COMPARTMENTALIZATION, ADAPTIVE EVOLUTION AND THERAPEUTIC RESPONSE OF HIV-1 IN THE GASTROINTESTINAL TRACT (GIT) OF AFRICAN PATIENTS INFECTED WITH SUBTYPE C: IMPLICATIONS FOR THE ENHANCEMENT OF THERAPEUTIC EFFICACY

PHETOLE WALTER MAHASHA

B.Sc. (Life Sciences), B.Sc. Honors (Biochemistry), M.Sc. in Medical Biochemistry (Medical Sciences)

This thesis is presented in partial fulfillment of the requirements for a Doctor of Philosophy in Medical Immunology in the Department of Immunology, Faculty of Health Sciences at the University of Pretoria, Pretoria, South Africa.

Sponsored by: European Commission (EC), Sante 2007/147-790; National Research Foundation (Unlocking the Future 61509 and NRF_SFH20070717000335); Medical Research Council (MRC) and Poliomyelitis Research Foundation (PRF, Grant #: 08/35).

University of Pretoria

June 2014

ABSTRACT [500 words]

Background: Due to its continuous exposure to food antigens and microbes, the gastrointestinal tract (GIT) is in a constant state of low level immune activation and contains an abundance of activated $CCR5^+CD4^+$ T lymphocytes, the primary target HIV-1. As a result, the GIT is a site of intense viral replication and severe $CD4^+$ T cell depletion, a process that begins during primary HIV-1 infection and continues at a reduced rate during chronic infection in association with increased production of pro-inflammatory cytokines, a breakdown in the epithelial barrier, microbial translocation, systemic immune activation and the continued recruitment and infection of new target cells. AntiRetroviral Therapy (ART) is only partially effective in reversing these pathogenic changes. Despite the importance of the GIT in HIV-1 pathogenesis, and as a reservoir of persistent virus during ART, little is known about the diversity of HIV-1 in the GIT, or how different tissues in the GIT respond to ART.

Objectives: Primary objectives of this thesis were to: 1) characterize the diversity of HIV-1 RNA variants in different parts of the GIT; 2) determine whether there is compartmentalized evolution of HIV-1 RNA variants in the GIT and whether these variants are likely to have different biological properties; 3) investigate the impact of ART on immune restoration in the GIT.

Methods: A prospective study of the duodenum, jejunum, ileum and colon of African AIDS patients with chronic diarrhea and/or weight loss, sampled before and during 6 months of ART. RNA extracted from gut biopsies was reverse transcribed and PCR amplified. *Env* and *gag* PCR fragments were cloned, sequenced and subjected to extensive phylogenetic analysis; *pol* PCR fragments were analyzed for drug resistance. $CD4^+$, $CD8^+$ and $CD38^+CD8^+$ T cells levels in biopsies collected at baseline (duodenum, jejunum, ileum and colon) and after 3 (duodenum) and 6 (duodenum and colon) months of ART were quantified by flow cytometry and immunohistochemistry, plasma and tissue VL by the Nuclisens assay.

Results: Viral diversity varied in different regions of the GIT with *env* HIV-1 RNA variants being significantly more diverse than *gag* variants. *Gag* HIV-1 RNA variants were widely dispersed among all tissue compartments. Some *env* variants formed tight monophyletic clusters of closely related viral quasispecies, especially in the colon, a finding that is suggestive of compartmentalized viral replication and adaptive evolution. $CD4^+$ T cell and VL levels were significantly lower, while $CD8^+$ including activated $CD38^+CD8^+$ T cell levels were higher in the duodenum and jejunum versus the colon. After 6 months of ART, a significant but incomplete recovery of $CD4^+$ T cells was observed in the colon but not in the duodenum. Failed restoration of $CD4^+$ T cells in the duodenum was associated with non-specific enteritis and $CD8^+$ T cell activation.

Conclusions: These results advance our understanding of the GIT as a host-pathogen interface by providing new insights into the diversity, evolution and dissemination of HIV-1 variants in the GIT. Strategies aimed at decreasing immune activation, especially in the small intestine, may be highly beneficial in enhancing the therapeutic efficacy of ART.

Key words: HIV-1; Gastrointestinal tract; Antiretroviral therapy; viral RNA quasispecies; genetic diversity; CD4 reconstitution; intestine; Africa; immune activation.

DECLARATION

This work was conducted as a part of a large collaborative study funded by several granting agencies. I, the undersigned, declare that all of the genetic and phylogenetic studies are my own original work and that I made major contributions to the immunological work in Chapter 4. This work has not, or previously in the past, been submitted at any University for any degree.

June 2014

…… ……

Phetole W. Mahasha (Student number: 28587805) *Date*

TABLE OF CONTENTS:

FIGURE LEGENDS:

PAGES:

CHAPTER 1

CHAPTER 2

PAGES:

CHAPTER 3

D: Representative Maximum Likelihood tree showing extensive inter-mixing of *gag* HIV-1 RNA quasispecies among different compartments in the GIT of patient HIV019.

E: Representative Maximum Likelihood tree showing extensive inter-mixing of *gag* HIV-1 RNA quasispecies among different compartments in the GIT of patient HIV021.

F: Representative Maximum Likelihood tree showing extensive inter-mixing of *gag* HIV-1 RNA quasispecies among different compartments in the GIT of patient HIV029.

CHAPTER 4

Figures: **160-163**

4.1: Depletion of $CD4^+$ T cells in cART-naïve patients with late-stage HIV-1 infection was more severe in the small than in the large intestine.

4.2: Increased frequency of $CD8^+$ T cells and activated $CD38^+CD8^+$ T cells in the upper small intestine compared to the colon of cART-naïve AIDS patients.

4.3: Prior to the initiation of cART, HIV RNA levels were higher in the colon than in the upper small intestine of African AIDS patients.

4.4: Impaired $CD4^+$ T cell restoration and sustained $CD8^+$ T cell activation in the duodenum, but not in the colon, of NRTI/NNRTI-treated AIDS patients.

ABBREVIATIONS:

%	percentage
^{o}C	degree celcius
3TC	Lumivudine
AFB	Acid Fast Bacilli
AIDS	Acquired Immunodeficiency Syndrome
APOBEC3G	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G
ART	Antiretroviral Therapy
AZT	Azidothymidine
CA	Capsid protein, p24
CCR5 (R5)	CC-Chemokine Receptor-5
CD	Crohn's Disease
CD4	Cluster of Differentiation 4
CDC	Centre for Disease Control
cDNA	Complimentary DioxyriboNuceic Acid
CHI	Chronic HIV infection
CNS	Central Nervous System
CPX	Complex
CRFs	Circulating inter-subtype Recombinant Forms
CTL	Cytotoxic T-Lymphocyte
CXCR4 (X4)	CXC-Chemokine Receptor-4
d4T	Stavudine

DC	Dendritic Cell
ddC	Zalcitabine or 2'-3'-dideoxycytidine
ddI	Didanosine
DNA	DeoxyriboNucleic Acid
dNTP	Deoxynucleotide Triphosphate
EDTA	Ethylenediaminetetraacetic Acid
EFV	Efavirenz
Env	Envelop
ESCRT	Endosomal Sorting Complexes Required for Transport
F81	Felsentein 1981
F84	Felsentein 1984
FACS	Fluorescence-Activated Cell Sorting
FDA	Food and Drug Administration
Gag	Group-specific antigen
GALT	Gut-Associated Lymphoid Tissue
GIT	Gastrointestinal Tract
gp120	Glycoprotein 120
gp41	Glycoprotein 41
GTR	General Time Reversible
GVHD	Graft versus-host disease
HAART	Highly Active Antiretroviral Therapy
HCV	Hepatitis C
HIV-1	Human Immunodeficiency Virus type 1

HIV-2	Human Immunodeficiency Virus type 2
HKY85	Hasegawa, Kishino and Yano 1985
HLA	Human Leukocyte Antigen
HTLV-III	Human T-cell lymphotropic virus type 3
IBD	Inflammatory Bowel Diseases
I-domain	Interaction domain
IEL	Intraepithelial Lymphocytes
IgA	Immunoglubulin A
IL	Interleukin
IN	Integrase enzyme
IRIS	Immune Reconstitution Inflammatory Syndrome
JC69	Jukes and Cantor 1969
K80	Kamura 1980
LAV	Lymphoadenopathy-associated virus
L-domain	Late domain
LN	Lymph Node
LPS	Lypopolysaccharise
LTR	Long Terminal Repeat
MA	Matrix protein, p17
MAdCAM-1	Mucosal vascular Addressin Cell-Adhesion Molecule 1
M-domain	Membrane targeting domain
ME	Minimum Evolution
MEGA	Molecular Evolutionary Genetics Analysis

$MgSO_4$	Magnesium Sulfate
MHC	Major Histocompatibility Complex
ML	Maximum Likelihood
MMLV	Moloney Murine Leukemia Virus
MP	Maximum Parsimony
MRCA	Most Recent Common Ancestor
M-tropic	Macrophage tropic
MVBs	Multi-vesicular bodies
NC	Nucleocapsid protein
Nef	Negative Regulatory Factor
NF-κB	Nuclear Factor-Kb
NF-IL6	Nuclear Factor-Interleukin 6
NNRTIs	Non-nucleoside Reverse Transcriptase Inhibitors
NRTIs	Nucleoside Reverse Transcriptase Inhibitors
NSI	Non-Syncytium Inducing
NVP	Nevirapine
OIs	Opportunistic Infections
PBMCs	Peripheral Blood Mononuclear Cells
PCP	*Pneumocystis carini* pneumonia
PCR	Polymerase Chain Reaction
PEPFAR	President's Emergency Plan for AIDS Relief
PIC	Pre-Integration Complex
PIs	Protease Inhibitors

PMTCT	Prevention of Mother-To-Child Transmission
Pol	Polymerase
$PR55^{Gag}$	Gag polyprotein
PHI	Primary HIV infection
PR	Protease
Rev	Regulator of Virus
RNA	RiboNucleic Acid
RNase H	Ribonuclease H
RNP	RiboNucleoProtein complex
RT	Reverse Transcriptase
SGS	Single genome sequences
SI	Syncytium-Inducing
SIVs	Simian Immunodeficiency Viruses
SLPI	Secretory Leukocyte Protease Inhibitor
TAMs	Thymidine Analogue Mutations
TAR	Trans-Activation Response region
Tat	Trans-activator of Transcription
TDF	Tenofovir
TIBO	Tetrahydroimidazobenzodiazepin/Thiobenzimidazolone
TN93	Tamura and Nei 1993
TNF	Tumor Necrosis Factor
Trim5α	Tripartite motif-containing protein 5 alpha
tRNA	Transfer RiboNucleic Acid

T-tropic	T-cell tropic
UC	Ulcerative Colitis
UPGMA	Unweighted Pair Group Method with Arothmetic means
URFs	Unique Recombinant Forms
Vif	Viral infectivity factor
VL	Viral Load
Vpr	Viral protein R
Vpu	Viral protein U
VS	Virological Synapse
WHO	World Health Organization
ZDV	Zidovudine

ACKNOWLEDGEMENTS:

My sincere thanks are extended to:

Mother for always been there for me and guiding me to be the best I can be, thanks Mom.

My Wife (Plantinah Maleka) and my children (Thakgalo and Theeletso) for always believing and supporting me with my studies

Professor Sharon Cassol, my supervisor, for sharing her knowledge in HIV-1/AIDS and for her guidance and encouragement throughout the duration of my studies. Sharon I wouldn't have picked a better mentor you are the best and thank you so very much for you guidance, May God Bless you with life in abundance.

Prof. Tulio de Oliveira, my co-supervisor, for sharing his knowledge in bioinformatics, phylogenetic analyses and database management, it sure was exciting working with you Tulio and I really appreciate.

Dr. Chris Seebregts for providing expertise in database management

Drs. Edana Cassol and Susan Malfeld for providing immunological data and sharing their expertise and knowledge in HIV-1 immunopathology and flow cytometry

Drs. Theresa Rossouw and Schalk van der Merwe for their clinical expertise and dedication in obtaining clinical specimens

Dr. Siva Danaviah, Dr Justen Manasa, Mrs Prevashinee Padayachee, Hloniphile Mthiyane and everyone in the Africa centre laboratory at the University of Kwazulu Natal for their assistant with my experiments.

The endoscopy nurses of the Unitas and Pretoria East GI-units and to Dr. Mark Theron, the study anesthetist, for their dedication and commitment to patient care. Most of all, we thank our patients and volunteers. Without their participation and commitment this study would not have been possible.

The Medical Research Council (MRC), National Research Foundation (NRF), Poliomyelitis Research Foundation (PRF) and the EU Grant Sante 2007 174-790 for providing financial support for the study

Finally, I want to thank God for protecting and enabling me, and giving me good health and leading me through the journey of life and towards to my destiny.

ABSTRACTS AND PRESENTATIONS:

1) Rossouw T, **Mahasha P**, Malherbe G, Manasa J, van Dyk G., Cassol S, Seebregts C, de Oliveira T. SATuRN, the Southern African Treatment and Resistance Network: Application to the Management and Surveillance of HIV-1 Drug Resistance in a Public Health Setting in Pretoria. 5th South African AIDS Conference, Durban, South Africa, June 7-10, 2011.
2) Cassol E, Malfeld S, **Mahasah P**, Malherbe G, van der Merwe S, Cassol S, Seebregts C, Alfano M, Poli G, Rossouw T. Persistent microbial translocation and immune activation in HIV-1-infected South Africans on antiretroviral therapy. 2nd Trans-Disciplinary HIV and AIDS Research Indaba: Crossing Disciplinary Boundaries. University of Pretoria, South African, February 27-March 1, 2010.
3) Malherbe G, Cassol S, Cassol E, Kronova L, **Mahasha P**, Malfeld S, Nzeku M, Rossouw T. HIV drug resistance testing using the oligonucleotide ligation assay. 2nd Trans-Disciplinary HIV and AIDS Research Indaba: Crossing Disciplinary Boundaries. University of Pretoria, South African, February 27-March 1, 2010.
4) Rossouw T, Cassol E, Malfeld S, **Mahasha P**, Malherbe G, Seebregts C, Cassol S, van der Merwe S, Slavik T. Clinical Profiles of HIV-1-infected South Africans with chronic diarrhoea and/or wasting syndrome. 2nd Trans-Disciplinary HIV and AIDS Research Indaba: Crossing Disciplinary Boundaries. University of Pretoria, South African, February 27-March 1, 2010.
5) Malfeld S, Cassol E, Rossouw T, **Mahasha P**, Malherbe G, Anderson R, van der Merwe *S.* Effect of antiretroviral therapy (ART) on memory and effector T cells in the duodenal mucosa. HIV Immunobiology: From Infection to Immune Control (X4). Keystone, Colorado, March 22-27, 2009.
6) **Mahasha P**, Cassol E, Malfeld S, Malherbe G, van der Merwe S, Rossouw T, Cassol S, Seebregts C, de Oliveira T. Differential compartmentalization of viral reservoirs in the gastrointestinal tract (GIT) of HIV-1-infected patients with late-stage subtype C infection. 15th International Bioinformatics Workshop on Virus Evolution and Molecular Epidemiology, Rotterdam, The Netherlands, September 7-11, 2009.

PUBLICATIONS:

1) **Mahasha P**, Cassol E, Danaviah S, Bond R, Seebregts C, van der Merwe S, Rossouw T, de Oliveira T. Greater genetic diversity in the large vs. small intesine of antiretroviral therapy naïve south African patients with chronic diarrhoea and (or) weight loss (submitted).
2) Cassol E, Rossouw T, Malfeld S, **Mahasha P**, Slavik T, Seebregts C, Bond R, du Plessis J, Janssen C, Roskams, Nevens F, Alfano M, Poli G, van der Merwe SW. $CD14^{+}$macrophages in the colon of AIDS patients express pro-inflammatory cytokines and are linked to microbial translocation (submitted).
3) Cassol E, Malfeld S, **Mahasha P**, Bond R, Slavik T, Seebregts C, Poli G, Cassol S, van der Merwe S, Rossouw T. Impaired $CD4^{+}$ T cell restoration in the small vs. large intestine of HIV-1^{+} South Africans receiving combination antiretroviral therapy. Journal of Infectious Diseases 2013;**208**:1113–22
4) Cassol E, Malfeld S, **Mahasha P,** van der Merwe S, Cassol S, Seebregts C, Alfano M, Poli G, Rossouw T. Persistent Microbial Translocation and Immune Activation in HIV-1–Infected South Africans Receiving Combination Antiretroviral Therapy. Journal of infectious Diseases 2010; **202:**722-33.

CHAPTER 1: LITERATURE REVIEW:

1.0 THE UNFOLDING OF A PANDEMIC

1.0.1 *The discovery of HIV-1/AIDS*

The first cases of HIV-1/AIDS were detected in the summer of 1981 when young gay men in New York and Los Angeles began presenting with, and dying of, *Pneumocystis carini* pneumonia (PCP) and other rare opportunistic infections that their immune systems should have been able to fend off. Many of these men also developed a relatively rare and aggressive form of Kaposi's sarcoma, a cancer that is typically observed in middle-aged men living in the Mediterranean region [1; 2]. Later that year, a report was published describing a marked increase in prescriptions for the treatment of PCP and other severe opportunistic infections (OIs) [3]. This report is often cited as the beginning of the AIDS epidemic. Physicians were baffled by this new "gay plague" and because the condition was acquired and associated with a number of different OIs, it was given the name "acquired immunodeficiency syndrome" or "AIDS" by the Centre for Disease Control (CDC) in Atlanta, Georgia [4-6]. Early theories attributed AIDS to factors such as "lifestyle", immune overload, the use of amyl or butyl nitrate "poppers", a reaction to semen and multiple sex partners or an unrecognized fungal infection [7].

However, in late 1981 and 1982, the first cases of PCP were reported in injecting drug users (IDUs), hemophiliacs and among individuals of Haitian origin. Soon thereafter, cases were discovered outside of America [8-13] and it became apparent that AIDS was an infectious disease transmitted by body fluids and exposure to contaminated blood and blood products. In 1983, Barré-Sinoussi *et al*., isolated a retrovirus from a lymph node taken from a patient with lymphadenopathy in the laboratory of Jean-Claud Chermann in collaboration with the research group of Luc Monagnier at the Pasteur Institute in Paris. The virus, named lymphoadenopathy-associated virus (LAV) was suspected to be the cause of AIDS [14]. Soon thereafter, Robert Gallo's group in America reported that a virus, which he called human T-cell lymphotropic virus type 3 (HTLV-III), appeared to be the causative agent of AIDS [15; 16]. Subsequent events in the evolution of the global HIV-1/AIDS pandemic (reviewed in 7) included the identification of CD4 as the viral receptor (1984), the development of a licensed FDA approved blood test (1985); renaming of the virus from LAV/HTLVIII to HIV (1986), the recognition that "Slims" disease in Africa caused by HIV (1986); the introduction of AZT as the first antiretroviral drug (1987); the introduction of AZT for the Prevention of Mother-To-Child Transmission (PMTCT, 1994), and the first use of Highly Active Antiretroviral

Therapy (HAART, 1996). Despite this rapid rate of discovery and ongoing progress in the development of new therapeutics, the number of HIV-1 infections has continued to expand at an alarming rate, especially in the developing world increasing from 20,303 cases at the end of 1985 [17] to 8-10 million in 1990; 14-15 million in 1993; 23 million in 1996; 34.3 million in 2000; > 40 million in 2006 and as many as 60 million infections in 2010 (**Figure 1.1.**).

1.0.3. HIV-1/AIDS in Africa

Africa is a continent plagued by war and ethnic conflict. However, it is the HIV-1/AIDS epidemic that is proving to the continent's most formidable foe. AIDS has already killed ten times more Africans than all of the armed conflicts on the continent, and the pandemic is now the leading cause of death among Africans of all ages [18]. Sub-saharan Africa remains the region most severely affected by the pandemic, accounting for 90% of all childhood and 67% of all adult infections worldwide. In 2007, 38% of AIDS deaths occurred in sub-Saharan Africa. In several African countries, between 30 and 50% of children have lost one or both parents to AIDS and, by the end of 2010, 40 million children were orphaned by the disease. New infections continue to occur at an estimated rate of >4 million per annum with the majority (45%) of these transmissions occurring in Africa and the developing world, in young people aged 15-24 years [19]. The HIV epidemic in South Africa is the largest in the world with an estimated 5.7 million South Africans living with the disease as of 2007 (**Figure 1.2.**). Although recent data suggests that the country's epidemic may be stabilizing [20]**,** there is no concrete evidence of a major change in HIV-causing behaviour.

In addition to being a public health emergency in its own right, the African AIDS epidemic, by generating a large population of immunosuppressed individuals, is fueling a re-emergence of tuberculosis and other infectious disease, decreasing life expectancy and reversing the gains that have been made in alleviating child mortality. At the individual level, AIDS is widening the gap between rich and poor – increasing medical costs, forcing households to sell their assets and increasing the number of dependants being supported by low (or no) income earners [21]. Although the impact on civil society is not yet fully understood, AIDS is also having a serious effect on education and the public service sector. School enrollment is declining as children and young adults drop out of school to look after families affected by the virus. By affecting young people, AIDS is also changing the age composition of the work force, leading to a loss of experienced staff and expertise, eroding

senior management and undermining social structures within the business community, health care sector, academic institutions, police force and military. On the Human Development Index, South Africa has fallen 35 places with AIDS being the single greatest contributor to this decline [22].

Although initial treatment efforts have been promising and similar to those in developed countries [23], long-term data of the efficacy of ART in Africa is scarce and there are insufficient financial and logistical resources. Given the sheer magnitude of the problem, it is unlikely that treatment alone will reverse the AIDS epidemic in Africa. In addition to the high costs of medications, life-long ART/HAART is complicated by issues of toxicity and drug resistance. More importantly, drugs are not a "cure" for AIDS. What is urgently needed for Africa and the developing world is a safe, affordable and efficacious vaccine (or microbicide) that can prevent new infections. Many scientists believe that the failure to develop an AIDS vaccine is due to a lack of scientific knowledge relating to the pathogenesis and mechanisms of HIV-1 genetic evolution. As described in this thesis, increased knowledge of HIV-1 pathogenesis and genetic evolution in the gastrointestinal tract is likely to be a key component of this critical new research.

1.1. THE VIRUS

1.1.1. The origins and salient characteristics

The key to understanding the origins of HIV came from the discovery that the human virus was closely related to simian immunodeficiency viruses (SIVs), primate lentiviruses that are present in more than 20 different species of African monkeys [24]. Phylogenetic analysis has revealed that two distinct types of HIV, HIV-1 and HIV-2, entered the human population approximately 50-80 years ago through multiple cross-species transmissions (or zoonoses). These studies have also shown that HIV-1 and HIV-2 have vastly different origins, evolutionary history and outcomes [25; 26]. HIV-1 is most closely related to $SIV_{cpz,}$ a virus found in geographically isolated chimpanzee communities of southern Cameroon [27-29]. By contrast, HIV-2 shares a close relationship with SIVsm, a virus that is found at high prevelance in sooty mangabeys in West Africa [30]. It has been postulated that HIV-1 originated from three independent cross-species transmissions, giving rise to pandemic (group M) and non-pandemic (group N and O) clades of HIV-1 (31). Recently, a putative

new group, designated P, was detected in several patients from Cameroon (32; 33). It is generally accepted that the three HIV-1 groups: Main (M), outlier (O) and non-major, non-outlier (N) arose from independent zoonotic events that occurred between 1915 and 1941, while HIV-2 is likely to have emerged from seven different zoonoses, presumably through the hunting and killing of these animals for food [27; 28; 34]. After their introduction into humans, each of these viruses then diversified into several genetic subtypes. Unlike HIV-1 group M, which has evolved into a devastating global pandemic, the disease impact of HIV-1 groups O and N, and HIV-2, is limited, both with respect of prevalence and geographic spread. HIV-2 infections are restricted to West Africa and regions of India and appear to be less transmissible and less pathogenic than HIV-1 [35; 36]. HIV-1 group O and N viruses are limited to a small number of people living in, or epidemiologically linked to Western and Central Africa [37].

The ability of HIV-1 to establish life-long infection has provided the virus with the capacity to continually evolve, both within an individual host and among different population groups [25; 38]. Some of the most salient features of HIV-1 include the heavy glycosylation of its envelope glycoprotein, Gp120, which allows the virus to escape from neutralizing antibodes; the fact that HIV-1 targets $CD4^+$ lymphocytes, a key cell involved in the integration of the immune response system; the ability of the HIV-1 to infect and become integrated into the genomic DNA of long-lived cellular reservoirs, and the potential of HIV-1 to mutate and evade both immune control and therapeutic measures [34; 39; 40].

1.1.2. Subtype diversification and geographic spread

HIV-1 is the most rapidly evolving RNA virus in recorded human history. As a result, the HIV-1 group M viruses that dominate the global pandemic, have evolved into nine major subtypes (A-D, F-H, J and K) and at least 61 different circulating inter-subtype recombinant forms (CRFs) [41; http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html, last accessed June 12, 2014] as well as numerous unique recombinant forms (URFs) [37; 42]. At least three epidemiologically unlinked full-length genomic sequences are needed to define a viral strain as a subtype or a CRF. Some subtypes, namely A and F, have two or three distinct lineages referred to as subtypes (A1, A2 and A3; F1 and F2). CRFs are defined by their specific structure and the subtypes they contain, i.e. A and G in CRF02_AG. When more than three subtypes are present, the designation “cpx” (complex) is given. These recombinant

strains are generated in individuals who are dually infected with more than one subytpe. The two most dominant CRF are CRF01_AE found at high prevalence in Southeast Asia and CRF02_AG that is found in West Africa **(Figure 1.3.)**. URFs are unique recombinant viruses that have a unique structure and subtype composition but that have been identified in only a single individual [37]. Subtype B is relatively widespread, dominating the epidemics in the Americas, Western Europe and Australia; subtype C, a virus that accounts for more than 50% of all global infections predominates in southern Africa, India, Nepal and China, fuelling the uncontrolled spread of HIV-1 throughout Botswana, Malawi, Zamibia, Namibia and South Africa; D is generally limited to East and Central Africa; an A/E mosaic is prevalent in Thailand, China, the Philippines and Central Africa; F occurs primarily in Central Africa, South America and Eastern Europe; G and A/G recombinants are found in Western and Eastern Africa as well as Central Europe; J is restricted to Central America and K has recently been identified in the Congo and Cameroon [37; 43; 44]. However, more subtypes and CRFs are constantly being discovered and increased global migration is shaping the emergence of new infection patterns. B, the most extensively studied subtype, constitutes <15% of the entire global pandemic.

1.1.3. Biological and public health consequences of HIV-1 diversity

In contrast to hepatitis C (HCV) for which there are clear-cut differences, not only in the geographic distribution of different viral subtypes but also in their transmission route and response to therapy, the relationships between HIV-1 subtype diversity, transmissibility and disease progression are not well defined and are complicated by factors such as host genetics, societal practices, mode of transmission and virological parameters, all of which are difficult to control. Subtype B is spread primarily through homosexual intercourse and intravenous drug abuse. In contrast, subtypes A, C, D and E, viruses that predominate in Africa and Asia are transmitted primarily through heterosexual sex or from mothers to their infants, either *in utero,* during labor and delivery**,** or through breast feeding. Although largely unconfirmed, several studies have suggested that transmissibility and AIDS progression differ depending on the infecting viral subtype [44; 45]. A study of pregnant women in Kenya suggested that C viruses may be more transmissible than subtypes A and D due to increased shedding of HIV-1-infected cells from the vaginal mucosa [46], while another study found that, relative to subtype B, C viruses exhibit higher rates of mother-to-child transmission [47]. With respect

to disease progression, a study conducted by Kanki et al. in female sex workers in Senegal found that women infected with subtypes C, D and G had increased rates of progression relative to those infected with A viruses [48]. A larger Ugandan study conducted in 2002, detected a more rapid progression to death among patients infected with subtype D compared to A [49]. Other studies have reported that the risk of NNRTI resistance following single dose nevaripine (NVP) to prevent mother to child transmission appears higher for subtype D than subtype A. Other investigators have reported that K65R resistance mutations develop more rapidly in patients infected with HIV-1 subtype C compared to other subtypes and that when administering tenofovir (TDF) to C-infected patients the reverse transcriptase gene needs to be careful monitoring for the selection of K65R mutations (50-52). A study from Kenya reported that patients infected with subtype D had a higher rate of morbidity and a faster CD4+ T cell decline relative to those infected with subtypes A and C [53], possibly due to the higher prevalence of dual-tropic (CCR5/CXCR4) isolates among patients infected with D viruses [54]. However, it is extremely difficult to design additional studies to confirm/disprove these findings. Such studies would have to be performed in African countries that have a high prevalence of multiple subtypes.

Although little data is currently available, the impact of HIV-1 genetic diversity on drug susceptibility and resistance has been somewhat easier to investigate. These studies have shown that the RT genes of HIV-1 group O and HIV-2 viruses are resistant to all drugs within the entire NNRTI class due to the presence of naturally occurring Y181C and Y181I mutations, respectively [55 -57]. Subtype F viruses are resistant to TIBO, but remain senstive to other NNRTIs including nevirapine and delaviridine, as well the NRTIs and PIs [58]. Similar to subtype B, C viruses appear to be sensitive to most drugs. However, some C isolates may exhibit inherent resistance to NNRTIs due to the presence of a G190A mutation [59]. At present, little is known about the contribution of recombination to either the viral fitness or the emergence of resistance. Analyses of the protease gene, has indicated that many non-B subtypes contain naturally-occurring polymorphisms that facilitate the development of PI resistance [44]. These findings are of particular concern due to the fact that suboptimal therapy and poor patient adherence to drug regimens is becoming increasingly common in the developing world where non-B subtypes predominate.

In addition to RT and protease, other genes vary considerably among subtypes, including the HIV-1 envelope gene, *gp120*, and the long terminal repeat (LTR) sequences containing transcriptional promoters of HIV-1 which vary in both sequence and number [60-

62]. The NF-κB binding site varies from just one site in E, to two in B and 3-4 in subtype C and the AP-1 transcription factor binding site exists as one site in subtypes C, E and G, as two in A and F and none in B and D [61; 63]. Subtype B, but not subtypes A, C, D or O, contains a site for the NF-IL6 transcriptional factor that activates the HIV-1 LTR in monocytic cells [60; 63]. Thus, it is not unexpected to find that NFκB, Rel-p65 and TNF activate the replication of C viruses to a greater extent than in subtypes B or E. Other important variations occur in the **regulatory (*nef, tat, rev* and *vpu*)** genes of HIV-1. Whether these variations translate into differences in replication kinetics and cellular tropism requires further investigation [64].

1.1.4. Structure and life cycle of HIV-1

HIV-1 virions are composed of a nucleoprotein core surrounded by a proteolipid envelope (Env) containing surface (Gp120) and transmembrane (Gp41) glycoproteins **(Figure 1.4)**. In addition to *env* gene, the HIV-1 genome codes for two other structural genes, *gag and pol* **(Figure 1.5.).** The core of the virus consists of an electron dense conical (p24) capsid containing two copies of viral genomic RNA held together by RNA-binding proteins and multiple complementary nucleic acid regions [65]. The core also contains enzymes (reverse transcriptase, RNase H and integrase), tRNA and "accessory" or "regulatory" proteins (Tat, Rev, Vif, Vpu, Vpr and Nef) required for viral infectivity, cDNA synthesis and virion assembly during budding from the cell surface [65]. Tat is a critical transcriptional activator of the LTR region [66], while Rev plays a major role in the transport of viral RNAs from the nucleus to the cytoplasm [67]. Vpu promotes CD4 degradation, Vif overcomes the inhibitory effects of host cell factors (namely APOBEC), Vpr promotes the infection of macrophages and Nef promotes the down regulation of surface CD4 and class I MHC expression, blocking apoptosis and enhancing virion infectivity (**Table 1**) [68].

Virus entry into susceptible target cells is mediated through a complex sequence of interactions between the gp120 subunit of env and cellular receptors, primarily CD4 and either CCR5 or CXCR4. Trimers of HIV-1 gp120 recognize and bind to CD4 inducing a conformational change in gp120 which then exposes the co-receptor binding sites. Depending on the tropism of the virus, which is determined by the V3 loop of gp120, the cellular co-receptor CCR5 or CXCR4 is engaged, initiating the process of fusion and pore formation [69;

70]. This triggers a subsequent conformational change in gp41, allowing HR1 and HR2 to interact and form a stable six-helix bundle structure. This completes the membrane fusion process and allows the viral core to enter the cell [71]. Envelopes that interact with CCR5 typically have a low number of positively charged amino acids, whereas CXCR4-using viruses usually contain V3 sequences with an excess of positively charged residues, primarily at amino acid positions 11 and/or 25 [72-74]. R5X4 viruses have a V3 charge pattern that is similar to that of X4 envelopes. Recent evidence suggests that CD4 is present in lipid rafts and that multiple receptor/co-receptor molecules must be recruited into actin-dependent cell surface structures such as ruffles and microvilli in order for virus-cell fusion to occur in an efficient manner [75].

Following release of the viral core into the cytoplasm, the genetic material must be delivered to the nucleus. This process consists of several steps: 1) the uncoating of the viral capsid and release of the viral ribonucleoprotein complex (RNP), 2) reverse transcription and synthesis of a double-stranded viral DNA intermediate, and 3) formation of a functional pre-integration complex (PIC) that is transported into the nucleus via the cell's microtubule transport system [76-78]. It is during this time that cellular restriction factors such as APOBEC3G and Trim5α can act to prevent cross-species transmission [79; 80]. Although the structure of the PIC has not been precisely defined, it appears that the ends of the viral DNA are "tethered" by an assembly of proteins that protect and prime the DNA for integration [81; 82]. Integration into the human genome is directed by cellular components of the PIC with insertion occurring in genes that are highly transcribed by RNA Polymerase II [83-86].

The 5'-LTR of HIV-1 closely resembles the promoter sequences found in the human genome. As a consequence, HIV-1 is able to use the transcriptional machinery of the host cell to promote viral gene expression and replication, a process that is profoundly affected by the activation status of the cell (87). The HIV-1 promoter region contains two modulatory enhancer and one core promoter region, located upstream of the transcription start site [88; 89]. A third regulatory element, the Trans-Activation Response (TAR) region, lies downstream of the start point. Interaction between the core promoter, SP1 and TATA proteins results in a basal level of transcription which is further enhanced, either by Tat/TAR interactions, or through the binding of host transcriptional factors to the modulatory enhancer, a region that contains at least 6 sequences that bind different cells factors such as NF-κB (90). The regulation of viral transcription in cells of the monocyte-macrophage lineage is

substantially more complex than in $CD4^+$ T cells and is dependent on, not only the activation status of the cell, but also its stage of differentiation. Depending on the timing of the stimulation (pre- or post-infection), exposure to pro-inflammatory and immune modulating cytokines such as TNF, can induce or suppress HIV-1 infection and replication, either at the level of virus entry or at the transcriptional and post-transcriptional levels [91; 92].

The maturation and budding of newly synthesized HIV-1 virions occurs at the cell membrane and is mediated by the HIV-1 Gag polyprotein, $PR55^{Gag}$ which directs the recruitment of all components required for the efficient assembly and release of new viral particles [93; 94]. All four domains of $Pr55^{Gag}$, matrix (MA), capsid (CA), nucleocapsid (NC) and p6, play an important role in this process [95-97]. Cleavage of $PR55^{Gag}$ by the viral protease is not required for particle formation, but is necessary for the virion to be infectious. Three domains $PR55^{Gag}$ are of particular importance: the membrane targeting (M), interaction (I) and late (L) domains. Following N-terminal myristoylation, the M-domain in the MA region interacts with acidic phospholipids on the inner leaflet of the plasma membrane targeting Gag to the membrane. The I domain, located in the NC, is responsible for Gag monomer interactions [98; 99]; separation of the viral and cellular membranes and virus budding is mediated by a PTAP motif located within the L-domain of p6 [100]. Mutations in this region lead to a marked inhibition/defect in virus budding from the plasma membrane. Other studies have suggested that HIV-1 assembly, particularly in macrophages, takes place primarily in endosomal compartments and that the release of viral particles follows an exosomal pathway in which virus-containing endosomes fuse with the plasma membrane releasing their contents [101-103]. *In vitro* studies of monocyte-derived macrophages suggest that these cells can synthesize and accumulate HIV-1 virions in late endosomes or multivesicular bodies (MVBs) that form part of a large, complex invaginated plasma membrane [104-106] and that these archived viruses (which remain stable for >6 weeks) can be subsequently transmitted to uninfected CD4+ T cells across a transient adhesive contact termed a Virological Synapse (VS) [107]. Virions transmitted in this manner are protected from effector elements of the host immune system. Of particular interest is the recent finding that TGS101, a member of the Class E vacuolar protein sorting machinery and a subunit of the Endosomal Sorting Complexes Required for Transport (ESCRT) is essential for sorting and the formation of MVBs [108; 109]. This information suggests that therapies aimed at disrupting ESCRT complexes may be particularly useful in inhibiting the cell-to-cell transmission of HIV-1.

1.2. HIV-1 GENETIC EVOLUTION

1.2.1. *Mechanisms of Evolution*

Unique features that facilitate the rapid genetic evolution of HIV-1 include: 1) its small genome (9 kb) with relatively few genes; 2) high mutation rates due to an error prone reverse transcriptase and the absence of proof-reading activity; 3) high production rates, in excess of 10^9 virions per day; and 4) long-term survival in large numbers of affected persons [110; 111]. The error frequency of the HIV-1 RT enzyme is estimated to be in the order of 3.4×10^{-5} subsitutions per site per replication cycle [104]. Since the HIV-1 genome is almost 10^4 base pairs in length and an estimated 10^{10} virions are produced each day, this leads to the generation of millions of variants within each individual in a single day. Another major factor contributing to viral diversity is HIV-1's propensity to undergo recombination [112], an event that may occur as frequently as three times per replication cycle. As a result, each HIV-1-infected individual harbors a uniquely diverse viral population referred to as the viral "quasi-species". "Inter-host" diversity is even more extensive [113; 114]. The time from the release of a virion until it infects a new cell and releases progeny of its own, the generation time of the virus, is approximately two days [111]. This high turnover and extreme plasticity, which is unprecedented in the world microbes, influences the virus' immunogenicity, transmissibility and infectivity, and is a major impediment to viral eradication and the development of an efficacious AIDS vaccine [115].

1.2.2. Viral fitness and within host Evolution

Intrahost genetic evolution is driven by a "survival of the fittest" concept. Each viral variant within a quasispecies has a unique fitness that is determined by interactions between the virus and its environment. Within a given microenvironment, different genes of the virus mutate at different rates depending on immune- and drug-induced selection pressures [116]. Positive (Darwinian) selection leads to the emergence of viral variants that are better suited to a given environment, whereas negative selection eliminates less fit variants [116; 117]. Thus, the viral population within a given individual is in a continuous state of expansion, competition and selection. As a result, HIV-1 can find multiple pathways to escape from immune- and drug-associated selection pressure, allowing it to reach peak viral fitness. For example, changes in N-linked glycoslyation sites in Env allow HIV-1 to escape from

neutralization, while mutations within CTL epitope sites of Env, Gag and Tat facilitate escape from CD8+ T cell responses [118-120]. During this process, a given viral quasispecies must continually balance its ability to escape immune elimination with its ability to survive, replicate and preserve its life-cycle. The marked heterogeneity of HIV-1, especially in the hypervariable (V1-V5) regions of *env*, is attributable in large measure to class I HLA-restricted CTL responses [121; 122]. Mutational escape from CTL selection pressure and neutralizing antibodies (which is often masked by the heavy glycosylation of the Env protein) is of lasting advantage only if the host is subsequently unable to mount an effective response to the escape epitopes. CTL and neutralization escape mutants emerge shortly after the onset of acute infection and can become the dominant variant during early chronic infection. Although some studies have found an association between the rate of viral diversification and disease progression, others have not, and it has been suggested that the level of diversity must exceed a certain threshold for disease progression to occur [123-126].

The impact of selection pressure on a viral gene or gene segment can be determined by measuring the ratio of synonymous to non-synonymous substitutions. A synonymous (or silent) mutation is defined as a substitution where the nucleotide change does not cause a change in the amino acid at that position; a non-synonymous mutation is one that leads to replacement with a different amino acid at that codon. Synonymous substitutions generally occur in the third position of the codon and do not change the amino acid composition. Such mutations are believed to be due to random genetic drift rather than selection pressure and, when fixed in a population**,** occur independently of environmental factors [127]. The selection pressure is assumed to be negative when the frequency (number) of non-synonymous mutations at a given codon position is less than the frequency of synonymous substitutions. For the most part, negative selection pressure prevails, particularly in genes that need to maintain their structure and function.

1.2.3. Viral Reservoirs, Compartmentalization and Sanctuary tissue sites

During early infection, HIV-1 colonizes different cell types, tissues and organs within the body. If viral trafficking between these regions is restricted, HIV-1 can evolve along different evolutionary pathways driven by differences in immune selection pressure and, in patients on ART, by disparities in the local concentration of antiviral drugs. If sufficiently restricted, each of these viral subpopulations can become genetically distinct or

compartmentalized [128; 129]. Some of these compartments can serve as “sanctuary” sites for continued viral replication during ART. Sanctuary sites can be divided into cellular sites including actively replicating and latently infected macrophages, dendritic cells and $CD4^+$ T cells, or into anatomical sites such as blood, lymph node, spleen, central nervous system (CNS) and genital tract [130-137]. Some of these cells and tissues are relatively inaccessible to antiretroviral drugs due to poor drug penetration, limited cell uptake or inefficient metabolism. These cells and tissues serve as sources of rebound virus and the evolution of drug resistance during therapeutic failure. Protease inhibitors, for example, exhibit poor penetration into the central nervous system and genital tracts due to the presence of a blood-brain or blood-testes barrier. Compartmentalized viral populations have distinct phenotypic, pathogenic and genetic properties. At the biological level, these populations can differ in cell tropism, NSI/SI phenotype, and ability to support productive infection. When assessed by sequence and phylogenetic tree analysis, compartmentalized viruses fall into separate distinct clusters of related sequences.

It is important to note that there are two types of sanctuary sites for HIV-1: cellular and anatomical. Cellular sites include memory $CD4^+$ T cells that have reverted to a dormant or near-dormant state; latently-infected macrophages which express low levels of HIV-1 for prolonged periods of times, and follicular dendritic cells which may capture and hold infectious virus on their cell surface for indeterminate lengths of time [138]. Anatomical sites, including the brain and testes, are tissue compartments that are highly impermeable to antiretroviral drugs [139]. Viral replication persists in “latently” infected macrophages and memory CD4+ T cells during ART and, in addition to actively replicating virus in sanctuary tissues, is a major cause of viral rebound and treatment failure [128; 138]. Archived sequences in a viral reservoir tend to be dispersed throughout the phylogenetic tree and to be more closely related to the most recent common ancestor (MRCA) than contemporaneous virus present in a non-reservoir [138]. A clear understanding of these differences is important since different approaches may be required to eliminate HIV-1 from anatomical vs. cellular reservoirs.

1.2.4. Quasispecies restriction during transmission

The selection pressures that accompany HIV-1 transmission are different from those that drive intra-host evolution [25]. The first of these factors relates to the “bottleneck” that

occurs during inter-host transmission, a phenomenon that greatly reduces the diversity of the transmission “swarm”. Although the degree of homogeneity may vary depending on the mode of transmission [140], a number of studies have shown the heterogeneity of the HIV-1 population in a new host is significantly less than that detected in the transmitting partner [141; 142]. This may be related to a number of different factors including the nature of the viral variants that are present in the transmission fluid (vaginal secretions, semen, breast milk), the density of target cells at the site of infection, as well as their co-receptor expression, the size of viral inoculum, the innate immune response and the ability of the virus to replicate in mucosal tissues [143; 144].

Selection pressure at the level of the V3 loop of gp120 appears to play an important role in limiting quasispecies diversity in the new host by favoring the preferential transmission of viral variants that recognize the CCR5 rather the CXCR4 co-receptor. Interestingly, individuals with a homozygous 32-base pair deletion in CCR5 (CCR5Δ32), an allele that is present in Caucasian populations at a frequency of 0.092, are protected from HIV-1 infection, while individuals who are heterozygous for CCR5Δ32 exhibit slower disease progression than those who do not have the deletion [145]. In addition, although not clearly defined, some viral variants may not readily adapt to the HLA haplotype of the new host. Alternatively, some CTL-escape mutations may be transmitted to individuals with the unmatched HLA background. Viruses with naturally occurring and drug-induced resistance mutations in RT and protease may also have lower transmission fitness and be less replication competent in a new host compared to drug-sensitive variants. Irrespective of the reasons, the evolutionary pressures during transmission are clearly different from those that occur within an individual host. Intra-host evolution does not select for viruses with enhanced transmissibility.

1.2.5. Virus evolution at the population level

Transmission bottlenecks also greatly reduce the influence of positive selection on HIV-1 evolution at the population levels. In contrast, genetic recombination, particularly between different HIV-1 subtypes, is an important contributor to inter-host variation. This is due to the fact that most infected cells harbor two or more different proviruses and that dual infections are common, especially in global regions with a high

prevalence of multiple subtypes. In some geographic regions, CRFs account for more than 25% of all HIV-1 infections and can have a profound effect on the analyses of phylogenetic relationships and the timing of evolutionary events, as well as demographic analyses and studies of natural selection. A set of robust tools that can accurately account for recombination events is needed to fully understand HIV-1 genetic evolution across different population groups. This will be extremely difficult**,** but may be possible through the use of transmission networks to infer phylogenetic relationships, an approach that allows individual sequences to have many different ancestors. Carefully designed studies are needed to better quantify the temporal, adaptive and lineage extinction relationships between recombinant viruses. The large number and growing frequency of CRFs makes this a critical area of future research, not only for diagnostics, but also for understanding the potential magnitude and future directions of the global AIDS pandemic, as well as for the design of appropriate treatment and vaccine strategies.

1.3. THE GASTROINTESTINAL TRACT (GIT)

1.3.1. Overview of the GIT

The mucosal surface of the GIT occupies a unique anatomical and physiological position, serving as a major structural and immunological barrier against microbial and viral pathogens from the outside world. In addition, the tightly apposed enterocytes that form the single cell layer of the mucosal epithelium absorb water and nutrients during the digestive process, a function that is critical to survival. This process occurs in the continued presence of an enormous density of potentially harmful microbial pathogens. It is estimated that up to 10^{14} bacteria exist in the lumen of the gut, a number that exceeds the total number of cells in the human body [146].

To combat the entry of harmful bacterial, parasitic and viral pathogens, the intestine has developed strong adaptive and effector immune functions, as well as a unique protective innate immune system. Specialized adaptations, at the level of the epithelium include the formation of tight junctions that seal the intercellular spaces between enterocytes, the coating of the apical surface with a layer of mucin and the secretion of antimicrobial factors such as defensins and lysozyme that help prevent direct access of microbes to the intestinal mucosa [147-149]. In addition, to intraepithelial lymphocytes (IEL), an effector cell population that

has immune regulatory functions, are the specialized M cells which transport antigens from the lumen to the immune inductive sites in the lamina propria of the GIT [150; 151). Other protective cells include lamina propria plasma cells that secrete non-complement activating IgA and $CD4^+CD25^+$ regulatory T cells (Tregs) that secrete large amounts of anti-inflammatory cytokines including IL-10 and TGF-β [152; 153].

The main lymphoid elements are Peyer's Patches in the small intestine and lymphoid aggregates in the large intestine as well as the mesenteric lymph nodes draining these tissues [150; 154]. These lymphoid structures, classified as inductive or education sites, are places where antigens are taken up and presented to B and T lymphocytes. After encountering antigen in the inductive compartment, lymphocytes migrate into the systemic circulation where they differentiate into mature effector memory cells. Effector cells then migrate back to the GIT where they neutralize both microbial and non-microbial foreign antigens [155]. This homing process involves interactions between α4β7-integrin on the surface of T cells and mucosal vascular addressin cell-adhesion molecule 1 (MAdCAM-1) on endothelial cells of the GIT; between αEβ7 on T cells and E-cadherin on intestinal epithelial cells and, in the small intestine, between CCR9 on T cells and CCL25 on epithelial cells [156-158].

The GIT also contains the largest reservoir of macrophages in the body. Intestinal macrophages, which represent 10-20% of all mononuclear cells found in the *lamina propria* [159; 160]. Macrophages are the first phagocytic cells of the innate immune system that interact with whole organisms, microbial products**,** and viruses that have breached the intestinal epithelium. They play a critical role in eliminating harmful pathogens and in the initiation and resolution of protective immune responses, as well as in the regulation of mucosal responses to commensal bacteria needed for the proper processing and digestion of food. To accomplish these opposing activities, intestinal macrophages have adapted a profound "anergic inflammatory" phenotypic while still retaining their avid phagocytic properties. In healthy $HIV\text{-}1^-$ individuals, "resident" intestinal macrophages do not express receptors for LPS (CD14), Fcα (CD89) or Fcγ (CD16, CD32, CD64), complement (CD11b/CD18, CD11c/CD18), growth factors (IL-2, IL3) or integrin **(**LFA-1**)** and, as a result, they do not produce pro-inflammatory cytokines in response to a wide range of immunological and microbial stimuli [154; 160; 161]. This allows for the effective clearance of harmful pathogens while preventing excessive immune activation in response to commensal bacteria. However, in inflammatory bowel diseases (IBD) such as Crohn's

disease (CD) and ulcerative colitis (UC), and in patients with advanced HIV-1 infection, a significant proportion of these macrophages become highly activated, acquiring both innate response (CD14, CD16) and co-stimulatory (CD80, CD86) receptors involved in the activation of T cells, as well as the ability to synthesis pro-inflammatory cytokines and chemokines. These phenotypic changes are likely to contribute to the persistent immune activation that leads to HIV-1 disease progression and therapeutic failure [162]. Another critical component of the innate immune system is the dendritic cell (DC). Immature DC sample antigens and bacteria and carry these particulate and antigenic substances into the draining lymph nodes and ultimately to distal lymph nodes for antigenic presentation to T cells, initiating an adaptive immune response [163; 164].

1.3.2. Role of the GIT in HIV-1 pathogenesis

Until recently, our understanding of HIV-1 pathogenesis was based almost entirely on immunological and viral load studies performed on peripheral blood. Blood constitutes only 2% of the body's total viral burden and contains only a small percentage of the body's total T cell population [165; 166]. In this traditional view of HIV-1, disease pathogenesis was divided into three phases: 1) an acute phase involving a rapid but transient decrease in $CD4^+$ T cells in association with the initiation of an antiviral immune response and a decrease in HIV-1 burden to a viral set point; 2) an asymptomatic or chronic phase of infection characterized by a gradual relatively indolent decline in peripheral blood $CD4^+$ T cells and 3) an symptomatic phase characterized by total collapse of the immune system and full-blown AIDS [167].

