DEEP SEQUENCING OF HIV-1 ENVELOPE DETERMINES CORECEPTOR USAGE AND PREDICTS VIROLOGIC RESPONSES TO ANTIRETROVIRAL TREATMENT WITH CCR5 ANTAGONIST MEDICATION

by

Luke Christopher Swenson

B.Sc., The University of British Columbia, 2008

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Experimental Medicine)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

May 2014

Abstract

Next-generation sequencing can be used to genotype an array of HIV variants within clinical specimens, in a process referred to as deep sequencing. When directed at the gene for HIV envelope, this approach can be used to generate a high-sensitivity overview of the viral tropism of the HIV quasispecies in an infected individual. Since HIV variants with tropism for the CXCR4 coreceptor are not susceptible to the CCR5 antagonist agent maraviroc, their detection is crucial in order to screen out patients unlikely to achieve viral load declines on maraviroc. There are several assays that test HIV tropism, but it has been unclear as to which are most useful in the clinical setting. The aim of this thesis is to compare next-generation sequencing using massively parallel pyrosequencing against several alternative tropism assays in a total of four large randomized clinical trials of maraviroc. The methodologies are refined and optimized in the early chapters and applied to clinical specimens from a total of 2864 HIV-infected individuals.

The concordance of next-generation sequencing with phenotypic tropism assays reached 87%. Relative to both the original Trofile phenotypic assay and standard population-based sequencing, deep sequencing had higher sensitivity to detect minority non-R5 HIV. Where assays gave discordant results, deep sequencing tended to outperform the comparator assay and was able to better discriminate maraviroc responders from non-responders. Next-generation sequencing had excellent performance in populations of both treatment-experienced and treatment-naïve individuals. It was consistently able to determine coreceptor usage, and to predict which patients would respond to maraviroc. It could be performed using either HIV RNA from blood plasma or HIV DNA from peripheral blood mononuclear cells. Additionally, longitudinal deep sequencing was performed on samples taken prior to maraviroc administration and again at treatment failure. Phylogenetic analyses confirmed that the non-R5 variants present at time of maraviroc treatment failure were derived from variants detected by deep sequencing before treatment was initiated. In conclusion, next-generation sequencing was applied to thousands of samples from phase III clinical trials, and was a superior screening tool to those originally used during trial enrollment. This thesis demonstrates the clinical utility of next-generation sequencing.

Preface

Versions or components of Chapters 1 through 6 have been published in the scientific literature. The candidate is the primary author on all seven of the published manuscripts, and their contents have been reprinted with permission from their respective journals.

This statement is to certify that the candidate was the major contributor to study design, collection and/or supervision of collection of laboratory data, and performed the majority of the data analysis. The candidate also researched and wrote all seven of the manuscripts.

Primary co-authors of these manuscripts include the thesis supervisor (Richard Harrigan), statistical and data analysts (Chanson Brumme, Dennison Chan, Chris Glascock, Art Poon, Conan Woods, Brian Wynhoven, and Xiaoyin Zhong), laboratory technicians (Celia Chui, Winnie Dong, David Knapp, Theresa Mo, and Andrew Moores), and external collaborators.

The details of these publications are as follows:

Swenson, L.C., Boehme, R., Thielen, A., McGovern, R.A., Harrigan, P.R. Genotypic determination of HIV-1 tropism in the clinical setting. *HIV Therapy* 2010; 4(3): 293-303. © 2010 Future Medicine, Inc.

Swenson, L.C., Däumer, M., Paredes, R. Next-Generation Sequencing to Assess HIV Tropism. *Current Opinion in HIV and AIDS* 2012; 7(5): 478-485. © 2012 Wolters Kluwer Health.

Swenson, L.C., Moores, A., Low, A.J., Thielen, A., Dong, W., Woods, C., Jensen, M.A., Wynhoven, B., Chan, D., Glascock, C., Harrigan, P.R. Improved Detection of CXCR4-Using HIV by V3 Genotyping: Application of Population-based and Deep Sequencing to Plasma RNA and Proviral DNA. *Journal of Acquired Immune Deficiency Syndromes* 2010; 54(5): 506-510. © 2010 Wolters Kluwer Health.

Swenson, L.C., Mo, T., Dong, W.W.Y., Zhong, X., Woods, C.K., Jensen, M.A., Thielen, A., Chapman, D., Lewis, M., James, I., Heera, J., Valdez, H., Harrigan, P.R. Deep Sequencing to Infer HIV-1 Co-Receptor Usage: Application to Three Clinical Trials of Maraviroc in Treatment-Experienced Patients. *Journal of Infectious Diseases* 2011; 203(2): 237-245.

Swenson, L.C., Dong, W.W.Y., Mo, T., Demarest, J., Chapman, D., Ellery, S., Heera, J., Valdez, H., Poon, A.F.Y., Harrigan, P.R. Use of Cellular HIV DNA to Predict Virologic Response to Maraviroc: Performance of Population-based and Deep Sequencing. *Clinical Infectious Diseases* 2013; 56(11): 1659-1666.

Swenson L.C., Chui C.K.S., Brumme C.J., Chan D., Woods C.K., Mo T., Dong W., Chapman D., Lewis M., Demarest J.F., James I., Portsmouth S., Goodrich J., Heera J., Valdez H., Harrigan P.R.. Genotypic Analysis of the HIV V3 Region in Virologic Non-Responders to Maraviroc-containing Regimens Reveals Distinct Patterns of Failure. *Antimicrobial Agents and Chemotherapy* 2013; 57(12): 6122-6130.

Swenson, L.C., Mo, T., Dong, W.W.Y., Zhong, X., Woods, C.K., Thielen, A., Jensen, M.A., Knapp, D.J.H.F., Chapman, D., Portsmouth, S., Lewis, M., James, I., Heera, J., Valdez, H., Harrigan, P.R.. Deep V3 Sequencing for HIV Type 1 Tropism in Treatment-Naïve Patients: A Reanalysis of the MERIT Trial of Maraviroc. *Clinical Infectious Diseases* 2011; 53(7): 732-742.

These studies received ethical approval from the Providence Health Care–University of British Columbia Research Ethics Board: H07-00987 & H07-01901.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

3TC — Lamivudine

A — Adenosine

AIDS — Acquired Immune Deficiency Syndrome

APOBEC3G — apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like-3G

ARV — Antiretroviral

ATP — Adenosine Triphosphate

AZT — Zidovudine

BID — Twice daily

C — Cytidine

CA — Capsid protein

CCR5 — C-C motif Chemokine Receptor 5

CD4 — Cluster of Differentiation 4

CTL — Cytotoxic T Lymphocyte

CRF — Circulating Recombinant Form

CXCR4 — C-X-C motif Chemokine Receptor 4

D/M — Dual/Mixed tropic

DNA — Deoxyribonucleic Acid

dNTP — Deoxyribonucleoside triphosphates

EFV — Efavirenz

emPCR — Emulsion Polymerase Chain Reaction

Env — Envelope

ESCRT — Endosomal Sorting Complex Required for Transport

ESTA — Enhanced Sensitivity Trofile Assay

FDA — U.S. Food and Drug Administration

FPR — False-Positive Rate

G — Guanidine

Gag — Group-specific antigen

GALT — Gut-Associated Lymphoid Tissue

gp41 — Glycoprotein 41 (kilodaltons)

gp120 — Glycoprotein 120 (kilodaltons)

gp160 — Glycoprotein 160 (kilodaltons)

GS-FLX — Genome Sequencer FLX

HAART — Highly Active Antiretroviral Therapy

HIV — Human Immunodeficiency Virus

HOMER — HAART Observational Medical Evaluation and Research cohort

IN — Integrase protein

IQR — Interquartile Range

LCB — Lower Confidence Bound

LTR — Long Terminal Repeat

MA — Matrix protein

MERIT — Maraviroc versus Efavirenz Regimens as Initial Therapy

MIP-1α/MIP-1β — Macrophage Inflammatory Protein 1 α or β

MOTIVATE — Maraviroc versus Optimized Therapy in Viremic Antiretroviral Treatment-Experienced

Patients

mRNA — messenger Ribonucleic Acid

MSM — Men who have Sex with Men

MVC — Maraviroc

NC — Nucleocapsid protein

Nef — Negative regulatory factor

NRTI — Nucleoside Reverse Transcriptase Inhibitor

NNRTI — Non-Nucleoside Reverse Transcriptase Inhibitor

n.s. — Not Significant

OTA — Original Trofile Assay

PBMC — Peripheral Blood Mononuclear Cell

PCR — Polymerase Chain Reaction

PDVF — Protocol-Defined Virologic Failure

PEP — Post-Exposure Prophylaxis

PI — Protease Inhibitor

PIC — Pre-Integration Complex

Pol — Polymerase

PR — Protease protein

PrEP — Pre-Exposure Prophylaxis

PSSM — Position Specific Scoring Matrix

pVL — Plasma Viral Load

QD — Once daily

R5 — CCR5-using

RANTES — Regulated-upon-Activation, Normal T Expressed and Secreted

Rev — Regulator of expression of virion proteins

RLU — Relative Light Units

RNA — Ribonucleic Acid

RT — Reverse Transcriptase protein

RT–PCR — Reverse Transcriptase Polymerase Chain Reaction

SDF-1 — Stroma-Derived Factor 1

SU — Surface protein

T — Thymidine

Tat — Trans-activator of transcription

TM — Transmembrane protein

V3 — Third Variable loop

Vif — Virion infectivity factor

Vpr — Viral protein R

Vpu — Viral protein unique

wOBTss — Weighted Optimized Background Therapy Susceptibility Score

X4 — CXCR4-using

Acknowledgments

There are many people to whom I am indebted for their help, input, conversations, collaborations, support, friendship, and generosity.

First, I thank the patients who participated in all the studies I had the privilege of exploring. I am humbled that thousands of people whom I have never met have made my research possible.

I am also indebted to the many co-authors who contributed to the studies contained in this thesis and who provided outstanding data, analyses, and insight.

I am extremely thankful for grants and financial support during my studies from the British Columbia Centre for Excellence in HIV/AIDS, the Canadian Institutes of Health Research, the Canadian Observational Cohort, the National Sciences and Engineering Research Council of Canada, and the University of British Columbia. I would also like to thank the industry colleagues and partners who provided access to samples and data, and who supported much of this research both financially and intellectually.

I am grateful to all the outstanding members, both past and present, of the British Columbia Centre for Excellence in HIV/AIDS Laboratory, Liliana Barrios, Carolyn Beatty, Zabrina Brumme, Dennison Chan, Peter Cheung, Celia Chui, Paul Crosson, Weiyan Dong, Winnie Dong, Erin Gillan, Chris Glascock, Alejandro Gonzalez-Serna, Rob Hollebakken, Susie Johnstone, Jeffrey Joy, Aram Karakas, David Knapp, Chris Lachowski, Richard Liang, Andrew Low, Eric Martin, Chelsea McCullough, Theresa Mo, Andrew Moores, Laurie Nichol, Carmond Ng, Emma Preston, Leslie Rae, Marjorie Robbins, Alissa Sadler, Jenna Spring, Emilie Stevens, Sharon Szeto, Lily Tam, Iris Tao, Jeremy Taylor, Art Poon, Birgit Watson, Conan Woods, Brian Wynhoven, and Xiaoyin Zhong.

Special thanks to my fellow graduate students Chanson Brumme, Vikram Gill, Guinevere Lee, and Rachel McGovern, who are brilliant, kind, and talented. I've truly enjoyed our time learning together as students.

Joel Kaake for being a constant and loyal friend. Amadou Issacs for his many years of support and partnership. Michael Moorhead, who helped me to get the summer job that would turn into a thesis.

My committee — Janet Chantler, Hélène Côté, and Julio Montaner — who gave me invaluable guidance and encouragement throughout my studies.

I feel extremely lucky to have been mentored by my supervisor, Richard Harrigan, who both challenged me and rooted for my success. Thank you, Richard.

Finally, I want to thank my beautiful family. Mom, Dad, Emily, Daniel, and James: thank you for your incredible support, steadfast belief, and enduring love.

Foreword

There was a point in time, not so very long ago by some estimates, when a miniscule virus brushed against a finger-like protein; twisting ever so slightly, it tethered a second protein, and slipped its way into the cell of a new species.

Since that first point, this viral entry has been endlessly and immeasurably repeated, in millions of people throughout the world. The way in which HIV enters cells has been a focus of basic and clinical research for decades. The study of that second protein it tethers — the HIV coreceptor — has far-reaching implications, from immune decline, to drug therapy and even hope for a cure.

By better understanding this early step in the life cycle of HIV, science and medicine have developed ways of slowing it down and even preventing the entry from occurring at all. Cutting-edge approaches have been applied to this line of research in the hopes that with each advance, and others like it, we come closer to the end of HIV.

Chapter 1: General Introduction & Thesis Objectives

1.1 Background

1.1.1 The Origins & Current State of the HIV Epidemic

There were an estimated 35.3 million people living with Human Immunodeficiency Virus (HIV) in 2013 [1]. Each year, 2.3 million people are estimated to become newly infected with HIV, and each year 1.6 million people die from Acquired Immune Deficiency Syndrome (AIDS) [1]. The first report of AIDS was published in 1981 [2], and its causative agent, HIV, was subsequently characterized in 1983 [3]. Since that time, there have been a myriad of studies characterizing almost every aspect of the virus, from its structure and biology to its epidemiology, treatment, and prevention.

There are two major types of HIV. HIV type 1 (HIV-1) comprises the bulk of the epidemic. HIV type 2 (HIV-2) is mostly isolated to Western Africa, with an estimated prevalence of 1 to 2 million people [4,5]. Both HIV types belong to the Lentivirus subfamily of the family *Retroviridae*. These are viruses with RNA genomes that replicate through a DNA intermediate [6]. HIV-1 can be further classified by phylogeny into three groups: M (main), O (outlier), or N (non-M/non-O) [7]. Group M can be further subdivided into subtypes or clades. These are lettered A through K (with some exceptions), or are labelled with more than one letter to indicate circulating recombinant forms [7,8]. Globally, subtype C is the most prevalent HIV-1 subtype, at close to half of all infections [9]. Subtypes A and B are the next most prevalent infections, at approximately 12% each. Subtype A is most common in Eastern Europe, Central Asia, and East Africa, while subtype B is most prevalent in Western Europe, the Americas, and Australia. Two circulating recombinant forms (CRF) exist at global prevalence between 5% and 8%. These are CRF01_AE, which is most common in South-East Asia, and CRF02_AG, which is most common in West Africa [9]. Central Africa — where the HIV epidemic is likely to have first crossed over into humans [10–12] — has the greatest diversity of subtypes [9].

The zoonotic origins of HIV-1 and HIV-2 have been mostly elucidated [13]. Both epidemics arose from from transmission of Simian Immunodeficiency Virus (SIV) infections of non-human primates [11,12,14]. HIV-1 is likely to have crossed over into humans in at least two cross-species transmission events of SIV_{cpz}, the virus that infects some chimpanzees (*Pan troglodytes troglodytes*), and likely gave rise to HIV-1 groups M, and N [12]. HIV-1 group O is related to SIV_{gor} and may have arisen from cross-species infection from gorillas (*Gorilla gorilla gorilla*) or chimpanzees [15,16], and a recent potential new HIV-1 group that is distinct from the other three also appears to have crossed over from gorillas [17]. The HIV-2 epidemic appears to have been derived from SIV_{sm}, which infects sooty mangabeys (*Cerocebus torquatus atys*) [14]. As is customary in the HIV field, hereafter, references to HIV can be assumed to be specifically to HIV-1 unless otherwise indicated.

1.1.2 The Structure & Replication Cycle of HIV-1

Virions of HIV-1 have a diameter of approximately 110 nm, with an internal protein capsid enveloped by a lipid bilayer. Inside the capsid are most of the proteins required for the early stages of infection, as well as two copies of an approximately 9 kilobase-long single-stranded RNA genome. These genome encodes a total of 15 viral proteins [18]. There are nine open reading frames: *gag, pol, vif, vpr, vpu, tat, rev, env,* and *nef*. The Gag, Pol, and Env polyproteins are each cleaved into their smaller active constituents. The *gag* gene encodes for the main viral structural components: the MA (matrix) protein, CA (capsid), NC (nucleocapsid), and p6. The *env* gene encodes the viral proteins of the outer membrane envelope: SU (surface), and TM (transmembrane). The primary viral enzymatic functions are performed by PR (protease), RT (reverse transcriptase), and IN (integrase). These are encoded by the *pol* gene. Additionally, there are six accessory proteins: Vif, Vpr, Nef, Tat, Rev, and Vpu, which have various specialized roles in viral replication [19]. The functions of each these aforementioned proteins will be addressed in turn as the replication cycle of HIV is described henceforth. An illustration of the replication cycle can be found in Figure 1.1. Steps targeted by antiretroviral agents are also denoted in the illustration. Treatment for HIV will be discussed in more detail in Section 1.3.

Figure 1.1: HIV-1 Replication Cycle & Drug Targets

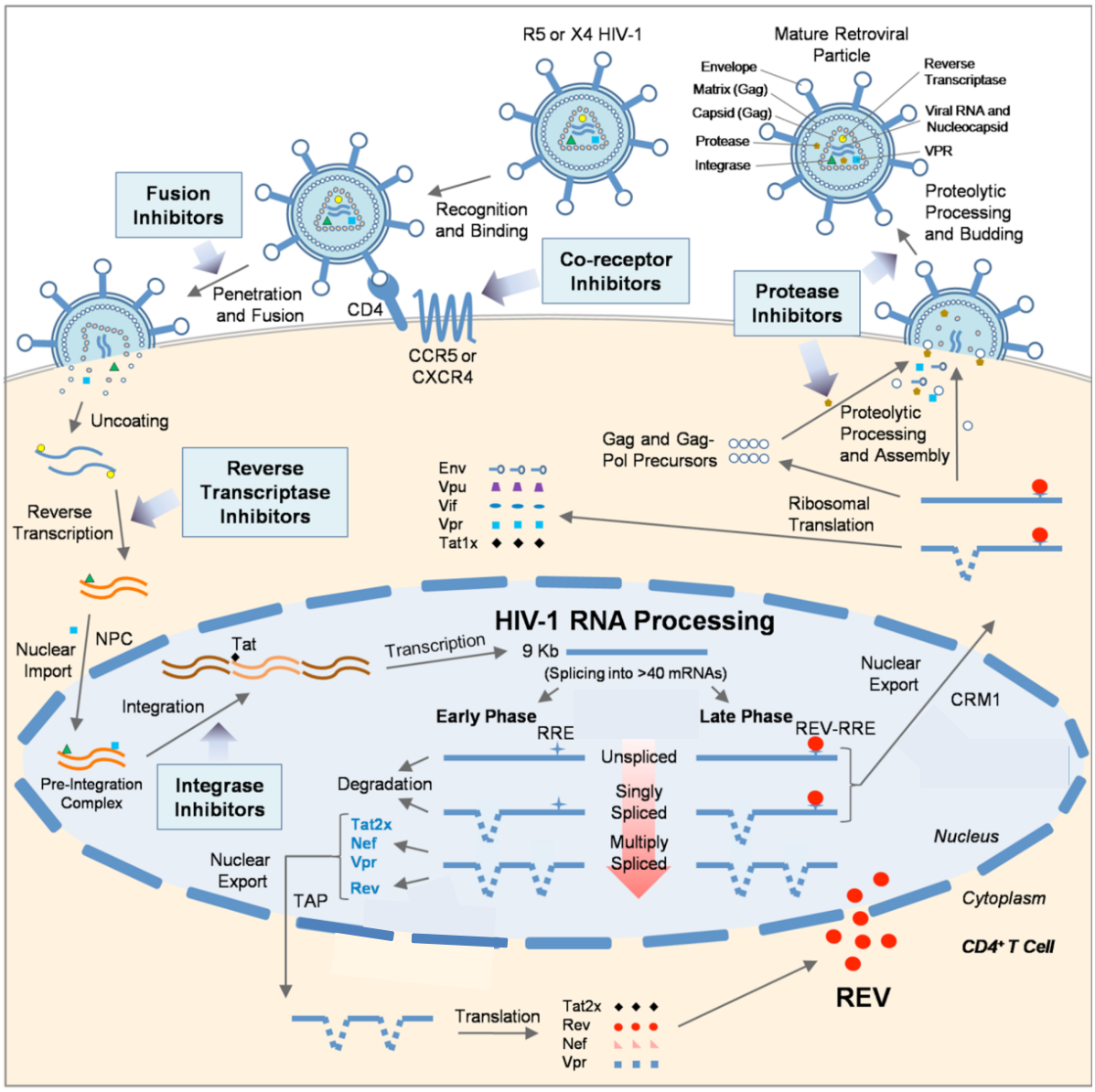


Figure 1.1: HIV-1 Replication Cycle & Drug Targets Depicted is the replication cycle of HIV-1 from cellular entry to progeny release and maturation. Specific antiretroviral classes are indicated adjacent to the step in the life cycle at which they target. This figure is adapted from Wong et al, PLoS Pathogens, with modification courtesy of Hélène Côté. This figure is used under the terms of the Creative Commons Attribution License. © 2013 Wong et al [20]. Wong RW, Balachandran A, Ostrowski MA, Cochrane A (2013) Digoxin Suppresses HIV-1 Replication by Altering Viral RNA Processing. PLoS Pathogens 9(3): e1003241.

Briefly, viral entry into a target cell begins when the viral SU glycoprotein (gp120), binds to the cellular surface receptor CD4, initiating a series of conformational changes in gp120 [21,22]. This exposes certain sites in the glycoprotein — a primary site being the variable V3 loop. Exposure of V3 leads to binding of a chemokine receptor, generally either CCR5 or CXCR4 [22–26]. Host restriction factors such as beta-defensins or RANTES may provide some protection against binding [27,28], but if coreceptor-binding occurs, the entry step can proceed. HIV entry will be described in further detail later in this chapter since this is the key step related to HIV tropism.

After conformational changes in gp120 associated with coreceptor binding occur, the hydrophobic fusion peptide of gp41 is exposed and inserts into the cellular membrane [29]. In addition to the fusion peptide of gp41, there is also the N-terminus-linked heptad repeat region (repeating patterns of seven amino acids), the C-terminus-linked heptad repeat and the transmembrane peptide. In trimeric gp41, the N heptads form a coiled-coil structure onto which the C heptads then bind, thus collapsing the overall structure of gp41 into a six-helix hairpin structure [29]. This conformational change in gp41 decreases the effective distance between its fusion and transmembrane peptides. In this way, the cellular membrane (within which the fusion peptide has been inserted), and the viral membrane (within with the transmembrane peptide exists) are brought into close proximity. This allows the viral membrane to fuse with the target cell membrane [22,30,31].

Fusion of the two membranes releases the viral contents into the cytoplasm. The most important content is the conical viral capsid which comprises multimers of the CA protein [32]. The capsid encases the viral genome and its proteins. A process called uncoating then occurs whereby the capsid shell begins to disassociate, releasing its contents. At this stage, the genome is released in the form of a pre-integration complex (PIC). The PIC comprises the RNA genome, closely associated with several viral proteins [33]. Its release triggers the next step in the life cycle: reverse transcription [22,34].

In order establish a productive infection of a host cell, HIV must convert its RNA-based genome into DNA. This is achieved by the enzyme reverse transcriptase (RT). RT contains two active sites: an N-

terminal RNA- and DNA-dependent DNA polymerase to convert the RNA genome into DNA, and a C-terminal RNase H to digest the RNA strand during complementary DNA strand synthesis [22,35]. First, a short stretch of single-stranded DNA is produced at the 5′ end of the viral RNA, and the complementary RNA is degraded [36]. Next, there is a "template switch" whereby this short stretch of DNA binds at the 3′ end of the viral RNA. Single-stranded DNA is produced complementary to the RNA, and the RNA is degraded by the RNase H activity of reverse transcriptase. DNA-dependent DNA polymerase activity of reverse transcriptase is then used to produce double stranded DNA. During this process, long terminal repeats (LTR) of repeated sequences of DNA are added flanking the viral genome, resulting in a DNA genome of approximately 10 kilobases in length [18]. The activity of reverse transcriptase is highly error-prone. This is a significant contributor to the high degree of variation seen for HIV [37]. The reverse transcription step is targeted by the host restriction factor APOBEC3G, which affects elongation of the reverse transcripts and introduces hypermutation in the growing viral DNA genome [38,39]. These activities are, in turn, antagonized by the viral protein Vif through induction of APOBEC3G degradation [40,41].

Upon successful completion of reverse transcription, the pre-integration complex with its now DNA genome is then imported into the cell nucleus, probably through binding of cellular nuclear import factors or nuclear import signals on the Vpr, IN, CA and/or MA components of the pre-integration complex [33,42,43]. The viral Integrase protein exposes short stretches of DNA at the 3′ ends of the viral genome LTRs. The exposed hydroxyl groups are then used to cut the host cell's chromosomal DNA and join the viral genome to the cellular DNA, with host cell enzymes completing the process [22,44]. Interestingly, the host cell factor, nuclear pore protein (Nup) 358 is recruited by the capsid protein in the pre-integration complex. Nup358 then actually targets the proviral DNA to regions of higher transcriptional activity, thus ensuring that the virus will be actively replicated [33]. The integration process thus results in a stable provirus within the cell's DNA. In some cases the cell may go into a resting state, resulting in a latently infected cell which can carry this provirus over a long period of time [45]. Alternatively, the replication process can continue with the manufacture of progeny virions from the integrated provirus.

Progeny virions are produced by transcription of the DNA provirus. A promoter within the LTR initiates the cell's own transcription process to create plus-strand RNAs from the integrated HIV genome. These plus-strand RNAs can act either as mRNAs for viral peptide synthesis or be encapsidated into progeny virions. The viral Tat protein is required for efficient elongation of the viral plus-strand RNA through recruitment of the cell's positive transcription elongation factor (P-TEF) b. Smaller transcripts can directly enter the cytoplasm from the nucleus. Longer transcripts are bound by Rev, which recruits the host nuclear export factor CRM1 to facilitate the entry of these transcripts into the cytoplasm [22]. The plus-strand RNAs produced are used as templates for making viral proteins, and full-length transcripts can be assembled into viral progeny. In the early phase of viral RNA production, the HIV RNA transcripts are spliced at multiple sites by host cell spliceosomes [19]. Translation of the viral mRNA is performed by the cell's ribosomes [46]. As the concentration of the viral Rev protein increases, it inhibits host cell splicing of viral RNA transcripts, resulting in singly-spliced and unspliced RNA transcripts which are used for translation of larger polyproteins, as well as for the genomic RNA [19]. Translated viral accessory proteins Nef, Vif, Vpu, and Vpr all contribute to modifying the cellular environment to ensure efficient viral replication, persistence, and release. This is achieved through such varied functions as downregulation of CD4 receptors, modulation of the cell cycle, antagonism of host restriction factors, and enabling of immune evasion [47].

The HIV envelope polypeptide, called gp160, is cleaved into gp120 and gp41 by the host protease, furin prior to being transported to the cell surface [48,49]. Translated Gag proteins accumulate on the inner surface of the cellular membrane [46]. A subset of Gag units exists as Gag-Pol polyproteins due to an infrequently occurring ribosomal frame shift [50]. The other viral proteins, as well as the RNA genomes, all co-associate with different portions of the Gag polyprotein to form the components of a complete assembled virion [51,52]. These Gag units also bind a class of cellular proteins called ESCRT complexes. These proteins are usually involved in the formation of vesicles, but are co-opted by the virus to induce budding of the viral progeny from the cell [22,53]. Budding is antagonized, however, by the host restriction factor tetherin, which harnesses the budding virions to the cell membrane, preventing their release [22,41]. This, in turn, is inhibited by HIV's Vpu protein through sequestration of tetherin to an intracellular site distal from the cell surface [41,54].

A lipid membrane derived from the former host cell now surrounds virions that have successfully budded, and these virions then begin a process of maturation. The viral Protease cleaves the Gag and Gag-Pol precursors into their active structural and enzymatic components [22,55,56]. The MA proteins assemble as a matrix on the inner surface of the viral lipid bilayer, while the CA proteins join to form the cone-shaped capsid surrounding the viral proteins and RNA genomes [57]. Once maturation is complete, the virion is infectious and able to initiate a new round of replication in another target cell.

1.1.3 HIV Target Cells

HIV acquisition generally occurs via sexual, perinatal, oral, or intravenous transmission of the virus from an infected individual to an uninfected individual [58,59]. Viral entry is dependent on the presence of CD4 as well as a coreceptor protein — typically CCR5 or CXCR4. Thus, generally only cells with these proteins on their surfaces can be infected by HIV. The cell type first encountered by the virus depends primarily on the route of transmission. In general, there are three main cell types which HIV infects: $CD4^+$ T-cells, macrophages, and dendritic cells [60]. Cells capable of being infected by HIV exist in the genital epithelium, gastrointestinal tract, placenta, central nervous system, and blood and lymphoid circulation. All of these are sites of exposure and transmission of the virus, and represent compartments of viral reservoirs [58,61]. Dendritic cells are not productively infected by HIV in the classical sense, but rather engulf virions and can then transmit them to uninfected cells which come in contact with them. Accordingly, HIV can exploit the normal behaviour of dendritic cells, whereby their function as antigen-presenting cells is hijacked and used to expose HIV target cells to the virus [58].

$CD4^+$ T-cells are the primary target of HIV. These are lymphocytes which coordinate the immune responses to infecting microbes and antigens through signalling to both the cellular and humoral immune systems [62]. $CD4^+$ T-cell counts decline over the course of HIV infection, and decreases in $CD4^+$ T cell counts are associated with increases in risk for opportunistic infections, AIDS, and mortality [63–67]. A $CD4^+$ cell count below 200 cells/mm^3 is associated with AIDS [68]. The mechanism by

which HIV causes depletion of $CD4^+$ T cells is not fully understood. However there are a number of proposed mechanisms involving both direct and indirect pathways. These include cytopathic death of infected cells primarily via pyroptosis, immune-mediated killing of infected cells, apoptosis of uninfected cells from contact with viral proteins, formation of syncytia between infected and/or uninfected cells, impaired production of new T cells or their progenitors, and chronic immune activation [60,69–72]. Apoptosis occurs in productively infected cells, whereas pyroptosis occurs in abortively infected non-permissive cells. The former is mediated by caspase-3 and is minimally immune activating, whereas the latter is mediated by caspase-1 and is extremely inflammatory. Thus, pyroptosis itself may be a driver of further immune activation, creating a cycle of further cell death and inflammation [69,70].

HIV also has a tendency to affect the very cells which respond to it. Memory $CD4^+$ T cells that are specific for HIV and whose function may better coordinate the immune response to the virus are actually preferentially infected by HIV compared to memory T cells without HIV specificity [73]. HIV-specific $CD4^+$ T cells become activated upon encountering viral antigens, remaining in prolonged close contact with HIV containing cells and free virus in the lymph nodes. Furthermore, detection of HIV antigen results in release of inflammatory chemokines which recruit HIV-specific cells to the site of infection. Thus, it has been hypothesized that this recruitment and extended proximity may allow the observed preferential infection of HIV-specific T cells [73]. In addition to $CD4^+$ T cells, HIV is capable of infecting two types of antigen-presenting cells: macrophages and dendritic cells [58,60,74]. Paradoxically, their immune function of antigen presentation may increase viral spread through their interactions with and signalling to T cells [60,75].

A major challenge of HIV infection is that the virus persists long-term in infected individuals. Cells which have been infected can enter a resting state after viral integration, such that they remain latently infected for years and can be subsequently reactivated to initiate new rounds of infection. Long-lived infected cells such as macrophages or resting $CD4^+$ T cells thus act as viral reservoirs and a barriers to curing HIV infection [60,76–79]. There are also several anatomical reservoirs in which HIV can be found [61]. The lymphoid organs, namely the spleen, lymph nodes, and gut-associated lymphoid

tissue (GALT), and their associated lymphocytes have been the focus of most research efforts. Lymphocytes present in systemic circulation are also studied since they are more straightforward to obtain from donors. However HIV replication also occurs in the central nervous system within macrophages and microglia, as well as in T cells and macrophages in the genital and gastrointestinal tracts [61]. Several efforts have been made both to study the latent reservoir as well as to activate it or decrease its size in attempts to purge latently infected cells and establish a functional cure [45,61,78,80,81].

1.1.4 Natural History & Pathogenesis of HIV-1 Infection in Humans

HIV exposure has a fairly low probability of resulting in productive transmission, and is largely dependent on the viral load of the donor [82–85]. HIV becomes established in the host during a phase known as acute infection. Following the transmission of HIV, there is an approximately 2 to 3 week incubation period during which the virus establishes itself in the lymphatic tissue [58,82]. Viral RNA can be detected in the blood plasma approximately 10 days after initial infection, and as the number of infected cells increase, so too does the viral load, reaching a peak viraemia after approximately 3 to 4 weeks [86]. Concomitant with this increase in viral load is a steep depletion in $CD4^+$ T cells due to HIV and CTL-mediated killing. Often the appearance of non-specific constitutional symptoms such as fever or rash which are similar to symptoms in more common viral illnesses [87,88]. The virus also becomes disseminated to various body tissues during this acute infection period.

The immune response during the acute phase is mostly dominated by the innate immune system, especially cytokine, dendritic cell and natural killer cell responses [86,89,90]. At the time of peak viraemia, infected individuals have viral loads often on the order of millions or hundreds of millions of HIV copies per millilitre [36]. Thereafter, the individuals pass through several stages (known as Fiebig stages) as progressively more viral infection markers become detectable by diagnostic assays [91]. The later Fiebig stages are marked by the presence of HIV-specific antibodies produced by B cells [86,92]. However, these antibodies tend to be non-neutralizing, and along with the extensive B cell dysregulation associated with HIV infection, the humoral immune system tends to be poor at containing infection [92].

Following peak viraemia, there is a gradual decrease in viral load until it reaches a relatively stable viral set point averaging approximately 10^5 copies per millilitre. This decrease is temporally associated with elevated $CD8^+$ cytotoxic T lymphocyte (CTL) activity [86,93]. CTLs are capable of detecting HIV-infected $CD4^+$ cells through interactions with Human Leukocyte Antigen (HLA) class I molecules displaying viral peptides. The most common viral peptides targeted by HLA are derived from the HIV Nef, Gag, and Pol proteins [94]. Once detected, CTLs can lyse the infected cells [93,95,96]. Certain HLA alleles are associated with lower viral set points and slower HIV disease progression, while other alleles are associated with higher set points and more rapid progression [97]. This seems to be driven largely by differential responses to viral peptide epitopes. Viral epitopes are themselves bound and displayed to CTLs by different HLA molecules. There seems to be substantive immune control of HIV by this component of the cellular immune system. This is evidenced by the decline in viral load and influence on disease progression that are associated with CTL responses. However, HIV is able to counteract and evade these responses through mutations in the HLA-targeted epitopes (termed escape mutations), and through downregulation of HLA molecules on the surface of infected cells [97,98]. The general inadequacy of the immune system in controlling HIV infection allows expansion of the viral population and depletion of $CD4^+$ cells. Eventually, this leads to the manifestation of HIV-related disease and AIDS.

1.2 HIV-1 Coreceptor Usage

1.2.1 Chemokine Receptors

The human biological functions of the chemokine receptors CCR5 and CXCR4 are to serve as as normal cell signalling proteins activated by the presence of their respective chemokines. Chemokines are small proteins of 70 to 90 amino acids in length, and mediate leukocyte migration and activation at inflammation sites [99]. The natural ligands for CCR5 are macrophage inflammatory protein 1 α (MIP-1α), MIP-1β, and RANTES (regulated-upon-activation, normal T expressed and secreted). These are released by $CD8^+$ T lymphocytes. The natural ligand for CXCR4 is stroma-derived factor 1 (SDF-1)

[21,100,101]. Also known as CXCL12 [102], it is released by fibroblasts in several tissues and organs throughout the body [103]. Both chemokine receptors are seven transmembrane G-protein coupled receptors. This class of proteins acts via intracellular G proteins, which activate several pathways to affect leukocyte chemotaxis and activation at sites of inflammation. For instance, binding of SDF-1 to CXCR4 activates a diverse set of signalling pathways involved in chemotaxis, cell adhesion, transcriptional activation, and cell survival [104]. Interestingly, it has been demonstrated *in vitro* that HIV gp120 can bind to these chemokine receptors and cause similar signal transduction cascades. Furthermore, co-administration of the chemokines with HIV has been shown to inhibit HIV infection in cells [21,104].

