was then terminated by incubation with 1µL of RNase H at 37°C for 20 min.

Q5 Hot Start High-Fidelity DNA polymerase kit (New England Biolabs, England) was used to amplify the 5´ fragment by nested PCR. The first-round master mix consisted of 5X Q5 Reaction Buffer, 5X Q5 High GC Enhancer, 10 mM of deoxynucleoside triphosphate mix, 0.02 U/µL Q5 Hot Start High-Fidelity DNA polymerase and 10 µM of each primer (0682F and 6352R) (Table 2.1). The first round PCR conditions were as follows: initial denaturation at 98°C for 30 s, 30 cycles of denaturation at 98°C for 10 s, annealing at 66°C for 30 s, and extension at 72 °C for 4.5 min, with a final extension of 72°C for 10 min. Thereafter, the second round PCR was carried out using 1 µL of the first-round product, 5X Q5 Reaction Buffer, 5X Q5 High GC Enhancer, 10mM of deoxynucleoside triphosphate mix, 0.02 U/µL Q5 Hot Start High-Fidelity DNA polymerase and 10 µM of each primer 0776F & 6231R. The second

round PCR conditions were identical to those used for the first round PCR reaction, except for the extension step, which was for 4 min at 72 °C.

2.3.1.2 Amplification of 3´ fragment of HIV-1 genome

cDNA was synthesized using the SuperScript III kit (Invitrogen, Life Technologies Corporation Carlsbad, USA) as per manufacturer's instructions. The protocol for cDNA synthesis was adapted from literature (53). Briefly, 10 µL of RNA template was added to the first cDNA master mix which contained deionized water, 10mM of deoxynucleoside triphosphate mix and 10 µM of gene-specific primer (OFM19) and incubated at 65°C for 5 min then 4°C for 1 min. Thereafter, 7µL of the second master mix containing 5X Strand Buffer, 100mM DTT, 40U/µL RNaseOUT (Recombinant RNase Inhibitor) and 200U/µL Superscript III Reverse Transcriptase was added to the reaction, and incubated as follows: 50°C for 60 min, 55°C for 60 min, followed by 70°C for 15 min. Reverse transcription reaction was then terminated by incubation with 1µL of RNase H at 37°C for 20 min. All the primers used are listed in Table 2.1.

The 3´ fragment was amplified by nested PCR. The first round PCR was carried out with 1µL cDNA template and a master mix containing 5X Q5 Reaction Buffer, 5X Q5 High GC Enhancer, 10mM of deoxynucleoside triphosphate mix, 10µM of each primer (5550F & 9555R) (Table 2.1) and 0.02 U/µL Q5 Hot Start High-Fidelity DNA polymerase (New England Biolabs, England). The first round PCR conditions were as follows: initial denaturation at 98°C for 30 s, 30 cycles of denaturation at 98°C for 10 s, annealing at 62°C for 30 s and extension at 72 °C for 4 min, and final extension at 72°C for 10 min. Second round PCR was carried out using 1 µL first round product, 5X Q5 Reaction Buffer, 5X Q5 High GC Enhancer, 10mM of deoxynucleoside triphosphate mix, 10µM of each primer (5861F & 9555R) and 0.02 U/µL Q5 Hot Start High-Fidelity DNA polymerase. The second round PCR conditions were to the same as the first round except for an annealing temperature of 64°C for 30 s.

2.3.2 Amplification of HIV-1 Gag by SGA

Briefly, the first cDNA master mix was carried out with 10 µM of a gene-specific primer (GagD_R) (Table 2.1) (54), water and 10mM of deoxynucleoside triphosphate mix, combined with 10 µL template RNA. The reaction mix was heated to 65°C for 5 min and then incubated at 4ºC. Thereafter, 16 µL of the second cDNA master mix was added [5X First Strand buffer, 100mM of DTT, RNaseOUT (Recombinant RNase Inhibitor) and 200U/µL Superscript III Reverse Transcriptase (Invitrogen, Life Technologies Corporation Carlsbad, USA). The reverse transcription conditions used were 55°C for 60 min, followed by 70°C for 15 min and a final removal of remaining RNA by adding 2 µL of RNase H and incubating at 37°C for 20min.

The first round PCR reaction was carried out using with 1 µL cDNA template, dH_2O, 10X Buffer I (incl. 15mM $MgCI_2$), 10mM of deoxynucleoside triphosphate mix, 10pmol/µL of each primer (GagD_F and GagD_R) (Table 2.1) and 5U/µL Expand High Fidelity Polymerase (Thermo-Scientific, Carlsbad, USA). The first round PCR conditions were as follows: initial denaturation at 95°C for 2 min, 10 cycles of denaturation at 95°C for 15 s, annealing at 52°C for 30 s and extension at 72 °C for 1 min; 15 cycles of denaturation at 95°C for 15 s, annealing at 55°C for 30 s and extension at 72 °C for 1.5 min, and final extension at 72°C for 7 min. The second round PCR reaction was carried out using with 1 µL first round DNA template, dH_2O, 10X Buffer I (incl. 15mM MgCI2), 10mM of deoxynucleoside triphosphate mix, 10pmol/µL of each primer (GagA_F and GagC_R) (Table 2.1) and 5U/µL Expand High Fidelity Polymerase. The second-round cycling conditions were as follows: initial denaturation at 95°C for 2 min, 35 cycles of denaturation at 95°C for 15 s, annealing at 55°C for 30 s and extension at 72 °C for 1.5 min, and final extension at 72°C for 7 min.

2.3.3 Amplification of HIV-1 Env by SGA

Briefly, the first cDNA master mix was carried out with 10 µM of a gene-specific primer OFM19 (Table 2.1) (52), water and 10mM of deoxynucleoside triphosphate mix, combined with 10 µL template RNA. The reaction mix was heated to 65°C for 5 min and then incubated at 4ºC for 1 min. Thereafter, 7 µL of the second cDNA master mix was added to the reaction, [5X First Strand buffer, 100mM of DTT, RNaseOUT (Recombinant RNase Inhibitor) and 200U/ul Superscript III Reverse Transcriptase (Invitrogen, Life Technologies Corporation Carlsbad, USA)]. The reverse transcription reaction conditions used were 50°C for 60 min, 55°C for 60 min followed by 70°C for 15 min and a final removal of remaining RNA by adding 1 µL of RNase H and incubating at 37°C for 20 min.

The first round of nested PCR was performed by adding 1 µL cDNA to the first-round master mix which contained dH_2O, 10X Buffer, 15mM $MgSO_4$, 10mM of deoxynucleoside triphosphate mix, 10pmol/µL of each primer (OFM19 and VIF1) (Table 2.1) and 5 U/µL Platinum High Fidelity DNA Polymerase (Invitrogen, Life Technologies Corporation Carlsbad, USA). The first round PCR conditions were as follows: initial denaturation at 94°C for 4 min, 35 cycles of denaturation at 94°C for 15 s; annealing at 55°C for 30 s and extension at 68 °C for 4 min, and final extension at 68°C for 20 min. The second round PCR was carried out by adding 2.5 µL first round product to second round PCR master mix, which contained dH_2O, 5X Phusion Buffer (New England Biolabs, England), 10 mM of deoxynucleoside triphosphate mix, 10pmol/µL of each primer (Env1A and Env1M) (Table 2.1) and 5 U/µL Phusion High Fidelity Taq Polymerase (New England Biolabs, England). The second round PCR conditions were as follows: initial denaturation at 98°C for 30 s, 35 cycles of denaturation at 98°C for 10 s, annealing at 65°C for 30 s and extension at 72 °C for 4 min, and final extension at 72°C for 10 min.

2.3.4 Amplification of the protease and reverse transcriptase regions of the HIV-1 pol gene

The reverse transcriptase and protease regions of the *pol* gene (~1.6 kb) were amplified using the HIV-1 Genotyping Kit: Amplification Module (Applied Biosystems, Texas, USA) as per manufacturer's manual. RNA was denatured by heating at 65°C for 10 min, then immediately incubating the reaction at 4°C for 3 min. The Reverse Transcriptase PCR (RT-PCR) was carried out by adding 40 µL of the RT-PCR master mix which contained RT-PCR master mix and SuperScript III One-Step RT-PCR with Platinum Taq High Fidelity enzyme (Thermo-Scientific, Carlsbad, USA). The RT-PCR cycling conditions were as follows: reverse transcription at 50°C for 45 min, enzyme inactivation at 94°C for 2 min, 40 cycles of denaturation 94°C for 15 s: annealing at 50°C for 20 s and extension at 72°C for 2 min, and final extension at 72°C for 10 min.

Thereafter, nested PCR was carried out by combining 2 µL of RT-PCR products with nested PCR master mix and AmpliTaq Gold LD DNA Polymerase (Thermo-Scientific, Carlsbad, USA). The nested PCR conditions were as follows: initial denaturation at 94°C for 4 min, 40 cycles of denaturation 94°C for 15 s: annealing at 55°C for 20 s and extension at 72°C for 2 min, and final extension at 72°C for 10 min.

2.4 Electrophoresis

PCR products were visualized using a 1% agarose gel electrophoresis (Bioline, Atlanta, USA) and O'GeneRuler 1 kb DNA ladder (Invitrogen™, Carlsbad, USA). One percent agarose gels were made by dissolving a 50 g agarose tablet in 50 mL of TBE buffer. PCR products (3 µL) were diluted with 2 µL gel loading buffer (Sigma-Aldrich, Missouri, USA) containing a 1 in 50 dilution of GelRed nucleic acid gel stain (Biotium, California, USA) before loading. The Chemidoc™ imaging system and Image Lab™ software version 5.0 (BioRad, California, USA) were used to visualize the gel, and resulting amplicons of the correct band size were stored at -20ºC prior to sequencing.

2.5 Sanger sequencing

All *gag* and *env* SGAs generated were sequenced using ABI Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Massachusetts, USA) in 96-well plates. Each well contained the following: 0.4 µL Terminator Ready Reaction mix, 2.0 µL 5X Sequencing buffer, 2 µL 2 pM primer, 1 µL (1:10 dilution) template and 4.6 µL of deionised water. Sequencing conditions were as follows: initiation denaturation at 96°C for 1 min, 25 cycles of: denaturation at 96°C for 10s, annealing at 50°C for 5 s and extension 60°C for 4 min. Primers used for sequencing Gag and Env are shown in Table 2.1.

The reverse transcriptase and protease regions of the *pol* gene were sequenced using the HIV-1 Genotyping Kit: Cycle Sequencing Module (Applied Biosystems, Texas, USA) as per manufacturer's manual. Briefly, the sequencing reaction consisted of 18 µL of pre-made sequencing master mix [containing one of six primers (F1, F2, F3, R1, R2, R3)] and 2 µL of nested PCR product. The sequencing reaction conditions were as follows: 25 cycles of: denaturation at 96°C for 10s, annealing at 50°C for 5 s and extension 60°C for 4 min.

The sequencing clean-up was carried out as follows: 1 µL of 125 mM ethylenediaminetetraacetic acid (EDTA) pH 8.0 was added to each well and mixed by pipetting. A solution of sodium acetate (NaOAc) (26 µL) and 1 µL of 100% ethanol was then added to each well. The plate was then sealed, briefly vortexed and centrifuged at 3000 x g for 20 min. The supernatant was removed by inverting the plate on a paper towel and centrifuging at 150 x g for 5 min. The pellet was washed by adding 35µl of fresh, cold 70% ethanol, followed by centrifugation at 3000 x g for 5 min to wash the pellet. The residual supernatant was removed by centrifugation at 150 x g for 1 min and the plate was dried at 50°C for 5 min and stored at -20°C prior to sequencing.

2.6 Sequencing and phylogenetic analysis

The Sequencher software program version 5.3 (Gene Codes Corporation, Ann Arbor, USA) was used to assemble and manually edit the overlapping DNA fragments. Assembly parameters of a minimum 65% match and 13-base overlap were used. Sequences containing a mixture of variants were discarded. Geneious v11.1.5 (Biomatters, New Zealand) was used to perform pairwise sequence alignment for 85 *gag*, 88 *env* and 10 *pol* sequences. A South African subtype C reference sequence SK164B1 (Accession number: AY772699) obtained from the Los Alamos HIV sequence database was used because the study participants were from South Africa. Phylogenetic analysis was performed using the Mega version X software. Maximum-likelihood trees were automatically generated for *gag* and *env* using the HyPhy packaged software. The level of genetic diversity within gag and env sequences was calculated based on pairwise amino acid distances (55). Highlighter plots were constructed using the Los Alamos Database software in order to identify silent and non-silent mutations in DNA sequences (http://www.hiv.lanl.gov/). The identified non-silent mutations were then studied further for their significance as pertains to CTL escape or bnAb escape.

2.7 Cytotoxic T lymphocyte (CTL) escape analysis

The Gag CTL epitopes of each participant at T/F and blip were identified using the A-list published in the Los Alamos HIV immunology database (http://www.hiv.lanl.gov/content/immunology/variants/ctl_variant.html). The five amino acids flanking either side of each epitope were included in the analysis. The HXB2 subtype B reference sequence (Accession number: K03455), was used as a standard in identifying epitope positions in the participants' sequences and ensuring that codons in the variable regions were properly allocated to sequence positions. The sequences were aligned using CLUSTAL OMEGA alignment in the Geneious v11.1.5 software. Epitopes which were identical to the A-list were denoted as consensus and those which had developed mutations were denoted as variants. A percentage of consensus versus variant epitopes was calculated for each participant's T/F and blip epitopes, respectively. A CTL/CD8+ epitope variant and escape mutation list (http://www.hiv.lanl.gov/) was then used to further identify which variants were actually known CTL escape mutations. The HLA alleles of each participant were previously determined by next-generation sequencing (39, 40). The percentage of known CTL escape versus other variants was then calculated for each participant's T/F and blip epitopes, respectively. The respective percentages of CTL escape associated with the chronic and acute-treated participants' HLA alleles, were also enumerated.

2.8 Env antibody escape mutation analysis

A list of epitopes and escape mutations associated with broadly neutralising antibody sensitivity within the following *env* regions: MPER, V1V2 loop, V3 loop, CD4 binding site and gp120/gp41, was adapted from literature (43, 44, 46, 47). Shown in Table 2.2 is a list of the unique positions which were examined within each epitope. SGAs were analysed to determine whether they exhibited unique mutations associated with bnAb escape within these epitopes, at T/F, PreART and blip.

2.9 Drug resistance analysis

The pol amplicons generated by bulk sequencing were assembled and edited using the Sequencher software program version 5.3 (Gene Codes Corporation, Ann Arbor, USA). Amino acids codons containing ambiguities were noted. Geneious v11.1.5 (Biomatters, New Zealand) was used to perform pairwise sequence alignment. The sequences were automatically scanned and analysed for known ART drug resistance using a Stanford HIV drug resistance tool.