However, recent studies of Simian Immunodeficiency Virus (SIV)-infected rhesus macaques and other non-human primates, and to a lesser extent in HIV-1-infected humans, have shed new light on both the kinetics and mechanisms underlying disease progression. These studies have shown that the early targets of SIV and HIV-1 are mucosal memory $CD4^+$ T cells that carry the CCR5 HIV-1 co-receptor. The majority of these cells (in fact >80% of the body's entire T cell population) are found in the gastrointestinal tract. Within 24 hours of vaginal exposure, productive HIV-1/SIV infection is detectable in GIT, followed at day 5-10 by an intense burst of viral replication in the *lamina propria* and a massive (70-90%) depletion of the "recently" or "partially-activated" $CCR5^+$ memory $CD4^+$ T cells [168-172]. During the chronic phase of HIV-1/AIDS, the gut continues to serve as a reservoir of

treatment refractory virus and, because of its constant exposure to bacterial, viral and parasitic pathogens, plays an important role in the recruitment, activation and infection of new $CCR5^{+}CD4^{+}$ target cells [173-176].

At the immunological level, studies have shown that this early massive loss of $CD4^{+}$ T cells in the GIT is not, on its own, sufficient to cause progression to AIDS [167; 177; 178] suggesting that additional factors must be involved. It now appears, based on comparative studies of non-pathogenic and pathogenic SIV (and HIV-1) infection, that the key determinant of disease progression is the persistence of systemic immune activation during the chronic phase of infection. Recent studies suggest that this process may be driven, in large measure, by virus-induced damage to the integrity of the epithelial barrier of the GIT and increased translocation of microbial products from the lumen of the intestine into the systemic circulation [177-181). Antiretroviral therapy (ART) is only partially effective at suppressing microbe-driven immune activation [162; 182] and the extent of immune activation during ART is strongly predictive of therapeutic failure and progression to AIDS [183]. These findings, together with the inability of ART to completely suppress viral replication in the GIT, underscore the critical nature of understanding the mechanisms driving immune dysfunction in the intestine. At present, little is known about these mechanisms. This lack of knowledge is likely to be one of the reasons why we still do not have an effective AIDS vaccine or antiviral treatments that are capable of eradicating HIV-1 infection. At the virological level, peripheral blood studies have shown that during the time from initial infection to peak viremia (days 10-14), rampant viral replication has already generated a series of mutant viruses [184; 185]. Some of these mutations result in viruses that are resistant to neutralizing antibodies [119; 186; others give rise to viral variants that are no longer recognized by virus-specific CTLs [187; 188]. These early escape viruses, which are already captured in intracellular reservoirs, are the main obstacle to the prevention or the elimination of HIV-1 by ART and the host immune defense system [189; 190].

1.3.3. GIT as a site of virus entry and dissemination

During vertical transmission, the upper GIT serves as the portal of entry after the fetus has swallowed infected amniotic fluid *in utero*, or after the infant has ingested infected maternal blood and cervical secretions during labor and delivery or through the ingestion of contaminated milk during breastfeeding [191]. During homosexual transmission, HIV-1 is

acquired through an orogenital or anogenital route. The efficiency of transmission is related, not only to the infection status of the donor, but also to the susceptibility of the recipient [192; 193]. Knowledge of donor infectivity is based primarily on studies of heterosexual transmission in individuals infected with HIV-1 subtype B. Factors that affect transmission include: the nature (cell-free vs. cell-associated) and size of the viral inoculum, as well as the presence of pro-inflammatory cytokines and opportunistic pathogens that increase viral shedding [194; 195]. Factors that influence recipient susceptibility include the integrity of the GIT or vaginal surface, the genetic predisposition of the recipient, the presence of sexually-transmitted infections (STIs), and the presence of antiviral factors such as secretory leukocyte protease inhibitor (SLPI), HIV-1-specific CTLs and anti-HIV-1 IgA and IgM antibodies [196-198].

The exact mechanism(s) by which HIV-1 enters the mucosa has not been clearly determined. Trauma and breaches in the structure barrier of the epithelium may facilitate virus entry. Alternatively, the virus may enter through “leaky” tight junctions, or through a process of endocytotic uptake and microtubule-mediated transport through M cells [199]. Specialized M cells are located directly above Peyer’s patches in the small intestine and above lymphoid follicles of the colon and rectum. Following intracellular transport, virus (SIV/HIV-1) is delivered to lymphocytes and macrophages located within pockets at the base of the M-cells. In the case of heterosexual transmission, studies of rhesus macaques have shown that, within 24 hours of vaginal inoculation, productive HIV-1 infection is detectable in the colon, either at organized inductive sites in follicular aggregates or at diffusely distributed effector sites throughout the *lamina propria* [168]. This rapid dissemination of virus from the vagina to the GIT presumably occurs via the lymphatic system and/or peripheral blood. As discussed below, the initial target cells in the GIT are CD4+ memory T cells that lack traditional activation (CD69, CD25) and proliferation (Ki67) markers, but that express sufficient CCR5 to support productive infection. Intestinal macrophages and DC are also early targets of HIV-1 in the GIT.

1.3.4. Co-receptor use, Tropism, Syncytium-inducing properties

As described above, virus entry into target cells is mediated through a series of complex interactions between the HIV-1 envelope glycoprotein, gp120, and cell receptors, primarily CD4 and either CCR5 or CXCR4. This, in turn, leads to activation of gp41 and

fusion of the viral envelope with the cell membrane [200]. Co-receptor usage and the syncytium-inducing properties of HIV-1 evolve during the course of disease. Viruses isolated during early infection preferentially utilize CCR5 for viral entry, while those arising later, particularly in subtype B infection, often use CXCR4 for membrane fusion [201; 202]. Based on this information, HIV-1 has traditionally been classified as either macrophage tropic (M-tropic) or non-syncytium inducing (NSI) or T-cell tropic (T-tropic) and syncytium-inducing (SI). M-tropic/NSI variants predominantly use the chemokine receptor 5 (CCR5) for entry into macrophages and were designated R5 virus, while T-tropic/SI variants predominantly use CXC chemokine receptor 4 (CXCR4) for entry in T-lymphocytes and were designated X4 viruses with dual tropic (R5X4) viruses being capable of infecting both macrophages and $CD4^+$ T cells [203; 204]. However, this classification is not completely accurate, as macrophages can express CXCR4, and some primary X4 isolates can use this receptor, even though prototypic X4 strains are unable to infect macrophages [205]. In addition, as described below, partially activated $CD4^+$ T cells are found in the GIT and are primary targets of early HIV-1 infection, in addition to $CCR5^+$ macrophages. In reality, the determinants of cellular tropism and productive HIV-1 infection are likely to be substantially more complicated than simple co-receptor usage, especially in the case of macrophages (91). In addition to genetic variations in the HIV-1 envelope, co-receptor restriction may involve determinants of viral replication, host genetics and the activation status of the cell [206-208].

1.3.5. Biological and Genetic evolution of HIV-1 in the GIT

There are only a few studies describing the genetic and biological diversity of HIV-1 in the intestine. Two of these studies were conducted in Canada in homosexual/bisexual men infected with subtype B virus. The first study, performed on ART-naïve patients recruited in 1993 (prior to the introduction of ART), revealed that the GIT can harbor both SI and NSI phenotypes of HIV-1 and that, in patients with advanced disease, discordant SI/NSI phenotypes are detected in the gut vs. blood [209]. Since SI variants are only occasionally detected in subtype C-infected patients, the relevance of this information for southern Africa is uncertain. The second study was performed on biopsy specimens collected prior to 1997. The study included both treatment naïve patients and patients treated with AZT, DDI, DDC, D4T or 3TC, or combinations thereof. Sequencing of RT and Nef genes from the esophagus, stomach, duodenum and colorectum revealed that different regions of the gut contain distinct

HIV-1 quasispecies and that viral replication in the gut is compartmentalized (210). Another study by van Marle and his colleaques in 2010 also showed that there exist different antiretroviral resistant variants in the different compartments gut and that they are very variable [211]. This study more importantly indicated that drug resistance in blood can be very different to what is observed in the gut. For the first time a study for well-characterized patients either at the time of the primary infection, or during chronic infection and or with AIDS compared the genetic diversity of variants in blood with the genetic diversity of HIV DNA in the rectum [212]. This study showed for the first time that, from the time of primary infection viral genetic diversity exists in the rectum, even though there was no complete independent HIV-1 evolution observed as was seen previously in other studies with the variants of HIV in faeces [213] and more recent studies in the gut-associated lymphoid tissue [211]. Another recent study by Imamichi and his colleaques they observed from matching time points that HIV-1 sequences from pheripheral blood (plasma and the PBMCs), and gut compartments were genetically indistinduishable, indicating an equilibrium of HIV-1 quasispecies between the pheripheral blood and the gut. In this study they also found that HIV-1 variants detected in the gut compartment were present in both cell-associated RNA forms and proviral DNA. Finally, they found that these transcriptionally active forms of HIV-1 in the gut were in the plasma and PBMCs from matching time points [214]. Another study using the robust technique of single genome sequences (SGS) of full length HIV-1 env with known immunologic evidence of residual GALT immune activation despite clinical evidence of plasma HIV-1 RNA suppression were unable to identify evolved forms in the GALT or PBMC of individuals [215].

In this study they concluded that at a minimum currently available regimens of suppressive cART have the ability to abrogate de-novo rounds of HIV-1 replication in the GI lymphoid tissue in individuals initiating such therapy during primary infection, supported by the absence of evidence of HIV-1 variants evolution during cART. The forces contributing to genetic variation and regional compartmentalization in the GIT are not well understood but may include factors stemming from a “founder” effect in association with seeding of cellular reservoirs, followed by localized viral evolution. Drug and tissue-specific immune selection pressures may also shape the viral populations found in different regions of the GIT, as well enteric co-infections. In the setting of HIV-1, increased translocation of commensal bacteria may also exert immune selective pressures. Delineation of the factors driving genetic evolution in the GIT is of critical importance, not only for preventing/delaying disease

progression, but also for development of novel strategies aimed at preventing therapeutic failure and the emergence of drug resistance. Such studies are of particular importance in southern Africa, a region with an explosive epidemic of subtype C infections and where enteric co-infections are common and large numbers of HIV-1-infected patients with late-stage disease and $CD4^+$ T cell counts below <200 cells/μL have been prioritized for NRTI/NNRTI-based ART, in the absence of protease inhibitors (PIs).

1.3.6. Role of the GIT in treatment failure

To date, relatively few studies have assessed the effects of ART on the GIT [170; 216-221]. A study by Talal and his colleaques in 2001, which evaluated the virologic and immunological responses of antiretroviral theraphy in the gut-associated lymphoid tissue (GALT) comparing with those found in pheripheral blood concluded that HIV-1 replication was effectively suppressed by Antiretroviral theraphy in the GALT [222]. This study showed that there was a uniform increase of $CD4^+$ T cells percentages in the peripheral blood among all patients, while it was variable in the GALT. In another study which evaluated the kinetics of viral suppression, CD4_ T-cell restoration, gene expression, and HIV-specific CD8_ T-cell responses in longitudinal gastrointestinal biopsy and peripheral blood samples from patients initiating HAART during primary HIV infection (PHI) or chronic HIV infection (CHI) they discovered that suppression of the virus was more effective in GALT of PHI patients than CHI patients. Mucosal $CD4^+$ T-cell restoration was delayed compared to peripheral blood and independent of the time of HAART initiation.Their finding suggested that the level of restoration of the mucosal immune system during HAART was determined by viral replication suppression and control of inflammatory responses in the GALT [219]. In another study comparing the effect of long term therapy in a cohort of HIV-1 infected men in the sigmoid colon vs blood, they observed that $CD4^+$ T cell reconstitution in the sigmoid colon exceeded that in blood [213]. In this study they further observed similar degrees of reconstitution in the proportion of colonic $CD4^+$ T-cells expressing the HIV coreceptor and activation marker CCR5. Another study by Ciccone and her colleaques showed that cycling of $CD4^+$ and $CD8^+$ T cell populations decreased or normalized after prolonged ART and that the number of $CD4^+$ T cells was restored in the gut mucosa of HIV-infected patients. This study also showed that the gut $CD4^+$ T cell immune restoration was also reflected in β7 expression on peripheral blood $CD4^+$ T cells. They observed an association between cycling of gut $CD4^+$ T cell with plasma LPS levels but not with plasma or tissue HIV-RNA levels,

which indicated that gut T cell activation in the gut may be interconnected with microbial translocation. The finding from this study agreed with other studies, as they also observed that the percentage of $CD4^+$ T cells in both colon and ileum remained significantly lower than that of HIVuninfected controls despite effective prolonged ART [224]. In one study of acute vs. chronic infection, initiation of HAART during acute infection resulted in a near complete restoration of $CD4^+$ T cells, whereas when treatment was introduced during the chronic phase of infection, $CD4^+$ T cell restoration was delayed and incomplete [220]. Other studies have shown that the failure of HAART to fully restore $CD4^+$ T cells in the gut of chronically infected patients is due to residual viral replication in intestinal reservoirs in association with continued depletion of the CD4+ T cells [219]. Virological and clinical changes in response to HAART include reductions in gut VL and in the number of apoptotic cells, as well as a marked, but incomplete decrease in gastrointestinal symptoms [199]. In patients failing ART, drug resistance profiles in the intestine differ from those in the blood, suggesting there is compartmentalization of virus between these two anatomical compartments [226]. All of these studies were conducted in patients infected with subtype B and all involved the use of complex drug regimens containing PIs in addition to NRTIs and NNRTIs. It has been suggested that therapeutic monitoring of the GIT may be a more accurate measure of treatment outcome than either plasma VLs or peripheral blood $CD4^+$ T cell counts.

1.3.7. Immune activation and Inflammation in the GIT

Immune activation and inflammation, during HIV-1 infection involves several mechanisms related to viral replication either directly or indirectly [227]. The gastrointestinal mucosa is located at the interface between a sterile internal and a microbially contaminated external environment. It serves the conflicting needs of nutrient absorption and host defense, functions that require intimate contact with the external environment [199; 228]. As a result, the GIT has evolved a thin polarized epithelium and a large surface area characterized by mucosal folds, villi and microvilli. In the small intestine of healthy persons, antigen exposure comes primarily from the diet, whereas in the ileum and colon the antigenic load is further enhanced by an abundant and complex array of commensal microbes [229]. The organization of the epithelial barrier with its tight junctions helps reduce the risk of luminal antigens entering the mucosa (including commensal and pathogenic microbes), but it does not completely prevent this process [229-231]. Some food proteins and non-pathogenic

commensal bacteria enter the mucosa through breaks in the tight junctions, possibly at the villus tip where epithelial cells are shed, a process that is relatively limited in healthy individuals [232]. The task of mounting an effective immune response to invading pathogens while remaining relatively unresponsive to commensal microbes and food antigens poses a daunting challenge to the GIT. As a result, the gut mucosa has evolved an elaborate system to protect its host from infectious agents. This system consists of two anatomically separated but functionally linked components of the common mucosal immune system [228]. An afferent component, which contains elements involved in the initiation of the immune response including antigen presentation and lymphocyte proliferation, and an efferent component containing elements directly involved in antibody production and CTL responses. Structurally, the afferent system consists of distinct lymphoid follicles overlaid by an epithelial membrane containing M cells. These cells transcytose particulate antigens to antigen-presenting macrophages located at the basal surface of the epithelium [233; 234]. Macrophages are the first phagocytic cells to interact with microorganisms that have entered the intestinal mucosa. Intestinal macrophages have avid phagocytic and bactericidal activities that protect the host from pathogenic organisms and they regulate the immune response to commensal bacteria. Antigen-presenting DC survey the microenvironment by extending processes between gut epithelial cells. They sample both commensal and pathogenic microbes for subsequent transport and presentation to B and T lymphocytes in the spleen and lymph nodes [235; 236]. Cells of the efferent compartment are diffusely scattered throughout the epithelium and lamina propria of the intestine. In addition to CTL, IFN-γ-producing lymphoctyes and IgG/IgAsecreting plasma cells, antibody-dependent cytotoxicity may also play an important role in the adaptive immune response [228; 237-239]. Together, this defense system consisting of the innate immune system, the epithelial barrier and its associated mucous layer and the adaptive immune system, can effectively prevent or restrict the entry and propagation of commensal and pathogenic organisms, including HIV-1 [240].

Recently, in North America, there has been a dramatic increase in inflammatory bowel disease in the absence of overt microbial infection [229]. This finding suggests that some, as yet unknown factor, is perturbing the balance between the normal microflora of the gut and host immunity. It will be important to determine whether this change is due to the use of antibiotics and a treatment-associated reduction in commensal flora. It has been suggested that, under normal homeostatic conditions, the anti-inflammatory responses induced by commensal flora protect the intestinal epithelium from pathogenic insults [241; 242]. This

relationship, however, appears to extremely delicate and anything that perturbs either immune or epithelial homeostasis can lead to inflammation and life-long inflammatory conditions such as Crohn's disease and ulcerative colitis. Patients afflicted by this disease suffer from chronic diarrhea, weight loss and fatigue, in addition to other potential complications such as skin ulcers, arthritis and bile-duct inflammation [243].

HIV-1-associated enteropathy is a poorly-defined clinical condition in which chronic diarrhea, malabsorption and wasting occurs in the absence of a detectable enteric pathogen [228; 244]. Whether this condition is due to subclinical enteric infections, chronic inflammation or direct effects of HIV-1 on the epithelium of the GIT remain to be established. Although there is little evidence that HIV-1 infects enterocytes, various histological studies have reported that the intestinal mucosa of HIV-1-infected persons is characterized by chronic inflammation, villous atrophy, crypt hyperplasia, nuclear enlargement and apoptosis. It has also been reported that jejunal enterocytes have an over-developed smooth endoplasmic reticulum and decreased levels of mitochondria [245-247]. *In vitro* studies, performed on duodenal biopsies have shown that HIV-1-infected patients with diarrhea have decreased transepithelial resistance when compared to patients without diarrhea [248]. At the molecular level, the binding of gp120 to receptors on the basal surface of epithelial cells has been shown to cause decreased acetylation of tubulin, microtubular depolymerization and cytoskeletal rearrangements, changes that result in increased intestinal permeability, increased levels of intracellular calcium and diarrhea [249]. Other studies have reported that HIV-1 infection is associated with degenerative changes in the enteric nerves and in the vasculature of the lamina propria [250; 251] and that the protease inhibitors, saquinavir, ritonavir and nelfinavir (but not indinavir) cause damage the epithelial barrier [252]. Studies showing massive depletion of CD4+ T cells in the GALT during the first few weeks of infection have suggested that this loss in helper T cell function together with the damage caused to the epithelium may alter the antimicrobial properties of the gut allowing for the increased translocation of luminal microbes and microbial products [175; 179]. This phenomenon, known as microbial translocation, can occur in the absence of overt viremia and is a prominent feature of inflammatory bowel disease (IBD) [253]. Damage to the intestine during invasive intestinal surgery and during immunosuppressive conditioning for bone marrow transplantation can also increase the level of microbial translocation [253; 254]. This leads to both local and systemic immune activation and, in the case of transplant patients, an intensification of graft versus-host disease (GVHD) [248]. Quantification of

plasma LPS, a major component of the cell wall of Gram-negative bacteria [255] is routinely used to measure the extent of epithelial leakage and bacterial translocation [256; 257]. Based on preliminary data showing that microbial translocation may be a cause of systemic immune activation in chronic HIV-1 infection, a new model of HIV-1 pathogenesis is beginning to emerge. In this model, the level of immune competency that exists after the acute phase of infection, especially as it relates to the control microbial pathogens in the GIT, determines the rate of disease progression to AIDS [199]. It has been known for some time that markers of immune activation are better predictors of HIV-1 disease progression than CD4+ T cell counts and plasma viral load [258]. However, the factors driving immune activation, including polyclonal B cell activation, increased activation and turnover of T cells and increased levels of circulating pro-inflammatory cytokines and chemokines remain unclear [259-262].

1.4. ANTIRETROVIRAL THERAPY (ART)

1.4.1. Historical perspective and past success

Since ART was introduced 30 years ago, substantial progress has been made in both the treatment and prevention of HIV-1/AIDS, especially in resource-rich regions of the developed world where drugs are readily available. In the USA alone, it is estimated that more than 3 million years of life have been saved since 1989, as a direct result of progress in AIDS care and treatment, including prophylactic treatment for the prevention of opportunistic infections (OIs) and treatment with ART [257]. In addition, prevention of mother-to-child transmission (PMTCT) efforts averted an estimated 2,900 childhood infections, the equivalent to an additional 137,000 years of survival benefit [263-265] Cost-effective analyses have suggested that, overall, prevention efforts in the USA have averted between 204,000 and 1,585,000 infections at a cost that is less than that required to treat an individual patient for HIV-1-related disease [265; 266]. Similar levels of success have been observed in other resource-rich regions of the world, including Western Europe and Australia.

1.4.2. Challenge of art in Africa and the developing world

The Doha declaration in 2001, which allowed developing countries to circumvent patent rights, combined with funding from PEPFAR and the Global Fund for AIDS, has

greatly increased accessibility to ART in developing country settings. The WHO goal of providing treatment to 3 million people in low-income countries was achieved by the end of 2007, with the greatest increase in people on treatment being in sub-Saharan Africa [23]. Although early results have been promising with 78% (10,351/13,288), 76% (7413/9794) and 67% (3840/5690) of patients showing virological suppression after 6, 12 and 24 months, respectively, this represents only 31% of the people in need of treatment and there are still major challenges associated further scale-up of these global treatment programs (23). These challenges relate to the limited repertoire and high cost of drugs in the developing world, the lack of diagnostic and laboratory monitoring and the shortage of health care workers. As a result, treatment is often delayed until late in the course of infection (i.e. when the patient's CD4 T-cell count drops below 200 cells/μL). Such patients have a high incidence of concurrent co-infections and malignancies, such as tuberculosis and Kaposi's sarcoma, which cause chronic activation and dysregulation of the immune system [267; 268] making it difficult to control viral replication and dissemination. Such patients are also more likely to suffer from Immune Reconstitution Inflammatory Syndrome (IRIS). Of particular concern is the limited availability of viral load and resistance testing. As a result, the detection of treatment failure is often based on clinical criteria and CD4 T cell counts, determinants that have only a limited associated with virological failure. As a result, decisions relating to treatment switches have been difficult. Switching patients who do not experience virological failure is expensive and may limit future treatment options. This is of considerable concern since only one second-line drug regimen is available in most African countries. Conversely, keeping patients on a treatment regimen they are failing can lead to the emergence of viral variants that are drug resistant.

Most of the treatment regimens used in African have combined two NRTIs, zidovudine (ZDV) or stavudine (d4T) and lamivudine (3TC), with one NNRTI, either nevirapine (NVP) or efavirenz (EFV) (Anonymous – ref 269). As of March 2010 (23), only 36 publications and 6 abstracts had reported results on virological failure. Most of these studies were cross-sectional, with the duration of follow-up limited to 3-48 months. The median percentage of patients experiencing virological failure (as defined by a rebound viral load of <1,000) was 14%. This relatively low rate of resistance may reflect that fact that most patients (91%) were started directly on an NRTI/NNRTI-based drug regimen and had no prior exposure to monotherapy, or other sub-optimal treatments. The most prevalent drug resistance mutation, detected in 65% of 478 patients with genotypic resistance, was M184V,

a mutation associated with resistance to lamivudine and emtricitabine. Thymidine analogue mutations (TAMs) (M41L, D67N, K70R, L210W, T215Y/F, K219Q/E) were detected in 5-20% of patients, and K65R was present in 36 (5%) of patients with resistance. NNRTI mutations K103N, Y181C, G190A and V106/M were identified in 52%, 17%, 17% and 13% of patients [23]. The relatively low levels of TAMs and K65R mutations may be due, in part, the fact that the M184V mutation causes increased susceptibility to tenofovir (TDF), zidovudine and stavudine, as well as the limited use of tenofovir in the African setting [23]. At present, caution is needed in the interpretation of this data because of the relatively short follow-up (usually < 2 years), the heterogeneity of different study designs and the small number of patients who were subjected to resistance testing, often at a single time point. Serious efforts are needed to expand the drug repertoire and to procure newer antiretroviral drug classes, and to introduce regular virological monitoring that is affordable, accurate and easy to perform.

1.5. MOLECULAR PHYLOGENETICS

In the days of Charles Darwin, evolutionary relationships were inferred from the phenotypic differences and similarities between different taxa and species [270]. Molecular studies of evolution emerged early in the 20th century and gained impetus in the late 1950's and early 1960's. These studies have revolutionized our understanding, not only of the historical relationships among human populations, but also among animal, plant and microbial species, including the human immunodeficiency viruses, HIV-1 and HIV-2. As a result of continued progress in nucleic acid and amino acid sequencing, and in the development of bioinformatics and computer software programs, molecular phylogenetics has emerged as a powerful and sophisticated tool for rapidly estimating the evolutionary histories of large, complex DNA/RNA and protein datasets [271].

The parsimony method (PM) of inferring evolutionary relationships and building trees, developed by Edwards and Cavalli-Sforza in the early 1960s, was the first major breakthrough in phylogenetic analysis [272]. The PM method analyzes the genetic differences in a given dataset and searches for the most parsimonius tree with the least number of evolutionary changes. Since the work of Edwards *et al.*, several different algorithms have been developed to infer evolutionary trees. These include: the Pair Group

Method with Arithmetic means (UPGMA) [273; 274]; the Maximum Likelihood (ML) [275], the Fitch-Margoliash [276], the Neighbor-Joining (NJ) [277], the Minimum Evolution (ME) [278; 279], and the Bayesian [280 -283] methods.

1.5.1. Development of nucleotide substitution models

In 1969, Jukes and Cantor (JC69) [284], developed the first model of nucleotide substitution. The model assumed that the equilibrium frequency for each of the four nucleotides ($\pi_A = \pi_C = \pi_G = \pi_T = ¼$) was 25% and that, during evolution, each nucleotides had the same probabilitiy of being replaced by any other nucleotides. This model, although easy to use, is overly simplistic –it assumes that the only variable is the overall substitution rate. To overcome the sort comings of the JC69 method, improvements and extentions that allow for different rates of transversions and transitions, and for unequal nucleotide frequencies, were incorporated into subsequent models. Figure 1.7.7, is a hierarchical representation of the most important models of nucleotide substitution.

1.5.2 Inferring and constructing phylogenetics trees

There are many methods for inferring and building phylogenetic trees and no single method performs well for all types of trees and datasets. These methods can be divided into two main categories based on the type of data they use to infer tree topologies. These categories are referred to as the character data and distance-matrix methods [285; 286].

1) Character (discrete)-based methods

These methods use discrete characters such as sequence data. The scores for each tree are calculated by comparing each character in the entire sequence alignment, separately, one at a time. Commonly used character-based methods include maximum parsimony, maximum Likelihood, and Bayesian inference methods [287]. The tree scores are defined as the posterior probability for Bayesian inference, the minimum number of changes for maximum

parsimony, and the log-likelihood value for maximum Likelihood [288]. Character-based methods retain the original character status of the taxa and therefore, can be used to reconstruct the character state of the ancestral nodes. The following sections describe the different character-based methods used to infer phylogenetic trees.

i) Maximum parsimony method

Maximum parsimony, one of the first methods of tree inference, was introduced in the early 1960's [272]. As mentioned above, this method searches for the most parsimonious tree that requires the least number of evolutionary changes to explain the observed differences in a given dataset [271]. It minimizes the number of changes on a phylogenetic tree by assigning character states to interior nodes on the tree. The sum of the character (or site) lengths, defined by the minimum number of changes required for each site [289], is calculated and used to determine the score of the tree. Fitch, 1971 developed an algorithm for finding the minimum number of changes on a binary tree [288] and an algorithm for reconstructing the ancestral states to achieve the minimum change was developed by Hartigan, 1973 [290]. The most commonly used maximum parsimony methods are MEGA [291], PAUP* [292], and TNT [293].

The maximum parsimony method is easy to use and understand. The simplicity of the method helps when developing efficient computer programs. Other advantages include its ability to withstand rigorous mathematical analysis. Disadvantages relate to its lack of explicit assumptions. This makes it nearly impossible to incorporate knowledge of the evolutionary process into tree construction. Since the method does not correct for multiple substitutions at the same site, it can also give rise to incorrect trees with long branches grouped together by parsimony or model-based methods, a phenomenon known as long-branch attraction [289; 294].

ii) Maximum likelihood method

Maximum likelihood (ML), developed in the early 1920's by R. A. Fisher, is a statistical method for estimating unknown parameters in a model [295]. In 1981, Felsenstein developed the first algorithm for maximum likelihood analysis of DNA sequence data [296]. This advance, together with the development of more powerful computers and new software implementations, paved the way for increasingly realistic models of sequence evolution [278]

such as MOLPHY [297], PAUP* 4.0 [292], PHYLIP [298]. The continued development of faster computers has led to even more effective methods of constructing ML trees including PHyML [299], RAxML [300], GARLI [301] and ML in MEGA [291]. The latter method allows for easy access and utilitzation. In order to estimate nucleotide substitution, the ML method requires a probabilistic model of evolution. It evaluates competing hypotheses (trees and parameters) by selecting those with the highest likelihood and renders the observed data most plausible. The ML algorithm searches for the tree that maximizes the probability of observing the character states obtained for the given data set, tree topology and model of evolution inferred. Numerical optimization techniques are used to determine the combination of branch lengths and evolutionary parameters that correspond to the maximum likelihood.

Depending on search algorithm used, the likelihoods of a large number of trees topologies are searched and the tree that yields the highest likelihood is choosen as the best possible tree. The method involves optimization of the branch lengths in order to calculate the tree score for each candidate tree and a search in the space of the maximum likelihood tree. The ML method is similar to the maximum parsimony method in that it also assumes independent evolution of sites in the sequence. As a result, the likelihood is a product of the probabilities for different sites, where the probability at any particular site is an average of the character states at the ancestral nodes. Unlike maximum parsimony, which uses the optimal ancestral states [289; 285; 286], ML averages the likelihood over all possible states. The advantage of the ML method is that all of its model assumptions are explicit and thus, are easy to evalute and improve. Unlike parsimony, the method has a rich repertoire of sophisticated evolutionary models and can be used to study the process of sequence evolution. The main disadvantage of ML is that tree searches and the calculation of likehood are computationally demanding. Weak statistical results can also be obtained if the model is misspecified [289].

iii) Bayesian inference method

Bayesian inference is a character-based method that creates a stochastic model of evolution. Rather than searching for the single best tree, the method provides the most probable distribution for a set of plausible trees, or for different aspects of an evolutionary history. The method introduced in mid-to-late 1990's, by Yang and Rannala [281] and by Mau and Newton [282], and further improved by Li and colleaques [283], uses an optimality criterion to determine the distribution of all trees in the dataset and to search for a set of

hypotheses and plausible trees. The posterior distribution of the trees gives a confidence estimate of evolutionary relationships.

To use the Bayesian inference method, the investigator or researcher must specify a parameter (i.e. tree topology, substitution model or branch length) that can be used to determine the stochastic distribution of the trees. In the absence of appropriate biological information about the gene or species being studied, this prior belief will be vague or uninformative. However, if the predominant mode of nucleotide substitution, the rate of mutation or other information is known, an informative analysis can be performed [302].

Due to advances in computational biology over the past two decades, especially the Markov Chain Monte Carlo algorithms (MCMC algorithms), Bayesian inference gained in popularity. MCMC is a sampling technique that can be used when exploring tree space and obtain the posterior probabilities of trees [303]. The method starts by simulating a random set of parameters (i.e. random branch legths, random substitution parameters and a random tree) at a random spot in the tree space. Each step in the chain is analyzed using techniques such as SPR (Subtree Pruning and Regrafting) method or NNI (Nearest Neighbor Interchange) and depending on the information obtained, the tree is rearranged by changing of the branch lengths of the tree topology, followed by a change in the algorithm and the creation of a new topology based on the likelihood ratios and the parameter selected for prioritization. The new state/tree is accepted if the product of the likelihood and prior ratio is better than the old one. Following a specified number of iterations of parameter and the chain length, a sample of the tree and its branch lengths are saved. Continuous sampling and tree rearrangement occurs at every few steps in the chain until the end of the chain length. When the chain ends all data sampled is summarized into posterior distribution of parameters and trees.

Major advantages of the Bayesian method is that posterior probabilities for trees and clades can be easily interpreted and the use of prior probability allows for investigator input and more realistic substitution models. The main disadvantages are that MCMC involves heavy computation that posterior probabilities are often high and the selection of suitable models can involve challenging computation [289].

*2) **Distance-based (Matrix) Methods***

Distance-based methods were introduced in the 1960's by Edwards and Cavalli-Sforza [272; 273] and by Fitch and Margoliash [276]. In these methods, the distance between every pair of sequences in the dataset is calculated, and the resulting distance matrix is used to reconstruct the tree employing algorithms based on functional relationships among distance values. Methods used in distance-based methods include the Neighbor-Joining Method (NJ), which arranges operational taxonomic units (OTUs) [277]; the unweighted pair group method with arithmetic means (UPGMA), which uses a sequential clustering algorithm [273] and the Fitch and Magoliash method, which is an extention of UPGMA and also uses a sequential clustering algorithm but, unlike UPGMA, produces trees with an unrooted topology and does not rely on the assumption of a molecular clock [276]; and the Mininum Evolution (ME) Method, which a tree selection method is used to construct trees with shorter branch lengths [278]. An advantage of distance-based methods is their ability to analyze very large phylogenies.

i) UPGMA method

The UPGMA method, introduced by Sokal and Michener in the late 1950's [273], is the simplest and probably the oldest distance-based method of tree construction. The method was originally developed to construct taxonomic phenograms (i.e. trees based on phenotypic similarities) and later adapted to phylogenetic analyses and the construction of ultrametric trees. Ultrametric trees are rooted trees in which all of the end nodes are equidistant from the trunk of the tree, assuming a molecular clock (a constant evolutionary rate). It uses a sequential clustering algorithm and similarity scores to identify OTUs that are similar in sequence and have the shortest evolutionary distance. This process continues or is repeated until only two OTUs remain and are treated as a single new composite OTU [271]. The distance of the newly formed cluster is defined by the algorithm as the average of the distances of the original OTUs. This process assumes that the evolutionary rate from the node of the two clustered OTUs to each of the two OTUs is identical.

ii) Fitch-Margoliash method

This method, which was first introduced in late 1960s, [276], and is broadly viewed as the extention of the UPGMA method of tree inference. It uses a distance-matrix and a

sequential clustering algorithm to find a tree that minimizes the differences between pairwise genetic distances and the distances represented by the sum of branch lengths for each pair of taxa.

iii) Minimum Evolution method

The Minimum Evolution (ME) method was developed in the early 1970s by Kidd and Sgaramella-Zonta [304] with minor changes being introduced by Rzhetsky and Nei in the early 1990s [278; 279]. The method uses the estimated pairwise distances between OTUs, for each tree topology, to estimate branch lengths for each tree and to search for the tree that minimizes the sum (S) of the branch lengths. In this respect, ME is similar to the maximum parsimony (MP) method of tree inference. The main difference between the MP and ME methods, is that in ME method, the branch lengths are inferred from genetic distances, rather than from counting individual nucleotide substitutions over the length of the branch. The main disadavantages of ME is its reliance on an exhaustive search to find the best tree topology, a process that can be computationally intensive when working with large datasets. Heuristic searches that improve the speed of tree inference, such as SPR and NNI, are needed to overcome this limitation. The main advantage of ME method is that it is more efficient when compared with other distance-based methods.

iv) *Neigbor-Joining method*

The Neigbor-Joining (NJ) method was initially described in the late 1980s by Saitou and Nei [277] as a method that minimizes the sum of the least-squares branch lengths (i.e. adopts the minimum evolution criterion). The method, which was modified by Studier and Keppler [305], is the most widely used distance-based method. It uses a distance matrix to specify the genetic distance between each pair of taxa. Following the initial construction of a star-shaped tree, using a cluster algorithm, further analyses are performed by successively choosing pair of taxa to join together based on the distances [271]. NJ method of tree inference does not assume the use of a molecular clock and the tree is unrooted. The main advantage of NJ method is its ability to analysis large data sets that have low levels of sequence divergence. The main disadvantage is its inability to calculate distances when sequences or taxa are highly divergent [306].

1.6 RATIONAL, PROBLEM STATEMENT, HYPOTHESIS, AND AIMS OF THE CURRENT STUDY

1.6.1. Relevance to Existing Knowledge

The increasing prevalence of HIV-1 in sub-Saharan Africa and other regions of the developing world is an urgent public health priority that demands expanded access to antiretroviral therapy. It is becoming increasingly apparent that innovative new strategies will be required if the global AIDS pandemic is to be successfully controlled. As discussed below, an in-depth understanding of HIV-1 immune pathogenesis and genetic evolution in the GIT is critical to the design of more effective treatment strategies aimed at enhancing drug efficacy, suppressing viral replication in cellular reservoirs, and delaying (or preventing) the emergence of drug resistance. The gut-associated lymphoid tissue (GALT), because of its continuous exposure to food and microbial antigens, is in a constant state of "physiological" inflammation characterized by high levels of pro-inflammatory cytokines and an abundance of activated $CCR5^+$ memory $CD4^+$ T cells that are highly permissive for HIV-1 infection **(**168; 170-173; 219; 307; 308). As a result, the gastrointestinal tract (GIT) is a major site of viral replication and $CD4^+$ T cell depletion (304**)**. The majority (~60%) of $CCR5^+CD4^+$ T cells in the effector compartment of the GIT are lost during primary HIV-1 infection with the loss being more severe in the small vs. the large intestine **(**168; 170-172**)**. Several studies have shown that the severity of $CD4^+$ T cell depletion in the GIT (170; 172; 174) during the first few weeks of infection plays an important role in determining the rate of disease progression.

Studying the diversity of HIV-1 variants is an important step in understanding the complexity and nature of the intestinal reservoir, the extent of cross-infection with the systemic circulation and the genetic relationship of intestinal variants to those found in other tissues. Studies of the GIT have been hampered by the limited amount of material available in "pinch" biopsies and by difficulties associated with repeat sampling of multiple sites. Most studies have been confined to a single anatomical site and have involved the sequencing of proviral DNA isolated from patients treated with diverse drug regimens. Overall, these studies have suggested that there is no completely independent evolution of HIV-1 variants in different regions of the GIT and that the intestine is unlikely to be a major source of rebound virus during treatment interruption. Although complete separation of HIV-1 variants was not observed, studies by van Marle *et al*. showed significant tissue-specific clustering of *nef*

proviral DNA in the GIT of ART-naïve patients (210). A recent study of actively replicating virus indicated that, although there was a high degree of concordance between tissues, HIV-1 RNA in the colon and PBMCs of ART-treated patients contained resistance mutations that were not present in plasma RNA, suggesting that virus evolution in the GIT may play a role in the emergence of drug resistance (211).

Given that the time-frame between infection, amplification and dissemination is so compressed, it is virtually impossible to investigate these events directly in HIV-1 infected individuals. The majority of individuals are already in the later stages of infection when they present at the clinical with the flu-like symptoms of acute retroviral infection. As proposed below, using evolutionary and population genetics methods, it should be possible to study the archiving of transmitted viral sequences and their genetic evolution as a function of disease progression, and therapy-induced selection pressures. Many of these important virological and immunological events, which occur within intestinal and cervicovaginal membranes, are not reflected in the peripheral blood. To expand the knowledge base and expedite the development of effective treatment and prevention strategies that are safe, appropriate and affordable for Africa and the developing world, a comprehensive study of the GIT before and after the initiation of antiretroviral therapy has been embarked upon.

1.6.2. ***Problem statement***

It is anticipated that serial sequencing of viruses archived in cellular reservoirs of the GI will provide critical new information needed to unravel the evolutionary history of HIV-1 infection and to better understand host-virus interactions. This provides a clear setting of the research questions and a logical lead to the experiemental approach and methods used in the study.

1.6.3. Hypothesis

Based on the hypothesis that the gastrointestinal tract encodes the complete evolutionary and immunological history of HIV-1 infection, from the very onset of infection, we proposed a comprehensive study of the GI tract in Africa patients infected with subtype C viruses. In this study I focused on the compartmentalization, genetic evolution and adaptive history of HIV-1 in the different regions of the gastrointestinal tract. I expected that serial sequencing of the viruses archived in cellular reservoirs of the GI tract will provide the information needed to unravel the complete evolutionary history of the infection, from its

inception. The study was unique and innovative in that 1) we used a new double balloon technique to sample along the entire GI tract; 2) we systematically examined the impact of antiretroviral therapy on the containment (and clearance) of HIV-1 throughout the entire GI tract and 3) this study was conducted in a real-life African setting, where these virological events are likely to be super-imposed on a high pre-existing burden of gastrointestinal disease.

We, and others, have shown that there are marked differences in HIV-1 pathogenesis in the small vs. large intestine in terms of viral replication, $CD4^+$ T cell levels and CD4+ T cell restoration during ART and that this is due, at least in part, to tissue-specific differences in inflammation and immune activation. Based on these findings, we further hypothesized that different inflammatory milieus in the GIT might lead to the emergence of HIV-1 RNA variants with differing pathologies. Such a phenomenon would have important implications not only for interrupting disease progression but also for the optimization of treatment strategies.

1.6.4. Aims

The studies in this thesis were performed on a cohort of South African patients with late-stage HIV-1/AIDS. **Objectives**: Primary objectives of this thesis were to: 1) characterize the diversity of HIV-1 RNA variants in different parts of the GIT; 2) determine whether there is compartmentalized evolution of HIV-1 RNA variants in the GIT and whether these variants are likely to have different biological properties; 3) investigate the impact of ART on immune restoration in the GIT.

1.7. FIGURE LEGENDS

Figure 1.7.1: Global view of HIV-1 infection in 2009. Map showing the distribution and frequency of HIV-1/AIDS in different regions of the world relative to the total number of HIV-1 infections around the world as presented in the global AIDS Epidemic report for 2010.

Figure 1.7.2: HIV-1 prevelance (%) in African adults. Map showing the distribution and frequency of HIV-1/AIDS in sub-Saharan Africa relative to the total number of HIV-1

infections around the continent of Africa as presented in the USAIDS report on the global AIDS Epidemic for 2010.

Figure 1.7.3: HIV-1 subtypes and recombinant forms. Map showing the distribution of HIV-1 subtypes and Circulating Recombinant Forms (CRFs) at the global level. This map is based on prevalence data contained in the UNAIDS report on AIDS and as published by Taylor et al. N Eng J Med 2008; 358:1590-1602.

Figure 1.7.4: A schematic representation of the HIV virion, and its associated proteins.

Figure 1.7.5: A diagrammatic representation of the HIV-1 genome. All three structural genes (i.e. env, gag, and pol genes) and accessory and regulatory genes (Tat, Rev, Vif, Vpu, Vpr and Nef) are shown (http://hiv-web.lanl.gov/content/hivdb).

Figure 1.7.6: Schematic representation of the different basic steps involved in the construction of a phylogenetic tree.

Figure 1.7.7: A schematic representation/breakdown of the important models of nucleotide substitution available to date. JC69 was the first model of nucleotide substitution to be developed and each of the other models are an extension of the JC69 model. The K80 model [Kimura, 1980] allows for different Transition to Transversion (Ts/Tv) ratios, but keeps the base frequencies equal, while the F81 model [Felsenstein, 1981] allows for the base frequencies to very but keep the Ts/Tv ratio equal. The HKY model [Hasegawa *et al*, 1985] is basically a combination of the K80 and F81 model in that it allows for unequal base frequencies and Ts/Tv ratio. The TN93 (Tamura and Nei, 1993) is an extention of the F84 model [Felsentein, 1984; Kishino and Hasegawa, 1989], and allows different nucleotide substitution rates for purine (A ↔ G) and pyrimidines (C ↔ T) transitions. The GTR model [Rodriguez *et al*, 1990; Yang *et al*, 1994] is an extension of the HKY model, but allows each of the 6 parameters to have its own probability while not assuming any direction in the change over time.

1.8 TABLES LEGENDS:

Table 1.8.1: Summary of HIV-1 genes and proteins (accessory and regulatory), and their functions.

Table 1.8.2: Classification of phylogenetic tree-reconstruction and tree analysis methods according to the data and algorithmic strategies.