There is differential expression of these chemokine receptors depending on cell type. In general, CCR5 is primarily expressed on macrophages, dendritic cells, and activated or memory $CD4^+$ T-cells, while CXCR4 is primarily expressed on naïve $CD4^+$ T cells, B lymphocytes and several types of stem cells [101,103,105]. Interestingly, the ligands for CCR5 can also act as ligands to other chemokine receptors (e.g., CCR1, CCR4). Therefore release of the chemokines can activate several receptors on a variety of different cell types, resulting in an array of activities. In contrast, SDF-1 binds only to CXCR4, and thus has more specific effects solely on CXCR4-expressing cells [21,106].

There is a genetic polymorphism for CCR5 which results in a non-functional and non-expressed version of the protein (this will be discussed in more detail elsewhere in the thesis). Individuals homozygous for this allele appear to have a relatively normal phenotype overall [100]. However, they are over-represented in cohorts of patients with symptomatic West Nile virus infection, and are more likely to die from that infection than patients without the allele, so the defect is not completely benign [107]. In contrast, the CXCR4 protein appears to be essential for ontogenetic development, and it is lethal to 'knockout' the *CXCR4* gene in mice [101]. Small-molecule antagonists of both chemokine receptors have been developed, primarily for their anti-HIV activity [108–111], but also for anti-cancer procedures involving CXCR4 and mobilization of hematopoietic stem cells [102,109,112,113].

1.2.2 Interactions of the Viral Surface Glycoproteins with Cellular Surface Receptors & Coreceptors

The protein HIV uses to attach to human cells is called gp120. Also referred to as SU or surface protein, gp120 is approximately 510 amino acids in length. It is first synthesized at the endoplasmic reticulum as part of the polyprotein gp160, which includes both gp120 and the transmembrane protein, gp41. Within the Golgi, the gp160 precursor polypeptide oligomerizes and is glycosylated with N-linked oligosaccharides [114]. Glycosylation is essential for the correct conformation and stability of the final envelope trimer [114]. The glycoprotein is transported through the Golgi, where its oligosaccharides are trimmed and modified. Gp160 is then cleaved by the cellular endoprotease furin into its constituents: gp120 and gp41 [49,114]. Despite cleavage, these two glycoproteins remain co-associated through weak non-covalent interactions. Together, they assemble in trimers at the cell membrane, with gp41 components embedded within the phospholipid bilayer, and trimeric gp120 protruding away from the cell. During budding, they become incorporated into the viral membrane [57,114].

The gp120 glycoprotein itself can be divided into five conserved regions (C1-C5), which have relatively consistent sequences across many HIV isolates. Interspersed between the constant regions are five variable regions (V1-V5), which can vary widely in both their amino acid sequences and lengths. Gp120 can be roughly divided into three areas: the inner domain, the outer domain, and the bridging sheet linking the two domains. The inner domain is formed by the conserved regions C1 and C5, which are the main contact areas for gp41, while C2, C3 and C4 form a hydrophobic core within gp120 [114]. The outer domain contains higher glycan concentrations, which makes the outer domain less immunogenic, thus protecting the virus from neutralizing antibodies. The variable regions, especially V1/V2 and V3 are exposed on the trimer and provide partial shielding of the CD4 binding cavity within gp120 [114,115]. A very important region of gp120 is the CD4-binding site, which is a cavity located at the interface between the inner and outer domains and the bridging sheet. Within the CD4-binding site there is a crucial hydrophobic cavity which interacts with phenylalanine-43 of CD4 during cellular attachment [24,114–116]. Disruption of this interaction limits HIV entry [117,118].

In its native state, the envelope trimer exists with three of its variable regions V1/V2, and V3 at the apex of the structure [119,120]. These regions are present as large loops due to intramolecular disulphide bonds [115]. For example, the V3 loop is formed by disulphide bonding between its two cysteine resides at codon 1 and codon 35 of V3 [121]. Binding of the CD4 receptor to the trimer results in significant conformational changes wherein each gp120 monomer rotates outward [119]. The V1/V2 loops are rotated away from the apex of the spike and the V3 loop extends directly towards the target cell [119,120,122]. Thus, rearrangement of the trimer upon CD4 binding exposes additional sites important for viral entry (Figures 1.2 & 1.3). The fact that these epitopes are only exposed after CD4 binding means they are partially shielded from possible neutralizing antibodies [119].

The V3 loop is approximately 35 amino acids in length, and is the primary determinant of HIV-1 coreceptor tropism [123,124]. There are two functional domains of V3: the crown (an antiparallel beta sheet formed by codons 11 through 25), and the stem (formed by the amino acids surrounding the crown). It has been demonstrated that V3 interacts directly with CCR5 and CXCR4 [125–128], and variation in HIV *env* and V3 is a significant contributor to viral fitness [129–131]. The V3 stem likely binds to the N-terminus of the coreceptor, while the crown (the most important contributor to coreceptor tropism) interacts with the extracellular loops of the coreceptor [126].

Sequence analysis of HIV-1 V3 loops reveals that those derived from X4 viruses tend to have an overall more positive charge than those derived from R5 viruses, and such charges are especially common at codons 11, 24, and 25 [132,133]. This observation fits well with the putative V3 binding pockets on the HIV-1 coreceptors, since CXCR4 has a high density of negative charge at its V3 binding pocket, while CCR5 does not [134,135]. Thus, the direct evidence of the interaction between coreceptors and V3, as well as the association between coreceptor usage and V3 sequence variation, have led to the development of a number of tropism assays. Whether phenotype-based or genotype-based, these assays almost always include the V3 region.

Figure 1.2: Schematic of Native & Bound Structures of HIV-1 gp120

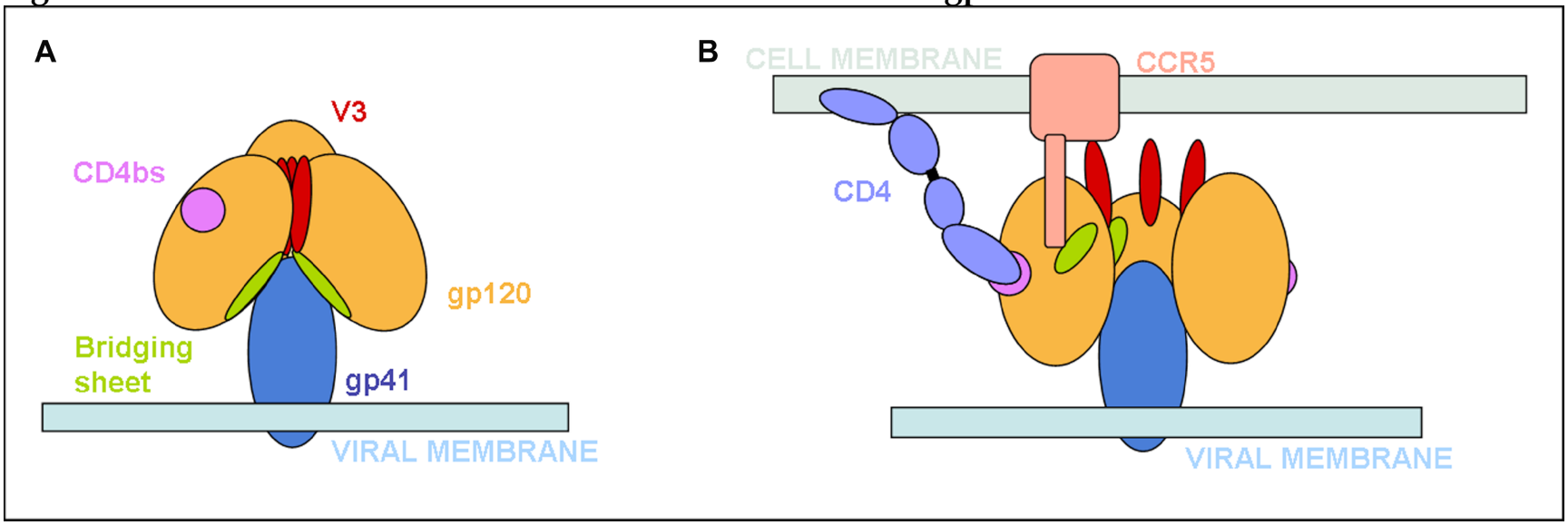


Figure 1.2: Schematic of Native & Bound Structures of HIV-1 gp120. (A) Depicts the native structure of trimeric HIV-1 gp120 prior to binding CD4 or the viral coreceptor. (B) Depicts the conformational changes in gp120 which occur following CD4 binding. The V3 regions are shown in red, and become exposed during the conformational change. The binding of V3 to the CCR5 coreceptor is also depicted. The HIV-1 gp120 trimer is shown in yellow, and the gp41 protein is shown in blue. The CD4 binding site of gp120 (CD4bs) is shown in pink. This figure is used under the terms of the Creative Commons Attribution License. © 2006 Zanetti et al [136]. Zanetti G, Briggs JAG, Grünewald K, Sattentau QJ, Fuller SD (2006) Cryo-Electron Tomographic Structure of an Immunodeficiency Virus Envelope Complex In Situ. PLoS Pathog 2(8): e83.

Figure 1.3: Density Maps of Native & Bound Structures of HIV-1 gp120

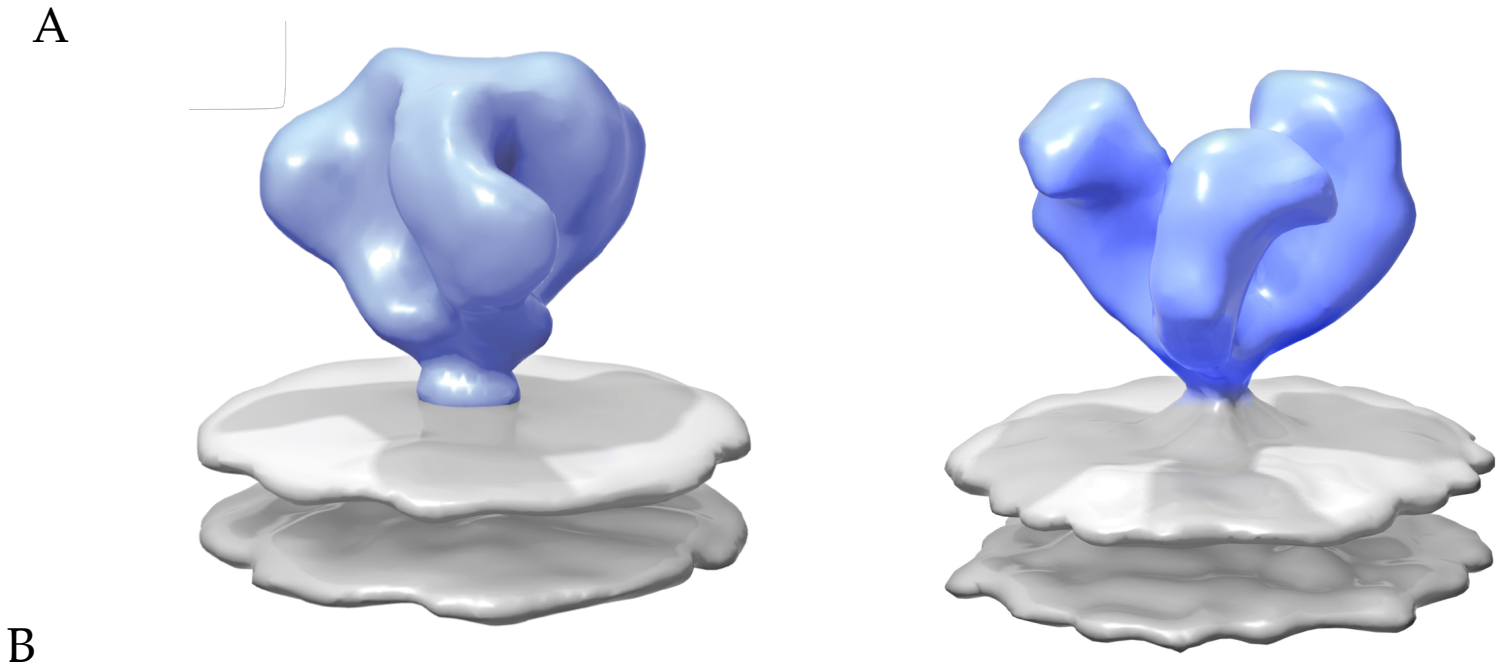


Figure 1.3: Density Maps of Native & Bound Structures of HIV-1 gp120 (A) Depicts a density map of the native structure of gp120 prior to CD4 or coreceptor binding. (B) Depicts a density map of HIV-1 gp120 when experimentally bound to constituents mimicking the receptor and coreceptor. These figures are used under the terms of the Creative Commons Attribution License. © 2010 White et al [120]. Adapted from White TA, Bartesaghi A, Borgnia MJ, Meyerson JR, de la Cruz MJV, Bess JW, Nandwani R, Hoxie JA, Lifson JD, Milne JLS, Subramaniam S. (2010) Molecular Architectures of Trimeric SIV and HIV-1 Envelope Glycoproteins on Intact Viruses: Strain-Dependent Variation in Quaternary Structure. PLoS Pathog 6(12): e1001249.

1.2.3 History of Cellular HIV-1 Tropism

Early in the study of the epidemic, the actual mechanism for cellular entry by HIV-1 was unknown [100]. In the same year as AIDS was first identified, it was discovered that $CD4^+$ T-cells were depleted in patients with AIDS-associated opportunistic infections [137]. Soon, HIV was found to have tropism for $CD4^+$ T-cells [138], though the presence of CD4 was found to be necessary but not sufficient for most cases of HIV entry [100]. Furthermore, it was discovered that certain strains of HIV were faster replicating and able to cause syncytia induction (SI isolates) of cells *in vitro*, while other slower replicating isolates were not able to induce syncytia (NSI isolates) [139]. These empirically derived phenotypes were correlated with more efficient replication in either T-cell lines or macrophages, respectively, such that they were also referred to as T-tropic or M-tropic isolates.

Furthermore, it was discovered that *in vivo,* progression to AIDS was faster for patients with SI versus NSI isolates [139]. It was also noted that M-tropic infections tended to be more common in early HIV infection, with the prevalence of T-tropic isolates increasing with the duration of infection [139]. The observation of differential cellular tropism suggested unidentified factors present in these cell types which rendered them capable of productive infection by one type of strain but not the other. Eventually, it was determined that these factors were the human proteins, CXCR4 or CCR5 [26,140–142]. The natural functions of these viral coreceptors are as receptors for certain chemokines, and HIV entry has been demonstrated to be inhibited in the presence of these chemokines [100].

Nomenclature for HIV tropism has a number of terms which overlap and have similar meanings but are used in different contexts. Broadly, T-tropic strains are also SI and use CXCR4, while M-tropic strains are NSI and use CCR5 [106]. Dual-tropic strains of HIV are capable of using either coreceptor. For convenience, viruses that use CXCR4 for entry are called X4, and those that use CCR5 are called R5 [100]. Occasionally, dual-tropic HIV is signified by R5X4. However, often R5X4 is grouped in with X4 viruses, or both are collectively called non-R5. Furthermore, an eclectic gallimaufry of alternative

proteins has been demonstrated to function as HIV coreceptors [143]. Many of these are also chemokine receptors though their relevance *in vivo* appears to be small with respect to HIV infection [144].

1.2.4 Clinical Relevance of HIV-1 Tropism

There are a number of effects that HIV-1 coreceptor usage has on viral pathogenesis and clinical outcomes during infection [145]. Studies of primary infection with subtype B HIV-1 have generally found that CCR5-using viruses tend to be more prevalent during this early phase of disease [146,147]. Furthermore, R5 HIV tends to be transmitted to the recipient even if non-R5 variants are present in the donor [148]. The reasons for this seemingly preferential transmission are unclear. Some studies have proposed a biological bottleneck such as the presence of CCR5-expressing cells such as Langerhans cells at the sites of sexual transmission. This biological bottleneck may select for CCR5-using viruses during transmission [149,150]. Moreover, there may be additional biological factors which allow for selective expansion of R5 HIV variants following transmission [145]. However, presence of non-R5 variants in recent infections has been documented in a minority of patients [151,152]. Alternatively, some authors have proposed that transmission is a stochastic event and therefore that transmission of R5 HIV is only more frequent because more donors tend to have R5 HIV in the first place [153].

As previously mentioned, human variation in the gene for CCR5 can also influence HIV infection [106,154,155]. A 32 base pair deletion within *CCR5*, denoted *CCR5 Δ32*, is relatively common (approximately 10% prevalence) in individuals of European descent [154,156,157]. Homozygosity for *CCR5 Δ32* is rarer (approximately 1% prevalence) but is associated with almost complete resistance to infection by HIV [156,157]. However, this mutation is not protective against HIV capable of using CXCR4. Infection of homozygotes with X4 HIV has been documented but only rarely [158,159]. Individuals heterozygous for *CCR5 Δ32* do not seem to be significantly protected against infection, but may have a slower disease course, though a higher risk for carrying X4 variants [154,155,160]. Finally, the observation that homozygotes are mostly protected against HIV infection lends further evidence that HIV exposures tend to be with CCR5-using variants.

Regardless of the reason for the predominance of CCR5-using HIV in early infection, CXCR4-using variants can evolve over the course of disease, in a process known as coreceptor switching, and quasispecies using both or either coreceptor may coexist within an individual. Non-R5 variants likely evolve from R5 variants over the course of infection [145,161]. This process may be gradual or sudden, and the evolution from R5 to X4 may traverse a fitness valley which may be an obstacle to coreceptor switching [162–164]. Thus, there are constraints on the evolution of HIV from CCR5 to CXCR4 usage, which one group has commented requires the virus "to make the right amino acid substitution in the right place at the right time" [131] – not a simple matter.

HIV virions with different coreceptor usage can infect different target cells more efficiently, such that there may be compartmentalization of HIV coreceptor usage depending on the cell type [105,145,165,166]. The evolution of CXCR4-using HIV is associated with an accelerated decline in $CD4^+$ T cells, and faster disease progression [145,167,168]. This is probably related to an ability to infect a larger proportion of $CD4^+$ T-cells, and also due to the higher cytotoxicity of CXCR4-using viruses compared to their CCR5-using counterparts [105,169]. The ability of HIV to use CXCR4 is also an independent factor which increases the risk of progression to AIDS by approximately seven-fold [139,167,170]. Non-R5 HIV infection has also been linked to higher mortality in some studies, though not in others [139,155,171–173].

The discovery and characterization of the HIV coreceptors, as well as their association with clinical outcomes, led to the development of a number of assays designed to assess the coreceptor usage of HIV obtained from clinical isolates. Phenotypic assays were the first types developed to determine tropism. Later, by linking the results of these phenotypes with the *envelope* genotypes of multiple HIV variants, there was an emergence of genotypic tropism assays [174,175]. This pattern echoes that seen in the earlier stages of HIV drug resistance research. Assays were originally developed to determine whether phenotypic resistance was present, and these assay results were later correlated with mutation profiles which could be used to infer the resistance phenotype [176–184].

1.2.5 Phenotypic Assays to Detect HIV-1 Coreceptor Usage

The earliest assay for HIV tropism was based on detection of the syncytium-inducing (SI) phenotype through co-culture of patient derived peripheral blood mononuclear cells (PBMC) with MT-2 cells *in vitro* [185]. This was an indirect phenotypic assay in that it could not determine which coreceptor was being used. Indeed, the HIV-1 coreceptors had not even been discovered at the time of the development of the MT-2 assay. However, it was later concluded that viruses with an SI phenotype tend to use CXCR4, while those with an NSI phenotype tend to use CCR5 [186]. A number of other phenotypic assays which are more direct measures of coreceptor usage have since been developed and described [175,187–193]. Generally, these phenotypic assays tend to use patient-derived viruses or recombinant viruses which are then used to infect reporter cell lines that constitutively express CD4 and either CCR5 or CXCR4.

The most widely used phenotypic tropism assay has been the Trofile assay [194]. This assay has been modified to improve the detection of low-prevalence X4 variants in clinical samples, and is now known as the Enhanced Sensitivity Trofile Assay (ESTA) [195]. In the assay, gp160 envelope sequences are amplified from patient plasma samples by using reverse transcriptase polymerase chain reaction (RT-PCR). These gp160 sequences are then inserted into pseudotyped viral vectors containing a luciferase reporter gene. These are then used to infect cell lines that express either CCR5 or CXCR4. Luciferase activity is measured in relative light units (RLUs), and is elevated upon successful infection of a cell line by the patient-derived pseudoviruses. A coreceptor antagonist drug is then added and infection is confirmed by a subsequent decrease in RLUs. High luciferase activity in the CCR5 cell line or CXCR4 cell line is indicative of R5 or X4 virus, respectively. Activity in both cell lines indicates a dual-tropic or mixed-tropic viral population, which is designated D/M [194]. Disadvantages of phenotypic assays include their high labour intensity and long turn-around time, which ranges from approximately two weeks for Trofile to up to five weeks for the MT-2 assay [191]. Accordingly, there have been attempts to reproduce phenotypic assay results using genotypic methods.

1.2.6 Sequencing-Based Genotypic Assays to Determine HIV-1 Tropism Using Bioinformatic Interpretation

At their core, genotypic tropism assays typically involve amplification of a portion of the HIV envelope gene, followed by sequencing, and interpretation of whether that sequence is likely to have X4 or R5 behaviour. The sequence input has traditionally been obtained by population-based (Sanger) sequencing. The sequence must then be interpreted manually by a technician. Interpretation involves identification of secondary peaks in the sequence chromatogram, which are indicative of a quasispecies mixture in the sample. Alternatively, there is automated software called RECall which can perform this sequence interpretation step with high accuracy and without manual intervention. Use of RECall for sequence interpretation increases both productivity and reliability [196]. An interpretation system or bioinformatic algorithm can then be used to infer the likely sequence tropism.

Early, more rudimentary genotypic assay interpretation systems tended to be "rulesbased", and focused on identifying basic, positively charged amino acids at certain codons in V3 and/or the net charge of the amino acid sequence [133]. In addition to codons 11, 24 and 25, many other codons within the V3 loop also have different amino acid compositions. Variation in V3 amino acid substitutions tends to cluster with the use of one coreceptor or another. As a result, a number of bioinformatic algorithms have been developed which use this genotypic information to infer a phenotype [197–202]. Arguably the two most commonly used algorithms have been the Position Specific Scoring Matrix (PSSM) [203] and geno2pheno [197,204]. These algorithms are trained on a set of samples with both known phenotype results and envelope sequences. These phenotypes may be from a Trofile-like assay such as for geno2pheno and $PSSM_{X4/R5}$ or may be from the MT-2 assay as in $PSSM_{SI/NSI}$ or $PSSM_{C}$ [197,203–205]. A sequence submitted for interpretation by the algorithms is then assessed for its similarity to typically X4 or typically R5 sequences.

Bioinformatic algorithms can be assembled and trained in a number of ways, each with its own strengths and weaknesses [197]. Support vector machines (SVM), artificial neural networks, linear regression models, position-specific scoring matrices (PSSM), and rules-based algorithms have been

most commonly used for bioinformatic prediction of HIV tropism [197,204]. Together, these represent statistical learning methods which fit a model based on a set of training data. This trained model is then applied to "unseen" datasets to predict the outcome of interest (in this case, HIV coreceptor usage). In essence, the output of the algorithms is a rating of how X4 or how R5 a sequence is likely to be. While this value is typically a continuous variable, a binary category can be applied using a pre-specified cutoff. For instance, geno2pheno gives an output known as a false-positive rate (FPR). The lower this value, the higher the likelihood is that the sequence is X4. Establishing a false-positive rate cutoff of 10, for example, would categorize any sequence with an FPR ≤10 as being X4, and any sequence >10 as being R5. Various cutoffs for these algorithms have been proposed, and no cutoff has been firmly established or widely adopted for any algorithm. Often a cutoff is established as a balance between sensitivity for correctly identifying an X4 sequence, and specificity for correctly identifying an R5 sequence.

The two most common algorithms for bioinformatic prediction of HIV tropism are $PSSM_{X4/R5}$ and geno2pheno. These algorithms were developed using position-specific scoring matrices, or support vector machines, respectively. Given that these are also the main algorithms used in the studies contained in this thesis, their design and performance will be reviewed here in additional detail with information obtained from Jensen et al (2003) and Sing et al (2007) [203,204].

1.2.7 The Development & Performance of the Bioinformatic Algorithms $PSSM_{X4/R5}$ & Geno2pheno

PSSM-based models are used to distinguish significant differences in the distribution of amino acids between groups of sequences which have been sorted by an empirically-obtained character – for example, phenotypic CCR5-capable or CXCR4-capable groups. The training set for $PSSM_{X4/R5}$ comprised a total of 213 V3 sequences obtained from 177 HIV-infected individuals. Phenotypic results for these sequences revealed 168 R5 sequences (79%), 17 X4 sequences (8%), and 28 dual-tropic sequences (13%) capable of using either coreceptor. The latter two sequence sets were combined into a single non-R5 category. This training set was used to generate a 35×20 matrix: 35 being the length of a typical V3 loop, and 20 being the number of amino acids. Each of the resulting 700 sites was

assigned a likelihood ratio score. This likelihood ratio is given by the natural log of the ratio between two frequencies: the frequency that a particular amino acid is seen at that position in the training set of non-R5 sequences divided by the frequency that the same amino acid is seen at that position in the R5 sequences. Thus, if a particular amino acid (for example an arginine at codon 11) is seen more frequently in non-R5 sequences than R5 sequences, then the likelihood ratio assigned to that particular site would be positive. Conversely, sites which are more common in R5 sequences receive a negative likelihood scores. Then, when a given V3 sequence is submitted to the matrix, each observed site receives a likelihood ratio score, and these are summed up to give a total PSSM score, with higher scores representing sequences more likely to be X4 and lower scores representing those more likely to be R5.

In order to classify sequences into the binary tropism categories, a PSSM cutoff value must be chosen, above which a sequence is classified as X4 and below which a sequence is classified as R5. For $PSSM_{X4/R5}$, a bootstrapping procedure was used whereby multiple sub-samplings of the dataset were used to generate new matrices (essentially "mini" PSSMs) that were then applied to the dataset as a whole. For each of these new, smaller matrices, an optimized cutoff was calculated which maximized the number of true positives (X4 variants higher than the cutoff), maximized the number of true negatives (R5 variants lower than the cutoff), and minimized the number of false positives and false negatives (R5 or X4 variants incorrectly falling above or below the cutoff, respectively). Each sub-sampling matrix gave an optimized cutoff value, and the distribution of these values was plotted. The 5^{th} and 95^{th} percentiles for the distribution of the optimized cutoffs were -6.96 and -2.88, respectively. Thus, for two sequences, one scoring below -6.96 or one scoring above -2.88, there is a very high likelihood that each sequence would be R5 or non-R5, respectively. In other words, these values represent cutoffs whereby a given score can be confidently used to predict tropism status of a sequence, while scores falling in between these cutoffs have less certain tropism associations and were classified according to the 11/25 rule. $PSSM_{X4/R5}$ was subsequently applied to two independent datasets for the purposes of validation. The algorithm was determined to have 84-89% sensitivity to correctly identify non-R5 sequences, and 96-100% specificity to correctly identify R5 sequences. A caveat of this performance is that it was on clonal sequences. Actual clinical samples are invariably

more heterogeneous, containing multiple quasispecies with different sequences and possibly different coreceptor usages. Therefore, applications to clinical samples will tend to have lower performance.

In addition to $PSSM_{X4/R5}$, the geno2pheno algorithm is also used in this thesis to infer the coreceptor usage of V3 sequences. Geno2pheno was developed using support vector machine methodology, which is a machine learning method that categorizes input examples (like sequences) based on a large number of features. Features that are found to be significant classifiers (e.g., an arginine at codon 11 of V3) are assigned weights based on their ability to classify inputs. The training set for the SVM which was used to generate geno2pheno comprised 1110 samples from 332 patients. A total of 769 samples had R5 phenotypes (69%), 210 had X4 phenotypes (19%), and 131 had dual-tropic phenotypes (12%). Similar to $PSSM_{X4/R5}$, there are approximately 700 features assessed in the SVM, each with a weight based on its ability to predict tropism status. To illustrate, significant X4 features in order of decreasing weight were the presence of the residues 13Y, 11R and 20V. Significant R5 features were the presence of residues 24G, 18S, and 13P.

Overall, the algorithm had 76.4% sensitivity and 92.5% specificity for inferring the coreceptor usage of clonal isolates. However, when assessed on 952 clinical isolates (containing HIV quasispecies mixtures), the SVM-based predictions performed worse compared to the clonal data: 39.8% sensitivity and 93.5% specificity. Interestingly, however, this performance could be improved by incorporating other features into the SVM model in addition to the amino acid features. Four additional "clinical" features which improved performance were: the $\log_{10}$ of the CD4 percentage, host heterozygosity for the *CCR5 Δ32* allele, number of ambiguous amino acids in V3, and the presence of insertions or deletions in V3 ("indels"). Lower CD4%, presence of *CCR5 Δ32*, higher numbers of ambiguous amino acids, and higher numbers of indels were all associated with a higher likelihood of having CXCR4-capable HIV. Including these parameters increased sensitivity on clinical specimens to greater than 60%. These additional features are biologically plausible markers of CXCR4 tropism. For instance, sequence diversity (i.e., increased numbers of amino acid ambiguities or indels) has been linked to X4 tropism and disease progression [206,207]. Additionally, CXCR4 propensity

is also associated with lower $CD4^+$ T cell counts, and therefore lower CD4 percentages [145,169]. Furthermore, *CCR5 Δ32* heterozygosity decreases the availability of CCR5 coreceptors for HIV entry, and it is plausible that this may manifest as a viral environment with increased selective pressure to change coreceptor usage. Indeed, such an association has been demonstrated in a separate study [155].

1.2.8 Alternative Genotypic Assays & Comparisons between Genotypic & Phenotypic Assays

Aside from population-based sequencing of HIV-1 envelope followed by sequence interpretation, there are a handful of other genotypic assays used for determining coreceptor tropism. For example, heteroduplex tracking assays use labelled probes to form duplexes with PCR-amplified envelope sequences, and these are then analyzed by electrophoresis to determine if X4 probes were bound [208,209]. Allele-specific PCR has also been used to selectively detect X4 or R5 sequences from patient samples using real-time PCR [210]. Additionally, there have been efforts to determine HIV-1 coreceptor usage from integrated or cell-associated HIV DNA in order to test patients with low or undetectable plasma viraemia [211–215].

Although most genotypic tropism assays tend to use V3 sequences as their input, other envelope regions outside of V3 are also important for HIV-1 tropism and may give additional insight. Amino acid substitutions in the other variable and constant regions of HIV-1 gp120 have been correlated with coreceptor tropism, as have substitutions in the gp41 transmembrane protein [216–220]. Furthermore, variation in peptide length and the number of N-linked glycosylation sites in the variable regions of gp120 tend to increase over time during HIV-1 infection. These factors have also been found to be associated with tropism [207]. Additionally, different HIV-1 subtypes can have different prevalences of X4 or R5 isolates, and bioinformatic algorithms may need to interpret the tropism of different subtypes in different manners [8,205,221,222]. There is even a tropism assay that is neither classically genotypic nor phenotypic. Called the maraviroc clinical test, patient tropism is assessed by response to short-term monotherapy with the CCR5 antagonist maraviroc [192,223–226]. In this scheme, an R5 result is given if by day eight of maraviroc monotherapy, the patient achieves a viral load <40 copies/mL or

at least a 1 $\log_{10}$ decrease in viral load from day one. The concordance of this approach with the Trofile assay has been measured as being 94% [224].

Since genotypic assays are designed to mimic phenotypic assay results, they are often assessed against phenotypic assay results in order to determine performance. Genotypic assays often perform very well on clonal sequences (i.e., not a mixture of quasispecies), and have given sensitivities as high as 99% [227]. However, clinical isolates derived from patients are composed of heterogeneous mixtures of diverse sequences, where X4 sequences may exist at low prevalence relative to R5 sequences [145]. Population-based sequencing has a reported detection threshold of approximately 20% of the viral population, meaning that quasispecies present at lower prevalence are not reliably detected [228]. Thus, while clonal sequences are often predicted easily by bioinformatic algorithms, clinical isolates containing minority X4 variants can often be misclassified as R5 by genotyping but still have an X4 biological phenotype [229]. This can result in genotypic approaches giving wildly varying sensitivities relative to phenotypic assays, even while using the same algorithm [187,190,229,230].

Furthermore, phenotypic assays themselves can give discordant results [190]. However, the traditional "gold standard" for determining HIV tropism has been the phenotypic assay. Minority X4 variants are often better detected by phenotypic assays than standard, population-based sequencing, and these variants can lead to treatment failure on CCR5 antagonists [175,195,229,231,232]. Therefore, a genotypic assay with higher sensitivity is likely to be needed in many clinical contexts and in order to better correlate with phenotypic assay results. A major advance in HIV genotyping has been the development of next-generation sequencing and its use for detecting low-level HIV quasispecies.

1.2.9 Next-Generation Sequencing to Detect Minority HIV-1 Variants

As alluded to in the previous section, genotypic tropism testing occasionally lacks sufficient sensitivity to detect minority non-R5 variants in clinical isolates. However, next-generation sequencing technology is an important alternative to traditional population-based sequencing methods. Next-generation sequencing allows for high-resolution detection of HIV-1 minority

variants, such that it can estimate the actual proportion of the viral population which may be of particular interest. For example, this methodology can be used to quantify subpopulations harbouring resistance mutations or CXCR4 tropism. This approach is also referred to as deep sequencing, since it can probe deep into the viral "swarm". Minority drug resistant or CXCR4-using variants have been shown to impact clinical outcomes on antiretroviral treatment, so their detection is an important tool in optimizing and personalizing therapy [233,234].

Currently, there are a number of next-generation sequencing platforms in development or operation. Many foundational studies in HIV-1 deep sequencing have been performed on platforms developed by 454 Life Sciences [232,235–241]. The next-generation instruments used to generate the results for this thesis are all pyrosequencing-based. Therefore, a more detailed description of this process will be given. In order to be sequenced on a Roche/454 Life Sciences Genome Sequencer FLX (GS-FLX) or GS Junior, a PCR-amplified DNA library must be further amplified in a step known as emulsion PCR. Emulsion PCR is performed by combining the DNA library with microbeads at a concentration such that approximately one DNA molecule is associated with one microbead [242]. This allows clonal amplification of DNA on the surface of microscopic beads, which increases the strength and purity of the available signal. The PCR primers used to create the library are designed with a short stretch of nucleotides complementary to oligonucleotides on the surface of the beads, which allows the beads to bind the DNA. The beads, DNA, and PCR reagents are mixed with oil and shaken to create an emulsion. This process generates small, independent micelle "microreactions" which are then subjected to thermal cycling, After emulsion PCR, these beads are then placed onto a plate with hundreds of thousands of small wells to hold their position. The sequence of the DNA on the beads is then determined by pyrosequencing [242].

The name pyrosequencing is derived from the pyrophosphate moiety that is released during sequencing. DNA complementary to the target DNA library on the beads is produced with successive washes of each of the four nucleotides (Thymidine (T), Adenosine (A), Cytidine (C), Guanidine (G)) [232]. This process is a type of "sequencing-by-synthesis", where the sequence of the DNA library is determined as it is being synthesized. The incorporation of a nucleotide (or nucleotides) by a DNA

polymerase results in the release of pyrophosphate. This is then converted to adenosine triphosphate (ATP) by ATP sulfurylase. The energy from the ATP is used by the luciferase enzyme to oxidize luciferin, causing the release of photons which are detected by a camera [242,243]. Each type of nucleotide is washed over the sequencing plate independently, and a wash containing apyrase degrades any remaining unincorporated nucleotides before the next nucleotide wash. Thus, light generated from any given nucleotide wash can be attributed to incorporation of that base only, thereby allowing the sequence to be determined. Since these nucleotides do not contain termination moieties, multiple nucleotides may incorporate into the nascent DNA strand if there is a stretch of the same type, known as a homopolymer. The number of bases is proportional to the light intensity; however, the camera is poor at distinguishing the number of nucleotides incorporated when the homopolymer exceeds approximately six bases in length [242]. Thus, this method has higher error rates within homopolymer regions [211,240,242,244,245].