Table 2.1 List of primers used in this study

PCR system	Amp. step	Primer	Sequence (5'-3')	Position (HXB2)	Product Size (bp)	Ref.
Near full length HIV-1 genome in 2-overlapping fragments	(5`)1st PCR	0682F	TCT CTC GAC GCA GGA CTC GGC TTG CTG ACT ACT GGT	0682-0708	5 670	(53)
		6352R	ACC CCA TAA TAG ACT GTR ACC CAC AA	6352-6324		
	2nd PCR	0776F	CTA GAA GGA GAG AGA GAT GGG TGC GAG	0776-0800	5 545	
		6321R	CTC TCA TTG CCA CTG TCT TCT GCT C	6321-6207		
	(3`)1st PCR	5550F	GGG TTT ATT ACA GGG ACA GCA GAG	5550-5574	4 005	
		9555R	ACT ACT TAG AGC ACT CAA GGC AAG CTT TAT TG	9555-9533		
	2nd PCR	5861F	GCA AAA CTA CTC TGG AAA GGT GAA GGG	5861-5884	3 771	
		OFM19	GCA CTC AAG GCA AGC TTT ATT GAG GCT TA	9632-9604		
Env	1st PCR	OFM19	GCA CTC AAG GCA AGC TTT ATT GAG GCT TA	9632-9604	4 732	(52)
		VIF1	GGG TTT ATT ACA GGG ACA GCA GA	4900-4923		
		Env1A	CAC CGG CTT AGG CAT CTC CTA TGG CAG GAA GAA	5954-5982	3 142	
	2nd PCR	Env1M	TAG CCC TTC CAG TCC CCC CTT TTC TTT TA	9096-9068		
	Sequencing	sq6rc2	GAA TTG GGT CAA AAG AGA CCT TTG GA	6839-6864	2 892	(52)
		ef00	AAA GAG CAG AAG ACA GTG GCA ATG A	6204-6228		
		bstq2+	CCT CAG CCA TTA CAC AGG CCT GTC CAA AG	6817-6845		
		sq5.5rc	CTA GGA GCT GTT GAT CCT TTA GGT AT	7979-8004		
		sq13f(2)c	TAT ATA AAT ATA AAG TGG TAG AAA TTA AGC	7672-7701		
		sq14fc	ACT CAC GGT CTG GGG CAT TA	7925-7944		
		sq3r(2)c	GCT ATG GTA TCA AGC GAC TAA TAG CAC TC	8651-8680		
		env1M	TAG CCC TTC CAG TCC CCC CTT TTC TTT TA	9096-9068		
Gag	1st PCR	Gag D_F	TCT CTA GCA GTG GCG CCC G	626-644	1 776	(54)
		Gag D_R	AAT TCC TCC TAT CAT TTT TGG	2382-2402		
	2nd PCR	Gag A_F	CTC TCG ACG CAG GAC TCG GCT T	683-704	1 673	
		Gag C_R	TCT TCT AAT ACT GTA TCA TCT GC	2334-2356		
	Sequencing	Gag A_F	CTC TCG ACG CAG GAC TCG GCT T	683-704	1 673	(54)
		Gag A_R	ACA TGG GTA TCA CTT CTG GGC T	1282-1303		
		Gag B_F	CCA TAT CAC CTA GAA CTT TGA AT	1226-1246		
		Gag B_R	CTC CCT GAC ATG CTG TCA TCA T	1825-1846		
		Gag C_F	CCT TGT TGG TCC AAA ATG CGA	1748-1768		
		Gag C_R	TCT TCT AAT ACT GTA TCA TCT GC	2334-2356		

Table 2.2 List of Env epitope positions associated with sensitivity to bnAb

Env epitope positions				
MPER*	**gp120/gp41**	**V3**	**V1V2**	**cD4 binding site**
N671	N88A	L165	N130	97
W672L	230	295	156	99
F673L	234	297	160	197
674	241	300	161	N234
676	262	301	162	262
N677	276	304	163	275
W680G	279	307	164	N276D
N683R	448	323	L165	277
	N611A	325	166	T278A
	N637A	328	167	N279D
		330	168	280
		332	169	281
		334	Q170K	282
		337	171	364
		339	172	365
		340	173	D368A
		341	177	E370A
		343	181	371
		344	187	372
			197	425
			332	429
				455
				R456W
				457
				G458Y
				459
				461
				462
				463
				471
				D474
				476

*each column contains a list of known bnAb epitope positions (43, 44, 46, 47)

CHAPTER 3: RESULTS

3.1 Study participants

The young women enrolled in the FRESH cohort were from the high-risk area of Umlazi, Durban these participants were aged 18-23 years, HIV-negative, sexually active, not anaemic (haemoglobin ≥10 g/L), pregnant or chronically ill, unemployed, and not attending school. The stored blood plasma samples used in this study were obtained from four HIV-1 subtype C infected participants from the FRESH Acute Infection Cohort. These participants had acquired HIV-1 subtype C infection, initiated treatment, and were virally suppressed for at least 170 days (median 584 days), before experiencing a viral blip. The criteria used to identify the viral rebound or blip was a viral load of ≥2 000 copies/mL following a period of viral suppression (≤20 copies/mL). Sample availability and ease of amplification influenced the sample size and selection criterion. The clinical characteristics of the selected participants are shown in Table 3.1.

Table 3.1 Clinical characteristics of FRESH study participants

Participant ID	Fiebig stage[a]	Treatment phase[b]	Peak viremia[c] (cps/mL)	Viral load at the blip time point (cps/mL)	Length of viral suppression (days)
12733-0048-**036**[d]	II	Chronic	57,000,000	12,000	747
12733-0387-**272**	II	Chronic	100,000,000	13,000	266
12733-1700-**1284**	I	Acute	11,000	32,000	233
12733-1825-**1388**	II	Acute	14,000	9,100	170

[a]Stage of early infection when T/F samples were collected
[b]Phase of infection when ART was initiated
[c]Highest pre-treatment viral load
[d]Particiant ID hereafter denoted by the last set of numbers shown in bold

The first blood plasma samples were collected from the participants during the second Fiebig stage of early HIV infection (Fiebig II), with the exception of PID 1284 whose sample was obtained during Fiebig I (Table 3.1). Due to changes in treatment guidelines over time, two of the participants (PID 036

and 272) initiated treatment in the chronic phase of infection (CD4 count ≤350), while PID 1284 and 1388 were tested and immediately treated in the acute phase of infection. Consequently, the chronic treated participants experienced high viral loads prior to treatment initiation. PID 036 reached a peak pre-treatment viral load of 57 million copies/mL. Treatment was initiated 665 days post first viral load detection and resulted in viral suppression for 747 days, followed by a viral blip of 12,000 copies/mL. PID 272 initiated ART 497 days post detection, having reached a pre-ART peak viremia of 100 million copies/mL. A period of 266 days of viral suppression was then followed by a viral blip of 13,000 copies/mL.

The acute-treated participants experienced relatively lower peak viral loads as compared to the chronic-treated participants, due to early treatment. PID 1284 initiated treatment and reached a peak viremia of 11,000 copies/mL, 1 day post detection. A viral blip of 32,000 copies/mL was preceded by a 233 day of viral suppression. PID 1388 initiated treatment and reached a peak viral load of 14,000 copies/mL, 1 day post detection. This participant was virally suppressed for 170 days before experiencing a viral blip of 9,100 copies/mL. The chronic-treated participants were virally suppressed for longer than the acute-treated participants (median 507 vs 202 days). Three time points were studied in the chronic-treated participants: transmitter/founder (T/F) (~ 3 days post detected seroconversion), pre-treatment (PreART) (latest available sample before treatment initiation date) and viral blip. Only the TF and blip were studied in the acute-treated participants as they were immediately tested and treated. Figure 3.1 shows the viral load trajectory over time for each of the four participants.

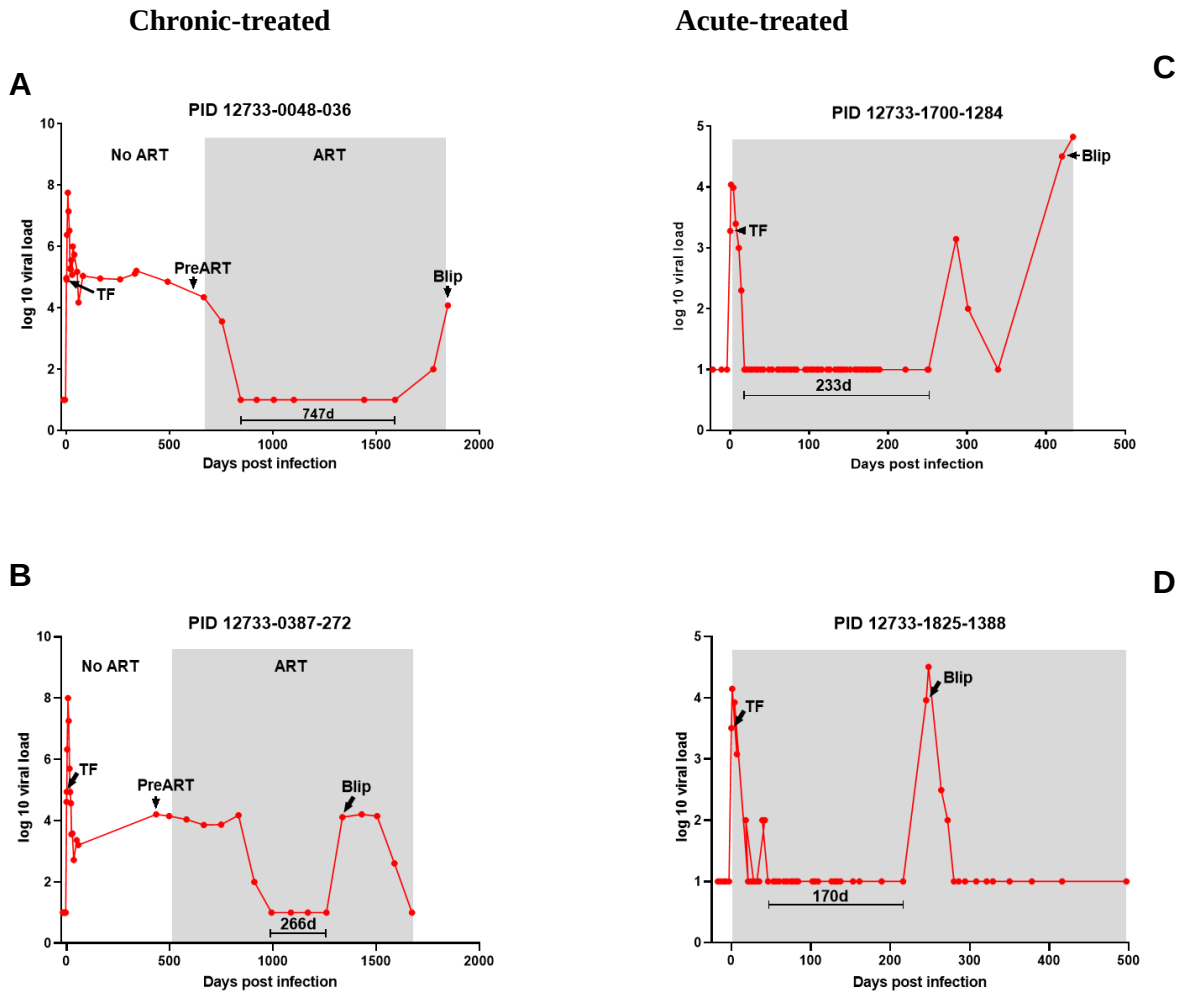


Figure 3.1 Longitudinal clinical data of four participants who experience viral blips following suppressive HIV ART

The log viral load was plotted against days post infection for the chronic-treated participants **(A and B)** and the acute-treated participants **(C and D)**. The grey block indicates the period of time for which the participant had been on ART at the start of this investigation. The transmitted/founder virus which established the infection was denoted as **(TF)**, the rebound virus was denoted as **(blip)** and the virus prior to treatment initiation was denoted as **(PreART)**. The time points indicated by arrows were the time points studied. The number of days (d) for which a participant was virally suppressed are shown.

3.2 Drug resistance testing

HIV treatment interruption and subsequent drug resistance are the most frequent cause of viral rebound. As such, the *protease* and *reverse transcriptase* regions of the *pol* gene were amplified, and bulk sequenced at each of the studied time points in order to determine whether each participant had developed drug resistance. All the participants were initiated on the FDC pill, however only the acute-treated participants also used raltegravir (RAL). This is because while RAL is associated with less side effects and a rapid decrease in viral load, it has been associated with treatment failure in those who had experienced high viral loads (≥100,000 copies/mL) prior to treatment (55).

Shown in Figure 3.2 is a gel image of some of the *pol* amplicons generated by bulk sequencing (~1.6kb). The resulting generated sequences were edited and analysed for drug resistance mutations using the Stanford HIV drug resistance database tool. The results shown in Table 3.2 indicate that the two acute-treated participants, PID 1284 and 1388, did not develop drug resistance against their ART regimen as they remained susceptible to all the drugs in their regimen for both the TF virus as well as the blip-derived virus. Of the chronic-treated participants, PID 036, did not develop drug resistance at any of the time points tested. However, PID 272 developed resistance to the drug efavirenz (EFV) at the blip time point. This was due to a K103N non-polymorphic mutation associated with NNRTI drugs such as EFV and may have contributed to the occurrence of the viral blip. This mutation was not present in the TF virus and was also not identified at the time point prior to ART initiation suggesting that this mutation developed over time.

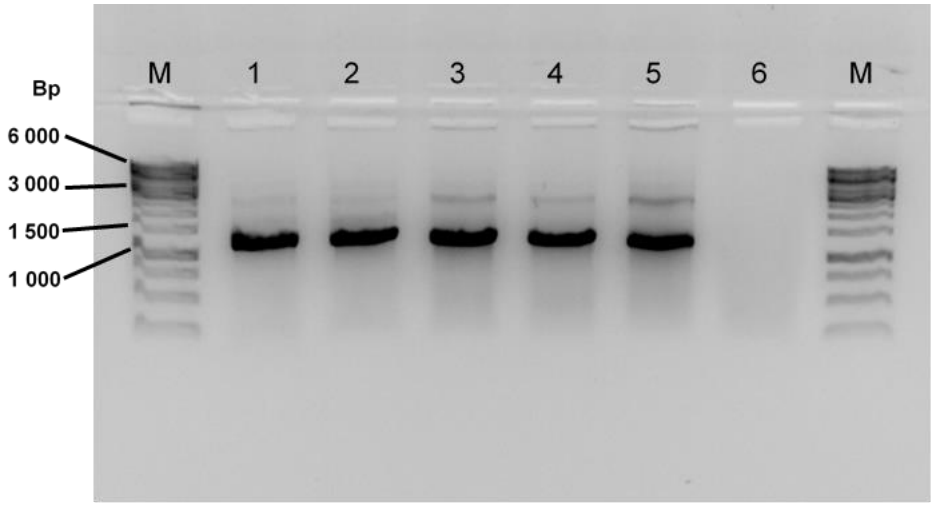


Figure 3.2 Gel image showing *pol* PCR amplicons

Gel electrophoresis image of PCR products generated by amplifying *pol* at the blip time point of PIDs 036, 272, 1284 and 1388. Lane **(M)** is a 1kb DNA ladder, **(1-4**) represents *pol* amplicons generated (~1.6kb) for PIDs 036, 272, 1284 and 1388, respectively, (5) is a positive control and (6) is the negative control.

Table 3.2 Participant susceptibility to antiretroviral treatment over time

PID	Time point	ART regimen	Susceptibility to ART:				Mutation
			TDF[b]	FTC[c]	EFV[d]	RAL[e]	
036	TF	FDC Pill[a]	√	√	√		
	PreART		√	√	√		
	Blip		√	√	√		
272	TF	FDC Pill	√	√	√		
	PreART		√	√	√		
	Blip		√	√	×		K103N
1284	TF	FDC + RAL	√	√	√	√	
	Blip		√	√	√	√	
1388	TF	FDC + RAL	√	√	√	√	
	Blip		√	√	√	√	

[a]fixed-dose combination pill contains: TDF, FTC, EFV
[b]tenofovir
[c]emtricitabine
[d]efavirenz
[e]raltegravir
√susceptibility to a drug
×drug resistance

3.3 Amplification of HIV-1 by Single Genome Amplification (SGA)

In order to determine the level of viral diversity at the blip time point, amplicons were generated by SGA. The generation of SGAs increased the probability of amplifying a wider variety of the variants circulating at each time point. The T/F, PreART and blip SGAs were generated for the chronic-treated participants and the T/F and blip SGAs were generated for the acute-treated participants. Second round PCR products were visualised by electrophoresis on a 1% agarose gel and 30% positive PCR bands were identified as SGAs. Sequenced SGAs from each participant's blip time point were compared to the SGA sequences generated from the earlier time points.

In the first approach, NFLGs were generated in two fragments in order to amplify *gag* and *pol* (5' fragment) and *env* (3' fragment). Using this approach, we amplified ten 5' fragments and six 3' fragments as SGAs from the blip time point for PID 036. For PID 272, one 5' fragment and six 3' fragments of T/F SGAs were generated, as well as four 3' blip SGAs. Six 5' and twenty 3' T/F SGAs were amplified for PID 1388. However, despite several attempts, further attempts to generate more SGAs for all participants at each time point using this approach were futile. Notwithstanding these challenges, 7 full *gag*, 7 partial *pol* and 4 full *env* SGAs were successfully generated and sequenced for PID 1284 at the blip time point and 3 full *env* at T/F. Shown in Figure 3.3A and B are representative gels of SGAs generated by the first approach, for PID 1284 blip (5' fragment) and PID 272 blip (3' fragment), respectively.

Due to the challenges of the first approach, a second approach was employed, in which *gag* and *env* SGAs were individually amplified using gene-specific protocols. Representative gel images of the SGAs generated by the second approach are shown in Figure 3.4, for PID 1284 T/F *gag* (A) and PID 272 blip *env* (B), respectively. Table 3.2 shows the total number of Gag and Env SGAs generated and sequenced per PID at each time point. The following SGAs were generated and successfully sequenced for PID 036 as follows: 6 *gag* and 19 *env* SGAs were amplified from the earliest collected positive sample denoted as the T/F. Two *gag* and 7 *env* SGAs were generated from the sample collected prior to treatment initiation denoted as PreART. At the viral blip time point, 9 *gag* and 7 *env* SGAs were generated. For PID 272, 12 *gag* and 21 *env* SGAs were generated and sequenced from the T/F sample. Six *gag* and 4 *env* SGAs were amplified from the PreART sample while 9 *gag* and 5 *env* SGAs were amplified from the viral blip sample. For PID 1284, amplicons were generated using both the gene-specific and NFLG approach as follows: 7 *gag* SGAs were generated using the first approach (5' fragment), whereas 13 *env* SGAs were generated using the gene-specific approach from the viral blip sample. From the T/F sample of PID 1284, 12 *gag* SGAs were amplified using the gene-specific approach and 3 *env* SGAs were generated using the first approach (3' fragment) at T/F. For PID 1388,

14 *gag* and 3 *env* SGAs were generated from the T/F sample. At the blip time point, 5 *gag* and 8 *env* SGAs were generated.