1.9. Figures

Figure 1.7.1:

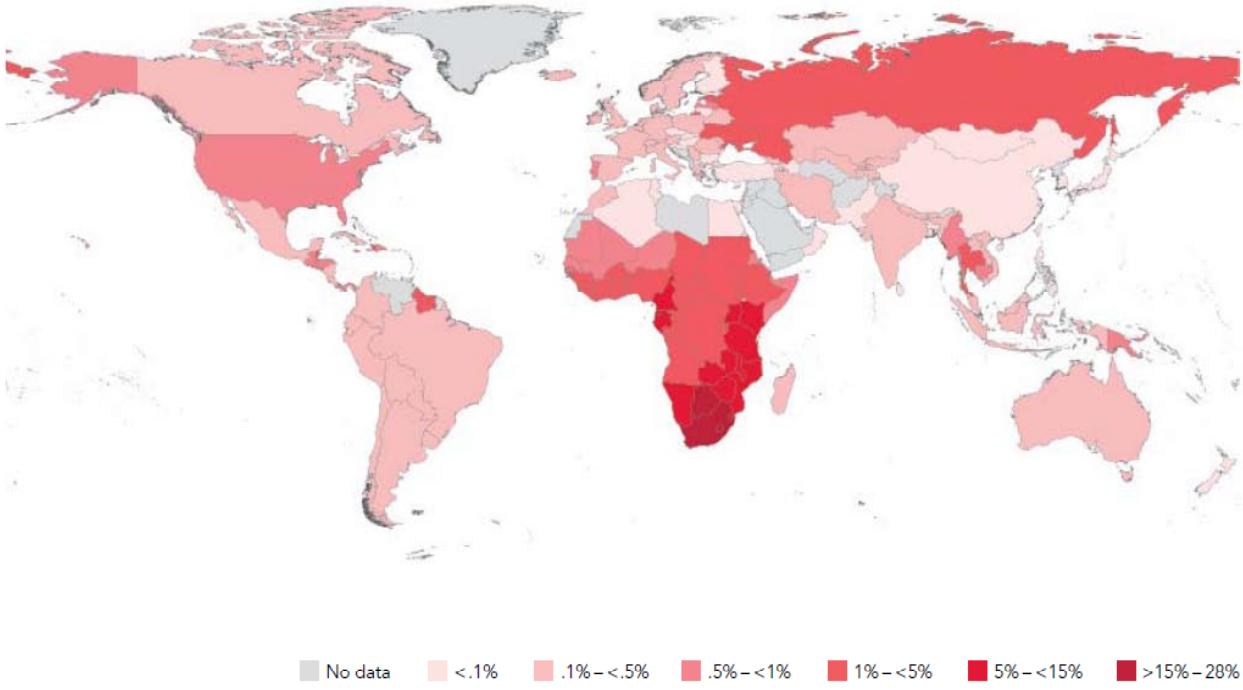


Figure 1.7.2:

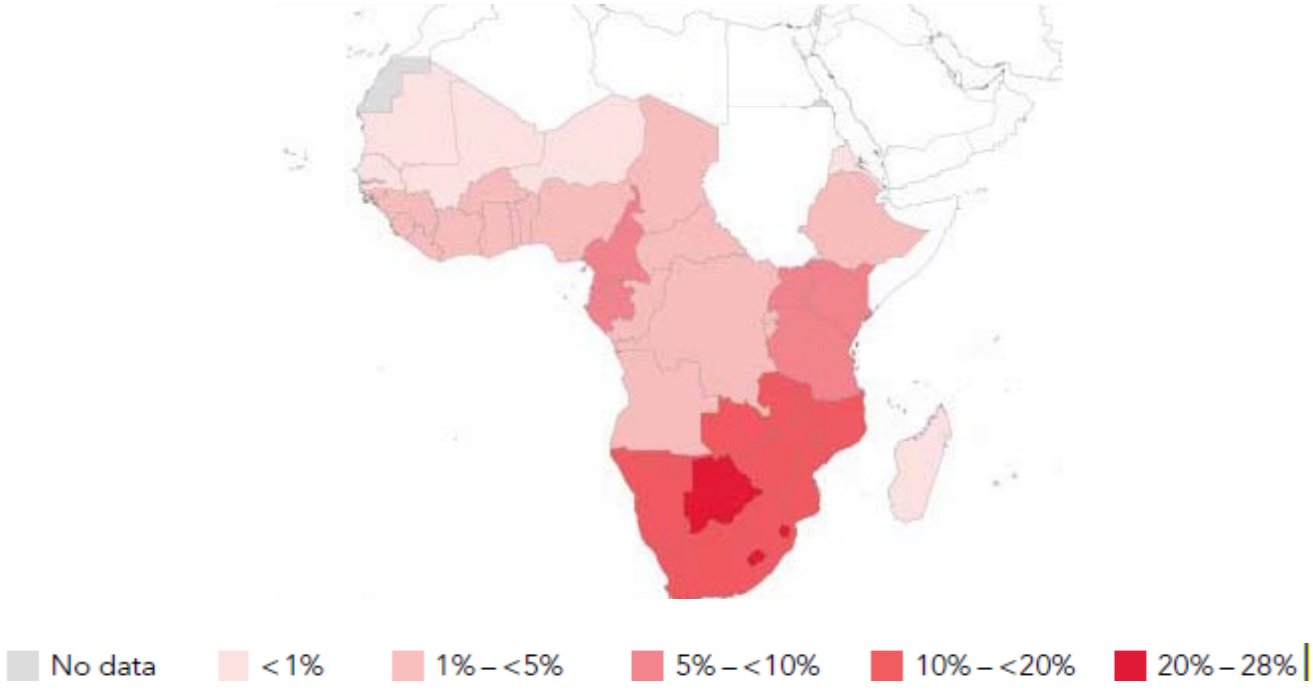


Figure 1.7.3:

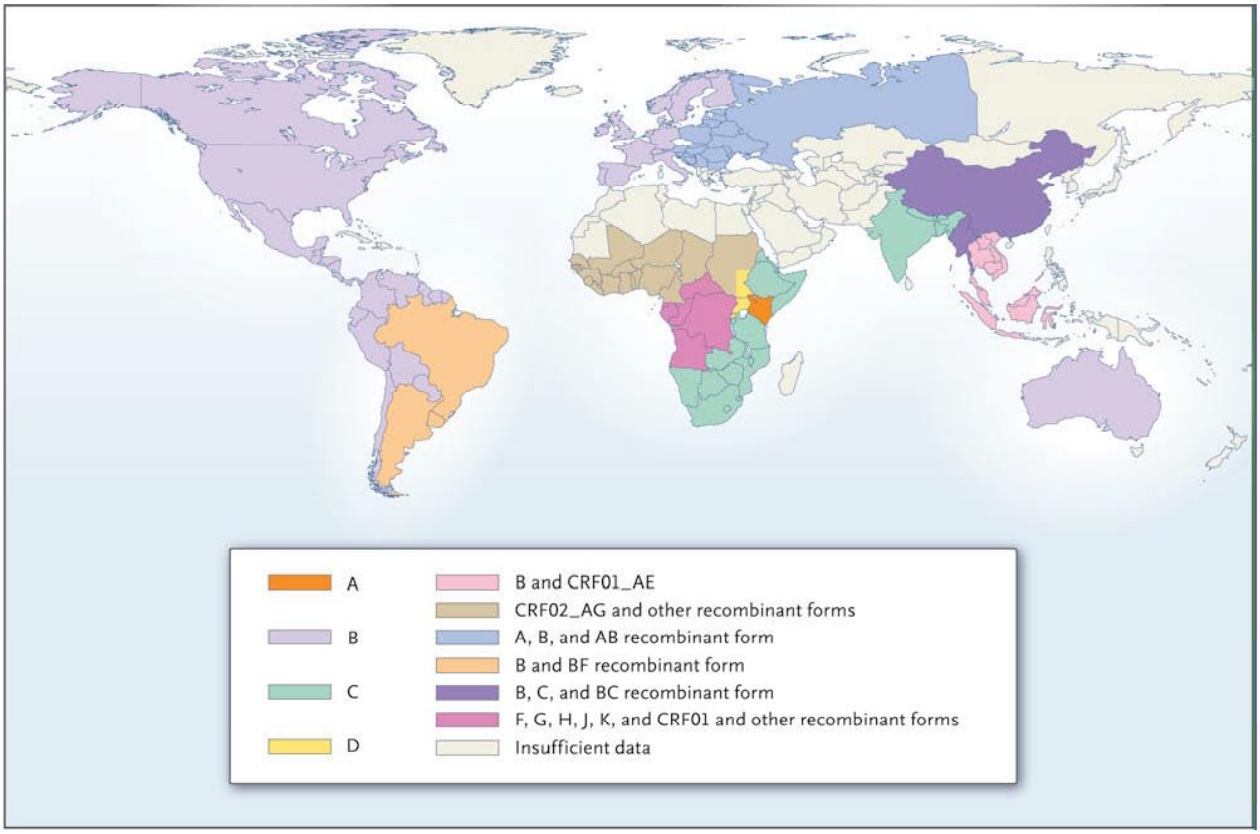
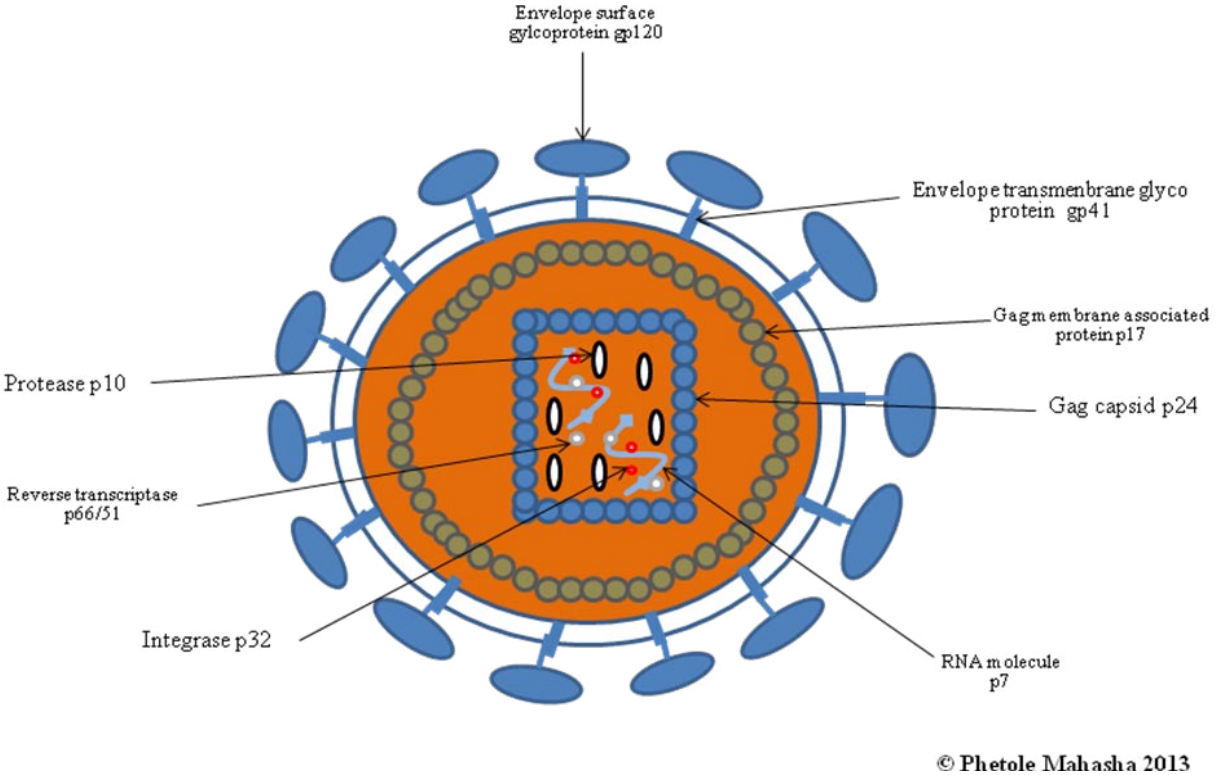


Figure 1.7.4:



© Phetole Mahasha 2013

Figure 1.7.5:

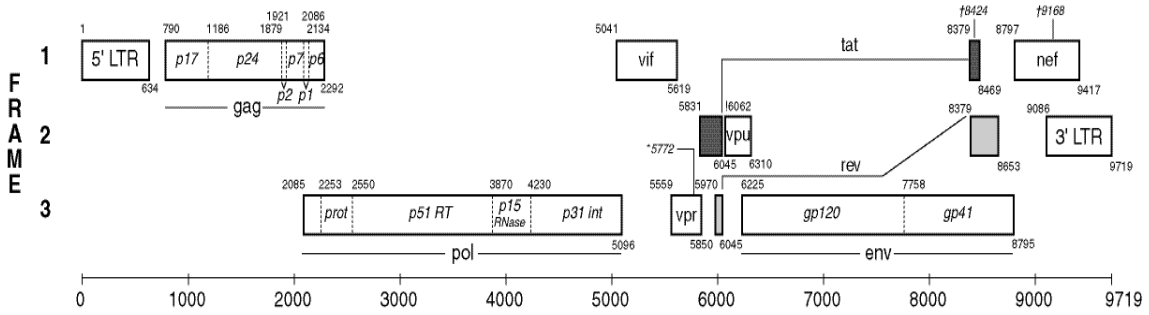


Figure 1.7.6:

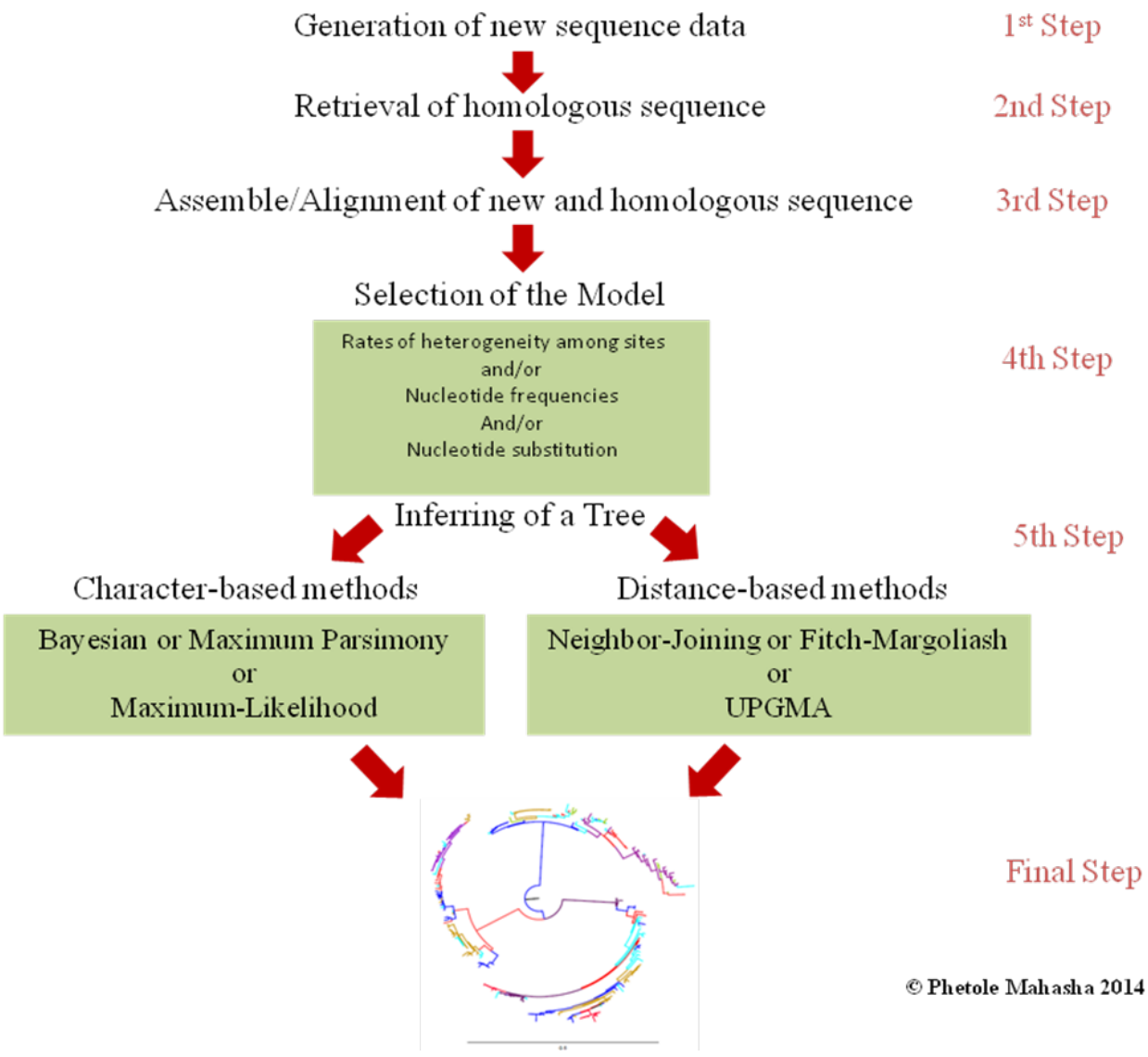
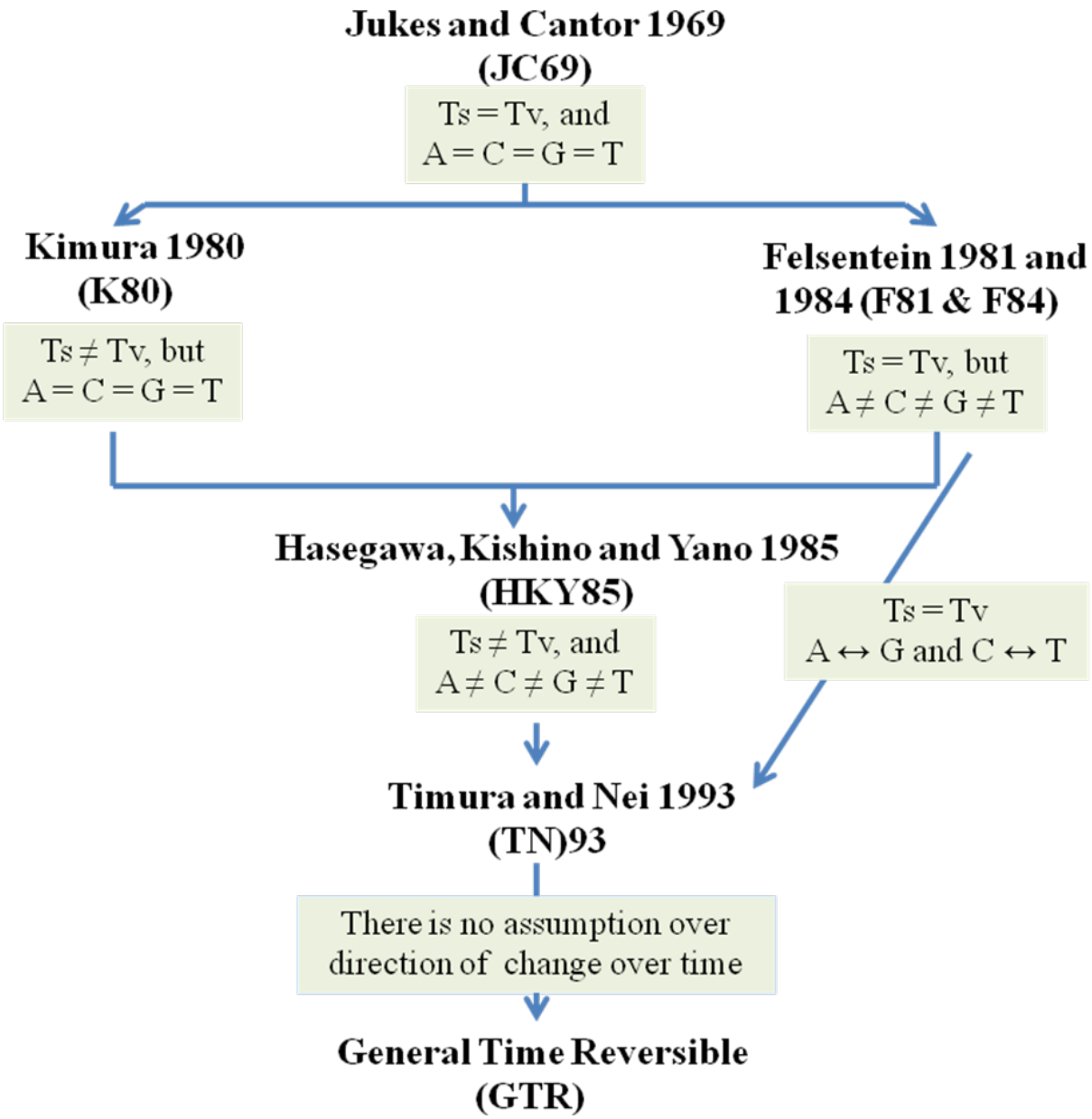


Figure 1.7.7:



1.10. Tables:

Table 1.8.1:

Gene	Protein	Type	Function
tat	Tat (p14/p16)	regulatory	Viral transcription transactivator
Rev	Rev (p19)	regulatory	Upregulation expression of *Gag, Pol* and *Env;* and also downregulates itself and Tat
nef	Nef (p25/p27)	accessory	Downregulates CD4 and MHC-I, prevents apoptosis
vif	Vif (p23)	accessory	Promotes virion maturation and infectivity, inhibits APOBEC function
vpr	Vpr	accessory	Involved in nuclear entry, prevents cell division
vpu	Vpu	accessory	Down modulates CD4 in ER, promotes virion

Table 1.8.2:

Method	Optimality search criterion or goodness-of-fit criterion	Clustering algorithm
Character-based methods	Maximum Parsimony (MP)	
	Maximum-Likelihood (ML)	
	Bayesian Inference	
Distance-based methods	Fitch-Margoliash	UPGMA
	Minimum Evolution (ME)	Neighbor-Joining (NJ)

1.11. REFERENCES:

1. Hymes, K. B., Greene, J. B., Marcus, A., Ballard, H., and William, D. C., Laubenstein, L. J. 1981. 'Kaposi's sarcoma in homosexual men: A report of eight cases. *Lancet* **2**:598-600.
2. Center for Disease Control (CDC). 1981. Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men. New York City and California. *MMWR Morb. Mortal. Wkly Rep*. **30**:305-8.
3. Center for Disease Control (CDC). 1981. Pneumocystis pneumonia—Los Angeles. *MMWR Morb. Mortal. Wkly Rep*. **30**:250-2.
4. Gottlieb M. S., Schroff R., Schanker H. M., Weisman, J. D., Fan, P. T., Wolf, R. A. and Saxon, A. *et al*. 1981. `Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency'. *N Engl J Med*. **305**:1425-31.
5. Marx, J. L. 1982. "New disease baffles medical community". *Science* **217** (4560): 618–21.
6. Center for Disease Control (CDC). 1982. Kaposi's Sarcoma (KS), Pneumocystis Carinii Pneumonia (PCP), and Other Opportunistic Infections (OI): Cases Reported to CDC as of July 8'.
7. Greene W. C. 2007. A history of AIDS: Looking back to see ahead. *European Journal of Immunology*, **37:** S94-S102.
8. Masur H., Michelis M. A., Greene J. B., Onorato I., Stouwe R. A., Holzman R. S., Wormser G., Brettman L., Lange M., Murray H. W., and Cunnigham-Rundles, S. 1981. An Outbreak of community acquired Pneumocystis carinii pneumonia: initial manifestation of cellular immune dysfunction'. *N Engl J Med*. **305** (24):1431-1438.
9. Dubois, R. M., Braitwaite, M. A., Mikhail, J. R. et al., 1981. 'Primary Pneumocystis Carinii and Cytomegalovirus Infections', *The Lancet*, **ii**, 1339.
10. Vilaseca, J., Arnau, J. M., Bacardi, R., Mieras, C., Serrano, A., and Navarro, C. 1982. *Kaposi's sarcoma* and *toxoplasma gondii* drain abscess in a Spanish homosexual. *Lancet* **1**:572.
11. Rozenbaum, W., Coulaud, J. P., Saimot, A. G., Klatzmann, D., Mayaud, C., and Carette, M. F. 1982. Multiple opportunistic infections in a male homosexual in France. *Lancet* **1**:572-3.
12. Francioli, P., Vogt, M., Scandelin, J., Clement, F., Russi, E., Delacretaz, F., Perret, C., and GlauSer, M. P. 1982. [Acquired immunologic deficiency syndrome, opportunistic

infections and homosexuality. Presentation of 3 cases studied in Switzeland]. *Schweiz Med. Wochenschr*. **112**:1682-7.

13. Serwadda, D. Mugerwa, R. D., Sewankambo, N. K., Lwegaba, A., Carswell, J. W., Kirya, G. B., Bayley, A. C., Downing, R. G., Tedder, R. S., Clayden, S. A. Weiss R. A. and Dalgleish A. G. 1985. Slim disease: a new disease in Uganda and its association with HTLV-III infection. *Lancet*, **2**: 849–852.
14. Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L. 1983. "Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)". *Science* 220 (4599): 868–71.
15. Gallo, R. C., S. Z. Salahuddin, M. Popovic, G. M. Shearer, M. Kaplan, B. F. Haynes, T. J. Palker, Redfield, R.; Oleske, J.; Safai, B.; White, G., Foster, P., and Markham, P. D. 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224**:500-503.
16. Levy, J. A., Hoffman, A. D., Kramer, S. M., Landis, J. A., Shimabukuro, J. M., and Oshiro, L. S. 1984. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* **225**:840-2.
17. World Health Organisation by the end of 1985 (Bureau of Hygiene & Tropical Diseases (1986) 'AIDS newsletter' Issue 1 January 30th.
18. AmfAR Special Report: AIDS in Africa http://amfar.org/cgi-bin/programs/globali/record.html? record=34.
19. 2008 Report on the Global AIDS Epidemic.
20. Department of Health of South Africa, 2010.
21. UNAIDS 2006, Global Report on AIDS.
22. Global AIDS Report on AIDS, 2006.
23. Barth, R. E., van der Loeff, M. F., Schuurman, R., Hoepelman, A. I., and Wensing, A. M. 2010. Virological follow-up of adult patients in antiretroviral treatment programmes in sub-Saharan Africa: a systematic review. *Lancet Infect Dis* **10**:155-166.
24. Hahn, B. H., G. M. Shaw, K. M. De Cock, and P. M. Sharp. 2000. AIDS as a zoonosis: scientific and public health implications. *Science* **287**:607-614.
25. Rambaut, A., Posasa, D., Crandall, K. A., and Holmes, E. C. 2004. The causes and consequences of HIV evolution. *Nature Reviews Genetics* **5**, 52-61.

26. Salemi, M., Strimmer, K., Hall, W. W., Duffy, M., Delaporte, E., Mboup, S., Peeters, M., and Vandamme, A. M. 2001. Dating the common ancestor of SIVcpz and HIV-1 group M and the origin of HIV-1 subtypes us *FASEB J* **15**: 276-8.

27. Keele, B. F., Van Heuverswyn, F., Li, Y., Bailes, E., Takehisa, J., Santiago, M. L., Bibollet-Ruche, F., Chen, Y., Wain, L. V., Liegeois, F., Loul, S., Ngole, E. M., Bienvenue, Y., Delaporte, E., Brookfield, J. F., Sharp, P. M., Shaw, G. M., Peeters, M., and Hahn, B. H. 2006. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* **313**:523-526.

28. Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B. H., Wolinsky, S., and Bhattacharya, T. 2000. Timing the ancestor of the HIV-1 pandemic strains. *Science* **288**:1789-96.

29. Gao, F., E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp, and Hahn, B. H. 1999. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* **397**:436-441.

30. Gao, F., Yue, L., White, A. T., Pappas, P. G., Barchue, J., Hanson, A. P., Greene, B. M., Sharp, P. M., Shaw, G. M., and Hahn, B. H. 1992. Human infection by genetically diverse SIVSM-related HIV-2 in West Africa. *Nature.* **358** (6386):495–499.

31. Lihana, R.W., Ssemwanga, D., Abimiku, A. and Ndembi N. 2012. Update on HIV-1 diversity in Africa: Decade in Review. *AIDS Rev*. **14**: 83-100

32. Plantier, J-C., Leoz, M., Dickerson, J.E., De Oliveira, F., Cordonnier, F., Lemée, V., Damond, F., Robertson, D. and Simon, F. 2009. A new human immunodefieciency virus derived from gorillas. *Nat. Med.* **15**: 871-872.

33. Vallari, A., Holzmayer, V., Harris, B., Yamaguchi, J., Ngansop, C., Makamche, F., Mbanya, D., Kaptué, L., Ndembi, N., Gürtler, L., Devare, S. and Brennan, C.A. 2011. Confirmation of putative HIV-1 group P in Cameroon. J. Virol. 85(3): 1403-1407.

34. Taylor, B. S., Sobieszczyk, M. E., McCutchan, F. E., and Hammer, S. M. 2008. The challenge of HIV-1 subtype diversity. *N Engl J Med* **358**:1590-1602.

35. Janssens, W., Buve, A., and Nkengasong, J. N. 1997. The puzzle of HIV-1 subtypes in Africa. *AIDS* **11**:705-12.

36. Toure-Kane, C., Montavon, C., Faye, M. A., Gueye, P. M., Sow, P. S., Ndoye, I., Gaye-Diallo, A., Delaporte, E., Peeters, M., and Mboup, S. 2000. Identification of all HIV type 1 group M subtypes in Senegal, a country with low and stable seroprevalence. *AIDS Res. Hum. Retrovir*. **16**:603-609.

37. McCutchan, F. E. 2006. Global epidemiology of HIV. *J.Med Virol*. **78**:S7-S12.

38. Stevenson, M. 2003."HIV-1 pathogenesis". *Nat Med*, **9**:853-60.

39. Sagar, M., Wu, X., Lee, S., and Overbaugh, J. 2006. Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity. *J Virol* **80**:9586-98.

40. Frost, S. D. W., Wrin, T., Smith, D.M., Pond, S. L. K, Liu, Y., Paxinos, E., Chappey, C., Galovich, J., Beauchaine, J., Petropolos, C. J., Little, S. J., and Richman, D. D. 2005. Neutralizing antibody responses drive the evolution of human immunodeficiency virus type 1 envelope during recent HIV infection. *Proc Natl Acad Sci USA* **102**:18514-9.

41. McClutchan, F. E., Carr, J. K., Murphy, D., Piyasirisilp, S., Gao, F., Hahn, B., Yu, X. F., Beyrer, C., and Birx, D. L. 2002. Precise mapping of recombination breakpoints suggests a common parent of two BC recombinant HIV type 1 strains circulating in China. *AIDS Res Hum Retroviruses* **18**:1135-1140.

42. Hemelaar, J., Gouws, E., Ghys, P. D., and Osmanov, S. 2006. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* **20**:W13-23.

43. Stebbing, J., and Moyle, G. 2003. The clades of HIV: their origins and clinical significance. *AIDS Rev* **5**: 205-13.

44. Spira, S., Wainberg, M. A., Loemba, H., Turner, D., and Brenner, B. G. 2003. Impact of clade diversity on HIV-1 virulence, antiretroviral drug sensitivity and drug resistance. *J Antimicrob Chemother* **51**:229-240.

45. Alaeus, A., Lidman, K., Bjorkman, A., Giesecke, J., Albert, J. 1999. Similar rate of disease progression among individuals infected with HIV-1 genetic subtypes A–D. *AIDS* **13**:901-7.

46. John-Stewart, G. C., Nduati, R. W., Rousseau, C. M., Mbori-Ngacha, D. A., Richardson, B. A., Rainwater, S., Panteleeff, D. D., and Overbaugh, J. 2005. Subtype C is associated with increased vaginal shedding of HIV-1. *J Infect Dis* **192**:492-6.

47. Renjifo, B., Gilbert, P., Chaplin, B., Msamanga, G., Mwakagile, D., Fawzi, W., and Essex, M. 2004. Preferential in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D. *AIDS* **18**:1629-1636.

48. Kanki, P. J., Hamel, D. J., Sankale, J. L., Hsieh, C., Thior, I., Barin, F., Woodcock, S. A., Gueye-Ndiaye, A., Zhang, E., Montano, M., Siby, T., Marlink, R., NDoye, I., Essex, M. E. and MBoup, S. 1999. Human immunodeficiency virus type 1 subtypes differ in disease progression. *J Infect Dis* **179**:68-73.

49. Kaleebu, P., French, N., Mahe, C., Yirrell, D., Watera, C., Lyagoba, F., Nakiyingi, J., Rutebemberwa, A., Morgan, D., Weber, J., Gilks, C., and Whitworth, J. 2002. Effect of human immunodeficiency virus (HIV) type 1 envelope subtypes A and D on disease progression in a large cohort of HIV-1-positive persons in Uganda. *J Infec Dis.* **185**:1244-1250.

50. Brenner, B.G., Oliveira, M., Donalla-Bell, F., Moise, D.D., Ntengwa, M., Frankel, F., Essex, M., and Wainberg, M.A. 2006. HIV-1 subtype C viruses rapidly develop K65R resistance to tenofovir in cell culture. AIDS 20(9): F9-F13.

51. Invernizzi, C.F., Coutsinos, D., Oliveira, M., Moise, D., Brenner, B.G., and Wainberg, M.A. 2009. Signatur nucleotide polymorphisms at positions 64 and 65 in Reverse Transcriptase favor the selction of K65R resistance mutation in HIV-1 subtype C.

52. Wainberg, M.A., Zaharatos, G.J., and Brenner, B.G. 2011. Development of antoretrovial drug resistance. N. Engl J Med. 365(7): 637-646.

53. Baeten, J. M., Chohan, B., Lavreys, L., Chohan, V., McClelland, R. S., Certain, L., Mandaliya, K., Jaoko, W., and Overbaugh, J. 2007. HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. *J Infect Dis* **195**:1177-1180.

54. Huang, W., Eshleman, S. H., Toma, J., Fransen, S., Stawiski, E., Paxinos, E. E., Whitcomb, J. M. Young, A. M., Donnell, D., Mmiro, F., Musoke, P., Guay, L. A., Jackson, J. B., Parkin, N. T., and Petropoulos, C. J. 2007. Co-receptor tropism in human immunodeficiency virus type 1 subtype D: high prevalence of CXCR4 tropism and heterogeneous composition of viral populations. *J Virol* **81**:7885-7893.

55. Quinones-Mateu, M. E., Albright, J. L., Mas, A., Soriano, V., and Arts, E. J. 1998. Analysis of *pol* gene heterogeneity, viral quasispecies, and drug resistance in individuals infected with group O strains of human immunodeficiency virus type 1. *J Virol* **72**: 9002-15.

56. Descamps, D., Collin, G., Loussert-Ajaka, I., Saragosti, S., Simon, F., and Brun-Vezinet, F. 1995. HIV-1 group O sensitivity to antiretroviral drugs. *AIDS* **9**: 977-8.

57. Descamps, D., Collin, G., Letourneur, F., Apetrei, C., Damond, F., Loussert-Ajaka, I., Simon, F., Saragosti, S., and Brun-Vézinet, F. 1997. Susceptibility of human immunodeficiency virus type 1 group O isolates to antiretroviral agents: In vitro phenotypic and genotypic analyses. *J Virol.* **71**:8893-8.

58. Apetrei, C., Descamps, D., Collin, G., Loussert-Ajaka, I., Damond, F., Duca, M., Simon, F., and Brun-Vézinet, F. 1998. Human immunodeficiency virus type 1 subtype F reverse transcriptase sequence and drug susceptibility. *J Virol*; **72**:3534-8.

59. Loemba, H., Brenner, B., Parniak, M. A., Ma'ayan, S., Spira, B., Moisi, D., Oliveira, M., Detorio, M., and Wainberg, M. A. 2002. Genetic Divergence of Human Immunodeficiency Virus Type 1 Ethiopian Clade C Reverse Transcriptase (RT) and Rapid Development of Resistance against Nonnucleoside Inhibitors of RT. *Antimicrob. Agents Chemother*. **46**:2087-94.

60. Montano, M. A., Novitsky, V. A., Blackard, J. T., Cho, N. L., Katzenstein, D. A., and Essex, M. (1997). Divergent transcriptional regulation among expanding human immunodeficiency virus type 1 subtypes. *J Virol*. **71**:8657-65.

61. Jeeninga, R. E., Hoogenkamp, M., Armand-Ugon, M., de Baar, M., Verhoef, K., and Berkhout, B. 2000. Functional Differences between the Long Terminal Repeat Transcriptional Promoters of Human Immunodeficiency Virus Type 1 Subtypes A through G. *J. Virol*. **74**:3740-3751.

62. Zacharova, V., Becker, M. L., Zachar, V., Ebbesen, P., and Goustin, A. S. 1997. DNA sequence analysis of the long terminal repeat of the C subtype of human immunodeficiency *AIDS Res Hum Retroviruses* **13**: 719-24.

63. Naghavi, M. M., Schwartz, S., Sonnerborg, A., and Vahlne, A. 1999. Long terminal repeat promoter/enhancer activity of different subtypes of HIV type 1. *AIDS Res Hum Retroviruses* **15**:1293-303.

64. Rosen, C. A., Haseltine, W. A., Lenz, J., Ruprecht, R., and Cloyd. M. W. 1985. Tissue selectivity of murine leukemia virus infection is determined by long terminal repeat sequences. *J Virol*. **55**:862-6.

65. Coffin, J., Hughes, H. S., and Varmus, S. 1998. *Retroviruses*. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY.

66. Gupta, S., and Mitra, D. 2007. Human immunodeficiency virus-1 Tat protein: immunological facets of a transcriptional activator. *Indian J Biochem Biophys* **44**:269-275.

67. Yu, X. G., Lichterfeld, M., Addo, M. M., and Altfeld, M. 2005. Regulatory and accessory HIV-1 proteins: potential targets for HIV-1 vaccines? *Curr Med Chem* **12**:741-747.

68. Malim, M. H., and Emerman, M. 2008. HIV-1 accessory proteins-ensuring viral survival in a hostile environment. *Cell Host Microbe* **3**:388-398.

69. Markosyan, R. M., Cohen, F. S., and Melikyan, G. B. 2003. HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation. *Mol Biol Cell* **14**:926-38.

70. Gomez, C., and Hope, T. J. 2005. The ins and outs of HIV replication. *Cell Microbiol.* **7**:621-6.

71. Melikyan, G. B., Markosyan, R. M., Hemmati, H., Delmedico, M. K., Lambert, D. M., and Cohen, F. S. 2000. Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion. *J Cell Biol* **151**:413-23.

72. Briggs, D. R., Tuttle, D. L., Sleasman, J. W., and Goodenow, M. M. 2000. Envelope V3 amino acid sequence predicts HIV-1 phenotype (co-receptor usage and tropism for macrophages). *AIDS* **14**:2937-2939.

73. Jensen, M. A., Li, F. S., van 't Wout, A. B., Nickle, D. C., Shriner, D., He, H. X. , McLaughlin, S., Shankarappa, R., Margolick, J. B., and Mullins, J. I. 2003. Improved co-receptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences. *J Virol* **77**:13376-13388.

74. Resch, W., Hoffman, N., and Swanstrom, R. 2001. Improved success of phenotype prediction of the human immunodeficiency virus type 1 from envelope variable loop 3 sequence using neural networks. *Virology* **288**:51-62.

75. Steffens, C. M., and Hope, T. J. 2004. Mobility of the Human Immunodeficiency Virus (HIV) Receptor CD4 and Coreceptor CCR5 in Living Cells: Implications for HIV Fusion and Entry Events. *J. Virol.* **78**: 9573-9578.

76. Dismuke, D. J., and Aiken, C. 2006. Evidence for a functional link between uncoating of the human immunodeficiency virus type 1 core and nuclear import of the viral preintegration complex. *J Virol* **80**: 3712-20.

77. Forshey, B. M., von Schwedler, U., Sundquist, W. I. and Aiken, C. 2002. Formation of a Human Immunodeficiency Virus Type 1 Core of Optimal Stability Is Crucial for Viral Replication. *J. Virol.* **76**, 5667–5677.

78. McDonald, D., Wu, L.; Bohks, S.M., Kewalramani, V.N., Unutmaz, D., and Hope, T.J. 2003). Recruitment of HIV and its receptors to dendritic cell-T cell junctions. *Science* **300**: 1295-7.

79. Rose, K. M., Marin, M., Kozak, S. L., and Kabat, D. 2004. The viral infectivity factor. (Vif) of HIV-1 unveiled. *Trends Mol Med* **10**:291-7.

80. Bieniasz, P. D. 2004. Intrinsic immunity: a front-line defense against viral attack. *Nat. Immunol.* **5**:1109-15.

81. Chen, H., Wei, S. Q., and Engelman, A. 1999. Multiple integrase functions are required to form the native structure of the human immunodeficiency virus type I intasome. *J Biol Chem* **274**:17358-17364.

82. Khiytani, D. K., and Dimmock, N. J. 2002. Characterization of a human immunodeficiency virus type 1 pre-integration complex in which the majority of the cDNA is resistant to DNase I digestion. *J Gen Virol* **83**:2523-2532.

83. Bushman, F. 2002. Targeting retroviral integration? *Mol Ther* **6**:570-571.

84. Bushman, F. D. 2003. Targeting survival: integration site selection by retroviruses and LTR-retrotransposons. *Cell* **115**:135-138.

85. Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F. 2002. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**:521-529.

86. Bukrinsky, M. I., Sharova, N., Dempsey, M. P., Stanwick, T. L., Bukrinskaya, A. G., Haggerty, S., and M. Stevenson. 1992. Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *Proc Natl Acad Sci USA.* **89**:6580-6584.

87. Hiscott, J., Kwon, H., and Genin, P. 2001. Hostile takeovers: viral appropriation of the NF-kappaB pathway. *J Clin Invest* **107**:143-151.

88. Gaynor R. 1992. Cellular transcription factors involved in the regulation of HIV- I gene expression. *AIDS* **6**:47-63.

89. Antoni, B. A., Stein, S. B., and Rabson, A. B. 1994. Regulation of human immunodeficiency virus infection: implications for pathogenesis. *Adv. Virus Res.* **43**:53–145.

90. White, M. K., Gorrill, T. S., and Khalili, K. 2006. Reciprocal transactivation between HIV-1 and other human viruses. *Virol.* **352**: 1-13.

91. Cassol, E., Cassetta, L., Rizzi, C., Alfano, M., and Poli, G. 2009. M1 and M2a polarization of human monocyte-derived macrophages inhibits HIV-1 replication by distinct mechanisms. *J Immunol.* **182**:6237-6246.

92. Cassol, E., Cassetta, L., Alfano, M., and Poli, G. 2010. Macrophage polarization and HIV-1 infection. *J Leukoc Biol.* **87**:599-608.

93. Freed, E. O. 1998. HIV-1 gag proteins: diverse functions in the virus life cycle. *Virol.* **251**:1-15.

94. Wilk, T., Gross, I., Gowen, B. E., Rutten, T., de Haas, F., Welker, R., Krausslich, H. G., Boulanger, P., and Fuller, S. D. 2001. Organization of immature human immunodeficiency virus type 1. *J Virol.* **75**:759-771.

95. Cimarelli, A.,. Sandin, S., Hoglund, S., and Luban, J. 2000. Basic residues in human immunodeficiency virus type 1 nucleocapsid promote virion assembly via interaction with RNA. *J Virol.* **74**:3046-3057.

96. Zhang, Y., and Barklis, E. 1997. Effects of nucleocapsid mutations on human immunodeficiency virus assembly and RNA encapsidation. *J Virol.* **71**:6765-6776.

97. Bryant, M., and Ratner, L. 1990. Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proc Natl Acad Sci USA.* **87**:523-527.

98. Sandefur, S., Smith, R., Varthakavi, V., and Spearman, P. 2000. Mapping and characterization of the N-terminal I domain of HIV type 1 Pr55 (Gag). *J Virol* .**74**:7238-49.

99. Derdowski, A., Ding, L., and Spearman, P. 2004. A Novel Fluorescence Resonance Energy Transfer Assay Demonstrates that the Human Immunodeficiency Virus Type 1 $Pr55^{Gag}$ I Domain Mediates Gag-Gag Interactions *J Virol.* **78**: 1230-42.

100. Freed, E.O. 2002. Viral late domains. *J Virol.* **76**: 4679-87.

101. Booth, A. M., Fang, Y., K. Fallon, J., Yang, J. M., Hildreth, J. E., and Gould, S. J. 2006. Exosomes and HIV Gag bud from endosome-like domains of the T cell plasma membrane. *J Cell Biol.* **172**:923-935.

102. Perlman, M., and Resh, M. D. (2006). Identification of an intracellular trafficking and assembly pathway for HIV-1 gag. *Traffic* **7**:731-745.

103. Gould, S. J., Booth, A. M., and Hildreth, J. E. 2003. The Trojan exosome hypothesis. Proc Natl Acad Sci *USA.* **100**:10592-10597.

104. Pelchen-Matthews, A., Kramer, B., and Marsh, M. 2003. Infectious HIV-1 assembles in late endosomes in primary macrophages. *J Cell Biol.* **162**:443-455.

105. Raposo, G., Moore, M., Innes., D., Leijendekker, R., Leigh-Brown, A., Benaroch, P., and Geuze, H. 2002. Human macrophages accumulate HIV-1 particles in MHC II compartments. *Traffic* **3**:718-729.

106. Groot, F., Welsch, S., and Sattentau, Q. J. 2008. Efficient HIV-1 transmission from macrophages to T cells across transient virological synapses. *Blood* **111**:4660-4663.

107. Sharova, N., Swingler, C., Sharkey, M., and Stevenson, M. 2005. Macrophages archive HIV-1 virions for dissemination in trans. *EMBO J.* **24**:2481-2489.

108. Garrus, J. E., von Schwedler, U. K., Pornillos, O. W., Morham, S. G., Zavitz, K. H., Wang, H. E., Wettstein, D. A., Stray, K. M., Côté, M., Rich, R. L., Myszka, D. G., and Sundquist, W. I. 2001. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **107**:55-65.

109. Raiborg, C., Rusten, T. E., and Stenmark, H. 2003. Protein sorting into multivesicular endosomes. *Curr Opin Cell Biol.* **15**(4):446-55.

110. Preston, B. D., Poiesz, B. J., and Loeb, L. A. 1988. Fidelity of HIV-1 reverse transcriptase. *Science* **242**:1168-71.

111. Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M., and Ho, D. D. 1997. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271**:1582-86.

112. Jetzt, A. E., Yu, H., Klarmann, G. J., Ron, Y., Preston, B. D., and Dougherty, J. P. 2000. High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J. Virol.* **74**:1234–1240.

113. Nowak, M. A., and May, R. M. 1992. Coexistence and competition in HIV infections. *J Theor Biol* **159**: 329-342.

114. Nowak, M. A., and May, R. M. 1992. Evolutionary games and spatial chaos. *Nature.* **359**: 826-829.

115. Barouch, D. H. 2008. Challenges in the development of an HIV vaccine. *Nature* **455**:613-9.

116. Domingo, E., and Holland, J. J. 1997. RNA virus mutations and fitness for survival. *Annu Rev Microbiol.* **51**:151-78.

117. Domingo, E., Escarmis, C., Sevilla, N., Moya, A., Elena, S. F., Quer, J., Novella, I. S., and Holland, J. J. 1996. Basic concepts in RNA virus evolution. *FASEB J.* **10**:859-64.

118. Philips, R. E., Rowland-Jones, S., Nixon, D. F., Gotch, F. M., Edwards, J. P., Ogunlesi, A. O., Elvin, J. G., Rothbard, J. A., Bangham, C. R., Rizza, C. R., and Mcmichael, A. J. 1991. Human immunodeficiency virus genetic mutation that can escape cytotoxic T cell recognition. *Nature* **354**:453-9.

119. Richman, D. D., Wrin, T., Little, S. J., and Petropoulos, C.J. 2003. Rapid evolution of the neutralizing antibody response to. HIV type 1 infection. *PNAS* **100**: 4144-9.

120. Frost, S. D. W., Wrin, T., Smith, D. M., Kosakovsky Pond, S. L., Liu, Y., Paxinos, E., Chappey, C., Galovich, J., Beauchaine, J., Petropoulos, C. J., Little, S. J., Richman, D. D. 2005. Neutralizing antibody responses drive the evolution of human immunodeficiency

virus type 1 envelope during recent HIV infection. *Proc.Natl.Acad.Sci.* **102**: 18514–18519.

121. Goulder, P. J. R., Phillips, R. E., Colbert, R. A., McAdam, S., Ogg, G., Nowak, M. A., Giangrande, P., Luzzi, G., Morgan, B., Edwards, A., McMichael, A. J., and Rowland-Jones, S. 1997. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nature Med.* **3**: 212-7.

122. Moore, C. B., John, M., James, I. R., Christiansen, F. T., Witt, C. S., and Mallal, S. A. 2002. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* **296**: 1439-1443.

123. Delwart, E. L., Pan, H., Sheppard, H. W., Wolpert, D., Neumann, A. U., Korber, B., and Mullins, J. I. 1997. Slower evolution of human immunodeficiency virus type 1 quasispecies during progression to AIDS. *J Virol.* **71**:7498-508.

124. Wolinsky, S. M., Korber, B. T., Neumann, A. U., Daniels, M., Kunstman, K. J., Whetsell, A. J., Furtado, M. R., Cao, Y., Ho, D. D. and Safrit, J. T. 1996. Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection. *Science* **272**, 537-42.

125. Markham, R. B., Wang, W.-C., Weisstein, A. E., Wang, Z., Munoz, A., Templeton, A., Margolick, J., Vlahov, D., Quinn, T., Farzadegan, H., and Yu, X. F. 1998. Patterns of HIV-1 evolution in individuals with differing rates of CD4 T cell decline. *PNAS.* **95**(21): 12568 - 12573.

126. McNearney, T., Hornickova, Z., Markham, R., Birdwell, A., Arens, M., Saah, A., and Ratner, L. 1992. Relationship of human immunodeficiency virus type 1 sequence heterogeneity to stage of disease. *PNAS* **89** (21) 10247-10251.

127. Sala, M., and Wain-Hobson, S. 2000. Are RNA viruses adapting or merely changing? *J Mol Evol*. **51**(1):12-20.

128. Zarate, S., Pond, S. L. K., Shapshak, P., and Frost, S. D. W. 2007. Comparative Study of Methods for Detecting Sequence Compartmentalization in Human Immunodeficiency Virus Type 1. *J Virol.* **81**: 6643–6651.

129. Bordería, A. V., Codoñer, F. M., and Sanjuán, R. 2007. Selection promotes organ compartmentalization in HIV-1: evidence from gag and pol genes, *Evol.* **61:** 272–279.

130. Clarke, J. R., White, N. C., Weber, J. N. 2000. HIV Compartmentalization: pathogenesis and clinical implications. *AIDS Rev.* **2**:15-22.

131. Haddad, D. N., Birch, C., Middleton, T., Dwyer, D. E., Cunningham, A. L., and Saksena, N. K. 2000. Evidence for late stage compartmentalization of HIV-1 resistance

mutations between lymph node and peripheral blood mononuclear cells. *AIDS.* **14**:2273-81.

132. Diem, K., Nickle, D. C., Motoshige, A., Fox, A., Ross, S., Mullins, J. I., Corey, L., Coombs, R. W., and Krieger, J. N. 2008. Male genital tract compartmentalization of human immunodeficiency virus type 1 (HIV). *AIDS Res Hum Retrovir.* **24**: 561–571.

133. Peters, P. J., Bhattacharya, J., Hibbitts, S., Dittmar, M. T., Simmons, G., Bell, J., Simmonds, P., and Clapham, P. R. 2004. Biological analysis of human immunodeficiency virus type 1 R5 envelopes amplified from brain and lymph node tissues of AIDS patients with neuropathology reveals two distinct tropism phenotypes and identifies envelopes in the brain that confer an enhanced tropism and fusigenicity for macrophages. *J Virol.* **78**:6915–26.

134. Potter, S., Lemey, P., Achaz, G., Chew, C., Vandamme, A., Dwyer, D., and Saksena, N. 2004. HIV-1 compartmentalization in diverse leukocyte populations during antiretroviral therapy. *J leuk biol.* **76**:562-70.

135. Salemi, M., Lamers, S. L., Yu, S., de Oliveira, T., Fitch, W. M., McGrath, M. S. 2005. Phylodynamic analysis of human immunodeficiency virus type 1 in distinct brain compartments provides a model for the neuropathogenesis of AIDS. *J Virol.* **79**:11343–52.

136. Kemal, K. S., Foley, B., Burger, H., Anastos, K., Minkoff, H., Kitchen, C., Philpott, S. M., Gao, W., Robison, E., Holman, S., Dehner, C., Beck, S., Meyer III, W. A., Landay, A., Kovacs, A., Bremer, J., and Weiser. B. 2003. HIV-1 in genital tract and plasma of women: Compartmentalization of viral sequences, coreceptor usage, and glycosylation. *PNAS* **100** (22):12972-12977.

137. Schnell, G., Price, R. W., Swanstrom, R., and Spudich, S. 2010. Compartmentalization and clonal amplification of HIV-1 variants in the cerebrospinal fluid during primary infection. *J Virol* **84**:2395-407.

138. Nickle D. C., Jensen, M. A., Shriner, D., Frenkel, L. M., Mittler, J., and Mullins, J. I. 2003. Evolutionary indicators of HIV-1 reservoirs and compartments. *J Virol* **77**: 5540-6.

139. Lowe, S.H., Sankatsing, S.U.C., Repping, S., Van der Veen, F., Reiss, P., Lange J.M.A. and Prins, J.M. (2004). Is the male genital really a sanctuary site for HIV? Arguments that it is not. AIDS 18:1353-1362

140. Dickover, R. E., Garratty, E. M., Plaeger, S., and Bryson, Y. J. 2001. Perinatal transmission of major, minor, and multiple maternal human immunodeficiency virus type 1 variants in utero and intrapartum. *J Virol* **75**:2194-2203.

141. Delwart, E., Magierowska, M., Royz, M., Foley, B., Peddada, L., Smith, R., Heldebrant, C., Conrad, A., and Busch, M. (2002). Homogeneous quasispecies in 16 out of 17 individuals during very early HIV-1 primary infection. *AIDS* **16**:189-95.

142. Zhu, T., Mo, H., Wang, N., Nam, D. S., Koup, R. A., and Ho, D. D. 1993. Genotypic and phenotypic characterization of HIV-1 in patients with primary infection. *Science* **261**:1179-81.

143. Overbaugh, J., Kreiss, J., Poss, M., Lewis, P., Mostad, S., John, G., Nduati, R., Mbori-Ngacha, D., Martin, H. Jr, Richardson, B., Jackson, S., Neilson, J., Long, E. M., Panteleeff, D., Welch, M., Rakwar, J., Jackson, D., Chohan, B., Lavreys, L., Mandaliya, K., Ndinya-Achola, J., and Bwayo, J. 1999. Studies of human immunodeficiency virus type 1 mucosal viral shedding and transmission in Kenya. *J Infect Dis.* **179** Suppl 3:S401-4.

144. Blauvelt, A., Glushakova, S., and Margolis, L. B. 2000. HIV-infected human Langerhans cells transmit infection to human lymphoid tissue ex vivo. *AIDS* **14**: 647-51.

145. Huang, Y., Paxton, W. A., Wolinsky, S. M., Neumann, A. U., Zhang, L., He, T., Kang, S., Ceradini, D., Jin, Z., Yazdanbakhsh, K., Kunstman, K., Erickson, D., Dragon, E., Landau, N. R., Phair, J., Ho, D. D., and Koup, R. A. 1996. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. *Nature Medicine* **2**:1240-1243.

146. Macpherson, A. J., and Harris, N. L. 2004. Interactions between commensal intestinal bacteria and the immune system. *Nat Rev Immunol* **4**:478-85.

147. Madara, J. L., Nash, S., Moore, R., and Atisook, K. 1990. Structure and function of the intestinal epithelial barrier in health and disease. *Monogr Pathol*: 306-24.

148. Frey, A., Giannasca, K. T., Weltzin, R., Giannasca, P. J., Reggio, H., Lencer, W. I., and Neutra, M.R. 1996. Role of the glycocalyx in regulating access of microparticles to apical plasma membranes of intestinal epithelial cells: implications for microbial attachment and oral vaccine targeting. *J Exp Med* **184**:1045–59.

149. Corfield, A. P., Myerscough, N., Longman, R., Sylvester, P., Arul, S., and Pignatelli, M. 2000. Mucins and mucosal protection in the gastrointestinal tract: new prospects for mucins in the pathology of gastrointestinal disease. *Gut* **47**:589-94.

150. Mehandru, S. 2007. The gastrointestinal tract in HIV-1 infection: questions, answers, and more questions! *The PRN Notebook.* **12**:1-10.

151. Boismenu, R., and Havran, W. L. 2001. Intraepithelial gamma delta T cells exposed by functional genomics. *Genome Biol* **2** (11): 1031-1031.4.

152. Crottet, P., and Corthésy, B. 1998. Secretory Component Delays the Conversion of Secretory IgA into Antigen-Binding Competent F(ab')2: A Possible Implication for Mucosal Defense. *J Immunol.* **161**:5445-5453.

153. Groux, H., O'Garra, A., Bigler, M., Rouleau, M., Antonenko, S., de-Vries, J. E., Roncarolo, M. G. 1997. A CD4+ T-cell subset inhibits antigen-specific T-cell responses and prevents colitis. *Nature* **389**: 737-42.

154. Platt, A. M., and Mowat, A. M. 2008. Mucosal macrophages and the regulation of immune responses in the intestine. *Immunol Lett.* **119**:22-31.

155. Cheroutre, H., and Madakamutil, L. 2004. Acquired and natural memory T cells join forces at the mucosal front line. *Nat Rev Immunol.* **4**:290–300.

156. Berlin, C., Berg, E. L., Briskin, M. J., Andrew, D. P., Kilshaw, P. J., Holzmann, B. Weissman, I. L., Hamann, A., and Butcher, E. C. 1993. Alpha 4 βeta 7 integrin mediate lymphocyte binding to the mucosal vascular adressin Mad-CAM1. *Cell* **74**:185-95.

157. Cepek, K. L., Shaw, S. K., Parker, C. M., Russell, G. J., Morrow, J. S., Rimm, D. L., and Brenner, M. B. 1994. Adhesion between epithelial cells and T lymphocytes mediated by E-cadherin and the αEβ7 integrin. *Nature* **372**:190-193.

158. Briskin, M., Winsor-Hines, D., Shyjan, A., Cochran, N., Bloom, S., Wilson, J., McEvoy, L. M., Buitcher, E. C., Kassam, N., Mackay, C. R., Newman, W., and Ringler, D. J. 1997. Human Mucosal addressin cell adhesion molecule-1 is preferentially expressed in intestinal tract and associated lymphoid tissue. *Am J Pathol.* **151**:97-100.