In addition to these pyrosequencing platforms however, a number of alternative next-generation sequencing instruments are in use or development [246–249]. These employ different principles in signal amplification and sequencing chemistry but have in common the capacity for simultaneous DNA sequencing from multiple samples. For example, there are several sequencers manufactured by Illumina which are being increasingly used for HIV applications. For these platforms, bridge PCR amplifies the template, creating a clonal cluster tethered to a solid chip [246]. Next, sequencing-by-synthesis occurs by washing four reversibly-terminating fluorescent-labelled nucleotides. This results in a single nucleotide binding to a growing DNA strand that is complementary to the input strand. Imaging of the fluorescent signal of the bound nucleotide takes place, followed by cleavage of the fluorescent tag and 3′ termination moiety to allow further extension and sequencing [247,250].

The SOLiD platform uses emulsion PCR, where amplification occurs on beads within microreactor oil micelles [247,251], followed by sequencing-by-synthesis using a DNA ligase. A total of 16 probes is used, with each carrying a two base-pair stretch and labelled with one of four fluorescent tags. During sequencing, one probe at a time is ligated to the complementary DNA strand. Since there are four different probes possible for each tag, ligation is repeated a number of times, each time beginning one

base pair downstream of the previous starting site. Thus, each base in the sequence is interrogated at least twice, allowing for deconvolution of the actual sequence and resulting in potentially lower error rates [247,250,252].

Real-time sequencing approaches are able to observe DNA synthesis directly during sequencing. The Pacific Biosciences platform uses immobilized DNA polymerase molecules which are recorded in real-time while they synthesize DNA with fluorescently labeled deoxyribonucleoside triphosphates (dNTPs) [253]. Another real-time sequencing approach from Oxford Nanopore also looks extremely promising [244,246,254]. Finally, sequencing approaches that do not use light or fluorescence have also been developed. For instance, the Life Technologies Ion Torrent platform detects changes in pH caused by hydrogen ions that are released when nucleotides bind to a growing DNA strand during sequencing [247,255]. The read lengths for this approach are continuing to increase [256].

When compared, next-generation sequencing platforms all have roughly comparable performance [248]. Any or all of these platforms have the potential to revolutionize sequencing in HIV and other diseases and applications. Next-generation sequencing may have special significance in terms of HIV coreceptor usage since minority non-R5 variants have been found to commonly exist within majority R5 populations [232]. Minority variants have also been shown to be relevant to development of resistance to a number of antiretroviral medications [234,257,258]. While treatment for HIV has made great advances, resistance to antiretrovirals remains a key barrier to successful therapy.

1.3 Treatment of HIV

1.3.1 Antiretroviral Treatment of HIV Infection & Development of Resistance

In the approximately 15 years after it was first identified, AIDS became the leading cause of death amongst adults aged 25 to 44 in the United States [259,260]. By 1985, an antibody test was developed to screen for HIV infection, but vaccine candidates and treatment for the virus were slower to come to fruition [261,262]. The first nucleoside reverse transcriptase inhibitor (NRTI), zidovudine, was approved

in 1987 by the U.S. Food and Drug Administration (FDA). NRTIs lack a 3'-hydroxyl group and therefore act as chain terminators to the nascent DNA strand during reverse transcription [263]. Prior to 1996, a handful of additional antiretroviral agents were approved for use by the FDA. These included four NRTIs, as well as the first protease inhibitor (PI), a class of antiretrovirals which bind to the catalytic domain of the viral protease, preventing its function [262,263]. However, monotherapy with a single antiretroviral agent proved to have disappointingly limited efficacy [264]. There is extremely rapid production of virus during infection untreated, or suboptimally treated infection. It is estimated that approximately 1 billion virions are produced each day in untreated individuals [265]. This fact, in combination with the error-prone nature of HIV reverse transcriptase led to the development of drug resistance in many patients [182,183].

During this decade, additional antiretrovirals were developed, including a new class: the non-nucleoside reverse transcriptase inhibitors. NNRTIs inhibit the viral reverse transcriptase protein by binding near its catalytic domain and allosterically inhibiting it. As new agents within existing classes continued to be developed [262,263,266], they were ultimately tested in combinations. Finally, around the mid-1990s, results of trials of triple therapy were released. These trials showed that combination regimens could both lower plasma viraemia and slow disease progression [18,262,267–269], with the former being a surrogate predictor of the latter [66,270]. However, as with monotherapy, a barrier to successful treatment with highly active antiretroviral therapy (HAART) was development of drug resistance.

Testing for HIV drug resistance can be used to identify alternative therapeutic options for patients failing their antiretroviral therapy regimen. Both cell-based phenotypic and sequencing-based genotypic assays can be used to test for HIV resistance [177,271,272], and the utility of genotypic resistance testing has been demonstrated in randomized clinical trials [176,181,273]. There are currently six classes of antiretroviral medications in clinical usage: the NRTIs, NNRTIs, PIs, fusion inhibitors, co-receptor antagonists, and integrase inhibitors (Figure 1.1) [67]. For each drug class there are specific drug-associated resistance mutations which can be selected for by suboptimal antiretroviral treatment [274]. Naturally occurring polymorphisms and envelope variation are also associated with decreased susceptibility to certain antiretrovirals [275,276]. Screening for antiretroviral drug resistance is

recommended by most treatment guidelines prior to beginning therapy and again if therapy failure occurs [67,277–280]. Screening for HIV tropism/coreceptor usage is also recommended prior to administration of the CCR5 antagonist maraviroc [67].

1.3.2 Coreceptor Usage & Antiretroviral Therapy

Patients receiving antiretroviral therapy are more likely to harbour X4 variants than those who are treatment-naïve [281]. This observation is likely driven by $CD4^+$ T-cell count, since treatment-experienced patients often have lower current or nadir $CD4^+$ cell counts, both of which are associated with viral CXCR4 usage [155,281]. Patients with X4 HIV may be less likely or may take longer to achieve virologic suppression on antiretroviral therapy than those with R5 HIV [173,282]. Once suppressed by therapy, the HIV population tends to be slower to evolve or change coreceptor usage [282,283]. Additionally, certain antiretrovirals have increased or decreased activity against HIV depending on the viral tropism [145].

The effect of HIV tropism on antiretroviral susceptibility is most significant for antiretrovirals that directly target the coreceptors [108,110,111,116,284,285]. There are several small-molecule antagonists of the coreceptors that have been tested in HIV applications [109,110,285,286]. Currently, the only coreceptor antagonist approved is the CCR5 antagonist, maraviroc. This is a small-molecule agent which binds to the CCR5 protein and changes its conformation, thereby reducing the ability of HIV to infect cells. Maraviroc has been shown to have good tolerability and high efficacy against HIV [108,287–290]. However, the drug has suboptimal activity against strains which are capable of using CXCR4. Therefore a tropism test must be performed prior to treatment [67,291]. Undetected X4 variants present prior to treatment with maraviroc can compromise its antiretroviral activity. This is because maraviroc therapy selects for those variants, potentially leading to treatment failure [232,238,290,292]. These non-R5 viruses often exist as a minority of the total viral population. Therefore sensitive detection of X4 variants is likely to be a key factor in therapy success with maraviroc and other CCR5 antagonists. This principle can also be applied to early detection of antiretroviral resistance.

1.3.3 Relevance of Minority Variants for Antiretroviral Resistance & Viral Coreceptor Usage

Antiretroviral efficacy is compromised by the presence of HIV with resistance or reduced susceptibility to components of the treatment regimen [234,258]. Due to the error-prone nature of its reverse transcriptase enzyme, HIV exists as a swarm of quasispecies which differ in their genetic makeup [293].This fact, combined with the high rate of HIV replication, can lead to a diverse set of variants, any of which may carry a resistance associated mutation simply by unfortunate coincidence. Resistance to antiretroviral agents is generally associated with genetic changes within the actual target of the drug. Since these targets are usually all essential for efficient HIV replication, mutations within the genes for these targets can be associated with reduced viral fitness in the absence of treatment [294]. The reduced fitness of resistant variants limits their population size within an individual, which further underscores the potential importance for minority species detection. In addition to natural accumulation of potential resistance mutations, minority drug resistant variants may be transmitted from one individual to another (though this has been disputed [295]) or they may be transmitted as the dominant variants but decline to minority levels in their new untreated host [296,297].

Several studies have demonstrated that minority drug resistant variants exist in antiretroviral-naïve patients [296–300], though other investigators have postulated that such variants are only spuriously detected [295]. Some minority mutations such as the NRTI resistance mutation K65R may be naturally more common in the context of certain genetic backgrounds (e.g., subtype C) due to there being only one nucleotide change between the wildtype and the mutant codons [301]. Indeed, sensitive minority species detection has confirmed that such variants are more common in subtype C viruses than other subtypes [302], which raises the possibility of a higher propensity to develop full resistance to many NRTIs. In addition to treatment-naïve patients, those with treatment experience are also at risk for minority drug resistance both on and off therapy [303–308].

Upon exposure to antiretroviral medications, resistant variants gain a selective advantage over susceptible quasispecies and expand in population size in a classical darwinian selection process [183]. The selection of drug resistant variants appears to be very rapid. Even after a single dose of

antiretroviral medication (such as during administration of single-dose nevirapine to prevent mother-to-child-transmission) there is sufficient selection pressure to increase the population size of resistant variants [309] and even lead to eventual therapy failure if later treated with nevirapine or other NNRTIs [310]. Importantly, the presence of minority variants before drug exposure has been associated with emergence of resistance and treatment failure in several independent studies [233,296,298,303,304,307]. Though some have not found such an association [297,300].

CXCR4-usage by HIV is similar to resistance in that it compromises treatment with CCR5 antagonists. However, HIV tropism also contrasts with resistance because there is a constant selective pressure that is present within the host immune and lymphatic systems. The ability of an individual virion to use both CCR5 and CXCR4 (i.e., dual tropism) could be seen to have a theoretical selective advantage due to an ability to infect a wider range of cells. However, there are conflicting arguments as to whether this may be the case [145,149,222,311,312], especially since CXCR4-usage is associated with a decline in target $CD4^+$ T cells [169,172,173]. Nevertheless, CXCR4-usage does evolve in a proportion of HIV-infected patients over time. However, perhaps as a result of this balance between advantage and disadvantage, this happens in a minority of patients, and X4 variants tend to exist as minorities within a majority CCR5-using viral population [153,281,313–315]. Again, as with minority resistant variants, minority X4 subpopulations can lead to antiretroviral treatment failure [232,316,317].

Consequently, there is a strong rationale to apply deep sequencing in order to detect minority X4 variants. Additionally, the methodologies which are applied to sequencing the HIV *envelope* gene can be readily extended to other parts of the virus, potentially enabling detection of other minority variants with reduced susceptibility to antiretrovirals. This thesis can therefore be considered as a specific application of a very flexible, multi-application tool which has implications that extend far beyond viral coreceptor usage.

1.4 Thesis Overview

1.4.1 Thesis Organization & Objectives

This thesis is organized into seven chapters. Chapter 1 introduces the HIV epidemic as a whole. It contains general information about the viral composition, life cycle, target cells, pathogenesis, and treatment. There is also an overview of HIV coreceptor usage, tropism assays, next-generation sequencing, and the clinical relevance of these topics. Chapters 2 through 6 address the primary objectives of the thesis, as detailed below. Chapter 7 summarizes the results of the research, discusses these results, and comments on their implications. Figures and tables are found after their first mention in the text. References for all chapters are presented in a Bibliography at the end of the thesis. The general hypothesis of these studies is that deep sequencing is equivalent or superior to alternative tropism assays and that it can be used to better assess HIV coreceptor usage and its implication in treatment with maraviroc.

The general aim of this thesis is to establish the clinical relevance of next-generation sequencing in HIV applications, with an emphasis on those relating to HIV tropism and treatment with CCR5 antagonists.

Chapter 2 introduces the methodology and performance of next-generation sequencing, while also comparing various other tropism assays. Chapters 3 through 5 all focus on three trials of treatment-experienced patients. Each chapter focuses on a different aspect of these trials. Chapter 3 applies deep sequencing to HIV RNA from plasma. Chapter 4 investigates the utility of next-generation sequencing in assessing the cellular compartment in HIV-infected individuals. Chapter 5 comprises a detailed longitudinal analysis of HIV envelope sequences prior to and after drug selection pressure by maraviroc. Finally, Chapter 6 applies deep sequencing to an entirely different treatment population of antiretroviral-naïve patients. The thesis thus addresses a number of aspects of HIV infection and treatment which can be assessed and predicted with next-generation sequencing.

1.4.2 Overview of Data Sources

The primary data sources for this thesis have been four large clinical trials of maraviroc conducted internationally. Two trials, Maraviroc versus Optimized Therapy in Viraemia Antiretroviral Treatment-Experienced patients (MOTIVATE) 1 and MOTIVATE 2, investigated maraviroc or placebo, both with optimized background regimens in therapy-experienced patients with CCR5-using HIV-1 as determined by the phenotypic Trofile assay. A separate but related trial, A4001029 had an identical study design to the other two, but enrolled patients with phenotypically assessed non-R5 HIV-1 infections. The fourth trial was conducted in treatment-naïve patients who were randomized to receive the NRTIs zidovudine and lamivudine, plus either maraviroc or the NNRTI, efavirenz.

The samples examined in this thesis were those drawn from patients when they were screened for eligibility in the trials, or were drawn on the first day of treatment (baseline). Additional samples were also obtained from later time-points from three of these trials for patients who were not responding optimally to therapy. The sample types included were either blood plasma or peripheral blood mononuclear cells from patients participating in these trials. Additional samples were also obtained from a subset of the HAART Observational Medical Evaluation and Research (HOMER) cohort, which is based in British Columbia.

Ethical approval for all of the studies presented in this thesis was granted by the Providence Health Care/University of British Columbia Research Ethics Board. Versions of Chapters 2 through 6 have all been published in several international, peer-reviewed journals: *the Journal of Acquired Immune Deficiency Syndromes, the Journal of Infectious Diseases, Clinical Infectious Diseases,* and *Antimicrobial Agents and Chemotherapy.* The candidate is the lead author on all of these manuscripts, and is lead author on two additional review articles, which are included in part in Chapter 1, and were published in *HIV Therapy* and *Current Opinion in HIV and AIDS*.

Chapter 2: Improved Detection of CXCR4-Using HIV by V3 Genotyping: Application of Population-Based & Deep Sequencing to Plasma HIV RNA & Proviral HIV DNA

2.1 Background & Introduction

HIV gains entry into a cell through the use of its envelope protein, gp120. During entry, it binds to the human CD4 receptor and a coreceptor — either CXCR4 (X4 HIV) or CCR5 (R5 HIV) [100]. The coreceptor used by a virus to gain cellular entry is referred to as its tropism or coreceptor phenotype (e.g., X4, R5, or dual-tropic). More advanced disease progression is often associated with CXCR4 tropism and detectable X4 viral load [100]. Furthermore, with the emergence of the CCR5 antagonist antiretroviral drug class (e.g., maraviroc [108]), coreceptor usage has become more clinically relevant, since the efficacy of these CCR5 antagonists is dependent on the CCR5 coreceptor being almost exclusively used by the patient's virus [291,316]. Many tests are available to screen for tropism, each with its own advantages and disadvantages [175,191]. One of the most commonly used tropism tests is the Trofile coreceptor assay (Monogram Biosciences) [194] and the Enhanced Sensitivity Trofile assay (ESTA) [195,318].

Genotypic screening methods for determining coreceptor usage have the potential to be faster, less expensive and more easily standardized than current phenotypic methods [197]. This approach is possible because viral tropism is reflected in the genetic sequence of the gp160 protein, with its third variable domain, or V3 loop, being particularly predictive of coreceptor usage [319]. Bioinformatic algorithms use V3 sequence data to predict coreceptor phenotype [197,203,204]. By analogy, drug resistance testing is routinely performed by inferring a phenotype (in this case, degree of virologic response to a drug) from genotypic sequence data. However, standard, population-based V3 sequencing may lack sensitivity for minority X4 HIV [229].

In a clinical setting, it may be useful to detect the presence of minority species at low concentrations, such as low levels of X4 virus that have the potential to emerge following treatment with a CCR5 antagonist. Although previous efforts to determine viral tropism through more traditional genotypic methods have appeared inadequate [229], generating sequences through independent triplicate PCR amplification and/or by deep sequencing may improve upon the sensitivity of these results. Triplicate PCR amplification increases the probability that X4 minority species will be amplified from a given sample extract. Deep sequencing can be performed on a Roche/454 Life Sciences Genome Sequencer FLX System and generates data for many individual variants within a given sample, including X4 variants. Deep sequencing allows amplification and increased detection of rare subpopulations, thereby increasing the threshold for detection of low-level viral populations.

An additional drawback of current tropism assays is that it is not currently possible to screen patients who wish to switch to a CCR5 antagonist for reasons of tolerability or otherwise, but who currently have viral loads below 1000 copies/ml or, indeed, undetectable viral loads. This stems from a limitation of assays based on plasma HIV, such as the Trofile assay, which requires plasma viral loads greater than 1000 HIV RNA copies/mL. As an alternative, proviral DNA can be amplified from peripheral blood mononuclear cells (PBMCs) and the coreceptor usage of these species can be inferred using methods similar to those performed on plasma HIV RNA. Interestingly, such methods may be performed even when pVL is undetectable, so it may be advantageous to use genotypic prediction methods to determine viral tropism from proviral DNA in patients with low or undetectable plasma viral loads.

The aims of the current study were to improve detection of X4 HIV using a number of genotypic tropism methods, and to compare them to phenotypic Trofile assay results. Standard, population-based sequencing and deep sequencing were performed on triplicate amplifications of the HIV V3 region. Amplifications were made from both viral RNA in plasma and from integrated proviral DNA in peripheral blood mononuclear cells (PBMC), where plasma viral load was undetectable. Tropism was inferred using bioinformatic algorithms, and the results from these various genotypic methods

were compared to those of the Trofile assay. This approach was then validated in an independent dataset of screening samples from the MOTIVATE 1 and MOTIVATE 2 trials of maraviroc [288,289,291].

2.2 Materials & Methods

2.2.1 Cohort Description & Patients

V3 loop sequence variation was assessed in samples from a cohort of antiretroviral-naïve, chronically infected individuals initiating antiretroviral therapy. The primary study group represents a subset (N=63 patients) of the well-characterized HAART Observational Medical Evaluation and Research ("HOMER") cohort [320]. Individuals were included in the present study by convenience, based on the availability of a peripheral blood sample for PCR amplification and a documented Trofile assay result [155]. Ethical approval was granted by the Providence Health Care/University of British Columbia Research Ethics Board.

2.2.2 Extraction & Population-Based Sequencing

HIV RNA was extracted from previously frozen plasma samples, and HIV DNA was extracted from buffy coat samples, both using a NucliSENS easyMAG (bioMerieux). Both RNA sequencing methods (i.e., population-based, and deep sequencing) followed the same procedures up to and including first round PCR, but differed in later steps, such as using different second round PCR primers. The region encoding the HIV V3 loop was amplified independently in triplicate by nested RT-PCRs set up simultaneously from extracts using a multichannel pipette — the additional effort is minimal compared to a single PCR.

Triplicates were chosen mostly arbitrarily as a compromise between potential increased probability for amplification of X4 HIV and a procedure that is clinically feasible for a technician to perform on a single PCR plate. Sequencing was performed in the 5' and 3' directions on an ABI 3730 automated

sequencer as previously described [173]. All primers and thermal cycler protocols for all methods are available in Appendices I and II.

2.2.3 Deep Sequencing & Emulsion PCR Methods

Deep sequencing on the Roche/454 Life Sciences Genome Sequencer FLX (GS-FLX) is a sensitive sequencing technique able to detect low-level subpopulations of virus and generate thousands of sequences from a given sample [232,248,321]. Second round PCR primers were designed with fusion primers to fuse to the emulsion PCR beads required by the pyrosequencing technique. Also included were 12 unique multiplex "barcode" sequence tags to enable the identification of samples after the sequencing was complete.

After PCR amplification, the concentrations of the PCR products were quantified using a Quant-iT Picogreen dsDNA Assay Kit (Invitrogen) and a DTX 880 Multimode Detector (Beckman Coulter). After quantitation, they were combined in equal proportions (2×10^{12} DNA molecules from each triplicate sample), purified with Agencourt Ampure PCR Purification beads (Beckman Coulter), and subsequently re-quantified. Note that only samples from which all three triplicates were successfully amplified were combined. Otherwise, RT-PCR amplification was re-performed until three triplicates were available. This increased the probability that CXCR4-using minority species would be amplified and sequenced from a given sample's RNA extract.

Following its quantification, the combined, purified amplicon "library" was then diluted to a concentration of 2×10^5 molecules per millilitre, and combined at a ratio of 0.6 molecules to 1 DNA capture microbead used for emulsion PCR (emPCR). This ratio of less than 1:1 increased the likelihood that a single microbead would bind a single DNA amplicon, such that subsequent amplification by emulsion PCR would generate homogeneous clones on each bead.

Two separate emPCR amplifications were performed: one for the forward sequencing direction, and one for the reverse. Along with the DNA amplicons, the microbeads are mixed with amplification

buffer, primer and enzyme, as well as with oil. These components are shaken with a TissueLyser (Qiagen/Retsch) to allow formation of oil microreactor micelles around the beads. This process gives mostly independent reaction sites to allow clonal amplification of the DNA amplicon, thus increasing the signal for subsequent pyrosequencing.

The emulsions are then broken and the beads are washed with isopropanol, followed by two 454 washing reagents. The microbeads are then enriched with magnetic beads such that only beads coated in DNA are carried forward into subsequent steps. After enrichment, an annealing step is performed to anneal the pyrosequencing primers onto the bead-bound DNA amplicons. The DNA beads are then added onto a picotitre plate (divided into 4 regions) at a density of 2.5×10^5 beads per region, as quantified with a Z1 Coulter Particle Counter (Beckman Coulter). Control beads (to generate quality scores for the sequencing run), packing beads (to hold the DNA beads in place in the picotitre plate) and enzyme beads are all added over the plate as well. After this point, the plate is prepared for deep sequencing on the GS-FLX.

The sequence amplified on each bead was determined by pyrosequencing on the GS-FLX. This process generated ~200 base pairs of data in each direction per amplicon, with a typical V3 loop consisting of 105 base pairs (35 amino acids). Truncated reads (defined as sequences missing ≥4 bases at the 5′ or 3′ end) were not included in the analysis. In total, 12 HOMER plasma samples underwent deep sequencing with the GS-FLX.

Proviral HIV DNA V3 sequences were assessed in a similar manner in 26 HOMER subjects with non-R5 HIV and 14 with R5 HIV. These subjects were receiving highly active antiretroviral therapy (HAART) [155] and had undetectable plasma viral loads. Sample material consisted of PBMCs from the buffy coat fraction of centrifuged whole blood. Patient phenotypic tropism had previously been determined using Trofile prior to initiating HAART. Nested PCR was performed for bulk sequencing on an ABI 3730 sequencer. Deep sequencing used the same first round primers as bulk sequencing, with different second round primers. Proviral DNA from a total of 12 buffy coat samples from the HOMER cohort underwent deep sequencing on the GS-FLX.

2.2.4 Sequence Analysis & Coreceptor Usage Determination by Bioinformatic Algorithms

After sequencing on the ABI 3730, data were analyzed using the custom software, RECall [196] with no manual intervention. RECall has been shown to have ~99% concordance with human calls [196]. Nucleotide mixtures were automatically called if the secondary peak height exceeded 12.5% of the dominant peak height. Sequences were aligned to HIV-1 subtype B reference strain HXB2 (Genbank Acc. No. K03455) using a modified NAP algorithm [322].

HIV tropism was predicted from V3 genotype using position specific scoring matrices ($PSSM_{X4/R5}$) [203] and/or $geno2pheno_{[coreceptor]}$ [197] scoring. Non-genotypic factors such as $CD4^+$ cell counts were not included in the bioinformatic analysis. Results were compared to the Trofile data as a reference. Standard sequencing replicates with PSSM values below the predetermined cut-off of -6.96 were called R5, while those with scores greater than or equal to -6.96 were called X4 [203]. The geno2pheno method [197,204] used a 5% false-positive rate, with samples also categorized as R5 or X4.

Where sequence ambiguity occurred due to the presence of nucleotide mixtures, the permutation with the highest PSSM score was used to assign the score for a given replicate. In other words, the "most" X4 residues were retained from a sequence where multiple variants contributed to the consensus population sequence. This method was used in order to increase the sensitivity for detection of X4 variants [229]. A similar system was used for geno2pheno. Where triplicate data differed, the most highly X4 (e.g., maximum PSSM score) replicate was assigned to a sample. Thus, R5 samples had all 3 replicates inferred as R5, and X4 samples had one or more replicates inferred as X4. For deep sequencing, each V3 variant detected received a tropism classification using the same two algorithms. This allowed the proportion of X4 virus within the sample to be determined, and samples were classified according to this parameter. Sensitivity was defined as the prediction of CXCR4-usage which correlated with the original Trofile assay.

2.2.5 Independent Validation

The performance of the triplicate method was also assessed in a blinded independent dataset (N=278) of screening samples from the Maraviroc versus Optimized Therapy in Viremic Antiretroviral Treatment-Experienced Patients (MOTIVATE) studies [288,289]. All patients in MOTIVATE consented to other tropism testing being performed on their samples. A sister trial, the A4001029 study had a patient criterion of non-R5 (X4, Dual/Mixed, non-reportable) virus. Samples from this trial were excluded since these samples could not be blinded.

The MOTIVATE screening samples were amplified in triplicate and bulk-sequenced on an ABI3730. As with the HOMER subset, MOTIVATE samples showing evidence of CXCR4-usage in their standard sequencing results were classified as X4. Subjects with R5 (by Trofile) at both screening and baseline (just prior to treatment with study medication) were classified as "confirmed R5s". A subset of 11 MOTIVATE samples were then amplified independently in triplicate as above and underwent deep sequencing on a GS-FLX.

2.3 Results

2.3.1 Standard Sequencing of V3 to Infer Tropism

Standard, population-based sequencing of triplicate amplifications of the V3 loop, in combination with PSSM tropism inference, gave approximately 81% concordance with Trofile. Of samples called R5 by Trofile, 31 of 34 (91%) were also identified as R5 by standard sequencing. Of those called X4 by Trofile, 20 of 29 (69%) were inferred as X4 (Figure 2.1).

Often, there was notable variation among the three independent amplifications performed for each sample. Indeed, 12/63 samples (19%) had at least one replicate indicate a different tropism than the others.

Figure 2.1: PSSM Scores from Standard Population-based Sequencing of Independent Triplicate PCRs of V3 Amplified from Plasma HIV RNA

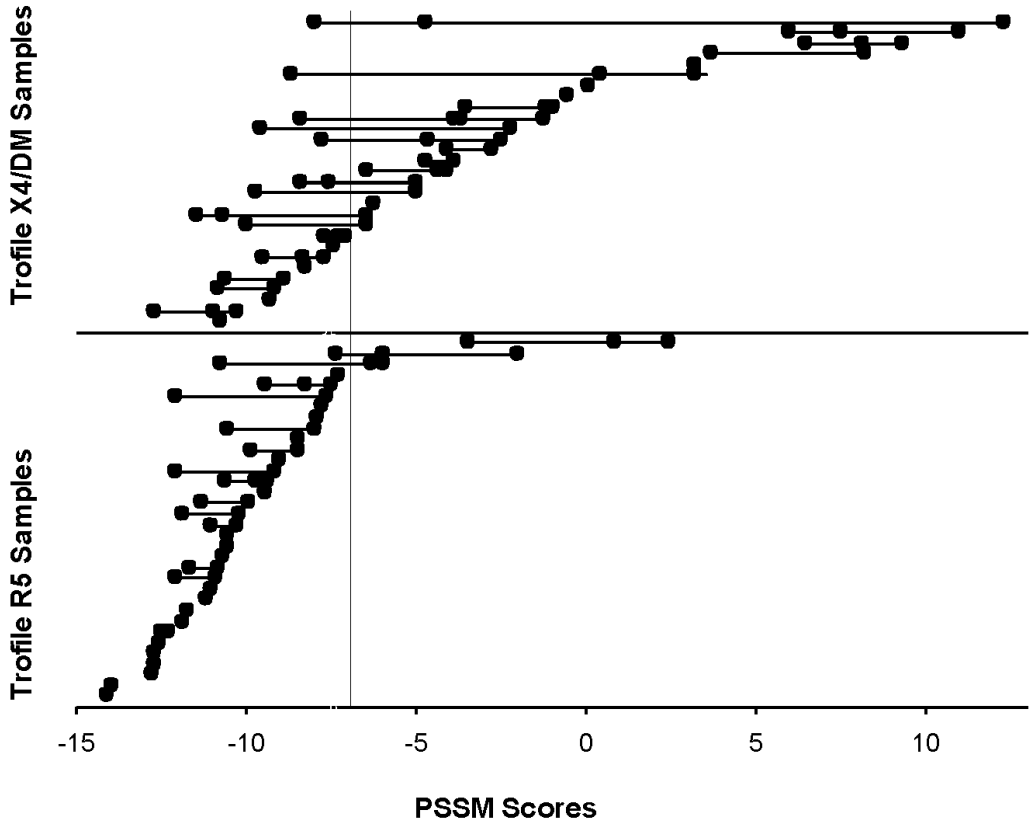


Figure 2.1: PSSM Scores from Standard Population-based Sequencing of Independent Triplicate PCRs of V3 Amplified from Plasma HIV RNA. The horizontal axis represents the possible PSSM scores, with scores to the left of the vertical line (-6.96) indicating R5 virus, and scores to the right of -6.96 indicating X4 virus. Samples are arranged by their Trofile screening result, with the Trofile X4/DM samples in the upper region of the figure, and the Trofile R5 samples in the lower region. Closed circles indicate the PSSM score of each of three replicate amplifications of the V3 loop for different samples. Horizontal lines span the range of the three scores in order to give an indication of the diversity within a respective sample. Where three circles are not visible, this is either because of a failed amplification or because the points overlap because of similar or identical PSSM scores.

Almost half of Trofile X4 or Dual/Mixed samples that were called X4 by standard sequencing had at least one amplification that would have been classified as R5 if the triplicate approach had not been used (9/20 samples, 45%). It is currently unknown as to what effect replicate testing of clinical samples would have had on Trofile assay results.

Receiver operator characteristic (ROC) curves were plotted using either the maximum of three triplicates, or a "singleton" approach using only the first of the triplicates. The area under the curve (AUC) of the ROC curve for the triplicate approach was 0.874 versus 0.828 for the singleton approach, indicating improved performance of the triplicate method over a single amplification. Overall, comparing the result (R5 or X4) by standard sequencing with PSSM, and using the Trofile result as a reference, the sensitivity for the HOMER samples was 69% and specificity was 91%. In comparison, keeping specificity constant, singleton approach would have given 48% sensitivity, and a duplicate approach, 59% sensitivity, relative to Trofile.

2.3.2 Proviral DNA to Infer Tropism

PBMC samples were retrieved from patients who were currently on HAART (without CCR5 antagonist medication), had undetectable plasma viral loads at the time of sampling, and for whom a pre-therapy plasma sample and Trofile assay result were available. Of 46 samples initially attempted, 40 samples yielded successful amplifications, giving an 87% amplification rate. Proviral DNA samples were amplified in triplicate, bulk sequenced on an ABI 3730, and the sequences were inferred as R5 or X4 by PSSM (Figure 2.2).

For samples called non-R5 by Trofile from plasma RNA, 20/26 (77%) had evidence of X4 HIV DNA in their peripheral blood mononuclear cells. A total of 10/14 samples (71%) called R5 by Trofile had R5 sequences in their corresponding proviral DNA, giving sensitivities and specificities of 77% and 71%, respectively, for PSSM; or 77% and 93%, respectively, for geno2pheno (data not shown). The mean proviral DNA PSSM score of each sample was also correlated to the mean pre-treatment RNA PSSM scores ($r^2 = 0.35$).

Figure 2.2: PSSM Scores from HIV Proviral DNA by Standard Population-Based Sequencing

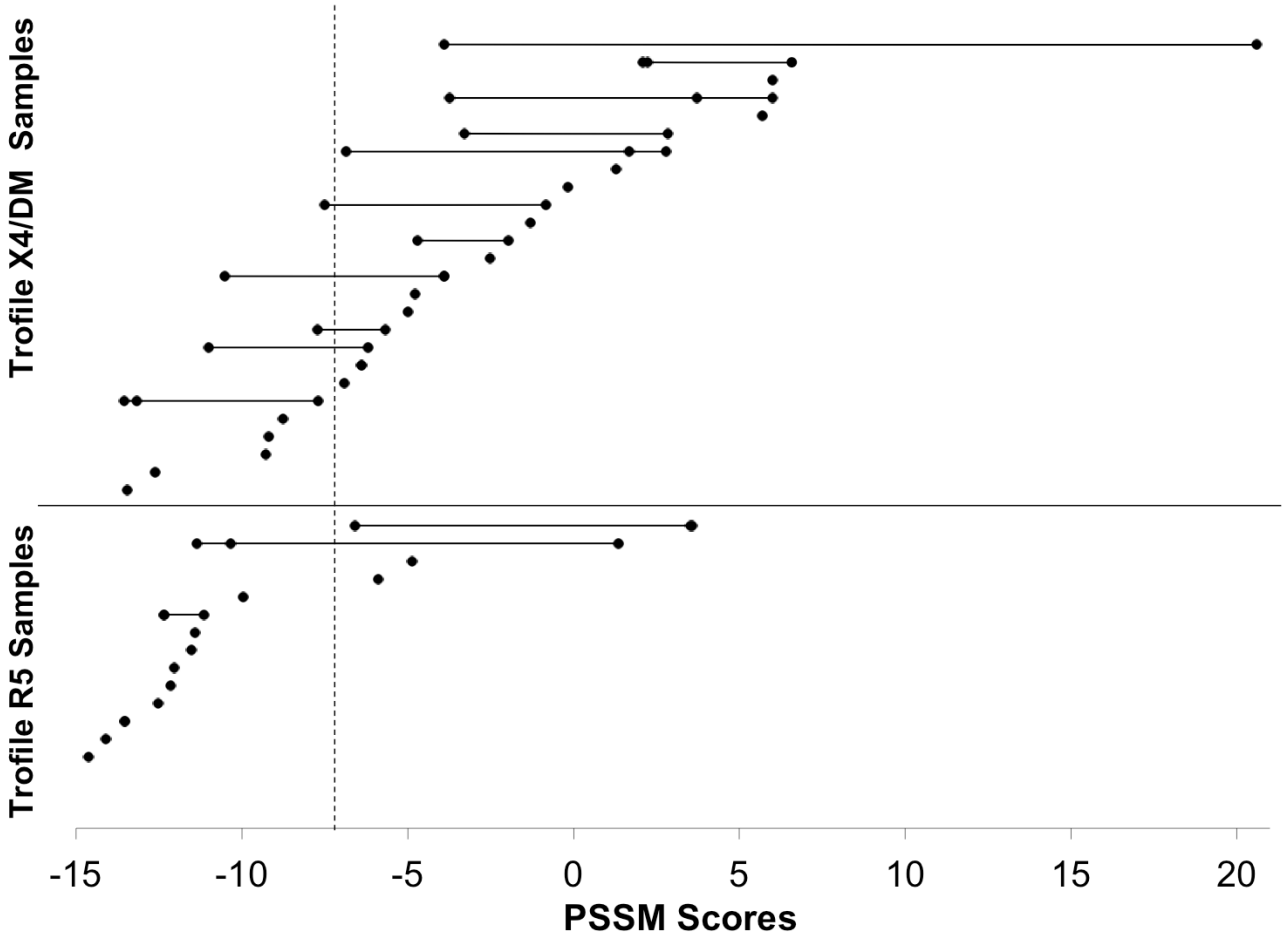


Figure 2.2: PSSM Scores from HIV Proviral DNA by Standard Population-Based Sequencing. Possible PSSM scores on the X-axis. Scores left of the dashed vertical line (-6.96) indicate R5 virus; scores to the right indicate X4 virus. Samples are arranged by their Trofile status. Note that the Trofile result is based on plasma RNA and *not* proviral DNA. Closed circles indicate the PSSM score of each V3 amplification for different samples. Horizontal lines span the range of the three scores for each sample. Where three circles are not visible, this is either because of a failed amplification or because the circles overlap due to similar or identical PSSM scores.