Table 3.3 Gag and Env SGA sequences generated per participant over time

PID	Time point	Viral load (copies/mL)	Gag SGAs	Env SGAs
036	T/F	2,400,000	6	19
	PreART	22,000	2	7
	Blip	12,000	9	7
272	T/F	87,000	12	21
	PreART	16,000	6	4
	Blip	13,000	9	5
1284	T/F	11,000	12	3
	Blip	32,000	7	13
1388	T/F	3,200	14	3
	Blip	9,100	5	8

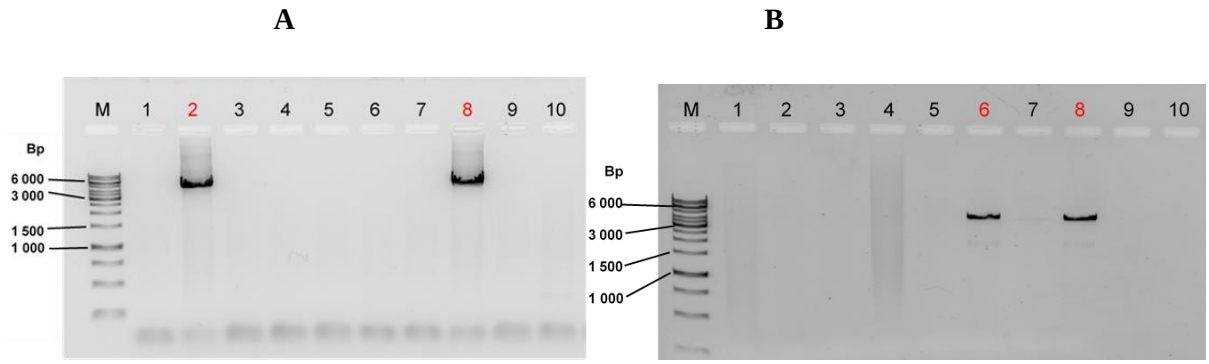


Figure 3.3 Gel image showing half genome single genome amplicons

Gel electrophoresis images of PCR products generated by amplifying near full-length HIV-1 genomes using the two overlapping fragments approach to obtain a 5' fragment and a 3' fragment of the genome. In both images, the empty lanes indicate negative PCR results and Lane **(M)** is a 1kb DNA ladder. **(A)** shows 2 positive amplicons in lanes 2 and 8 (5' fragments ~5.5kb representing the region 5' LTR to *vpu*), **(B)** shows 2 positive amplicons in lanes 6 and 8 (3' fragments ~3.3kb representing the *vif* to 3' LTR region).

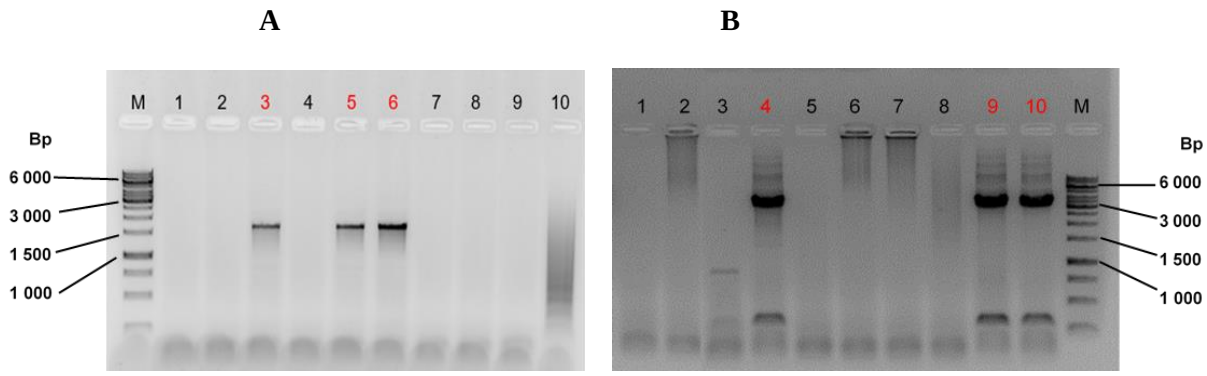


Figure 3.4 Gel image *gag* and *env* single genome amplicons

Gel electrophoresis images of PCR products generated by amplifying *gag* and *env* SGAs using a gene-specific approach. In both images, the empty lanes indicate negative PCR results and Lane **(M)** is a 1kb DNA ladder. **(A)** shows 3 positive *gag* amplicons in lanes 3, 5 and 6 (~1.5kb), **(B)** shows 3 positive *env* amplicons in lanes 4, 9 and 10 (~3kb).

3.4 HIV-1 sequence analysis

3.4.1 Phylogenetic relatedness of Gag-derived sequences

Phylogenetic analysis was conducted using the MEGA vX software to draw maximum-likelihood trees. The trees were constructed using the *gag* SGA sequences generated from all four participants, at the T/F, PreART and blip time points, in order to determine whether there was a change in viral diversity between time points. We constructed two phylogenetic trees to distinguish between the acute-treated participants and the chronic-treated participants (Figure 3.5 and 3.6). Highlighter plots were also drawn using the Highlighter tool from the Los Alamos database in order to determine whether the genetic differences observed in the phylogenetic tree translated to silent or non-silent mutations (Figures 3.5 and 3.6). A participant-specific consensus T/F sequence was generated and used as a reference in the identification of mutations at T/F, PreART and blip.

All the HIV-1 *gag* sequences generated from the two acute-treated participants 1284 and 1388, clustered with the subtype C reference sequence, indicating that both participants were infected with an HIV-1 subtype C virus. For PID 1284, the T/F sequences generated were identical to each other, with the exception of 3 sequences which each had 1 mutation. Only 1 of these mutations was non-silent, indicating an amino acid change. Of the 7 blip sequences generated, 3 were identical to the T/F, 2 had the same silent mutation (mutation which does not result in an amino acid change) and the other 2 had an identical silent mutation. The T/F sequences of PID 1388 were genetically identical with the exception of 1 sequence which had a silent mutation. The blip sequences were identical to the T/F sequences with the exception of 1 sequence which had a silent mutation which was different to that

observed at T/F. Overall, it was observed that PID 1388 had homogeneous T/F and viral blip populations. The T/F and blip populations of PID 1284 were generally homogeneous with the exception of a non-silent mutation at each time point.

The HIV-1 *gag* sequences of both chronic-treated participants (PIDs 036 and 272), clustered with the subtype C reference sequence, indicating that both participants were infected with subtype C viruses. The identical PID 272 T/F sequences were indicative of a homogeneous viral population. The T/F sequences clustered separately from the PreART and blip populations which intermingled together. The PreART and blip sequences were observed to have both silent and non-silent mutations, however, no two sequences were identical. Out of the 6 common mutations which were maintained from PreART to blip, only 1 was a non-silent mutation. For PID 036, the T/F viral population was homogeneous as indicated by the sequences which were identical to each other, and clustered separately from the PreART and blip. The PreART and blip viral populations did not intermingle, despite having 3 silent and 4 non-silent mutations maintained over time. The significance of this observation however could not be verified due to the small number of PreART (n=2) sequences analysed.

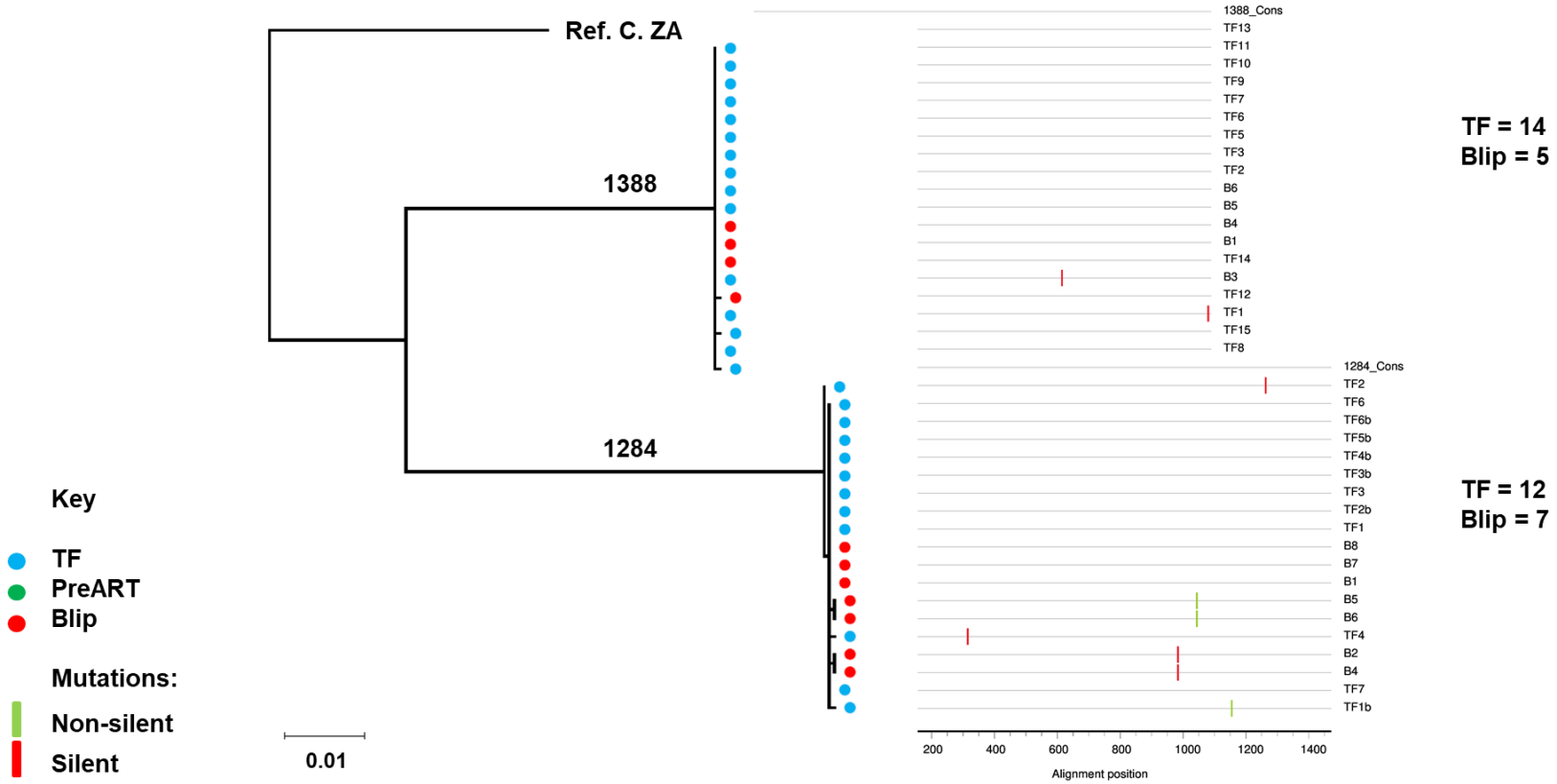


Figure 3.5 Phylogenetic analysis of acute-treated participants' Gag sequences

Maximum-likelihood tree showing *gag* SGA sequences (left) from two acute-treated participants, (PID 1388 and 1284) and the corresponding highlighter plots indicating the silent and non-silent mutations (right). The South African subtype C reference sequence (accession no: AY772699) was used for this study as the participants were from South Africa. The reference sequence and the highlighter plot software used were from the Los Alamos database. The participant ID numbers are indicated on the corresponding tree branch, each sequence was assigned a colour-coded circle to indicate its timepoint; T/F (blue), blip (red). In the highlighter plots all sequences were compared to the consensus sequence (cons) of all T/F, blank grey lines were identical to the consensus, silent and non-silent mutations were denoted as red and green ticks, respectively. The number of sequences per timepoint is indicated to the right of each highlighter plot.

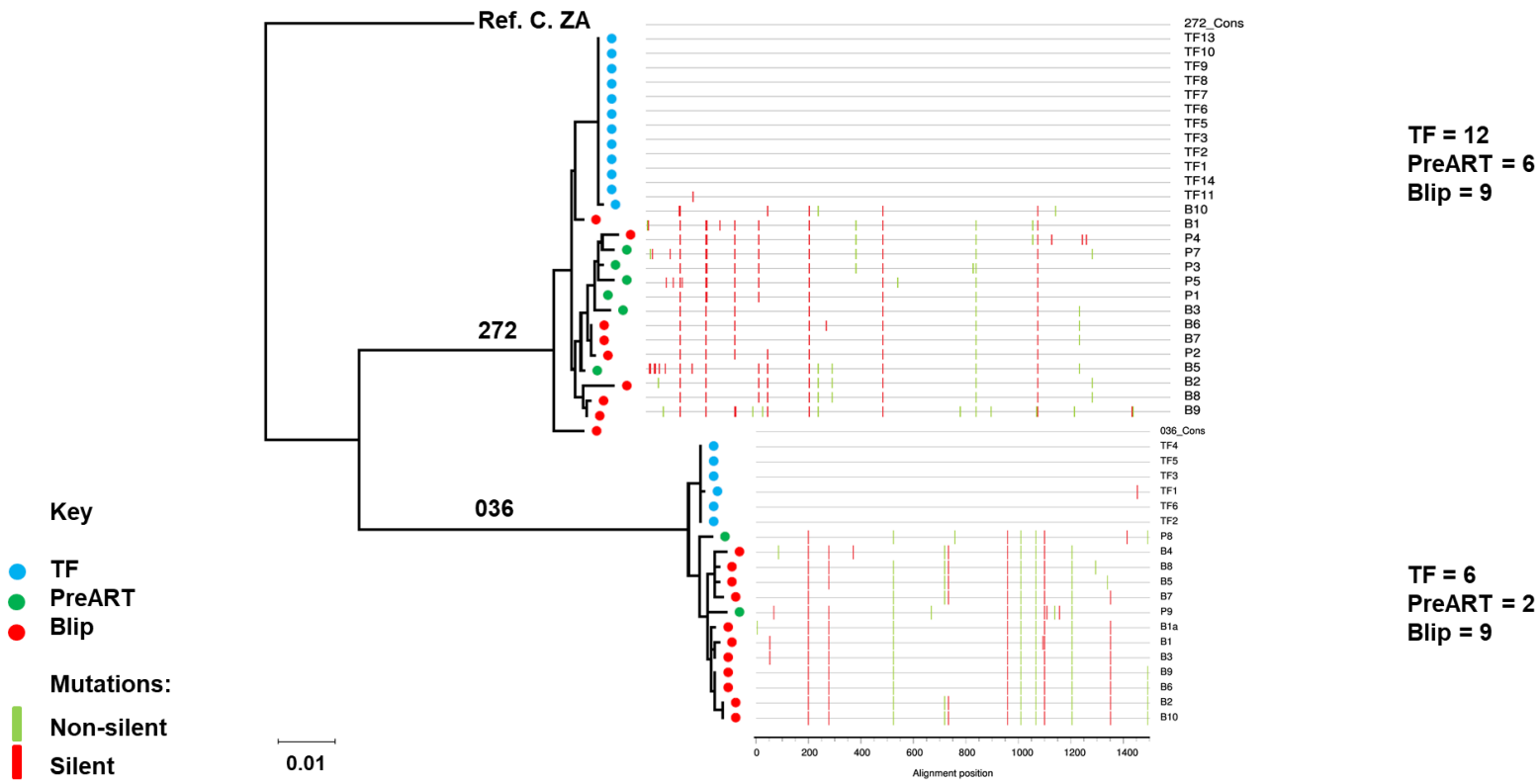


Figure 3.6 Phylogenetic analysis of chronic-treated participants' Gag sequences

Maximum-likelihood tree showing *gag* SGA sequences (left) from two chronic-treated participants, (PID 036 and 272) and the corresponding highlighter plots indicating the silent and non-silent mutations (right). The South African subtype C reference sequence (accession no: AY772699) was used for this study as the participants were from South Africa. The reference sequence and the highlighter plot software used were from the Los Alamos database. The participant ID numbers are indicated on the corresponding tree branch, each sequence was assigned a colour-coded circle to indicate its timepoint; T/F (blue), PreART (green), blip (red). In the highlighter plots all sequences were compared to (cons) the consensus sequence of all T/F, blank grey lines were identical to the consensus, silent and non-silent mutations were denoted as red and green ticks, respectively. The number of sequences per timepoint is indicated to the right of each highlighter plot.

3.4.2 Phylogenetic relatedness of Env-derived sequences

The phylogenetic relatedness of the *env* SGA sequences was determined by drawing maximum likelihood trees using MEGA vX software. The Env SGA sequences derived from the T/F, PreART and blip time points of each participant, were used in the construction of the phylogenetic trees. Two trees were constructed in order to distinguish and compare the genetic diversity of the *env* sequences in the acute-treated versus chronic-treated participants (Figure 3.7 and 3.8).. A highlighter tool from the Los Alamos database was used to determine whether the genetic differences observed in the phylogenetic tree were silent or non-silent mutations (Figures 3.7 and 3.8). In this analysis, a participant-specific consensus T/F sequence was generated and used as a reference in the identification of mutations at T/F, PreART and blip.

The env sequences of both acute-treated participants PID 1284 and 1388, clustered with the reference sequence, indicating that both participants were infected with an HIV-1 subtype C virus. The T/F and blip sequences of PID 1284 clustered together on the same branch indicating genetic similarity. The 3 T/F sequences generated were relatively identical, however 2 of them had silent mutations. Out of the 13 blip sequences, 4 were identical to the T/F consensus, 6 had at least 1 non-silent mutation while the remaining 3 had only silent mutations. The T/F sequences of PID 1388 were identical, except for a few non-silent mutations in 2 out of the 3 sequences. It was observed that most of the blip sequences (5/8) clustered on the same branch as the T/F, indicating a close genetic relationship. Three of these sequences were identical to the T/F consensus while the other 2 had silent mutations. The blip sequences which clustered separately from the T/F (3/8) had developed non-silent mutation. Overall, the T/F and blip sequences of the acute-treated participants were genetically similar, however the significance of our observations was limited by the number of T/F sequences (n=3) analysed for both participants.