159. Bull, D. M., and Bookman, M. A. 1997. Isolation and functional characterization of human intestinal mucosal lymphoid cells. *J Clin Invest.* **59**(5):966–974.

160. Schenk, M., and Mueller, C. 2007. Adaptations of intestinal macrophages to an antigen-rich environment [Review]. *Semin Immunol* **19**:84-93.

161. Smythies L. E., Sellers M., Clements R. H., Mosteller-Barnum M., Meng G., Benjamin W. H., Orenstein J. M., and Smith P. D. 2005. Human intestinal macrophages display profound inflammatory anergy despite avid phagocytic and bacteriocidal activity. *J. Clin. Invest.* **115**: 66–75.

162. Cassol, E., Malfeld, S., Mahasha, P., van der Merwe, S., Cassol, S., Seebregts, C., Alfano, M., Poli, G., and Rossouw, T. 2010. Persistent Microbial Translocation and Immune Activation in HIV-1 Infected South Africans Receiving Combination Antiretroviral Therapy. *J infec Dis.* **202**: 723-733.

163. Cook, D. N., Presser, D. M., Forster, R., Zhang, J., Kuklin, N. A., Abbondanzo, S. J., Niu, X. D., Chen, S. C., Manfra, D. J., Wiekowski, M. T., Sullivan, L. M., Smith, S. R.,

Greenberg, H. B., Narula, S. K., Lipp, M., and Lira, S. A. 2000. CCR6 mediates dendritic cell localization, lymphocyte homeostasis, and immune responses in mucosal tissue. *Immunity* **12**:495-503.

164. Huang, F. P., Platt, N., Wykes, M., Major, J. R., Powell, T. J., Jenkins, C. D., and MacPherson, G. G. 2000. A discrete subpopulation of dendritic cells transports apoptotic intestinal epithelial cell. *J Exp Med* **191**: 435-44.

165. MacDonald T.T., and Spencer J. (1994) In *Gastrointestinal and Hepatic Immunology*, R. H. Heatley, Ed. (Cambridge Univ. Press, Cambridge) pp. 1–23.

166. Veazey R.S., Rosenzweig M., Shvetz D.E., Pauley D.R., DeMaria M., Chalifoux L.V., Johnson R.P., and Lackner A.A. (1997). Characterization of gut-associated lymphoid tissues (GALT) of normal rhesus macaques. *Clin. Immunol. Immunopathol*. **82**:230-242.

167. Douek, D. 2007. HIV disease progression: immune activation, microbes, and a leaky gut. *Topics in HIV Med* **15**(4):114-7.

168. Li, Q., Duan, L.,. Estes, J. D., Ma, Z. M., Rourke, T., Wang, Y., Reilly, C., Carlis, J., Miller, C. J., and Haase. A. T. 2005. Peak SIV replication in resting memory CD4+ T cells depletes gut lamina propria CD4+ T cells. *Nature* **434**:1148-1152.

169. Miller, C. J., Li, Q., Abel, K., Kim, E. Y., Ma, Z. M., Wietgrefe, S., La Franco-Scheuch, L., Compton, L., Duan, L., Shore, M. D., Zupancic, M., Busch, M., Carlis, J., Wolinsky, S., and Haase A. T. 2005. Propagation and dissemination of infection after vaginal transmission of simian immunodeficiency virus. *J Virol.* **79**:9217-9227.

170. Guadalupe, M., Reay, E., Sankaran, S., Prindiville, T., Flamm, J., Mcneil, A., and Dandekar, S. 2003. Severe CD4+ T-cell depletion in gut lymphoid tissue during primary human immunodeficiency virus type 1 infection and substantial delay in restoration following highly active antiretroviral therapy. *J Virol*. 77:11708-17.

171. Mattapallil, J. J., Douek, D. C., Hill, B., Nishimura, Y., Martin, M., and Roederer. M. 2005. Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection. *Nature* **434**:1093-1097.

172. Mehandru, S., Poles M. A., Tenner-Racz, K., Horowitz, A., Hurley, A., Hogan, C., Boden, D., Racz, P., and Markowitz, M. 2004. Primary HIV-1 infection is associated with preferential depletion of CD4+ T lymphocytes from effector sites in the gastrointestinal tract. *J Exp Med*. 200:761-770.

173. Veazey, R. S., DeMaria, M., Chalifoux, L. V., Shvetz, D. E., Pauley, D. R., Knight, H. L., Rosenzweig, M., Johnson, R.P., Desrosiers, R.C., and Lackner, A. A. 1998.

Gastrointestinal tract as a major site of CD4+ T cell depletion and viral replication in SIV infection. *Science* **280**:427-431.

174. Brenchley, J. M., Schacker, T. W., Ruff, L. E., Price, D. A., Taylor, J. H., Beilman, G. J., Nguyen, P. L., Khoruts, A., Larson, M., Haase, A. T., and Douek, D. C. 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *J. Exp. Med.* **200**:749-759.

175. Brenchley, J. M., Price, D. A., and Douek, D. C. 2006. HIV disease: fallout from a mucosal catastrophe? *Nat Immunol* **7**:235-9.

176. Grossman, Z., Meier-Schellersheim, M., Paul, W. E., and Picker, L. J. 2006. Pathogenesis of HIV infection: what the virus spares is as important as what it destroys. *Nat Med.* **12**(3):289-95.

177. Paiardini, M., Frank, I., Pandrea, I., Apetrei, C., and Silvestri, G. 2008. Mucosal immune dysfunction in AIDS pathogenesis. *AIDS Rev*. **10**(1):36-46.

178. Benito J. M. 2008. New Insights into the Mechanisms Driving Immune Activation in HIV. *AIDS Rev* **10**:254-5.

179. Brenchley, J. M., Price, D. A., Schacker, T. W., Asher, T. E., Silvestri, G., Rao, S., Kazzaz, Z., Bornstein, E., Lambotte, O., Altmann, D., Blazar, B. R., Rodriguez, B., Teixeira-Johnson, L., Landay, A., Martin, J. N., Hecht, F. M., Picker, L. J., Lederman, M. M., Deeks, S. G., and Douek, D. C. 2006. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat Med.* **12**(12):1365-71.

180. Estes, J. D., Harris, L. D., Klatt, N. R., Tabb, B., Pittaluga, S., Paiardini, M., Barclay, G. R., Smedley, J., Pung, R., Oliveira, K. M., Hirsch, V. M., Silvestri, G., Douek, D. C., Miller, C. J., Haase, A. T. Lifson, J., and Brenchley, J. M. 2010. Damaged Intestinal Epithelial Integrity Linked to Microbial Translocation in Pathogenic Simian Immunodeficiency Virus Infection. *PloS Pathogens* **6 (8):** e1001052. doi:10.1371/journal.ppat.1001052.

181. Brenchley, J. M., and Douek, D. C. 2008. HIV infection and the gastrointestinal immune system. *Mucosal Immunol***. 1**:23-30.

182. Wallet, M. A., Rodriguez, C. A., Yin, L., Saporta, S., Chinratanapisit, S., Hou, W., Sleasman, J. W., and Goodenow, M. M. 2010. Microbial translocation induces persistent macrophage activation unrelated to HIV-1 levels or T-cell activation following therapy. *AIDS* **24**:1281-90.

183. Hazenberg, M. D., Otto, S. A., van Benthem, B. H. B., Roos, M. T. L., Coutinho, R. A., Lange, J. M. A., Hamann, D., Hamann, D., Prins, M., and Miedema, F. 2003.

Persistent immune activation in HIV-1 infection is associated with progression to AIDS. *Cell*, ***17***(13): 1881-8.

184. Pope, M., and Haase, A. T. 2003. Transmission, acute HIV-1 infection and the quest for strategies to prevent infection. *Nature Medicine*. **9**:847-52.

185. Haase, A.T. 2005. Perils at mucosal front lines for HIV and SIV and their hosts. *Nat Rev Immunol*. **5**: 783-92.

186. Zeitz, M., Greene, W. C., Peffer, N. J., and James, S. P. 1988. Lymphocytes isolated from the intestinal lamina propria of normal nonhuman primates have increased expression of genes associated with T-cell activation. *Gastroenterol*. **94**:647–655.

187. Borrow, P., Lewicki, H., Wei, X., Horwitz, M. S., Peffer, N., Meyers, H., Nelson, J.A., Gairin, J.E., Hahn, B. H., Oldstone, M. B., and Shaw, G. M. 1997. Anti-viral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. *Nat Med* **3:** 205-11.

188. Allen, T. M., O'Connor, D. H., Jing, P., Dzuris, J. L., Mothe, B. R., Vogel, T. U., Dunphy, E., Liebl, M. E., Emerson, C., Wilson, N., Kunstman, K. J., Wang, X., Allison, D. B., Hughes, A. L., Desrosiers, R. C., Altman, J. D., Wolinsky, S. M., Sette, A., and Watkins, D. I. 2000. Tat-specific cytotoxic T lymphocytes select for SIV escape variants during resolution of primary viraemia. *Nature* **407**:386—90.

189. Chun, T. W., Stuyver, L., Mizell, S. B., Ehler, L. A., Mican, J. A., Baseler, M., Lloyd, A. L., Nowak, M. A., and Fauci, A. S.1997. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *PNAS USA*. **94**(24): 13193-7.

190. Persaud, D., Zhou, Y., Siliciano, J. M., and Siliciano, R. F. 2003. Latency in human immunodeficiency virus type 1 infection: no easy answers. *J Virol*. **77**: 1659-65.

191. Li, L., and Smith, P. D. 1996. Immunobiology of mucosal HIV-1 infection. *Curr Opin. Gastroenterol*. **12**:560-3.

192. Page-Shafer, K., Shiboski, C. H., Osmond, D. H., *et al*. 2002. Risk of HIV infection attributable to oral sex among men who have sex with men and in the population of men who have sex with men. *AIDS* **16**:2350-2.

193. Vittinghoff, E., Douglas, J., Judson, F., McKirnan, D., MacQueen, K., and Buchbinder, S. P. 1999. Per-contact risk of human immunodeficiency virus transmission between male sexual partners. *Am J Epidemiol*. **150**:306-11.

194. Quinn, T. C., Wawer, M. J., Sewankambo, N., Serwadda, D., Li, C., Wabwire-Mangen, F., Meehan, M. O., Lutalo, T.L., and Gray, R.H.. 2000. Viral load and

heterosexual transmission of human immunodeficiency virus type 1. *N Engl J Med* **342**:921-9.

195. Lawn, S. D., Subbarao, S., Wright, T. C., Evans-Strickfaden, T., Ellerbrock, T. V., Lennox, J. L., Butera, S. T., and Hart, C. E. 2000. Correlation between HIV-1 RNA levels in the female genital tract and immune activation resulting from ulceration of the cervix. *J Infect Dis* **181**:1950-6.

196. Telzak, E. E., Chiasson, M. A., Bevier, P. J., Stoneburner, R. L., Castro, K. G., and Jaffe, H. W. 1993. HIV-1 seroconversion in patients with and without genital ulcer disease. *Ann Intern Med* **119**:1181-6.

197. Van de Perre, P., Simonon, A., Karita, E., Butera, J. B., Hitimana, D. G., Mukamabano, B., Van Goethem, C., Lepage, P., Dabis, F., and Msellati, P. 1993. Infective and anti-infective properties of breastmilk from HIV-1-infected women. *Lancet* **341**: 914-8.

198. John, G., Nduati, R.W., Mbori-Ngacha, D.A., Richardson, B.A., Panteleeff, D.D., Mwatha, A., Overbaugh, J., Bwayo, J., Ndinya-Achola, J. O., and Kreiss, J. K. 2001. Correlates of mother-to-child transmission of HIV-1 in the Kenyan breastfeeding clinical trial. *J Infect Dis.* **183**:206-12.

199. Kotler, D.P. 2005. HIV infection and the gastrointestinal tract. *AIDS* **19**:107-17.

200. Berger, E. A., Murphy, I. M., and Farber, J. M. 1999. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annu Rev Immunol* **17**:657-700.

201. Connor, R. I., Sheridan, K. E., Ceradini, D., Choe, S., and Landau. N. R. 1997. Change in Coreceptor Use Correlates with Disease Progression in HIV-1–Infected Individuals. *J. Exp. Med.* **185**(4): 621–628.

202. Hu, Q., Barry, A. P., Wang, Z., Connolly, S. M., Peiper, S. C., and Greenberg, M. L. 2000. Evolution of the human immunodeficiency virus type 1 envelope during infection reveals molecular corollaries of specificity for coreceptor utilization and AIDS pathogenesis. *J Virol.* **74**:11858–11872.

203. Bjorndal, A., Deng, H., Jansson, M., Fiore, J.R., Colognesi, C., Karlsson, A., Albert, J., Scarlatti, G., Littman, D. R., and Fenyo. E. M. 1997. Coreceptor usage of primary human immunodeficiency virus type 1 isolates varies according to biological phenotype. *J. Virol.* **71**:7478-87.

204. De Roda Husman, A. M., and Schuitemaker, H. 1998. Chemokine receptors and the clinical course of HIV-1 infection. *Trends Microbiol* **6**:244-9.

205. Naif, H. M., Cunningham, A. L., Alali, M., Li, S., Nasr, N., Buhler, M. M., Schols, D., Clercq, E., and Stewart, G. 2002. A human immunodeficiency virus type 1 isolate froman infected person homozygous for CCR5Δ32 exhibits dual tropism by infecting macrophages and MT2 cells via CXCR4. *J Virol* **76**:3114-3124.

206. Neil, S., Martin, F., Ikeda, Y., and Collins, M. 2001. Postentry restriction to human immunodeficiency virus-based vector transduction in human monocytes. *J. Virol.* **75**:5448-56.

207. Naif, H. M., Li, S., Alali, M., Chang, J., Mayne, C., Sullivan, J., and Cunningham, A. L. 1999. Definition of the stage of host cell genetic restriction of replication of human immunodeficiency virus type 1 in monocytes and monocyte-derived macrophages by using twins. *J Virol.* **73**:4866-81.

208. Eisert, V., Kreutz, M., Becker, K., Königs, C., Alex, U., Rübsamen-Waigmann, H., Andreesen, R., and von Briesen, H. 2001. Analysis of cellular factors influencing the replication of human immunodeficiency virus type I in human macrophages derived from blood of different healthy donors. *Virology*. **286**:31-44.

209. Al-Mulla, W., Church, D., and Gill, M. J. 1997. Phenotypic variations and switches in HIV isolated from the blood and gastrointestinal tissues of patients with HIV-1 infection. *J Med Virol.* **52**:31–34.

210. van Marle, G., Gill, M. J., Kolodka, D., McManus, L., Grant, T., and Church, D. L. 2007. Compartmentalization of the gut viral reservoir in HIV-1 infected patients. *Retrovirol.* **4**:87.

211. van Marle, G., Church, D.L., Nunweiler, K.D., Cannon, K., Wainberg, M.A., and Gill, M.J. 2010. Higher levels of Zidovudine resistant HIV in the colon compared to blood and other gastrointestinal compartments in HIV infection. Retrovirology **7**:74

212. Avettand-Fenoel, V., Hocqueloux, L., Müller-Trutwin, M., Prazuck, T., Melard, A., Chaix, Marie-Laure., Agoute, E., Michau, C., and Rouzioux, C. 2011. Greater Diversity of HIV DNA Variants in the Rectum Compared to Variants in the Blood in Patients Without HAART. J Med Virol **83**:1499–1507.

213. van der Hoek, L., Goudsmit, J., Maas, J., Sol, C.J. 1998. Human immunodeficiency virus type 1 in faeces and serum: evidence against independently evolving subpopulations. J Gen Virol **79**:2455–2459.

214. Imamichi, H., DeGray, G., Dewar, R.L., Mannon, P., Yao, M., Chairez, C., Sereti, I., and Kovacs, J.A. 2011. Lack of compartmentalization of HIV-1 quasispecies between the Gut and peripheral blood compartments. *J infec Dis* **204**:309-14.

215. Evering, T.H., Mehandru, S., Racz, P., Tenner-Racz, K., Poles, M.A., Figueroa, A., Mohri1, H., and Markowitz, M. 2012. Absence of HIV-1 evolution in the Gut-associated lymphoid tissue from patients on combination antiviral therapy initiated during primary infection. *PLoS Pathogens* **8:**2

216. Poles, M. A., Boscardin, W., Elliott, J., Taing, P., Fuerst, M. M. P., Mcgowan, I., Brown, S., and Anton, P.A. 2006. Lack of Decay of HIV-1 in Gut-Associated Lymphoid Tissue Reservoirs in Maximally Suppressed Individuals. *Aquir. Immune Defic. Syndr.* **43**(1), 65-68.

217. Mehandru, S., Poles, M.A., Tenner-racz, K., Jean-pierre, P., Manuelli, V., Lopez, P., Shet, A., Low, A., Hiroshi, M., Boden, D., Racz, P., Markowitz, M. 2006. Lack of mucosal immune reconstitution during prolonged treatment of acute and early HIV-1 infection. *PLoS Med.* **3**(12):2335-2348.

218. George, M.D., Reay, E., Sankaran, S., and Dandekar, S. 2005. Early antiretroviral therapy for simian immunodeficiency virus infection leads to mucosal $CD4^+$ T-cell restoration and enhanced gene expression regulating mucosal repair and regeneration. *J Virol* **79**: 2709–2719.

219. Guadalupe, M., Sankaran, S., George, M.D., Reay, E., Verhoeven, D., Shacklett, B. L., Flamm, J., Wegelin, J., Prindiville, T., and Dandekar, S. 2006. Viral Suppression and Immune Restoration in the Gastrointestinal Mucosa of Human Immunodeficiency Virus Type 1-Infected Patients Initiating Therapy during Primary or Chronic Infection. *J Virol* **80**:8236-8247.

220. Torre, D., Speranza, F., and Martegani, R. 2005. Impact of highly active antiretroviral therapy on organ-specific manifestations of HIV-1 infection. *HIV Med.* **6**:66-78.

221. Kotler, D.P., Shimada, T., Snow, G. Winson, G., Chen, W., Zhao, M., Inada, Y., and Clayton, F. 1998. Effect of combination antiretroviral therapy upon rectal mucosal HIV RNA burden and mononuclear cell apoptosis. *AIDS* **12**:597-604.

222. Talal, A.H., Monard, S., Vesanen, M., Zheng, Z., Hurley, A., Cao, Y., Fang, F., Smiley, L., Johnson, T., Kost, R., and Markowitz, M.H. 2001. Virologic and Immunologic effect of antiretroviral theray on HIV-1 in Gut-associated lymphoid tissue. *AIDS* **26**:1-7.

223. Sheth, P.M., Chege, D., Shin, L.Y.Y., Huibner, S., Yue, F-Y., Loutfy, M., Halpenny, R., Presad, D., Kovacs, C., Chun, T-W., Kandel, G., Ostrowski, M., and Kaul, R. 2008. Immune reconstitution in the sigmoid colon after long-term HIV therapy. *Mucosal Immunol.* **1**(5):382-8.

224. Ciccone, E.J., Read, S.W., Mannon, P.J., Yao, M.D., Hodge, J.N., Dewar, R., *et al*. 2010. Cycling of gut mucosal CD4R T cells decreases after prolonged antiretroviral therapy and is associated with plasma LPS levels. *Mucosal Immunol.* **3**:172–181.

225. Baker J., Larson, M., Khoruts, A., Haase, A., and Schacker, T. 2006. $CD4^+$ T cell quantification within gut lamina propria of HIV patients after HAART. Abstract #41, 13th Conference on Retroviruses and Opportunistic Infection (CROI) Feb 5-8th, Denver, Colorado.

226. Poles, M.A., Elliott, J., Vingerhoets, J., Michiels, L., Scholliers, A., Bloor, S., Larder, B., Hertogs, K., and Anton, P.A. 2001. Despite high concordance, distinct mutational and phenotypic drug resistance profiles in human immunodeficiency virus type 1 RNA are observed in gastrointestinal mucosal biopsy specimens and peripheral blood mononuclear cells compared with plasma. *J infect. Dis.* **183**(1), 143-8.

227. Appay, V., and Sauce, D. 2008. Immune activation and inflammation in HIV-1 infection: causes and consequences. *J Pathol* **214**:231-241.

228. Brandtzaeg, P. 1989. Overview of the mucosal immune system. *Curr Top Microbiol Immunol* **146**:13-25.

229. Macdonald, T.T., and Monteleone, G. 2005. Immunity, inflammation, and allergy in the gut. *Science* **307**:1920-1925.

230. Berg, R.D. 1995. Bacterial translocation from the gastrointestinal tract. *Trends Microbiol* **3**:149-154.

231. Husby, S., Jensenius, J.C. and Svehag, S.E. 1985. Passage of undegraded dietary antigen into the blood of healthy adults. Quantification, estimation of size distribution, and relation of uptake to levels of specific antibodies. *Scand J Immunol* **22**:83-92.

232. Neutra, M.R., Mantis, N.J., and. Kraehenbuhl, J.P. 2001. Collaboration of epithelial cells with organized mucosal lymphoid tissues. *Nat Immunol* **2**:1004-1009.

233. Amerongen, H.M., Weltzin, R., Farnet, C.M., Michetti, P., Haseltine, W.A., and Neutra, M.R. 1991. Transepithelial transport of HIV-1 by intestinal M cells: a mechanism for transmission of AIDS. *J Acquir Immune Defic Syndr* **4**:760-765.

234. Fotopoulos, G., Harari, A., Michetti, P., Trono, D., Pantaleo, G., and Kraehenbuhl, J.P. 2002. Transepithelial transport of HIV-1 by M cells is receptormediated. *Proc Natl Acad Sci USA* **99**:9410-9414.

235. Rescigno, M., Urbano, M., Valzasina, B., Francolini, M., Rotta, G., Bonasio, R., Granucci, F., Kraehenbuhl, J.P., and Ricciardi-Castagnoli, P. 2001. Dendritic cells express

tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nat Immunol* **2**:361-367.

236. Niess, J.H., Brand, S., Gu, X., Landsman, L., Jung, S., McCormick, B.A., Vyas, J.M., Boes, M., Ploegh, H.L., Fox, J.G., Littman, D.R., and Reinecker, H.C. 2005. CX3CR1-mediated dendritic cell access to the intestinal lumen and bacterial clearance. *Science* **307**:254-258.

237. Hayday, A., Theodoridis, E., Ramsburg, E., and Shires J. 2001. Intraepithelial lymphocytes: exploring the Third Way in immunology. *Nat Immunol* **2**:997-1003.

238. Brandtzaeg, P., and Pabst, R. 2004. Let's go mucosal: communication on slippery ground. *Trends Immunol* **25**:570-577.

239. Cheroutre, H. 2004. Starting at the beginning: new perspectives on the biology of mucosal T cells. *Annu Rev Immunol* **22**:217-246.

240. Smith, P.D., Ochsenbauer-Jambor, C., and Smythies, L.E. 2005. Intestinal macrophages: unique effector cells of the innate immune system. *Immunol Rev* **206**:149-159.

241. Haller, D., Holt, L., Kim, S.C., Schwabe, R.F., Sartor, R.B., and Jobin, C. 2003. Transforming growth factor-beta 1 inhibits non-pathogenic Gram negative bacteria-induced NF-kappa B recruitment to the interleukin-6 gene promoter in intestinal epithelial cells through modulation of histone acetylation. *J Biol Chem* **278**:23851-23860.

242. Akhtar, M., Watson, J.L., Nazli, A., and McKay, D.M. 2003. Bacterial DNA evokes epithelial IL-8 production by a MAPK-dependent, NF-kappaBindependent pathway. *FASEB J* **17**:1319-1321.

243. Podolsky, D. K. 2002. Inflammatory bowel disease. *N Engl J Med* **347**:417- 429.

244. Simon, D., and Brandt, L.J. 1993. Diarrhea in patients with the acquired immunodeficiency syndrome. *Gastroenterology* **105**:1238-1242.

245. Batman, P.A., Miller, A.R., Forster, S.M., Harris, J.R., Pinching, A.J., and Griffin, G.E. 1989. Jejunal enteropathy associated with human immunodeficiency virus infection: quantitative histology. *J Clin Pathol* **42**:275-281.

246. Batman, P. A., Kotler, D.P., Kapembwa, M.S., Booth, D., Potten, C.S., Orenstein, J.M., Scally, A.J., and Griffin, G.E. 2007. HIV enteropathy: crypt stem and transit cell hyperproliferation induces villous atrophy in HIV/Microsporidia-infected jejunal mucosa. *AIDS* **21**:433-439.

247. Heise, C., Dandekar, S., Kumar, P., Duplantier, R., Donovan, R.M., and Halsted, C.H. 1991. Human immunodeficiency virus infection of enterocytes and mononuclear cells in human jejunal mucosa. *Gastroenterology* **100**:1521-1527.

248. Stockmann, M., Fromm, M., Schmitz, H., Schmidt, W., Riecken, E.O., and Schulzke, J.D. 1998. Duodenal biopsies of HIV-infected patients with diarrhoea exhibit epithelial barrier defects but no active secretion. *AIDS* **12**:43-51.

249. Clayton, F., Kapetanovic, S., and Kotler, D.P. 2001. Enteric microtubule depolymerization in HIV infection: a possible cause of HIV-associated enteropathy. *AIDS* **15**:123-124.

250. Griffin, G.E., Miller, A., Batman, P., Forster, S.M., Pinching, A.J., Harris, J.R., and Mathan, M.M. 1988. Damage to jejunal intrinsic autonomic nerves in HIV infection. *AIDS* **2**:379-382.

251. Mathan, M.M., Griffin, G.E., Miller, A., Batman, P., Forster, S., Pinching, A., and Harris, W. 1990. Ultrastructure of the jejunal mucosa in human immunodeficiency virus infection. *J Pathol* **161**:119-127.

252. Bode, H., Schmidt, W., Schulzke, J.D., Fromm, M., Zippel, T., Wahnschaffe, U., Bendfeldt, K., Riecken, E.O., and Ullrich, R. 1999. The HIV protease inhibitors saquinavir, ritonavir, and nelfinavir but not indinavir impair the epithelial barrier in the human intestinal cell line HT-29/B6. *AIDS* **13**:2595-2597.

253. Caradonna, L., Amati, L., Magrone, T., Pellegrino, N.M., Jirillo, E., and Caccavo, D. 2000. Enteric bacteria, lipopolysaccharides and related cytokines in inflammatory bowel disease: biological and clinical significance. *J Endotoxin Res* **6**:205-214.

254. Cooke, K.R., Olkiewicz, K., Erickson, N., and Ferrara, J.L. 2002. The role of endotoxin and the innate immune response in the pathophysiology of acute graft versus host disease. *J Endotoxin Res* **8**:441-448.

255. Takeda, K., Kaisho, T., and Akira, S. 2003. Toll-like receptors. *Annu Rev Immunol* **21**:335-376.

256. Wyatt, J., Vogelsang, H., Hubl, W., Waldhoer, T., and Lochs, H. 1993. Intestinal permeability and the prediction of relapse in Crohn's disease. *Lancet* **341**:1437-1439.

257. Schietroma, M., Cappelli, S., Carlei, F., Di Giuro, G., Agnifili, A., Recchia, C.L., Antonellis, M., and Amicucci, G. 2006. [Intestinal and systemic endotoxaemia after laparotomic or laparoscopic cholecystectomy]. *Chir Ital* **58**:171-177.

258. Giorgi, J.V., Hultin, L.E., McKeating, J.A., Johnson, T.D., Owens, B., Jacobson, L.P., Shih, R., Lewis, J., Wiley, D.J., Phair, J.P., Wolinsky, S.M., and Detels, R. 1999. Shorter

survival in advanced human immunodeficiency virus type 1 infection is more closely associated with T lymphocyte activation than with plasma virus burden or virus chemokine coreceptor usage. *J Infect Dis* **179**:859-870.

259. Lane, H.C., Masur, H., Edgar, L.C., Whalen, G., Rook, A.H., and Fauci, A.S. 1983. Abnormalities of B-cell activation and immunoregulation in patients with the acquired immunodeficiency syndrome. *N Engl J Med* **309**:453-458.

260. Hellerstein, M., Hanley, M.B., Cesar, D., Siler, S., Papageorgopoulos, C., Wieder, E., Schmidt, D., Hoh, R., Neese, R., Macallan, D., Deeks, S., and McCune, J.M. 1999. Directly measured kinetics of circulating T lymphocytes in normal and HIV-1-infected humans. *Nat Med* **5**:83-89.

261. Hazenberg, M.D., Stuart, J.W., Otto, S.A., Borleffs, J.C., Boucher, C.A., de Boer, R.J., Miedema, F., and Hamann, D. 2000. T-cell division in human immunodeficiency virus (HIV)-1 infection is mainly due to immune activation: a longitudinal analysis in patients before and during highly active antiretroviral therapy (HAART). *Blood* **95**:249-255.

262. Valdez, H., and Lederman, M.M. 1997. Cytokines and cytokine therapies in HIV infection. *AIDS Clin Rev*: 187-228.

263. Walensky, R. P., Paltiel, A. D., Losina, E., Mercincavage, L. M., Schackman, B. R., Sax, P. E., Weinstein, M. C., and Freedberg, K. A. 2006. The survival benefits of AIDS treatment in the United States. *J Infect Dis.* **194**:11-9.

264. Mofenson, L. M. Committee on Pediatric AIDS. 2000. Technical report: perinatal human immunodeficiency virus testing and prevention of transmission. *Pediatrics.* **106**:E88.

265. Vermund, S. H. 2006. Millions of life-years saved with potent antiretroviral drugs in the United States: a celebration, with challenges. *J Infect Dis*. **194**:1-5.

266. Valdiserri, R. O., Ogden, L. L., and McCray, E. 2003. Accomplishments in HIV prevention science: Implications for stemming the epidemic. *Nat Med* **9**: 881-6.

267. Clerici, M., Butto,' S., Saresella, M., Lukwiya, M., Biasin, M., Pastori, C., Declich, S., Trabattoni, D., Fracasso, C., Rizzardini, G., Piconi, S., Ferrante, P., and Lopalco, L. 2000. Immune activation in Africa is enviromentally-driven and is associated with upregulation of CCR5. *AIDS* **14**: 2083-92.

268. Cassol, E., Page, T., Mosam, A., Friedland, G., Jack, C., Lalloo, U., Kopetka, J., Patterson, B., Esterhuizen, T., and Coovadia, H. M. 2005. Therapeutic response of HIV-1

subtype C in African patients co-infected with either Mycobacterium tuberculosis or humanherpesvirus-8. *J Infect Dis.* **191**:324-32.

269. Anonymous. 2006. Antiretroviral therapy for HIV infection in adults and adolescents: recommendations for a public health approach. WHO, ed, Geneva.

270. Darwin, C. 1859. On the origin of species, ed. L.J. Murray.

271. Dowell, K. 2008. Molecular phylogenetics: An introduction to computational methods and tools for analyzing evolutionary relationships. *Math* **500**:1-18.

272. Edwards, A.W.F., and Cavalli-Sforza, L.L. 1963. Reconstruction of evolution. Annals of *Human Genet*. **27**:105 – 106.

273. Sokal, R. and Michener, C. 1958. A statistical method for evaluating systematic relationships. University of Kansas Science Bulletin. **38**:1409 – 1438.

274. Murtagh F. 1984. Complexities of Hierarchic Clustering Algorithms: the state of the art. *Comput. Stats Quart.* **1**:101–113.

275. Edwards, A.W.F., and Cavalli-Sforza, L.L. 1964. Reconstruction of evolutionary trees. Phenetic and Phylogenetic Classification. Heywood, V.H. and McNeil, J. ed. *System Assoc. Publ*. No.6. London.

276. Fitch, W.M. and Margoliash, E. 1967. Construction of phylogenetic trees. *Sci*. **155**(3760): 279 –284.

277. Saitou, N. and Nei, M. 1987. The Neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**(4):406 - 425.

278. Rannala, B. and Yang, Z. 1996. Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *J. Mol. Evol.* **43**:304 – 311.

279. Yang, Z. and Rannala, B. 1997. Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo Method. *Mol. Biol. Evol.* **14**:717 - 724.

280. Mau, B. and Newton, M.A. 1997. Phylogenetic inference for binary data on dendrograms using Markov chain Monte Carlo. *J.Comp. Graph Stats.* **6**:122 – 131.

281. Li S, Pearl D, and Doss H. 2000. Phylogenetic tree reconstruction using Markov chain Monte Carlo. *J. Amer. Stats. Ass*. **95**:493 – 508.

282. Yang, Z. And Rannala, B. 2012. Molecular phylogenetics: principles and practice. *Nat. Rev. Genet.* **13**:303-314.

283. Lemey, P., Salemi, M. and Vandamme, A-M. (eds). 2009. The phylogenetic handbook: A practical approach to phylogenetic analysis and hypothesis testing, Published by Cambridge University Press. © Cambridge University Press.

284. Jukes, T. and Cantor, C.R. 1969. Evolution of protein molecules. In: Mammalian protein Metabolism, edited by Munro HN, New York, Academic Press. 21-132.

285. Salemi M and Vandamme A-M. 2003. The Phylogenetic Handbook: A practical approach to DNA and protein phylogeny. Cambridge University Press. Cambridge, United Kingdom.

286. Hall, B.G. 2008. Phylogenetic Trees Made Easy: A How-To Manual. 3de Edition. Sunderland Associates, Inc. Massachusetts, United States of America.

287. Holder, M. and Lewis, P.O. 2003. Phylogeny Estimation: Traditional and Bayesian approaches. Nature Reviews: Genetics. **4**:275 – 284.

288. Fitch, W.M. 1971. Towards defining the course of evolution: Minimum change for a specific tree topology. *Syst. Zool.* **20**:406-416.

289. Rzhetsky, A. and Nei, M. 1992. A simple method for estimating and testing minimum-evolution trees. *Mol. Biol. Evol.* **9**:945-967.

290. Hartingan, J. 1973. Minimum evolution fits to a given tree. *Biomet*. **29**:53-65.

291. Tamura K., Peterson D., Peterson N., Stecher G., Nei M., and Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol. Evol* **28**:2731-2739.

292. Swofford, D.L 2000. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, PA.

293. Goloboff, P. Farris, J. And Nixon, K. 2008. TNT. Willi Henning Society. Retrieved May 2008, from http://www.zmuc.dk/public/phylogeny/TNT/.

294. Felsentein, J. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* **27**:401-410.

295. Fisher, R.A. 1922. On the mathematical foundations of theoretical statistics. *Philos. Trans. Roy. Soc. London Ser.* A **222**:309-368.

296. Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**:368 – 376.

297. Adachi, J. and Hasegawa, M. 1996. MOLPHY, Pfrograms for molecular phylogenetics based on maximum likelihood, (ed. Tokyo: Institute of Statistical Mathematics).

298. Felsenstein, J., *PHYLIP (Phylogeny Inference Package)*. 2005, Department of Genome Sciences, University of Washington, Seattle.

299. Guindon, S. and Gascuel, O. 2003. PhyML original version: *"A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.". Syst. Biol.* **52**(5):696-704. Was improved by Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol*. 2010. **59**(3):307 - 321.

300. Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinfo.* **22**(21):2688-90.

301. Zwickl, D.J. 2006**.** Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.

302. Suchard, M.A., Weiss, R.E., Sinsheimer, J.S. 2001. Bayesian Selection of Continuous-Time Markov Chain Evolutionary Models. *Mol. Biol. Evol.* **18**(6):1001 - 1013.

303. Yang, Y. 2008. Computational Molecular Evolution. Oxford Series in Ecology and Evolution. Oxford University Press.

304. Kidd, K.K. and Sgaramella-Zonta, L.A. 1971. Phylogenetic analysis: Concepts and methods. *Am. J. Human Genet*. **23**: 235-253.

305. Studier, J.A. and Keppler, K.J. 1998. A note on the neighbor-joining method of Saitou and Nei. *Mol. Biol. Evol.* **5**:729-31.

306. Gascuel, O. and Steel, M. 2006. Neighbor-joining revealed. *Mol. Biol. Evol.* **23**(11):1997 – 2000.

307. Chase A., Zhou Y., and Siliciano R.F. (2006). HIV-1-induced depletion of CD4+ T cells in the gut: mechanism and therapeutic implications. *Trends Pharmacol. Sci.* **27** (1):4-7.

308. Centlivre M., Sala M., Wain-Hobson S., and Berkhout B. (2007). In HIV-1 pathogenesis the die is cast during primary infection. *AIDS* **21**:1-11.

CHAPTER 2: MATERIALS AND METHODS

2.1. STUDY POPULATION

A total of 38 ART-naïve HIV-1-infected patients were recruited at the Comprehensive Care, Management and Treatment Clinic in Pretoria, South Africa. As summarized in **Table 2.1.** all of these patients had advanced HIV-1 infection, high plasma viral loads (VL), low peripheral blood $CD4^+$ T cell counts (<250 cells/µl)], chronic diarrhoea of greater than 4 weeks duration (i.e. increased baseline stools, more than 2 loose watery stools/day, or stools weight greater than 200 g/day) and (or) unintentional weight loss (>10%) of unknown etiology. Patients were eligible for the study if they were >18 years of age, were cART-naïve and tested negative for *Mycobacterium tuberculosis* in sputum smears and negative for enteric pathogens in stool specimens. All 38 patients were started on a standardized first-line drug regimen consisting of two nucleoside reverse transcriptase inhibitors (NRTIs: stavudine + lamivudine) and one non-NRTI (NNRTI: efavirenz or nevirapine) as specified in the South African Treatment Guidelines. [1]. Biopsies were collected at enrollment (duodenum, jejunum, ileum, colon) and at 3 (duodenum) and 6 (duodenum, colon) months after the initiation of cART. Monthly blood samples were collected for $CD4^+$ and $CD8^+$ T cell counts and viral load (VL) testing. The control group (n=5) consisted of HIV-1 seronegative volunteers who tested pathogen negative and showed no evidence of disease during screening for colorectal cancer. Due to the limited amount of material present in “pinch biopsies” not all studies were performed on all patients. A subset of duodenal, ileal, jejunal, and colonic biopsies collected from 10 representative patients, prior to the introduction of ART, was subjected to sequence and phylogenetic analysis. The study protocol was approved by the Faculty of Health Sciences Research Ethics Committee of the University of Pretoria. The study project was registered under the application number 71/2006 and was reviewed on an annual basis. Prior to enrollment, written informed consent was obtained from all study participants.

2.2. SAMPLE COLLECTION

Biopsies of the duodenum and jejunum were collected using a Fujinon double balloon enteroscope and Radial III biopsy forceps (Boston Scientific, Natick, MA, USA). Endoscopic biopsies of the colon and terminal ileum were obtained at the same session through the anal route using a standard colonoscope. The Endoscopy Procedure: Endoscopy was performed by

a Gastroenterologist under conscious sedation (Midazolam or Propafol), using a Fujinon double balloon enteroscope. Biopsies were obtained at 40-50cm intervals starting from the proximal small bowel until the terminal ileum was reached using a single use biopsy forceps at every site. Five biopsies were taken per site. All endoscopic abnormalities were recorded and additional biopsies were taken for diagnostic purposes where deemed appropriate by the Gastroenterologist. At three and six months repeat upper GI endoscopies were performed following the initiation of HAART therapy using a standard endoscope (gastroscope) of the proximal small bowel. A total of 8-10 "pinch" biopsies were randomly collected from each site. Biopsies were either snap frozen for sequence and HIV RNA (VL) analysis or placed on ice in RPMI 1640/10% FCS and immediately processed for flow cytometry, or fixed in 10% formalin for immunohistochemistry. Blood specimens, collected in EDTA, were separated into plasma and peripheral blood mononuclear cells (PBMC) by Ficoll-Hypaque (Sigma-Aldrich, St Louis, MO, USA) density gradient centrifugation. Plasma aliquots and an aliquot of the PBMCs were stored at -80^{o}C until required for sequence and VL analysis. The remaining PBMCs were processed for flow cytometry and $CD4^{+}$/$CD8^{+}$ T cell subset quantification using standard methods.

2.3. PLASMA AND TISSUE-ASSOCIATED VIRAL LOADS

Biopsy (4-12 mg) samples were minced with a razor blade, placed in lysis buffer (Nuclisens Isolation Kit; BioMerieux, Hazelwood, MO, USA) and agitated overnight at room temperature. RNA was extracted using the Magnetic Extraction Reagent Kit from BioMerieux. HIV-1 RNA in biopsy and plasma extracts was quantified using the Nuclisens Easy Q HIV-1 v.1.2 kit and expressed as copies per tissue/g of tissue or mL of plasma.

2.4. $CD4^{+}$ AND $CD8^{+}$ T CELL SUBSET DETERMINATIONS

Multi-parameter flow cytometry was performed on PBMCs and mononuclear cells (MNC) isolated from intestinal biopsies. Biopsies were digested with collagenase type IV (0.5 mg/mL: Sigma, St. Louis, MO, USA) in RPMI medium (Bio Whittaker, Walkersville, MD, USA) (30 min, 37^{o}C) and passed through a 70-µm cell strainer (Becton Dickinson Labware, NJ). The resultant single suspensions were then stained with various combinations of anti-human CD4, CD8, CD3, CD45 and CD38 monoclonal antibodies (mAb, Beckman

Coulter, Miami,, FL, USA) as described previously. The percentage of $CD4^+$, $CD8^+$ and activated $CD38^+CD8^+$ T cells was determined on Beckman Coulter FC500 using Flow Jo (Tree Star, Ashland, OR, USA) software.

2.5. REVERSE TRANSCRIPTION OF HIV-1 RNA

2.5.1. Synthesis of env cDNA

Total RNA was extracted from biopsy and PBMC samples using the Magnetic Extraction Kit (BioMerieux, Hazelwood, MO, USA) and reversed transcribed into *env* cDNA using Superscript III (Invitrogen, San Diego, CA, USA) and oligo (dT). The RNA template was heated to 65^{o}C for 5 minutes to denature secondary structure, cooled to 4^{o}C for 2 min and reverse transcribed for 60 min at 50^{o}C in a reaction mixture containing 1.0 µL oligodT, RT Enzyme (SuperScript III RT and RNaseOUT) and RT Reaction Mix (1x buffer, $MgCl_2$, dNTPs) (Invitrogen, San Diego, CA, USA) followed by heating to 85^{o}C for 5 min and chilling on ice to terminate the reaction. RNase H was added and the reaction was incubated an additional 20 min at 37^{o}C to degrade the RNA template. This approach amplified only mRNA variants of HIV-1 and allowed us to differentiate between cell-associated mRNA and RNA in viral particles. The cDNA was purified using the Invitrogen Purification Kit (Invitrogen, San Diego, CA, USA).

2.5.2. Synthesis of gag cDNA

Total RNA, extracted using the BioMerieux method, was reverse transcribed into *gag* cDNA using Superscript III. The reaction conditions were the same as those used for env, except that the reaction involved the use of a sequence specific reverse primer, KVL065 (Table 2.1). Following RNase H digestion (20 min, 37^{o}C), the cDNA was purified using the GeneJET RNA Purification Kit (Fermentas Life Sciences, Burlinton, Ontario, Canada).

2.5.3. Synthesis of pol cDNA

BioMerieux extracted RNA was reverse transcribed into *pol* cDNA using the MMLV Reverse Transcriptase 1st Strand cDNA Synthesis Kit (Epicentre Biotechnologies, Madison,

WI, USA) and the sequence specific reverse primer, IN3 [2] (Table 2.1). Total RNA was heated to 65^{o}C for 5 min, cooled on ice and reverse transcribed at 37^{o}C for 60 min in a 23µL reaction mixture containing MMLV RT enzyme, ScriptGuard^{TM} RNase inhibitor, MMLV RT mix (1x buffer, DTT, dNTPs), followed by heating for 85^{o}C for 5 minutes and chilling on ice to terminate the reaction Following RNase H digestion (20 min, 37^{o}C), the cDNA was purified using the GeneJET RNA Purification Kit (Fermentas Life Sciences, Burlinton, Ontario, Canada).

The reason why different PCR amplification kits were used for *env* and *gag* versus *pol*, was due to failure of amplication of *gag* and *env* using the 1st Strand cDNA Synthesis kit (Epicentre Biotechnologies), which lead me to use a different kit in this case the Superscript III from Invitrogen.

2.6. PCR AMPLIFICATION OF ENV, GAG AND POL cDNAs

2.6.1. Nested PCR amplification of env gene region

Amplification of a 667-bp segment of the *env* C2V5 gene region was performed in a nested PCR reaction using MK605 and CD4R2 as outer and ES7 and ES8 as inner primers (3) (**Table 2.2).** Both PCR steps were carried out in a final volume of 25 µl, containing 1x PCR buffer, 2.5 mM $MgCl_2$, 0.25 mM each deoxynucleoside triphosphate, 0.05 pmol of each primer, and 0.02 U of Platinum Tag polymerase (Invitrogen, Carlsbad, CA, USA). Thermocycling conditions were 94°C for 2 min, followed by 45 cycles at 95°C for 30 s, 58°C for 20 s, and 72°C for 2 min, and a final extension of 72°C for 10 min. Amplified DNA was visualized by a 1% agarose gel electrophoresis, and the DNA was purified using the Invitrogen Purification Kit and a Microcon (Amicon) spin column, and sequenced on an automated 3100 genetic analyzer (Applied Biosystems Inc., Foster City, CA, USA) using M13 sequencing primers and a Big-Dye terminator cycle sequencing kit.

2.6.2. Nested PCR amplification of gag gene region

Amplification of 1997-bp segment of the gag gene was performed using KVL064 and KVL065 as outer and KVL66 and KVL67 as inner primers (**Table 2.2**). Both PCR reactions

were carried out in a final volume of 25 µl containing 1x PCR buffer, 2.0 mM $MgS0_4$, 0.25 mM dNTPs, 1.25 pmoles of outer or inner primers and 0.5 U of Platinum Taq High Fidelity DNA polymerase (Invitrogen, San Diego, CA, USA). The PCR conditions were 94°C for 2 min, followed by 30 cycles at 94°C for 30 s, 58°C for 20 s, and 72°C for 2 min, with a final extension at 72°C for 10 min. PCR products were quantified by agarose gel electrophoresis and purified using the GeneJET™ Gel Extraction Kit (Fermentas Life Sciences, Burlinton, Ontario, Canada).

The *pol* and *gag* genes were selected for study because of their structural stability and relatively limited intra-host evolution in the absence of ART and, in the case of *pol*, for its importance in the emergence of drug resistance; the *env-C2V5* genomic region was selected for its extensive intra- and inter-host variability and genetic evolution in response to immune selection pressure.

2.6.3. Nested PCR amplification of the pol gene region

Amplification of a 1856-bp segment of *pol* cDNA was performed using G25Rev and IN3 as outer and AV150 and POLM4 as inner primers (**Table 2.2**). Both PCRs contained 1x PCR buffer with 2 mM $MgCl_2$, 0.2 µM of each dNTP, 1.25 pmoles of each primer and 0.5 Units of Excel Taq DNA polymerase (ExSel, JMR Holdings Inc, London, UK). Thermocycling conditions were 94°C for 2 minutes, followed by 35 cycles at 94°C for 10 s, 50°C for 30 s, and 70°C for 2 minutes and an extension cycle at 70°C for 7 minutes. PCR products were quantified by agarose gel electrophoresis and purified using the GeneJET™ Gel Extraction (Fermentas Life Sciences, Burlinton, Ontario, Canada) and Purelink PCR Purification (Invitrogen, San Diego, CA, USA) kits.

2.7. CLONING AND SEQUENCING

Purified RT-PCR products were cloned into the pCU®4 TOPO TA vector (Invitrogen, San Diego, CA, USA) and transformed into XL 10-Gold Ultra-competent cells (Agilent Technologies, Santa Clara, CA, USA). Plasmids containing the expected inserts, as confirmed by restriction digestion with FastDigest EcoRI (Fermentas Life Sciences, Burlinton, Ontario, Canada), were purified from the bacterial lysates using the Wizard® SV

96 Plasmid DNA Purification System (Promega Corporation, Madison, WI, USA). The plasmid DNA was quantified using a Nanodrop spectrophotometer and sequenced on an automated ABI 3100 DNA analyzer using PRISM Big-Dye terminator v3.1 reagents (ABI, Foster City, CA, USA) and the following primers: T7 and ES8 for *env*, and KVL066, KVL067, KVL078, KVL079, KVL080, GAI, AV103 and AV159 for *gag*. The conditions used for cycle sequencing were 96^{o}C for 10 s, 50^{o}C for 5 s and 60^{o}C for 4 min. The same conditions were used for population sequencing of *gag* and *pol* RT-PCR products. Sequencing of the pol gene fragments was performed using the primers AV150, POLM4, SQ5F, SQ12R, SQ6F, and 13R(2)C(2) (**Table 2.2.**). A genomic map of the primers used in this study is shown in **Figure 2.1.**

2.8. SEQUENCE ALIGNMENT, EDITING AND GENETIC SUBTYPING

To rule out contamination between samples, each new sequence was compared to other sequences amplified at the same time and to HIV-1 sequences previously amplified in our laboratory. Chromatograms were imported and edited in CLC DNA workbench 5 (http://clc-dna-workbench.softpedia.com) [4] and/or Geneious version 5.3.6 (http://www.geneious.com) [5; 6]. Sequences were aligned using CLUSTAL W and manually edited using Bioedit version 7.0.4.1 [7] and/or Se-AL (http:tree.bio.ed.ac.uk) [8]. Gap-stripped sequences were compared to other ZA subtype C strains found in the Los Alamos subtype database (http://hiv-web.lanl.gov/content/hivdb/ SUBTYPE_REF/align. html) [9]. The Rega HIV-1 subtyping tool version 3.0 (http://www. bioafrica.net/rega-genotype/html/subtypinghiv.html) [10] and COMET (http://comet. retrovirology.lu/index.php) [11], automated tools were used to subtype all sequences.

2.9. PHYLOGENETIC ANALYSES

A broad range of evolutionary and positive selection pressure programs was used to construct phylogenetic relationships among sequences amplified from different regions of the GIT and to compare these sequences to those derived from PBMCs. These programs included codon-based substitution Maximum Likelihood (ML) models implemented in PhyML [12] and the Bayesian tree inference program [13; 14]. The latter method uses a Markov Chain Monte Carlo (MCMC) sampling technique to assess virus compartmentalization between different tissues. ML and Bayesian tree construction was performed using the Gamma heterogeneity alpha parameter. The SPRs method and a BioNJ starting tree implemented in

PhyML were used to optimize ML tree topology. Bootscanning methods (100 replicates) were used to assess the reliability of the internal nodes on the ML trees. Bootstrap values >70% were considered significant. The reproducibility of Bayesian trees was assessed using four MCMC chains of 10,000,000 generations sampled every 1,000 generations. The first 10% of the sampled trees and estimated parameters were excluded from the final analysis (i.e. burning). Tracer (http://beast.bio.ed.ac.uk) [15], was used to visualize the results of the estimated parameters. All the parameters had effective sample sizes greater than 300, suggesting good mixing and sampling of the trees. TreeAnnotator (http://beast.bio.ed.ac.uk) [15], was used to generate the consensus tree. Posterior probability support for the evaluation of the internal branches was estimated using FigTree.