2.3.3 Deep Sequencing of HIV RNA & DNA

A subset of patients with matching plasma RNA and proviral DNA (N=12) samples were sequenced using the GS-FLX. The RNA was extracted from plasma samples drawn prior to initiation of antiretroviral therapy, while the DNA was extracted from buffy coat samples drawn after patients achieved undetectable viraemia, after a median of 36.5 months (IQR: 30.5–39) on antiretroviral therapy. These samples were assessed according to the percentage of X4 virus comprising their deep sequencing results. A total of 4 of the 12 patients (33%) had very similar proportions of X4 virus in their plasma RNA and proviral DNA, (within ~1% of each other). For the remaining 8 samples, the percent X4 in RNA and proviral DNA differed by a range of 6-72%, with proviral DNA tending to harbour a higher percentage of X4 variants (median 46% X4 in DNA versus 8% in RNA). Table 2.1 shows the comparison of these 12 patients across the various tropism methods.

Table 2.1: Deep Sequencing of HIV RNA & DNA Compared with Standard Population-Based Sequencing & the Trofile Assay

Sample	Trofile Result	Standard Sequencing		Deep Sequencing	
		Plasma RNA	Proviral DNA	Plasma RNA (%X4)	Proviral DNA (%X4)
1	R5	R5	R5	0.05	0.6
2	R5	R5	R5	0.4	0.4
3	R5	**X4**	R5	0.4	0.1
4	R5	R5	**X4**	1.9	**74.3**
5	**D/M**	R5	**X4**	**6.8**	**53.2**
6	**D/M**	**X4**	R5	**7.6**	1.2
7	**D/M**	**X4**	**X4**	**8.7**	**48.4**
8	**D/M**	R5	**X4**	**15.9**	**49.4**
9	**D/M**	**X4**	**X4**	**19.5**	**39.8**
10	**D/M**	R5	**X4**	**71.0**	**99.4**
11	**D/M**	**X4**	**X4**	**84.5**	**43.1**
12	**D/M**	**X4**	**X4**	**99.8**	**99.6**

Table 2.1: Deep Sequencing of HIV RNA & DNA Compared with Standard Population-Based Sequencing & the Trofile Assay. Deep sequencing was performed on 12 plasma (RNA) samples and 12 PBMC (DNA) samples, and the percent X4 in each sample was compared to the corresponding Trofile result and maximum PSSM score by standard sequencing. Standard sequencing classified samples as X4 if the maximum PSSM score was ≥-6.96. All non-R5 results are bolded. For deep sequencing, samples are bolded if >2% X4 virus was detected by the GS-FLX. D/M — Dual/Mixed Tropic

Overall, the deep sequencing percent X4 from pre-treatment plasma RNA, and post-suppression proviral DNA were well correlated (r^2=0.44), and also corresponded very well to the pre-treatment Trofile results. Using RNA and DNA, respectively, 4/4 and 3/4 samples called R5 by Trofile had less than 2% X4 virus within their deep sequencing results, while 8/8 and 7/8 samples called non-R5 by Trofile had greater than 2% X4 virus comprising their deep sequencing results. Standard sequencing of RNA and DNA gave X4 calls in 11/12 samples (92%) that had 20% or more X4 by deep sequencing, and gave R5 calls in 9 of 12 (75%) samples with less than 20% X4, consistent with the typical sensitivity of standard sequencing in reliably detecting minority species. Indeed, the presence of low-level (below 20%) X4 variants could explain 3/4 (75%) Trofile-non-R5 samples which were apparently misclassified as R5 by standard sequencing.

2.3.4 Results of Independent Validation

The sensitivity and specificity of the current approach were ascertained on a blinded, independent sample set (N=278) from the MOTIVATE trials, which tested maraviroc in treatment-experienced individuals. A previous attempt to determine tropism by standard sequencing (not in triplicate) with PSSM methods had only 24% sensitivity with 97% specificity when compared to Trofile [229]. The independent validation of the current method with bioinformatic analysis using PSSM yielded substantially increased sensitivity (75%) with only a modest decline in specificity (83%). Compared to PSSM, geno2pheno methods yielded a slightly worse sensitivity (61%) but had higher specificity (93%), though there was limited power to distinguish either algorithm as superior.

For "confirmed" R5 samples (i.e., those called R5 by Trofile at time of screening and again at baseline just before starting maraviroc), genotyping had sensitivities and specificities of 71% and 95%, respectively, for geno2pheno and 75% and 82%, respectively, for PSSM. A subset of 11 MOTIVATE samples also underwent deep sequencing. Many of these samples had at least a minority of non-R5 sequences present in their deep sequencing results. There was a wide distribution of PSSM scores present within these samples, representing a viral population that often comprised both CXCR4- and CCR5-using variants . For 10 of 11 samples (91%), the PSSM score of the most common deep

sequencing variant resulted in the same tropism classification as the Trofile assay classification (Figure 2.3).

2.4 Discussion & Conclusions

Clinical samples were assessed using standard, population-based sequencing and deep sequencing of the HIV V3 region, and the results show higher sensitivity for detecting CXCR4-using virus in samples than previously achieved [229]. Of additional significance was the use of proviral DNA to infer viral tropism in treated patients with undetectable plasma viral loads. Deep sequencing also seemed to be a good predictor of Trofile results, with a cut-off of 2% X4 giving good concordance with the original Trofile assay.

Relatively few studies performed prior to this study used proviral DNA [323] or deep sequencing [232,317] to infer tropism. Genotypic tropism testing from proviral DNA suggests the possibility of screening for those with suppressed viraemia who may wish to switch to CCR5 antagonists for reasons such as tolerability, whereas the Trofile assay requires a viral load of at least 1000 copies/mL [195]. With the above outlined approach, most patients harbouring X4 virus can be quickly screened out as being ineligible to receive maraviroc.

The differences in the results of the earlier study versus the current one may be attributable to a number of factors, especially: triplicate amplification, better sequencing technology, and automatic base-calling. The use of independent triplicate amplifications here may be able to amplify a greater proportion of minority species due to the inherently stochastic nature of PCR. This was evidenced in the variability amongst the replicates, with approximately 20% of samples yielding replicates with different inferred coreceptor usage. Also lending evidence to the utility of the triplicate approach was the larger area under the curve for the ROC curve plotting the triplicate approach versus that of a singleton approach (0.874 versus 0.828, respectively). Triplicate PCRs also demonstrated greater overall sensitivity for detecting CXCR4-using variants (69% for the triplicate approach versus 48% for a singleton approach).

Figure 2.3: Distribution of PSSM Scores for Variants Detected by Deep Sequencing of Plasma Samples

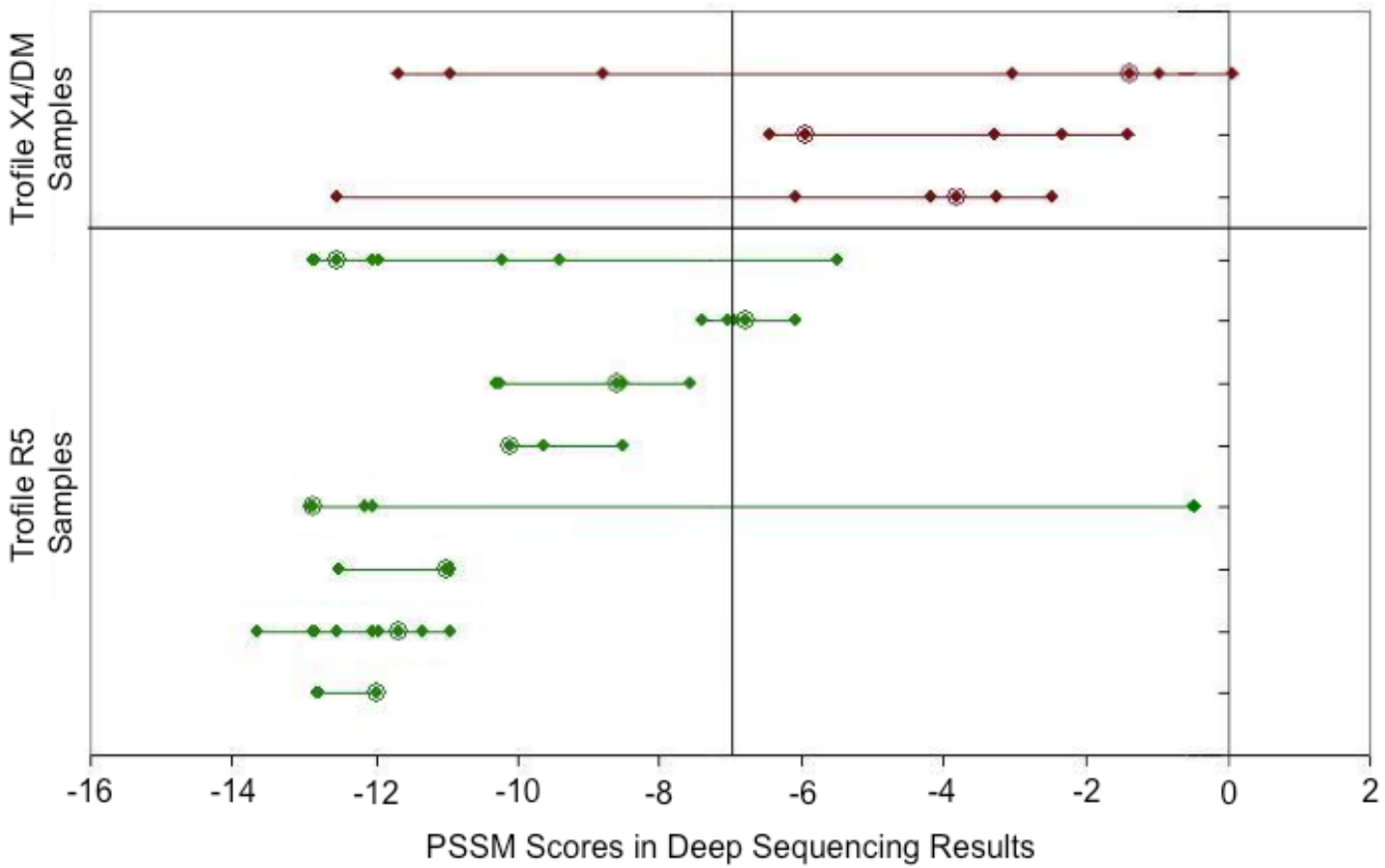


Figure 2.3: Distribution of PSSM Scores for Variants Detected by Deep Sequencing of Plasma Samples There was a large degree of variation in PSSM scores within samples, according to deep sequencing. The large circle indicates the PSSM score of the most common variant detected by deep sequencing for each sample. Diamonds indicate the PSSM scores of the rest of the ten most common variants detected. The horizontal axis represents the possible PSSM scores, with scores to the left of the dashed vertical line (-6.96) indicating R5 variants, and scores to the right of -6.96 indicating X4 variants. Samples are arranged by their Trofile classifications, with X4/DM calls by Trofile on the top and coloured in red, and R5 calls below and coloured in green.

The sequencing technology has improved, with the ABI 3730 used for the current study, and either the ABI 3100 or 3700 used for the earlier study. The Roche/454 Life Sciences Genome Sequencer-FLX also represents a further advance in sequencing technology. An additional advantage in the current method is the automated nature of the sequence analysis. Due to analysis by RECall, sequence data underwent no manual intervention such as base-calling. Use of RECall made this method quick and efficient while bypassing the inherently inconsistent and labour-intensive process of manual sequence analysis by a technician.

Some limitations of the sample population should be noted. The first 63 available samples as organized by sample identifier were arbitrarily chosen, resulting in a potentially unknown selection bias. Furthermore, the clinical test set from British Columbia is composed of 97.5% clade-B virus, thus skewing the results in favour of methods trained primarily on clade-B [203,204]. The PSSM algorithm used here may not be readily extendable to non-clade B sequences. It should also be noted that the original Trofile assay was used for these analyses but not the enhanced sensitivity Trofile assay, ESTA. The current results may have differed if these genotypic methods were compared to ESTA. For instance, some of the Trofile R5 samples may have yielded X4 or Dual/Mixed results if tested by ESTA, which may have decreased specificity.

Interestingly, the deep sequencing genotypic methods were able to detect low levels of inferred-X4 virus in almost all samples that underwent deep sequencing, regardless of their Trofile assay results. Indeed, some samples which had R5 results by Trofile gave entirely different results by deep sequencing, which revealed a majority of X4 variants. Subsequent chapters will explore how such discordant patients (e.g., Trofile R5/deep sequencing X4) perform virologically on CCR5 antagonist therapy. These results do suggest, however, that tropism may not always fall into discrete categories of R5 or X4, and that the proportion of a patient's virus that uses either coreceptor may be variable.

The "true" sensitivity of genotypic tropism testing is confounded by the "gold standard" against which these tests are compared. Numerous studies have compared a variety of genotypic and

phenotypic tests, each yielding varying sensitivities and specificities [187,190,230]. Concordance even between phenotypic tropism assays is not necessarily 100% (e.g., [190]). Furthermore, depending on the tests used, genotypic sensitivity for X4 variants has ranged from as low as 10% [187] for genotypic predictors, such as the 11/25 charge rule [324], to ~70% for support vector machines [190] and geno2pheno [187]. Even the same algorithm used on different datasets can yield vastly different sensitivities [230]. Because of the wide range of sensitivities reported from genotypic testing, the results reported here should be taken in this context.

This study yielded improved detection of X4 HIV in clinical samples and was a better predictor of viral tropism than many previous attempts at genotypic approaches to determining coreceptor usage. These data also suggest that deep sequencing technology and genotypic analysis of proviral DNA may prove useful in the determination of coreceptor usage. With the above outlined approach, most patients harbouring X4 virus can be quickly screened out.

Using the Trofile call as the reference may, however, be problematic. Ultimately, the best indication against which results should be compared is the virologic outcome of patients who receive CCR5 antagonist medication. Clinical outcome, and not other assays, may be the best candidate for the "gold standard" of comparison [325]. Therefore, the next chapter of this thesis tests the ability to predict outcomes on maraviroc in treatment-experienced patients entering three large clinical trials of the CCR5 antagonist. Methods developed above are optimized and refined in the following chapter, with the primary focus being the performance of deep sequencing situated in a clinical context.

Chapter 3: Deep Sequencing to Infer HIV-1 Coreceptor Usage: Application to Three Clinical Trials of Maraviroc in Treatment-Experienced Patients

3.1 Background & Introduction

Human Immunodeficiency Virus Type 1 (HIV-1) enters and infects a target cell by an interaction of its envelope glycoprotein, gp120, with the cellular CD4 receptor and a co-receptor: CCR5 or CXCR4 [100,116,326,327]. CCR5 antagonists such as maraviroc inhibit HIV entry via CCR5. These agents work by allosterically altering the conformation of CCR5 at the cell surface, thereby disrupting its interaction with HIV gp120 [100,108,286]. However, CCR5 antagonists have suboptimal activity against viral populations capable of using CXCR4 [291,316]. Accordingly, before clinical use of CCR5 antagonists, a tropism test is performed to rule out the presence of detectable non-CCR5-tropic (non-R5) virus.

Some of the most widely-used coreceptor tropism tests have been the recombinant, phenotypic Trofile assay (Monogram Biosciences) [194], or its newer iteration, the Enhanced Sensitivity Trofile assay ("ESTA") [195]. Despite their wide use, there are some practical limitations to these assays, including a long turnaround time, restricted geographic access, and large sample volume required [328]. Genotypic tropism testing is an alternative method [329] that is possible because the sequence of the third variable (V3) loop of HIV gp120 is the principal determinant of tropism [23,123,220,324,330,331], allowing tropism inference using bioinformatic algorithms such as $PSSM_{X4/R5}$ [203] and $geno2pheno_{[coreceptor]}$ [197,204].

However, genotypic assays based on standard, population-based V3 sequencing have often had apparently poor sensitivity for detection of non-R5 HIV (e.g., [229]), especially when such species comprise minorities in the viral population below ~20%, the reliable sensitivity of standard sequencing [180,228]. In comparison, next-generation deep sequencing approaches have much higher sensitivity and can detect minority HIV variants at much lower levels [240,332], including minority non-

R5 subpopulations [232]. Consequently, this method can capture a detailed "cross-section" of co-receptor-usage across a patient's viral population, and quantify the prevalence of non-R5 HIV within the patient.

Presented here is an extensive study of deep V3 sequencing as a tool for predicting virologic outcomes on maraviroc-based therapy in treatment-experienced patients in the Maraviroc versus Optimized Therapy in Viremic Antiretroviral Treatment-Experienced Patients (MOTIVATE) 1 and 2 studies. These were randomized, phase 3, placebo-controlled studies of maraviroc in treatment-experienced patients with R5 HIV [288,289]. Patients were originally screened using the original Trofile assay. Of those screened out due to non-R5 HIV, approximately 20% (186/955) entered the A4001029 trial [291]. This trial also assessed maraviroc versus placebo but in patients with non-R5 HIV results. Deep sequencing was retrospectively tested on a total of 1827 blinded screening samples from these three clinical trials, and assessed for its ability to predict virologic responses in maraviroc recipients.

3.2 Materials & Methods

3.2.1 Trial Patients, Samples & Amplification Methods

Briefly, the V3 loop of HIV gp120 was amplified independently in triplicate by nested RT-PCR methods from a total of 1827 screening samples from the three trials. These were then sequenced by either (A) standard, population-based sequencing [325], or (B) deep sequencing [333]. Refer to Chapter 2 for detailed methodologies. The current study focuses on the deep sequencing data, hereafter referred to as genotyping.

In total, 1093 of 1827 patients examined in the current study were randomized into the three arms of the MOTIVATE (R5) and A4001029 (non-R5) trials (Figure 3.1). Informed consent was obtained from all individuals. Treatment arms were maraviroc once-daily (QD), maraviroc twice-daily (BID), or placebo, plus an optimized background therapy of 3 to 6 agents, based on treatment history and resistance testing [288,289].

Figure 3.1: Sample & Patient Distribution

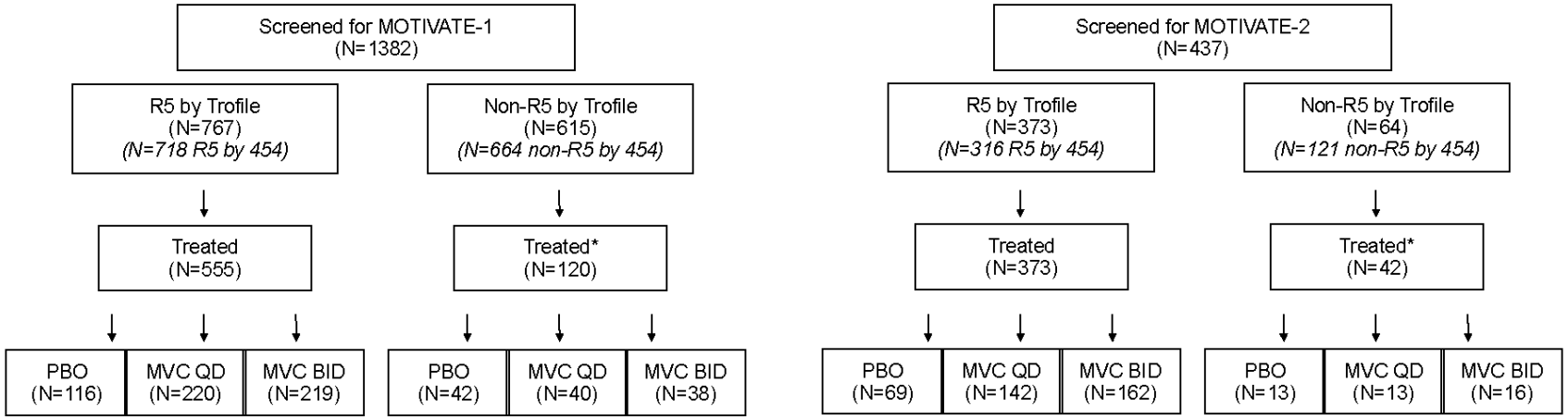


Figure 3.1: Sample & Patient Distribution Of the current study population, patients were screened for entry into either MOTIVATE-1 (76%) or MOTIVATE-2 (24%). In both studies, a majority of patients had R5 HIV at screening by the original Trofile assay, while a minority had non-R5 results. A subset of these patients were enrolled into the trials and received treatment in one of three arms: placebo (PBO), maraviroc once-daily (MVC QD), or maraviroc twice daily (MVC BID). *Patients with non-R5 HIV at screening were treated in the A4001029 study, with 8 patients being screened and enrolled into A4001029 directly.

Note that all phenotypic screening results were performed using the original Trofile assay (approximate 10% non-R5 cutoff [334]) and not ESTA (0.3% cutoff [195]). Figure 3.1 shows the distribution of patient samples tested in the current study. The primary analysis was based on all patients who entered any study (MOTIVATE-1, MOTIVATE-2 or A4001029). Critically, this included all treated patients for whom Trofile gave a non-R5 result. For additional analyses with respect to tropism assessments by both assays, the patients screened for MOTIVATE-1 (including Trofile-non-R5 patients), but who did not enter a study were also included. However, only patients entering the studies could be examined for virologic responses.

HIV RNA was extracted from 500 μL of plasma per sample using a NucliSENS easyMAG (bioMérieux). Three independent one-step RT-PCR amplifications were performed with 4μL of extract per amplification, followed by a second-round amplification using customised primers that included a V3-specific PCR primer and a multiplex barcode (for distinguishing between samples). All primers and are listed in Appendix III.

3.2.2 Emulsion Polymerase Chain Reaction & Pyrosequencing

After PCR amplification, PCR amplicon concentrations were quantified using a Quant-iT Picogreen dsDNA Assay Kit (Invitrogen) and a DTX-880 Multimode Detector (Beckman Coulter). These were combined in equal proportions ($2x10^{12}$ DNA molecules/amplification), purified with Agencourt Ampure PCR Purification beads (Beckman Coulter), and re-quantified. The purified products were then diluted to $2x10^5$ molecules/mL, and combined at a ratio of 0.6 molecules to one emulsion PCR (emPCR) microbead. Oil and emPCR buffer components were shaken with a TissueLyser (Qiagen/Retsch) to allow formation of microreactor micelles around the beads. After emPCR, the beads were washed and enriched for DNA-coated beads as per the manufacturer's instructions. These were added onto a picotitre plate at 2.5×10^5 beads in each of four regions, and underwent sequencing with a Genome Sequencer-FLX (Roche/454 Life Sciences).

3.2.3 Optimizing Bioinformatic Cutoffs for Deep Sequencing

Because of the limited prior experience in the literature with deep sequencing, an attempt was made to first identify an optimal cutoff for both the geno2pheno and PSSM bioinformatic algorithms [335–337]. Using a large dataset of 1875 samples (1827 for the PSSM analyses), Receiver Operating Characteristic (ROC) curves [338] were generated to compare the performance of deep sequencing relative to the original Trofile assay. A variety of FPR cutoffs were tested, ranging from 1.0 to 6.5 for geno2pheno and a variety of scores ranging from -1.0 to -6.5 for PSSM.

The area under the curve (AUC) of the ROC curves was also calculated for each cutoff and plotted as shown in Figure 3.2. Cutoffs which optimized performance relative to the original Trofile assay were selected. For geno2pheno, this was a cutoff of 3.5, whereby sequences with FPRs above 3.5 were classified as R5 (Figure 3.3: Panel 1). PSSM had an optimized cutoff of -4.75, whereby sequences with scores less than -4.75 were classified as X4 (Figure 3.3: Panel 2). These particular cutoffs for genp2pheno and PSSM had maximal AUCs of 0.8911 (95% confidence interval: 0.8755 – 0.9067) and 0.8915 (95% confidence interval: 0.8765 – 0.9065), respectively.

Additionally, the cutoff for the percentage of non-R5 variants was explored in order to optimally predict virologic outcomes. A random 75% of the dataset was used for training and exploration of cutoffs, and once these were established, they were tested in a validation dataset comprising the remaining 25%. For this approach, virologic success was defined as a ≥2 $\log_{10}$ decline in viral load from baseline and/or a viral load <50 copies/mL at week 8. Figure 3.4 shows the results of these analyses. A cutoff of 2% non-R5 variants was able to correctly predict virologic success with sensitivities of 84% for geno2pheno and 81% for PSSM. This cutoff was also found to give the best performance in Chapter 2.

3.2.4 Bioinformatic Algorithms for Inferring Tropism from Genotypic Data

Deep sequencing generated read-lengths of approximately 250 base pairs of data in each direction. A typical V3 loop was 105 base pairs long (35 amino acids). Truncated reads (missing 4 or more bases at either end of V3) were excluded from the analysis, as were samples producing fewer than 750 usable reads. Genotyping generated a mean of over 3000 V3 sequences per sample. The tropism of each sequence was interpreted by the $PSSM_{X4/R5}$ or geno2pheno bioinformatic algorithms (optimized cut-offs for non-R5: $PSSM_{X4/R5}$, ≥-4.75; geno2pheno, ≤3.5). The selection of these cutoffs is detailed in Figures 3.2 – 3.4. The overall sample tropism was expressed as the proportion of non-R5 sequences within the sample's viral population. Patients with samples harbouring ≥2% non-R5 variants were classified as having non-R5 HIV, while those with <2% were classified as having R5 HIV.

Figure 3.2: Areas under Receiver Operating Characteristic Curves for Various Bioinformatic Cutoffs

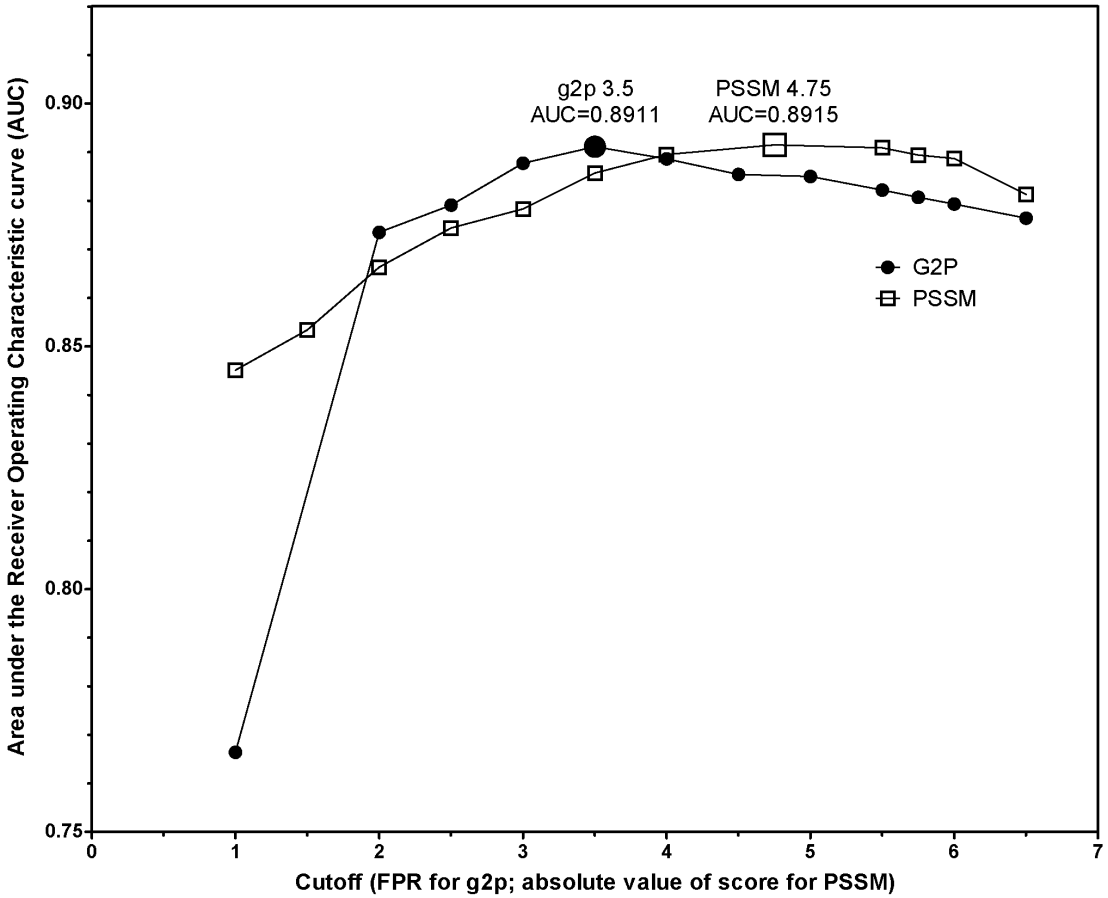


Figure 3.2: Areas under Receiver Operating Characteristic Curves for Various Bioinformatic Cutoffs. The areas under the receiver operating characteristic curves (AUC) were calculated for various bioinformatic cutoffs, and their values are plotted here. The cutoffs which maximized the AUC were a geno2pheno false-positive rate cutoff of 3.5, and a PSSM score cutoff of -4.75. For display purposes in the figure, the PSSM scores have been changed to their absolute values rather than negative values. Maximal AUCs indicated that the highest numbers of samples were classified "correctly" using the original Trofile assay as a comparator.

Figure 3.3: Receiver Operating Characteristic Curves for Optimal Geno2pheno & PSSM Cutoffs — Prediction of Original Trofile Assay Results

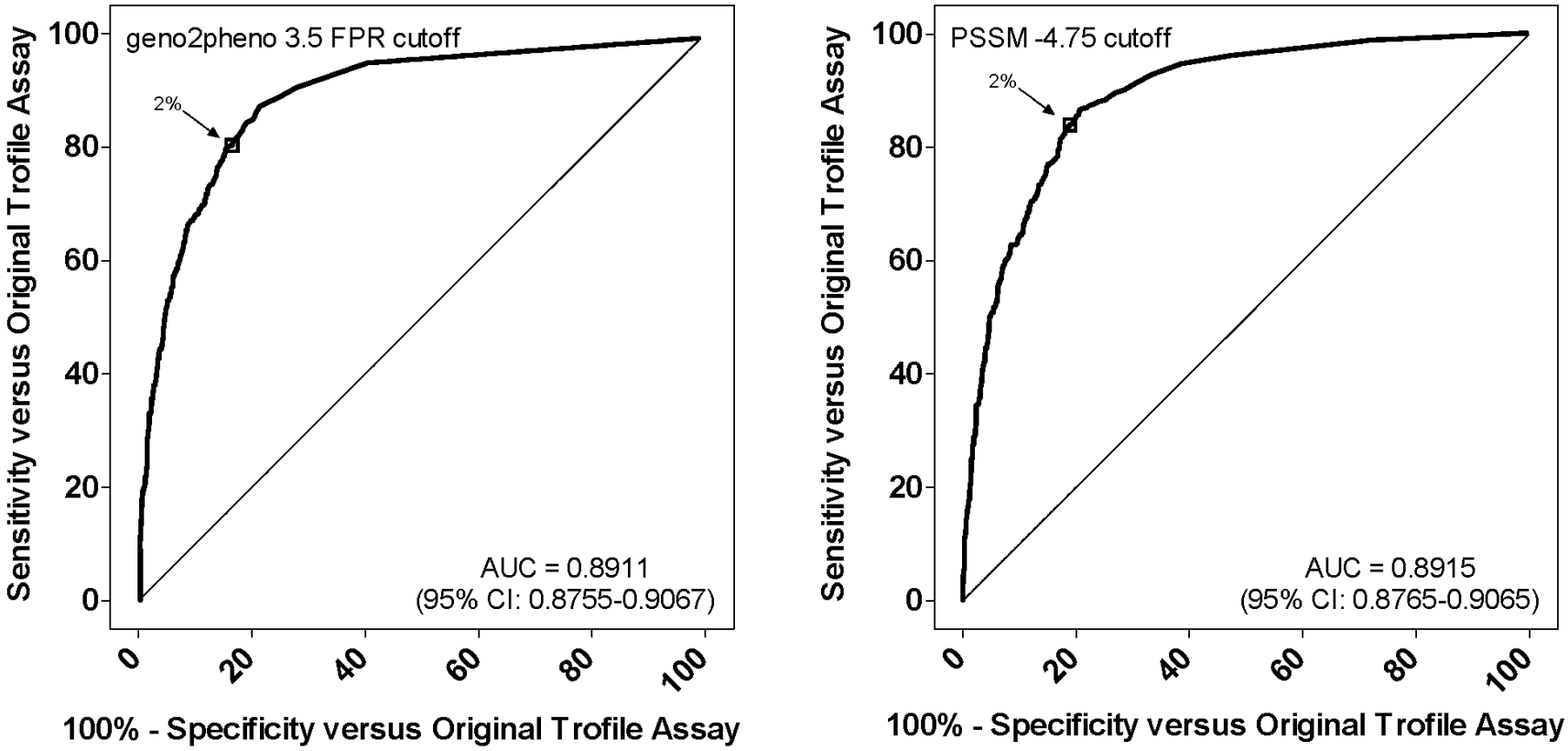


Figure 3.3: Receiver Operating Characteristic Curves for Optimal Geno2pheno & PSSM Cutoffs — Prediction of Original Trofile Assay Results. The optimal geno2pheno cutoff was a false-positive rate of 3.5%. The optimal PSSM cutoff was a score of -4.75. The areas under the curve (AUC) and their 95% confidence intervals (95% CI) are displayed under their respective curves. The specific point of 2% non-R5 variants is emphasized on each graph.

Figure 3.4: Receiver Operating Characteristic Curves for Optimal Geno2pheno & PSSM Cutoffs — Prediction of Week Eight Virologic Success

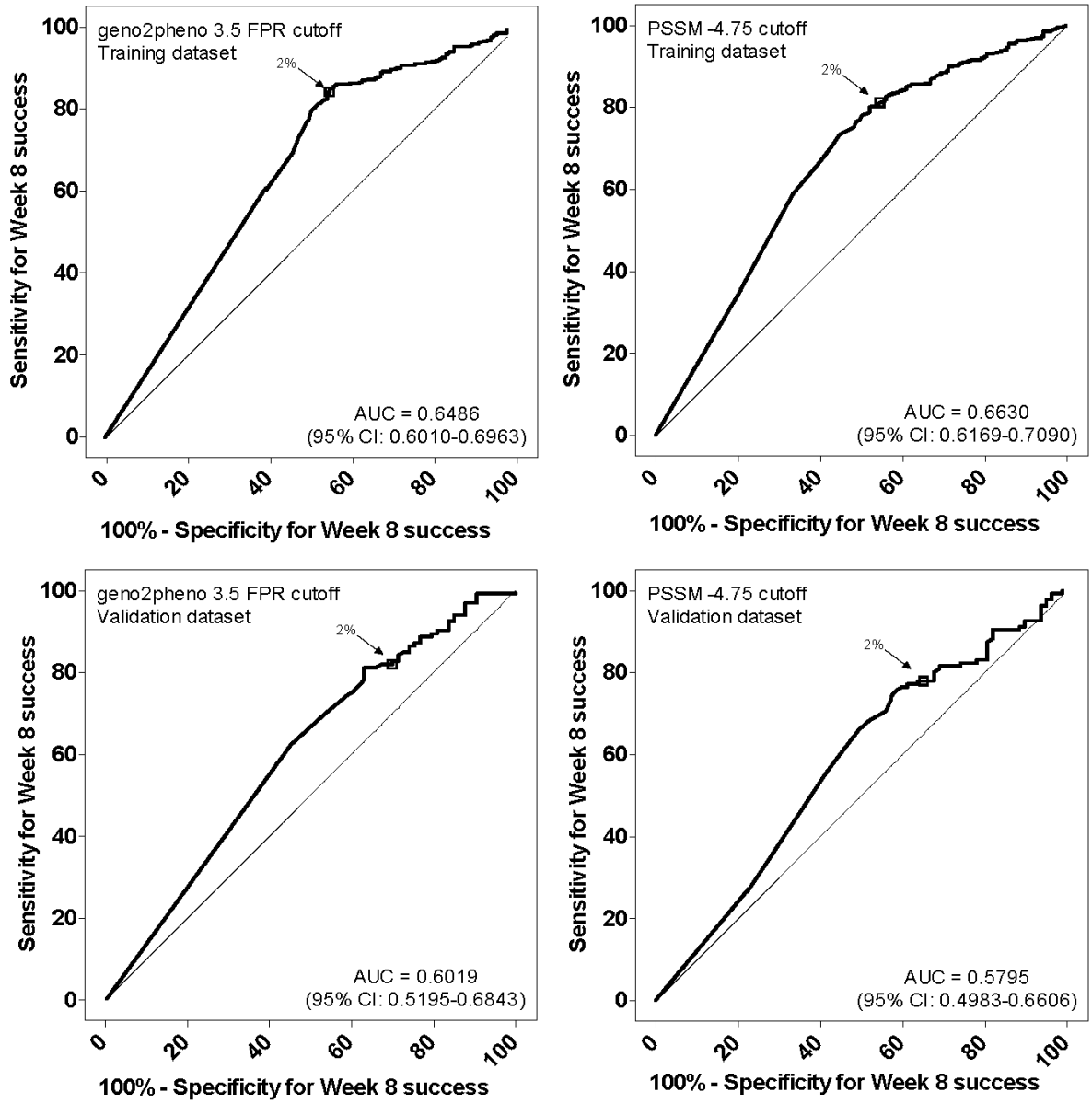


Figure 3.4: Receiver Operating Characteristic Curves for Optimal Geno2pheno & PSSM Cutoffs — Prediction of Week Eight Virologic Success. Receiver Operating Characteristic (ROC) curves for predicting week 8 virologic success on maraviroc in the training dataset (upper two panels) and validation dataset (lower two panels). The areas under the curve (AUC) and their 95% confidence intervals (95% CI) are displayed under their respective curves. The specific point of 2% non-R5 variants is emphasized on each graph. Performance for predicting virologic success was lower than performance predicting original Trofile assay results, likely due to the fact that clinical outcomes are inherently more difficult to predict than assay results. The number of samples included in the validation dataset was 684 for geno2pheno and 648 for PSSM. The number of samples in the validation dataset was 206 for geno2pheno and 200 for PSSM. Clinical sensitivity to predict virologic success was 84% for geno2pheno in the training set and 82% in the validation set. For PSSM these were 81% and 78%. The specificities (corresponding roughly to correct prediction of virologic non-response) were lower at 46%/30% for the geno2pheno training/validation datasets, and 45%/35% for the PSSM training/validation datasets.