In an analysis of the chronic-treated participants, it was observed that both PID 036 and 272 were infected by an HIV-1 subtype C virus, as each participant's sequences clustered with the reference. The T/F sequences of PID 272 clustered separately from the PreART and blip sequences. The T/F population was mostly homogeneous, as 16 out of 21 sequences were identical, 3 had a non-silent mutation and the other 2 had silent mutations. None of these mutations were maintained at PreART or blip. The PreART and blip clustered separately from one another, however 3 non-silent and 11 silent mutations which were present at PreART, were maintained at the blip. For PID 036, a similar pattern was observed in which the T/F, PreART and blip clustered separately from one another. The T/F population was homogeneous with 15/19 sequences being identical, 2/19 having a silent mutation and the remaining 2 sequences having silent mutations which were not observed at the later time points. Of the several mutations were observed at PreART, 4 non-silent and 12 silent mutations were also present at the blip.

No two PreART or blip sequences were identical as each sequence had different mutations. Overall, it was observed that the T/F env sequences of the chronic-treated participants, were genetically different from the viral blip sequences, whereas the blip and PreART sequences were closely related.

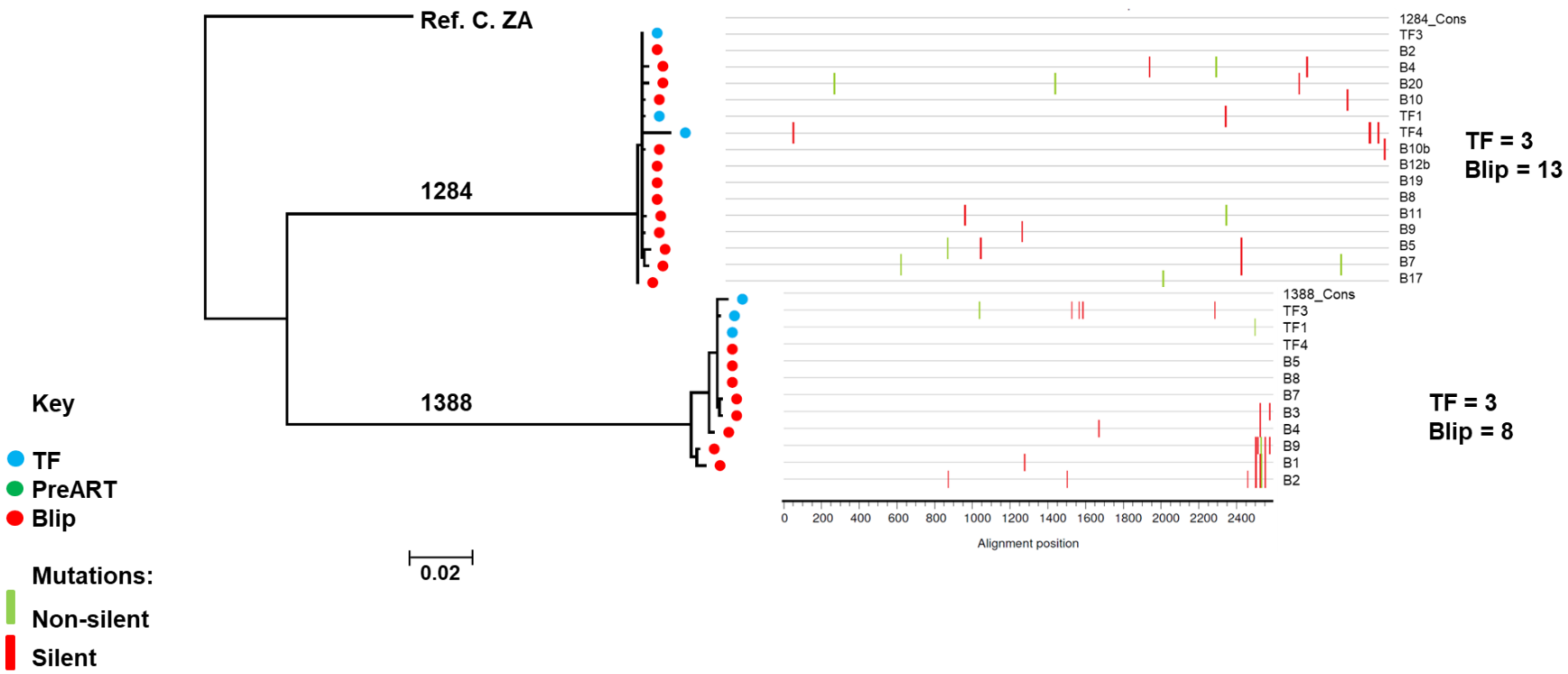


Figure 3.7 Phylogenetic analysis of acute-treated participants' Env sequences

Maximum-likelihood tree showing *env* SGA sequences (left) from two acute-treated participants, (PID 1284 and 1388) and the corresponding highlighter plots indicating the silent and non-silent mutations (right). The South African subtype C reference sequence (accession no: AY772699) was used for this study as the participants were from South Africa. The reference sequence and the highlighter plot software used were from the Los Alamos database. The participant ID numbers are indicated on the corresponding tree branch, each sequence was assigned a colour-coded circle to indicate its timepoint; T/F (blue), blip (red). In the highlighter plots all sequences were compared to the consensus sequence (cons) of all T/F, blank grey lines were identical to the consensus, silent and non-silent mutations were denoted as red and green ticks, respectively. The number of sequences per timepoint is indicated to the right of each highlighter plot.

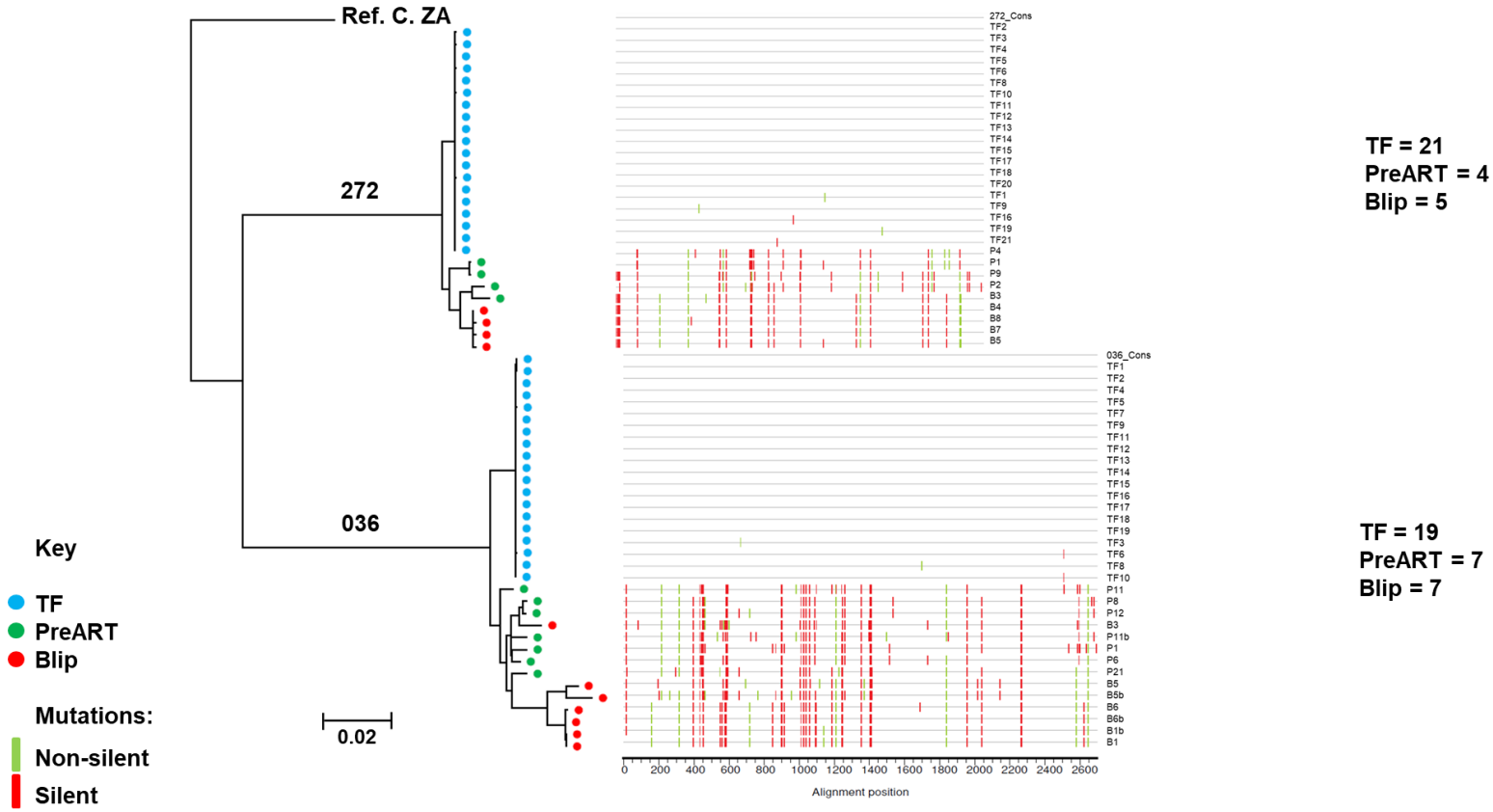


Figure 3.8 Phylogenetic analysis of chronic-treated participants' Env sequences

Maximum-likelihood tree showing *env* SGA sequences (left) from two chronic-treated participants (PID 036 and 272) and the corresponding highlighter plots indicating the silent and non-silent mutations (right). The South African subtype C reference sequence (accession no: AY772699) was used for this study as the participants were from South Africa. The reference sequence and the highlighter plot software used were from the Los Alamos database. The participant ID numbers are indicated on the corresponding tree branch, each sequence was assigned a colour-coded circle to indicate its timepoint; T/F (blue), PreART (green), blip (red). In the highlighter plots all sequences were compared to (cons) the consensus sequence of all T/F, blank grey lines were identical to the consensus, silent and non-silent mutations were denoted as red and green ticks, respectively. The number of sequences per timepoint is indicated to the right of each highlighter plot.

3.5 Genetic diversity analysis

The extent of genetic diversity present within *gag* and *env* sequences was quantified at each timepoint using the MEGA vX software as shown in Table 3.4. The within group distance diversity score, quantified the level of diversity between sequences from the same timepoint. It was observed that both of the acute-treated participants had diversity scores of 0% for both *gag* and *env* sequences at T/F and blip. This was indicative of the low level of genetic diversity between the sequences at each time point.

A diversity score of 1% was observed for the *gag* and *env* sequences of PID 036 at the PreART time point. A diversity score of 1% was also observed for *env* at the blip time point. These diversity scores indicated that the circulating viremia represented by the generated sequences at each time point, were not genetically identical to each other, due to viral evolution over time. PID 272 was shown to have a 1% diversity score at T/F *gag* and PreART *env*, indicating the presence of variants within each group of sequences. It was noted however, that the definitive significance of these observations could be limited by the number of sequences generated for each time point.

Table 3.4 Percentage genetic diversity at TF, pre-treatment, and blip time points

PID	Treatment phase	Within group distance[a] (Gag)			Within group distance[a] (Env)		
		TF	PreART	Blip	TF	PreART	Blip
1284	Acute	0.0[b]	-	0.0	0.0	-	0.0
1388	Acute	0.0	-	0.0	0.0	-	0.0
036	Chronic	0.0	1.0	0.0	0.0	1.0	1.0
272	Chronic	1.0	0.0	0.0	0.0	1.0	0.0

[a]Refers to diversity at each time point
[b]All diversity scores shown are in percentages

3.6 Cytotoxic T lymphocyte escape analysis

The HIV-1 Gag sequences were further analysed to determine whether mutations had developed, which would enable the virus at the blip time point to evade a CTL response. The epitopes which were identical to the Los Alamos Immunology A-list epitope sequence were denoted as consensus (cons) and those with mutations as variants. Shown in Table 3.5 is a list of Gag CTL epitopes which developed mutations at PreART and or blip which were not previously present at T/F, for all participants irrespective of their HLA alleles. The A-list epitope was denoted in blue, with the flanking amino acids in black and the

CTL escape in red. The dots were used to indicate the amino acids which were identical to the A-list epitope and each participant's selecting HLA alleles were highlighted in yellow.

In an analysis of PID 1284 epitopes, we observed 2 epitopes (HLA-A*11:01 associated AK11 and HLA-B*07:02 associated GL9) where a mutation developed at the blip which was not previously present at T/F. Within the AK11 epitope, all 12 T/F sequences and 5/7 blip sequences were denoted as cons as they were identical to the A-list epitope. Only 2/7 blip sequences had a G7E non-CTL escape mutation. A G1E non-CTL escape mutation developed in 2/7 blip sequences within the GL9 epitope, while 12/12 T/F and 5/7 blip sequences remained cons. For PID 1388, 13/13 TF sequences had a known CTL escape mutation, K9R within the NL11 epitope. This mutation is selected for by the participant's HLA B*08:01 which could suggest early immune escape. This mutation was maintained in 5/5 blip sequences however, 1/5 blip sequences had an additional I7F non-CTL escape mutation.

Four PID 036 epitopes (HLA-A*03:01 RK9, HLA-B*07:02 RK9, HLA-B*51:01 RK9, HLA-A*11:01 TR9) had known CTL escape mutations at T/F which were maintained in all the PreART and blip. These epitopes were not targeted by the participant's HLA alleles, suggesting that the mutations may have been the result of transmitted escape. Within the HLA-A*03:01 KK9 epitope, the known CTL escape mutation K1R which was present in 6/6 T/F, 3/3 PreART and 9/11 blip sequences, reverted to cons in 2/11 blip sequences. The CTL escape mutation E1D which had been present in 6/6 T/F within the HLA-B*08 EL9 epitope, was replaced by an E1N non-CTL escape mutation in 2/3 PreART and 10/11 blip sequences. In the HLA-B*35:01 epitope, the K4Q CTL escape mutation was maintained in all the T/F, PreART and blip sequences with the exception of 1/11 blip sequence which had a K4P non-CTL escape mutation. These epitopes were not associated with the participant's HLA alleles. Interestingly, in epitope HLA-A*02:01 VV9 (which is not targeted by the participant's HLA alleles), one CTL escape mutation in 6/6 T/F (V9T) was replaced by another CTL escape mutation (V9A) in 3/3 PreART and 11/11 blip sequences. In the TW10 epitope which is targeted by PID 036's HLA-B*57:01/B*58:01, the CTL escape mutation (G9A) which was present in 6/6 T/F, 3/3 PreART and 5/11 blip, was substituted for another CTL escape mutation (G9T) in 6/11 blip sequences, suggesting potential immune pressure at the blip.

The following PID 272 epitopes had known CTL escape mutations at T/F which were maintained in the PreART and blip sequences: HLA-B*40:02 GI9, HLA-B*40 DP10, HLA-A*03:01 KK9, HLA-B*07:02 RK9, HLA-B*51:01 RY8, HLA*B08:01 GK9, HLA-A*24:02 KW9, HLA-B08:01 EV9, HLA-A*26:01/A*26:02/A*26:03 EL9, HLA-Cw*01:02 VL8, HLA-B*53 TL9. These epitopes were

not targeted by the participant's HLA alleles, suggesting that the mutations may have been the result of transmitted escape. A V9I CTL escape mutation which was not present in any of the 12 T/F sequences, was observed within the HLA-B*44:15 epitope, in 6/6 PreART and 9/9 blip sequences. The same CTL escape mutation was also observed at PreART and blip in the following epitopes which are not targeted by the participant's HLA allele: HLA-B57:01 KF11, HLA-A*26:01/A*26:02/A*26:03 EL9 and HLA-Cw*01:02 VL8. This would suggest the random development of CTL escape mutations. Within the IF11 epitope associated with the participant's HLA-A*30, a V2I CTL escape mutation which was not present in 12/12 T/F, was observed in 6/6 PreART and 9/9 blip sequences, suggesting that there was immune pressure prior to treatment initiation. Within the participant's HLA-B*57:01/B*58:01-associated TW10 epitope, a CTL escape mutation (G9A) was maintained in all T/F, PreART and blip sequences. An additional CTL escape mutation accumulated in this epitope in all 6 PreART and 9 blip sequences. This was an indication of immune pressure before treatment.