The BEAST software package 1.6 (http://beast.bio.ed.ac.uk) [15], was used to investigate migration patterns and estimate the most recent common ancestor (MRCA). The initial tree was generated using a simple, constant population size coalescent model. The rate of nucleotide substitution was estimated using a general time reversible (GTR) substitution model with gamma distribution. An uncorrected relaxed clock with gamma distribution and an initial chain of 10^8 was used to determine the time of divergence of each compartment. A burning of 10% was applied with the initial 10% of data points being excluded from the analysis. The log file output was viewed in Tracer and was deemed suitable for analysis if the estimated sample sizes were greater than 100. Tree Annotator was used to generate a consensus tree, which was visualized in FIGTREE.

2.10. GENETIC DIVERSITY

Pairwise distances among HIV-1 RNA quasispecies in different regions of the GIT were calculated using Kimura 2 parametric model of nucleotide substitution implemented in MEGA [16, 17]. The gamma distribution model was used to estimate the rate of variation among sites and a bootstrap of 1000 replications was used to estimate the mean distances and standard errors (S.E.). The mean diversity for the different GIT compartments (plasma, duodenum, ileum, jejunum, right and left colon) was calculated for patients (patient HIV019, 021, and 029).

2.11. AMINO ACID SUBSTITUTION, SELECTION PRESSURE AND FUNCTIONAL MOTIFS

Env and *gag* sequences were translated, aligned and used to calculate the rate of synonymous to non-synonymous amino acid substitution as a measure of natural selection pressure at the protein level. Various methods [i.e. Web Position Specific Scoring Matrix (PSSM) algorithm (http://indra.mullins.microbiol.washington.edu/pssm/_index2.html) [18; 19], the geno2pheno prediction algorithm (http://coreceptor.bioinf.mpi-inf.mpg.de/index.php [20; 21], MEGA [16; 17], were used to compare the HIV-1 quasispecies amplified from the duodenum, jejunum, ileum and colon to the consequence sequence for that compartment and to the subtype C consensus sequence for southern Africa. High rates of non-synonymous mutation were taken as evidence of adaptive change in response to tissue-specific selection pressure. Prediction of co-receptor usage (CCR5 vs. CXCR4) among HIV-1 quasispecies in different regions of the GIT was based on amino acid analysis of the V3 loop using the subtype C specific Web PSSM genotypic interpretation algorithm [19].

2.12. DRUG RESISTANCE ANALYSIS

RT and protease sequences were screened for the presence of drug resistant mutations (NRTI, NNRTI and PI) using the HIV Drug Resistance Database from Stanford University site (http://hivdb.stanford.edu/hiv/) [22].

2.13. NUCLEOTIDE SEQUENCE ACCESSION NUMBERS

The following accession numbers were assigned for the sequences obtained in this study: KF905653- KF905865, for *env* gene sequences; KF905866- KF906024, for *gag* gene sequences; KF906025- KF906069, for *pol* gene sequences.

2.14. IMMUNOPHENOTYPING

Multiparameter flow cytometry was performed on mononuclear cells (MNC) isolated from intestinal biopsies. MNC and GIT samples were digested with collagenase type IV (0.5 mg/mL: Sigma, St. Louis, MO) in RPMI medium (30 min, 37^{o}C) and passed through a 70-

μm cell strainer (Becton Dickinson Labware, NJ). Residual tissue fragments were re-digested and the pooled, single cell suspensions were washed to remove collagenase. MNC were stained (30 min, 4°C) with various combinations of anti-human CD4, CD8, CD3, CD45 and CD38 monoclonal antibodies (mAb, Beckman Coulter). The percentage of $CD4^+$, $CD8^+$ and activated $CD38^+CD8^+$ T cells was determined by gating on the lymphocyte population defined by forward and side-scatter characteristics (confirmed to be $CD45^+CD3^+$ cells by back gating) followed sequential gating for either CD4 or CD8 and CD38. A minimum of 5,000 lymphocyte events were collected per tube on a Beckman Coulter FC500. Results were analyzed using Flow Jo (Tree Star, Ashland, OR).

2.15. HISTOLOGICAL ANALYSIS

H&E, periodic acid-Schiff diastase (PAS), Ziehl-Neelsen and Giemsa staining were performed to exclude fungal and parasitic organisms (Cryptosporidia, Isopora and Microsporidia) and AFB. Quantification of T cell subsets was performed on tissue sections incubated with murine anti-human CD4 or CD8 mAbs (Dako, Denmark), visualized with DAB and counter-stained with haematoxylin. The number of $CD4^+$ and $CD8^+$ lymphocytes in five representative high power fields of 0.80 mm^2 (400X) was counted and the results reported as the average number of cells/0.8 mm^2.

2.16. STATISTICAL ANALYSES

All statistical analyses were performed in Prism 5 from GraphPad Software (La Jolla, CA). Paired observations were compared using Wilcoxon matched pair tests or the Friedman test with Dunns post-test. Linear correlations were assessed using Spearman's rank correlation coefficient. Two-tailed p values >0.05 were considered significant.

2.17. FIGURE LEGENDS:

Figures:

Figure 2.1: A genomic map of HIV-1 indicating the location of the PCR and sequencing primers used in this study.

Tables:

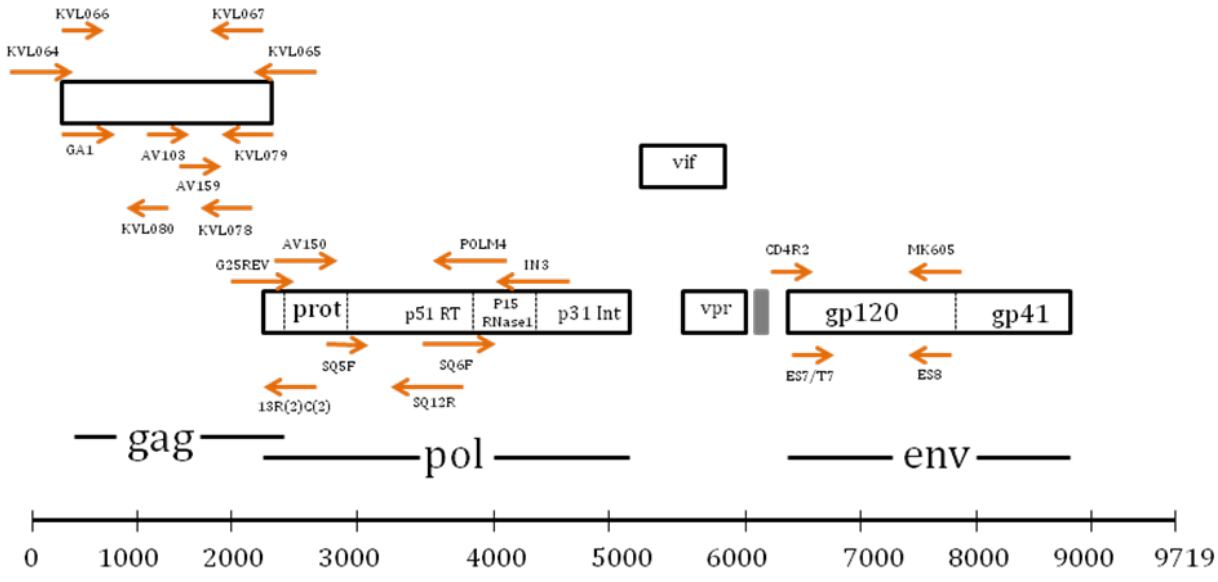
Table 2.1: Clinical Characteristics of HIV-1^{+} Subjects (n=38).

Table 2.2: Primers used in the study (synthesized by IDT).

3. FIGURES AND TABLES

Figures:

Figure 2.1:



© Phetole Mahasha
2013

Tables:

Table 2.1:

Gender	Male	15
	Female	23
Age (years)	Mean ± S.D.	36 ± 9
	Median (Range)	33 (23 - 58)
CD4⁺ T cell count (cells/μL)	Mean ± S.D.	71 ± 59
	Median (Range)	57(6 - 224)
Percentage CD4⁺ T cells	Mean ± S.D.	8.6 ± 5.5
	Median (Range)	7.4 (0.4 – 22.7)
Plasma VL (HIV-1 RNA copies/mL)	Mean ± S.D.	309663 ± 533487
	Median (Range)	125000 (160 - 3000000)
Diarrhoea	Absent	11
	Present	27
Weight loss	Absent	6
	<10%	6
	>10%	26

Table 2.2:

ENV	Primer Sequence (5'-3')	Region	Nested PCR and Sequencing	Primer Position in HXB2	References
MK605	5AATGTCAGCACAGTACAATGTACAC	Outside Forward	Nested PCR	6945–6969	Gordon *et al*, 2003
CD4R2	TATAATTCACTTGTCCAATTGTCC	Outside Reverse	Nested PCR	7652–7675	Gordon *et al*, 2003
ES7	CTGTTAAATGGCAGTCTAGC	Inside Forward	Nested PCR & Sequencing	7002–7021	Gordon *et al*, 2003
ES8	CACTTCTCCAATTGTCCCTCA	Inside Reverse	Nested PCR & Sequencing	7648–7668	Gordon *et al*, 2003
T7	TAATACGACTCACTATAGGG	Forward	Sequencing	***pCR4-TOPO vector:*** 328–347	**TOPO TA Cloning® Kit for Sequencing from Invitrogen**
GAG	**Primer Sequence (5'-3')**	**Region**	**Nested PCR and Sequencing**	**Primer Position in HXB2**	**References**
KVL064	GTTGTGTGACTCTGGTAACTAGAGATCCCTCAGA	Outside Forward	Nested PCR	570–603	Van Laethem *et al*, 2006
KVL065	TCCTAATTGAACYTCCCARAARTCYTGAGTTC	Outside Reverse	Nested PCR	2797–2828	Van Laethem *et al*, 2006
KVL066	TCTCTAGCAGTGGCGCCCGAACAG	Inside Forward	Nested PCR & Sequencing	626–649	Van Laethem *et al*, 2006
KVL067	GGCCATTGTTTAACYTTTGGDCCATCC	Inside Reverse	Nested PCR & Sequencing	2597–2623	Van Laethem *et al*, 2006
KVL078	CATTCCTGGCTTTAATTTTACTG	Reverse	Sequencing	2575–2597	Van Laethem *et al*, 2006
KVL079	GTGTCCTTCCTTYCCACATTTCC	Reverse	Sequencing	2030–2052	Van Laethem *et al*, 2006
KVL080	CCCATTCTGCAGCTTCCTCATTG	Reverse	Sequencing	1404–1426	Van Laethem *et al*, 2006
GA1	GACGCAGGACTCGGCTTGCT	Forward	Sequencing	688–707	Van Laethem *et al*, 2006
AV103	GCCATATCACCTAGAACTTT	Forward	Sequencing	1225–1244	Van Laethem *et al*, 2006
AV159	GGGGTTAAATAAAATAGTAAG	Forward	Sequencing	1593–1613	Van Laethem *et al*, 2006
POL	**Primer Sequence (5'-3')**	**Region**	**Nested PCR and Sequencing**	**Primer Position in HXB2**	**References**
G25Rev	GCAAGAGTTTTGGCTGAAGCAATGAG	Outside Forward	Nested PCR	1867–1892	Pillay, 2006 (NHLS Document)
IN3	TCTATVCCATCTAAAAATAGTACTTTCCTGATTCC	Outside Reverse	Nested PCR	4212–4246	Pillay, 2006 (NHLS Document)
AV150	GTGGAAAGGAAGGACACCAAATGAAAG	Inside Forward	Nested PCR & Sequencing	2036–2063	Pillay, 2006 (NHLS Document)
POLM4	CTATTAGCTGCCCCATCTACATA	Inside Reverse	Nested PCR & Sequencing	3892–3915	Pillay, 2006 (NHLS Document)
SQ5F	AAACAATGGCCATTAACAGAAGAGA	Forward	Sequencing	2613–2637	Vergne *et al.*, 2000 and Rousseau *et al.* 2006
SQ12R	TGTATGTCATTGACAGTCC	Reverse	Sequencing	3304–3322	Vergne *et al.*, 2000 and Rousseau *et al.* 2006
SQ6F	CTTTGGATGGGTTATGAACT	Forward	Sequencing	3231–3250	Vergne *et al.*, 2000 and Rousseau *et al.* 2006
13R(2)C(2)	GTGTTATATGGATTTTCAGGCCC	Reverse	Sequencing	2700-2722	Vergne *et al.*, 2000 and Rousseau *et al.*, 2006

4. REFERENCES

1. National Department of Health, Republic of South Africa. Clinical guidelines for the management of HIV and AIDS in adults and adolescents. 2004 (www.sanac.org.za).

2. Pillay V (2006) In-house assay for pol sequencing to monitor ARV Drug Resistance National Health Laboratory Service (South Africa), Q-Pulse4/docs/active/ NIC0085: 1-12.

3. Gordon, M., De Oliveira, T., Bishop, K., Coovadia, H. M., Madurai, L., Engelbrecht, S., Janse van Rensburg, E., Mosam, A., Smith, A., and Cassol, S. 2003. Molecular Characteristics of HIV-1 Subtype C Viruses from KwaZulu-Natal, South Africa: Implications for Vaccine and Antiretroviral Control Strategies. *J. Virol.* **77**(4): 2587-2599.

4. http://clc-dna-workbench.softpedia.com

5. http://www.geneious.com/

6. Rozen, S., and Skaletsky, H.J. (2000). Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386.

7. Thompson, J. D., Higgins, D. G., and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucl. Ac. Res.* **22**(22): 4673-4680.

8. http:tree.bio.ed.ac.uk

9. http://hiv-web.lanl.gov/content/hivdb/SUBTYPE_REF/align. html

10. http://www.bioafrica.net/rega-genotype/html/subtypinghiv.html

11. Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., Gascuel O. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0." *Systematic Biology*, 59(3):307-21. (http://www.atgc-montpellier.fr/phyml/

12. http://comet. retrovirology.lu/index.php

13. Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17:754-755.

14. Ronquist, F., and Huelsenbeck, J. P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.
15. http://beast.bio.ed.ac.uk
16. Kimura M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16:1111-20.
17. Tamura K., Peterson D., Peterson N., Stecher G., Nei M., and Kumar S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol and Evol 28*: 2731-2739.
18. http://indra.mullins.microbiol.washington.edu/pssm/index2.html.
19. Jensen, M.A., Li, F.S., van `t Wout, A.B., Nickle, D.C., Shriner, D., *et al*. (2003). Improved coreceptor usage prediction and genotype monitoring of R5-to-X4 transition by motif analysis of human immunodefiency virus type 1 env V3 loop sequences. *J Virol* 77: 13376-13388.
20. http://coreceptor.bioinf.mpi-inf.mpg.de/index.php.
21. Lengauer, T., Sander, O., Sierra, S., Thielen, A., and Kaiser, R. (2007). Bioinformatics prediction of HIV coreceptor usage. *Nat Biotechnol.* **25**(12):1407-10.
22. http://hivdb.stanford.edu/hiv/

CHAPTER 3:

DIFFERENTIAL DISTRIBUTION AND EVOLUTION OF HIV-1 RNA VARIANTS IN THE GASTROINTESTINAL TRACT OF ANTIRETROVIRAL NAÏVE AFRICAN AIDS PATIENTS WITH DIARRHEA AND (OR) WEIGHT LOSS

TABLE OF CONTENT: **PAGES:**

3.0. INTRODUCTION

The gastrointestinal tract (GIT) is a complex organ composed of multiple tissue compartments with different structural and functional elements and diverse cell types. Because of its continuous exposure to a wide range of microbes and food antigens, the GIT is in a constant state of "physiological" inflammation and contains a large number of activated $CCR5^+$ memory $CD4^+$ T cells, the primary target of HIV-1 [1-6]. As a result, the GIT is a site of intense viral replication and severe $CD4^+$ T cell depletion during primary HIV-1 infection [2; 3; 7]. This process continues at a lesser pace throughout the chronic phase of infection and is associated with progressive immunological and structural damage to the gut associated lymphoid tissue (GALT), a loss of epithelial integrity, increased microbial translocation and both local and systemic immune activation [8-10]. Highly active ART (HAART) is only partially effective in restoring the immunological and structural integrity of the intestinal mucosa, even when initiated during primary infection ([11-13].

Despite the importance of the GIT in the pathogenesis of HIV-1 infection, and as a reservoir of treatment refractory virus, little is known about the genetic diversity and evolution of HIV-1 variants in different regions of the GIT and whether these variants are associated with differing pathologies. There is also a paucity of information about the dissemination and genetic relationships between intestinal variants and those found in other tissues. This is due, in large measure, to the difficulties associated with the sampling of multiple tissue sites and the limited amount of material that is available in "pinch" biopsies. Most studies of HIV-1 genetic variation have been confined to 1 or 2 anatomical sites in the GIT [14-18] and have involved the sequencing of proviral HIV-1 DNA isolated from patients on ART [15; 17-20]. Proviral DNA contains archived and defective genetic material and may not accurately reflect the full spectrum of replication competent viral strains, while studies of patients on suppressive ART with low or undetectable levels of viral replication and decreased chemokine levels may minimize the trafficking and evolution of HIV-1 variants [16; 18]. A recent study by Avettand-Fenoel *et al*, 2011 [17] found that prior to the introduction of ART, HIV-1 DNA variants in the rectum were more diverse than those in blood but that in patients on suppressive HAART there was an "equalization" of the dynamics of HIV-1 variants in the rectum and blood**.**

Understanding the diversity and evolution of HIV-1 in the GIT is a key step toward

understanding the emergence of HIV-1 variants with increased pathogenicity. However, as recent studies have shown, there are marked differences between the small vs. large intestine in terms of viral replication, $CD4^+$ T cell depletion and levels of immune activation and inflammation, all of which may impact on the evolution of viral diversity [21-23]. The aim of this study was to characterize HIV-1 RNA variants in different regions of the GIT (duodenum, jejunum, ileum, right and left colon) of ART-naïve patients with chronic diarrhea and (or) unexplained weight loss and to assess whether the proteins encoded by these RNA variants were likely to differ in their biological properties. This information may have important implications not only for interrupting disease progression but also for the optimization of improved strategies to assess and optimize therapeutic efficacy.

3.1. MATERIALS AND METHODS (Described in detail in Chapter 2).

3.2. RESULTS

3.2.1. Patient characteristics

The 10 patients in this study were ART-naïve and had advanced -1 infection (WHO stage III or IV). The mean $CD4^+$ T cell count (± SD) at the time of presentation was 102 ± 65 cells/μL; the mean VL (± SD) was 258,801 ± 375,428 copies of HIV-1 RNA/mL of plasma and the mean age (± SD) was 37 ± 10.6 years. The study included equal numbers of male (n=5) and female (n = 5) patients. Symptoms leading to endoscopy were diarrhea in 8 (80%) and unexplained weight loss >10% in 2 (20%) patients. Non-specific mild-to-moderate enteritis or enterocolitis was detected in at least one site in 70% of patients, at a similar frequency in all four intestinal sites (70%, 70%, 80% and 70% in the duodenum, jejunum, ileum and colon, respectively).

3.2.2. HIV-1 subtyping and genetic diversity

The Rega HIV-1 version 3.0 (http://www.bioafrica.net/rega-genotype/html/subtypinghiv. html), and COMET HIV-1/2 (http://comet.retrovirology.lu/index.php) automated tools were used to determine subtype.

These programmes, together with SCUEAL (www.datamonkey. org/dataupload_ scueal.php), 24] have been shown to be the best performing tools for HIV-1 subtyping and break-point analysis [25; 26]. Subtype analysis was performed on 45 RT PCR *pol* sequences and 350 cloned viral quasispecies derived from the *env* and *gag* gene regions of 5 different sites in the GIT of patients HIV019, HIV021 and HIV029. All sequences grouped as subtype C when analyzed by both REGA and COMET, and in the case of *pol*, also when analyzed by the SCUEAL method. SCUEAL was not used in the analysis of *gag* and *env* as it is applicable to the *pol* gene only. The finding that matched *env-gag* sequence pairs typed as subtype C suggests that none of our patients were infected with a C/D, or other intersubtype recombinant virus. Subtyping results were supported by bootstrap values of > 95%.

The Kimura-2 parametric model of nucleotide substitution implemented in MEGA [27; 28] was used to calculate the extent of viral diversity among the 350 cloned viral sequences obtained from the duodenum (n = 61), jejunum (n = 72), ileum (n = 73), right colon (n = 80) and left colon (n = 64) colon (**Table 3.2**) of patients HIV019, HIV021 and HIV029. These 3 representative patients exhibited varying levels of inter- and intra-patient genetic diversity. The mean overall pairwise distances among *env* HIV-1 RNA variants in the GIT ranged from 2.27% (S.E., 0.35%) in patient HIV019 to 5.99% (S.E., 0.74%) and 4.25% (S.E., 0.54%) in patients HIV021 and HIV029, respectively (**Figure 3.1A**). A comparison of inter-patient diversities indicated that these differences were statistically significant at $p<0.001$ and $p<0.0001$ levels. The pairwise distances among *gag* HIV-1 RNA variants were substantially lower with mean levels ranging from 1.20% (S.E., 0.14%) in patient HIV019 to 1.96% (S.E., 0.16%) in patient HIV021 and 1.22% (S.E., 0.14%) in patient HIV029 (**Figure 3.1 B)**. For each of the 3 patients, the diversity among *env* variants was significantly greater than among *gag* variants ($p<0.001$ for patient HIV019; $p<0.0001$ for patients HIV021 and HIV029) (**Figure 3.1 C**).

I next assessed the genetic diversity within different anatomical compartments of the GIT (ie. the duodenum, jejunum, ileum, right colon and left colon). A total of 131 cloned *env* (n= 72) and *gag* (n = 59) sequences were analyzed for patient HIV019. The corresponding number of sequences analyzed for patients HIV021 and HIV029 were 110 (*env* = 66; *gag* = 44) and 109 (env = *65*; *gag* = 44), respectively. The number of clones analyzed for each individual compartment is shown in **Table 3.2**. Depending on the patient, inter-compartmental distances ranged from 0.20% to 5.80% among *env* and from 0.94% to 2.12%

among *gag* HIV-1 RNA variants (**Table 3.3**). In patient HIV029, the diversity (± S.E.) among plasma variants of HIV-1 RNA (5.68 ± 0.72% for env; 1.16 ± 0.16% for *gag*) was equivalent to, or slightly greater than, that of variants in the ileum (4.60 ± 0.57% for env; 1.16 ± 0.18% for gag) and jejunum (4.46% ± 0.59% for env; 1.17% ± 0.16 % for gag). With few exceptions (ie. the jejunum of patient HIV021 and the right colon of patient HIV029), the diversity among *env* quasispecies was significantly greater than among *gag* variants in the same tissue compartment **(Table 3.3**; **Figure 3.2 A-F**). In two of the three patients, HIV019 and HIV029, *env* diversity was significantly greater in the small intestine (and plasma) than in the left and/or right colon (**Figure 3.2 A-F)**.

3.2.3. Phylogenetic relationships among HIV-1 RNA quasispecies in different compartments of the GIT

Phylogenetic analysis indicated that all *pol*, *env* and *gag* sequences clustered by individual patients with no cross-contamination of sequences (bootstrap >70%). This pattern was observed irrespective of whether Maximum Likelihood methods or Bayesian methods were used for tree contstruction. A representative ML tree is shown in **Figure 3.3**.

In the next set of analyses, I investigated the compartmentalization of HIV-1 RNA variants in the GIT in patients HIV019, HIV021 and HIV029 (**Figures 3.4- 3.9**). Partial or total inter-dispersion of viral variants was observed among different anatomical compartments of the GIT, and between the GIT and plasma (patient 29), for both *env* and *gag* quasispecies. The migration and exchange of variants among different tissue compartments was most evident in the small intestine (duodenum, jejunum and ileum), with less extensive trafficking and mixing of viral quasispecies occurring in the right and left colon. In addition to mixed phylogenetic clusters, distinct monophyletic clusters of closely related sequences were detected in the GIT of all 3 patients, predominantly in the right and left colon and to a lesser extent in the small intestine (in the duodenum of patient HIV019 and the jejunum and ileum of patient HIV021). Monophyletic clustering was more evident for the *env* than for the *gag* encoding gene region. The presence of tight monophyletic clusters suggests that, in addition to extensive migration between different compartments in the GIT, there is localized tissue-specific evolution of the HIV-1 *env* gene, especially in the colon.

A characteristic feature of the *env* gene of patient HIV019 (**Figure 3.4**) was the presence of distinct monophyletic clusters in the duodenum (Clusters 1 and 6), left colon (Cluster 2) and right colon (Clusters 3, 4 and 5). The right colon, the compartment with the lowest level of genetic diversity, exhibited the highest level of monophyletic clustering. Greater than 93% (14/15) of sequences in the right colon fell into clusters 3, 4 and 5 with posterior probabilities ranging from 0.45 to 0.99. Most of the *env* sequences in the small intestine were inter-dispersed among the duodenum, jejunum and ileum. One large mixed jejunal/ileal cluster (with a posterior probability score >0.95) was detected on an external branch of the tree (Cluster 7). This cluster was external to Cluster 6 suggesting it may have originated in the duodenum. However, another 3 ileal sequences clustered basal to the combined duodenal and jejunal-ileal clusters with a high posterior support (0.945). Thus, it is not possible to conclusively infer the direction of origin but only to confirm that there was exchange of viral variants between these 3 compartments. Two sequences from the left colon were also basal to the duodenal/jejunal-ileal cluster suggesting that the left colon is also involved in the migration of viral variants between these compartments. There was little evidence of trafficking from other regions of the GIT into, or out of, the right colon.

A notable feature of patient HIV021 was the relationship between *env* sequences in the ileum and the right and left colon (**Figure 3.5**). These sequences were found in 4 monophyletic clusters that segregated together on a single branch at the top of the tree. The basal cluster (Cluster 1) consisted of 6 sequences from the left colon that were supported with a posterior probability score of 1.0. The most external cluster (Cluster 2), with a posterior probability score of 1.0, was composed of 12 sequences found in the right colon. This cluster was closely linked to 2 monophyletic IL clusters (Clusters 3 and 4), one of which was basal to the main R/IL cluster (Cluster 4). The posterior probabilities for these clusters were 1 and 0.29, respectively. A fifth near-monophyletic cluster was located external to a large cluster of highly inter-dispersed sequences derived from the duodenum, ileum, right and left colon. This external cluster (Cluster 5) was composed 17 closely related sequences present in the jejunum (n =16) and duodenum (n = 1).

As observed for patients HIV019 and HIV021, the *env* gene of patient HIV029 also exhibited a dual pattern of tight monophyletic clustering and extensive inter-mixing of viral quasispecis, primarily between the ileum and other tissue compartments (**Figure 3.6**). The majority of *env* variants from the duodenum (9/13) (Cluster 1) and the left colon (11/13)

(Clustser 2), and all variants from the right colon (14/14) (Cluster 3), fell into distinct monophyletic (or near-monophyletic clusters) with the duodenal cluster (Cluster 1) containing a single jejunal sequence. The posterior probability scores for these clusters were of 1.0, 0.48 and 0.35, respectively. Sequences from the ileum were basal to all 3 clusters. In addition, a large number of sequences in the ileum clustered with sequences in the plasma, jejunum and, to a lesser extent, the duodenum suggesting there was extensive exchange of HIV-1 RNA variants between the ileum and other tissue compartments including peripheral blood.

In contrast to *env*, a similar clustering was not observed for the *gag* encoding gene region (**Figures 3.7** to **3.9**). With the exception of a few small clusters, *gag* HIV-1 RNA variants were highly inter-dispersed not only within the small intestine, but also between the small intestine, the colon and plasma suggesting that inter-compartmental trafficking was the major factor contributing to the heterogeneity of *gag* HIV-1 RNA in different tissues of the GIT.

3.2.4. Amino acid substitutions associated with altered protein function

3.2.4. (a) Amino acid polymorphisms linked to drug resistance

To determine whether any of the nucleotide substitutions in different parts of the GIT were associated with amino acid changes related to drug resistance, we analyzed our *pol* RT PCR dataset for the presence of reverse transcriptase (RT) and protease (PR) resistance mutations using the Standord HIV-1 Drug-Resistance Database (http://hivdb.Stanford.Edu/hiv/). No primary drug resistance mutations were detected in any of the 5 GIT sites, or in plasma samples. However, we did detect substitutions at amino acid positions associated with drug resistance. One of these substitutions, RT D67G, was detected in HIV-1 RNA obtained from the colon, but not from the plasma or small intestine of patient HIV010 (**Table 3.4)**. In contrast to D67N, which confers low-level resistance to zidovudine (ZDV) and stavudine (d4T), the effects of D67G on drug susceptibility are not known. A second highly unusual substitution, RT L210V, was detected in HIV-1 RNA derived from the ileum of patient HIV011 (**Table 3.4**). Similarly, the impact of L210V on drug susceptibiity is largely unknown. However, L210W is a well-known resistance mutation that, when present in combination with RT M41L and T215Y, contributes to resistance to most NRTIs. The two

remaining polymorphisms, PR T74S and PR V77I, were detected in HIV-1 RNA from the plasma, duodenum, jejunum, ileum and right colon of patient HIV011. PR T74S is present in up to 10% of ART-naive patients infected with subtype C and is associated with decreased susceptibility to nelfinavir (NFV). PR V77I, is a common to compensatory mutation that is selected by NFV. Other polymorphisms were present at positions that alter the conformation and 3-dimensional structure of the protease enzyme.

Many of these polymorphisms were located in the hinge region (M36I/R41I/H69K/L89M) and at the N-terminus (T12S/I15V/L19I) of protease. These polymorphisms are prevalent in HIV-1 C viruses and, in our study, were present in the viral RNA extracted from most, but not all, patients and compartments of the GIT (**Table 3.4**). The first series of amino acid substitutions is associated with increased catalytic activity of the protease enzyme [29]; the second series has been linked to an I93L polymorphism associated with increased (hyper) sensitivity to lopinavir [30]**.** I93L was detected in 89% of GIT and 88% of plasma samples. The mean number of substitutions in the RT gene relative to HXB_2 was 23.2 ± 2.3 (range 18-37) with 86.0% of sequences having 21 or more substitutions relative to HXB_2. Some substitutions such as D123D/G/N/S and T286A exhibited extensive inter-patient and inter-compartmental variation while others including V35T, V245Q, T200A and I293V were highly conserved among different patients and different tissues of the GIT. There was no evidence for the preferential presence of any specific RT polymorphism in a given tissue compartment. The most frequent amino acid substitutions are shown in **Table 3.4.**

3.2.4. (b) Amino acid substitutions linked to differences in co-receptor usage

To determine whether HIV-1 RNA variants in different regions of the GIT differed in their immunological and co-receptor binding properties, the V3 loop sequences of patients HIV019, HIV021 and HIV029 were aligned, translated (**Figure 3.10 A, B and C**) and analyzed for substitutions at positions that play a critical role in antibody neutralization and CCR5/CXCR4 co-receptor usage. The primary neutralization domain, a GPGQ tetramer located at the tip of the V3 loop, was conserved in 97.0% (197/203) of GIT and 100% (9/9) of plasma samples. Six tetramer motifs, four in patient HIV021 (GPGH in the duodenum; GPGR and 2 RPGQ in the right colon) and two in patient HIV029 (RPGQ in the duodenum; RPGY in the left.colon) had an arginine (R) or histidine (H) substitution in the first or last

position of the tetramer. The classical V3 signature associated with CCR5 usage (S11, 22A and 25D) was detected in 99.0% of 203 GIT and 100% of 9 plasma samples. One HIV-1 RNA variant in the jejunum of patient HIV019 contained a glycine (G) residue at position 25; a second variant in the ileum of patient HIV021 contained a threonine (T) instead of an alanine (A) at position 22. An analysis of tropism using three different aligorithms, the Web Position Specific Scoring Matrix (PSSM) tool (SINSI matrix), the Geno2pheno and the CoRSeqv3C co-receptor predictor aligorithms indicated that 100%, 92.4% and 83.1% of all variants in the GIT of patients HIV019 (n = 72), HIV021 (n = 66) and HIV029 (n = 65), respectively, were CCR5-tropic and non-synctium inducing (NSI). The most sensitive of method, CoRSeqv3C, identified 5 CXCR4-tropic variants in the duodenum (n = 1), left colon (n = 1) and right colon (n = 3) of patient HIV021. These variants included the 4 viral quasispecies with a basic amino acid substitution in the tetrameric GPGQ tip of the V3 loop and a fifth variant that had a basic (arginine) residue instead of a cysteine at position 1 of the V3 loop. A total of 11 CXCR4 tropic variants were identified in samples obtained from the duodenum (n=1), jejunum (n = 5), ileum (n =1) and plasma (n = 3) of patient HIV029 using the CoRSeqv3C method. All of these variants contained a basic amino acid substitution (lysine, K) at position 32 of the V3 loop (n=9) or in the V3 loop. No CXCR4-using variants were detected in patient HIV019 using any of the 3 testing algorithms.

3.3. DISCUSSION

There are still many unanswered questions relating to the role of the GIT in the genetic evolution and dynamics of HIV-1 infection. It has also been difficult to compare the data generated in different studies due to inter-study differences in sampling sites and the stage of HIV-1 infection, as well as the confounding effects of different treatment regimens. Most studies of genetic variation in the GIT have been performed on HIV-1 proviral DNA rather than actively replicating RNA. DNA contains a significant amount of archived and defective genetic material that has accumulated over time and thus, may not provide accurate information on contemporaneous viral sequences. To our knowledge, this is the first study to examine the genetic heterogeneity of HIV-1 RNA in the intestine of well-characterized African patients with AIDS, chronic diarrhoea and (or) weight loss. Despite the small sample size, consistent and statistically significant differences were observed in the trafficking, heterogeneity and evolution of HIV-1 RNA quasispecies in different compartments of the

GIT. The study, conducted at five different sites in the GIT of ART-naïve patients, may help explain some of the apparent discrepancies in the current literature.

Consistent with a pre-HAART study of proviral DNA by **van Marle *et al.*, 2007**, I found that different HIV-1 RNA quasispecies populate different regions of the GIT and that genetic diversity and patterns of viral trafficking and clustering were dependent on the region of the viral genome selected for study. Of interest in this study, was the higher overall diversity of *env* compared to the *gag* gene region and the higher level of *env* variability in the small intestine compared to the right and left colon. In contrast, no statistically significant differences were observed in *gag* diversity among different regions of the small (duodenum, jejunum and ileum) or large (right and left colon) intestine. Van Marle and his colleagues found that, compared to the *RT* gene region, *nef* DNA quasispecies were more diverse and exhibited higher levels of diversity in the colorectum than in the esophagus, stomach and duodenum [19]. Taken together, these findings suggest that different genes are under different selection pressures and that these pressures vary in different parts of the GIT. The factors contributing to HIV-1 diversity in the GIT have not been clearly defined but may include site-specific differences in immune activation, the specificity of the $CD8^+$ CTL cell response and factors related to cell tropism and virus entry, all of which would be expected to have a greater impact on *env* than on the *gag* encoding gene region. In this study I observed that *env* diversity was greatest in tissues where there was a large influx of HIV-1 RNA variants from other tissue compartments and that it was unexpectedly low in tissues where the majority of the variants were present in a single large cluster (ie. the jejunum of patient HIV021 and the right and left colon of patient HIV029).

In addition to differences in viral diversity, the phylogenetic trees constructed for the *env* and *gag* encoding gene regions were clearly distinct. *Gag* trees showed a pattern of extensive dissemination of HIV-1 RNA variants, both between different regions of the GIT (duodenum, jejunum, ileum, right colon and left colon) and between the GIT and plasma. This pattern that is suggestive of widespread dissemination of newly produced viral particles and (or) infected cells between tissue compartments. *Env* trees, on the other hand, exhibited a more complex pattern containing large clusters of closely related HIV-1 RNA quasispecies, in addition to mixed clustering of sequences from different gut compartments. Although there was little evidence of completely independent tissue-specific evolution, the tight clustering of *env* quasispecies, especially in the colon, was suggestive of distinct foci of virus production

that could contribute to the emergence of new viral variants with altered pathogenicity. These findings suggest that sequencing of multiple genes may be required to fully understand the dynamics of HIV-1 infection in the GIT. Analysis of a conserved gene may be needed to obtain accurate information on trafficking; while analysis of a variable gene that is under strong selection pressure is more likely to be more informative with respect to virus evolution and compartmentalization. The observed clustering env sequences in the small intestine of some patients suggests that it may also be important to sequence viral quasispecies from the upper, as well as, the lower GIT. The reason for the more consistent clustering of *env* quasispecies in the upper relative to the lower intestine is not known. We have previously reported that, compared to the small intestine, the colon has more $CD4^+$ T target cells, higher levels of viral replication and higher proportion of $CD8^+$ T cells that are ART responsive and are presumed to be anti-HIV-1 specific (as discussed in Chapter 4). These conditions would be expected to error prone replication and promote genetic evolution.

Analysis of inferred amino acid sequences revealed that some of the nucleotide polymorphisms were associated with altered protein function. As expected in ART-naïve patients, no primary drug resistance mutations were detected at sites of drug binding or at active sites of the RT and protease enzymes. However, using population-based sequencing of *pol* RT PCR products, a method that is commonly used to screen for resistance in plasma, we detected a number of accessory polymorphisms that have been linked to altered responsiveness to ART. As reported for primary resistance mutations in pre-HAART patients [20] and HAART-treated patients with plasma viral loads >1,000 HIV-1 RNA copies/mL [14], some RT and protease polymorphisms were differentially distributed in different sites of the GIT; others were distributed across all tissue compartments including plasma. Although population sequencing may fail to detect minor variants in the virus population, our results suggest that there are differences in the predominate variant(s) of the *pol* gene in different GIT sites. Carefuly designed longitudinal studies will be required to determine whether these variants play a role in the emergence of drug resistance.

Although HIV-1 subtype C viruses generally use the CCR5 chemokine receptor (along with CD4) for virus entry into susceptible target cells, genotypic changes can occur that allow C viruses to use CXCR4 as an entry receptor. The switch to CXCR4 usage is associated with accelerated $CD4^+$ T cell depletion and progression to AIDS [31; 32]. For this reason, and because testing for co-receptor usage is recommended before prescribing a

chemokine receptor blocker [33], we determined the co-receptor genotype of our V3 loop quasispecies using three different bioinformatics programmes. I detected a number of polymorphisms involving positively charged amino acid substitutions in the V3 loop that are known to play a role in the switch to CXCR4 co-receptor usage [34]. These CXCR4-using quasispecies were detected in 2 of 3 patients and, as observed for *pol* sequences, were differentially distributed in different GIT sites. As for *pol*, longitudinal studies will be required to determine whether these CXCR4 variants contribute to HIV-1 disease pathogenesis.

To my knowledge this is the first genetic study of actively replicating HIV-1 RNA quasispecies in the intestine of patients infected with HIV-1 subtype C, a viral subtype that is particularly well adapted to HIV-1 replication in the GIT [35; 36]. Strengths of my study relate to homogeneity of the study cohort with respect to stage of disease, the focus on viral RNA rather than proviral DNA and the sampling of multiple tissue sites. Limitations relate to the small sample size and the potential biases introduced by population-based or "bulk" PCR cloning (i.e. Taq induced recombination, nucleotide misincorporation) [44; 45]. However, this does not appear to be a problem when analyzing samples obtained from ART-naïve patients with a high HIV-1 copy number, as in the current study [46].

In contrast to ART-naïve and pre-HAART patients [19; 40], and viremic patients on HAART [14], recent studies of patients on highly suppressive HAART have reported a lack of HIV-1 evolution and compartmentalization in the GIT and suggested that the GIT is not a source of rebound virus during therapeutic failure [16-18,41]. The reason(s) for the failure to detect virus evolution in patients on fully suppressive HAART is not known but may be related to the low levels of viral replication, or to $CD4^+$ T cell recovery and decreased activation of latently infected cells. Alternatively, virus transmission from long-lived latently infected cells may be occurring *via* a cell-cell route that does not involve exposure to the immune system [42]. Of interest in this regard, is a study by **Avettand-Fenoel *et al.*, 2011**; showing that in ART-naïve patients rectal variants of HIV-1 *env* DNA were more diverse than variants in peripheral blood. However, in patients on ART, there was an "equalization" of the genetic diversity between these two compartments. I found that inter-compartmental trafficking between different sites in the GIT, and between the GIT and plasma, was an important contributor to HIV-1 diversity in different compartments of gut. A dramatic reduction in plasma HIV-1 RNA, to <50 copies/mL, would be expected to contribute to this

equalization process by altering the dynamics of virus migration into, and (or) out of, the GIT.

An improved understanding of the mechanisms driving the replication, evolution and trafficking of HIV-1 RNA variants between different tissue compartments is needed to better understand HIV-1 pathogenesis and design improved strategies for the monitoring of therapeutic efficacy.

3.4. CONCLUSIONS

Studies of HIV-1 genetic diversity in the GIT have, thus far, been inconclusive and confounded by differences in methodology and study design and in the duration and administration of ART. Our study, conducted in a relatively homogeneous cohort of African AIDS patients with diarrhea and/or weight loss, confirms and extends studies of HIV-1 DNA diversification and dissemination in the intestine of ART-naïve and pre-HAART patients. Additional longitudinal studies of ART-naïve patients are needed to better understand the dynamics of HIV-1 diversification in the GIT, and to serve as a baseline for assessing novel therapeutic strategies aimed at controlling, or possibly eliminating, viral replication and genetic evolution in the GIT.

3.5. FIGURES LEGENDS

Figures:

Figure 3.1: Mean overall pairwise distances among rectal HIV-1 RNA variants in the GIT of ART-naïve patients with HIV-1/AIDS. *Env* (A) and *gag* (B) HIV-1 RNA sequences were more diverse [higher mean total distance (d)] in the GIT of patient 21 compared to patients 19 and 29. (C) Significant differences were also detected within each patient. For each of the 3 patients studied, the mean pairwise distances between *env* HIV-1 RNA clones in the GIT was greater than the distances calculated for the *gag* encoding region. Genetic distances between viral clones were computed using the Kimura model. Mean values (horizontal line) and statistical significance are indicated (*p<0.05, p**<0.01, p***<0.001).

Figure 3.2: Mean pairwise distances among HIV-1 RNA variants within different compartments of the GIT. *Env* and *gag* RT-PCR products derived from the duodenum ((DUO), jejunum (JEJ), ileum (ILE), right colon (RC) and left colon (LC)] of patients 19, 21 and 29 were cloned, sequenced and analyzed for genetic diversity using the Kimura model. With few exceptions (the jejunum of patient 21 and the right colon of patient 29), *env* HIV-1 RNA sequences were more diverse than *gag* HIV-1 RNA variants derived from the same tissue compartment. Mean values (horizontal line) and statistical significance are indicated (*p<0.05, p**<0.01, p***<0.001).

Figure 3.3: Representative phylogenetic tree showing patient-specific clustering of HIV-1 pol sequences amplified from different regions of the GIT. All of the clusters identified by the Bayesian approach were confirmed by Maximum Likelihood. Patient specific clusters and IDs (HIV008, HIV010, HIV011, HIV014, HIV015, HIV019, HIV021, HIV027, HIV029 and HIV031) are color coded and indicated by the brackets on the right side of the tree.

Figure 3.4: Representative Bayesian tree showing the phylogenetic relationships among C2V5 envelope sequences derived from different tissue compartments in the GIT of patient HIV019. The tree was constructed using a general time reversible (GTR) with a likelihood of 3281.4 and a γ-distribution α value of 5929.5. In addition to extensive trafficking and exchange of HIV-1 RNA variants among different sites in the GIT, especially in the small intestine, clustering of *env* sequences by tissue compartment was observed within the duodenum, right and left colon suggesting localized evolution of HIV-1 RNA quasispecies within these different tissue sites. All of these tissue-specific clusters were confirmed by Maximum Likelihood analysis (**Appendix A**). The tree is color coded as follows: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right colon (R), yellow; left colon (L), dark blue.

Figure 3.5 Representative Bayesian tree showing the phylogenetic relationships among C2V5 envelope sequences derived from different tissue compartments in the GIT of patient HIV021. The tree was constructed using a GTR with a likelihood >5103 and a γ-distribution α value of 3411.5. Of particular note was the close relationship between HIV-1 RNA variants in the ileum and the colon suggesting that there was extensive exchange of viral quasispecies between compartments in the lower GIT. Also of note was the clustering of *env* sequences by tissue compartment within the jejunum and within the right and left colon of this patient, a finding that is suggestive of localized evolution of HIV-1 RNA variants in

these compartments of the GIT. All clusters were confirmed by Maximum Likelihood analysis (**Appendix B**). The tree is color coded as follows: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right colon (R), yellow; left colon (L), dark blue.

Figure 3.6: Representative Bayesian tree showing the phylogenetic relationships among C2V5 envelope sequences derived from different tissue compartments in the GIT of patient HIV029. The tree was constructed using a GTR with a likelihood > 2704.6 and a γ-distribution α value of 1392.3. In addition to extensive inter-mixing of HIV-1 RNA variants between different compartments, especially in the small intestine and between the small intestine and plasma, distinct clusters of closely related sequences were detected in the left and right colon and the dudodenum. As shown in **Appendix C,** these clusters were confirmed by Maximum Likelihood analysis. The tree is color coded as follows: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right colon (R), yellow; left colon (L), dark blue.

Figure 3.7: Representative Bayesian tree showing extensive inter-mixing of *gag* HIV-1 RNA sequences among different compartments in the GIT of patient HIV019. The tree was constructed using a GTR with likelihood >336 and a γ-distribution α value of 3420.2. A similar pattern of viral dispersion was observed in trees constructed using the Maximum Likelihood method (**Appendix D**). The tree is color coded as follows: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right colon (R), yellow; and left colon (L), dark blue.

Figure 3.8: Representative Bayesian tree showing extensive inter-mixing of *gag* HIV-1 RNA sequences among different compartments in the GIT of patient HIV021. Tree construction was performed using a GTR with a likelihood >3666.3 and a γ-distribution α value of 4217.6. Color codes are as follows: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right colon (R), yellow; and left colon (L), dark blue. Trees constructed by the Maximum Likehood method showed a similar pattern of extensive inter-mixing of *gag* quasispecies among different tissue compartments (**Appendix E**).

Figure 3.9: Representative Bayesian tree showing extensive inter-mixing of *gag* HIV-1 RNA quasispecies among different compartments in the GIT of patient HIV029. The tree was constructed using a GTR with a likelihood >5473.2 and a γ-distribution α value of 3037.3. Color codes are: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right

colon (R), yellow; left colon (L), dark blue and plasma (PLA), light blue. A similar pattern of viral dispersion was observed in trees constructed using the Maximum Likelihood method (**Appendix F**).

Figure 3.10: Amino acid substitutions in the V3 loop of the envelope gene of patients HIV019, HIV021 and HIV029 relative to the consequence sequence for subtype C viruses (A) and to the subtype B reference strain, HXB2 (B).

Tables:

Appendix:

Appendix A: Representative Maximum likelihood BioNJ tree showing the phylogenetic relationships among C2V5 envelope sequences derived from different tissue compartments in the GIT of patient HIV019. The tree was constructed using the Subtree Pruning and Regrafting (SPR) with a likelihood of 2853.9 and a gamma shape parameter value of 0.188. In addition to extensive trafficking and exchange of HIV-1 RNA variants among different sites in the GIT, especially in the small intestine, clustering of *env* sequences by tissue compartment was observed within the duodenum, right and left colon suggesting localized evolution of HIV-1 RNA quasispecies within these different tissue sites. The tree is color coded as follows: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right colon (R), yellow; and left colon (L), dark blue.

Appendix B: Representative Maximum Likelihood tree showing the phylogenetic relationships among C2V5 envelope sequences derived from different tissue compartments in the GIT of patient HIV021. The tree was constructed using the Subtree Pruning and Regrafting (SPR) with a likelihood of 2470.4 and a gamma shape parameter value of 0.274. Of particular note was the close relationship between HIV-1 RNA variants in the ileum and the colon suggesting that there was extensive exchange of viral quasispecies between compartments in the lower GIT. Also of note was the clustering of *env* sequences by tissue compartment within the jejunum and within the right and left colon of this patient, a finding that is suggestive of localized evolution of HIV-1 RNA variants in these compartments of the GIT. The tree is color coded as follows: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right colon (R), yellow; and left colon (L), dark blue.

Appendix C: Representative Maximum Likelihood tree showing the phylogenetic relationships among C2V5 envelope sequences derived from different tissue compartments in the GIT of patient HIV029. The tree was constructed using the SPR with a likelihood of 2668.3 and a gamma shape parameter value of 0.171. In addition to extensive inter-mixing of HIV-1 RNA variants between different compartments, especially in the small intestine and between the small intestine and plasma, distinct clusters of closely related sequences were detected in the left and right colon and the duodenum. The tree is color coded as follows: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right colon (R), yellow; left colon (L), dark blue and plasma (PLA), light blue.

Appendix D: Representative Maximum Likelihood tree showing extensive inter-mixing of *gag* HIV-1 RNA quasispecies among different compartments in the GIT of patient HIV019. The tree was constructed using the SPR with a likelihood of 4775.4 and a gamma shape parameter value of 0.081. The tree is color coded as follows: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right colon (R), yellow; and left colon (L), dark blue.

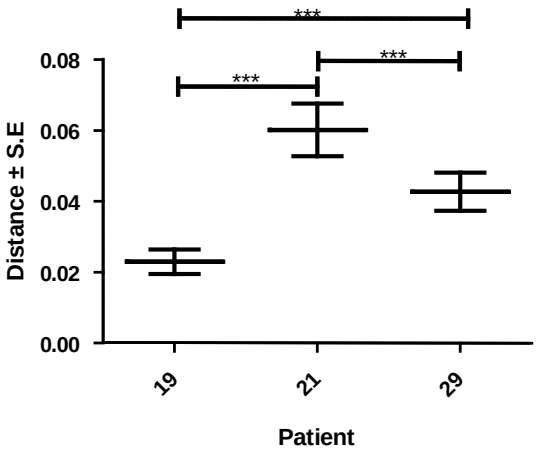
Appendix E: Representative Maximum Likelihood tree showing extensive inter-mixing of *gag* HIV-1 RNA quasispecies among different compartments in the GIT of patient HIV021. Tree construction was performed using the SPR with a likelihood of 6145.8 and a gamma shape parameter value of 0.215. Color codes are as follows: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right colon (R), yellow; and left colon (L), dark blue.

Appendix F: Representative Maximum Likelihood tree showing extensive inter-mixing of *gag* HIV-1 RNA quasispecies among different compartments in the GIT of patient HIV029. The tree was constructed using the SPR with a likelihood of 4477.3 and a gamma shape parameter value of 0.015. Color codes are: jejunum (J), purple; duodenum (D), red; ileum (I), light blue; right colon (R), yellow; left colon (L), dark blue and plasma (PLA), light blue.

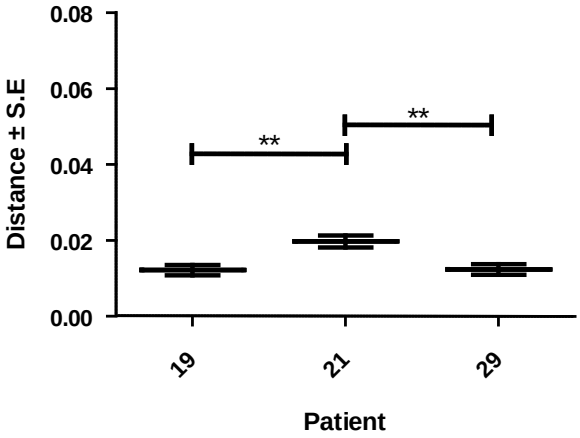
3.6. FIGURES AND TABLES:

Figure 3.1: A, B, and C.