These cutoffs were established by optimizing to week 8 virologic response in a random 75% of the dataset, and testing on the remaining 25%, as shown in section 3.2.3 [335–337]. Results for all patients are presented in the text. This 2% cutoff also approaches the likely level of reproducibility for PCR-based methods.

3.2.5 Population-Based Sequencing as a Comparator

Additional second-round PCR amplifications were also performed prior to standard population-based sequencing on an ABI 3730XL DNA analyzer, according to previously-described methods [325]. The cut-offs used were -4.25 for $PSSM_{X4/R5}$ and 5.75 for geno2pheno [335,336].

3.2.6 Data Analysis

Where clinical data were missing for patients enrolled in the clinical trials, the last observation was carried forward, except for the analysis of the proportion of patients with pVL <50 copies/mL, where a missing result was considered >50 copies/mL. Data from MOTIVATE-1, -2 and A4001029 were pooled, and both maraviroc arms were combined into a single group. Analyses were restricted to patients with tropism results from both assays.

Clinical parameters examined included: the median change in $\log_{10}$-transformed HIV plasma viral load (pVL) from baseline; the proportion of patients with a pVL <50 HIV RNA copies/mL; and time to a tropism switch. The performance of genotyping on these parameters was compared against that of the original Trofile assay. The above results could not be extensively compared to ESTA response rates in this population, as ESTA results were not available from most patients in these studies. However a subset of patients with ESTA results is examined in the results, as is a subset with replicate deep sequencing results performed by an independent laboratory. Comparisons of the performance between laboratories were assessed through X-Y correlation and Bland-Altman plots.

3.3 Results

3.3.1 Tropism Screening by Deep Sequencing Relative to Trofile & Population-Based Sequencing

Overall, genotyping identified 1037 samples (57%) as R5 and 790 (43%) as non-R5 using the $PSSM_{X4/R5}$ algorithm. PSSM and geno2pheno had approximately 90% concordance with one another. For ease of presentation, results will be shown for the $PSSM_{X4/R5}$ algorithm. When screened with Trofile, 1141 samples (62%) were called R5, and 686 (38%) were called non-R5 (Dual-/Mixed-tropic or X4). Global concordance of genotyping with Trofile-defined tropism was 82%.

Of the 686 samples identified as non-R5 by Trofile, genotyping agreed in 575 (84%) of cases. An additional 215 samples were further identified as non-R5 by genotyping. Using Trofile as a reference, sensitivity of genotyping was 84%, and specificity was 81%. Using genotyping as a reference, the sensitivity of original Trofile assay was 73%, and specificity was 89%. Comparing population-based against deep sequencing, overall concordance was 80%, with 64% sensitivity and 93% specificity. The sensitivity of both Trofile and population-based sequencing was lower when the proportion of non-R5 variants in the viral population was lower according to deep sequencing (Figure 3.5).

Of interest, deep sequencing detected at least some non-R5 HIV in >90% of patients (1700/1827), regardless of tropism classification. Samples with R5 HIV by Trofile had a median of 0.1% variants identified as non-R5 at screening, according to deep sequencing results. However, non-R5 levels below 1-2% likely have low reproducibility and should be considered with caution. The Trofile non-R5 group excluding Dual/Mixed samples (i.e., only "pure" X4 by Trofile, N=39) had a median of 93% non-R5-virus.

Patients screened R5 by Trofile but non-R5 by genotyping had 12% non-R5 HIV present at screening (N=167), much higher than where both assays indicated R5: 0.1% (N=926). This suggests that the original Trofile assay did not reliably detect patients with low level non-R5 variants present at screening (Figure 3.5).

Figure 3.5: Sensitivity of Population-Based V3 Sequencing & Original Trofile Assay with Deep Sequencing as the Reference

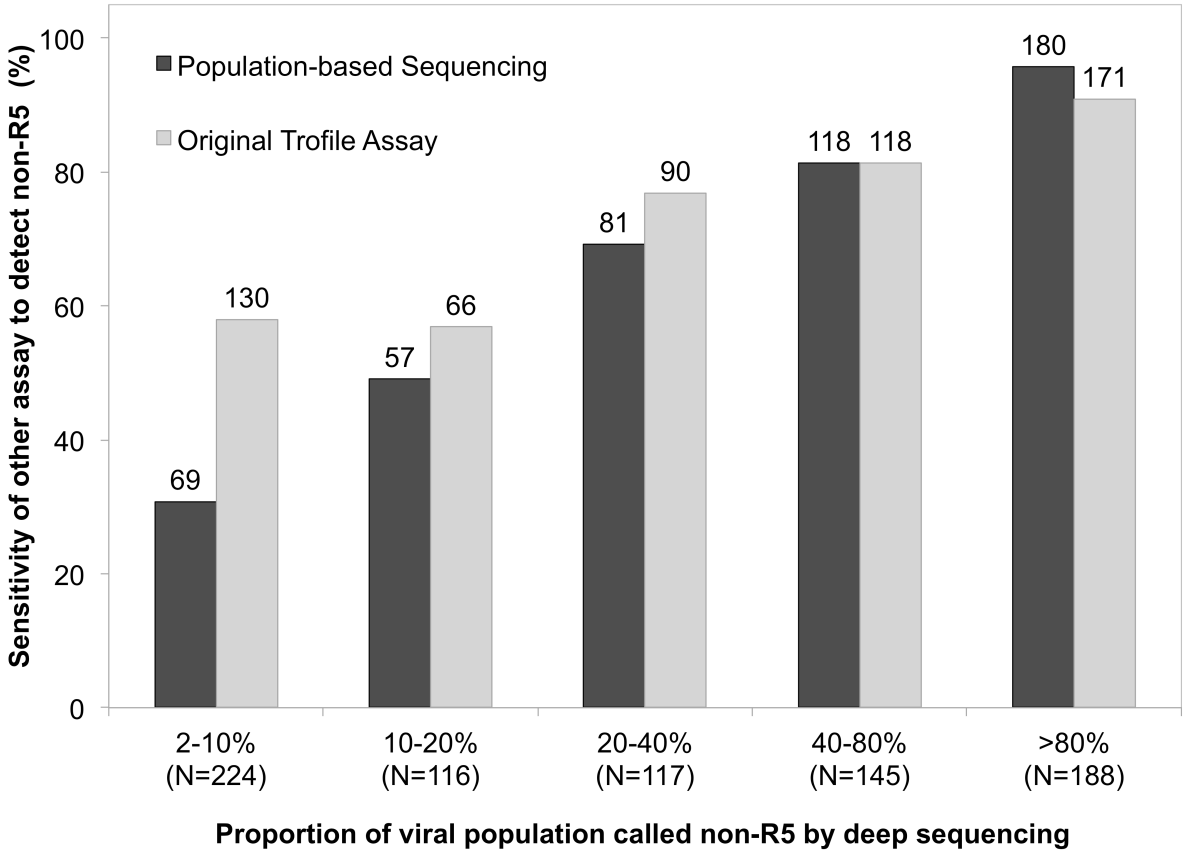


Figure 3.5: Sensitivity of Population-Based V3 Sequencing & Original Trofile Assay with Deep Sequencing as the Reference. The ability of the population-based V3 sequencing (black bars) and original Trofile (grey bars) to identify screening samples as non-R5 that deep sequencing had identified as having ≥ 2% non-R5 virus, stratified by different proportions of non-R5 virus identified in the deep sequencing result. Both alternative assays seemed to have decreased sensitivity for non-R5 HIV when such variants were present at lower proportions of the viral population. When non-R5 was present at 2-10% according to deep sequencing, 31% (69/224) & 58% (130/224) of samples were also called non-R5 by population-based sequencing & Trofile, respectively. These were 49% (57/116) & 57% (66/116) in the 10-20% group; 69% (81/117) & 77% (90/117) in the 20-40% group; 81% (118/145) & 81% (118/145) in the 40-80% group; and 96% (180/188) & 91% (171/188) in the >80% group.

This was consistent with the finding that 8% of patients had non-R5 results at baseline despite R5 Trofile results at screening [288], and consistent with ESTA results for the MERIT trial of maraviroc [290].

Of the 1827 patients screened, 1093 actually entered the maraviroc (N=851) or placebo (N=242) arms of the trials. Baseline characteristics of patient groups screened by both methods are presented in Table 3.1. The R5 groups by either method were similar in terms of baseline viral load and $CD4^+$ cell

count, as were the non-R5 groups. Amongst maraviroc recipients, genotyping identified over twice as many patients than Trofile as being unlikely to respond to maraviroc (N=240 versus 111).

Table 3.1: Baseline Characteristics of Treated Population, Stratified by Tropism Status by Genotype and Phenotype

	Geno R5 (N=775)	Trofile R5 (N=925)	Geno non-R5 (N=318)	Trofile non-R5 (N=168)
Baseline pVL, median log_{10} HIV-RNA copies/mL	4.85	4.88	5.04	5.07
Median CD4 cell count, cells/mm^3	177	168	72	54
Median percent non-R5 variants in deep sequencing screening result, % (IQR)	0.1% (0 – 0.2%)	0.1% (0 – 0.7%)	19% (7-54%)	28% (7-63%)

Table 3.1: Baseline Characteristics of Treated Population, Stratified by Tropism Status by Genotype & Phenotype.
Geno non-R5, identified as having non-R5 virus by genotyping; Geno R5, identified as having R5 virus by genotyping; IQR, interquartile range; pVL, plasma viral load

3.3.2 Early Virologic Response to Maraviroc

Screening genotype was a predictor of response to maraviroc-based antiretroviral therapy in treatment-experienced patients. Maraviroc recipients screened with R5 HIV by genotyping had consistently better virologic outcomes than those screened non-R5. Using a number of parameters, genotyping performed similarly to, or marginally out-performed the original Trofile assay in predicting virologic response. Virologic performance was slightly better in the maraviroc BID arm compared to the QD arm (data not shown), but these arms have been pooled to simplify the presentation of the results.

The median pVL change from baseline to week 8 of treatment was examined in order to minimize the number of patients who had discontinued the study due to reasons such as treatment failure, or loss-to-follow-up, but with sufficient time to measure the efficacy of maraviroc in patients. Maraviroc recipients screened R5 by genotyping had a combined median week 8 decrease in pVL from baseline

of 2.4 $\log_{10}$ (Inter-quartile Range [IQR]: 1.7 – 2.9; N=611). This was twice as large as the 1.4 $\log_{10}$ decline (IQR: 0.2 – 2.7; N=240) for patients classified non-R5 by genotyping.

Results where missing patients were censored, or where data were restricted to only those screened for MOTIVATE-1, were largely similar (data not shown). As mentioned in Section 3.2.3 on optimizing bioinformatic cutoffs, clinical sensitivity to predict virologic success on maraviroc was approximately 80% when patients were screened with a 2% non-R5 variant cutoff (Figure 3.4).

Using Trofile, the corresponding week 8 viral load declines were similar: 2.4 $\log_{10}$ (IQR: 1.3 – 2.8; N=740) for R5 patients versus 1.3 $\log_{10}$ (IQR: 0.3 – 2.7; N=111) for non-R5 patients. For placebo recipients, the week 8 pVL declines were modest (0.5 – 0.8 $\log_{10}$) and similar regardless of genotypic tropism. Median pVL responses on maraviroc and placebo over the course of the studies are shown in Figures 3.6A and 3.6B, where prediction of virologic outcomes can be compared between the genotypic and phenotypic assays.

3.3.3 Longer-Term Virologic Efficacy

The efficacy of maraviroc was sustained to week 48 in patients identified by genotyping as having R5 HIV. The primary endpoint for the MOTIVATE trials was the percentage of patients with viral loads <50 copies/mL at week 48. The proportion of patients achieving virologic suppression <50 copies/mL was assessed throughout the study for both the maraviroc and placebo arms. Maraviroc recipients with R5 HIV at screening were more likely to achieve a pVL <50 copies/mL at week 48 compared to the non-R5 group. In total, 49% (301/611) of the R5 group, and 26% (62/240) of the non-R5 group had virologic suppression at week 48 when screened by genotyping. By Trofile, these were 46% (337/740) and 23% (26/111), respectively (Figure 3.7).

Figure 3.6: Median Change in Plasma Viral Load from Baseline in the Maraviroc & Placebo Arms

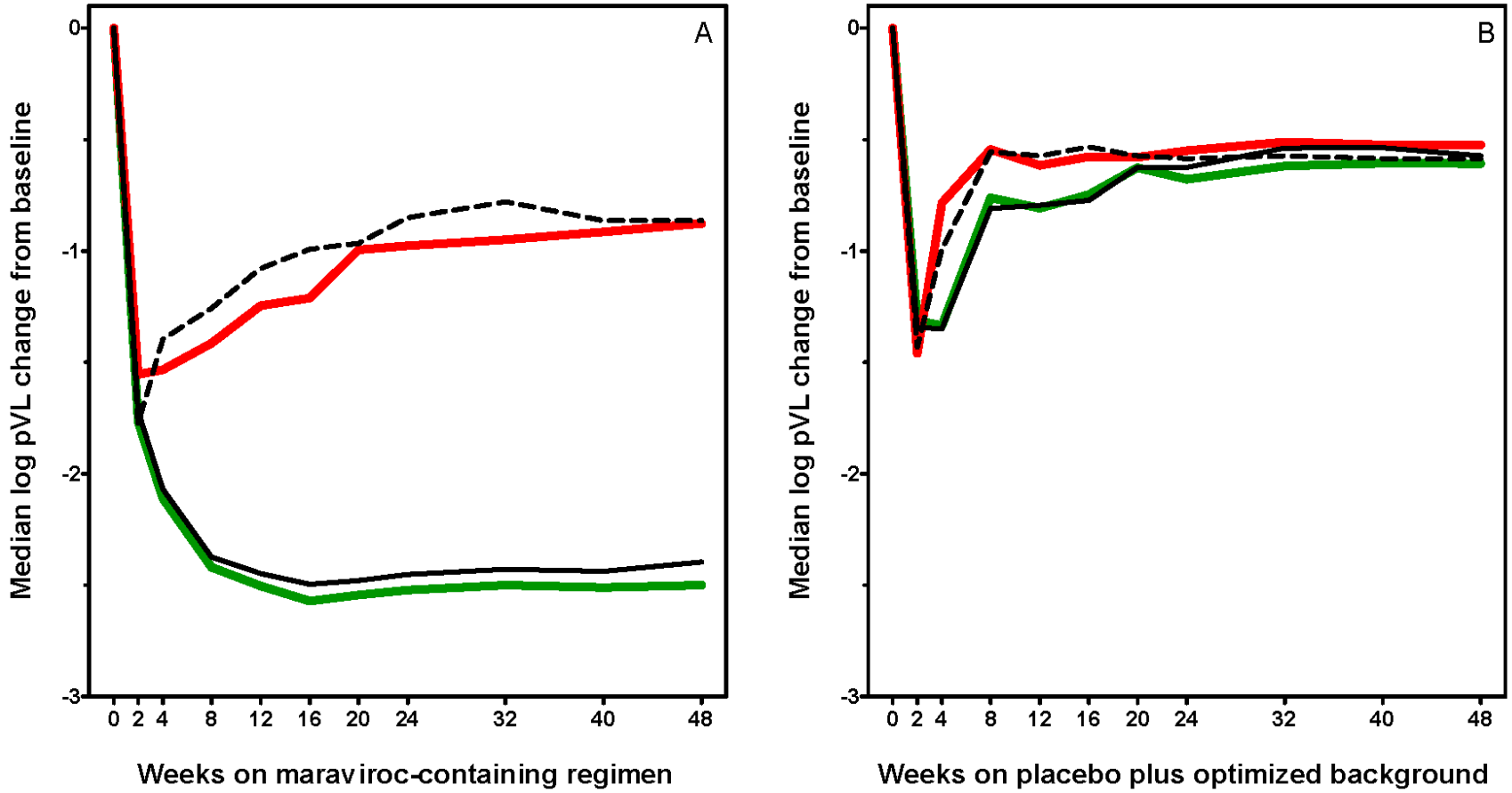


Figure 3.6: Median Change in Plasma Viral Load from Baseline in the Maraviroc & Placebo Arms. Left: Panel A shows the responses in the maraviroc arms. Patients screened as R5 by either genotyping or Trofile had much larger median pVL declines from baseline relative to patients screened as non-R5. Green and red lines correspond to deep sequencing R5 (N=611) and non-R5 (N=240) groups respectively, while solid black and dotted black lines correspond to Trofile R5 (N=740) and non-R5 (N=111) groups. Right: Panel B shows the pVL declines from baseline for patients receiving placebo were similar to MVC-receiving patients identified as non-R5 or non-R5, and were small regardless of screening tropism or assay used. Green and red lines correspond to deep sequencing R5 (N=164) and non-R5 (N=78) groups respectively, while solid black and dotted black lines correspond to Trofile R5 (N=185) and non-R5 (N=57) groups.

Figure 3.7: Percentage of Patients with Viral Loads Less than 50 HIV RNA Copies/mL in the Maraviroc & Placebo Arms

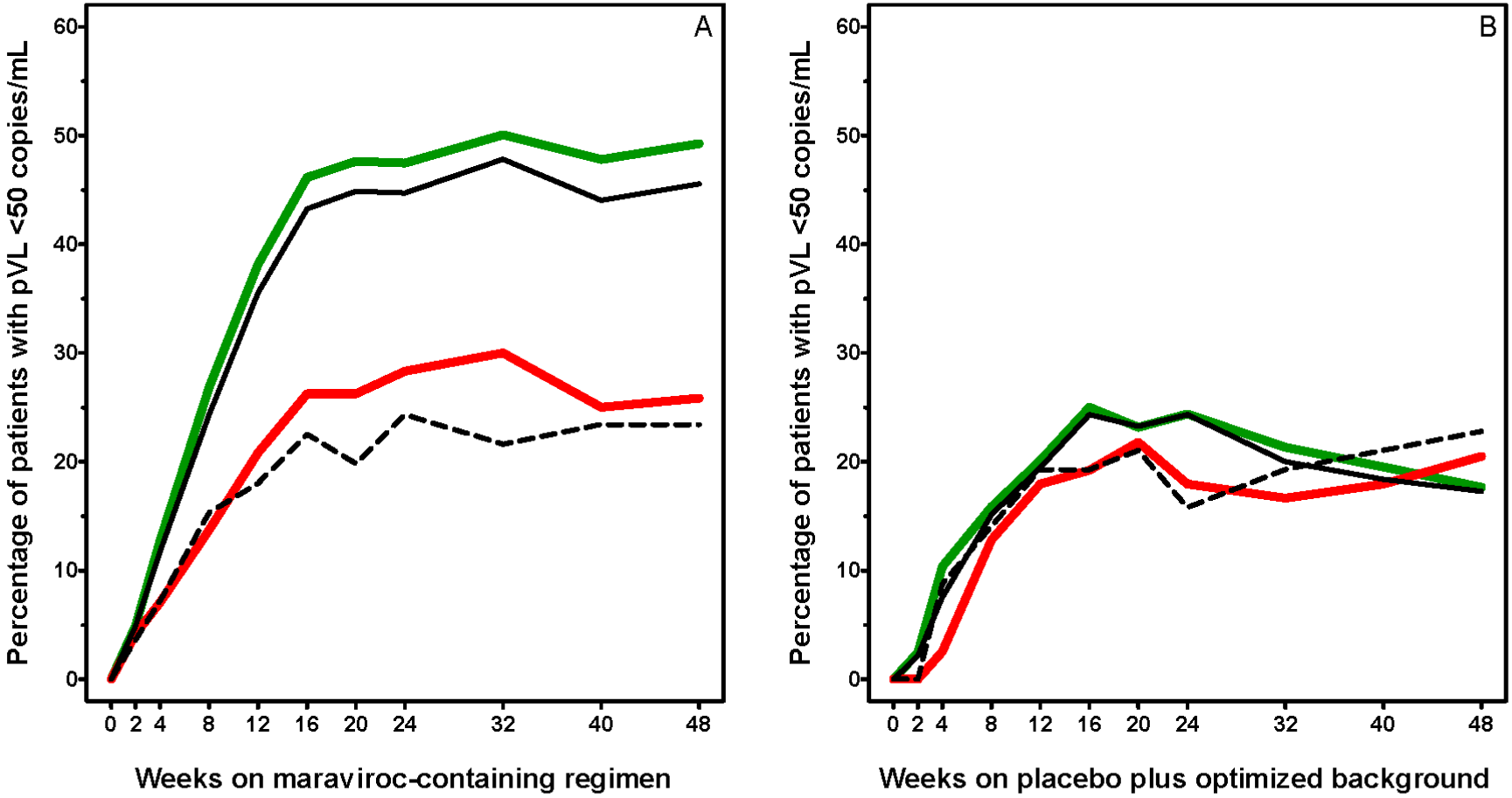


Figure 3.7: Percentage of Patients with Viral Loads Less than 50 HIV RNA Copies/mL in the Maraviroc & Placebo Arms. Left: Panel A shows the proportion with virologic suppression in the maraviroc arms. A higher proportion of maraviroc recipients screened by either method as R5 had a pVL <50 HIV RNA copies/mL compared to the non-R5 patients. Green and red lines correspond to deep sequencing R5 (N=611) and non-R5 (N=240) groups respectively, while solid black and dotted black lines correspond to Trofile R5 (N=740) and non-R5 (N=111) groups. Right: Panel B shows the proportion with virologic suppression in the placebo arms. Green and red lines correspond to deep sequencing R5 (N=164) and non-R5 (N=78) groups respectively, while solid black and dotted black lines correspond to Trofile R5 (N=185) and non-R5 (N=57) groups.

The genotypic non-R5 group could be divided roughly in half, with 127 patients having low-prevalence (2-20%) non-R5, and 113 having >20% non-R5 virus. The group of patients with 2-20% non-R5 according to deep sequencing had minority non-R5 variants that were not reliably detected by standard population-based sequencing methods (Figure 3.3). Importantly, this group of patients had a poor response to maraviroc, with 27% (34/127) of patients achieving virologic suppression at week 48, similar to the non-R5 group as a whole (26%) and to patients with >20% non-R5 (25%; 28/113). The virologic responses of placebo recipients were similarly low to maraviroc recipients identified as having non-R5 HIV, ranging from 17-23% depending on tropism or assay.

Interestingly, the virologic outcomes of maraviroc recipients showed a general inverse relationship with the percentage of non-R5-virus present at screening according to genotyping. Patients with 0% non-R5 had the best success, showing a week 8 $\log_{10}$ pVL decline of 2.6, with 65% (58/89) of patients having week 48 virologic suppression. Patients with between 0 and 1% non-R5 had slightly poorer outcomes: week 8 pVL $\log_{10}$ decline of 2.4, 48% (234/491) with virologic suppression. This declined again in patients with between 1-2% non-R5: 2.1 $\log_{10}$, 29% (9/31). Patients with >2% non-R5 (i.e., the genotyping non-R5 group) all showed similar low virologic responses, as detailed above.

3.3.4 Changes in Viral Tropism

As a separate endpoint, patients were analysed according to whether they experienced a change in their Trofile result from R5 to non-R5 (a tropism "switch") over the course of the studies, as measured by Kaplan-Meier analysis. This parameter is both clinically relevant for maraviroc-based therapy, and functioned as a measure separate from changes in viral load measurements. Amongst those patients originally screened as R5 by Trofile, those identified to be non-R5 by genotyping were almost twice as likely to have non-R5 HIV emerge by week 24 compared to patients screened as R5 by both methods (Figure 3.8). A total of 40% (72/180) of maraviroc recipients who switched tropism were identified by genotyping as having ≥2% non-R5 virus. Tropism switches occurred in 18% (111/612) of the group called R5 by genotyping, lower than in those called R5 by Trofile alone: 25% (180/724).

Figure 3.8: Time to Change in Tropism from R5 to Dual/Mixed or X4 in the Maraviroc & Placebo Arms

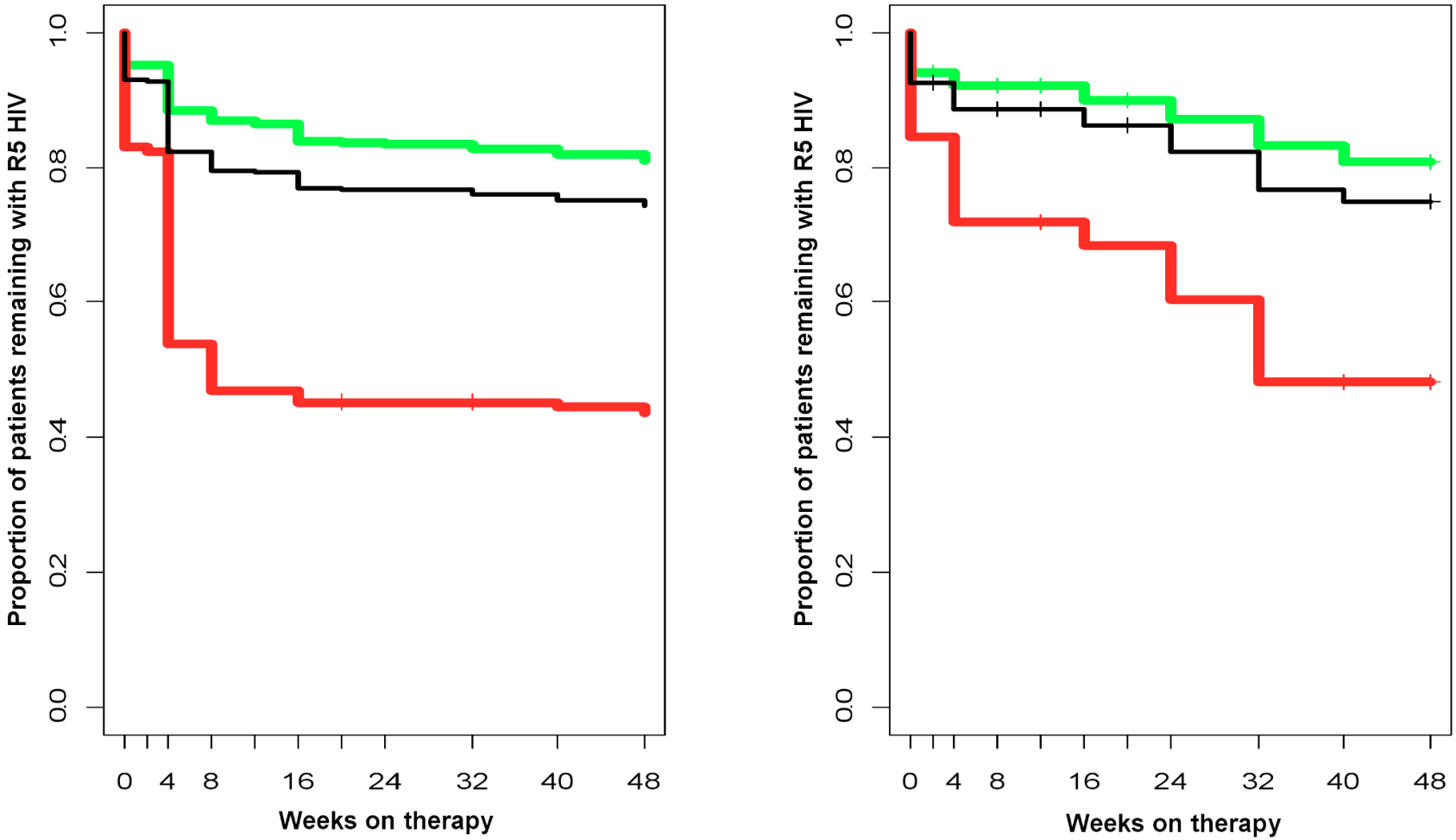


Figure 3.8: Time to Change in Tropism from R5 to Dual/Mixed or X4 in the Maraviroc & Placebo Arms. Left: The change in tropism in the MVC arms where all patients were R5 at screening by Trofile and switched tropism to DM or X4 over the course of the studies according to the Trofile assay. Green and red lines correspond to genotyping R5 (N=605) and non-R5 (N=135) groups respectively, while the solid black lines correspond to the Trofile R5 (N=740) group. Right: The change in tropism in the placebo arms for patients with Trofile R5 results at screening. Green and red lines correspond to genotyping R5 (N=153) and non-R5 (N=32) groups respectively, while the solid black lines correspond to the Trofile R5 (N=185) group.

Amongst patients who switched tropism, maraviroc recipients classified as R5 by Trofile but non-R5 by genotyping had been on treatment for a mean of 4.6 weeks before Trofile gave a non-R5 result, over twice as quickly (9.7 weeks) as where both tests indicated R5 (Figure 3.8).

3.3.5 Response Stratified by Background Drug Activity

Patients were also classified according to a weighted optimized background therapy susceptibility score (wOBTss). In general, wOBTss was defined as the number of active drugs in the background regimen at baseline, with nucleoside reverse transcriptase inhibitors scoring 0.5 [339]. Genotyping was predictive of virologic success on maraviroc-based therapy regardless of wOBTss. Maraviroc was successful in either of the R5 groups where the wOBTss was between 1 and 2. The proportions of these patients with a week 48 pVL <50 copies/mL were 58% (179/311) and 53% (205/389) when screened by genotyping and Trofile, respectively. The predictive ability of genotyping was more pronounced at more compromised background regimens. The proportions with undetectable viral loads were 33% (81/232) versus 29% (78/271) of R5-classified patients with wOBTss <1 by genotyping or Trofile, respectively (Figure 3.9).

3.3.6 Discordance Amongst Bioinformatic Algorithms

There was a high degree of concordance using alternative bioinformatic algorithms. Geno2pheno and $PSSM_{SINSI}$ were compared to $PSSM_{X4R5}$ in terms of their ability to predict various virologic outcomes. Detailed analyses of these comparisons are presented in Tables 3.2 – 3.4. In general, where the algorithms disagreed on tropism classifications, the virologic responses were intermediate between the concordant R5 and concordant non-R5 groups. This indicated that no algorithm clearly outperformed the others.

Figure 3.9: Median Change in Plasma Viral Load from Baseline in Patients with R5 Virus Stratified by Their Weighted Optimized Background Therapy Susceptibility Score (wOBTss)

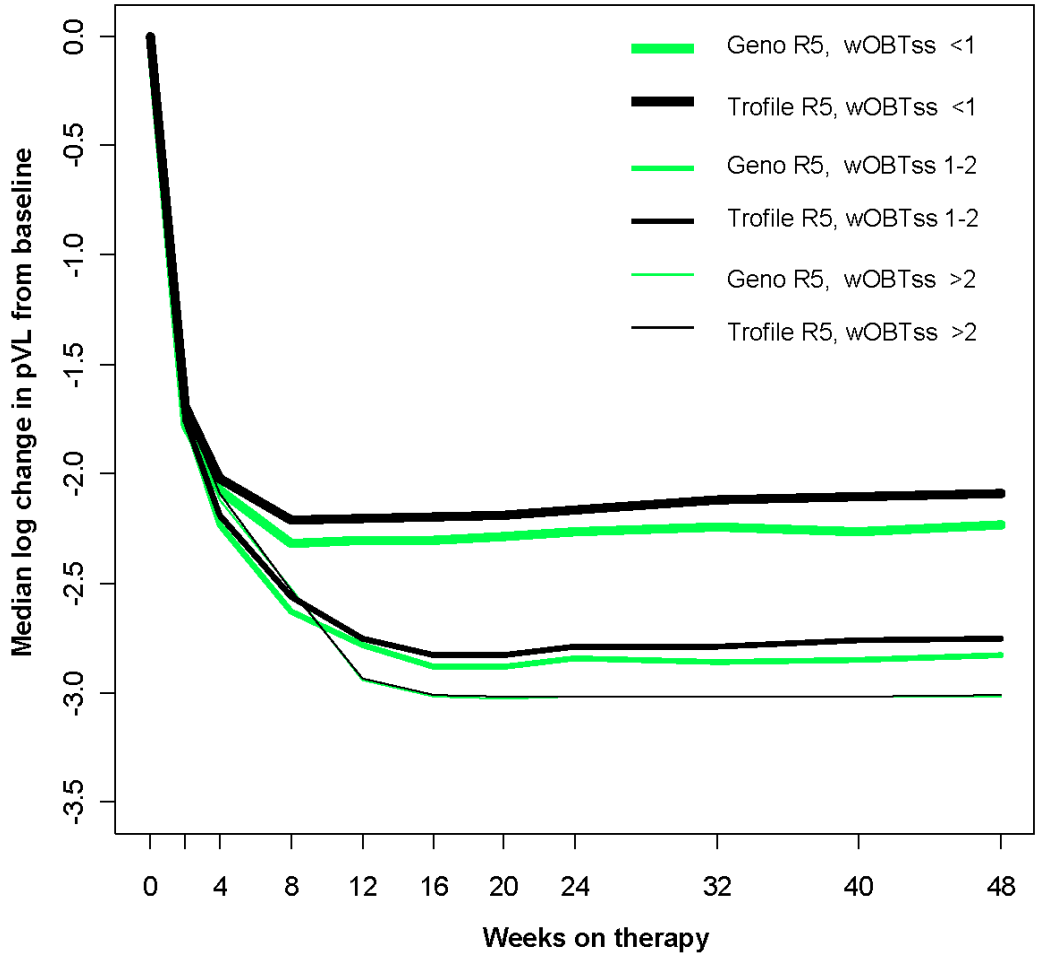


Figure 3.9: Median Change in Plasma Viral Load from Baseline in Patients with R5 Virus Stratified by Their Weighted Optimized Background Therapy Susceptibility Score (wOBTss). Maraviroc-treated patients screened as R5 by genotyping (green lines) or Trofile (black lines). Patients screened as R5 by either method who also had wOBTss > 2 (N=68 or 80, respectively) showed the largest pVL declines from baseline. Patients with wOBTss 1-2 (N=151 or 197) showed intermediate pVL decline and patients with wOBTss≤1 (N=392 or 463) showed poorer changes in pVL. The wOBTss > 2, 1-2, and ≤ 1 groups are indicated by thin, intermediate, and thick lines, respectively.