Table 3.5 Gag CTL escape variants per participant over time

Participant	Time point	Days post detection	Epitope[a]	Sequence frequency	HXB2 position	HLA association	Mutation
1284			EEMMTACQGVGGPGHKARVLA				
	T/F		.....	12/12	344-364	A*1101	
	Blip	420	..........E.....	2/7			
			.....	5/7			
			CQGVGGPGHKARVLAEAMS				
	TF		.....	12/12	350-368	B*0702	
	Blip	420	.....E.....	2/7			
			.....	5/7			
1388			VQNANPDCKTILKALGPAAT				
	TF		A.....R....G..[b]	13/13	323-342	B*0801[c]	K9R
	Blip	245	A.....F.R....G..	1/5			
			A.....R....G..	4/5			
036			LDRWEKIRLRPGGKKYKL				
	TF		..K..R.....H.M.	6/6	13-31	A*0301	K1R
	PreART	665	..K..R....T....H.M.	1/3			
			..K..R.....H.M.	2/3			
	Blip	1847	..K..R.....H.M.	9/11			
			..K.....H.M.	2/11			
			DRWEKIRLRPGGKKYKLK				
	TF		.K..R.....H.M..	6/6	14-32	B*2705	
	PreART	665	.K..R....T....H.M..	1/3			
			.K..R.....H.M..	2/3			
	Blip	1847	.K..R.....H.M..	9/11			
			.K.....H.M..	2/11			
			RWEKIRLRPGGKKKYKLKH				
	TF		K..R.....H.M...	6/6	15-33	A*0301	K9H
	PreART	665	K..R....T....H.M...	1/3			
			K..R.....H.M...	2/3			
	Blip	1847	K..R.....H.M...	9/11			
			K.....H.M...	2/11			
			EKIRLRPGGKKKYKLKHIV				
	TF		.R.....H.M.....	6/6	17-35	B*0702	K7H, K9MK
	PreART	665	.R....T....H.M.....	1/3			

		.R.....H.M.....	2/3			
Blip	1847	.R.....H.M.....	9/11			
		...........H.M.....	2/11			
		EKIRLRPGGKKKYKLKHL				
TF		.R.....H.M...I	6/6	17-34	B*5101	
PreART	665	.R....T....H.M...I	1/3			
		.R.....H.M...I	2/3			
Blip	1847	.R.....H.M...I	9/11			
		...........H.M...I	2/11			
		YNTVATLYCVHQRIEIKD				
TF		...........EG.DVR.	6/6	79-96	A*1101	Q7E, R8K
PreART	665	...........EG.NVR.	2/3			
		...........EG.DVR.	1/3			
Blip	1847	...........EG.NVR.	10/11			
		...........EG.DVR.	1/11			
		CVHQRIEIKDTKEALDKIEE				
TF		...EG.DVR...........	6/6	87-106	B*08	E1D, I2V
PreART	665	...EG.NVR...........	2/3			
		...EG.DVR...........	1/3			
Blip	1847	...EG.NVR...........	10/11			
		...EG.DVR...........	1/11			
		AADTGNSSKVSQNYPIVQN				
TF		...-.---Q..........	6/6	119-137	B*3501	K4Q
PreART	665	...-.---Q..........	3/3			
Blip	1847	...-.---P..........	1/11			
		...-.---Q..........	10/11			
		DIAGTTSTLQEQIGWMTNNPP				
TF		............A...S...	6/6	235-255	B*5701, B*5801	G9A, G9T
PreART	665	............A...S...	3/3			
Blip	1847	............T...S...	6/11			
		............A...S...	5/11			
		KARVLAEAMSQVTNSATIMM				
TF		...........TG.A-N...	6/6	359-378	B*0201	V9T, V9A
PreART	665	...........AG.A-NV..	1/3			
		...........AG.A-N...	1/3			
		...........AG.T-N...	1/3			
Blip	1847	...........AG.A-N...	10/11			

			.....I.AG.A-N...	1/11			
			KDCTERQANFLGKIWPSYK				
	TF		..................H.	6/6	424-442	B*4801, B13	
	PreART	665	..................H.	1/1			
	Blip	1847	..................H.	10/11			
			............R....H.	1/11			
			ERQANFLGKIWPSYK-GRPGN				
	TF		.............H.-.....	6/6	428-447	A*0201	
	PreART	665	.............H.-.....	1/1			
	Blip	1847	.............H.-.....	10/11			
			........R....H.-.....	1/11			
272			SVLSGGELDRWEKIRLRPG				
	TF		...R.EK..K..R......	12/12	6-24	B*4002	E7K
	PreART	434	...R.EK..K..R......	5/6			
			...R.EK..K..R..F...	1/6			
	Blip	1337	...R.EK..K..R......	8/9			
			...KSEK.NK....K....	1/9			
			SGGELDRWEKIRLRPGGKKKY				
	TF		R.EK..K..R.........H.	12/12	9-29	B40	
	PreART	434	R.EK..K..R..F......P.	1/6			
			R.EK..K..R......A..H.	1/6			
			R.EK..K..R.........H.	4/6			
	Blip	1337	R.EK..K..R.........H.	8/9			
			KSEK.NK..R.K.......H	1/9			
			LDRWEKIRLRPGGKKKYKL				
	TF		..K..R.........H.M.	12/12	13-31	A*0301	K1R
	PreART	434	..K..R..F......P.M.	1/6			
			..K..R......A..H.M.	1/6			
			..K..R.........H.M.	4/6			
	Blip	1337	..K..R.........H.M.	8/9			
			.NK..R.K.......H.M.	1/9			
			DRWEKIRLRPGGKKKYKLK				
	TF		.K..R.........H.M..	12/12	14-32	B*2705	
	PreART	434	.K..R..F......P.M..	1/6			
			.K..R......A..H.M..	1/6			
			.K..R.........H.M..	4/6			

Blip	1337	.K..R.....H.M..	8/9			
		NK..R.K.....H.M..	1/9			
		RWEKIRLRPGGKKKYKLKH				
TF	434	K..R.....H.M...	12/12	15-33	A*0301	R1K,K9H
PreART		K..R..F.....P.M...	1/6			
		K..R.....A..H.M...	1/6			
	1337	K..R.....H.M...	4/6			
Blip		K..R.....H.M...	8/9			
		K..R.K.....H.M...	1/9			
		EKIRLRPGGKKKYKLKHIV				
TF		.R.....H.M.....	12/12	17-35	B*0702	K9H
PreART	434	.R.....A..H.M....I	1/6			
		.R..F.....P.M....I	1/6			
		.R.....H.M....I	4/6			
Blip	1337	.R.....H.M....I	8/9			
		.R.K.....H.M....I	1/9			
		EKIRLRPGGKKKYKLKHL				
TF		.R.....H.M...I	12/12	17-35	B*5101	K9H
PreART	434	.R.....A..H.M...I	1/6			
		.R..F.....P.M...I	1/6			
		.R.....H.M...I	4/6			
Blip	1337	.R.....H.M...I	8/9			
		.R.K.....H.M...I	1/9			
		IRLRPGGKKKYKLKHIVWA				
TF		.....H.M.....	12/12	19-37	B*0801	
PreART	434	..F.....P.M....I.P	1/6			
		.....A..H.M....I..	1/6			
		.....H.M....I..	4/6			
Blip	1337	.....H.M....I..	8/9			
		.K.....H.M....I..	1/9			
		PGGKKKYKLKHIVWASREL				
TF		.....H.M.....	12/12	23-41	A*2402	
PreART	434	.....P.M....I.P....	1/6			
		..A..H.M....I.....	1/6			
		.....H.M....I.....	4/6			
Blip	1337	.....H.M....I.....	8/9			
		.....H.M...LI.....	1/9			

		KYKLKHIVWASRELERFAV				
TF		H.M.....L	12/12	28-46	Cw*0804	
PreART	434	P.M....I.P.....L	1/6			
		H.M....I.....L	5/6			
Blip	1337	H.M...LI.....L	1/9			
		H.M....I.....L	8/9			
		YKLKHIVWASRELERFAVNPG				
TF		.M.....L...	11/12	29-49	A30	
		.M.....L.H.	1/12			
PreART	434	.M....I.P.....L...	1/6			
		.M....I.....L...	5/6			
Blip	1337	.M...LI.....L...	1/9			
		.M....I.....L...	8/9			
		LKHIVWASRELERFAVNPG				
TF		.....L...	12/12	31-49	B*3501	
		.....L.H.	1/12			
PreART	434	....I.P.....L...	1/6			
		....I.....L...	5/6			
Blip	1337	...LI.....L...	1/9			
		....I.....L...	8/9			
		QTGSEELRSLYNTVATLYC				
TF		...T...K.....	12/12	69-87	B*0801	R3K, Y6K
PreART	434	...T...K.....	6/6			
Blip	1337	...T..PK.....	1/9			
		...T...K.....	8/9			
		YNTVATLYCVHQRIEIKD				
TF		.....E...V..	12/12	79-96	A*1101	Q7E
PreART	434	.....K...V..	5/6			
		.....V..	1/6			
Blip	1337	.....K...V..	4/9			
		.....E...V..	4/9			
		.....KG..V..	1/9			
		WVKVIEEKAFSPEVIPMFS				
TF		.....T	12/12	155-173	B*4415	
PreART	434	.....I....T	6/6			
Blip	1337	.....I....T	9/9			
		WVKVIEEKAFSPEVIPMFSA				
TF		.....T.	12/12	155-174	B*5703	
PreART	434	.....I....T.	6/6			

Blip	1337	.....I....T.	9/9			
		KVIEEKAFSPEVIPMFSALSE			B*5701,	
TF		.....T....	12/12	157-177	B*5703, B63	V7I
PreART	434	.....I....T....	6/6			
Blip	1337	.....I....T....	9/9			
		IEEKAFSPEVIPMFSALSE				
TF	434	.....T....	12/12	159-177	B57	V5I
PreART	1337	.....I....T....	6/6			
Blip		.....I....T....	9/9			
		KAFSPEVIPMFSALSEGAT				
TF		.....T.....	12/12	162-180	A*2601,	V2I, S7T
PreART	434	.....I....T.....	6/6		A*2602, A*2603	
Blip	1337	.....I....T.....	9/9			
		AFSPEVIPMFSALSEGAT				
TF		.....T.....	12/12	163-180	Cw*0102	
PreART	434	.....I....T.....	6/6			
Blip	1337	.....I....T.....	9/9			
		LSEGATPQDLNTMLNTVGG				
TF		.....	12/12	175-193	B*0702,B*3910,	
PreART	434	.....	6/6		B*4201,B*8101,	
Blip	1337	.....	8/9		Cw*0802	
		.....I.....	1/9			
		LSEGATPQDLNMMLNTVGG				
TF		.....T.....	12/12	175-193	B53	
PreART	434	.....T.....	6/6			
Blip	1337	.....T.....	8/9			
		.....IT.....	1/9			
		LSEGATPYDINQMLNTVGG				
TF		.....Q.L.T.....	12/12	175-193	B*5301	
PreART	434	.....Q.L.T.....	6/6			
Blip	1337	.....Q.L.T.....	8/9			
		.....Q.LIT.....	1/9			
		DIAGTTSTLQEQIGWMTNNPP				
TF	1	.....A.....	12/12	235-255	B*5701, B*5801	G9A, T3N
PreART	434	.....N.....A.....	6/6			
Blip	1337	.....N.....A.....	9/9			

[a]*gag* CTL epitope (blue), [b]known CTL escape variant (red), [c]participant HLA allele (yellow)

Second, we sought to evaluate to what extent the observed CTL escape was influenced by immune pressure. Patient-derived epitopes sequences which were identical to HXB2 reference were denoted as consensus and those with mutations as variants. The distribution of consensus versus variants within all T/F Gag epitopes, was initially analysed irrespective of HLA alleles for all four participants. It was noted that a large majority of the CTL epitopes analysed (85.6%) contained variants at this early stage of infection (Figure 3.9). The high level of variants present in the TF virus would either suggest transmitted escape or early immune selections. To determine this, we quantified the level of CTL escape. Interestingly, more than half of these variants are known CTL escape variants (54.5%) as defined by the LANL Immunology database. Further analysis was conducted in order to determine how many of these CTL escape mutations were associated with the participants' HLA alleles. We found that only 10% of this known CTL escape was associated with the participants' HLA alleles.

We then conducted a similar analysis for each individual participant, comparing the distribution of variants, known CTL escape variants and HLA-associated CTL escape between the T/F and blip Gag epitopes, regardless of HLA alleles. When comparing the distribution of consensus versus variants at T/F, for the acute treated participant, PID 1284 (Figure 3.10), it was observed that 90% of the CTL epitopes contained variants, suggesting that the variants were transmitted. More than half (55.6%) of these variants were found to be known CTL escape variants, however only 5% of this known CTL escape was actually associate with the participant's HLA alleles. When the same analysis was conducted for the blip Gag epitopes of PID 1284, there was no change observed in the percentage of variants, known CTL escape or HLA-associated CTL escape, as compared to the T/F. This suggested that the transmitted CTL escape was maintained at the blip time point. Interestingly, the same trend was observed in the other acute-treated participant 1388 (Figure 3.11), as the distribution of variants, known CTL escape and HLA-associated CTL escape at T/F was identical to that at the blip time point. The majority of Gag epitopes were found to contain variants (87.5%), half of these variants were known CTL escape (50%) and but only 5% of the CTL escape was associated with the participant's HLA allele.

In an analysis of chronic-treated participant, PID 036, it was observed that the majority of the Gag CTL epitopes contained variants at TF (82.5%) and the percentage of these variants increased to 83.8% at the blip time point (Figure 3.12). While more than half of the variants at T/F (54.5%) were known CTL escape, only 16.7% of this known escape was associated with the participant's HLA alleles, suggesting that most CTL escape was transmitted escape. At the blip, it was observed that the percentage of known CTL escape decreased to 50.7% due to the reversion of CTL escape mutations in 2 epitopes which were not targeted by the participant's HLA alleles. This suggests that the reversion occurred at random. Despite this decrease in known CTL escape at the blip, the percentage HLA-associated CTL escape

increased to 17.6%. This was because while the number of CTL escape mutations decreased, the number of HLA-associated CTL escape mutations remained constant from T/F to blip. For PID 272, the majority of the CTL epitopes contained variants at both T/F (82.5%) and blip (83.8%), respectively (Figure 3.13). The majority of these variants were known CTL escape mutations at both T/F (53.0%) and blip (61.2%), with the marked increase observed at the blip. Of this known CTL escape, 11.4% was associated with the participant's HLA alleles at T/F and this increased to 12.2% at the blip, likely due to immune pressure.

Furthermore, we analysed the HLA-associated CTL epitopes of the chronic-treated participants and observed that variants accumulated, in these epitopes over time. The epitope A-list from the Los Alamos Immunology database was used to identify mutations in the HLA-associated *gag* CTL epitopes of each participant. Each mutation within an HLA-associated epitope was denoted as 1 variant, regardless of whether or not the variant was a known CTL escape mutation. The sum of variants which accumulated at each time point were plotted for each participant (Figure 3.14A). For PID 036, the number of variants at T/F (n=2) were maintained at PreART and then increases to 3, at the blip within the following epitopes: HLA-B*57:01/B*58:01 TW10, HLA-A*02:01 FK10. For PID 272, only 1 variant was present at T/F, however, 3 more mutations had accumulated in the HLA-associated epitopes at PreART and blip, within the HLA-A*30 IF11 and HLA-B*57:01/B*58:01 TW10 epitopes. Shown in Figure 3.14B is an example of the accumulation of variants within the HLA-B*57:01/B*58:01 TW10 epitope of PID 272. Two CTL escape mutations (T3N and G9A) accumulated the participant's HLA B*57:01/B*58:01-associated epitope. This accumulation of CTL escape variants was likely due to immune pressure prior to treatment initiation.

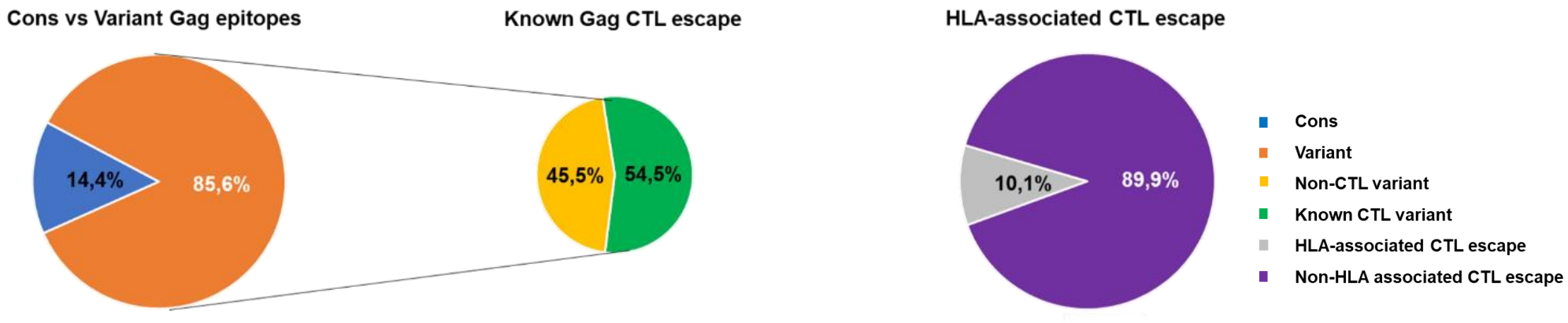


Figure 3.9 CTL escape distribution in the T/F Gag epitopes

The overall distribution of variants in the T/F *gag* epitopes of all four participants is shown. The percentage of epitope which were wild type was denoted as cons (blue) and the mutations as variant (orange). The percentage of variants which corresponded to known CTL escape were (green) and non-CTL variants (yellow). The percentage of known CTL escape which was associated with the participants' HLA-alleles was indicated (grey) and the non-HLA associated CTL escape (purple).