A. *env*



B. *gag*



C. *env* versus *gag*

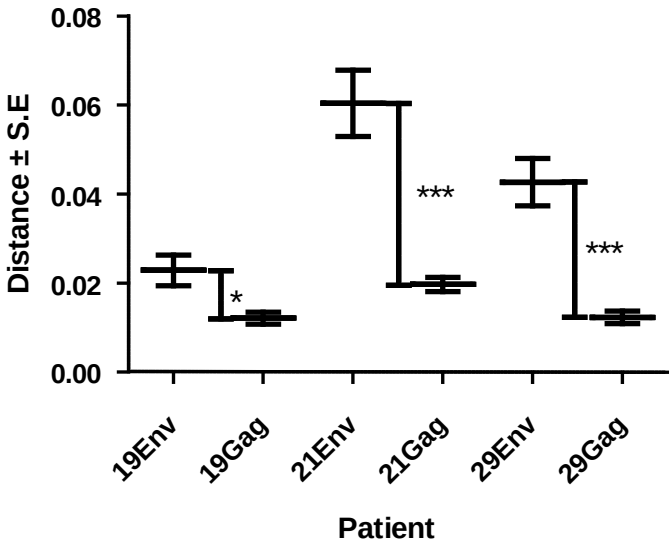
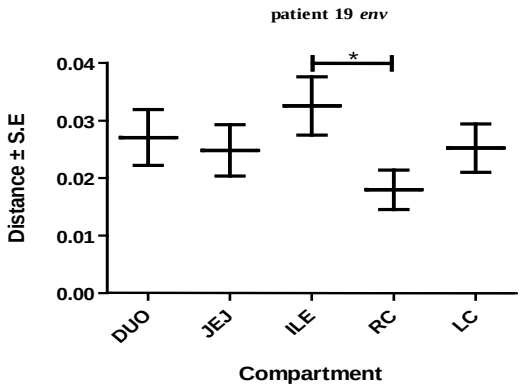
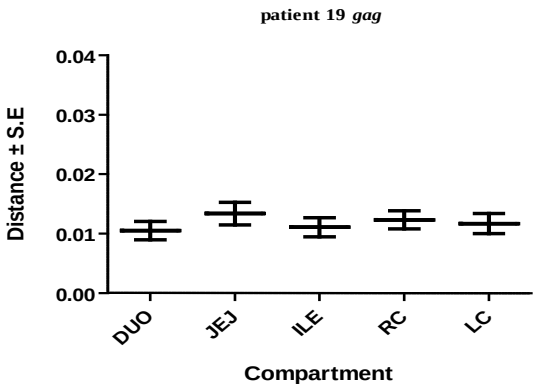


Figure 3.2: A, B, C, D, E, and F

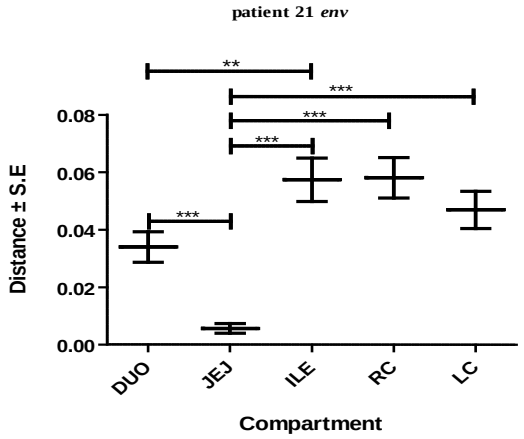
A.



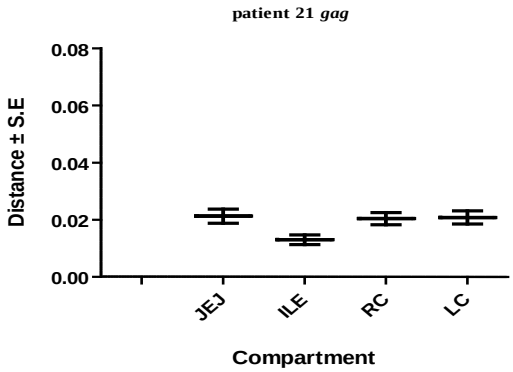
B.



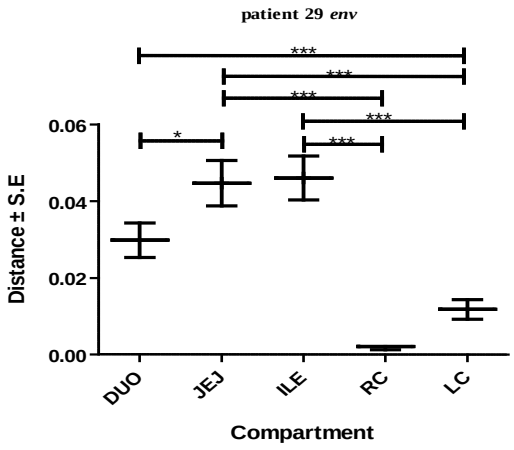
C.



D.



E.



F.

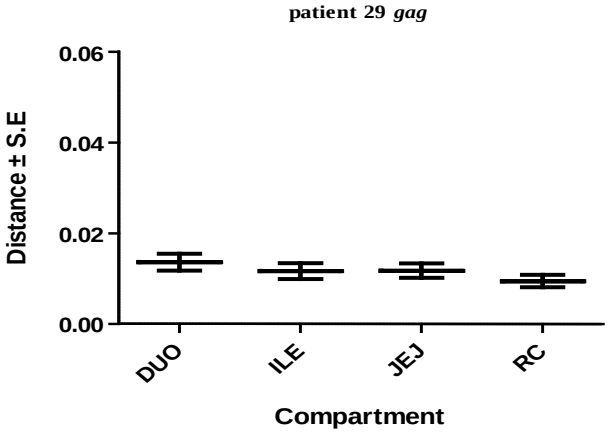


Figure 3.3:

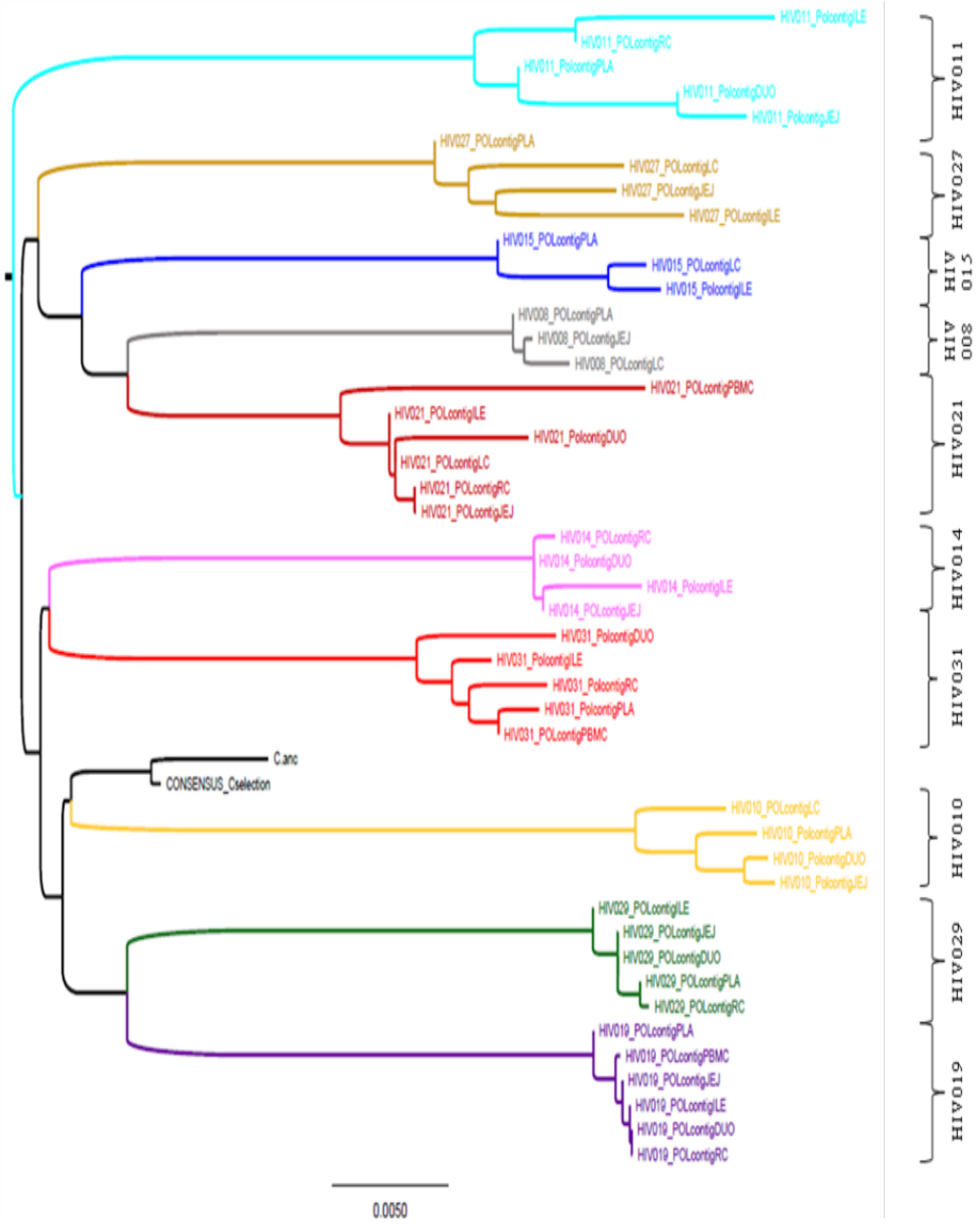


Figure 3.4.:

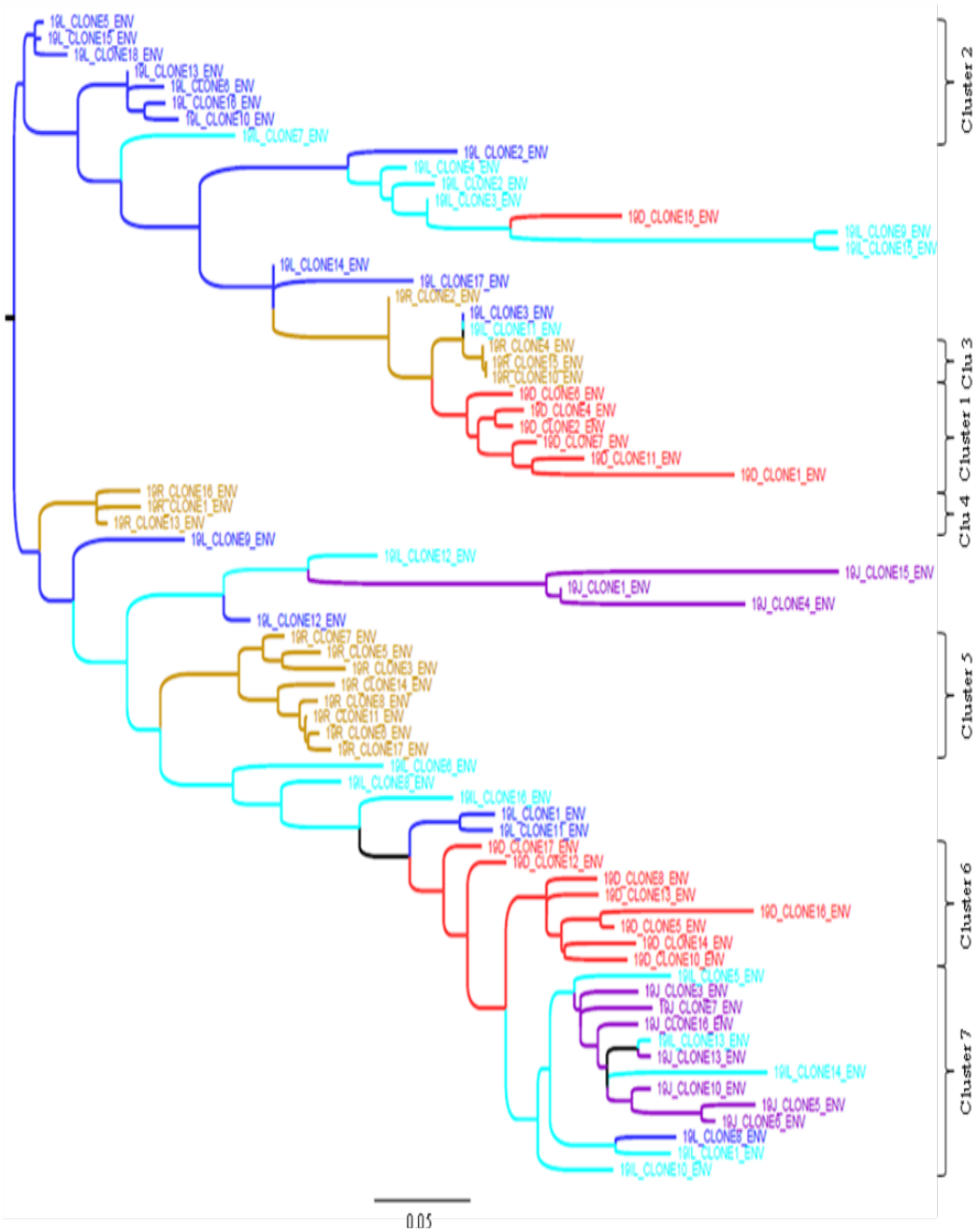


Figure 3.5.:

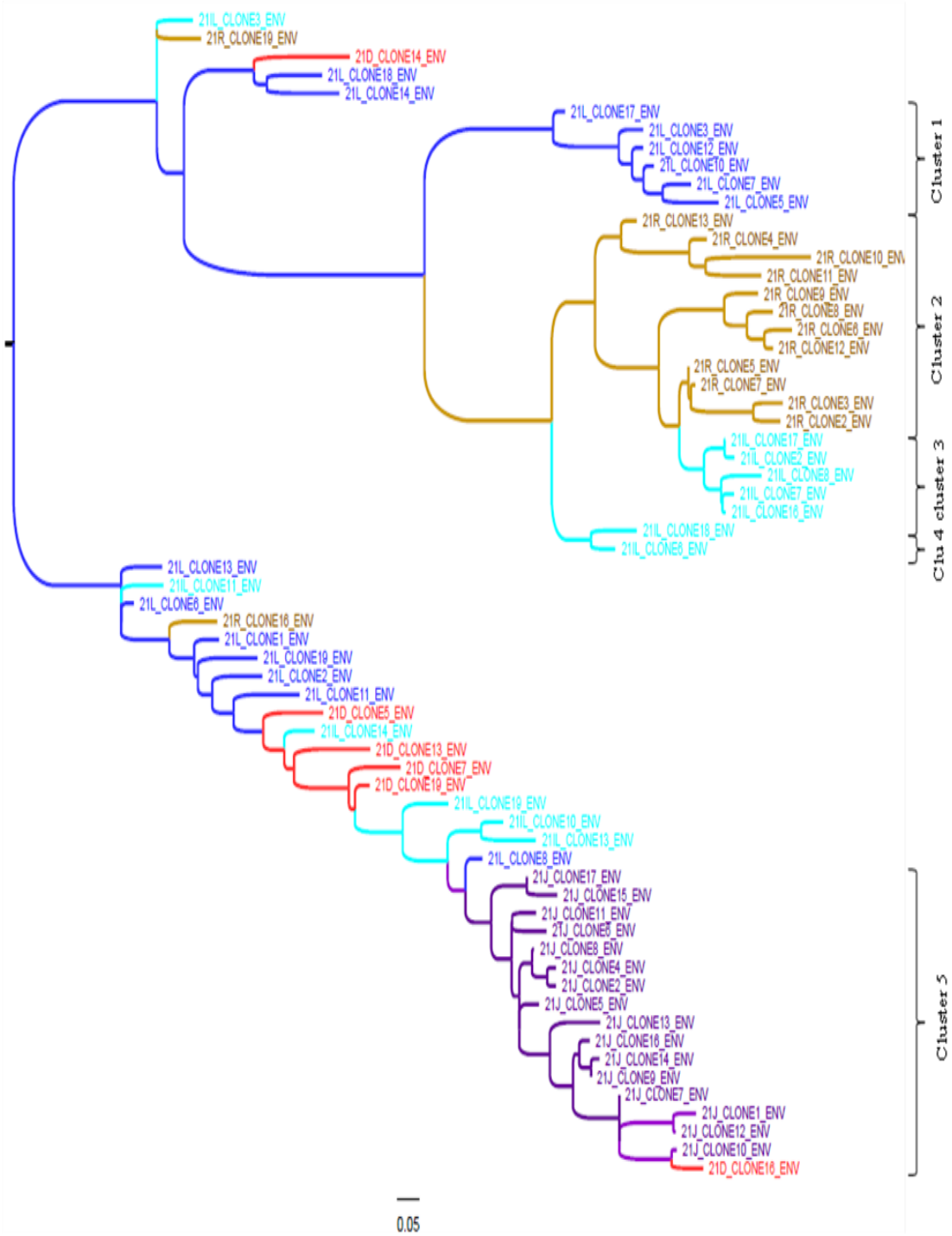


Figure 3.6.:

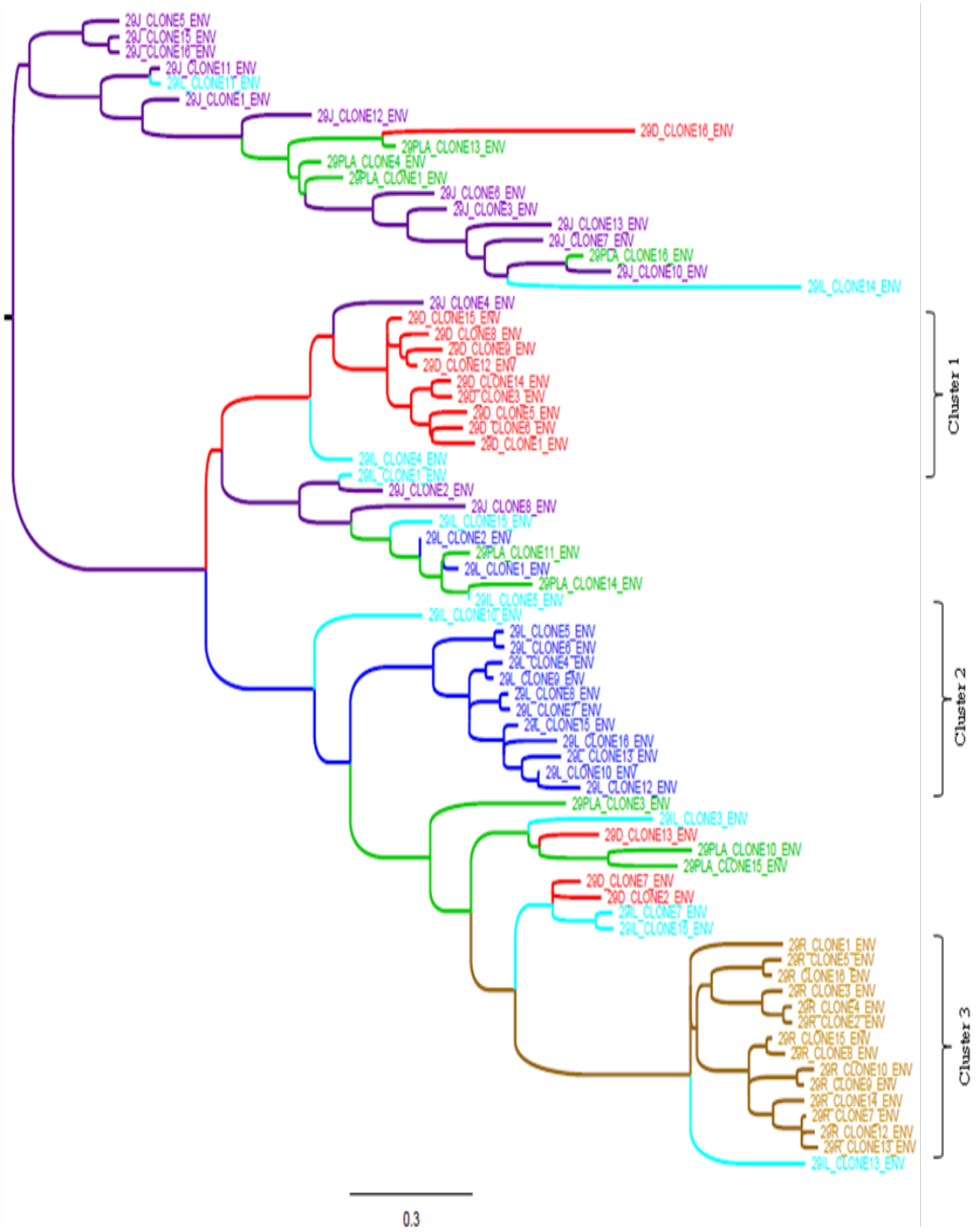


Figure 3.7:

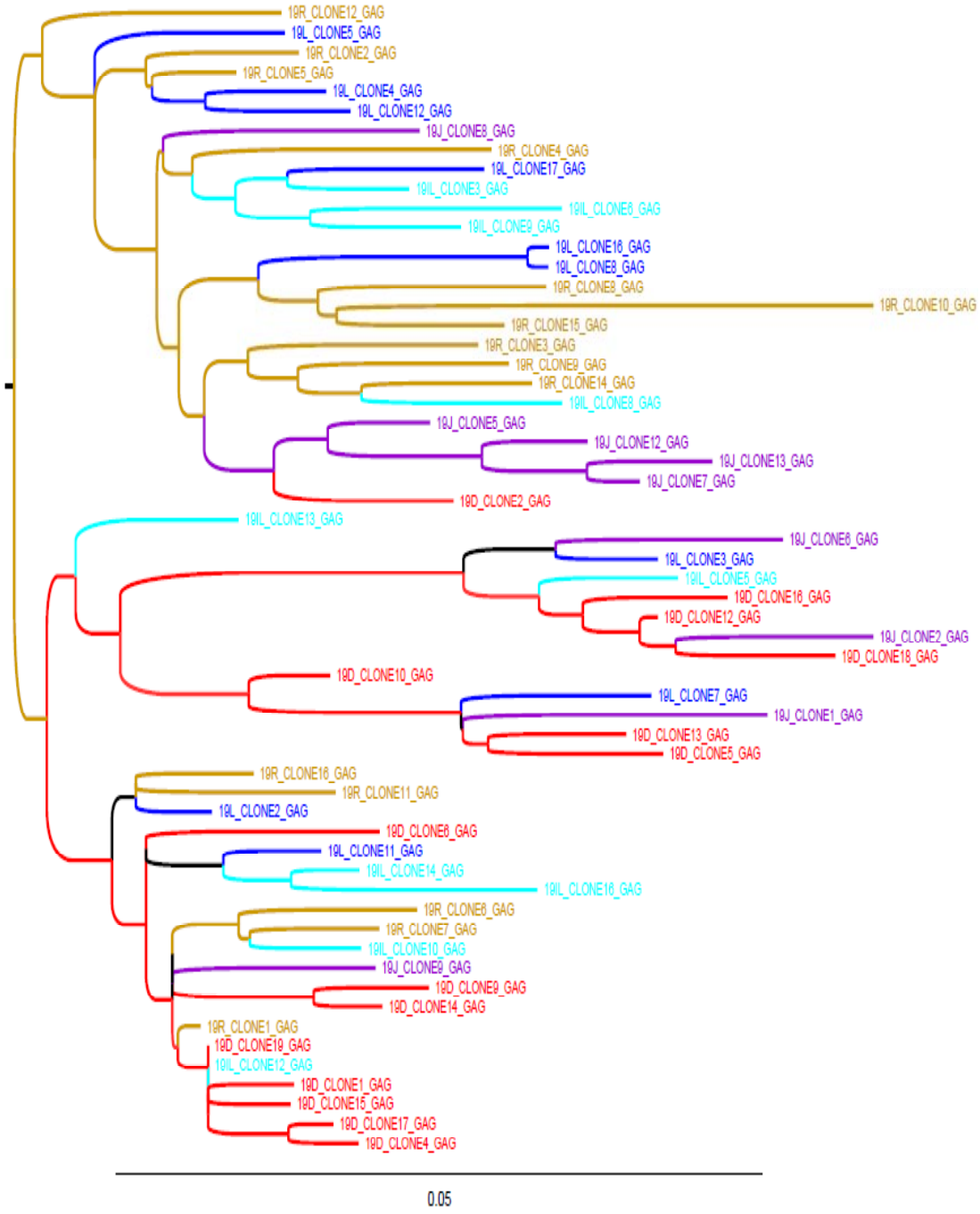


Figure 3.8:

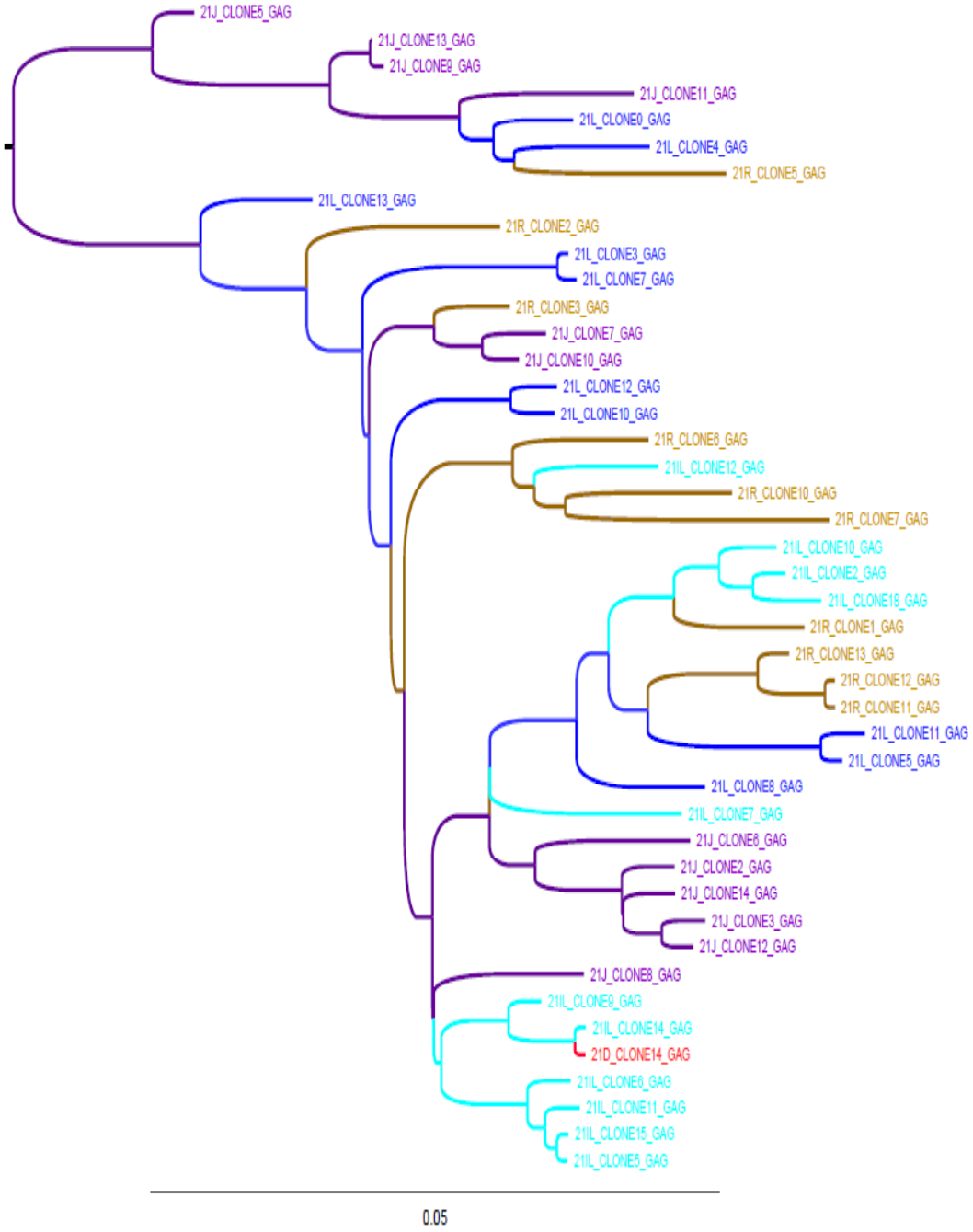
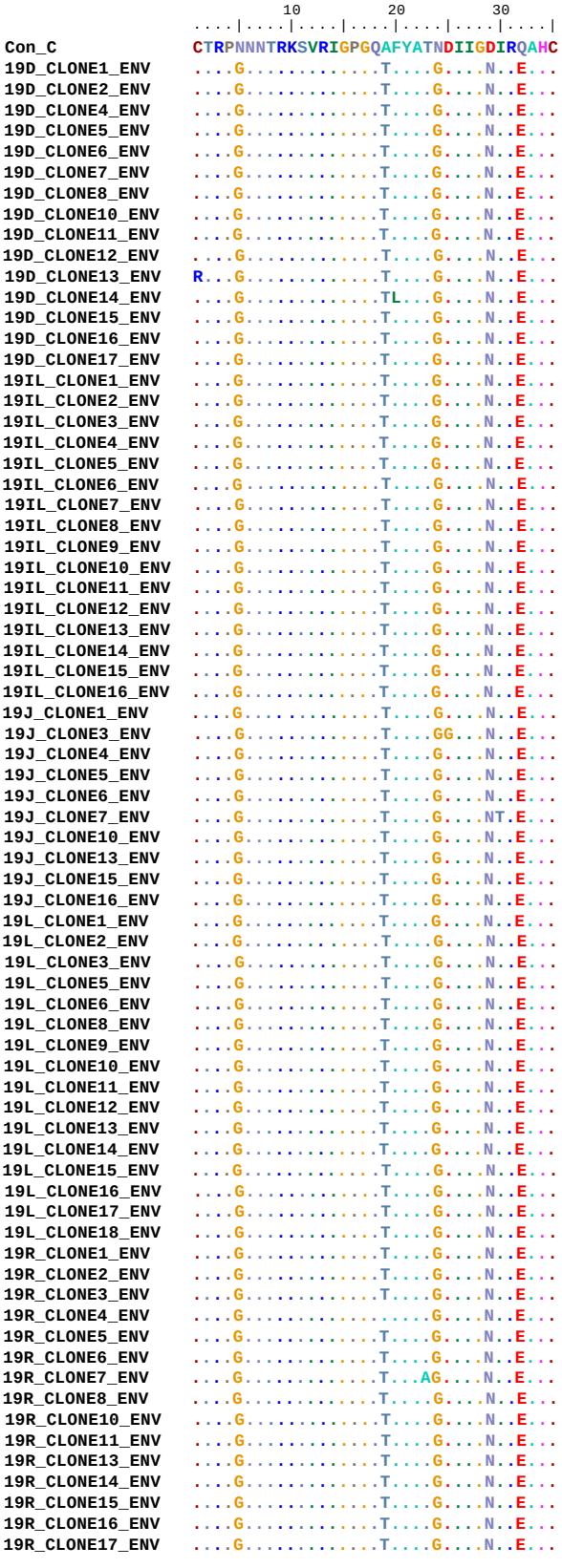


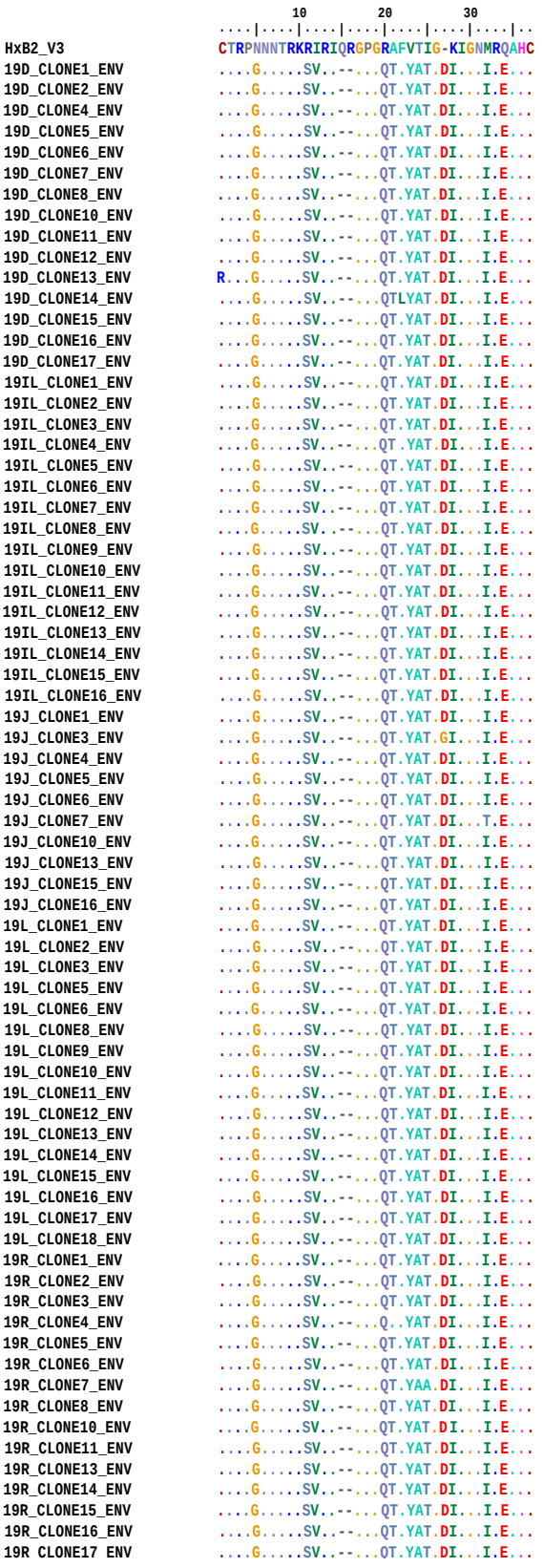
Figure 3.9:

Figure 3.10.: A, B, and C:

A. Consensus C

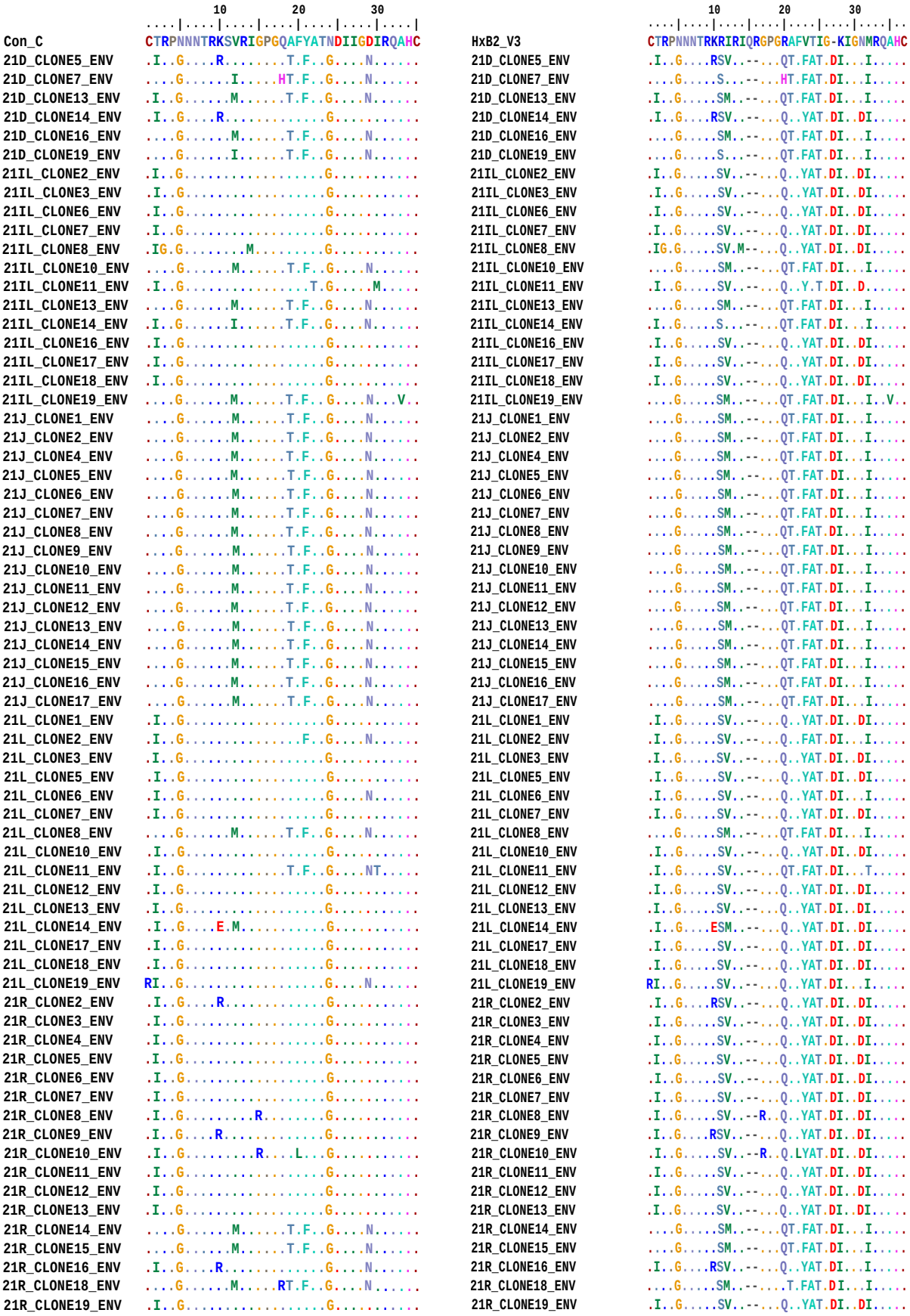


HxB2

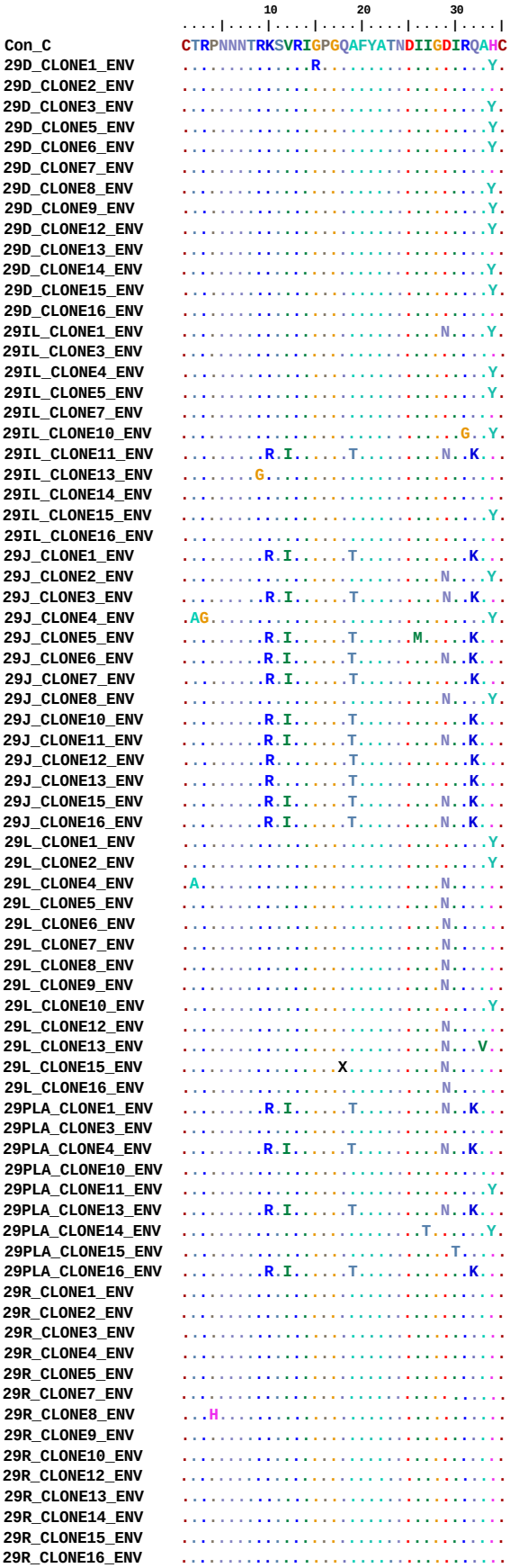


B. Consensus C

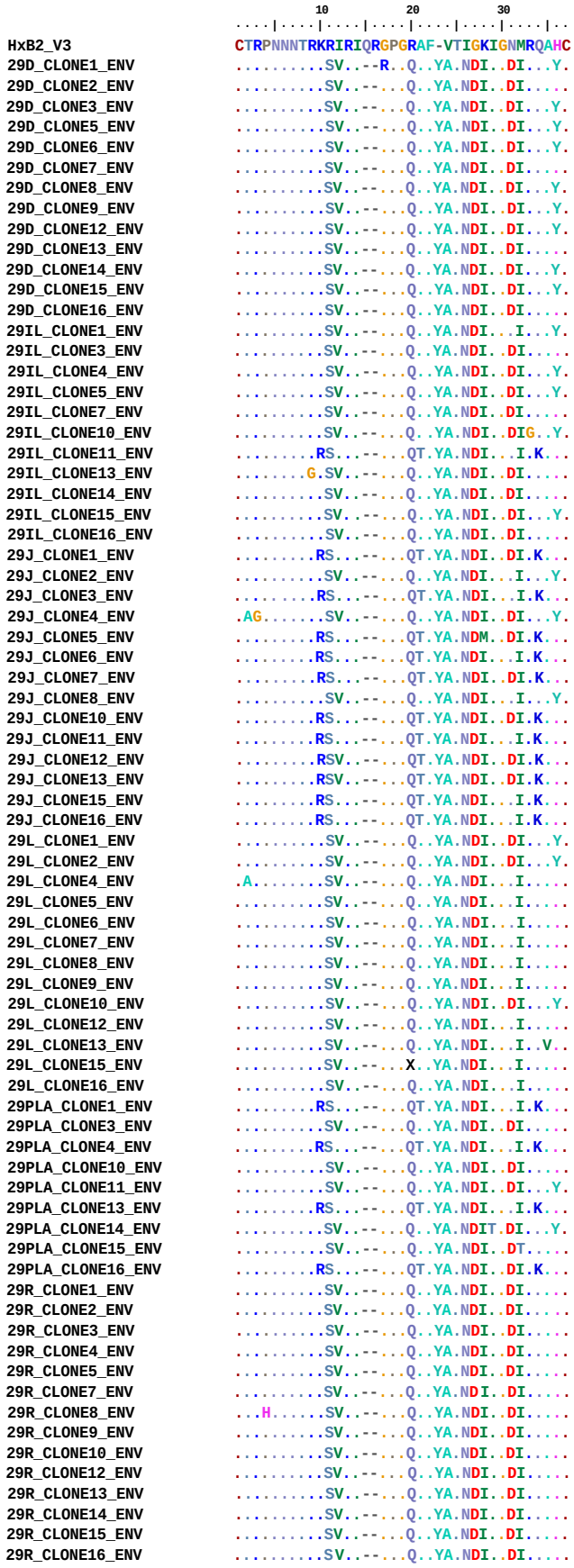
HxB2



C. Consensus C



HxB2



3.7. TABLES:

Table 3.1.: Clinical and Demographic Characteristics of the HIV-1^{+} Study Population

Gender	Male	5
	Female	5
Age (years)	Mean ± S.D.	37 ± 10.61
	Median (Range)	33 (24 - 58)
$CD4^{+}$ T cell count (cells/μL)	Mean ± S.D.	102 ± 65
	Median (Range)	104 (6 - 181)
Percentage $CD4^{+}$ T cells	Mean ± S.D.	9.7 ± 6.8
	Median (Range)	7.1 (0.78 – 22.7)
Plasma VL (HIV-1 RNA copies/mL)	Mean ± S.D.	258801 ± 375428
	Median (Range)	104000 (410 - 1100000)
Plasma VL (log_{10} HIV-1 RNA copies/mL)	Mean ± S.D.	4.65 ± 1.24
	Median (Range)	5.0 (1.24 – 6.04)
Diarrhoea	Absent	2
	Present	8
Weight loss	Absent	2
	<10%	6
	>10%	2

* All 10 patients had heterosexually-acquired HIV-1 subtype C infection.

Table 3.2.: Number of HIV-1 RNA quasispecies analyzed per patient and per compartment

Patient	**No. *env* HIV-1 RNA quasispecies analyzed per compartment**	**No. *gag* HIV-1 RNA quasispecies analyzed per compartment**	**Total no. HIV-1 RNA clones obtained**
HIV019	Duodenum 15 Jejunum 10 Ileum 16 Right Colon 15 Left Colon 16 Total : 72	Duodenum 15 Jejunum 9 Ileum 10 Right Colon 15 Left Colon 10 Total: 59	**131**
HIV021	Duodenum 6 Jejunum 16 Ileum 15 Right Colon 14 Left Colon 15 Total: 66	Duodenum 1 Jejunum 12 Ileum 11 Right Colon 10 Left Colon 10 Total: 44	**110**
HIV029	Duodenum 13 Jejunum 14 Ileum 11 Right Colon 14 Left Colon 13 Total: 65	Duodenum 11 Jejunum 11 Ileum 10 Right Colon 12 Left Colon 0 Total: 44	**109**
Total	**203**	**147**	**350**

Table 3.3: Molecular diversity of *env* and *gag* HIV-1 RNA in the GIT of patients 19, 21 and 29

		Mean % distances (SE)	
Patient	**Compartment**	**Envelope**	**Gag**
HIV019			
	Duodenum	2.70 (0.48)	1.05 (0.15)
	Jejunum	2.48 (0.45)	1.34 (0.19)
	Ileum	3.25 (0.51)	1.12 (0.16)
	Right colon	1.79 (0.34)	1.23 (0.15)
	Left colon	2.52 (0.42)	1.17 (0.17)
HIV021			
	Duodenum	3.39 (0.53)	ND
	Jejunum	0.56 (0.17)	2.12 (0.25)
	Ileum	5.73 (0.76)	1.29 (0.17)
	Right colon	5.80 (0.71)	2.03 (0.22)
	Left colon	4.68 (0.65)	2.07 (0.23)
HIV029			
	Duodenum	2.97 (0.45)	1.36 (0.19)
	Jejunum	4.46 (0.59)	1.17 (0.16)
	Ileum	4.60 (0.57)	1.16 (0.18)
	Right colon	0.20 (0.08)	0.94 (0.14)
	Left colon	1.17 (0.26)	ND
	Plasma	5.68 (0.72)	1.16 (0.16)

With the exception of the jejunum of patient 21 and right colon of patient 29, the *env* gene region was more diverse (higher mean total distance) than the *gag* encoding regions. *Env* sequences in the small intestine (duodenum, ileum) of patients 19 and 29 were also more diverse than those in the right and left colon. These observations suggest that different selection pressures (or selection factors) were acting on the *env* gene in different regions of the GIT. ND, not done.

Table 3.4: Frequency of Most Common Protease and Reverse Transcriptase Amino Acid Substitutions in the GIT and Plasma of Patients Infected with HIV-1 Subtype C.

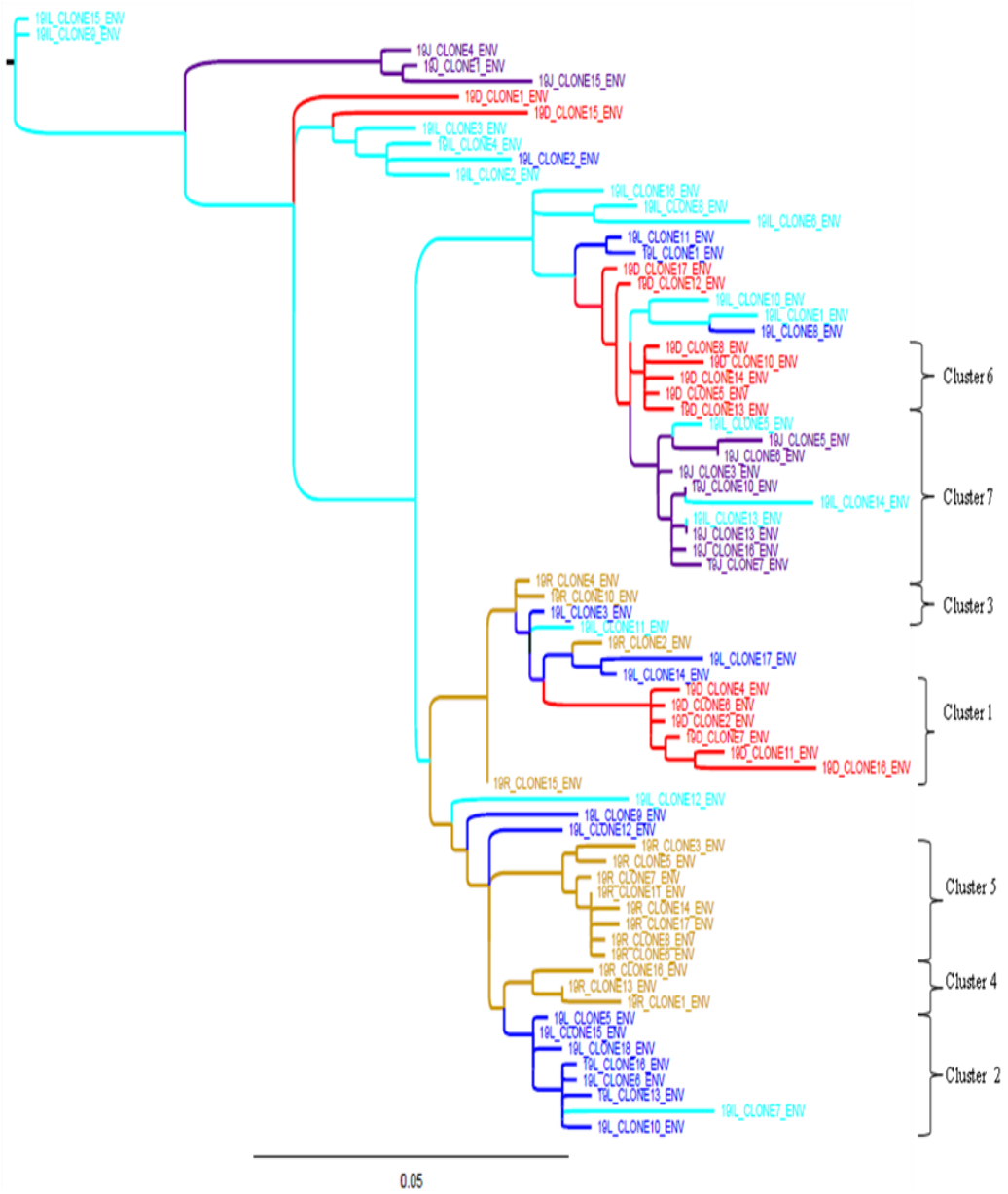
Protein	Amino acid	% of Samples					
		DUO	JEJ	ILE	RC	LC	PLA
Protease (PR)							
	H69K	100	100	100	100	100	100
	M36I[a]	100	88	100	100	100	100
	L89M[b]	100	100	100	83	80	75
	R41K[c]	86	75	75	83	80	88
	L19I/T[d]	100	88	88	100	80	88
	I93L	86	88	88	83	100	88
	I15V[e]	86	75	75	67	60	88
	T12S[f]	70	63	63	67	60	75
Reverse Transcriptase (RT)							
	V35T	100	100	100	100	100	100
	T39D/E/K/T	100	100	100	100	100	100
	K43D	100	100	100	100	100	100
	K122E	100	100	100	100	80	100
	K173A/T	100	100	100	100	100	100
	T200A	100	100	100	100	100	100
	Q207E/D/E/G/K/R/T	100	100	100	100	100	100
	R211A/E/G/K/N	86	100	100	100	100	100
	I293V	100	100	88	100	80	88
	E36A/E	100	75	50	67	60	63
	V292I	86	88	100	100	100	88
	V245E/K/Q/V	86	88	100	100	100	88
	A272P	86	88	88	83	100	88
	E291A/D/E	71	88	88	83	100	100
	D123D/G/N/S	57	63	50	83	20	50
	D177E	57	63	63	83	40	50
	T286A	43	38	38	50	40	25
	K277K/R/S	38	50	63	67	63	50

[a]Present as a mixed M36I/M residue in the duodenum of patient 11. [b]Present as a mixed L89L/M residue in patient 31; present in the duodenum but not in the right colon or plasma of patient 31; not detected in any GIT or plasma samples from patient 8. [c]Detected in the plasma but not in the jejunum, ileum or left colon of patient 31; not detected in samples from patient 27. [d]L19I/T was present in 5.9% of GIT and 12.5% of plasma samples. [e]Present as a mixed I15I/V residue in the duodenum of patient 10 and the ileum of patient 31; detected in the right colon and plasma but not in the duodenum or right colon of patient 31; not detected in samples from patient 14, or in the jejunum, left colon or plasma of patient 10. [f]Present as a mixed T12S/T residue in the ileum of patient 15.