Table 3.2: Discordance amongst Bioinformatic Algorithms in the Maraviroc Arms — $PSSM_{X4/R5}$ versus Geno2pheno

Concordance or Discordance $PSSM_{X4R5}$/geno2pheno	N	Week 8 log_{10} pVL decline (IQR)	%<50c/mL @ week 48 (N)	% Changing tropism (N)
R5/R5	583	2.43 (1.79 – 2.86)	50% (289/583)	17% (97/577)
R5/Non-R5	26	2.06 (0.33 – 2.59)	38% (10/26)	54% (14/26)
Non-R5/R5	55	2.19 (0.68 – 2.79)	36% (20/55)	26% (12/46)
Non-R5/Non-R5	179	0.98 (0.12 – 2.61)	22% (40/179)	70% (62/89)

Table 3.2: Discordance amongst Bioinformatic Algorithms in the Maraviroc Arms — $PSSM_{X4/R5}$ versus Geno2pheno.
Virologic outcomes for patients with discordant results between algorithms. $PSSM_{X4R5}$ (non-R5 if ≥2% scored ≥-4.75) versus geno2pheno (non-R5 if ≥2% scored ≤3.5)

Table 3.3: Discordance amongst Bioinformatic Algorithms in the Maraviroc Arms — $PSSM_{X4/R5}$ versus $PSSM_{SI/NSI}$

Concordance or Discordance $PSSM_{X4R5}$/$PSSM_{SINSI}$	N	Week 8 log_{10} pVL decline (IQR)	%<50c/mL @ week 48 (N)	% Changing tropism (N)
R5/R5	557	2.43 (1.78 – 2.86)	49% (275/557)	17% (92/553)
R5/Non-R5	52	2.34 (0.87 – 2.78)	46% (24/52)	38% (19/50)
Non-R5/R5	17	2.44 (0.47 – 3.05)	24% (4/17)	27% (4/15)
Non-R5/Non-R5	217	1.27 (0.21 – 2.63)	26% (56/217)	58% (70/120)

Table 3.3: Discordance amongst Bioinformatic Algorithms in the Maraviroc Arms — $PSSM_{X4/R5}$ versus $PSSM_{SI/NSI}$.
Virologic outcomes for patients with discordant results between algorithms. $PSSM_{X4R5}$ (non-R5 if ≥2% scored ≥-4.75) versus $PSSM_{SINSI}$ (non-R5 if ≥2% scored ≥-3.5)

Table 3.4: Discordance amongst Bioinformatic Algorithms in the Maraviroc Arms — $PSSM_{SI/NSI}$ versus Geno2pheno

Concordance or Discordance $PSSM_{SINSI}$/geno2pheno	N	Week 8 log_{10} pVL decline (IQR)	%<50c/mL @ week 48 (N)	% Changing tropism (N)
R5/R5	559	2.43 (1.76 – 2.86)	49% (273/559)	16% (91/554)
R5/Non-R5	15	2.39 (1.48 – 2.70)	40% (6/15)	36% (5/14)
Non-R5/R5	79	2.35 (1.56 – 2.79)	46% (36/79)	26% (18/69)
Non-R5/Non-R5	190	0.98 (0.11 – 2.60)	23% (44/190)	70% (71/101)

Table 3.4: Discordance amongst Bioinformatic Algorithms in the Maraviroc Arms — $PSSM_{SI/NSI}$ versus Geno2pheno. Virologic outcomes for patients with discordant results between algorithms. $PSSM_{SINSI}$ (non-R5 if ≥2% scored ≥-3.5) versus geno2pheno (non-R5 if ≥2% scored ≤3.5)

3.3.7 Assay Discordance

Where screening assays differed, virologic outcomes on maraviroc slightly favoured the genotyping results. Amongst the discordant patients, where Trofile indicated R5 but genotyping identified >2% non-R5 virus (N=135), median log_{10} pVL declines were lower, at 1.8 log_{10}. For comparison, the concordant non-R5 group (N=105) had 1.2 log_{10} pVL decline. Where genotyping screened patients as having R5 virus, but Trofile screened as non-R5, the median week 8 pVL decline was 2.6 log_{10} (N=6), similar to the concordant R5 group: 2.4 log_{10} (N=605).

When deep sequencing was compared to population-based sequencing, virologic outcomes again favoured the deep sequencing results. The results over 48 weeks of treatment, with patients grouped by assay concordance or discordance are shown in Figure 3.10. Furthermore, there were a large number of patients with 50% or lower non-R5 variants, which could potentially be difficult to detect by standard population-based sequencing. Importantly, these patients had poor virologic outcomes on maraviroc compared to those with low non-R5 prevalence of <2% (Figure 3.11).

Figure 3.10: Where Assay Results Were Discordant, the Virologic Responses Tended to Favour Deep Sequencing Results

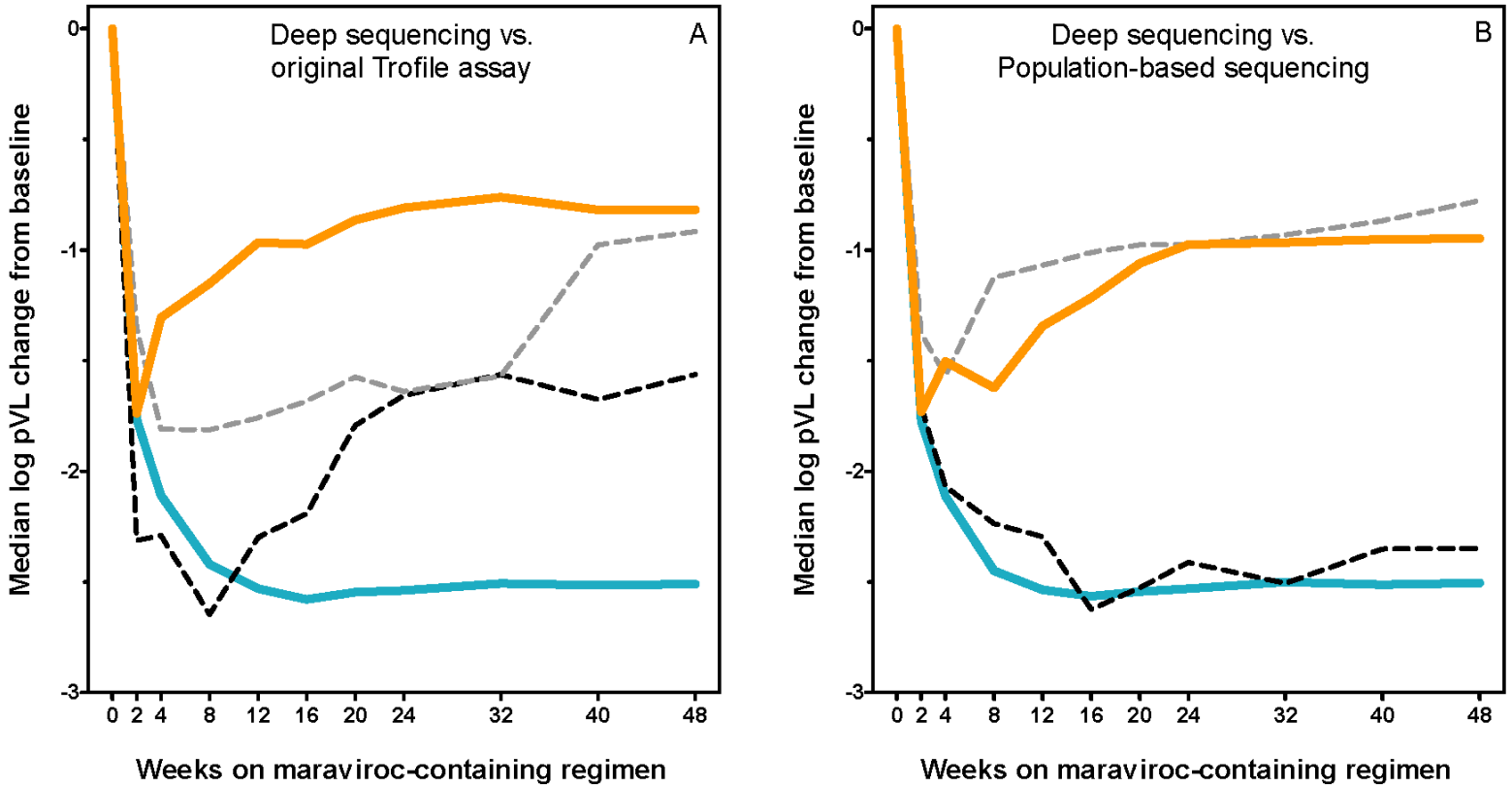


Figure 3.10: Where Assay Results Were Discordant, the Virologic Responses Tended to Favour Deep Sequencing Results. Panel A shows virologic responses to maraviroc for patients stratified by whether deep sequencing and the original Trofile assay gave concordant or discordant results. The turquoise line represents patients with concordant R5 results (N=605), the orange line represents concordant non-R5 results (N=105). The dashed black line represents patients where deep sequencing indicated R5 but Trofile indicated non-R5 (N=6), and the dashed grey line represents patients with non-R5 by deep sequencing but R5 by Trofile (N=135). Panel B shows virologic responses to maraviroc for patients stratified by whether deep sequencing and population-based sequencing gave concordant or discordant results. The turquoise line represents patients with concordant R5 results (N=573); the orange line represents concordant non-R5 results (N=127). The dashed black line represents patients where deep sequencing indicated R5 but population-based sequencing indicated non-R5 (N=38), and the dashed grey line represents patients with non-R5 by deep sequencing but R5 by population-based sequencing (N=113). Overall, the discordant lines tended to favour the deep sequencing classifications, especially when compared to population-based sequencing.

Figure 3.11: Virologic Response of Maraviroc Recipients as a Function of the Proportion of Non-R5 HIV at Screening.

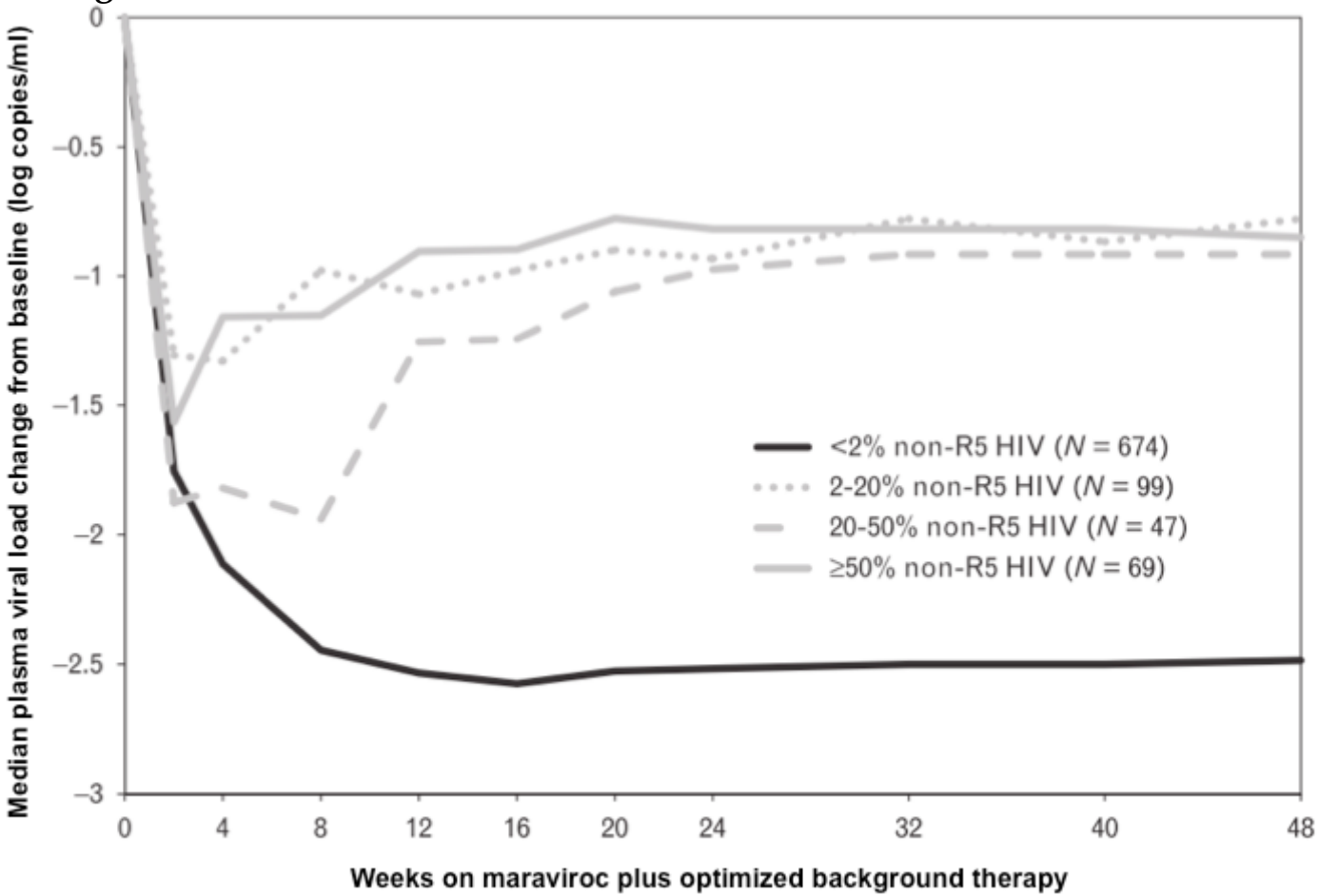


Figure 3.11: Virologic Response of Maraviroc Recipients as a Function of the Proportion of Non-R5 HIV at Screening. The median plasma viral load change from baseline is shown for groups of treatment-experienced maraviroc recipients. Patients are grouped according to the percentage of non-R5 HIV detected at screening by deep sequencing using a geno2pheno false-positive rate of 3.5. Patients with less than 2% non-R5 HIV had a median pVL decline from baseline of approximately 2.5 $\log_{10}$ copies/mL at week 48. Patients with greater than 2% non-R5 all had similarly poor responses to maraviroc, with a median pVL decline of approximately 1 $\log_{10}$. Patients with 2-20% non-R5 are less likely to have non-R5 detected using population-based sequencing, but still had poor virologic responses. This figure is adapted with permission from Swenson et al, Current Opinion in HIV and AIDS. 2012; 7(5): 478-485.

Having either or both assays indicate non-R5 was a poor prognostic indicator of longer-term maraviroc response. At week 48, the proportions of patients with suppressed viraemia were: 27% (36/135) for the Trofile R5_geno non-R5 group and 0% (0/6) for the Trofile non-R5_geno R5 group. Week 48 suppression was twice as high in the concordant-R5 group versus the concordant-non-R5 group: 50% (301/605) versus 25% (26/105), respectively. In the Trofile-R5_geno-non-R5 group, 55% (74/135) of maraviroc recipients changed tropism, much higher than the concordant-R5 group: 18% (111/605). Most patients screened by Trofile as non-R5 remained so over the course of the study period, regardless of concordance with genotyping.

3.3.8 Comparison with Independent Replication by an External Laboratory

In order to assess the reproducibility of this method, a subset of 310 samples from this dataset were also processed by an independent laboratory (Quest Diagnostics Nichols Institute, California). For this comparison, the geno2pheno algorithm was used to infer tropism rather than $PSSM_{X4/R5}$.

This dataset was selected to represent an unbiased subpopulation of enrollees into the MOTIVATE and A4001029 trials [340,341]. This was achieved by restricting to those patients entering the maraviroc arms of MOTIVATE and A4001029 until the latter study had completely enrolled. Since at that point, patients with non-R5 original Trofile assay results were no longer able to enroll into a trial, then by restricting to this subset of patients, the population is not biased by initial tropism screening.

Inter-laboratory concordance was excellent between the two implementations of deep sequencing, with 92% of samples receiving the same tropism classification (Kappa=0.84). The percentage of non-R5 variants detected by the laboratories was also highly correlated in the samples, with a Pearson's correlation coefficient of 0.92, and a median difference of only 0.3% (interquartile range: 0-3%) (Figure 3.12). Bland-Altman analysis demonstrated that there was limited bias in the percentage of non-R5 variants detected by either laboratory (Figure 3.13).

Figure 3.12: Correlation of Deep Sequencing Between Two Independent Laboratories

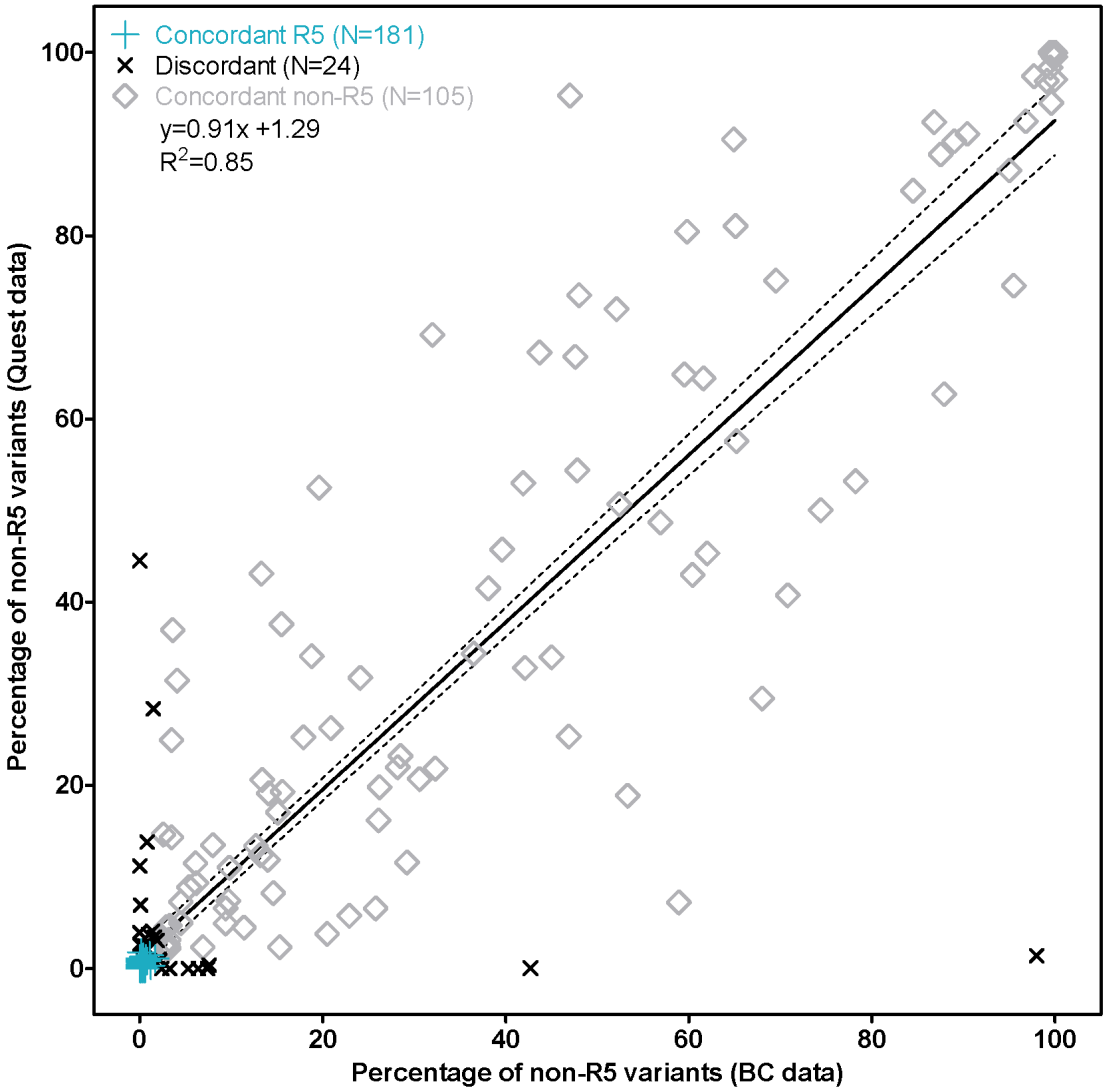


Figure 3.12: Correlation of Deep Sequencing Between Two Independent Laboratories. The correlation between both laboratories in terms of the percentage of non-R5 variants was high (R^2=0.85). Points are marked by whether they represent concordant or discordant results between laboratories (see key). The line of best-fit for the linear regression is shown as a black line, and bordered by the 95% confidence intervals, shown with dashed lines. The slope of the line of best-fit was significantly non-zero ($p<0.0001$).

Figure 3.13: Bland-Altman Plot Comparing the Assay Results from Two Independent Laboratories

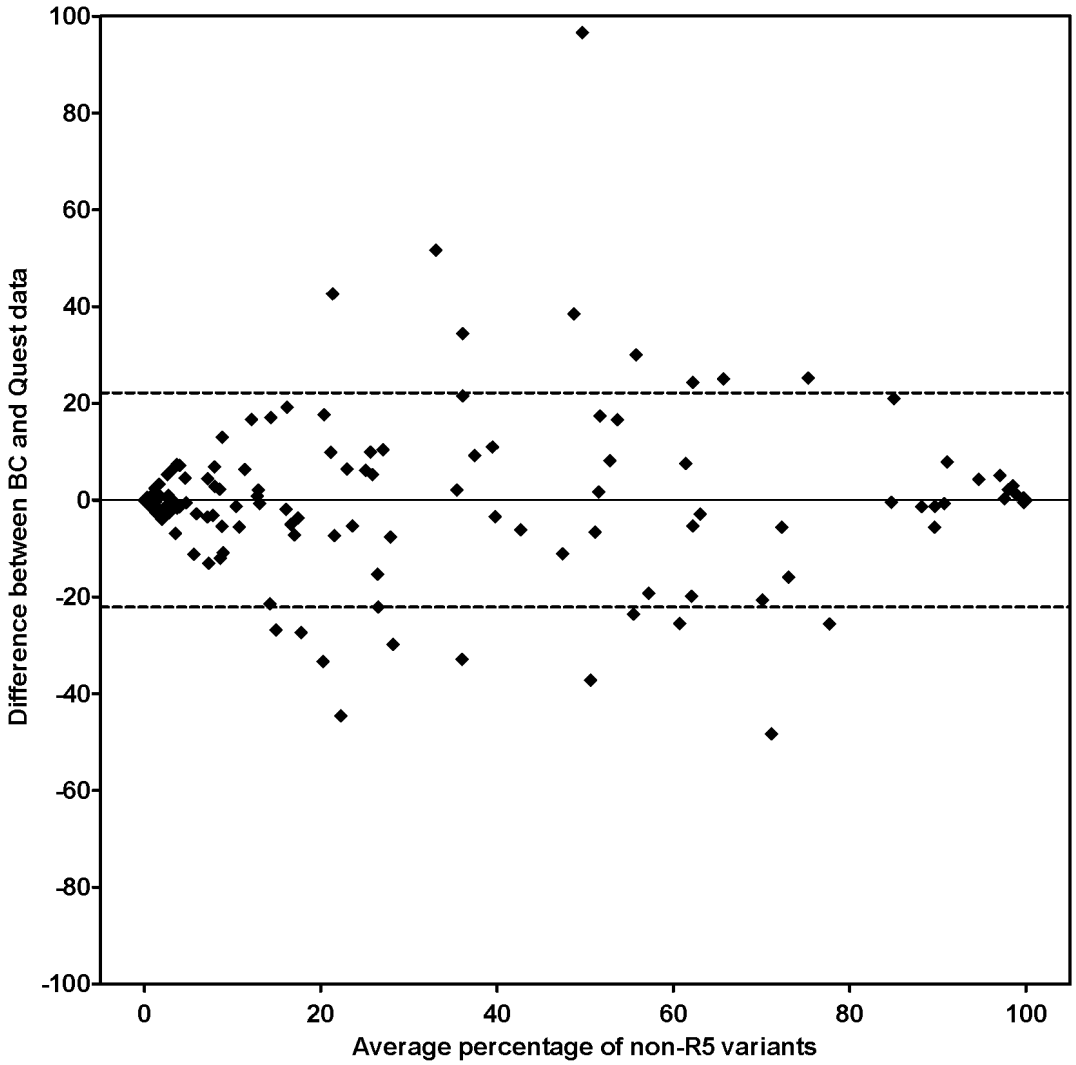


Figure 3.13: Bland-Altman Plot Comparing the Assay Results from Two Independent Laboratories.
There was minimal bias in the results between the laboratories according to the Bland-Altman analysis (bias=0.048%). A total of 95% of the data points were within ±22% of one another (dashed lines represent the 95% limits of agreement).

3.3.9 Comparison to the Enhanced Sensitivity Trofile Assay

Of the 310 samples processed by the two independent laboratories, 294 also had results from the Enhanced Sensitivity Trofile Assay (ESTA) at screening. Because ESTA has replaced the original Trofile assay, the performance of deep sequencing was compared to ESTA. Relative to ESTA, deep

sequencing had 76% sensitivity (87 of 114 samples), 89% specificity (160 of 180 samples) and 82% concordance (247 of 294 samples).

Furthermore, both assays predicted subsequent virologic outcomes on maraviroc equally well. Patients classified with R5 HIV had plasma viral load declines from baseline to week 8 of 2.4 $\log_{10}$ copies/mL and 2.4 $\log_{10}$ copies/mL by ESTA and deep sequencing, respectively. Those with non-R5 HIV at screening had week 8 declines of 0.9 $\log_{10}$ copies/mL and 0.9 $\log_{10}$ copies/mL, respectively (Figure 3.14). Where assays gave discordant results, virologic outcomes were intermediate between the concordant R5 and non-R5 groups, with neither assay seeming to give superior predictions to the other (Figure 3.15).

3.4 Discussion & Conclusions

This study represents one of the largest clinical applications of next-generation sequencing technology to date. Deep V3 sequencing was able to detect and quantify low prevalence sub-populations of CXCR4-using HIV within a large set of clinical isolates. This method was predictive of virologic response to a maraviroc-containing regimen and matched or surpassed the predictive ability of the original Trofile assay on a number of parameters, including the proportion of patients achieving a plasma viral load <50 copies/mL, and the likelihood of switching tropism while receiving maraviroc.

Retrospectively screening with genotyping led to over twice as many maraviroc recipients being identified with non-R5 HIV. Trofile-R5 patients screened as non-R5 by genotyping were more likely to change their Trofile result to Dual/Mixed or X4 during the trials, suggesting earlier non-R5 detection by deep sequencing versus the original Trofile assay. Of maraviroc recipients who experienced tropism switches, genotyping would have identified 40% as non-R5. Thus, the high sensitivity of deep sequencing was able to account for a substantial portion of tropism switches as being due to the presence of low-level non-R5 variants that were not detected by Trofile.

Figure 3.14: Similar Predictions of Plasma Viral Load Changes on Maraviroc by Deep Sequencing & the Enhanced Sensitivity Trofile Assay

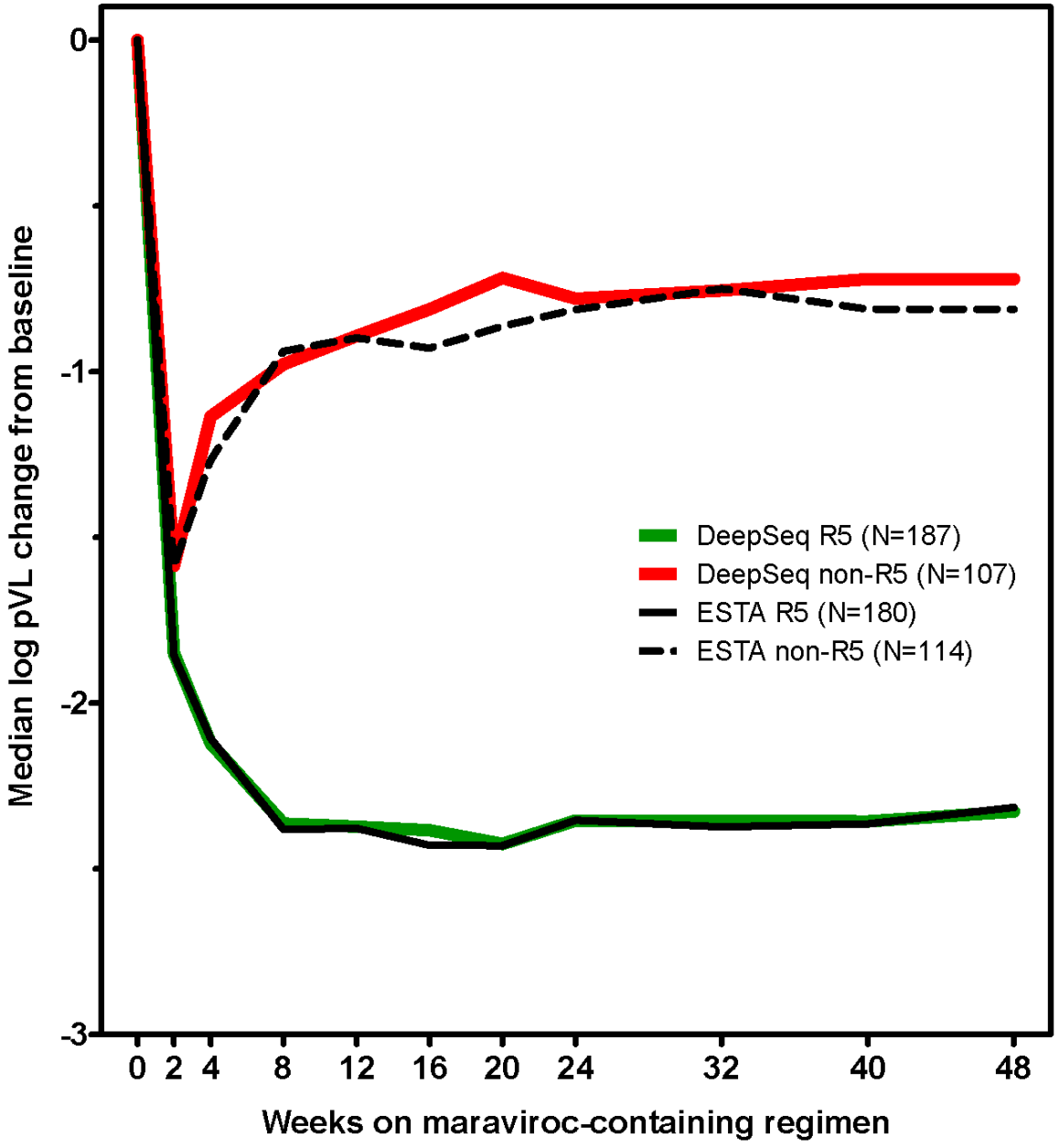


Figure 3.14: Similar Predictions of Plasma Viral Load Changes on Maraviroc by Deep Sequencing & the Enhanced Sensitivity Trofile Assay. The median viral load changes from baseline are shown for maraviroc recipients, stratified by their deep sequencing genotype results or Enhanced Sensitivity Trofile Assay phenotype results. The median viral load change for patients classified as having R5 HIV by deep sequencing is shown as a green line, and the patients with R5 by ESTA are shown with a solid black line. Patients with non-R5 results by genotype or phenotype are shown with red lines or dashed black lines, respectively.

Figure 3.15: Intermediate Plasma Viral Load Changes Where Deep Sequencing & the Enhanced Sensitivity Trofile Assay Gave Discordant Results

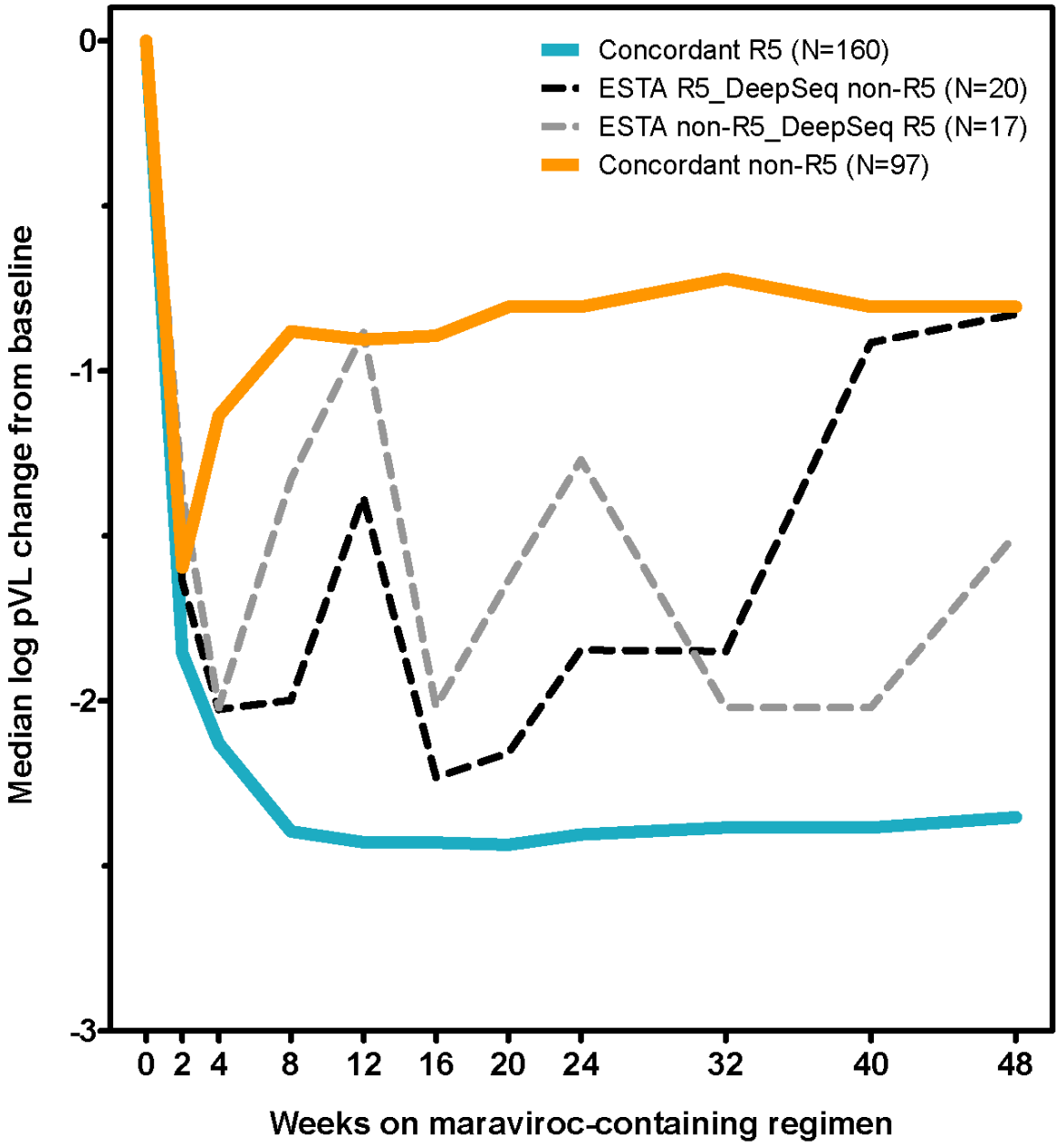


Figure 3.15: Intermediate Plasma Viral Load Changes Where Deep Sequencing & the Enhanced Sensitivity Trofile Assay Gave Discordant Results. Virologic responses to maraviroc for patients stratified by whether deep sequencing and the Enhanced Sensitivity Trofile assay (ESTA) gave concordant or discordant results. The turquoise line represents patients with concordant R5 results (N=160); the orange line represents concordant non-R5 results (N=97). The dashed black line represents patients where ESTA indicated R5 but deep sequencing indicated non-R5 (N=20), and the dashed grey line represents patients with non-R5 by ESTA but R5 by deep sequencing (N=17). Where the assays gave discordant results, neither assay seemed to have consistently superior performance to the other.

The emergence of minority non-R5 variants detected by deep sequencing but not Trofile has been shown previously following treatment with maraviroc [232]. In the MERIT trial of maraviroc in treatment-naïve patients, low X4 sensitivity of the original Trofile assay was determined to be a primary reason that maraviroc demonstrated inferiority to efavirenz [290]. When this trial was retrospectively re-analysed with ESTA, more patients were identified as harbouring non-R5 virus [290].

A potential benefit of deep sequencing over standard population-based sequencing is that the latter cannot reliably detect variants present below approximately 20% of the viral population [180,228] , while deep sequencing can reliably detect quasi-species present at much lower levels, as shown here. Only 37% of maraviroc recipients with 2-20% non-R5 variants were identified by population-based V3 sequencing (and only 58% were identified by Trofile), yet these patients still showed suboptimal virologic responses. Thus, like ESTA, deep sequencing may represent an enhanced sensitivity tropism test, able to detect minority non-R5 variants. However, the added clinical benefit of capturing low-prevalence non-R5 variants should be weighed against the accessibility and relative affordability of standard sequencing. Importantly, population-based sequencing had >80% concordance with deep sequencing in this same dataset.