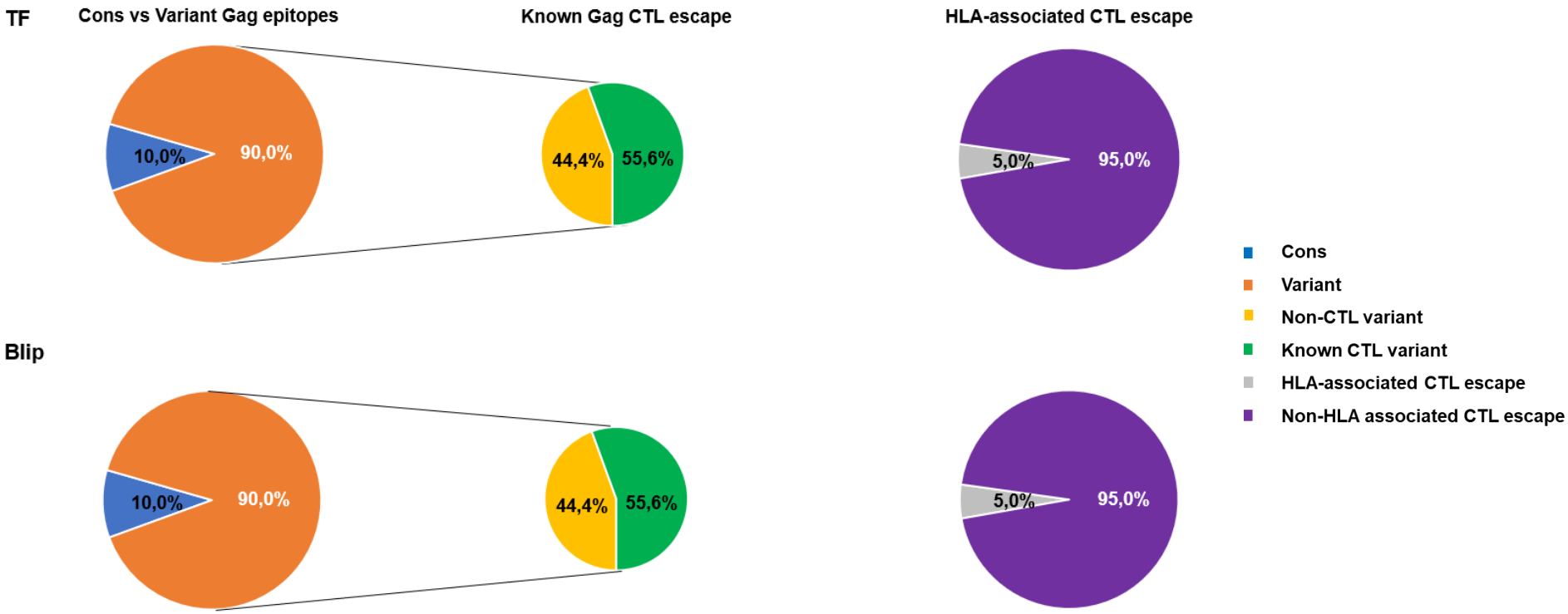


Figure 3.10 CTL escape distribution in the T/F versus blip Gag epitopes of PID 1284

The distribution of variants in the T/F and blip *gag* epitopes of two acute-treated participants is shown. The percentage of epitope which were wild type was denoted as cons (blue) and the mutations as variant (orange). The percentage of variants which corresponded to known CTL escape were (green) and non-CTL variants (yellow). The percentage of known CTL escape which was associated with the participants' HLA-alleles was indicated (grey) and the non-HLA associated CTL escape (purple).

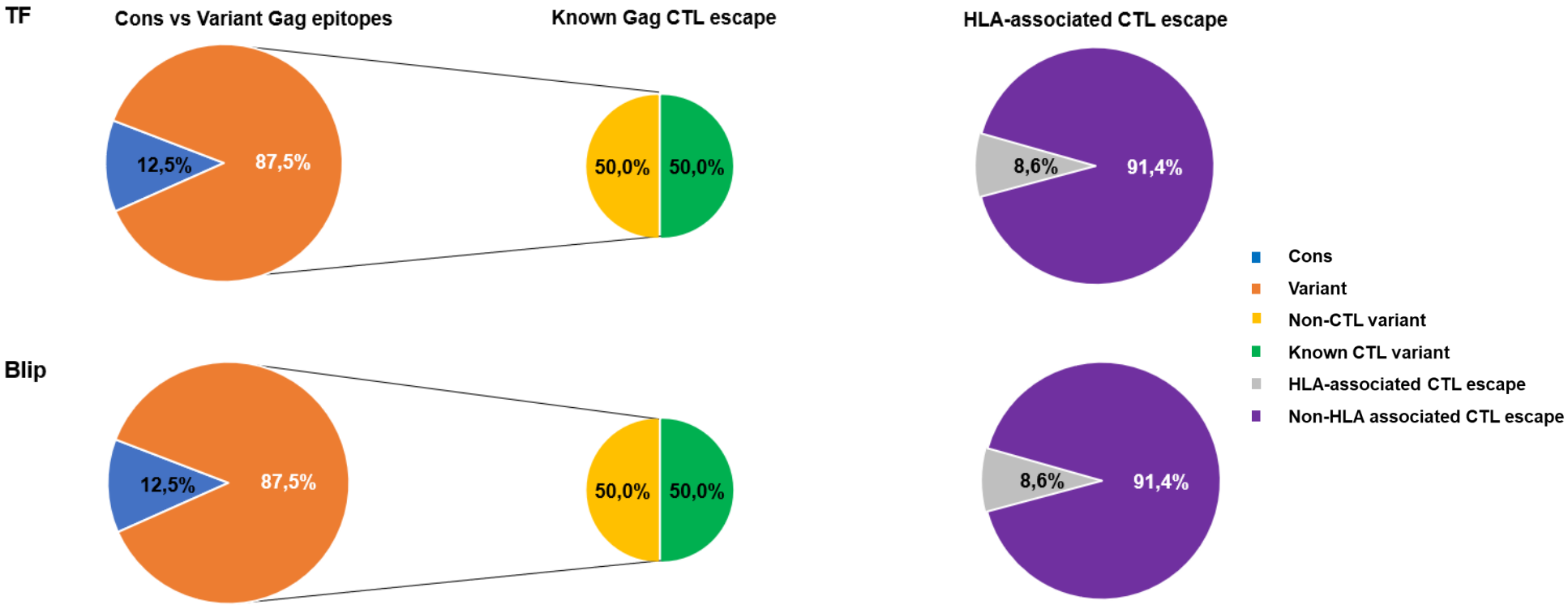


Figure 3.11 CTL escape distribution in the T/F versus blip Gag epitopes of PID 1388

The distribution of variants in the T/F and blip *gag* epitopes of two chronic-treated participants is shown. The percentage of epitope which were wild type was denoted as cons (blue) and the mutations as variant (orange). The percentage of variants which corresponded to known CTL escape were (green) and non-CTL variants (yellow). The percentage of known CTL escape which was associated with the participants' HLA-alleles was indicated (grey) and the non-HLA associated CTL escape (purple).

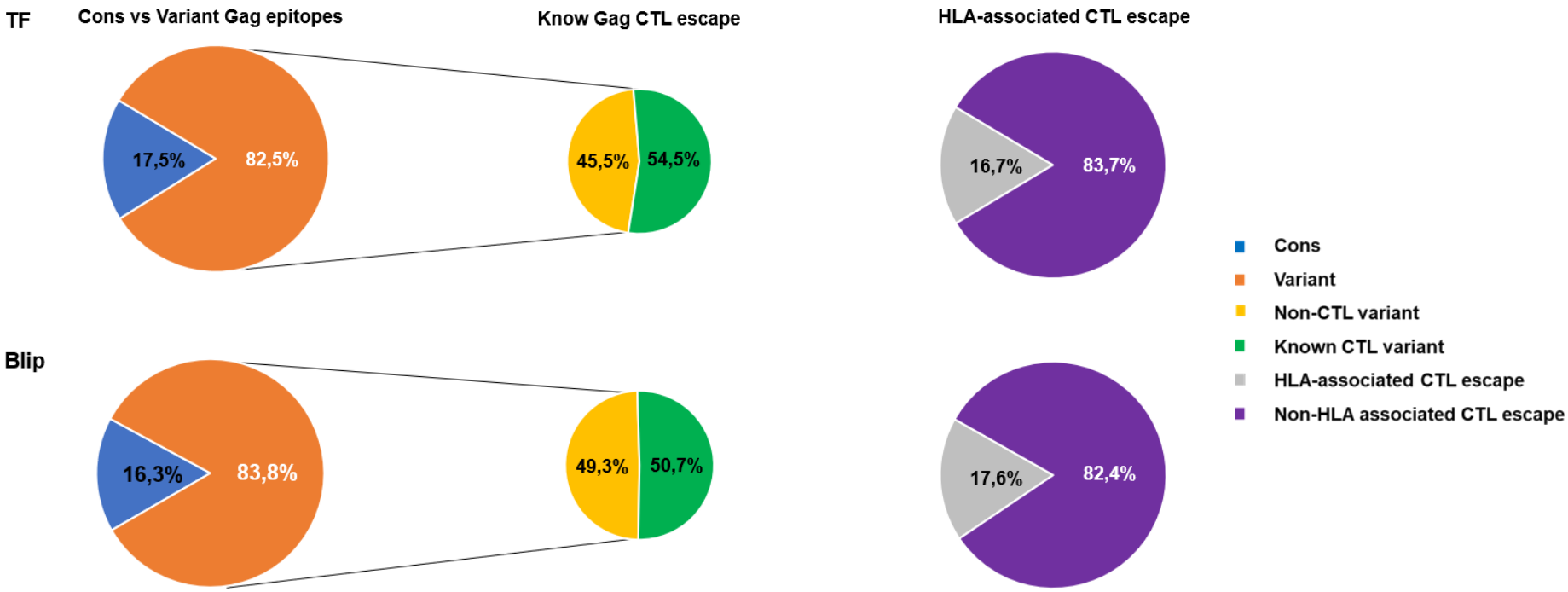


Figure 3.12 CTL escape distribution in the T/F versus blip Gag epitopes of PID 036

The distribution of variants in the T/F and blip *gag* epitopes of two chronic-treated participants is shown. The percentage of epitope which were wild type was denoted as cons (blue) and the mutations as variant (orange). The percentage of variants which corresponded to known CTL escape were (green) and non-CTL variants (yellow). The percentage of known CTL escape which was associated with the participants' HLA-alleles was indicated (grey) and the non-HLA associated CTL escape (purple).

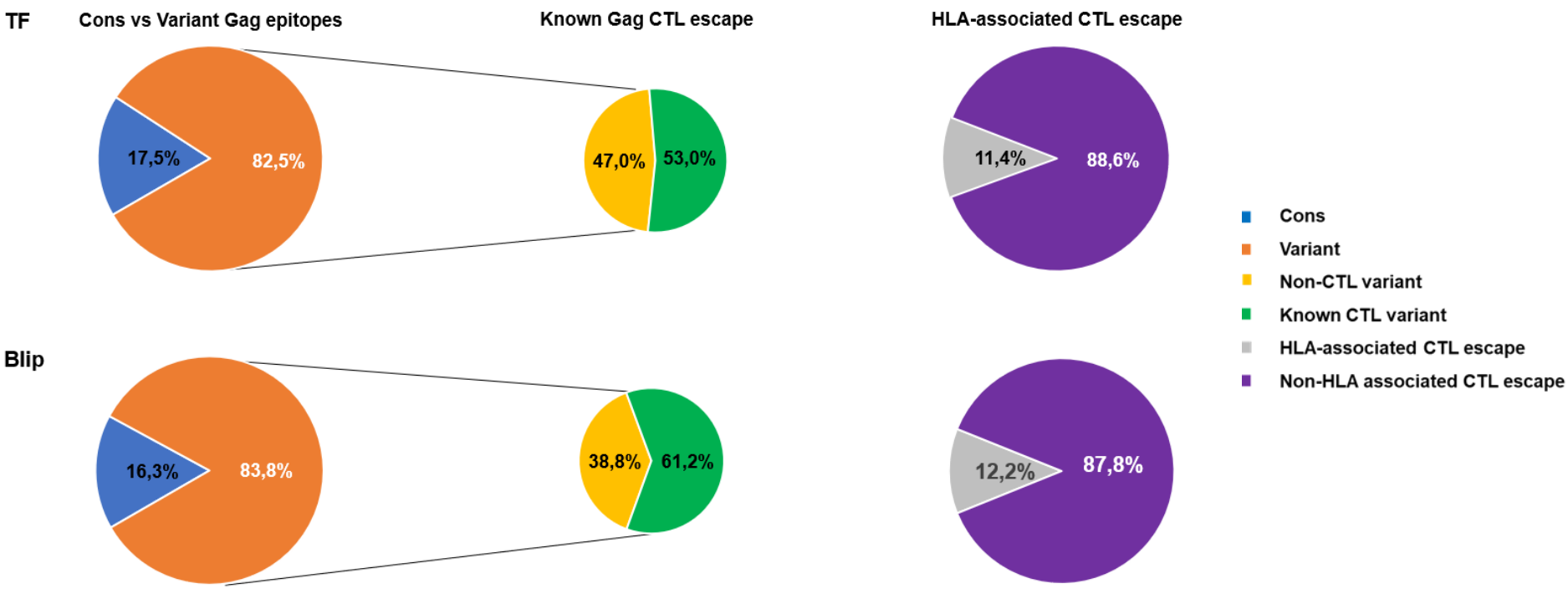
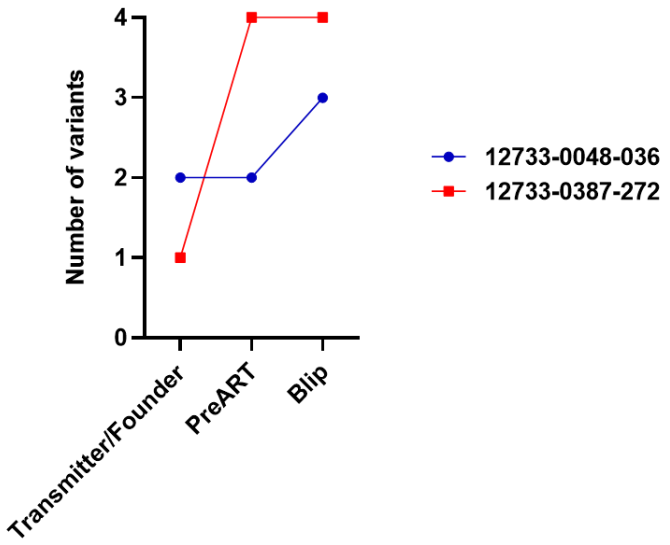


Figure 3.13 CTL escape distribution in the T/F versus blip Gag epitopes of PID 272

The distribution of variants in the T/F and blip *gag* epitopes of two chronic-treated participants is shown. The percentage of epitope which were wild type was denoted as cons (blue) and the mutations as variant (orange). The percentage of variants which corresponded to known CTL escape were (green) and non-CTL variants (yellow). The percentage of known CTL escape which was associated with the participants' HLA-alleles was indicated (grey) and the non-HLA associated CTL escape (purple).

A



B

Timepoint	Days post detection	Epitope	Sequence frequency
		DIAGTTSTLQEQIGWMTNNPP	
TF	1	..............A.....	12/12
PreART	434	.......N.....A.....	6/6
Blip	1337	.......N.....A.....	9/9

HLA type: B*58:01

Figure 3.14 An accumulation of variants within HLA-associated Gag CTL epitopes in two chronic-treated participants

The distribution of variants which accumulate within the HLA-associated Gag CTL epitopes of the chronic-treated participants over time. The number of variants that accumulate over time **(A)** per PID are denoted in **(blue)** for PID 036 and **(red)** for PID 272. The variants which accumulate in the TW10 epitope of PID 272 over time, are shown **(B)**.

3.7 Broadly neutralizing antibody escape mutations

The bnAb binding sites of the *env* sequences were analysed in order to ascertain whether mutations associated with bnAb escape were present at each time point, for all four participants. The escape mutations were identified based on literature. The HXB2 reference sequence was used to assign position numbers to the CAP45 HIV subtype C sequence (Accession number: DQ435682), to which all the *env* SGA sequences were compared. The CAP45 South African reference sequence was selected because the participants were from South Africa. A dot was assigned to an epitope which was identical to CAP45, and mutations were indicated by the letter symbolizing the new amino acid. An escape mutation was denoted by an asterisk as being unique if it differed from the mutation or amino acid observed at earlier or later time points. Shown in Table 3.6-10 are the mutations which develop over time within the 5 neutralizing antibody binding sites, for all 4 participants.

In the CD4 binding site, a K97R mutation developed at the PreART time point of PID 036 in 2/10 sequences, while the rest of PreART and blip sequences remained identical to the T/F. This minor PreART variant may exhibit differential sensitivity to bnAbs as compared to T/F and blip. Other minor variants which occurred in ≤20% of the PreART sequences had the following mutations: N234T, K275Q, T281A, S461G, D463K, N474D. Of these mutations only T281A and S461G were retained at the blip in 1/7 sequences. The K97E mutation that had been present in 20/20 T/F sequences of PID 272, reverted to wildtype in all PreART and blip sequences, suggesting immune pressure prior at PreART. A glycine (G) was added at PreART and blip to the previously vacant position 461, while a D463T mutation was observed in only 1/3 PreART sequences. In the acute-treated participants 1284 and 1388, it was observed that the mutations which were present at T/F were maintained over time, as shown at positions 281, 364, 463, 474 and 476 in both participants.