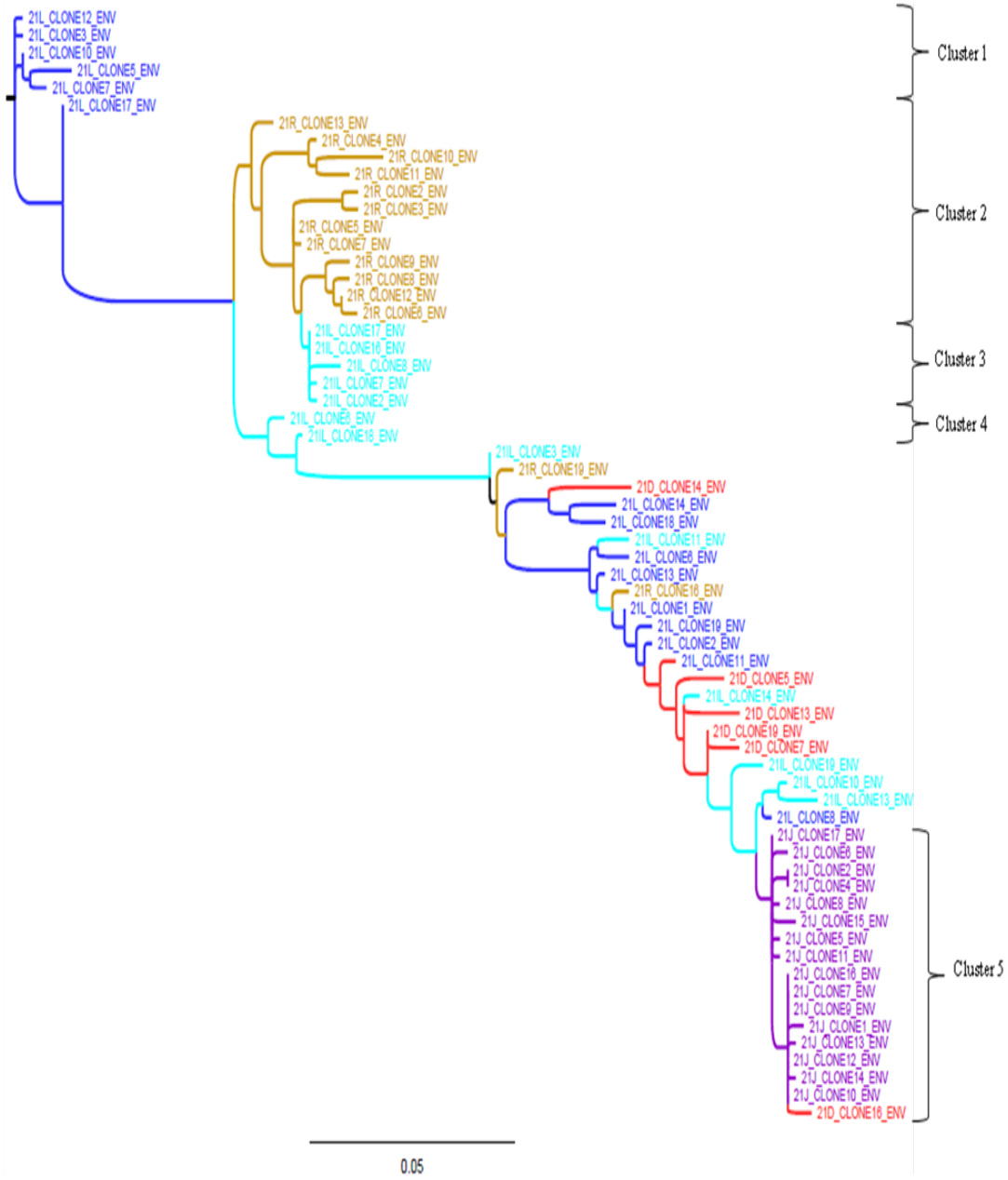
3.8. APPENDIX A TO F:

Appendix A

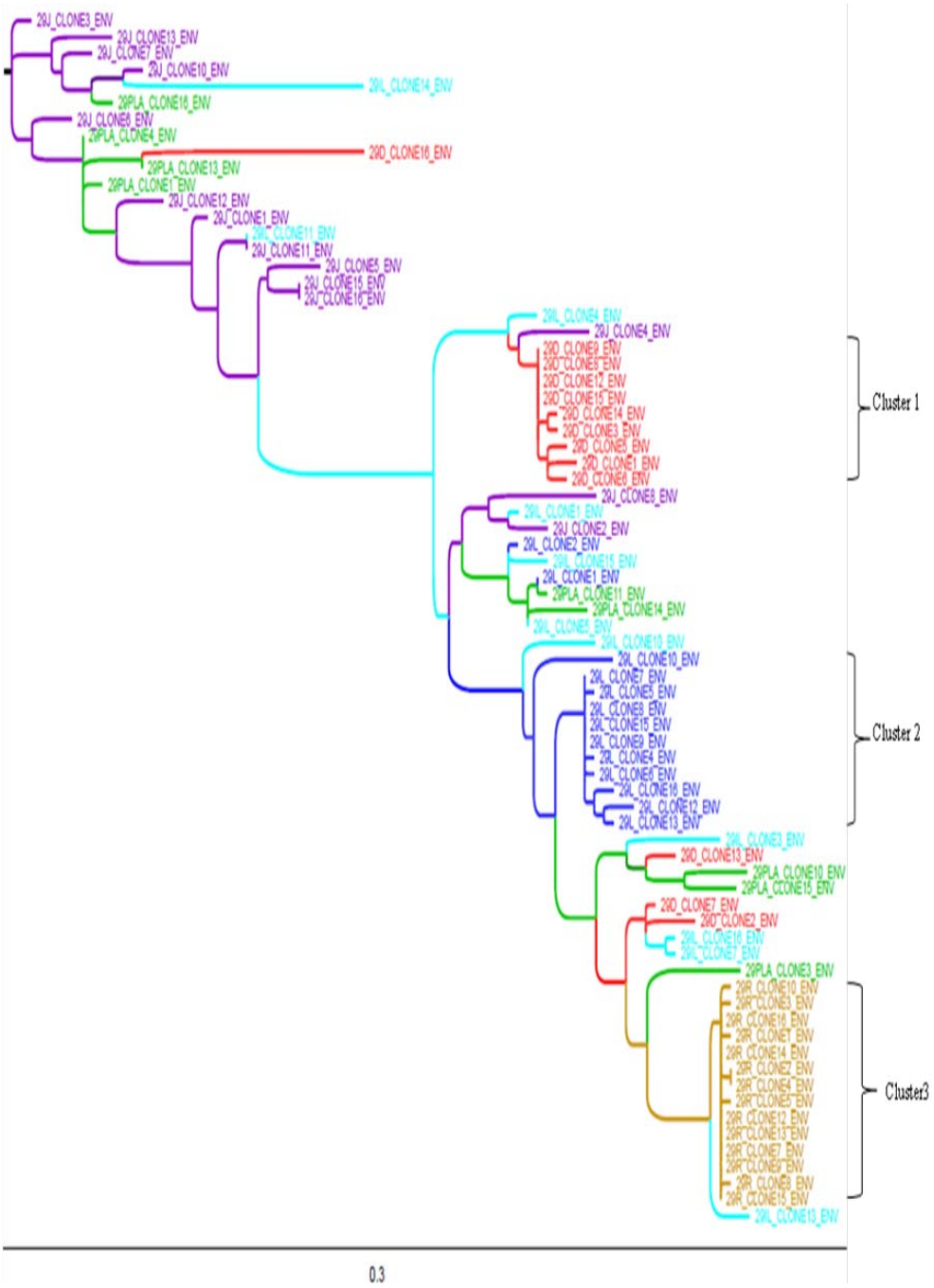
:



Appendix B:

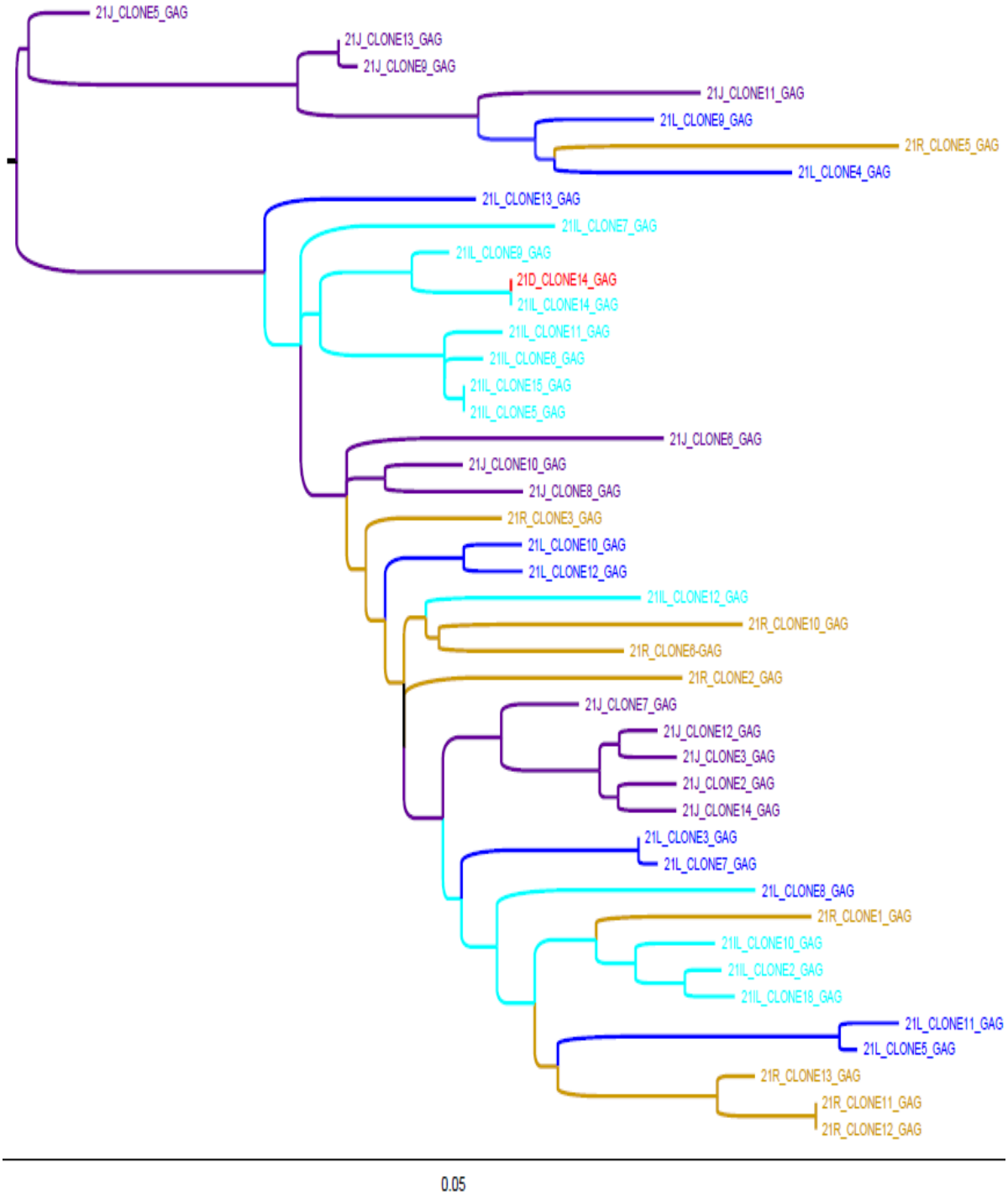


Appendix C:

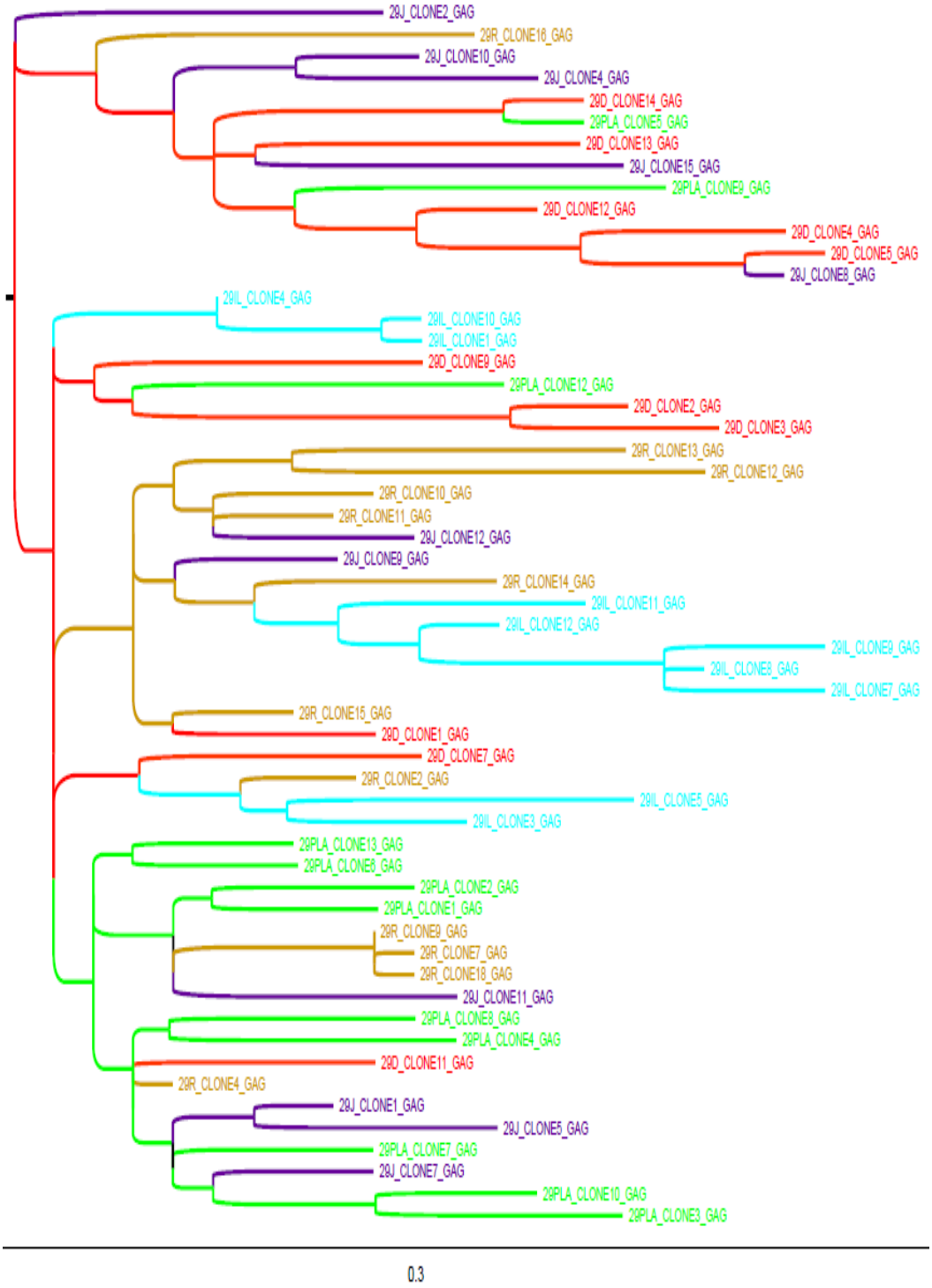


Appendix D:

Appendix E:



Appendix F:



3.9. REFERENCES

1. Veazey, R. S., De Maria, M., Chalifoux, L.V., Shvetz, D. E., Pauley, D. R., Knight, H. L., Rosenzweig, M., Johnson, R. P., Desrosiers, R. C., and Lackner, A. A. 1998. Gastrointestinal tract as a major site of CD4+ T cell depletion and viral replication in SIV infection. *Science* **280**: 427-431.
2. Guadalupe, M., Reay, E., Sankaran, S., Prindiville, T., Flamm, J., Mcneil, A., Dandekar, S. 2003. Severe CD4 T-cell depletion in Gut lymphoid tissue during primary Human Immunodeficiency Virus Type 1 infection and substantial delay in restoration following Highly Active Antiretroviral Therapy**.** *Society* **77**: 11708-11717.
3. Mehandru, S., Poles, M. A., Tenner-Racz, K., Horowitz, A., Hurley, A., Hogan, C., Boden, D., Racz, P., and Markowitz, M. 2004. Primary HIV-1 infection is associated with preferential depletion of $CD4^+$ T lymphocytes from effector sites in the gastrointestinal tract**.** *J Exp Med* **200**:761-770.
4. Brenchley, J. M., Schacker, T. W., Ruff, L. E., Price, D. A., Taylor, J. H., Beilman, G. J., Nguyen, P. L., Khoruts, A., Larson, M., Haase, A. T., and Douek, D. C. 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *J Exp Med* **200**:749-759.
5. Mattapallil, J. J., Douek, D. C., Hill, B., Nishimura, Y., Martin, M., and Roedere, M. 2005. Massive infection and loss of memory CD4C T cells in multiple tissues during acute SIV infection**.** *Nature* **434**:1093–1097.
6. Guadalupe, M., Sankaran, S., George, M. D., Reay, E., Verhoeven, D., Shacklett, B. L., Flamm, J., Wegelin, J., Prindiville, T., and Dandekar, S. 2006. Viral suppression and immune restoration in the gastrointestinal mucosa of human immunodeficiency virus type 1-infected patients initiating therapy during primary or chronic infection**.** *J. Virol.* **80**: 8236-47.
7. Centlivre, M., Sala, M., Wain-Hobson, S., and Berkhout, B. 2007. In HIV-1 pathogenesis the die is cast during primary infection. *AIDS* **21**:1-11.
8. Brenchley, J. M., Price, D. A., and Douek, D. C. 2006. HIV disease: fallout from a mucosal catastrophe? *Nat. Immunol.* **7**(3): 235-239.
9. Paiardini, M., Frank, I., Pandrea, I., Apetrei, C., and Silvestri, G. 2008. Mucosal immune dysfunction in AIDS pathogenesis. *AIDS Rev*. **10**(1):36-46.

10. Brenchley, J. M., and Douek, D. C. 2008. HIV infection and the gastrointestinal immune system. *Mucosal Immunol*. 1:23-30.

11. Mehandru, S., Poles, M. A., Tenner-racz, K., Jean-pierre, P., Manuelli, V., Lopez, P., Shet, A., Low, A., Hiroshi, M., Boden, D., Racz, P., and Markowitz, M. 2006. Lack of mucosal immune reconstitution during prolonged treatment of acute and early HIV-1 infection. *PLoS Med* **3**(12):2335-2348.
12. Poles, M. A., Boscardin, W. J., Elliott, J., Taing, P., Fuerst, M. P., McGowan, I., Brown, S., and Anton, P. A. 2006. Lack of decay of HIV-1 in gut-associated lymphoid tissue reservoirs in maximally suppressed individuals. *JAIDS* **43**(1):65-68.
13. Cassol, E., Malfeld, S., Mahasha, P., van der Merwe, S., Cassol, S., Seebregts, C., Alfano M., Poli, G., and Rossouw, T. 2010. Persistent Microbial Translocation and Immune Activation in HIV-1 Infected South Africans Receiving Combination Antiretroviral Therapy**.** *J infec Dis* **202**: 723-733.

14. Poles, M. A., Elliott, J., Vingerhoets, J., Michiels, L., Scholliers, A., Bloor, S., Larder, B., Hertogs, K., and Anton, P. A. 2001. Despite high concordance, distinct mutational and phenotypic drug resistance profiles in human immunodeficiency virus type 1 RNA are observed in gastrointestinal mucosal biopsy specimens and peripheral blood mononuclear cells compared with plasma**.** *J infe. Dis* **183**(1), 143-8.
15. Avettand-Fenoel, V., Hocqueloux, L., Müller-Trutwin, M., Prazuck, T., Melard, A., Marie-Laure, C., Agoute, E., Michau, C., and Rouzioux, C. 2011. Greater Diversity of HIV DNA Variants in the Rectum Compared to Variants in the Blood in Patients Without HAART. *J Med Virol* **83**:1499–1507.
16. Imamichi, H., DeGray, G., Dewar, R. L., Mannon, P., Yao, M., Chairez, C., Sereti, I., and Kovacs, J. A. 2011. Lack of compartmentalization of HIV-1 quasispecies between the Gut and peripheral blood compartments**.** *J inf Dis* **204**:309-14.
17. Lerner, P., Guadalupe, M., Donovan, R., Hung, J., Flamm, J., Prindiville, T., Sankaran-Walters, S., Syvanen, M., Wong, J. K., George, M. D., and Dandekar, S. 2011. Gut mucosal viral reservoir in HIV infected patients is not the major source of rebound plasma viremia following HAART interruption. *J Virol* **85**:4772–4782.
18. Evering, T. H., Mehandru, S., Racz, P., Tenner-Racz, K., Poles, M. A., Figueroa, A., Mohril, H., and Markowitz, M. 2012. Absence of HIV-1 evolution in the Gut-associated lymphoid tissue from patients on combination antiviral therapy initiated during primary infection. *PLoS Pathogens* **8:**e1002506.

19. van Marle, G., Gill, M. J., Kolodka, D., McManus, L., Grant, T., Church, D. L. 2007. Compartmentalization of the gut viral reservoir in HIV-1 infected patients. *Retrovirol* 4:87.
20. van Marle, G., Church, D. L., Nunweiler, K. D., Cannon, K., Wainberg, M. A., Gill, M. J. 2010. Higher levels of Zidovudine resistant HIV in the colon compared to blood and other gastrointestinal compartments in HIV infection**.** *Retrovirol* **7**:74
21. Schneider, T., Jahn, H. U., Schmidt, W., Riecken, E. O., Zeitz, M., and Ullrich R. 1995. Loss of CD4 T lymphocytes in patients infected with human immunodeficiency virus type 1 is more pronounced in the duodenal mucosa than in the peripheral blood. *Gut* **37**:524–9.
22. Yukl, S. A., Gianella, S., Sinclair, E., Epling, L., Li, Q., Duan, L., Choi, A. L., Girling, V., Ho, T., Li, P., Fujimoto, K., Lampiris, H., Hare, C. B., Pandori, M., Haase, A. T., Günthard, H. F., Fischer, M., Shergill, A. K., McQuaid, K., Havlir, D. V., Wong, J. K. 2010**.** Differences in HIV burden and immune activation within the gut of HIV-positive patients receiving suppressive antiretroviral therapy. *J Infect Dis* **202**:1553-1561.
23. Cassol, E., Malfeld, S., Mahasha, P., Bond, R., Slavik, T., Seebregts, C., Poli, G., Cassol, S., van der Merwe, S., and Rossouw, T. 2013. Impaired $CD4^+$ T cell restoration in the small vs. large intestine of $HIV\text{-}1^+$ South Africans receiving combination antiretroviral therapy. *J Infect Dis* **208**:1113–22.
24. Kosakovsky Pond, S. L., Posada, D., Stawiski, E., Chappey, C., Poon, A. F., Hughes, G., Fearnhill, E., Gravenor, M. B., Leigh Brown, A. J., and Frost, S. D. 2009. An evolutionary modelbased algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1. *PLoS Comp Biol* **5**:e1000581.
25. Lessells, R. J., Katzenstein, D. K., de Oliveira, T. 2012. Are subtype differences important in HIV drug resistance? *Curr Opinion Virol* **2**(5):636-43.
26. Peña, A.C.P., Faria, N.R., Imbrechts, S., Libin, P., Abecasis, A.B., Deforche, K., Gomez, A., Camacho, R.J., De Oliveira, T., and Vandamme, A-M. 2012. Performance of the subtyping tools on the surveillance of HIV-1 epidemic: Comparison between Rega version 3 and six other automated tools to identify pure subtypes and circulating recombinant forms. 17^{th} International Bioinformatics Workshop on Virus evolution and molecular epidemiology University of Belgrade, Faculty of Medicine, Belgrade, Serbia, 27-31 August. Abstract presentation

27. Kimura, M .A. 1980. Simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**:1111-20.
28. Tamura K., Peterson D., Peterson N., Stecher G., Nei M., and Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol and Evol* **28**: 2731-2739.
29. Velazquez-Campoy, A., Todd, M. J., Vega, S., Freire, E. 2001. Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. *PNAS* **98**: 6062-6067.
30. Gonzalez, L. M. F., Brindeiro, R. M., Tarin, M., Calazans, A., Soares, M. A, Cassol, S., Tanuri, A. 2003. The subtype C of human immunodeficiency virus type 1 protease presents an in vitro hypersusceptibility to protease inhibitor Lopinavir**.** *Antimicrob. Agents Chemother* **47:**2817-2822.
31. Connor, R. I., Sheridan, K. E., Ceradini, D., Choe, S., and Landau. N. R. 1997. Change in Coreceptor Use Correlates with Disease Progression in HIV-1–Infected Individuals. *J. Exp. Med.* **185**(4): 621–628.
32. Moore, J. P., Kitchen, S. G., Pugach, P., and Zack, J.A. 2004. The CCR5 and CXCR4 Coreceptors Central to Understanding the Transmission and Pathogenesis of Human Immunodeficiency Virus Type 1 Infection. *AIDS Res and Hum Retroviruses* **20** (1): 111–126.
33. Vandekerckhove, L. P., Wensing, A. M., Kaiser, R., Brun-Vezinet, F., Clotet, B., De Luca, A., Dressler, S., Garcia, F., Geretti, A. M., Klimkait, T., Korn, K., Masquelier, B., Perno, C. F., Schapiro, J. M., Soriano, V., Sonnerborg, A., Vandamme, A. M., Verhofstede, C., Walter, H., Zazzi, M., and Boucher, C. A. 2011. European guidelines on the clinical management of HIV-1 tropism testing. *Lancet Infect Dis.* **11** (5), 394-407.
34. McGovern, R. A., Thielen, A., Portsmouth, S., Mo, T., Dong, W., Woods, C. K., Zhong, X., Brumme, C. J., Chapman, D., Lewis, M., James, I., Heera, J., Valdez, H., and . Harrigan, P. R. 2012. Population-Based sequencing of the V3-loop can predict the virological response to Maraviroc in treatment-naive patients of the MERIT Trial. *J Acq Imm Def Synd* **61**(3):279.
35. Sala, M., Centlivre, M., and Wain-Hobson, S. 2006. Clade-specific differences in active viral replication and compartmentalisation. *Curr Opin HIV AIDS* **1**:108–114.

36. Centlivre, M., Sommer, P., Michel, M., Ho Tsong, F. R., Gofflo, S., Valladeau, J. *et al*. 2006. The HIV-1 clade C promoter is particularly well adapted to replication in the gut in primary infection. *AIDS* **20**:657-666.
37. Meyerhans, A., Vatanian, 1. P. and Wain-Hobson, S. 1990. DNA recombination during PCR. *Nucl. Acid Research* **18**, 1687-1690.
38. Yang, Y. L., Wang, G., Dorman, K., and Kaplan, A. H. 1996. Long polymerase chain reaction amplification of heterogeneous HIV type 1 templates produces recombination at a high frequency. *AIDS Res. Hum. Retroviruses* **12**:303–306.
39. Jordan, M. R., Kearney, M., Palmer, S., Shao, W., Maldarelli, F., Coakley, E. P., Chappey, C., Wanke, C., Coffin, J. M. 2010. Comparison of standard PCR/cloning to single genome sequencing for analysis of HIV-1 populations**.** *J. Virol Methods* **168**:114–120.
40. Mens, H., Kearney, M., Wiegand, A., Shao, W., Schonning, K., Gerstoft, J., Obel, N., Maldarelli, F., Mellors, J. W., Benfield, T., Coffin, J. M.. 2010. HIV-1 continues to replicate and evolve in patients with natural control of HIV infection. *J Virol* **84**: 12971–12981.
41. Anderson, J. A., Archin, N. M., Ince, W., Parker, D., Wiegand, A., Coffin, J. M., Kuruc, J., Eron, J., Swanstrom, R., Margolis, D. M. 2011. Clonal sequences recovered from plasma from patients with residual HIV-1 viremia and on intensified antiretroviral therapy are identical to replicating viral RNAs recovered from circulating resting CD4+ T cells. *J. Virol.* **85:**5220-5223.
42. Sigal, A., Kim, J. T., Balazs, A. B., Dekel, E., Mayo, A., Milo, R., Baltimore, D. 2011. Cell-to-cell spread of HIV permits ongoing replication despite antiretroviral therapy. *Nature* **477:** 95–98.

CHAPTER 4

DIFFERENCES IN HIV-1 REPLICATION AND IMMUNE ACTIVATION IN THE THERAPEUTIC RESPONSE OF HIV-1 IN THE SMALL VS. LARGE INTESTINE

TABLE OF CONTENTS

4.0. INTRODUCTION

HIV-1 infection is characterized by a profound depletion of $CCR5^+$ memory $CD4^+$ T cells in the *lamina propria* of the gastrointestinal tract (GIT) [1-3]. This loss has been attributed to the direct killing of these cells by HIV-1 and to bystander apoptosis of uninfected cells [4; 5]. Depletion of $CD4^+$ T cells in the GIT has been implicated in the immune dysfunction and loss of epithelial integrity that leads to increased microbial translocation, immune activation and disease progression [6-8].

There are still many unanswered questions relating to the loss of intestinal $CD4^+$ T cells and HIV-1 pathogenesis, particularly since intestinal $CD4^+$ T cells are also depleted in non-pathogenic SIV infection [9; 10]. Recent studies suggest differential depletion of Th17 $CD4^+$ T cells may account for this discrepancy as these cells are preferentially depleted in HIV-1 and pathogenic SIV infection and preserved in non-pathogenic SIV infection [11]. Importantly, Th17 cells play a key role in regulating intestinal immunity and in maintaining the mucosal epithelium [12; 13]. Strategies aimed at restoring $CD4^+$ T cells in the GIT may be critical for re-establishing epithelial integrity and reducing chronic inflammation and immune activation.

Investigations to determine whether cART can restore $CD4^+$ T cell levels in the GIT have, thus far, been inconclusive. Most studies have been cross-sectional and performed at a single site in a small number of patients. Overall, these studies have suggested that $CD4^+$ T cell restoration in the GIT is incomplete and occurs more slowly than in blood with many patients exhibiting limited restoration despite prolonged suppressive treatment [14-17]. Other studies have reported near-complete $CD4^+$ T cell reconstitution in the jejunum of patients treated during acute infection [18], and in the ileum [19] and colon [19; 20] of patients on long-term suppressive cART (>4 years). Further investigation is needed to determine whether these discrepancies are due to differences in the timing, duration and type of treatment or tissue-specific differences in homing receptors, immune activation and/or collagen deposition.

To assess how cART affects immune reconstitution in different regions of the GIT; $CD4^+$, $CD8^+$, $CD38^+CD8^+$ T cell and HIV-1 RNA levels in intestinal biopsies of African patients with late-stage infection, chronic diarrhea, and/or weight loss were compared. $CD4^+$ T cell restoration on cART was impaired in the upper small intestine (duodenum, jejunum) compared to the colon and blood. Failed $CD4^+$ T cell restoration was associated with non-

specific enteritis and sustained $CD8^+$ T cell activation. Treatment strategies that target inflammation and immune activation in the upper GIT may provide significant clinical and immunological benefit.

4.1. MATERIALS AND METHODS REFER TO CHAPTER 2.

4.2. RESULTS

4.2.1. Patient characteristics

The demographic and clinical characteristics of patients population are summarized in **Table 2.1**. Symptoms leading to endoscopy were diarrhoea in 27 (71.1%) and/or unexplained weight loss in 26 (68.4%) patients. Infectious agents were present in duodenum and jejunum of 6 patients (Cryptosporidium in 4 and Giardia in 1 patient). Active Schistosoma infection was identified in the colon of one patient with mild acute colitis. Thirty-one patients had no evidence of opportunistic infections as assessed by stool analysis and histology. Non-specific mild-to-moderate enteritis or colitis was detected in at least one site in 92% patients, at similar frequencies in all intestinal sites (77%, 70%, 70% and 67% in the duodenum, jejunum, ileum and colon, respectively). Of the 38 HIV-1^+ patients studied, 71% and 66% were sampled at 3 and 6 months after the initiation of cART, respectively. Ninety-two percent had good immunological responses (>3 fold increase in blood $CD4^+$ T cell counts) and 75% had complete virological responses after 6 months of cART (plasma VL<50 copies/mL). Only 5 patients continued to suffer from chronic diarrhea at the 6-month time point.

4.2.2. $CD4^+$ T cell depletion was more severe in the small vs. large intestine

$CD4^+$ T cells percentages and absolute numbers were significantly reduced in the GIT of HIV-1^+ patients compared to controls. Compared to levels of 36.8±8.7% (duodenum) and 39.3±10.4% (colon) in controls, $CD4^+$ T cell percentages in cART-naïve HIV-1^+ patients ranged from very low levels in the duodenum and jejunum (3.8±2.8% and 2.6±3.2%, respectively) to somewhat higher levels in the ileum (5.1±5.4%). $CD4^+$ T cell percentages in

the colon (10.3±8.2%) of HIV-1$^+$ patients were significantly higher than in the duodenum and jejunum (p=0.002 and 0.001, respectively) and comparable to those in blood (8.4±5.7%, p=0.547) (**Figure 4.1.A)**. Memory cells are made up the vast majority (>90%) of CD4$^+$ T cells (**Figure 4.1.B**). Histological analysis revealed a similar, statistically significant decrease in the absolute number of CD4$^+$ T cells in HIV-1$^+$ patients compared to controls with the decrease being more severe in the duodenum (78±10 vs. 8±5 cells/0.80 mm^2, p<0.001) than the colon (65±10 vs 17±6 cells/0.8 mm^2, p<0.001)(**Figure 4.1.C and 4.1.D)**. No significant correlations were detected between CD4$^+$ T cell counts or percentages in the gut compared to blood. Compared to controls, CD4$^+$ T cell levels in the duodenum and colon of HIV-1$^+$ patients were reduced by 89.5% and 73.7%, respectively, when assessed by flow cytometry and by 83.3% and 69.4%, respectively, when assessed by immunohistochemistry (**Figure 4.1.E)**.

4.2.3. CD8$^+$ T cell levels were higher in the small vs. large intestine

The percentages of CD8$^+$ T cells in the duodenum and jejunum of cART-naive HIV-1$^+$ patients were significantly higher than in the colon (71.1±14.4% and 77.5±17.3% vs. 53.8±18.3%, respectively, p=0.0069 and p=0.0006) with the ileum displaying intermediate levels (66.9± 7.1%)(**Figure 4.2.A)**. CD8$^+$ T cell percentages in the colon were comparable to those in blood (53.8±18.3% vs. 56.1±8.4%, respectively). The absolute numbers of CD8$^+$ T cells were also significantly higher in the duodenum compared to the colon (101±9.4 vs. 74±5.7 cells per 0.80 mm^2, p=0.012) (**Figure 4.2.B and 4.2.C)**. Although not statistically significant, the percentages of activated CD38$^+$ CD8$^+$ T cells were slightly higher in the duodenum and jejunum than the colon (35.9±21.42% and 40.9±23.1% vs. 30.1±17.75%, p=0.10 and p=0.09, respectively) (**Figure 4.2.D)**. In the colon, but not in the duodenum, jejunum or ileum, the percentage of activated CD38$^+$CD8$^+$ T cells was positively correlated with the amount of tissue HIV-1 RNA (r =0.50; p=0.04).

4.2.4. HIV-1 RNA levels were higher in the colon than in the duodenum and jejunum

All 38 (100%) patients had detectable HIV-1 RNA in the duodenum, jejunum, ileum, and colon prior to the initiation of cART (range from 3.89 to 9.31 $\log_{10}$ HIV-1 RNA copies/g

of tissue). Mean HIV-1 RNA levels were significantly higher in the colon than the duodenum and jejunum (7.2±1.2 vs. 6.4±1.0 and 6.1±1.1 log_{10} HIV-1 RNA copies/g of tissue, respectively, p=0.0067 and p=0.0004, respectively) with intermediate levels in the ileum (6.7±0.95 log_{10} HIV-1 RNA copies/g of tissue p=0.1192) (**Figure 4.3.A)**. Normalization by absolute $CD4^+$ T cell count indicated that higher levels of HIV-1 RNA in the colon were associated with an increased frequency of $CD4^+$ T cells rather than an increase in the amount of virus production per cell (**Figure 4.3.B)**.

4.2.5. NRTI/NNRTI treatment resulted in partial restoration of $CD4^+$ T cells in the blood and colon, but not in the duodenum

To assess the efficacy of short-term cART on $CD4^+$ T cell restoration, $CD4^+$ T cells levels in peripheral blood and longitudinal biopsy samples (duodenum, colon) collected prior to, and following 3 and 6 months of cART were compared. These analyses were limited to patients who had complete virological responses in plasma (<50 copies of HIV-1 RNA/mL) and no evidence of opportunistic infection (n=23). The percentages of $CD4^+$ T cells in the blood and colon of cART-treated HIV^+ patients increased from 8.4±5.7% and 10.3±8.2%, respectively, at baseline to 17.8±10.1% (2.1 fold increase, p=0.01) and 15.8±7.4% (1.5 fold increase, p=0.0003), respectively 6 months after the initiation of cART (**Figure 4.4.A)**. Blood $CD4^+$ T cell counts increased on average 3.5-fold after 6 months of cART. Despite these increases, the frequency of $CD4^+$ T cells in these compartments remained lower than in uninfected controls (17.8±10.1% vs. 38.0±9.4% in blood, p=0.0042, 15.8±7.4 vs. 36.3±10.4 % in the colon, p=0.0006). $CD4^+$ T cell depletion was more persistent in the duodenum, with no measurable increase detected after 6 months of cART (mean frequency of 3.7 ± 2.8% at baseline vs. 3.9 ± 3.1% at month 6 (p=0.5623) (**Figure 4.4.B)**.

4.2.6. Failed restoration of $CD4^+$ T cells was associated with persistent inflammation and $CD8^+$ T cell activation

To investigate whether cART resulted in decreased inflammation and $CD8^+$ T cell activation, infiltration of MNC was examined including macrophages, lymphocytes, eosinophils and/or plasma cells and assessed $CD38^+CD8^+$ T cell levels in longitudinal

duodenal and colonic biopsies collected prior to, and after 3 (duodenum) and 6 (duodenum, colon) months of cART. Prior to cART, 67% and 77% of patients had mild-to-moderate levels of non-specific inflammation in the colon and duodenum, respectively. After 6 months of cART, only 44% of HIV-1^+ patients had sustained non-specific colitis (colon) compared to 72% with non-specific enteritis (duodenum). Similarly, the proportion of activated $CD8^+$ T cells in the colon decreased from 30.1±17.8 at baseline to 18.9±8.7% (p=0.048) after 6 months of cART. In the colon, the delta change in HIV-1 RNA was positively correlated with the decrease in activated $CD38^+CD8^+$ T cells (r =-0.38; p=0.03). Although not statistically significant, $CD38^+CD8^+$ T cell percentages in the duodenum were somewhat higher after 3 (48.6±23.4%) and 6 (44.1±23.8%) months of cART than at baseline (35.9±21.4%; p=0.43 and 0.58, respectively). This occurred despite a marked decrease in HIV-1 RNA (from 6.4±1.0 at baseline to 4.2±0.8 and 4.9±1.1 $\log_{10}$ copies/g tissue at months 3 and 6, p=0.0005 and p=0.0134 respectively). This decrease was comparable to that observed in the colon where HIV-1 RNA decreased from 7.2±1.2 at baseline to 5.15±1.2 $\log_{10}$ copies/g of tissue at month 6 (p=0.0005).

4.3. **DISCUSSION**

The impact of cART on $CD4^+$ T cell reconstitution in the intestine of HIV-1^+ patients is a topic of considerable debate. Contributing to the confusion are the inter-study differences in sampling sites, treatment regimens, time on cART and stage of disease, as well as limitations inherent in cross-sectional studies. To our knowledge, this is the first longitudinal study to examine the extent of $CD4^+$ T cell depletion and restoration in the GIT of African patients with late-stage disease and unexplained chronic diarrhoea and/or weight loss. In the developing world, individuals presenting with $CD4^+$ T cell counts <200 cells//mm^3 are fast-tracked for treatment with a standardized NRTI-NNRTI drug regimen [21; 22].

Prior to cART, a profound decrease was detected in $CD4^+$ T cells and a corresponding moderate increase in $CD8^+$ T cells in all sites of the GIT regardless of the method used to quantify T cell subsets, flow cytometry or immunohistochemistry. The severity of depletion exceeded that reported for patients with acute/early HIV-1 infection, presumably a reflection of the continued loss of $CD4^+$ T cells during disease progression [14; 23]. One of the most striking findings was the profound depletion of $CD4^+$ T cells in the small (duodenum,

jejunum, ileum) vs. large intestine (colon) and peripheral blood. Other studies have reported that $CD4^+$ T cell depletion is more severe in the duodenum [23], ileum [1; 24] and colon [3; 24] than in the blood of treatment-naive patients and in the duodenum vs. the colon and rectum [25] of patients on suppressive cART.

To determine whether the distinct immune environments support different levels of HIV-1, we compared total HIV-1 RNA levels (encompassing both cell-associated and virion RNA) in the duodenum, jejunum and ileum with those in the colon. HIV-1 RNA was readily detectable throughout the GIT of cART-naïve patients, at levels exceeding those in blood (assuming 1.0 g of tissue is equivalent to 1.0 mL of plasma). As observed for $CD4^+$ T cells, HIV-1 RNA increased from moderate levels in the duodenum and jejunum to intermediate levels in the ileum and significantly higher levels in the colon. Similarly, in cART-treated patients, studies have shown that the amount of HIV-1 proviral DNA per $CD4^+$ T increases when descending through the GIT toward the colon and rectum [25].

The positive correlation observed between activated $CD38^+CD8^+$ T cells and HIV-1 RNA in the colon of treatment-naïve patients, together with the parallel decrease in both parameters in response to cART suggests that $CD8^+$ T cell activation is closely linked to viral replication. Other investigators have reported that, in cART-treated patients with maximally suppressed viral replication, the degree of $CD8^+$ T cell activation in the colon correlates with the size of the residual HIV-1 DNA reservoir [20]. In the duodenum, the lack of a correlation between $CD38^+CD8^+$ T cells and HIV-1 RNA and the persistence of activated $CD38^+CD8^+$ T cells during cART, despite a marked decrease in HIV-1 RNA, suggests that factors other than, or in addition to, HIV-1 (co-infections, microbial translocation, disruption of the microenvironment, persistent immune activation that is not fully reversed by cART) contribute to $CD8^+$ T cell activation in this tissue.

In agreement with a growing body of evidence [14; 26-28], these results also suggest that persistent inflammation and $CD8^+$ T cell activation are associated with lower $CD4^+$ T cell gains. This view is supported by studies showing an inverse correlation between activated $CD38^+CD8^+$ T cells and $CD4^+$ T cell levels in blood [29; 30] and between the activated $Ki67^+CD8^+$ T cells and the severity of $CD4^+$ T cell depletion in the lower GIT [24]. Although the mechanism(s) are not fully understood, activated $CD8^+$ T cells have the capacity to delay or prevent $CD4^+$ T cell recovery by producing pro-inflammatory cytokines and chemokines, suppressing $CD4^+$ T cell proliferation and inducing $CD4^+$ T cell apoptosis [18; 31; 32].

Chronic inflammation and immune activation may also lead to increased collagen deposition and fibrotic damage [33; 34]. Fibrotic deposition of collagen occurs in Peyer's patches of the ileum and the extent of deposition/damage is predictive of $CD4^+$ T cell restoration [35; 36]. There are no studies of collagen deposition in the duodenum, jejunum or colon, regions of the GALT that are devoid of Peyer's patches, known to exist. Other investigators have reported that genes involved in inflammation and cell activation are up-regulated in the jejunum of patients who exhibit poor mucosal $CD4^+$ T cell responses [28]. Further studies are needed to determine whether these changes are confined to the small intestine.

Another factor that could contribute to poor $CD4^+$ T cell reconstitution in the duodenum relates to the CCL25/CCR9 axis of $CD4^+$ T cell recruitment. This axis plays an important role in lymphocyte homing to the small intestine, but not to the colon. CCL25 is a chemokine produced by epithelial cells of the small intestine [37; 38]. Its receptor, CCR9, is expressed on lymphocytes that traffic to the small intestine [39; 40]. HIV-1^+ patients produce less CCL25 and as a result many $CCR9^+CD4^+$ T cells remain in the circulation rather than re-populating effector sites in the small intestine [41]. Reduced homing of $CCR9^+CD4^+$ T cells has been associated with mucosal damage, increased microbial translocation and systemic immune activation, changes that may contribute to poor immune restoration [42].

The strengths of the study, in comparison to previous studies, relate to the homogeneity of our patient cohort (with respect to treatment, stage of infection, and the absence of overt opportunistic infections in the majority of patients) and the repeat sampling of multiple tissue sites from the same patients before and after the initiation of cART. The study focused on the effects of short-term cART since the interest was in identifying early predictors of immunological failure. Although there is no consensus as to the best time to assess immunological responses [43; 44], studies performed on peripheral blood after 3 and 6 months of cART have been shown to be predictive of immune restoration in the long-term [45]. Limitations of these studies relate to the small size of our control group and the limited number of immune and virological correlates examined. Furthermore, despite efforts to exclude patients with enteric co-infections (stool cultures and histological staining for AFB, fungal and parasitic infections); couldn't exclude the possibility that other co-infections may affect our results. In addition, cell numbers did not allow us to determine the antigen specificity and clonal diversity of our $CD8^+$ T cells. Further studies are also required to determine whether these results can be generalized to patients without diarrhea and weight loss and patients of non-African origin.

This study demonstrates that, despite the late stage of infection, gradual restoration of $CD4^+$ T cell levels is possible in the colon and blood of African AIDS patients on short-term cART. It also underscores the distinct differences in the small vs. large intestine with regards to $CD4^+$ T cell depletion and reconstitution and highlights the need for detailed comparative studies to better delineate the mechanisms leading to immunological failure in the upper small intestine. An understanding of these mechanisms may lead to the development of highly targeted treatment strategies that enhance $CD4^+$ T cell recovery. If inflammation and immune activation in the small intestine are driven by local factors not directly related to HIV-1 (tissue damage, loss of epithelial integrity, non-specific activation of $CD8^+$ T cells), it is unlikely that drugs targeting HIV replication will be maximally effective with respect to immune recovery. This view is supported by studies showing that treatment intensification with raltegravir, a potent integrase inhibitor, did not have any significant effect on $CD4^+$ T cell levels or immune activation in the GIT, or on HIV-1 RNA levels in plasma [46-48]. cART protocols that include immune modulating and anti-inflammatory agents may offer an attractive alternative approach.

4.4. **FIGURE LEGENDS**

Figure 4.1: Depletion of $CD4^+$ T cells in cART-naïve patients with late-stage HIV-1 infection was more severe in the small than in the large intestine. A. Proportion of T cells that express CD4, as measured by flow cytometric analysis of peripheral blood mononuclear cells (PBMC) or total cell suspensions isolated from the gastrointestinal tract (GIT) of cART-naïve AIDS patients. Mean values and statistical significance among the different sites are indicated (*p<0.05, p**<0.01, ***p<0.001). B. The proportions of naïve and memory CD4 T cells in the various intestinal sites in a subset of 12 patients prior to the initiation of cART. Memory CD4+ T cells were defined as cells that express CD45RO and naïve cells were defined as CD45RO- and CD27+. C. Representative immunohistochemical staining for CD4 (brown) in the duodenum and colon of a cART-naïve AIDS patient. D. Decrease in the absolute number of CD4+ T cells in the duodenum and colon of cART-naïve AIDS patients relative to uninfected controls as assessed by immunohistochemistry (p**<0.01). E. Comparison of the decrease in $CD4^+$ T cells in cART-naïve AIDS patients vs. uninfected controls as assessed by two different methods - flow cytometry (percentage of $CD4^+$ T cells) vs. histology (absolute number of $CD4^+$ T cells).

Figure 4.2: Increased frequency of $CD8^+$ T cells and activated $CD38^+CD8^+$ T cells in the upper small intestine compared to the colon of cART-naïve AIDS patients. A. Proportion of T cells that express CD8, as measured by flow cytometric analysis of peripheral blood mononuclear cells (PBMC) or total cell suspensions isolated from the GIT of treatment-naïve AIDS patients. Mean values and statistical significance among the different sites are indicated (*$p<0.05$, ***$p<0.001$). B. Increase in the absolute number of $CD8^+$ T cells in the duodenum and colon of treatment-naïve AIDS patients compared to uninfected controls, as quantified by immunohistochemistry. C. Representative immunohistochemical staining for CD8 (brown) in the duodenum and colon of a treatment-naïve AIDS patient. D. Proportion of $CD8^+$ T cells expressing the activation marker, CD38, as assessed by flow cytometry of cell suspensions isolated from different regions of the GIT of cART-naïve AIDS patients.

Figure 4.3: Prior to the initiation of cART, HIV RNA levels were higher in the colon than in the upper small intestine of African AIDS patients. A. HIV RNA copy numbers in cART-naïve AIDS patients (n=38) were measured in plasma and tissue extracts prepared from different sites in the GIT. Biopsy samples collected from different sites in the GIT were minced with a razor blade, extracted overnight in guanidinium lysis buffer and processed for HIV RNA quantification using the Nuclisens Easy Q assay. Results were reported as $\log_{10}$ HIV RNA copies/g of tissue, or per mL of plasma. B. HIV-1 RNA levels in the duodenum and colon normalized by their respective absolute $CD4^+$ T cell counts.