Most patients called R5 by genotyping still had low levels (<2%) of detectable non-R5 virus. Despite this, good virologic responses to maraviroc were seen in this population. It is possible that the background antiretrovirals were able to suppress the non-R5 variants in these patients. Alternatively, it may be the case that a minimum threshold of non-R5 HIV must be surpassed before treatment with CCR5 antagonists is compromised. The data reported here may indicate that this threshold is approximately 2% of the viral population. The activity of the background regimen is also likely a major factor.

Some limitations of this study and the deep sequencing method in general should be acknowledged. A major limitation is the relatively high cost of deep sequencing. Also, the labour in preparing samples for pyrosequencing is intensive and complex. Third, there may be correlates of tropism

outside V3 [218,342,343], which this method is unable to capture. This study was retrospective in nature, and a randomized clinical trial with exclusion of patients screened non-R5 by deep sequencing instead of Trofile may have yielded different results. However, the results in an unbiased subset of patients presented in Sections 3.3.8 and 3.3.9 suggest that the performance of deep sequencing remained excellent even without pre-screening.

Surprisingly, there are limited ESTA data available for these studies. This is a concern, since maraviroc is primarily prescribed for treatment-experienced patients yet ESTA, the most commonly used tropism assay, has not been formally validated in MOTIVATE. This is a limitation for the interpretation of the results of the current study, however this and other studies have found good concordance between these methods [340,341]. The pre-screening of patients with Trofile also limited the number of treated non-R5 patients examined in this study. These limitations coupled with the good performance of the above detailed method support a prospective trial evaluating the relative merits of genotypic and phenotypic approaches. Such a trial would also establish the true sensitivities and specificities of both approaches [344].

Overall, despite the study's limitations, deep sequencing showed good performance in predicting a variety of clinical parameters including viral load declines, likelihood of achieving virologic suppression, and time to a tropism switch. This large study establishes deep sequence analysis of the HIV envelope V3 loop as an extremely promising tool for identifying treatment-experienced individuals who could receive clinical benefit from CCR5 antagonist-containing therapy regimens.

The above chapter exclusively examined next-generation sequencing using RNA amplified from blood plasma of HIV-infected patients. In a number of settings, however, current HIV RNA is not available due to the success of antiretroviral therapy at suppressing plasma viraemia [345]. HIV exists in the cells of its host as both integrated and cell-associated (e.g., episomal) DNA. Therefore, there are efforts aimed at evaluating the utility of cellular HIV DNA for determining HIV coreceptor usage [214,344,346,347]. While this approach was introduced briefly in Chapter 2, it is expanded on and examined more in-depth in the following chapter.

Chapter 4: Use of Cellular HIV DNA to Predict Virologic Responses to Maraviroc: Performance of Population-Based & Deep Sequencing

4.1 Background & Introduction

Recent advances in HIV treatment and curative strategies have led to the need for sensitive and accurate HIV tropism assays. The CCR5 antagonist antiretroviral drug class, including maraviroc [288] and others [111,318], are most successful when used in patients with solely CCR5-using (R5) HIV, as determined with phenotypic [290,291] or genotypic tropism assays [325,348,349]. Additionally, fledgling attempts at establishing long-term remission from HIV disease have been developed [350,351], such as using zinc-finger nucleases for disruption of the CCR5 gene. These are related to the successful cure by stem cell transplantation from a *CCR5 Δ32*-homozygous [156] donor to an HIV-infected patient [352]. These curative approaches will probably require pre-screening with tropism assays to identify candidate patients with exclusively R5 HIV. This is because a viral population which uses CXCR4 would likely be unaffected by reducing CCR5 protein levels.

Sustained suppression of plasma viraemia with advances in antiretroviral therapy improves patient outcomes [345] but precludes resistance and/or tropism testing from plasma due to low HIV copy numbers. Given the impracticalities and clinical consequences of treatment interruptions [353], tropism testing from HIV RNA in successfully treated patients is not possible. Still, some patients may wish to incorporate maraviroc into an already successful antiretroviral regimen to manage side effects or simplify the regimen. For such patients, a more feasible approach may be tropism testing from HIV DNA.

HIV DNA is the product of successful infection of cells by HIV [214,333]. Tropism testing from HIV DNA involves PCR amplification of a portion of the envelope gene followed by phenotypic testing in cell lines or genotypic testing by sequencing. Generally, the V3 loop of HIV gp120 is the main target of

such approaches. Tropism testing from HIV DNA can allow patients to switch a component of their antiretroviral regimen to a CCR5 antagonist without an interruption to their existing treatment.

The aim of this study was to assess the performance of cell-based genotypic tropism testing approaches in a large group of patients entering three clinical trials of the CCR5 antagonist, maraviroc. All patients were viraemic at the baseline testing visit, allowing parallel testing and comparison of plasma and cell-based tropism assays, as well as examination of actual virologic outcomes to the medication. Both population-based and deep sequencing approaches were applied in both compartments, giving a total of four different genotypic tropism tests. The abilities of these four methods to predict subsequent virologic response to maraviroc were compared with each other, and also with the phenotypic, plasma-based original Trofile assay at the same time point.

4.2 Materials & Methods

4.2.1 Samples & Patient Composition

Peripheral blood mononuclear cell (PBMC) samples were obtained at baseline from 181 maraviroc recipients in the MOTIVATE-1 (N=48), MOTIVATE-2 (N=48) and A4001029 (N=85) studies [288,291]. Note that these samples were deliberately selected to include a large proportion of non-R5 Trofile results (N=89, 49%). An approximate 1:1 ratio of R5 to non-R5 Trofile results mitigated any population skewing through over-enrichment for CCR5-tropic samples — a criticism of past studies [354].

A total of 156 (86%) had matching tropism results from plasma available. The baseline time point was day 0 of treatment with maraviroc. All participants were antiretroviral therapy-experienced and received at least one dose of maraviroc (once or twice daily) plus an optimized background therapy during the trials. Patients were screened and their plasma samples were periodically tested while on-treatment using the original Trofile assay. Results using the enhanced sensitivity Trofile assay (ESTA) or DNA-based Trofile assay were not available.

4.2.2 V3 Amplification & Sequencing

V3 amplification and sequencing were performed by similar methods to those previously published and reported in the above chapters [348,349]. Briefly, triplicate nested RT-PCR was used to amplify the V3 region from HIV RNA in plasma. For HIV DNA, 500 μL of PBMC samples were extracted with automated methods, followed by triplicate nested PCR targeting V3. Deep sequencing with a Roche/454 Life Sciences Genome Sequencer FLX was performed using the second-round PCR products, which had multiplex tags, allowing 48 samples to be sequenced in each direction per run.

The median read depths obtained were 2799 reads per DNA sample (interquartile range [IQR]: 2057 – 3623) and 2088 (IQR: 1783 – 2579) reads per RNA sample. A description of the data processing pipeline for deep sequencing is included in Appendix IV. In addition, a second round PCR amplification was also performed using the same triplicate amplified template. These PCR products underwent standard, population-based sequencing on an ABI 3730 XL DNA analyzer according to previously described methods [325].

4.2.3 Tropism Prediction

The geno2pheno algorithm generates a false positive rate (FPR) for each input sequence [204]. Those scoring above a certain pre-selected cutoff are classified as R5. The false positive rate (FPR) cutoff for geno2pheno tropism assignments was set previously [325,336,348]. Optimization of these cutoffs was performed using a random 75% of plasma screening samples from the maraviroc treatment-experienced trials, and was validated on the remaining 25%. Cutoffs were chosen in order to distinguish maximal differences between early response and non-response to maraviroc at week 8 of treatment. The maximum percentage of non-R5 variants allowed for a sample to be classified as R5 was also optimized and validated in a similar manner [335].

A sample was considered R5 if the lowest of three population-based V3 sequences had a geno2pheno FPR cutoff greater than 5.75 [325]. For deep sequencing, a sample was considered R5 if fewer than 2% of the variants detected fell below an FPR of 3.5 [348]. Population-based sequencing required a higher FPR

cutoff than deep sequencing likely due to the reduced sensitivity of population-based sequencing to detect minority variants. There was an additional exploration of alternative geno2pheno cutoffs in the current study.

4.2.4 Ethics Statement

Written, informed consent was obtained from all individuals, including consent to allow other tropism testing to be performed on their samples. The University of British Columbia-Providence Health Care Research Ethics Board reviewed the research project and granted ethical approval. All data were analyzed anonymously.

4.2.5 Data Analysis

Patients were grouped according to the R5 or non-R5 result by each tropism assay. Concordance was calculated from the number of samples with identical tropism calls by any two assays. The period from baseline to week 24 was examined for all patients. Each assay was assessed in its ability to predict responses to maraviroc plus optimized background therapy. Patients classified as having R5 HIV would be expected to have larger virologic responses to maraviroc compared to patients classified as having non-R5 HIV.

Within compartments (plasma or PBMCs), data were restricted to samples with results by all available assays; this was 181 PBMC samples and 156 plasma samples. When the two compartments were compared, analyses were restricted to the 156 samples with results by all five assays. Tests for statistical significance included the Mann-Whitney U test for comparisons of median pVL declines on maraviroc, the Fisher's exact test for comparisons of the proportion of patients achieving virologic responses at weeks 8 and 24, and the log-rank test for differences in median time to change in phenotypic tropism.

4.3 Results

4.3.1 Prediction of Virologic Efficacy on Maraviroc

Matched plasma and PBMC baseline samples were assessed by two genotypic methods and the original Trofile assay in plasma, giving a total of five tropism assays to compare. Both the short-term (to 8 weeks) and long-term (to 48 weeks) virologic efficacy of maraviroc in patients were assessed as primary analyses. Patients were deemed to be correctly classified as having R5 HIV if they were virologic responders to maraviroc-based therapy. Patients were stratified by whether they had R5 or non-R5 results by each of the five assays at baseline. Generally, these five tropism methods were all similarly predictive of virologic responses to the study medication, regardless of the specific approach or compartment (Figure 4.1).

Short-term virologic responses by week 8 were examined, as in a previous study of deep sequencing [348]. A response to maraviroc-based therapy was defined as a plasma viral load decline ≥ 2 $\log_{10}$ copies/mL from baseline to week 8, or having an undetectable viral load at week 8. Odds ratios of success for groups identified as R5 versus non-R5 ranged from 3.2 for deep sequencing in PBMCs to 9.4 for deep sequencing in plasma.

At week 24, the percentage of patients with undetectable viraemia ranged from 42-47% amongst the R5 groups, which was 16-22% higher than the non-R5 groups. By all assays, patients with R5 HIV had approximately 2 times the odds of virologic suppression at week 24 compared to those with non-R5. Virologic success of R5 groups was statistically significantly higher than non-R5 groups by all assays at week 8. These groups also had significantly different virologic suppression by week 24, as classified by all assays ($p<0.05$), except the Trofile assay (which had a trend towards significance) (Table 4.1).

Figure 4.1: Percentage of Patients with Plasma Viral Loads below 50 Copies/mL Was Similar by All Cellular or Plasma-Based Methods

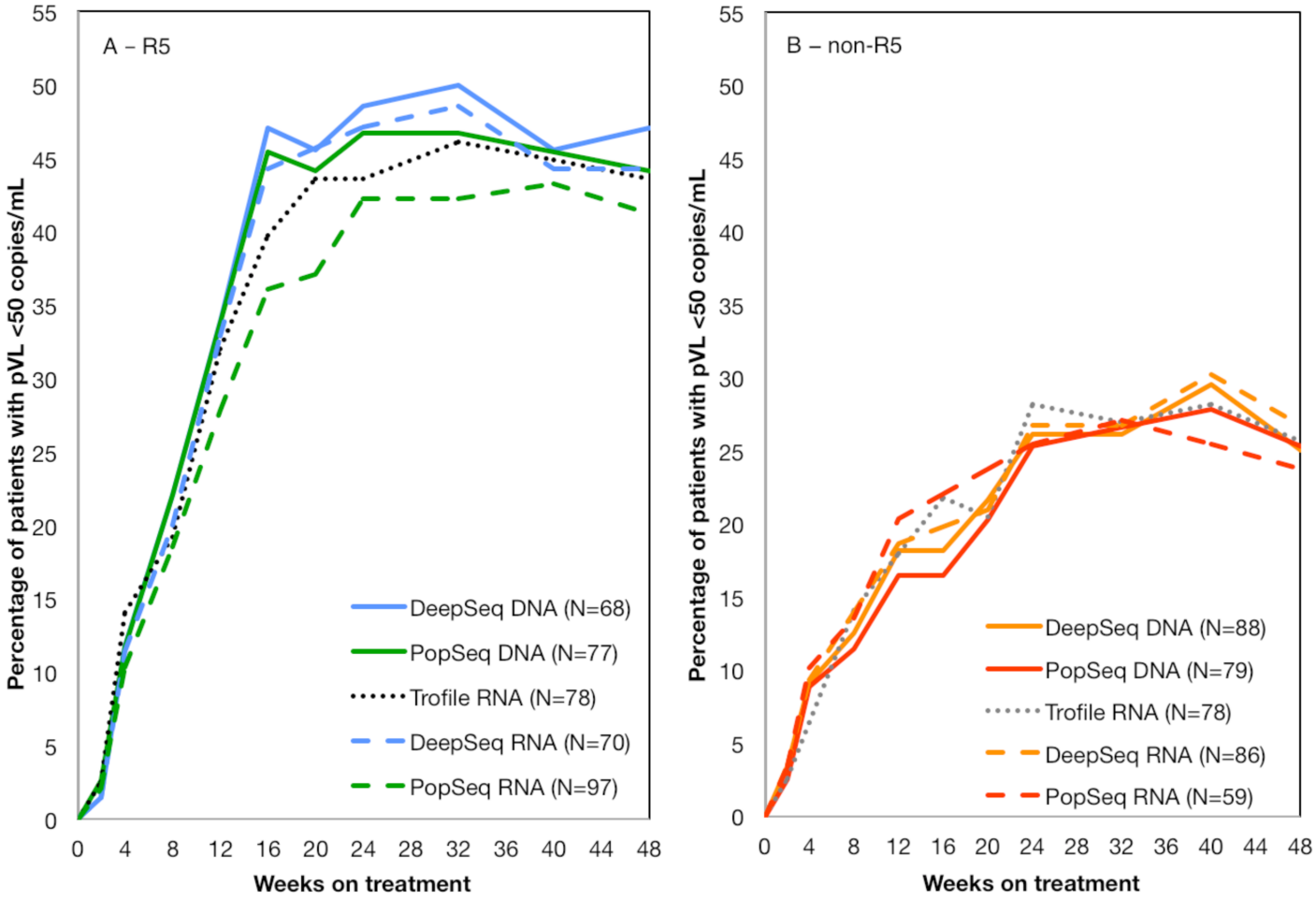


Figure 4.1: Percentage of Patients with Plasma Viral Loads below 50 Copies/mL Was Similar by All Cellular or Plasma-Based Methods. Panel A shows the percentage of patients classified with R5 HIV at baseline who had undetectable plasma viral loads over 24 weeks of treatment with maraviroc. Panel B shows the same but for patients classified with non-R5 HIV. Results are restricted for all groups to the 156 patients with baseline tropism results by all five methods. Solid lines represent the DNA-based assays, dashed lines represent the RNA-based assays, and the Trofile assay results are represented with dotted lines. DeepSeq — deep sequencing; PopSeq — population-based sequencing; pVL — plasma viral load

Table 4.1: Short- & Long-Term Virologic Responses to Maraviroc as Predicted by All Tropism Assays in the Full Dataset

	Log decline in pVL from baseline to week 8, median (IQR)		p-value	Percentage of patients with virologic response at week 8, % (n/N)		p-value	Percentage of patients with pVL <50 copies at week 24, % (n/N)		p-value	Odds ratio of week 8 success (R5 vs non-R5)	Odds ratio of week 24 suppression (R5 vs non-R5)
	R5 group	Non-R5 group		R5 group	Non-R5 group		R5 group	Non-R5 group			
PopSeq (DNA)	2.4 (1.8 – 2.7)	1.1 (0.2 – 2.7)	<0.01	83% (64/77)	52% (41/79)	<0.001	47% (36/77)	25% (20/79)	<0.01	4.6	2.6
DeepSeq (DNA)	2.4 (1.5 – 2.8)	1.7 (0.3 – 2.7)	0.01	81% (55/68)	57% (50/88)	<0.01	47% (32/68)	26% (23/88)	0.01	3.2	2.5
Trofile (RNA)	2.4 (1.9 – 2.8)	1.1 (0.2 – 2.6)	<0.001	83% (65/78)	51% (40/78)	<0.01	44% (34/78)	28% (22/78)	0.07	4.8	2.0
PopSeq (RNA)	2.3 (1.4 – 2.7)	1.0 (0.2 – 2.7)	0.01	77% (75/97)	51% (30/59)	<0.01	42% (41/97)	25% (15/59)	<0.05	3.3	2.1
DeepSeq (RNA)	2.4 (2.1 – 2.8)	1.0 (0.2 – 2.6)	<0.001	90% (63/70)	49% (42/86)	<0.01	47% (33/70)	27% (23/86)	<0.05	9.4	2.4

Table 4.1: Short- & Long-Term Virologic Responses to Maraviroc as Predicted by All Tropism Assays in the Full Dataset. The plasma viral load (pVL) responses to maraviroc plus OBT are shown for patients classified as having R5 and non-R5 HIV by all 5 tropism methods. The median pVL decline from baseline to week 8 is shown as well as the percentage of patients with a week 8 virologic response — defined as a pVL decline $\geq 2 \log_{10}$ or a pVL of <50 copies/mL at week 8. The percentage of patients with undetectable viraemia is also shown. The odds ratios of week 8 and week 48 virologic responses are also displayed. This table is restricted to the 156 samples which had results for all assays.

4.3.2 Similar Performance Regardless of Background Regimen Activity

These analyses were re-examined in a subset of 81 patients who had compromised treatment background activity, and for whom maraviroc would be expected to have the largest impact. This approach can best distinguish differences in assay performance by minimizing the contribution of background antiretroviral agents to the activity of maraviroc.

These patients, who had a weighted optimized background therapy susceptibility score (wOBTss) ≤1, had similar responses to the overall study group. In fact, the difference between the R5 and non-R5 groups was exaggerated in this subset. The patients classified as having R5 HIV by any of the tropism assays had much higher responses to maraviroc than did those with non-R5 HIV classifications (Table 4.2).

4.3.3 Prediction of Future Tropism Changes on Maraviroc

Virologic failure on maraviroc is often accompanied by a "switch" in tropism from R5 to non-R5 [289]. Changes in phenotypic tropism by Trofile were also examined over the course of the trial. Analyses were restricted to those patients who had R5 Trofile results at both screening and baseline, leaving a total of 84 patients with DNA results and 71 with RNA results. Patients with non-R5 results by the genotypic assays were significantly more likely to have subsequent Dual/Mixed or X4 results by Trofile in plasma over the course of the study (Figure 4.2).

4.3.4 Diagnostic Performance of Tropism Assays

The performance characteristics of all four sequencing approaches were compared against the phenotypic original Trofile assay. Concordance of DNA-based tropism testing in cells was assessed relative to tropism classifications by the Trofile assay in the matching plasma sample.

Table 4.2: Short- & Long-Term Virologic Responses to Maraviroc in Patients with Compromised Background Regimens

wOBTss ≤1	Log decline in pVL from baseline to week 8, median (IQR)		p-value	Percentage of patients with virologic response at week 8, % (n/N)		p-value	Percentage of patients with pVL <50 copies at week 24, % (n/N)		p-value	Odds ratio of week 8 success (R5 vs non-R5)	Odds ratio of week 24 suppression (R5 vs non-R5)
	R5 group	Non-R5 group		R5 group	Non-R5 group		R5 group	Non-R5 group			
PopSeq (DNA)	2.3 (1.2 – 2.7)	0.4 (+0.2 – 2.3)	<0.01	77% (34/44)	32% (12/37)	<0.01	41% (18/44)	14% (5/37)	<0.05	7.1	4.4
DeepSeq (DNA)	2.3 (1.1 – 2.7)	0.5 (+0.2 – 2.5)	<0.01	77% (30/39)	38% (16/42)	<0.01	41% (16/39)	17% (7/42)	<0.05	5.4	3.5
Trofile (RNA)	2.3 (1.0 – 2.8)	0.4 (+0.2 – 1.9)	<0.01	75% (33/44)	35% (13/37)	<0.01	39% (17/44)	16% (6/37)	<0.05	5.5	3.3
PopSeq (RNA)	2.2 (0.9 – 2.7)	0.3 (+0.2 – 0.9)	<0.01	70% (40/57)	25% (6/24)	<0.01	37% (21/57)	8% (2/24)	<0.05	7.1	6.4
DeepSeq (RNA)	2.3 (1.6 – 2.8)	0.4 (+0.2 – 1.5)	<0.01	85% (34/40)	29% (12/41)	<0.01	45% (18/40)	12% (5/41)	<0.05	13.7	5.9

Table 4.2: Short- & Long-Term Virologic Responses to Maraviroc in Patients with Compromised Background Regimens. This table shows a subset of patients (N=81) with a weighted optimized background sensitivity score (wOBTss) ≤1. The plasma viral load (pVL) responses to maraviroc plus OBT are shown for patients classified as having R5 and non-R5 HIV by all 5 tropism methods. The median pVL decline from baseline to week 8 is shown as well as the percentage of patients with a week 8 virologic response — defined as a pVL decline ≥2 $\log_{10}$ or a pVL of <50 copies/mL at week 8. The percentage of patients with undetectable viraemia is also shown. The odds ratios of week 8 and week 48 virologic responses are also displayed.

Figure 4.2: All Genotypic Tropism Testing Methods Predicted Future Phenotypic Tropism Changes While Receiving Maraviroc

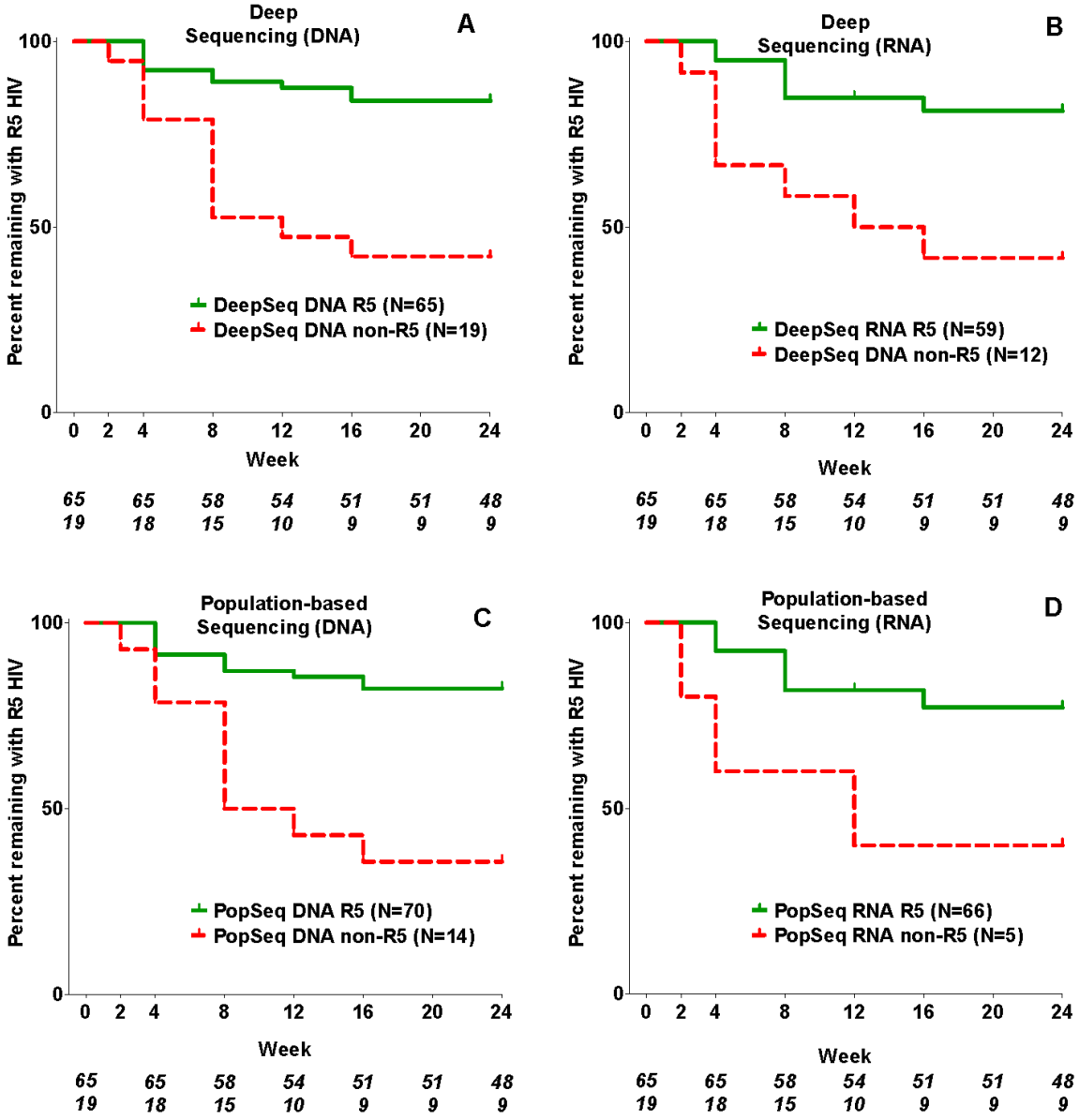


Figure 4.2: All Genotypic Tropism Testing Methods Predicted Future Phenotypic Tropism Changes While Receiving Maraviroc. Kaplan Meier plots show the percentage of patients who change phenotypic tropism status by Trofile over the course of the study. Analysis is restricted to patients with R5 Trofile results at both screening and baseline. Patients are stratified by their baseline tropism results by each of the four genotypic assays, with the R5 groups in green and the non-R5 groups in red. Panels A and B depict deep sequencing results in the PBMC and plasma compartments, respectively. Panels C and D depict the population-based sequencing results. The p-values by the log-rank test for the differences in median time to tropism change between groups were all <0.05. Ticks represent censored observations. pVL — plasma viral load

Genotypic DNA tropism testing by population-based sequencing had 80% concordance with Trofile; DNA-based deep sequencing had 77% concordance. The corresponding RNA-based approaches had concordance with Trofile of 80% and 86%, respectively. The sensitivity, specificity, and concordance of all approaches are shown in Table 4.3.

Table 4.3: Performance of Sequencing-Based Tropism Assays from Peripheral Blood Mononuclear Cells & Plasma Relative to the Original Trofile Assay in Plasma

Method	Sensitivity, % (n/N)	Specificity, % (n/N)	Concordance, % (n/N)
PopSeq DNA	78% (69/89)	83% (76/92)	80% (145/181)
DeepSeq DNA	81% (72/89)	74% (68/92)	77% (140/181)
PopSeq RNA	68% (53/78)	92% (72/78)	80% (125/156)
DeepSeq RNA	91% (71/78)	81% (63/78)	86% (134/156)

Table 4.3: Performance of Sequencing-Based Tropism Assays from Peripheral Blood Mononuclear Cells & Plasma Relative to the Original Trofile Assay in Plasma. The sensitivity (percentage of correct non-R5 results), specificity (percentage of correct R5 results), and concordance (overall percentage of correct results) are shown for all sequencing approaches relative to the original Trofile assay (not ESTA). PopSeq — Population-based Sequencing; DeepSeq — Deep Sequencing.

4.3.5 Compartmental Differences

Since plasma approaches have been more thoroughly examined in the literature, the ability of the DNA-based approaches were compared against their corresponding RNA-based approaches. Relative to RNA, population-based and deep sequencing from DNA had sensitivities of 88% and 86%, and specificities of 72% and 80%. Overall rates of concordance between plasma and PBMCs were 78% using population-based sequencing and 83% using deep sequencing.

To address the possibility that the bioinformatic algorithm cutoffs previously optimized for plasma were not optimized in the cellular compartment, exploratory analyses of additional geno2pheno cutoffs were undertaken. The virologic responses of patients were evaluated using a range of FPR cutoffs: 2, 3.5, 5.75, 10, 20, and 50, in a group of 93 patients with a weighted optimized background therapy susceptibility score <1 (i.e., fewer than 1 drug in addition to maraviroc in their background regimens). This analysis again confirmed that a cutoff in the range of approximately 5.75 to 10 was able to distinguish the largest difference in week 8 pVL declines between tropism groups (Figure 4.3). Thus, the poorer performance of tropism testing in the cellular compartment was likely not an issue with bioinformatic cutoffs having been optimized in a plasma-based context (Figure 4.3).

4.3.6 Virologic Responses with Screening by DNA-Based versus RNA-Based Approaches

For each sequencing approach, the relative performance of the plasma or PBMC predictions was assessed. The plasma compartment tended to outperform the PBMC compartment in its predictability (Table 4.1). For both deep and population-based sequencing, patients identified as having R5 HIV in both compartments had virologic declines on maraviroc of approximately 2.5 $\log_{10}$ copies/mL by week 24. In contrast, where both compartments indicated non-R5 HIV, the median viral load declines were approximately 1 $\log_{10}$.

At week 24, the median pVL decline was ~1.5 $\log_{10}$ where plasma indicated R5 but PBMCs indicated non-R5, while the median decline was ~0.5 $\log_{10}$ where plasma indicated non-R5 but PBMCs indicated R5 HIV. This suggests that testing from the plasma compartment was able to correctly identify more patients as maraviroc responders or non-responders than testing from the cellular compartment (Figure 4.4).

Figure 4.3: Effect of Geno2pheno Cutoffs on Prediction of Response to Maraviroc in Patients with Compromised Background Regimens

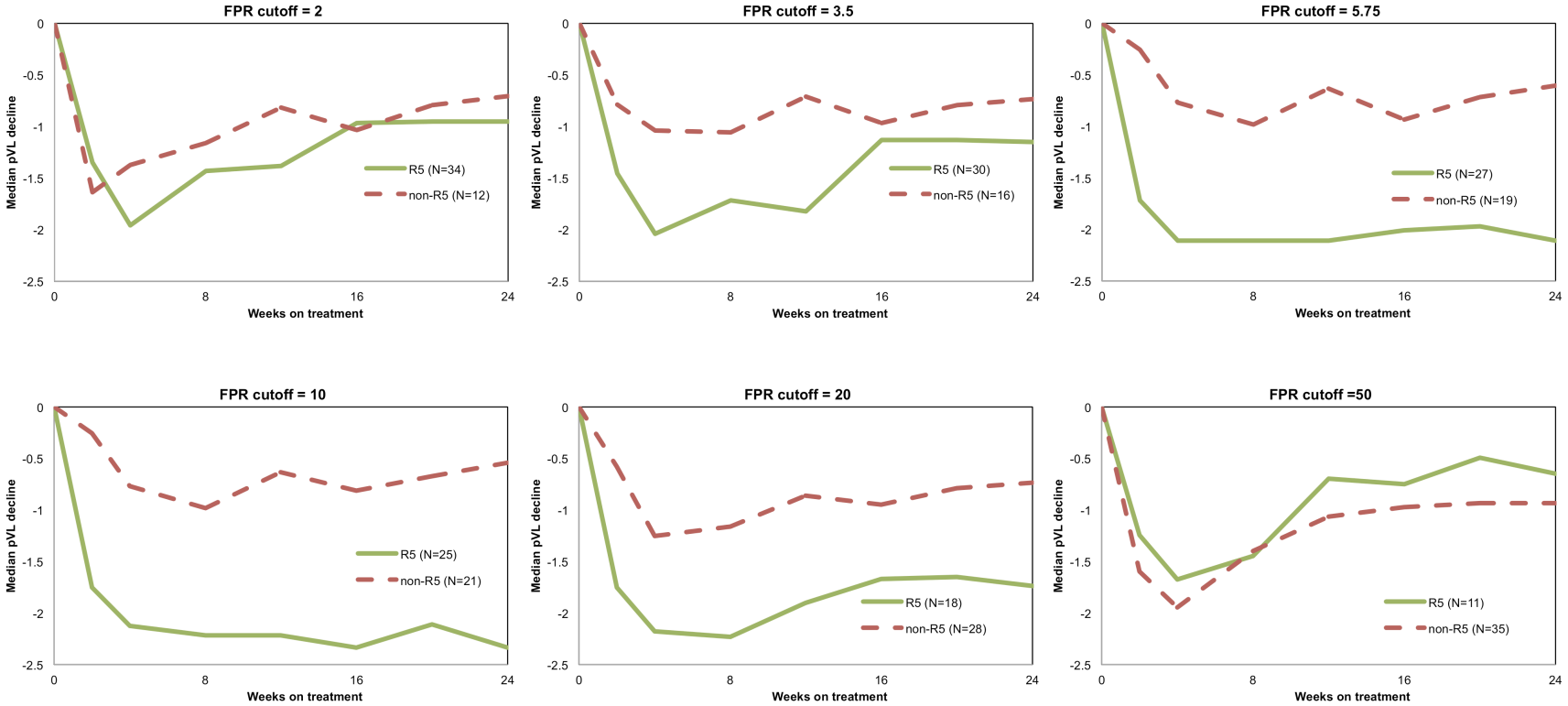


Figure 4.3: Effect of Geno2pheno Cutoffs on Prediction of Response to Maraviroc in Patients with Compromised Background Regimens. Median pVL decline on maraviroc in patients with compromised background regimens, defined as a weighted Optimized Background Susceptibility Score (wOBTss) <1. Population-based sequencing results were interpreted using FPR cutoffs ranging from 2 to 50. Patients with FPRs less than or equal to each respective cutoff are classified as having non-R5 HIV and are indicated with the dashed red lines. Patients with FPRs above the cutoff are classified as having R5 HIV and are indicated with the solid green lines. A geno2pheno FPR cutoff of 5.75 – 10 seems to best discriminate between responders and non-responders to maraviroc. Note that a cutoff of 3.5 is optimized for deep sequencing and not population-based sequencing.