At the gp120/gp41 binding site, a minor variant which had an N234T mutation was observed in 1/10 of PID 036 PreART sequences, however it was not present in any of the blip sequences. At position 448 the glutamic acid (E) residue at T/F was replaced by an asparagine (N) in all the PreART and blip sequences. An E279P mutation occurred in 1/5 PID 272 blip sequences, while other positions associated with bnAb escape such as 88, 611 and 637 remained wild type at T/F, PreART and blip. Other mutations also included N276E and N448V, which occurred in 1/5 PID 272 blip sequences. In the acute-treated participants 1284 and 1388, the N230D, N276D and N234S, S448N mutations which were present at the T/F of each respective participant, were retained in all the PreART and blip sequences.

The MPER binding site mutations which were present in the T/F sequences of each participant were maintained at PreART and or blip. There were no variants which developed over time for any participant at any position, as the MPER region is relatively conserved. The V1V2 N-linked glycan at N160 and

N332 were conserved for all the participants. Other conserved positions included were T163, T164, D167, Y173 and Y177, in all participants. The N-linked N156 glycan was conserved for all participants except PID 1388 where it was replaced by arginine (R) at both T/F and blip. The L165M mutation was observed for PID 036 at the pre-ART time point in 1 out of 10 sequences. The N-linked N130 glycan for PID 036 T/F was replaced by an aspartic acid (D) residue in all the pre-ART and blip sequences. Other interesting mutations were the replacement of the N-linked N187 glycan with threonine (T) and serine (S) in 60 and 20% of preART sequences, then 29 and 14% of the blip sequences, respectively. With the exception of the aforementioned V1V2 site mutations, the mutations which were present in the participants' T/F sequences were retained at pre-ART and blip at the following positions: 161, 162, 166, 168, 170, 171, 172, 181, 197.

The V3 loop N-linked glycans at N301 and N332 were conserved in all the participants. The N-linked glycan at N334 was replaced by serine (S) in all the participants except PID 1284 where it was conserved. In PID 036 minor variants were observed in 1/9 sequences having an R165M and I297V mutation. A Q328K mutation which occurred in 2/9 PreART was retained in 3/7 blip sequences. The following mutations which occurred in the 3 PID 272 PreART sequences were not retained at the blip: E337D, H343N and H344L. The definitive significance of these mutations could not be ascertained due to low sequence frequency (n=3). Although minor variants were observed at the blip of acute-treated participants PID 1284 (I323V, Q343K) and PID 1388 (TK297), generally the mutations which were present at T/F were retained at the blip as shown at positions 165, 300, 334, 337, 339,340 and 344.

Table 3.6 CD4 binding site neutralization escape mutations

PID	Time point	Sequence frequency	97	99	197	N234	262	275	N276D	277	T278A	N279D	280	281	282	364	365	D368A	E370A	371	372	425	429	455	R456W	457	G458Y	459	461	462	463	471	D474	476
HXB2			.	.	.	.	.	V	.	F	.	D	.	A	.	S	.	.	.	I	V	.	K	.	.	.	.	.	S	N	.	.	D	R
CAP45			K	D	N	N	N	E	N	L	T	N	N	I	K	P	S	D	E	V	T	N	G	T	R	D	G	G	T	R	N	G	N	K
036	TF	19/19	.	.	D	.	.	K	.	.	.	D	.	T	.	S	T	.	.	I	.	.	Q	V	.	.	.	.	S	T	D	.	.	.
	PreART	5/10	.	.	D	.	.	K	.	.	.	D	.	T	.	S	T	.	.	I	.	.	Q	V	.	.	.	.	S	T	.	.	.	.
		2/10	R*	.	D	.	.	K	.	.	.	D	.	T	.	S	T	.	.	I	.	.	Q	V	.	.	.	.	S	T	.	.	.	.
		1/10	.	.	D	.	.	K	.	.	.	D	.	T	.	S	T	.	.	I	.	.	Q	V	.	.	.	.	G	T	K	.	D	.
		1/10	.	.	D	.	.	Q	.	.	.	D	.	A	.	S	T	.	.	I	.	.	Q	V	.	.	.	.	S	T	.	.	.	.
		1/10	.	.	D	T	.	K	.	.	.	D	.	T	.	S	T	.	.	I	.	.	Q	V	.	.	.	.	G	T	.	.	.	.
	Blip	4/7	.	.	D	.	.	R	.	.	.	D	.	T	.	S	T	.	.	I	.	.	Q	V	.	.	.	.	S	T	.	.	.	.
		1/7	.	.	D	.	.	K	.	.	.	D	.	T	.	S	T	.	.	I	.	.	Q	V	.	.	.	.	G	T	.	.	.	.
		1/7	.	.	D	.	.	K	.	.	.	D	.	T	.	S	T	.	.	I	.	.	Q	V	.	.	.	.	S	T	.	.	.	.
		1/7	.	.	D	.	.	K	.	.	.	D	.	A	.	S	T	.	.	I	.	.	Q	V	.	.	.	.	S	T	.	.	.	.
272	TF	20/20	E	.	.	.	.	.	.	.	.	E	.	T	.	S	.	.	.	I	.	.	A	.	.	.	.	-	-	N	.	.	E	R
	PreART	2/3	.	.	.	.	.	.	.	.	.	E	.	T	.	S	.	.	.	I	.	.	A	.	.	.	.	.	G	N	.	.	E	R
		1/3	.	.	.	.	.	.	.	.	.	E	.	T	.	S	.	.	.	I	.	.	A	.	.	.	.	.	G	N	T	.	E	R
	Blip	4/4	.	.	.	.	.	.	.	.	.	E	.	T	.	S	.	.	.	I	.	.	A	.	.	.	.	.	G	N	.	.	E	R
1388	TF	3/3	E	.	.	S	.	.	.	.	.	.	.	A	.	H	.	.	.	I	.	.	K	.	.	.	.	.	N	E	T	E	D	R
	Blip	8/8	E	.	.	S	.	.	.	.	.	.	.	A	.	H	.	.	.	I	.	.	K	.	.	.	.	.	N	E	T	E	D	R
1284	TF	2/3	.	.	.	.	.	.	D	.	G	.	.	A	.	S	.	.	.	.	.	.	.	.	.	.	.	E	.	-	S	.	D	R
		1/3	.	.	.	-	-	-	-	-	-	-	-	-	-	S	.	.	.	.	.	.	.	.	.	.	.	E	.	-	S	.	D	R
	Blip	17/17	.	.	.	.	.	.	D	.	G	.	.	A	.	S	.	.	.	.	.	.	.	.	.	.	.	E	.	-	S	.	D	R

*escape mutation

Table 3.7 Gp120/gp41 binding site neutralization escape mutations

PID	Time point	Sequence frequency	N88A	230	234	241	262	276	279	448	N611A	N637A
HXB2			.	.	.	.	.	.	D	N	.	.
CAP45			N	N	N	N	N	N	N	S	N	N
036	TF	19/19	.	.	.	.	.	.	D	E	.	.
	PreART	9/10	.	.	.	.	.	.	D	N	.	.
		1/10	.	.	T	.	.	.	D	N	.	.
	Blip	7/7	.	.	.	.	.	.	D	N	.	.
272	TF	20/20	.	.	.	.	.	.	E	N	.	.
	PreART	2/3	.	.	.	.	.	.	E	N	.	.
		1/3	.	.	.	-	.	.	E	N	.	.
	Blip	4/5	.	.	.	.	.	.	E	N	.	.
		1/5	.	.	.	.	.	E	P*	V	.	.
1388	TF	3/3	.	.	S	.	.	.	.	N	.	.
	Blip	8/8	.	.	S	.	.	.	.	N	.	.
1284	TF	2/3	.	D	.	.	.	D	.	.	.	.
		1/3	.	-	-	-	-	-	-	.	.	.
	Blip	17/17	.	D	.	.	.	D	.	.	.	.

*escape mutation

Table 3.8 MPER neutralizing antibody binding site mutations

PID	Time point	Sequence frequency	N671	W672L	F673L	674	676	N677	W680G	N683R
HXB2			.	.	.	.	.	.	.	.
CAP45			N	W	F	N	T	N	W	K
036	TF	19/19	S	.	.	D	S	.	.	.
	PreART	10/10	S	.	.	D	S	.	.	.
	Blip	7/7	S	.	.	D	S	.	.	.
272	TF	20/20	S	.	.	S	.	.	.	R
	PreART	3/3	S	.	.	S	.	.	.	R
	Blip	5/5	S	.	.	S	.	.	.	R
1388	TF	3/3	.	.	.	S	.	K	.	.
	Blip	8/8	.	.	.	S	.	K	.	.
1284	TF	3/3	S	.	.	S	.	.	.	.
	Blip	17/17	S	.	.	S	.	.	.	.

Table 3.9 V1V2 neutralizing antibody binding site mutations

PID	Time point	Sequence frequency	N130	156	160	161	162	163	164	L165	166	167	168	169	Q170K	171	172	173	177	181	187	197	332
HXB2			K	.	.	.	S	.	S	I	R	G	.	V	.	.	E	.	.	I	D	.	.
CAP45			R	N	N	I	T	T	E	L	R	D	K	K	Q	K	A	Y	Y	V	N	N	N
036	TF	19/19	N	.	.	A	.	.	.	R	K	.	.	.	K	Q	V	.	.	I	.	D	.
	PreART	6/10	D	.	.	A	.	.	.	R	K	.	.	.	K	Q	V	.	.	I	T*	D	.
		2/10	D	.	.	A	.	.	.	R	K	.	.	.	K	Q	V	.	.	I	S*	D	.
		1/10	D	.	.	A	.	.	.	M*	K	.	.	.	K	Q	V	.	.	I	-	D	.
		1/10	D	.	.	A	.	.	.	R	K	.	.	.	K	Q	V	.	.	I	.	D	.
	Blip	4/7	D	.	.	A	.	.	.	R	K	.	.	.	K	Q	V	.	.	I	-	D	.
		2/7	D	.	.	A	.	.	.	R	K	.	.	.	K	Q	V	.	.	I	T	D	.
		1/7	D	.	.	A	.	.	.	R	K	.	.	.	K	Q	V	.	.	I	S	D	.
272	TF	20/20	N	.	.	.	.	.	.	.	.	.	.	.	.	.	E	.	.	I	-	.	.
	PreART	3/3	N	.	.	.	.	.	.	.	.	.	.	.	.	.	E	.	.	I	-	.	.
	Blip	5/5	N	.	.	.	.	.	.	.	.	.	.	.	.	.	E	.	.	I	-	.	.
1388	TF	3/3	N	R	.	T	K	.	.	V	.	.	R	T	.	.	E	.	.	I	S	.	.
	Blip	8/8	N	R	.	T	K	.	.	V	.	.	R	T	.	.	E	.	.	I	S	.	.

*escape mutation

Table 3.10 V3 loop neutralizing antibody binding site mutations

PID	Time point	Sequence frequency	L165	295	297	300	301	304	307	323	325	328	330	332	334	337	339	340	341	343	344
HXB2			I	N	T	.	.	.	.	.	N	.	.	.	S	K	.	N	.	K	.
CAP45			L	V	R	N	N	R	I	I	D	Q	H	N	N	T	N	R	T	E	Q
036	TF	19/19	R	E	I	.	.	.	.	.	.	.	Y	.	S	Q	K	K	.	Q	R
	PreART	5/9	R	E	I	.	.	.	.	.	.	.	.	.	S	Q	.	K	.	Q	R
		2/9	R	E	I	.	.	.	.	.	.	K	Y	.	S	Q	.	.	.	Q	R
		1/9	M	E	I	.	.	.	.	.	.	.	.	.	S	Q	.	K	.	Q	R
		1/9	R	E	V	.	.	.	.	.	.	.	.	.	S	Q	.	K	.	Q	R
	Blip	4/7	R	E	V	.	.	.	.	.	.	.	.	.	S	Q	.	K	.	Q	R
		3/7	R	E	I	.	.	.	.	.	.	K	Y	.	S	Q	.	K	.	Q	R
272	TF	20/20	.	.	V	.	.	.	V	.	.	.	.	.	S	E	E	K	.	H	R
	PreART	1/3	.	.	V	.	.	.	V	.	.	.	.	.	S	D	E	K	.	H	.
		1/3	.	.	V	.	.	.	V	.	.	.	.	.	S	K	E	K	.	H	L
		1/3	.	.	V	.	.	.	V	.	.	.	.	.	S	K	E	K	.	N	L
	Blip	4/4	.	.	V	.	.	.	V	.	.	.	.	.	S	K	E	K	.	H	R
1388	TF	3/3	V	E	T	G	.	.	.	.	.	.	.	.	S	N	S	.	.	.	.
	Blip	7/8	V	E	T	G	.	.	.	.	.	.	.	.	S	N	S	.	.	.	.
		1/8	V	E	K	G	.	.	.	.	.	.	.	.	S	N	S	.	.	.	.
1284	TF	2/3	.	N	T	S	.	.	.	.	.	.	.	.	.	K	.	E	.	Q	K
	Blip	15/17	.	N	T	S	.	.	.	.	.	.	.	.	.	K	.	E	.	Q	K
		1/17	.	N	T	S	.	.	.	.	.	.	.	.	.	K	.	E	.	K	K
		1/17	.	N	T	S	.	.	.	V	.	.	.	.	.	K	.	E	.	Q	K

CHAPTER 4: DISCUSSION

The purpose of this study was to determine the genetic characteristics of the rebound virus in HIV-1 subtype C infected patients, who experienced viral blips following a period of suppressive ART. The viral variants present at the blip timepoint were sequenced and compared to those variants from earlier timepoints namely, T/F and PreART. In this study the genes of interest were *gag, pol* and *env*. The main findings of this study were that early treatment initiation was associated with a lack of genetic diversity in the blip viruses, however late treatment was associated with an increase in the number of CTL escape variants in the blips. Late treatment was also associated with the development of bnAb escape mutations. With the exception of one chronic-treated participant, the viral blips were not associated with drug resistance.

A phylogenetic analysis of *gag* sequences revealed that the blip virus was genetically diverse from the T/F and PreART, in participants who had been treated during the chronic phase of infection as shown in Figure 3.6. This was likely due to significant viral evolution occurring prior to treatment initiation (46, 48, 56). While the scope of this study did not investigate the likely source of viral rebound, the similarities between the PreART and blip in the chronic-treated participants, may suggest that the viral reservoir in this case, may have been more likely maintained by clonal expansion of variants present prior to treatment, as opposed to ongoing replication during ART (32, 57). In contrast, a distinct lack of diversity was observed in the comparison of blip to T/F sequences, in those participants that received early treatment (Table 3.3). In these individuals, the T/F and blip were almost identical, as such their diversity scores were zero. This was the result of initiating treatment upon first detection, before the onset of viral evolution, thus limiting the size of the viral reservoir as has been shown in other studies (36, 38). The identical nature of blip and T/F sequences in the acute-treated participants, suggest the unlikelihood of ongoing replication within their viral reservoirs (28).

The development of vaccine therapies which induce an immune response is important for prevention and better treatment outcomes. Previous studies have shown the potential of using specific Gag peptides to induce an effective CTL response (26). Thus, it was necessary to determine whether the circulating rebound viruses, would be susceptible to CTLs or have developed mutations to escape a CTL response. In the acute-treated participants, the CTL escape variants present at T/F were maintained at the blip. This suggested that the use of a vaccine therapy as an adjunct treatment might be potentially effective in those who initiate treatment early (58). For example, such a vaccine is likely to be effective in early acute infection as there is no escape in the immunogenic epitopes, thus making the blip virus potentially

sensitive to adjunct treatment. Those who initiated treatment late, however, had developed 2.5% more CTL escape variants at the blip than at T/F, making these individuals less likely to benefit from a vaccine therapy (59).

A 0.4% increase in the percentage of HLA-associated CTL escape variants was observed at the blip as compared to the T/F sequences of the chronic-treated participants, whereas no changes were observed in the acute-treated participants. Viral evolution in the former, in the presence of immune pressure and the absence of treatment, resulted in more CTL escape, whereas in the latter, the lack of viral diversity between the T/F and blip, was evident in the identical responses to immune pressure. The CTL escape mutation T242N, which developed within the TW10 Gag epitope of PID 272, was in response to selection pressure imposed by the participant's HLA B*58:01 allele. This mutation has been shown to be associated with reduced replication capacity, which in a chronic-treated participant would relatively reduce the viral load, and be somewhat advantageous in the absence of treatment, resulting in slow disease progression (60).