Figure 4.4: Impaired $CD4^+$ T cell restoration and sustained $CD8^+$ T cell activation in the duodenum, but not in the colon, of NRTI/NNRTI-treated AIDS patients. A. Spaghetti plots of the kinetics of CD4+ T cell restoration, HIV RNA clearance and CD8 T cell activation in the duodenum, colon, and blood. Each line represents an individual patient. Analysis was limited to patients with a complete virological response at month 6 post-cART (plasma VL<50 copies/ml, n=23). B. Summary statisitics of the kinetics CD4+ T cell restoration, HIV RNA clearance and CD8 T cell activation in the duodenum, colon, and blood (n=23). Results are reported as the mean and standard deviation.

4.5. FIGURES:

Figure 4.1.

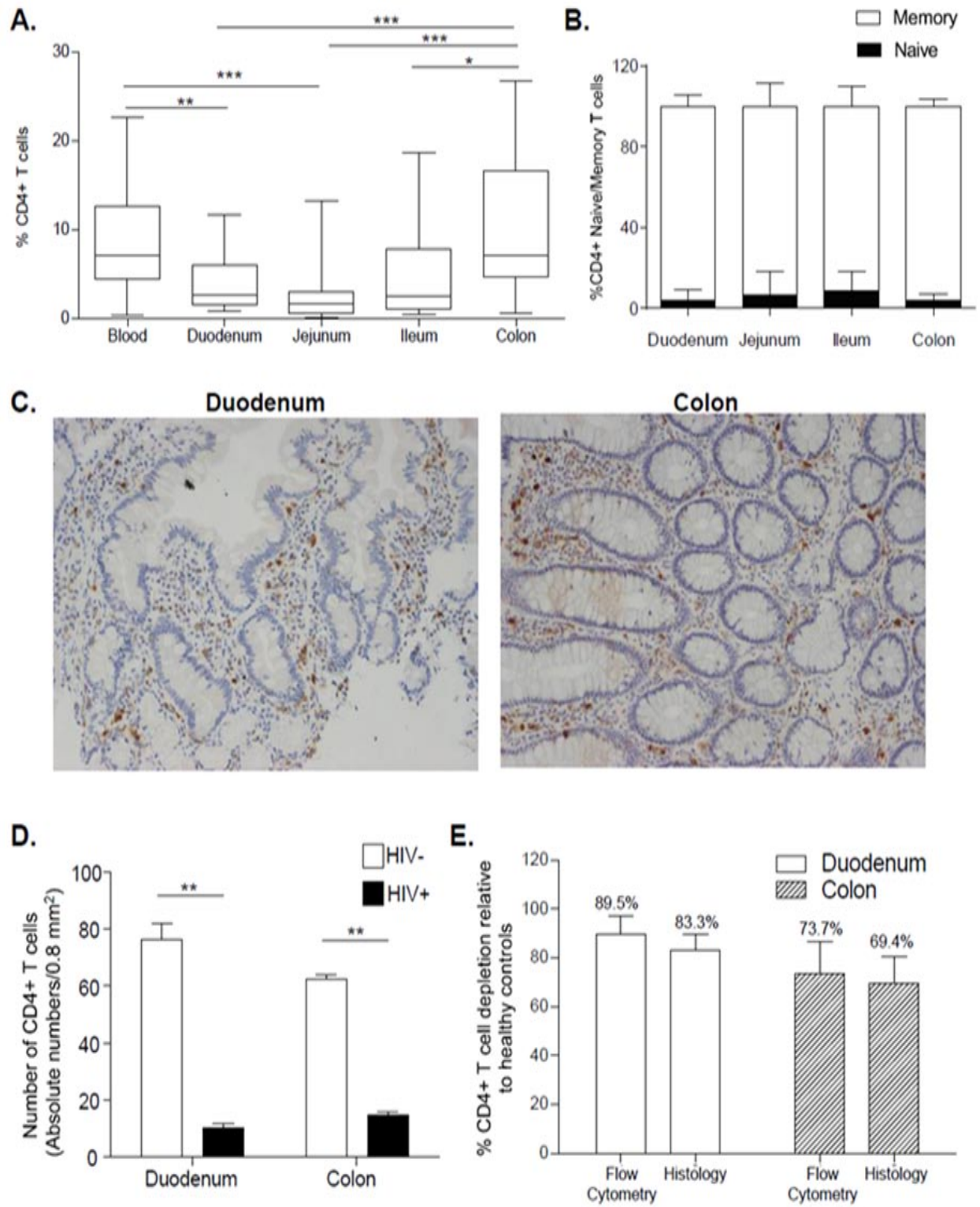


Figure 4.2.

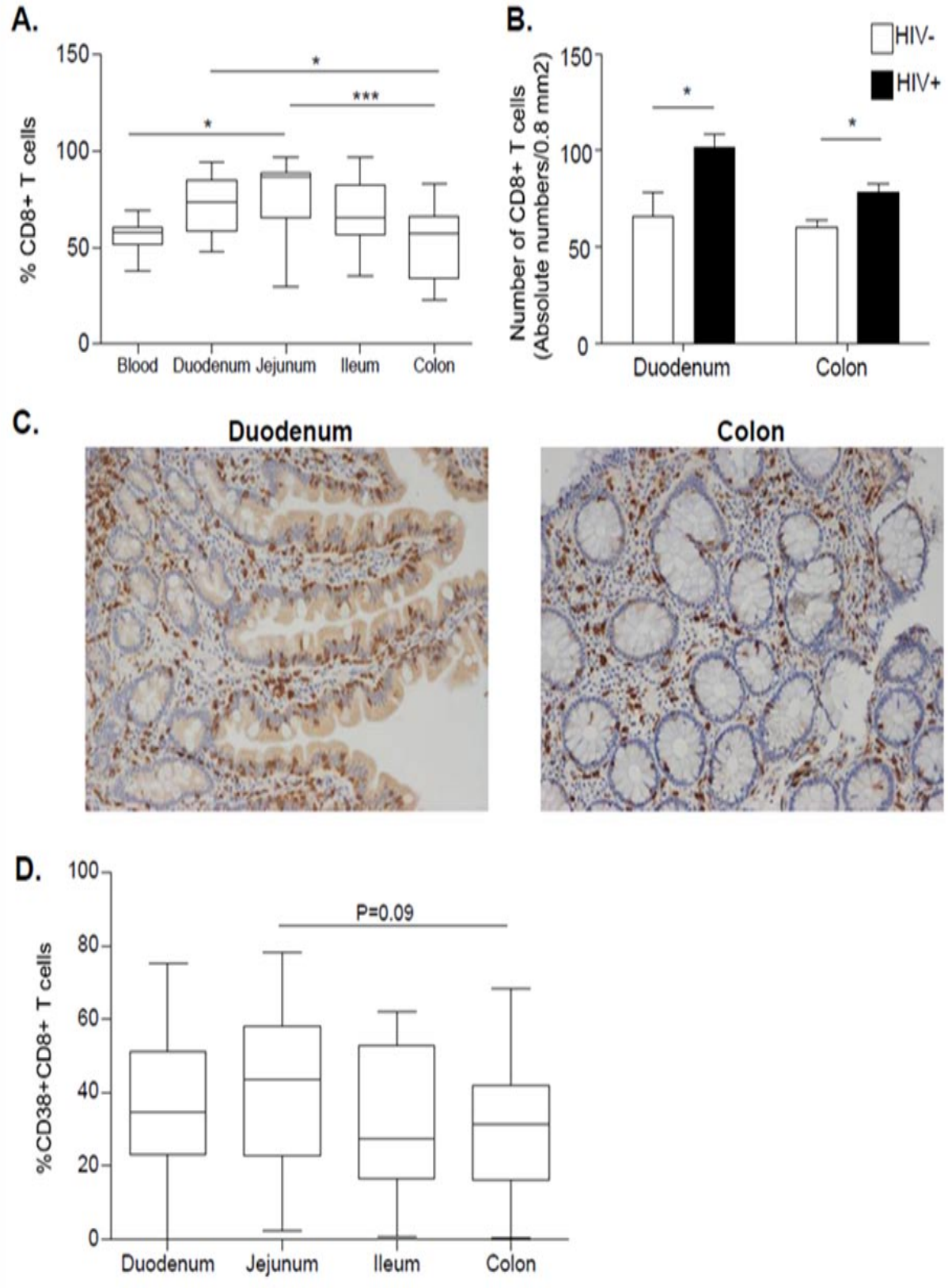


Figure 4.3.

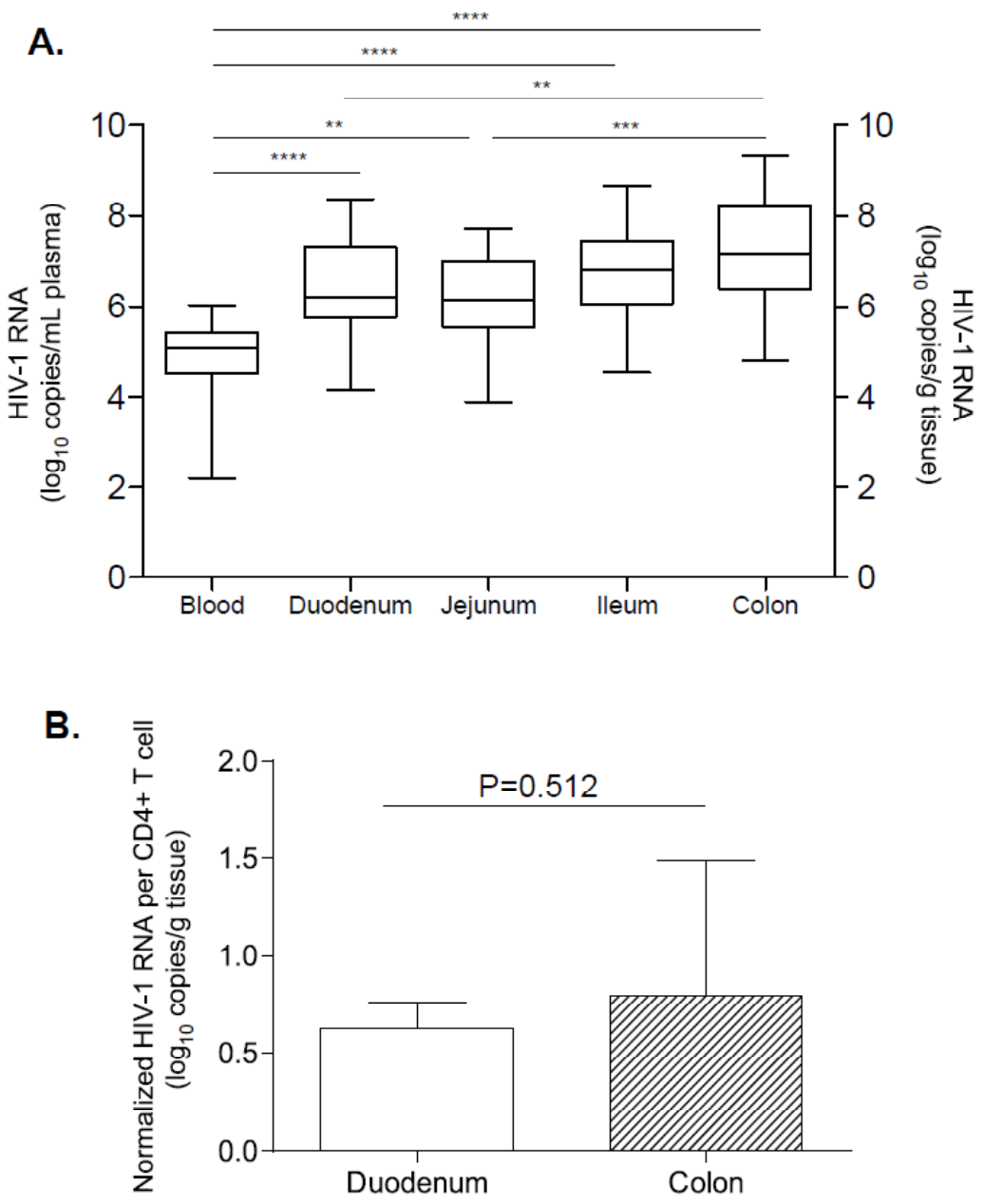
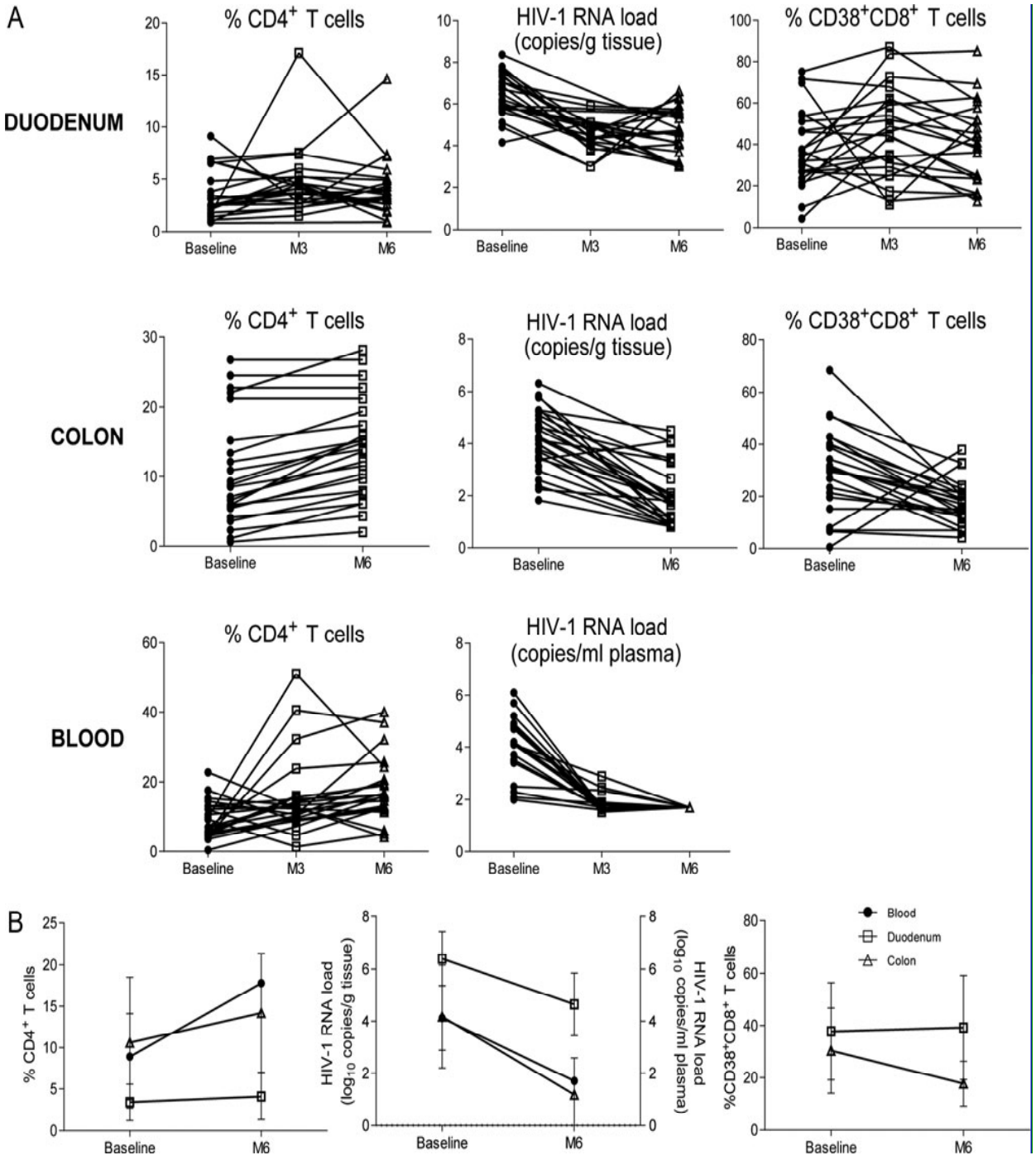


Figure 4.4.



4.6. REFERENCES

1. Brenchley, J. M., Schacker, T. W., Ruff, L. E., Price, D. A., Taylor, J. H., Beilman, G. J., Nguyen, P. L., Khoruts, A., Larson, M., Haase, A. T., and Douek, D. C. 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *J Exp Med.* **200**:749-759.

2. Fukazawa, Y., Miyake, A., Ibuki, K., Inaba, K., Saito, N., Motohara, M., Horiuchi, R., Himeno, A., Matsuda, K., Matsuyama, M., Takahashi, H., Hayami, M., Igarashi, T., and Miura, T. 2008. Small intestine $CD4^+$ T cells are profoundly depleted during acute simian-human immunodeficiency virus infection, regardless of viral pathogenicity. *J Virol*. **82**:6039-6044.

3. Mehandru, S., Poles M. A., Tenner-Racz, K., Horowitz, A., Hurley, A., Hogan, C., Boden, D., Racz, P., and Markowitz, M. 2004. Primary HIV-1 infection is associated with preferential depletion of CD4+ T lymphocytes from effector sites in the gastrointestinal tract. *J Exp Med*. **200**:761-770.

4. Mehandru, S., Poles, M. A., Tenner-Racz, K., Manuelli, V., Jean-Pierre, P., Lopez, P., Shet, A., Low, A., Mohri, H., Boden, D., Racz, P., Markowitz, M.. 2007. Mechanisms of gastrointestinal CD4+ T-cell depletion during acute and early human immunodeficiency virus type 1 infection. *J Virol*. **81**:599–612.

5. Li, Q., Duan L., Estes, J. D., Ma, Z. M., Rourke, T., Wang, Y., Reilly, C., Carlis, J., Miller, C. J., and Haase, A. T. 2005. Peak SIV replication in resting memory CD4+ T cells depletes gut lamina propria CD4+ T cells. *Nature* **434**:1148-51.

6. Brenchley, J. M., D. A. Price, T. W. Schacker, T. E. Asher, G. Silvestri, S. Rao, Z. Kazzaz, E. Bornstein, O. Lambotte, D. Altmann, B. R. Blazar, B. Rodriguez, L. Teixeira-Johnson, A. Landay, J. N. Martin, F. M. Hecht, L. J., Picker, M. M. Lederman, S. G. Deeks, and D. C. Douek. 2006. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat Med.* 12:1365-1371.

7. Paiardini, M., Frank, I., Pandrea, I., Apetrei, C., Silvestri, G. 2008. Mucosal immune dysfunction in AIDS pathogenesis. *AIDS Rev.* **10**: 36–46.

8. Brenchley, J.M., and Douek, D.C. 2008. HIV infection and the gastrointestinal immune system. *Mucosal Immunol.* **1**:23-30.

9. Gordon, S. N., Klatt, N. R., Bosinger, S. E , Brenchley, J. M., Milush, J. M., Engram, J. C., Dunham, R. M., Paiardini, M., Klucking, S., Danesh, A., Strobert, E. A., Apetrei, C., Pandrea, I. V., Kelvin, D., Douek, D. C., Staprans, S. I., Sodora, D. L., and Silvestri, G. 2007. Severe depletion of mucosal CD4+ T cells in AIDS-free simian immunodeficiency virus-infected sooty mangabeys. J. Immunol. 179**:**3026-3034

10. Pandrea, I. V., R. Gautam, R. M. Ribeiro, J. M. Brenchley, I. F. Butler, M. Pattison, T. Rasmussen, P. A. Marx, G. Silvestri, A. A. Lackner, A. S. Perelson, D. C. Douek, R. S. Veazey, and C. Apetrei. 2007. Acute loss of intestinal CD4+ T cells is not predictive of simian immunodeficiency virus virulence. *J. Immunol.* **179:**3035-3046.

11. Brenchley, J. M., Paiardini, M., Knox, K. S., Asher, A. I., Cervasi, B., Asher, T. E., Scheinberg, P., Price, D. A., Hage, C. A., Kholi, L. M., Khoruts, A., Frank, I., Else, J., Schacker, T., Silvestri, G., Douek, D. C. 2008. Differential Th17 CD4 T-cell depletion in pathogenic and nonpathogenic lentiviral infections. *Blood* **112**: 2826–2835.

12. Aujla, S. J., Dubin, P. I., and Kolls, J. K. 2007. Th17 cells and mucosal host defense. *Semin Immunol.* **19**:377-382.

13. Liang, S. C., Tan, X. Y., Luxenberg, D. P., Karim, R., Dunussi-Joannopoulos, K., Collins, M., Fouser, L. A. 2006. Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides. *J Exp Med.* **203**:2271–2279.

14. Mehandru, S., Poles, M. A., Tenner-racz, K., Jean-pierre, P., Manuelli, V., Lopez, P., Shet, A., Low, A., Hiroshi, M., Boden, D., Racz, P., Markowitz, M. 2006. Lack of mucosal immune reconstitution during prolonged treatment of acute and early HIV-1 infection. *PLoS Med.* **3**(12):2335-2348.

15. Gazzola, L., Tincati, C., Bellistre, G. M., Monforte, A., and Marchetti, G. 2009. The Absence of $CD4^+$ T Cell Count Recovery Despite Receipt of Virologically Suppressive Highly Active Antiretroviral Therapy: Clinical Risk, Immunological Gaps, and Therapeutic Options. *Clin. Infect. Dis.* **48**:328-337.

16. Tincati, C., Biasin, M., Bandera, A., Violin, G., Marchetti, L., Piacentini, L., Vago, G. L., Balotta, C., Moroni, M., Franzetti, F., Clerici, M, and Gori, A. 2009. Early initiation of highly active antiretroviral therapy fails to reverse

immunovirological abnormalities in gut-associated lymphoid tissue induced by acute HIV infection. *Antivir Ther*. **14**: 321–330.

17. Talal, A. H., Monard, S., Vesanen, M., Zheng, Z., Hurley, A., Cao, Y., Fang, F., Smiley, L., Johnson, T., Kost, R., and Markowitz, M. H. 2001. Virologic and Immunologic effect of antiretroviral theray on HIV-1 in Gut-associated lymphoid tissue. *AIDS* **26**:1-7.

18. Guadalupe, M., Reay, E., Sankaran, S., Prindiville, T., Flamm, J., Mcneil, A., and Dandekar, S. 2003. Severe CD4+ T-cell depletion in gut lymphoid tissue during primary human immunodeficiency virus type 1 infection and substantial delay in restoration following highly active antiretroviral therapy. *J Virol*. **77**:11708-17.
19. Ciccone, E. J., Read, S. W., Mannon, P. J., Yao, M. D., Hodge, J. N., Dewar, R., Chairez, C. L., Proschan, M. A., Kovacs, J. A., and Sereti, I. 2010. Cycling of gut mucosal CD4R T cells decreases after prolonged antiretroviral therapy and is associated with plasma LPS levels. *Mucosal Immunol*. **3**:172–181.
20. Sheth, P. M., Chege, D., Shin, L.Y.Y., Huibner, S., Yue, F-Y., Loutfy, M., Halpenny, R., Presad, D., Kovacs, C., Chun, T-W., Kandel, G., Ostrowski, M., and Kaul, R. 2008. Immune reconstitution in the sigmoid colon after long-term HIV therapy. *Mucosal Immunol*. **1**(5):382-8.

21. National Department of Health Republic of South Africa. 2004. Clinical guidelines for the management of HIV & AIDS in adults and adolescents. www.sanac.org.za.

22. WHO/UNICEF/UNAIDS. 2010. Joint reporting tool on the health sector response to HIV/AIDS 2010. Geneva, Switzerland.

23. Schneider, T., Jahn, H. U., Schmidt, W., Riecken, E.O., Zeitz, M., and Ullrich, R. 1995. Loss of CD4 T lymphocytes in patients infected with human immunodeficiency virus type 1 is more pronounced in the duodenal mucosa than in the peripheral blood. Gut **37**:524–9.
24. Gordon, S. N., Cervasi, B., Odorizzi, P., Silverman, R., Aberra, F., Ginsberg, G., Estes, J. D., Paiardini, M., Frank, I., Silvestri, G. 2010. Disruption of intestinal CD4+ T cell homeostasis is a key marker of systemic CD4+ T cell activation in HIV-infected individuals. *J Immunol*. **185**:5169–5179

25. Yukl, S. A., Gianella, S., Sinclair, E., Epling, L., Li, Q., Duan, L., Choi, A. L., Girling, V., Ho, T., Li, P., Fujimoto, K., Lampiris, H., Hare, C. B., Pandori, M., Haase, A. T., Günthard, H. F., Fischer, M., Shergill, A. K., McQuaid, K., Havlir, D. V., and Wong, J. K. 2010. Differences in HIV burden and immune activation within the gut of HIV-positive patients receiving suppressive antiretroviral therapy. *J Infect Dis.* **202**:1553-1561.

26. Hunt, P. W., Martin, J. N., Sinclair, E., Bredt, B., Hagos, E., Lampiris, H., Deeks, S. G**.** 2003. T cell activation is associated with lower CD4+ T cell gains in human immunodeficiency virus-infected patients with sustained viral suppression during antiretroviral therapy. *J Infect Dis* **187**(10): 1534–1543.

27. Mavigner, M., Delobel, P., Cazabat, M., Dubois, M., L'faqihi-Olive, F. E., Raymond, S., Pasquier, C., Marchou, B., Massip, P., Izopet, J. 2009. HIV-1 residual viremia correlates with persistent T-cell activation in poor immunological responders to combination antiretroviral therapy. *PLos One.* **4**(10), e7658.

28. Guadalupe, M., Sankaran, S., George, M. D., Reay, E., Verhoeven, D., Shacklett, B. L., Flamm, J., Wegelin, J., Prindiville, T., and Dandekar, S. 2006. Viral Suppression and Immune Restoration in the Gastrointestinal Mucosa of Human Immunodeficiency Virus Type 1-Infected Patients Initiating Therapy during Primary or Chronic Infection. *J Virol* **80**:8236-8247.

29. Giorgi, J. V., Liu, Z., Hultin, L. E., Cumberland, W. G., Hennessey, K., Detels, R. 1993. Elevated levels of $CD38^+CD8^+$ T cells in HIV infection add to the prognostic value of low $CD4^+$ T cell levels: Results of 6 years of follow-up. The Los Angeles Center, Multicenter AIDS Cohort Study. *J Acquir Immune Defic Syndr.* **6**:904.912.

30. Chun, T. W., Justement, J. S., Sanford, C., Hallahan, C. W., Planta, M. A., Loutfy, M., Kottilil, S., Moir, S., Kovacs, C., and Fauci. A. S. 2004. Relationship between the frequency of HIV-specific $CD8^+$ T cells and the level of $CD38^+CD8^+$ T cells in untreated HIV-infected individuals. *Proc Natl Acad Sci USA.* **101**: 2462-2469.

31. Nobel, A., Pestano, G. A., and Cantor, H. 1998. Suppression of immune responses by CD8 cells. I. Superantigen-activated CD8 cells induce unidirectional Fas-mediated apoptosis of antigen-activated CD4 cells. *J Immunol.* **160**: 559-565.

32. Smit-McBride, Z., Mattapallil, J. J., Villinger, F., Ansari, A. A., and Dandekar, S. 1998. Intracellular cytokine expression in the $CD4^+$ and $CD8^+$ T cells from intestinal mucosa of simian immunodeficiency virus infected macaques. *J. Med. Primatol*. **27:**129-140.

33. Schacker, T. W., Brenchley, J. MBellman, G. J., Reilly, C., Pambuccian, S. E., Taylor, J., Skarda, D., Larson, M., Douek, D. C. and Haase., A. T. 2006. Lymphatic tissue fibrosis is associated with reduced numbers of naive $CD4^+$ T cells in human immunodeficiency virus type 1 infection.*Clin. Vaccine Immunol.* **13:**556-560.

34. Estes, J. D., Haase, A. T., and Schacker, T. W. 2008. The role of collagen deposition in depleting $CD4^+$ T cells and limiting reconstitution in HIV-1 and SIV infections through damage to secondary lymphoid organ niche. *Semin Immunol*. **20**:181-186.

35. Schacker, T. W., Reilly, C., Beilman, G. J., Taylor, J., Skarda, D., Krason, D., Larson, M., Haase, A. T. 2005. Amount of lymphatic tissue fibrosis in HIV infection predicts magnitude of HAART-associated change in peripheral CD4 cell count. *AIDS* **19:**2169-2171.

36. Estes, J., Baker, J. V., Brenchley, J. M., Khoruts, A., Barthold, J. L., Bantle, A., Reilly, C. S., Beilman, G. J., George, M. E., Douek, D. C., Haase, A. T., and Schacker, T. W. 2008. Collagen deposition limits immune reconstitution in the gut. *J Infect Dis* **198**:456-464.

37. Kunkel, E. J., Campbell, J. J., Haraldsen, G., Pan, J., Boisvert, J., Roberts, A. I., Ebert, E. C., Vierra, M. A., Goodman, S. B., Genovese, M. C., Wardlaw, A. J., Greenberg, H. B., Parker, C. M., Butcher, E. C., Andrew, D. P., and Agace. W. W. 2000. Lymphocyte CC chemokine receptor 9 and epithelial thymus-expressed chemokine (TECK) expression distinguish the small intestinal immune compartment: epithelial expression of tissue-specific chemokines as an organizing principle in regional immunity. J. Exp. Med. 192**:**761-768.

38. Papadakis, K. A., Prehn, J., Moreno, S.T., Cheng, L., Kouroumalis, E. A., Deem, R., Breaverman, T., Ponath, P. D., Andrew, D. P., Green, P. H., Hodge, M. R., Binder, S. W., Targan, S. R. 2001. CCR9-positive lymphocytes and thymus-expressed chemokine distinguish small bowel from colonic Crohn's disease. *Gastroenterology*. **121**: 246-254.

39. Johansson-Lindbom, B., Svensson, M., Wurbel, M-A., Malissen, B., Ma´ rquez, G., Agace, W. 2003. Selective generation of gut-tropic T cells in gut associated lymphoid tissue (GALT); requirement for GALT dendritic cells and adjuvant. *J Exp Med.* **198**:963-969.

40. Mora, J. R., Bono, M. R., Manjunath, N., Weninger, W., Cavanagh, L. L., Rosemblatt, M., and Von Andrian, U. H. 2003. Selective imprinting of gut-homing T-cells by Peyer's patches dendritic cells. *Nature.* **424**: 88-93.

41. Mavigner, M., Cazabat, M., Dubois, M., L'Faqihi, F. E., Requena, M., Pasquier, C., Klopp, P., Amar, J., Alric, L., and Barange, K. 2012. Altered $CD4^+$ T cell homing to the gut impairs mucosal immune reconstitution in treated HIV-infected individuals. *J Clin Invest.* **122**(1):62-60.

42. Dronda, F., Moreno, S., Moreno, A., Casado, J. L., Perez-Elias, M. J., Antela, A. 2002. Long-term outcomes among antiretroviral naive human immunodeficiency virus infected patients with small increases in CD4+ cell counts after successful virologic suppression. *Clin Infect Dis*. **35**:1005–1009. .

43. Yeni, P.G., Hammer, S.M., Hirsch, M.S., *et al*. 2004. Treatment for adult HIV infection: 2004 recommendations of the International AIDS Society-USA Panel. *JAMA.* **292**:251-265.

44. Kaufmann GR, Furrer H, Ledergerber B, Perrin L, Opravil M, Vernazza P, Cavassini M, Bernasconi E, Rickenbach M, Hirschel B, Battegay, M., and Swiss HIV Cohort Study. 2005. Characteristics, determinants, and clinical relevance of CD4 T cell recovery to <500 cells/microL in HIV type 1-infected individuals receiving potent antiretroviral therapy. *Clin Infect Di.s.* **41** (3)**:** 361–372.

45. D'Arminio Monforte, A., Testori, V., Adornim F., Castelnuovo, B., Bini, T., Testa, L., Moscatelli, G., Chiesa, E., Rusconi, S., Abeli, C., Sollima, S., Musicco, M., Meroni, L., Galli, M., Moroni. M. 1999. CD4 cell counts at the third month of HAART may predict clinical failure. *AIDS.* **13**:1669-76.

46. Yukl, S. A., Gianella, S., Sinclair, E., Epling, L., Li, Q., Duan, L., Choi, A. L., Girling, V., Ho, T., Li, P., Fujimoto, K., Lampiris, H., Hare, C. B., Pandori, M., Haase, A. T., Günthard, H. F., Fischer, M., Shergill, A. K., McQuaid, K., Havlir,

D. V., and Wong, J. K. 2010. Differences in HIV burden and immune activation within the gut of HIV-positive patients receiving suppressive antiretroviral therapy. *J Infect Dis.* **202**:1553-1561.

47. Hatano, H. H., Hayes, T. L., Dahl, V., Sinclair1, E., Lee, T-H., Hoh1, R., Lampiris, H., Hunt, P. W., Palmer3, S., McCune, J. M., Martin, J. N., Busch, M. P., Shacklett, B. L., and Deeks, S. G. 2011. A randomized, controlled trial of raltegravir intensification in antiretroviral-treated, HIV-infected patients with a suboptimal CD4+ T cell response. *J Infect Dis.* **203**:960-968.
48. Asmuth, D. M., Ma, Z-M., Mann, S., Knight, T. H.; Yotter, T.; Albanese, A.; Melcher, G. P.; Troia-Cancio, P.; Hayes, T.; Miller, C. J.; and Pollard, R. B. 2012. Gastrointestinal-associated lymphoid tissue immune reconstitution in a randomized clinical trial of raltegravir versus non-nucleoside reverse transcriptase inhibitor-based regimens. *AIDS.* **26**: 1625-1634.

CHAPTER 5:

CONCLUSIONS AND PERSPECTIVES FOR ENHANCED THERAPEUTIC STRATEGIES AND THE PREVENTION OF DISEASE PROGRESSION

TABLE OF CONTENTS

5.0. DISCUSSION

Although major advances have been made in the treatment of HIV-1/AIDS, there is still no therapeutic "cure" for HIV-1 infection. Despite the dramatic improvement in quality and duration of life in individuals on suppressive HAART, the virus persists at extremely low levels (<50 copies/mL of plasma) and upon cessation of therapy, re-emerges and replicates to pre-treatment levels. In patients on suboptimal therapy, the virus typically emerges in a multi-drug resistant form [1]. Evidence that the latter is the case; comes from studies showing that treatment intensification with integrase inhibitors and other novel antiretroviral drugs does not lead to therapeutic enhancement, or to the evolution of drug resistance [2-4]. It is becoming increasingly apparent that chronic immune activation and persistent inflammtion are the major factors driving the persistence of HIV-1 in cryptic cellular reservoirs, and that elimination of these reservoirs will require an entirely new therapeutic approach– one aimed at controlling the harmful effects of persistent immune activation and unresolved inflammation [5-8].

To develop strategies capable of suppressing the harmful effects of chronic immune activation and persistent inflammation, we need to first understand the factors fuelling these abnormal immune responses, as well as the factors driving the evolution of more virulent viral quasispecies and the emergence of drug resistant variants of the virus. This issue is not merely of academic interest, it is of critical importance for the advancement of HIV-1 treatment efforts. A number of recent studies have shown that the gut-associated-lymphoid tissues (GALT) are disproportionally affected by HIV-1 and that HAART is only partially effective at restoring CD4+ T cell levels and controlling viral replication in the GIT, even after 10 years of continuous suppressive therapy [1; 5; 9-12]. Other studies have shown that the gut plays a major role, not only in the initial establishment and dissemination of infection, but also in the massive depletion of $CD4^+$ T cells and chronic immune dysregulation and activation that is characteristic of HIV-1 disease and progression to AIDS [14-16]. Despite the importance of the GIT in the pathogenesis of HIV-1/AIDS, there are still many gaps in our understanding of intestinal pathogenesis, and the information that is available was obtained in a variety of different settings and thus, the data are not comparable or maximally informative. The majority of studies have been conducted in North America and Europe in patients infected with HIV-1 B viruses, a viral subtype that accounts for only 10% of global infections [17; 18]. Most of these studies have involved a single site in the GIT and have been performed on HAART-treated patients. In addition, most genetic studies have been cross-

sectional and have involved the analysis of HIV-1 proviral DNA, a substrate that contains a large amount of archived and defective genetic material and thus, may not be truly representative of the replication-competent genetic reservoir [1; 10-12;19-21]. There is paucity of data on subtype C viruses, even though this viral subtype accounts for >50% of infections worldwide and that these viruses are particularly well-adapted to replicate in the GIT possibly due to the presence of an extra NFκB in the promoter region of the virus. This lack of information is due, in large part, to the ethical and technical challenges associated with repeat sampling multiple tissue sites, as well as the small amount of material that is available in "pinch" biopsy specimens.

To begin addressing some of these knowledge gaps, I have used a multi-pronged genetic and immunological approach to answer questions relating to the role of the GIT in the evolution, diversification and dissemination of HIV-1 RNA quasispecies in ART-naïve African AIDS patients with chronic diarrhoea and (or) weight loss and early restoration of $CD4^+$ T cell levels in patients on NRTI-NNRTI-based ART. The genetic study provides an important baseline for studying the contribution of the GIT to the evolution and emergence of drug resistance in patients infected with subtype C viruses. Distinct differences were observed in the diversity and phylogenetic clustering of *env* vs *gag* sequences. In this study there was higher overall diversity of the *env* gene region observed when compared to the *gag* gene region and the higher level of *env* variability in the small intestine, especially in the ileum, compared to the right and left colon. Though there was evidence of trafficking in the *gag* gene between all gut compartments and plasma, there was no statistically significant differences observed in the extent of *gag* diversity in different regions of the small (duodenum, jejunum, and ileum) or large (right and left colon) intestine. In the analysis for the *env* gene region trafficking was also observed, but in addition there was foci of localized genetic evolution, predominantly in the colon. These findings suggest that *gag* and *env* are under different selection pressures and that these selection pressures vary in different regions of the GIT. The factors fuelling this differential evolution of HIV-1 are not known but may be due to tissue-specific differences in the cellular immune response directed against the *env* and *gag* gene regions, or to factors related to virus entry and replication, or to differences in target cell phenotype and availability, all of which would be expected to have a greater impact on the *env* gene.

To determine whether any of the nucleotide substitutions in different parts of the GIT were associated with amino acid changes related to drug resistance, *pol* RT PCR dataset was

analyzed for the presence of reverse transcriptase (RT) and protease (PR) resistance mutations. There was no primary drug resistance mutations detected in any of the 5 GIT sites, or in plasma samples, though substitutions at amino acid positions associated with drug resistance were detected. The first and second series of amino acid substitutions were associated with increased catalytic activity of the protease enzyme, and was linked to an I93L polymorphism associated with increased (hyper) sensitivity to lopinavir, respectively as was shown in previous studies [22; 23]. Though no evidence was observed for the preferential presence of any specific RT polymorphism in a given tissue compartment, some substations exhibited extensive inter-patient and inter-compartmental variation while others were highly conserved among different patients and different tissues of the GIT.

When assessing the relationships between viral replication and immune activation, and the therapeutic response of HIV-1 in different regions of the GIT, focusing on the small vs. large intestine, a profound depletion of $CD4^+$ T cells in the small (duodenum, jejunum, ileum) vs. large intestine (colon) and peripheral blood was observed which was in agreement with what was reported in other studies of treatment-naive patients and of patients on suppressive cART [14; 15; 24-26]. My study demonstrates that, despite the late stage of infection, gradual restoration of $CD4^+$ T cell levels is possible in the colon and blood of African AIDS patients on short-term cART. It also underscores the distinct differences in the small vs. large intestine with regards to $CD4^+$ T cell depletion and reconstitution and highlights the need for detailed comparative studies to better delineate the mechanisms leading to immunological failure in the upper small intestine. An understanding of these mechanisms may lead to the development of highly targeted treatment strategies that enhance $CD4^+$ T cell recovery. If inflammation and immune activation in the small intestine are driven by local factors not directly related to HIV-1 (tissue damage, loss of epithelial integrity, non-specific activation of $CD8^+$ T cells), it is unlikely that drugs targeting HIV replication will be maximally effective with respect to immune recovery. This view is supported by studies showing that treatment intensification with raltegravir, a potent integrase inhibitor, did not have any significant effect on $CD4^+$ T cell levels or immune activation in the GIT, or on HIV-1 RNA levels in plasma [3; 4; 25]. In conclusion there was direct correlation between the analysis of both treatment-naive patients and patients on cART in this study, which showed a profound depletion of $CD4^+$ T cells in the small (duodenum, jejunum, and ileum) as compared to the large intestine (colon).

In particular, as discussed in Chapter 3 of this thesis, it will be important to better understand the effects of ART on reducing virus evolution and trafficking between different tissue compartments as a function of ART-induced suppression of viral replication. Such studies may be restricted to patients with varying levels of "partial" suppression since patients on fully suppressive HAART are inherently more difficult and less informative to the extremely low levels of HIV-1 RNA available for analysis.

The strengths of the work described in this thesis, in comparison to other studies, relate to homogeneity of the patient cohorts (with respect to treatment, stage of infection, and the absence of overt opportunistic infections in the majority of patients), the sampling of multiple tissue sites and the focus on patients infected with HIV-1 subtype C. Additional strengths in the case of genetic studies relate to the fact that these analyses were performed on contemperaneous strains of viral RNA rather than archived DNA, and in the case of immunlogical studies, to the fact that the same patients were sampled both before and during combination ART. Weaknesses of the studies, relate to the small number of patients and the limited amount of material that was available for comprehensive correlative analyses. As emphasized in this thesis, a full appreciation of the role of the GIT in HIV-1 disease pathogenesis will require serial longitudinal sampling of multiple tissue sites. Some of these issues may be best addressed in simian immunodeficiency (SIV) models of intestinal retroviral infection. Chapter 1 of this thesis contains a general overview of the history, virology, evolution of HIV-1 subtypes, immunology and treatment of HIV-1/AIDS, as well as a section on the role of the GIT in AIDS pathogenesis. Chapter 2 contains all of the methods and methods used for the analyses of the results. Chapter 3 contains novel information relating to the differential distribution and evoluation of HIV-1 RNA variants in different regions of the small vs. large intestine, and Chapter 4 describes major differences in the activation and inflammatory status of different regions of the GIT and how this influences $CD4^+$ T cell recovery during ART. These findings underscore the complexity of the GIT and the region differences in this large and dynamic lymphoid organ. The studies also provides new information that should prove useful, not only for therapeutic monitoring, but also for investigating the impact of ART on the evolution of drug resistance and more virulent HIV-1 quasispecies and for the design of strategies to reduce inflammation, reverse tissue damage and restore the integrity of the intestinal barrier.

5.1. REFERENCES

1. Guadalupe, M., Sankaran, S., George, M. D., Reay, E., Verhoeven, D., Shacklett, B. L., Flamm, J., Wegelin, J., Prindiville, T., and Dandekar, S. 2006. Viral suppression and immune restoration in the gastrointestinal mucosa of human immunodeficiency virus type 1-infected patients initiating therapy during primary or chronic infection. *J. Virol.* **80**: 8236

2. Yukl, S. A., Gianella, S., Sinclair, E., Epling, L., Li, Q., Duan, L., Choi, A. L., Girling, V., Ho, T., Li, P., Fujimoto, K., Lampiris, H., Hare, C. B., Pandori, M., Haase, A. T., Günthard, H. F., Fischer, M., Shergill, A. K., McQuaid, K., Havlir, D. V., and Wong, J. K. 2010. Differences in HIV burden and immune activation within the gut of HIV-positive patients receiving suppressive antiretroviral therapy. *J Infect Dis.* **202**:1553-1561.

3. Hatano, H. H., Hayes, T. L., Dahl, V., Sinclair1, E., Lee, T-H., Hoh1, R., Lampiris, H., Hunt, P. W., Palmer3, S., McCune, J. M., Martin, J. N., Busch, M. P., Shacklett, B. L., and Deeks, S. G. 2011. A randomized, controlled trial of raltegravir intensification in antiretroviral-treated, HIV-infected patients with a suboptimal CD4+ T cell response. *J Infect Dis.* **203**:960-968.

4. Asmuth, D. M., Ma, Z-M., Mann, S., Knight, T. H.; Yotter, T.; Albanese, A.; Melcher, G. P.; Troia-Cancio, P.; Hayes, T.; Miller, C. J.; and Pollard, R. B. 2012. Gastrointestinal-associated lymphoid tissue immune reconstitution in a randomized clinical trial of raltegravir versus non-nucleoside reverse transcriptase inhibitor-based regimens. *AIDS.* **26**: 1625-1634.

5. Brenchley, J. M., Price, D. A., and Douek, D. C. 2006. HIV disease: fallout from a mucosal catastrophe? *Nat Immunol* **7**:235–239.

6. Paiardini, M., Frank, I., Pandrea, I., Apetrei, C., and Silvestri, G. 2008. Mucosal immune dysfunction in AIDS pathogenesis. *AIDS Rev* **10**: 36–46.

7. Brenchley, J. M., and Douek, D. C. 2008. HIV infection and the gastrointestinal immune system. *Mucosal Immunol.* **1**:23-30.

8. Estes, J. D., Harris, L. D., Klatt, N. R., Tabb, B., Pittaluga, S., Paiardini, M., Barclay, G. R., Smedley, J., Pung, R., Oliveira, K. M., Hirsch, V. M., Silvestri, G., Douek, D. C., Miller, C. J., Haase, A. T., Lifson, J., and Brenchley, J. M. 2010. Damaged Intestinal Epithelial Integrity Linked to Microbial Translocation in Pathogenic Simian Immunodeficiency Virus Infections. *PLoS Pathog* **6**(8): e1001052, 1-15.

9. Guadalupe, M., Reay, E., Sankaran, S., Prindiville, T., Flamm, J., Mcneil, A., and Dandekar, S. 2003. Severe CD4 T-cell depletion in Gut lymphoid tissue during primary Human Immunodeficiency Virus Type 1 infection and substantial delay in restoration following Highly Active Antiretroviral Therapy. *Society* **77**: 11708-11717.

10. Li, Q., Duan L., Estes, J. D., Ma, Z. M., Rourke, T., Wang, Y., Reilly, C., Carlis, J., Miller, C. J., and Haase, A. T. 2005. Peak SIV replication in resting memory CD4+ T cells depletes gut lamina propria CD4+ T cells. *Nature* **434**:1148-51.

11. Mattapallil, J. J., Douek, D. C., Hill, B., Nishimura, Y., Martin, M., and Roedere, M. 2005. Massive infection and loss of memory CD4C T cells in multiple tissues during acute SIV infection. *Nature* **434**:1093–1097.

12. Mehandru, S., Poles, M. A., Tenner-racz, K., Jean-pierre, P., Manuelli, V., Lopez, P., Shet, A., Low, A., Hiroshi, M., Boden, D., Racz, P., and Markowitz, M. 2006. Lack of mucosal immune reconstitution during prolonged treatment of acute and early HIV-1 infection. *PLoS Med* **3**(12):2335-2348.

13. Veazey, R. S., De Maria, M., Chalifoux, L. V., Shvetz, D. E., Pauley, D. R., Knight, H. L., Rosenzweig, M., Johnson, R.P., Desrosiers, R.C., and Lackner, A. A. 1998. Gastrointestinal tract as a major site of CD4+ T cell depletion and viral replication in SIV infection. *Science* **280**: 427-431.

14. Mehandru, S., Poles, M. A., Tenner-Racz, K., Horowitz, A. Hurley, A., Hogan, C., Boden, D., Racz, P., and Markowitz, M. 2004. Primary HIV-1 infection is associated with preferential depletion of $CD4^+$ T lymphocytes from effector sites in the gastrointestinal tract. *J. Exp. Med.* **200**:761-770.

15. Brenchley, J. M., Schacker, T. W., Ruff, L. E., Price, D. A., Taylor, J. H., Beilman, G. J., Nguyen, P. L., Khoruts, A., Larson, M., Haase, A. T., and Douek, D. C. 2004. $CD4^+$ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *J. Exp. Med.* **200**:749–759.

16. Chase, A., Zhou, Y., and Siliciano, R. F. 2006. HIV-1-induced depletion of CD4C T cells in the gut: mechanism and therapeutic implications. *TRENDS in Pharmacological Sciences* **27**(1):4-7.

17. Walensky, R. P., Paltiel, A. D., Losina, E., Mercincavage, L. M., Schackman, B. R., Sax, P. E., Weinstein, M. C., and Freedberg, K. A. 2006. The survival benefits of AIDS treatment in the United States. *J Infect Dis* **194**:11-19.

18. Vermund, S. H. 2006. Millions of life-years saved with potent antiretroviral drugs in the United States: a celebration, with challenges. *J Infect Dis* **194**:1-5.

19. Poles, M. A., Boscardin, W. J., Elliott, J., Taing, P., Fuerst, M. P., McGowan, I., Brown, S., Anton, P. A. 2006. Lack of decay of HIV-1 in gut-associated lymphoid tissue reservoirs in maximally suppressed individuals. *JAIDS* **43**(1):65-68.

20. Estes, J., Baker, J. V., Brenchley, J. M., Khoruts, A., Barthold, J. L., Bantle, A., Reilly, C. S., Beilman, G. J., George, M. E., Douek, D. C., Haase, A. T., and Schacker, T. W. 2008. Collagen deposition limits immune reconstitution in the gut. *J Infect Dis* **198**:456-464.

21. Levesque, M. C., Moody, M. A., Hwang, K. K., Marshall, D. J., Whitesides, J. F., Amos, J. D., Gurley, T. C., Allgood, S., Haynes, B. B., Vandergrift, N. A., Plonk, S., Parker, D. C., Cohen, M. S., Tomaras, G. D., Goepfert, P. A., Shaw, G. M., Schmitz, J. E., Eron, J. J., Shaheen, N. J., Hicks, C. B., Liao, H. X., Markowitz, M., Kelsoe, G., Margolis, D. M., and Haynes, B. F. 2009. Polyclonal B cell differentiation and loss of gastrointestinal tract germinal centers in the earliest stages of HIV-1 infection. *PLoS Med* **6**:e1000107.

22. Velazquez-Campoy, A., Todd, M. J., Vega, S., and Freire, E. 2001. Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. *PNAS* **98**: 6062-6067.

23. Gonzalez, L. M. F., Brindeiro, R. M., Tarin, M., Calazans, A., Soares, M. A., Cassol, S., and Tanuri, A. 2003. The subtype C of human immunodeficiency virus type 1 protease presents an in vitro hypersusceptibility to protease inhibitor Lopinavir. Antimicrob. Agents Chemother. **47:**2817-2822.

24. Schneider, T., Jahn, H. U., Schmidt, W., Riecken, E. O., Zeitz, M., and Ullrich, R. 1995. Loss of CD4 T lymphocytes in patients infected with human immunodeficiency virus type 1 is more pronounced in the duodenal mucosa than in the peripheral blood. Gut **37**:524–9.

25. Yukl, S. A., Gianella, S., Sinclair, E., Epling, L., Li, Q., Duan, L., Choi, A, L., Girling, V., Ho, T., Li, P., Fujimoto, K., Lampiris, H., Hare C. B., Pandori, M,. Haase, A. T., Günthard, H. F., Fischer, M., Shergill, A. K., McQuaid, K., Havlir, D. V., and Wong, J. K. 2010. Differences in HIV burden and immune activation within the gut of HIV-positive patients receiving suppressive antiretroviral therapy. *J Infect Dis* **202**:1553-1561.

26. Gordon, S. N., Cervasi, B., Odorizzi, P., Silverman, R., Aberra, F., Ginsberg, G., Estes, J. D., Paiardini, M., Frank, I., and Silvestri, G. 2010. Disruption of intestinal CD4+ T cell homeostasis is a key marker of systemic CD4+ T cell activation in HIV-infected individuals. *J Immunol*. **185**:5169–5179.