Over 80% of the CTL escape mutations in both acute and chronic-treated participants were not associated with the participants' HLA alleles, suggesting that there may be other potential drivers of CTL escape beyond the scope of this study. The use of CTL-inducing vaccines may have potential in treating acute infection prior to significant viral evolution in a population whose transmitted CTL escape mutations are known. The toggling, reversion and increase of CTL escape mutations in the chronic-treated participants was likely a consequence of immune pressure in the absence of ART, and this would disqualify these patients from potentially benefiting from a CTL vaccine that the acute-treated patients would likely been sensitive to. It would therefore be prudent to introduce CTL vaccines together with ART in acute infection in order to realize the best outcomes.

The potential use of bnAbs as prophylaxis or adjunct treatment has been the subject of recent studies (61). In light of this it was crucial to determine whether the rebound virus would be susceptible to bnAbs, by first determining whether mutations had developed in epitopes previously associated with antibody escape. We observed that the acute-treated participants not only lacked genetic diversity within Env (Table 3.3), but they did not develop bnAb escape mutations in any of the five binding sites (Table 3.6-10). This again was a consequence of early treatment (59). It is, therefore, unlikely that the rebound virus of these acute-treated participants, would exhibit differential sensitivity to bnAbs compared to the T/F virus, in an in-vitro neutralisation assay. Genetic diversity was however, observed in the PreART and blip env sequences (Table 3.3) of the chronic-treated participants, owing to viral evolution prior to

treatment. The CD4 binding site, the K97R mutation which was observed at PreART of PID 036 and the gp120/gp41 site mutations, N276E and E279P, at the blip of PID 272 (Table 3.5-6), suggest that these viruses may not be susceptible to bnAbs targeting these sites. The MPER, V1V2 and V3 loop sites which did not develop mutations, are likely to be sensitive to bnAbs, however conducting neutralisation assays would provide confirmation.

The participants of this study were self-reportedly adhering to treatment, yet simultaneously experiencing viral rebound. It was thus important to determine whether these viral blips were not merely a consequence of drug resistance brought on by non-adherence. It was noted that only one participant who initiated treatment late (PID 272), developed drug resistance to efavirenz, a NNRTI drug, which was used in combination with tenofovir and emtricitabine (NRTIs). Efavirenz and other NNRTI drugs generally have a low threshold for drug resistance and a long half-life (62). A recent study showed that when a treatment regimen containing efavirenz, tenofovir and emtricitabine was stopped abruptly, the patients developed resistance to efavirenz due to the diminishing effective dose in-vivo during its long half-life (62, 63). It was then shown that drug resistance would not develop if the two NRTIs were taken for up to 7 days following the last efavirenz dose (62, 63). This would then suggest that PID 272 may have discontinued all their treatment for a period of a week or more, prior to the occurrence of the blip. In the case of this participant, the viral blip was most likely the result of drug resistance.

Limitations

This study had a few limitations, the first being that adherence to treatment was self-reported and could not be verified at the beginning, in order to rule out participants who may have been potentially drug resistant, hence drug resistance analysis was included. The generation of NFLG SGAs in two overlapping fragments was unsuccessful, thus the more effective gene-specific alternative was employed, albeit labour and cost intensive. The sample size used was small due to sample unavailability and the ease with which amplicons could be generated from low viral load samples. A small number of unique sequences was generated, from a small sample size, thus limiting the inferences that could be made from the study. A larger sample size study would be essential in validating these findings. Neutralization assays which would have validated the bnAb escape data, were not conducted.

Conclusion

In summation, this study showed that drug resistance was not the primary cause of viral blips amongst the majority of the participants, however it may be necessary to perform a similar study on a larger population in order to substantiate these findings. The early initiation of treatment was shown to be associated with a lack of genetic diversity between the T/F and blip viruses. The impact of starting treatment late was shown in the increased number of CTL escape variants in the rebound virus. This is an essential factor to consider in the development of induced immune response strategies. The rebound virus in those who initiated treatment early, was more likely to be susceptible to bnAbs, however further investigation would be required to validate these findings.

REFERENCES

1. Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. Cold Spring Harbor perspectives in medicine. 2011;1(1):a006841.
2. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. The early spread and epidemic ignition of HIV-1 in human populations. science. 2014;346(6205):56-61.
3. De Cock KM, Adjorlolo G, Ekpini E, Sibailly T, Kouadio J, Maran M, et al. Epidemiology and transmission of HIV-2: why there is no HIV-2 pandemic. Jama. 1993;270(17):2083-6.
4. Hemelaar J. The origin and diversity of the HIV-1 pandemic. Trends in molecular medicine. 2012;18(3):182-92.
5. UNAIDS. Fact sheet: Global HIV & AIDS Statistics. Geneva, Switzerland: United Nations; 2021.
6. Africa SS. Mid-year Population Estimates. Pretoria, South Africa; 2021.
7. Kim H, Tanser F, Tomita A, Vandormael A, Cuadros DF. Beyond HIV prevalence: identifying people living with HIV within underserved areas in South Africa. BMJ global health. 2021;6(4):e004089.
8. Ramjee G, Daniels B. Women and HIV in sub-Saharan Africa. AIDS research and therapy. 2013;10(1):1-9.
9. Karim SSA, Baxter C. HIV incidence rates in adolescent girls and young women in sub-Saharan Africa. The Lancet Global health. 2019;7(11):e1470-e1.
10. Hake A, Pfeifer N. Prediction of HIV-1 sensitivity to broadly neutralizing antibodies shows a trend towards resistance over time. PLoS computational biology. 2017;13(10):e1005789.
11. HIV/AIDS UNJPo. Global AIDS Update. Geneva, Switzerland: United Nations; 2018 18 July 2018.
12. Arts EJ, Hazuda DJ. HIV-1 antiretroviral drug therapy. Cold Spring Harbor perspectives in medicine. 2012;2(4):a007161.
13. Jain V, Hartogensis W, Bacchetti P, Hunt PW, Hatano H, Sinclair E, et al. Antiretroviral therapy initiated within 6 months of HIV infection is associated with lower T-cell activation and smaller HIV reservoir size. The Journal of infectious diseases. 2013;208(8):1202-11.
14. Lewin S, Vesanen M, Kostrikis L, Hurley A, Duran M, Zhang L, et al. Use of real-time PCR and molecular beacons to detect virus replication in human immunodeficiency virus type 1-infected individuals on prolonged effective antiretroviral therapy. Journal of virology. 1999;73(7):6099-103.
15. Furtado MR, Callaway DS, Phair JP, Kunstman KJ, Stanton JL, Macken CA, et al. Persistence of HIV-1 transcription in peripheral-blood mononuclear cells in patients receiving potent antiretroviral therapy. New England Journal of Medicine. 1999;340(21):1614-22.
16. Zhang L, Ramratnam B, Tenner-Racz K, He Y, Vesanen M, Lewin S, et al. Quantifying residual HIV-1 replication in patients receiving combination antiretroviral therapy. New England Journal of Medicine. 1999;340(21):1605-13.
17. Di Mascio M, Markowitz M, Louie M, Hogan C, Hurley A, Chung C, et al. Viral blip dynamics during highly active antiretroviral therapy. Journal of virology. 2003;77(22):12165-72.
18. Dornadula G, Zhang H, VanUitert B, Stern J, Livornese Jr L, Ingerman MJ, et al. Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy. Jama. 1999;282(17):1627-32.
19. Yerly S, Perneger TV, Vora S, Hirschel B, Perrin L. Decay of cell-associated HIV-1 DNA correlates with residual replication in patients treated during acute HIV-1 infection. Aids. 2000;14(18):2805-12.
20. Hamlyn E, Ewings FM, Porter K, Cooper DA, Tambussi G, Schechter M, et al. Plasma HIV viral rebound following protocol-indicated cessation of ART commenced in primary and chronic HIV infection. PLoS One. 2012;7(8):e43754.
21. Pennings PS. HIV drug resistance: problems and perspectives. Infectious disease reports. 2013;5(S1):21-5.
22. Zdanowicz MM. The pharmacology of HIV drug resistance. American journal of pharmaceutical education. 2006;70(5).

23. Kearney MF, Wiegand A, Shao W, Coffin JM, Mellors JW, Lederman M, et al. Origin of rebound plasma HIV includes cells with identical proviruses that are transcriptionally active before stopping of antiretroviral therapy. Journal of virology. 2015;90(3):1369-76.
24. Lorenzo-Redondo R, Fryer HR, Bedford T, Kim E-Y, Archer J, Pond SLK, et al. Persistent HIV-1 replication maintains the tissue reservoir during therapy. Nature. 2016;530(7588):51.
25. Bozzi G, Watters S, Simonetti F, Anderson E, Gorelick R, Shao W, et al. THAA0104LB [LINE SEPARATOR] No evidence of ongoing replication in tissue compartments during combination antiretroviral therapy. Journal of the International Aids Society. 2016;19.
26. Deng K, Pertea M, Rongvaux A, Wang L, Durand CM, Ghiaur G, et al. Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations. Nature. 2015;517(7534):381-5.
27. Kearney MF, Spindler J, Shao W, Yu S, Anderson EM, O'Shea A, et al. Lack of detectable HIV-1 molecular evolution during suppressive antiretroviral therapy. PLoS pathogens. 2014;10(3):e1004010.
28. Kearney MF, Wiegand A, Shao W, McManus WR, Bale MJ, Luke B, et al., editors. Ongoing HIV replication during ART reconsidered. Open forum infectious diseases; 2017: Oxford University Press.
29. Van Zyl GU, Katusiime MG, Wiegand A, McManus WR, Bale MJ, Halvas EK, et al. No evidence of HIV replication in children on antiretroviral therapy. The Journal of clinical investigation. 2017;127(10):3827-34.
30. Maldarelli F. HIV-infected cells are frequently clonally expanded after prolonged antiretroviral therapy: implications for HIV persistence. Journal of virus eradication. 2015;1(4):237-44.
31. Bui JK, Sobolewski MD, Keele BF, Spindler J, Musick A, Wiegand A, et al. Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir. PLoS pathogens. 2017;13(3):e1006283.
32. Anderson EM, Maldarelli F. The role of integration and clonal expansion in HIV infection: live long and prosper. Retrovirology. 2018;15(1):1-22.
33. Wang Z, Simonetti FR, Siliciano RF, Laird GM. Measuring replication competent HIV-1: advances and challenges in defining the latent reservoir. Retrovirology. 2018;15(1):1-9.
34. Bozzi G, Simonetti F, Watters S, Anderson E, Gouzoulis M, Kearney M, et al. No evidence of ongoing HIV replication or compartmentalization in tissues during combination antiretroviral therapy: Implications for HIV eradication. Science advances. 2019;5(9):eaav2045.
35. Crowell TA, Danboise B, Parikh A, Esber A, Dear N, Coakley P, et al. Pretreatment and acquired antiretroviral drug resistance among persons living with HIV in four African Countries. Clinical Infectious Diseases. 2020.
36. Joos B, Fischer M, Kuster H, Pillai SK, Wong JK, Böni J, et al. HIV rebounds from latently infected cells, rather than from continuing low-level replication. Proceedings of the National Academy of Sciences. 2008;105(43):16725-30.
37. Bachmann N, von Siebenthal C, Vongrad V, Turk T, Neumann K, Beerenwinkel N, et al. Determinants of HIV-1 reservoir size and long-term dynamics during suppressive ART. Nature communications. 2019;10(1):1-11.
38. Luo L, Wang N, Yue Y, Han Y, Lv W, Liu Z, et al. The effects of antiretroviral therapy initiation time on HIV reservoir size in Chinese chronically HIV infected patients: a prospective, multi-site cohort study. BMC infectious diseases. 2019;19(1):1-8.
39. Erlich RL, Jia X, Anderson S, Banks E, Gao X, Carrington M, et al. Next-generation sequencing for HLA typing of class I loci. BMC genomics. 2011;12(1):1-13.
40. Ndhlovu ZM, Chibnik LB, Proudfoot J, Vine S, McMullen A, Cesa K, et al. High-dimensional immunomonitoring models of HIV-1–specific CD8 T-cell responses accurately identify subjects achieving spontaneous viral control. Blood, The Journal of the American Society of Hematology. 2013;121(5):801-11.
41. Shan L, Deng K, Shroff NS, Durand CM, Rabi SA, Yang H-C, et al. Stimulation of HIV-1-specific cytolytic T lymphocytes facilitates elimination of latent viral reservoir after virus reactivation. Immunity. 2012;36(3):491-501.

42. Lynch RM, Tran L, Louder MK, Schmidt SD, Cohen M, DerSimonian R, et al. The development of CD4 binding site antibodies during HIV-1 infection. Journal of virology. 2012;86(14):7588-95.
43. Wu X, Wang C, O'Dell S, Li Y, Keele BF, Yang Z, et al. Selection pressure on HIV-1 envelope by broadly neutralizing antibodies to the conserved CD4-binding site. Journal of virology. 2012;86(10):5844-56.
44. Moore PL, Sheward D, Nonyane M, Ranchobe N, Hermanus T, Gray ES, et al. Multiple pathways of escape from HIV broadly cross-neutralizing V2-dependent antibodies. Journal of virology. 2013;87(9):4882-94.
45. Lynch RM, Tran L, Louder MK, Schmidt SD, Cohen M, Members CCT, et al. The development of CD4 binding site antibodies during HIV-1 infection. Journal of virology. 2012;86(14):7588-95.
46. Lynch RM, Wong P, Tran L, O'Dell S, Nason MC, Li Y, et al. HIV-1 fitness cost associated with escape from the VRC01 class of CD4 binding site neutralizing antibodies. Journal of virology. 2015;89(8):4201-13.
47. Burton DR, Poignard P, Stanfield RL, Wilson IA. Broadly neutralizing antibodies present new prospects to counter highly antigenically diverse viruses. Science. 2012;337(6091):183-6.
48. Freund NT, Wang H, Scharf L, Nogueira L, Horwitz JA, Bar-On Y, et al. Coexistence of potent HIV-1 broadly neutralizing antibodies and antibody-sensitive viruses in a viremic controller. Science translational medicine. 2017;9(373).
49. Ahmed Y, Tian M, Gao Y. Development of an anti-HIV vaccine eliciting broadly neutralizing antibodies. AIDS research and therapy. 2017;14(1):1-6.
50. Rademeyer C, Korber B, Seaman MS, Giorgi EE, Thebus R, Robles A, et al. Features of recently transmitted HIV-1 clade C viruses that impact antibody recognition: implications for active and passive immunization. PLoS pathogens. 2016;12(7):e1005742.
51. Dong KL, Moodley A, Kwon DS, Ghebremichael MS, Dong M, Ismail N, et al. Detection and treatment of Fiebig stage I HIV-1 infection in young at-risk women in South Africa: a prospective cohort study. The Lancet HIV. 2018;5(1):e35-e44.
52. Salazar-Gonzalez JF, Salazar MG, Learn GH, Fouda GG, Kang HH, Mahlokozera T, et al. Origin and evolution of HIV-1 in breast milk determined by single-genome amplification and sequencing. Journal of virology. 2011;85(6):2751-63.
53. Grossmann S, Nowak P, Neogi U. Subtype-independent near full-length HIV-1 genome sequencing and assembly to be used in large molecular epidemiological studies and clinical management. Journal of the International AIDS Society. 2015;18(1):20035.
54. Chopera DR, Woodman Z, Mlisana K, Mlotshwa M, Martin DP, Seoighe C, et al. Transmission of HIV-1 CTL escape variants provides HLA-mismatched recipients with a survival advantage. PLoS pathogens. 2008;4(3):e1000033.
55. Adolescents PoAGfAa. Guidelines for the Use of Antiretroviral Agents in Adults and Adolescencts with HIV. In: science HaH, editor. 2021. p. G-13.
56. Brumme ZL, John M, Carlson JM, Brumme CJ, Chan D, Brockman MA, et al. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. PloS one. 2009;4(8):e6687.
57. Liu R, Simonetti FR, Ho Y-C. The forces driving clonal expansion of the HIV-1 latent reservoir. Virology journal. 2020;17(1):1-13.
58. Altfeld M, Rosenberg ES, Shankarappa R, Mukherjee JS, Hecht FM, Eldridge RL, et al. Cellular immune responses and viral diversity in individuals treated during acute and early HIV-1 infection. The Journal of experimental medicine. 2001;193(2):169-80.
59. Simonetti FR, Kearney MF. influence of ART on HIV genetics. Current Opinion in HIV and AIDS. 2015;10(1):49.
60. Huang K-HG, Goedhals D, Carlson JM, Brockman MA, Mishra S, Brumme ZL, et al. Progression to AIDS in South Africa is associated with both reverting and compensatory viral mutations. PloS one. 2011;6(4):e19018.
61. Huang D, Tran JT, Olson A, Vollbrecht T, Tenuta M, Guryleva MV, et al. Vaccine elicitation of HIV broadly neutralizing antibodies from engineered B cells. Nature communications. 2020;11(1):1-10.

62. Bennett K, Rowley C, Lockman S, Powis K, Maruapula D, Ajibola G, et al. Drug resistance after cessation of efavirenz-based antiretroviral treatment started in pregnancy. Southern African journal of HIV medicine. 2020;21(1):1-4.
63. Chi BH, Sinkala M, Mbewe F, Cantrell RA, Kruse G, Chintu N, et al. Single-dose tenofovir and emtricitabine for reduction of viral resistance to non-nucleoside reverse transcriptase inhibitor drugs in women given intrapartum nevirapine for perinatal HIV prevention: an open-label randomised trial. The Lancet. 2007;370(9600):1698-705.