Figure 4.4: Patients with Discordant Tropism Results Between Compartments Had Virologic Responses Which Favoured the Plasma Prediction

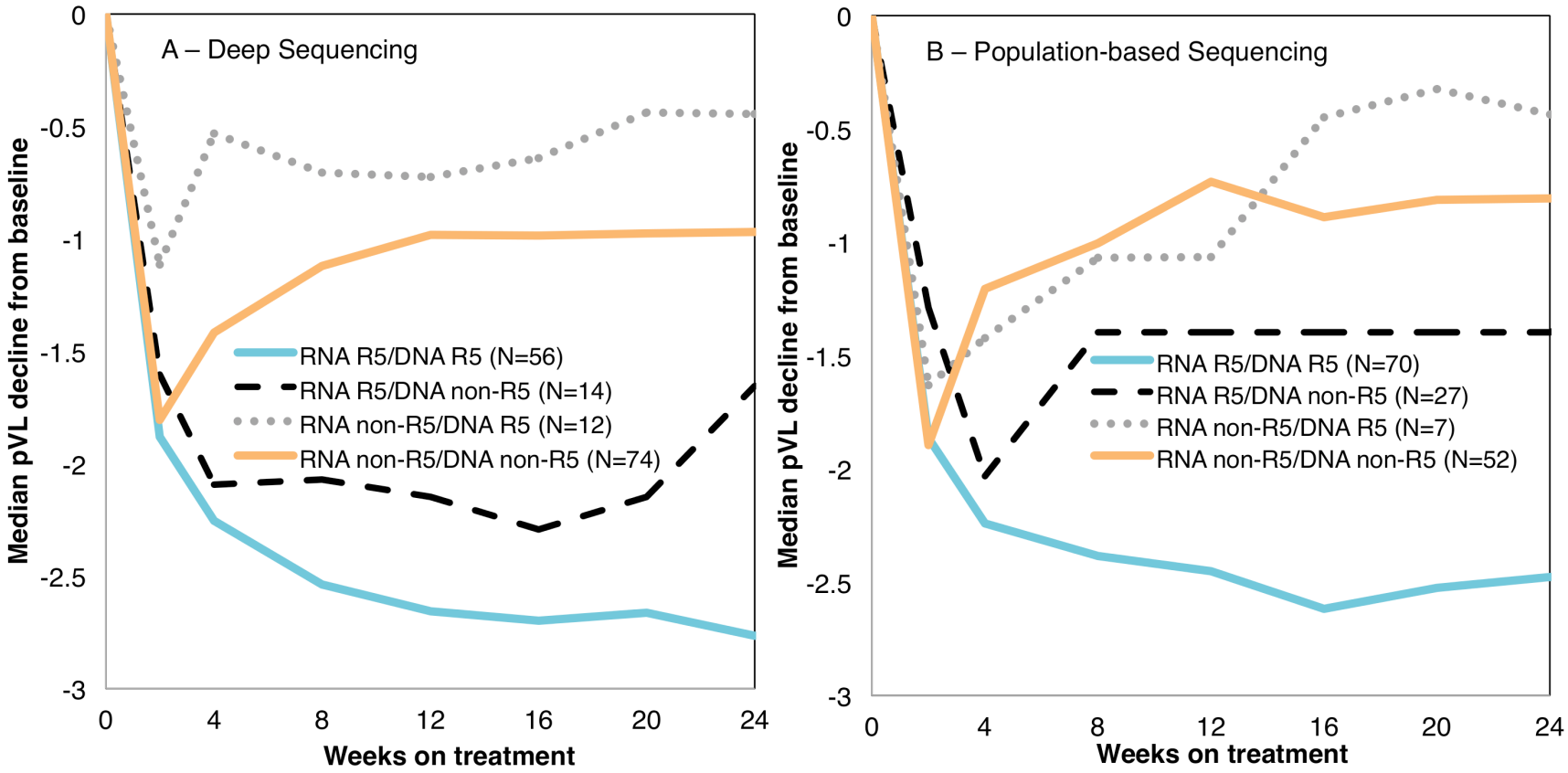


Figure 4.4: Patients with Discordant Tropism Results Between Compartments Had Virologic Responses which Favoured the Plasma Prediction. Panel A shows the deep sequencing results and Panel B shows the population-based sequencing results. Solid blue lines indicate the groups where RNA and DNA methods both indicated R5. Solid orange lines indicate the groups where RNA and DNA both indicated non-R5. Discordant groups are indicated by black dashed (RNA R5 and DNA non-R5), and grey dotted (RNA non-R5 and DNA R5) lines. pVL — plasma viral load

4.4 Discussion & Conclusions

This study examined the performance of tropism classifications from the cellular compartment in a large number of patients initiating maraviroc-based therapy. Two independent sequencing approaches from both the plasma and peripheral blood mononuclear cell compartments were tested against each other and against the plasma-based phenotypic original Trofile assay. While performance was fairly good in PBMCs, tropism predictions from the plasma compartment tended to outperform the DNA-based methods. Where results were discordant, longer-term virologic suppression was not predicted as well by DNA-based methods, suggesting some misclassification by DNA versus RNA approaches.

Nevertheless, this approach may be the only option for some patients, barring deliberate but inadvisable treatment interruptions [355] to raise viraemia to levels needed for RNA-based tropism testing. Additionally, the reasonably high concordance (~80%) between the plasma and cellular compartments should give some confidence that DNA-based approaches give useful clinical information. The use of DNA-based tropism testing is suggested in European guidelines [276], but better guidance will likely result from an ongoing clinical trial of DNA-based tropism testing in patients with suppressed viraemia [344].

This study confirms and expands on the recent results from Vitiello and colleagues, who examined a group of 20 patients switching a component of their antiretroviral regimens to maraviroc while virologically suppressed [356]. Both studies found that DNA tropism testing could be used to predict successful treatment with maraviroc, arguably the most clinically relevant outcome of a tropism test. Compared to past reports, this study found roughly similar, if slightly worse, diagnostic accuracy of cell-based genotypic tropism assays relative to phenotypic and plasma-based approaches [231,347,357,358].

Aside from the ability to assess diagnostic performance, the primary advantage of the current study is that the tropism classifications could be additionally evaluated for how well they predict virologic response to maraviroc in a real clinical setting. The ability to predict these responses acted as an

independent, objective confirmation of the diagnostic performance of these assays. An additional advantage of this study is the fact that paired plasma and PBMC fractions from the same blood draws could be compared for their ability to predict virologic response, whereas other studies have tended to compare later PBMC results with earlier pre-suppression plasma results [283,347,359]. This study also has the advantage of the design of the clinical trials from which these patients were drawn. The inclusion of patients enrolled in A4001029 who had non-R5 results at baseline but were still prescribed maraviroc gives additional confirmation on the utility of these methods.

There are some difficulties inherent to tropism testing from HIV DNA. The cellular "buffy coat" fraction of whole blood is not routinely collected or stored, nor are peripheral blood mononuclear cells routinely separated for analysis. The cellular compartment has also been found to have higher sequence variation [211], and higher prevalence of CXCR4-tropic HIV [165] compared to blood serum. Importantly, overestimation of CXCR4-usage may actually increase the likelihood of success with CCR5 antagonists, since more patients may be screened out by DNA-based approaches. However, the current results are not definitive in their support for this hypothesis, and a prospective clinical trial using DNA tropism testing has yet to be completed.

Low input copy number may also be an issue for testing from the cellular compartment compared to the plasma [360]. Quantitation of HIV DNA and cell number were not performed in this study, so potentially low copy numbers may have contributed to performance issues. However, routine quantitation of cells or DNA copies represents a fairly significant barrier to the implementation, availability, and turn-around time of DNA-based tropism testing. Input copy number may be accounted for by the use of PCR "tags" accompanying each DNA strand amplified, as recent work has shown [361]. However, this technique was not available at the time of testing.

A major strength of this study was having access to paired plasma and PBMC samples from the same time point. This enabled direct comparisons between compartments. Although these patients were viraemic, this sample set is ideal for comparing DNA-based approaches to RNA-based ones, since there is much more clinical experience with RNA-based approaches. However, the fact that patients

did not have suppressed viraemia at baseline requires extrapolation of these results to patients with undetectable viral loads who may switch to maraviroc. This should be noted, since such patients ultimately comprise the target group for DNA-based tropism approaches. Cellular HIV DNA copy number may decay with effective antiretroviral therapy [362,363], since lower replication may reduce the pool of HIV DNA. This could lead to difficulties in DNA-based tropism testing in aviraemic patients. However this issue could not be addressed with this sample set because all patients had detectable viral loads.

The relatively small number of patients who had tropism classifications that were discordant between compartments (e.g., R5 in plasma but non-R5 in PBMCs) meant it was not possible to definitively state that plasma predicts maraviroc response better than the cellular compartment. Conversely, the small number of patients with discordant results also reflects the reassuring fact that a large majority of patients in fact had concordant results between the compartments. Another potential limitation of this study is the fact that background therapy also affects response to maraviroc in addition to HIV tropism [364]. However, this would presumably affect all assays equally in their ability to predict virologic response to maraviroc-based therapy, so it should not have greatly skewed the results.

Despite the above-mentioned caveats, this study demonstrates promising potential for DNA-based tropism methods. That the cell-based classifications were not as clinically predictive as plasma-based ones should add a measure of caution to the routine use of this approach. However, the DNA-based testing was still able to discriminate between responders and non-responders to maraviroc-containing regimens, and despite some shortcomings, may be the best course of action prior to prescribing maraviroc in patients with suppressed viraemia.

This, and the previous chapter examined the utility of next-generation sequencing in evaluating HIV tropism and using the results to predict subsequent virologic failure over long-term follow-up. These chapters examined the performance of next-generation sequencing and V3 genotyping prior to CCR5 antagonist treatment and assessed this performance during follow-up by employing results from other assays and virologic tests such as phenotypic assays and plasma viral load tests. The next

chapter applies genotypic tropism and next-generation sequencing methods in a more detailed context with long-term, longitudinal follow-up. Changes in the *env* sequence which are associated with treatment failure are examined and correlated with pre-treatment deep sequencing results. The following chapter thus explores the specific genotypic factors associated with the virologic failure which the previous chapters have demonstrated can be predicted by next-generation sequencing.

Chapter 5: Genotypic Analysis of the HIV-1 V3 Region in Virologic Non-Responders to Maraviroc-Containing Regimens Reveals Distinct Patterns of Failure

5.1 Background & Introduction

Successful antiretroviral treatment with the CCR5 antagonist, maraviroc, requires a tropism test to confirm that the patient's HIV uses the CCR5 coreceptor for cellular entry (R5 HIV) rather than CXCR4 (non-R5 HIV) [288,289,291]. In the Phase III clinical trials of maraviroc, patients were screened for tropism status using the original Trofile phenotypic coreceptor assay (OTA), which has subsequently been replaced by the enhanced sensitivity Trofile assay (ESTA) [290,365]. Recent re-screening of clinical trials of maraviroc has confirmed the utility of genotypic approaches for the determination of HIV tropism [325,340,348,349,366]. Such approaches typically involve sequencing of the third variable (V3) region of the HIV envelope gene [200].

Bioinformatic algorithms such as geno2pheno [204] are then used to infer the phenotypic tropism that is likely associated with a V3 genotype. Geno2pheno converts an input V3 sequence into an output value in the form of a false-positive rate (FPR), ranging from 0 to 100. An FPR indicates how likely a sequence is to be incorrectly identified as non-R5. Therefore, sequences yielding low false-positive rates have a high likelihood of being non-R5.

Historically, population-based sequencing has been the most commonly used genotypic approach for predicting coreceptor usage [200]. However, more sensitive tropism determination methods can more accurately predict response to maraviroc [290]; thus, newer deep sequencing methods targeting the V3 loop are becoming increasingly common [232,248,321,348,349,367]. These next-generation sequencing approaches can identify low-level non-R5 subpopulations in clinical samples. Following treatment

with maraviroc, these minority non-R5 quasispecies may emerge to much higher prevalence, thereby compromising treatment efficacy [232,316].

There are several pathways by which patients may fail a maraviroc-containing therapy regimen. Most commonly, a minority non-R5 population in a patient's HIV population may expand under drug pressure, causing an overall change in observed tropism [289]. Less commonly, the viral population may retain its CCR5 tropism while evolving the ability to use maraviroc-bound CCR5 protein for cellular entry — a form of maraviroc resistance [368]. Thirdly, the viral population may develop resistance to the other agents in the background regimen in the absence of a change in susceptibility to maraviroc [364]; this may be associated with either R5 or non-R5 tropism. Furthermore, as with other agents, adherence, absorption, and other patient-associated and pharmacokinetic factors can also lead to therapy failure.

Early detection of tropism shifts or maraviroc resistance can accelerate the decision to substitute maraviroc with another antiretroviral agent and potentially prevent further accumulation of antiretroviral drug resistance to other agents in the regimen. Thus, patients in this study were sampled relatively soon after beginning maraviroc treatment to determine the utility of an early monitoring approach.

This study uses both population-based and deep sequencing approaches to assess changes in tropism and V3 sequence among treatment-experienced, R5-infected patients who experienced virologic failure while receiving maraviroc in the MOTIVATE-1 and -2 studies [288,289]. Patients from the A4001029 study which enrolled patients with non-R5 HIV [291] were not included in the current study. Thus, all patients studied were determined to have exclusively R5 HIV by the original Trofile assay (OTA). They are therefore a representative population of patients most likely to receive maraviroc. Phylogenetic methods were also used to assess whether sequences present at failure were derived from pre-existing minority subpopulations, and next-generation sequencing was used to assess changes in non-R5 prevalence after treatment with maraviroc.

Previous studies [232,316] have noted emergence of CXCR4-using virus from pre-existing subpopulations, and CCR5 antagonists have been known to inhibit R5-only, while selecting non-R5 subpopulations [241,289], with such shifts appearing to occur very quickly [317]. Furthermore, resistance to maraviroc has been associated with genotypic changes in the HIV envelope gene [368,369]. Thus, it was hypothesized that there are distinct mechanisms of failure that can be identified by population-based and/or deep sequencing of the HIV V3 region.

5.2 Materials & Methods

5.2.1 Patient & Sample Selection

A subset of patients was selected who had suboptimal responses to maraviroc in the MOTIVATE trials (N=181). Patients were selected such that approximately the same proportion had non-R5 OTA results at failure as was reported for the MOTIVATE trials overall (57% in MOTIVATE, 58% in the current study) [289]. All patients were treatment-experienced, 100% of patients had R5 results by OTA at screening, and 69% had R5 results by ESTA (124/181). All received maraviroc (once or twice daily) plus an optimized background regimen of three to six other antiretroviral agents. All individuals gave written informed consent, including consent to allow other tropism testing to be performed on their samples. The University of British Columbia—Providence Health Care Research Ethics Board reviewed the research project and granted ethical approval.

Sequencing was performed on samples from two time points: one prior to receiving maraviroc (the screening sample) and one while receiving treatment (the failure sample). This on-treatment failure sample was defined as the earliest available sample with a plasma viral load (pVL) greater than 500 HIV RNA copies/mL, and an OTA result. The screening sample was drawn approximately 6 to 8 weeks prior to beginning maraviroc; the failure sample was drawn a median of 4 weeks after beginning maraviroc (interquartile range [IQR]: 4-16 weeks), and a median of 2 weeks (IQR: 2-10 weeks) after the first viral load ≥500 copies/mL. While phenotypic tropism results were available for all samples, phenotypic maraviroc resistance assay results were not available for these samples. ESTA results were available at screening, but only OTA results were available at failure.

5.2.2 Genotypic Tropism Testing

The third variable (V3) loop of the HIV envelope gene was amplified with nested RT-PCR. The screening samples were amplified and sequenced in triplicate; the failure samples had a single sequence generated per sample. Standard, population-based sequencing was performed on all screening and failure samples, as previously described [325]. Deep V3 sequencing was also performed on all screening samples, plus a subset (N=73) of failure samples, with methods as previously described [348,349]. The 73 samples comprised the last batch of samples processed through population sequencing, with no targeted selection.

The tropism associated with the V3 loop sequences was inferred using the geno2pheno algorithm [204] with false-positive rate (FPR) cutoffs of 5.75 for population-based sequencing and 3.5 for deep sequencing [336,337], below which sequences were categorized as non-R5. These cutoffs had previously been optimized for predicting virologic response to maraviroc [336,337]. The percentage of non-R5 variants in the viral population was defined as the proportion of sequences scoring below or equal to an FPR of 3.5 as observed by deep sequencing, and previous studies have defined an R5 sample as having <2% non-R5 variants [348,349].

The screening and failure sequences were assessed for amino acid changes that may have appeared following maraviroc-based therapy, as well as for a change in the geno2pheno FPR value. Neighbour-joining phylogenetic trees were constructed with ClustalX using deep sequencing data at screening and failure population-based or deep sequencing data. Thus, it could be determined whether a sequence present at failure may have already been present prior to treatment with maraviroc. The change in the percentage of non-R5 variants between screening and failure was also examined using the deep sequencing results. Sample phenotypes were obtained using OTA at all time points, and ESTA at screening.

5.2.3 Statistical Analyses

Statistical analyses performed included the Mann-Whitney U test and Kruskal-Wallis test for testing the statistical significance of differences between medians (e.g., median plasma viral loads). The Fisher exact test was used for testing the statistical significance of differences in proportions (e.g., the proportion of patients who were R5 by ESTA at screening).

5.3 Results

5.3.1 Patient & Sample Composition

Patients in the current study were all treatment-experienced and received maraviroc once daily (89 patients, 49%) or twice daily (92 patients, 51%), as per randomization at study entry. Most patients (91%) were enrolled in the North American MOTIVATE-1 trial [370], with the remaining 9% in the MOTIVATE-2 trial, which had an identical study design. Of the 181 patients selected, 100 (55%) experienced virologic failure; 44 (24%) never achieved virologic suppression but completed 48 weeks of treatment; and 15 (8%) had a virologic rebound. Of the remaining 22 patients, 18 were lost to follow up, two died, one experienced an adverse event, and one was withdrawn due to pregnancy.

The mean age of subjects was 45 (range: 19-70), and the proportion of males in the study was 91% (165/181). These were similar to the maraviroc arms of MOTIVATE overall [288]. The proportion of patients reporting Black race or ethnicity was 19% (34/181), which was slightly elevated relative to the larger trial overall (14%). This was likely due to a higher number of maraviroc non-responders who reported Black race/ethnicity in MOTIVATE [289] (Table 5.1). As expected for a study on patients who experience failure of therapy, the patients in the current study had higher plasma viral loads, lower CD4 cell counts, and fewer active drugs in their background regimens than those in the MOTIVATE studies [288,348] overall.

Table 5.1: Patient Characteristics in the Current Study Compared to the MOTIVATE Studies Overall

			Current study population			
		MOTIVATE 1 & 2 MVC arms overall	Overall study population	Patients with non-R5 population-based genotype at failure	Patients with R5 population-based genotype at failure	p-value R5 vs non-R5
N		794	181	91	90	-
Age – mean (range)		46 (17–75)*	45 (19–70)	45 (19–70)	45 (34–69)	n.s.
Sex:	Males – % (n/N)	89% (745/840)*	91% (165/181)	93% (85/91)	89% (80/90)	n.s.
	Females – % (n/N)	11% (95/840)*	9% (16/181)	7% (6/91)	11% (10/90)	
Race:	White – % (n/N)	83% (699/840)*	80% (145/181)	84% (76/91)	77% (69/90)	n.s.
	Black – % (n/N)	14% (121/840)*	19% (34/181)	15% (14/91)	22% (20/90)	
R5 by ESTA at screening – % (n/N)		N/A	69% (124/181)	48% (44/91)	89% (80/90)	<0.001
R5 by original Trofile at failure – % (n/N)		43% (57/133)*	42% (76/181)	2% (2/91)	82% (74/90)	<0.001
Baseline pVL (log)		4.9 (4.4 – 5.3)	5.1 (4.7 – 5.5)	5.0 (4.6 – 5.4)	5.2 (4.8 – 5.6)	0.03
Failure pVL (log)		N/A	4.1 (3.5 – 5.0)	4.3 (3.6 – 5.1)	4.0 (3.3 – 4.9)	0.06
Baseline CD4 count (cells/mm^3)		168 (74 – 289)	72 (17 - 177)	64 (15 - 174)	79 (23 - 182)	n.s.
Number of active ARVs (wOBTss)		1.0 (0 – 2)	0.5 (0 – 1.0)	0.5 (0 – 1.0)	1.0 (0 – 1.5)	0.03
Geno2pheno FPR at screening		33.2 (14.7 – 56.5)	31.1 (12.5 – 55.3)	19.6 (6.9 – 41.4)	41.9 (21.8 – 65.0)	<0.001
Geno2pheno FPR at failure		N/A	5.3 (1.1 – 48.9)	1.1 (0.4 – 1.8)	48.9 (21.2 – 74.0)	<0.001
Percent non-R5 variants by deep sequencing (screening, N=181)		0% (0% – 0.3%)	0.1% (0% – 3.1%)	1.4% (0% – 15.2%)	0% (0% – 0.1%)	<0.001
Percent non-R5 variants by deep sequencing (failure, N=73)		N/A	0.8% (0% – 99.0%)	99.5% (94.8% – 99.9%)	0% (0% – 0.2%)	<0.001

Table 5.1: Patient Characteristics in the Current Study Compared to the MOTIVATE Studies Overall. The baseline patient characteristics in the current study, as well as those from the maraviroc (MVC) arms of the MOTIVATE-1 and -2 studies. Most values shown are median values with the interquartile range (IQR) in parentheses, unless otherwise indicated. The MOTIVATE column was derived from a previously published dataset (Swenson et al, JID 2011 [348]) comprising a majority of maraviroc recipients in the MOTIVATE studies (94%, 788/840 patients). Due to low numbers of patients, those with race/ethnicities other than White or Black are not included in the table. pVL — plasma viral load; wOBTss — weighted optimized background therapy sensitivity score; FPR — false-positive rate; ARVs — antiretrovirals. *Some values were derived from Gulick et al, NEJM 2005 [288], and from Fätkenheuer et al, NEJM 2005 [289] and each is marked with an asterix.

The failure sample was taken as the earliest available on-treatment sample with both a viral load >500 copies/mL and an OTA result from the same time-point. Samples with viral loads <500 copies were not tested by OTA, and were therefore excluded from the study. The median viral load at failure was 4.1 $\log_{10}$ copies/mL (IQR: 3.5 – 5.0 $\log_{10}$), ranging from a minimum of 670 copies/mL to a maximum of 10 million copies/mL.

Most patients (55%) in the study experienced protocol-defined virologic failure (PDVF) over the 48 weeks of the MOTIVATE trials. However, the samples tested were generally from earlier time points than the week where PDVF was met. The median time to PDVF was approximately 17 weeks and 25 weeks for the groups failing with non-R5 and R5 OTA phenotypes, respectively [289]. In comparison, the samples in the current study were from a median of 4 and 4 weeks, for those with non-R5 and R5 phenotypes, respectively, since the earliest available failure samples were intentionally selected.

5.3.2 Performance of Population-Based Genotyping for Determining HIV Tropism

At screening, all patients had R5 HIV by OTA, but 12% and 31% had non-R5 results by population-based sequencing and ESTA, respectively. At the failure time point, 91 patients had genotypic non-R5 results (50%) and 90 patients had genotypic R5 results (50%) by population sequencing. In comparison, the proportions reported by the phenotypic OTA were 105 non-R5 (58%) and 76 R5 (42%). Approximately half of the patients had non-R5 HIV by both genotypic and phenotypic methods at failure (89 patients, 49%). Of the remaining patients, 41% had R5 HIV by both methods

(N=74), and 10% (N=18) had discordant results (with 16/18 having R5 as determined by genotype but non-R5 by OTA).

5.3.3 Change in V3 Sequence & Geno2pheno Value after Maraviroc Treatment

For a number of patients, there were large changes in the geno2pheno false-positive rate following maraviroc treatment. The overall change in geno2pheno for the population as a whole is shown in Figure 5.1. The median FPR for all patients regardless of tropism status fell from 31.0 at screening to 5.3 at failure (Figure 5.1), owing to the large number of patients failing with non-R5 HIV. Importantly, these patients fell into two distinct categories: those who maintained essentially the same geno2pheno FPR, and those for whom a large decrease in the FPR value between screening and failure was observed (Figure 5.2). These categories of failure generally corresponded with failing with an R5 or non-R5 phenotype, respectively.

The overall drop in geno2pheno FPR was driven by an increase in the number of patients with non-R5 genotypes, with this number increasing over 4-fold from 21 patients at screening to 91 patients at failure (12% to 50%). Between screening and failure, the geno2pheno FPR fell by a median of 18.2 units (IQR: -38.0 – -5.7) for those patients with concordant non-R5 results (Figure 5.3). These patients had extremely low geno2pheno false-positive rates at failure, with a median FPR of 1.1 (IQR: 0.4 – 1.7). In comparison, the median FPR of these same patients at screening was 20 (IQR: 6.9 – 38). In contrast, there was negligible change in geno2pheno FPR in patients failing with concordant R5 results. For these patients, the median FPR change was 2.2 (IQR: -0.5 – 16) (Figure 5.3, Table 5.1).

Figure 5.1: Overall Change in Geno2pheno False-Positive Rate between Screening & Failure

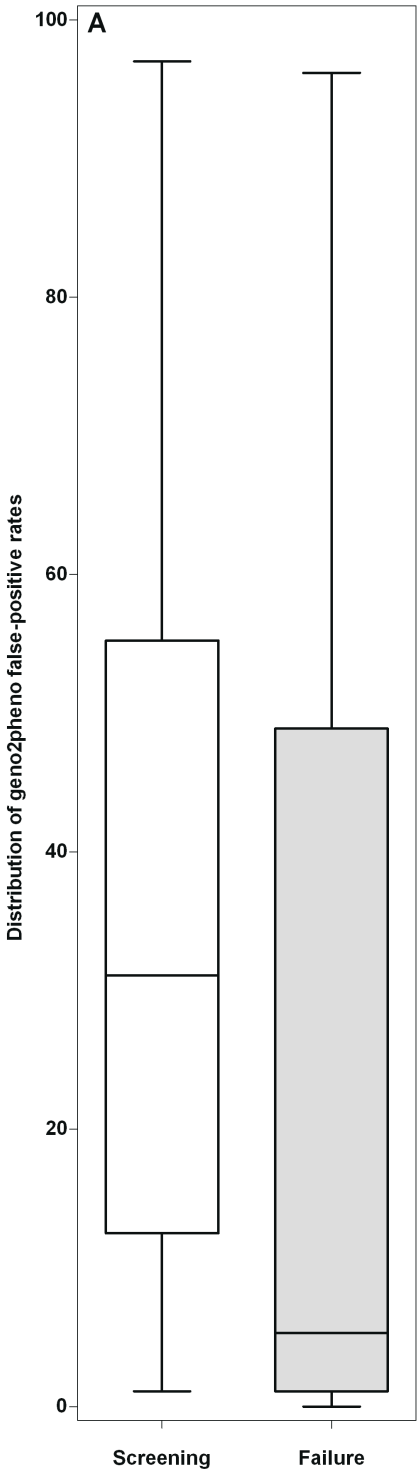


Figure 5.1: Overall Change in Geno2pheno False-Positive Rate between Screening & Failure. The overall decrease in the geno2pheno false-positive rate value between screening and failure. The distribution of geno2pheno false-positive rate (FPR) values is shown for the screening (left) and failure sequences (right). Boxes indicate the interquartile range of the values, with the median value indicated by a solid horizontal line. Whiskers correspond to 1.5 times the interquartile range.

Figure 5.2: Individual Geno2pheno False-Positive Rate Values at Screening & Failure

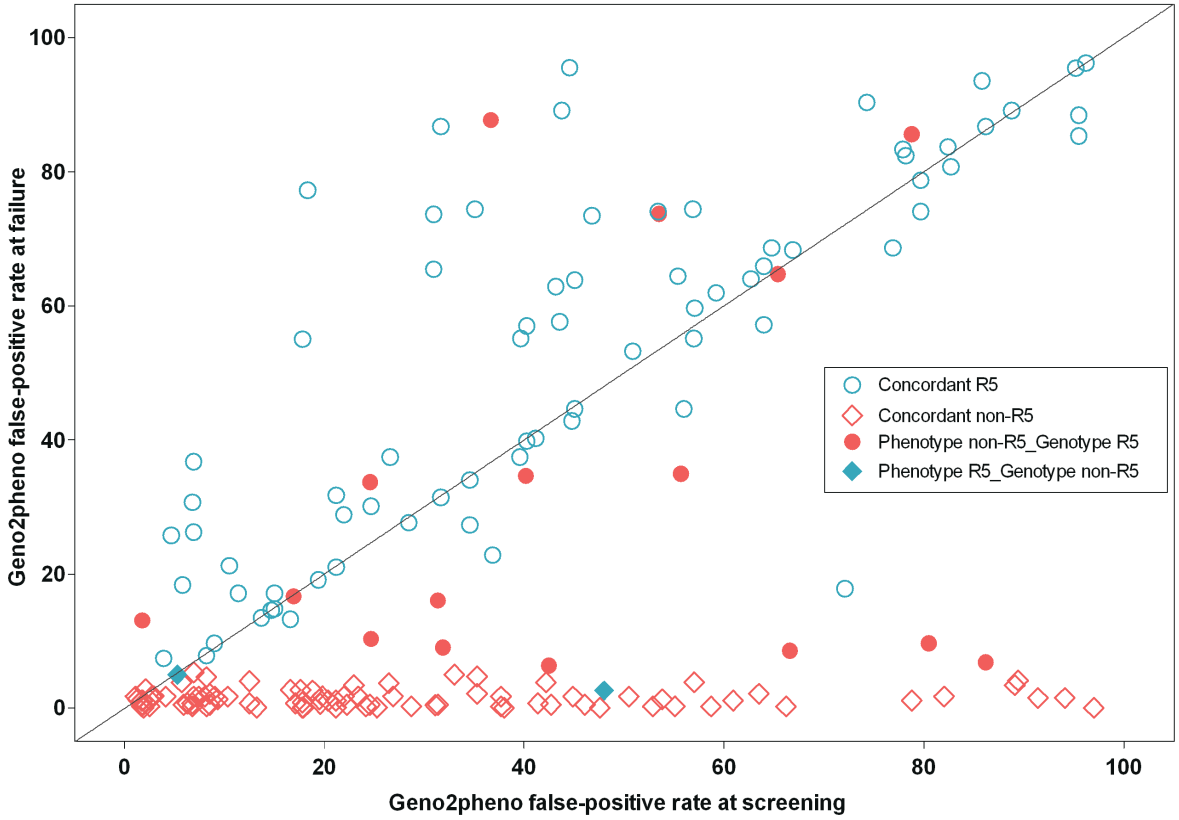


Figure 5.2: Individual Geno2pheno False-Positive Rate Values & Screening and Failure. A scatterplot of the geno2pheno false-positive rate (FPR) of all patients with coordinates at two time-points: screening value on the horizontal axis and failure value on the vertical axis. Points are marked by whether tropism results at failure were concordant between phenotype and genotype (see legend). The geno2pheno decreased by a large amount between screening and failure for patients in the non-R5 group, but changed very little for those in the R5 group.

Figure 5.3: Individual Changes in Geno2pheno False-Positive Rate Values between Screening & Failure

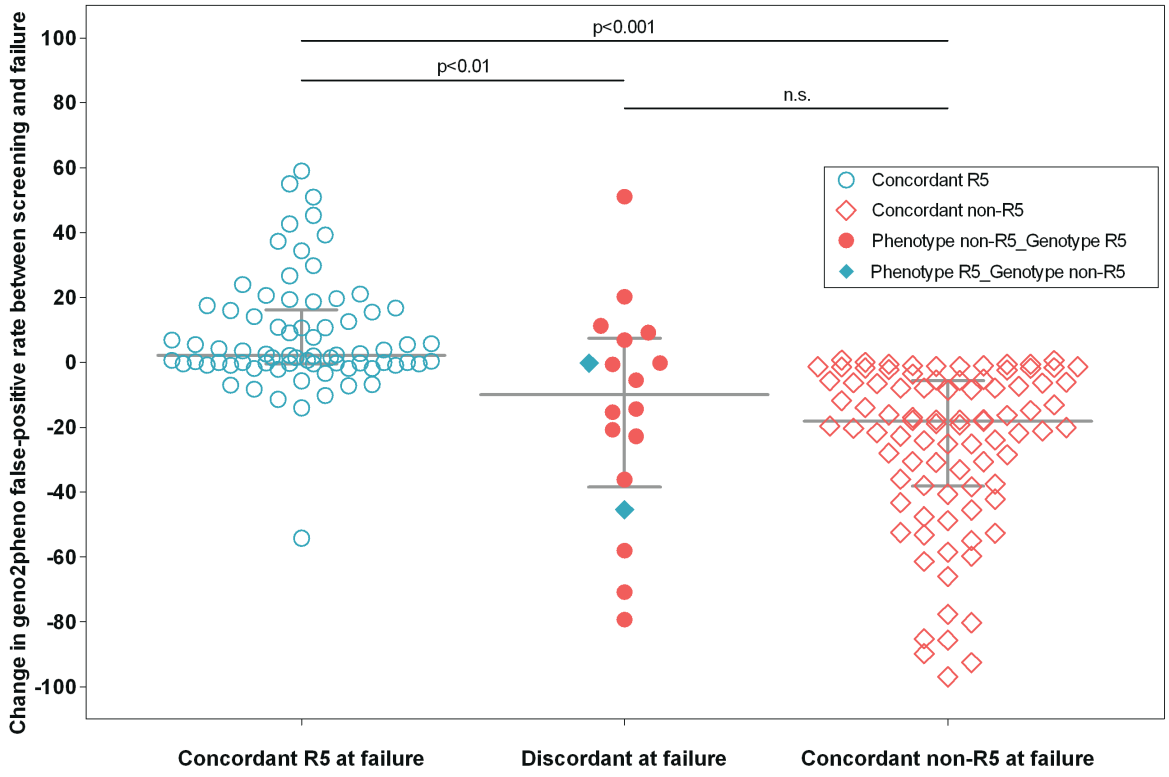


Figure 5.3: Individual Changes in Geno2pheno False-Positive Rate Values between Screening & Failure. The change in geno2pheno false-positive rate between screening and failure for each patient. Horizontal lines denote the median values, with error bars indicating the interquartile ranges. Patients with concordant R5 tropism at failure had a median FPR change of 2 (IQR: -1 – 16), versus a median decline of 18 FPR units (IQR: -38 – -6) in the concordant non-R5 group (p<0.001). Patients with discordant results at failure had an overall intermediate FPR decline (median FPR change = 10; IQR: -39 – 7). Points are marked by whether tropism results at failure were concordant between phenotype and genotype (see legend).

5.3.4 Classical Substitutions in Patients with Non-R5 HIV at Failure but Limited Evidence of Maraviroc Resistance in Those with R5 HIV

When the sequences from patients with non-R5 genotypes at failure were examined, not surprisingly, the most common emergent amino acid substitutions among patients with non-R5 genotypes were substitutions to basic amino acids. There were 3 primary codons where non-R5 substitutions occurred. These occurred as follows: 11R (36 patients, 40%), 13R (23 patients, 25%), and 25K (20 patients, 22%). Consequently, the 11/25 rule [133] could identify a substantial proportion of sequences as non-R5 (57/91, 63%) at the time of failure.

In contrast to patients with non-R5 failure genotypes, patients who maintained genotypic R5 HIV through the study period (N=90) exhibited no clear accumulation of mutations at the failure time point. Among these patients, the median geno2pheno FPR at screening and failure showed very little change, at 41.9 versus 48.9, respectively. In 21 patients (23%) no V3 amino acid changes were observed following maraviroc treatment, while substitutions in the remaining 70 patients (77%) were restricted to partial amino acid changes (mixtures).

Among these patients without a tropism change, the sites with the highest rates of substitutions were codons 10, 13, 14, 18, and 25. The most common substitutions at these positions were 10R, 13H/P, 14I/M, 18R, and 25D. Importantly, however, the prevalence of these substitutions was very low in this population, ranging between 8 and 14 samples (9 – 16%) depending on the substitution. Prevalence was low even for substitutions previously documented to be associated with maraviroc resistance

These may indeed be simply natural polymorphisms unrelated to maraviroc resistance. Furthermore, since most of these samples did not have phenotypic maraviroc resistance assay results, the ability to interpret the implications of these substitutions was limited, and many patients with R5 viruses may simply have been non-adherent, or had viruses that were resistant to other components of their regimens.

5.3.5 Change in Non-R5 Viral Population as Determined by Deep Sequencing

The viral population present prior to treatment with maraviroc was assessed by deep sequencing of the screening plasma samples. The deep sequencing data was then investigated for both the change in the percentage of non-R5 variants, as well as the phylogenetic relationship between the screening and failure V3 sequences.

All screening samples underwent deep V3 sequencing, as did a subset of 73 failure samples. At screening, the median percentage of non-R5 variants per patient was 0.1% (IQR: 0–3.1%) — reflective of the R5 phenotypes of all patients. However, a majority of patients had at least some level of non-R5 sequences present at screening (96/181, 53%), and including one patient in whom 99.9% of recovered sequences were interpreted as non-R5 at screening, who had R5 by OTA and ESTA at screening, but experienced virologic failure with Dual-Mixed tropism by OTA at week 4. A total of 50 patients (28%) had ≥2% non-R5 HIV according to their deep sequencing screening result. This was over twice as many as were detected at screening by population-based sequencing, despite all patients having R5 OTA phenotypes at screening.

Of the 50 patients with non-R5 variants present at ≥2% prevalence by deep sequencing, 42 (84%) were confirmed to have non-R5 at failure by population-based genotype. Where deep sequencing results were available at both time points (N=73), the overall median percentage of non-R5 variants rose slightly from 0% (IQR: 0 – 1.2%) at screening to 0.8% (0 – 98.8%) at failure. When these patients were restricted to those with non-R5 at failure by population-based sequencing, the median percentage of non-R5 variants rose to 99.4% (IQR: 95.4 – 99.9%) at failure.

Strikingly, distribution of non-R5 variants in patients treated with maraviroc was nearly completely dichotomous. According to deep sequencing results, the vast majority of patients (65/73, 89%) had treatment failure with either less than 5% non-R5 variants or greater than 95% non-R5 variants, with very few patients falling in between. As previously mentioned, the population-based sequencing results were also quite unambiguous in their interpretation. Of those with non-R5 population sequencing results at failure, over three-quarters had extremely low geno2pheno FPRs of 2 or lower (70/91, 77%), indicative of "highly" non-R5 virus [336] (Figure 5.4).

Figure 5.4: The Percentage of Non-R5 Variants in Deep Sequencing Results at Screening & Failure

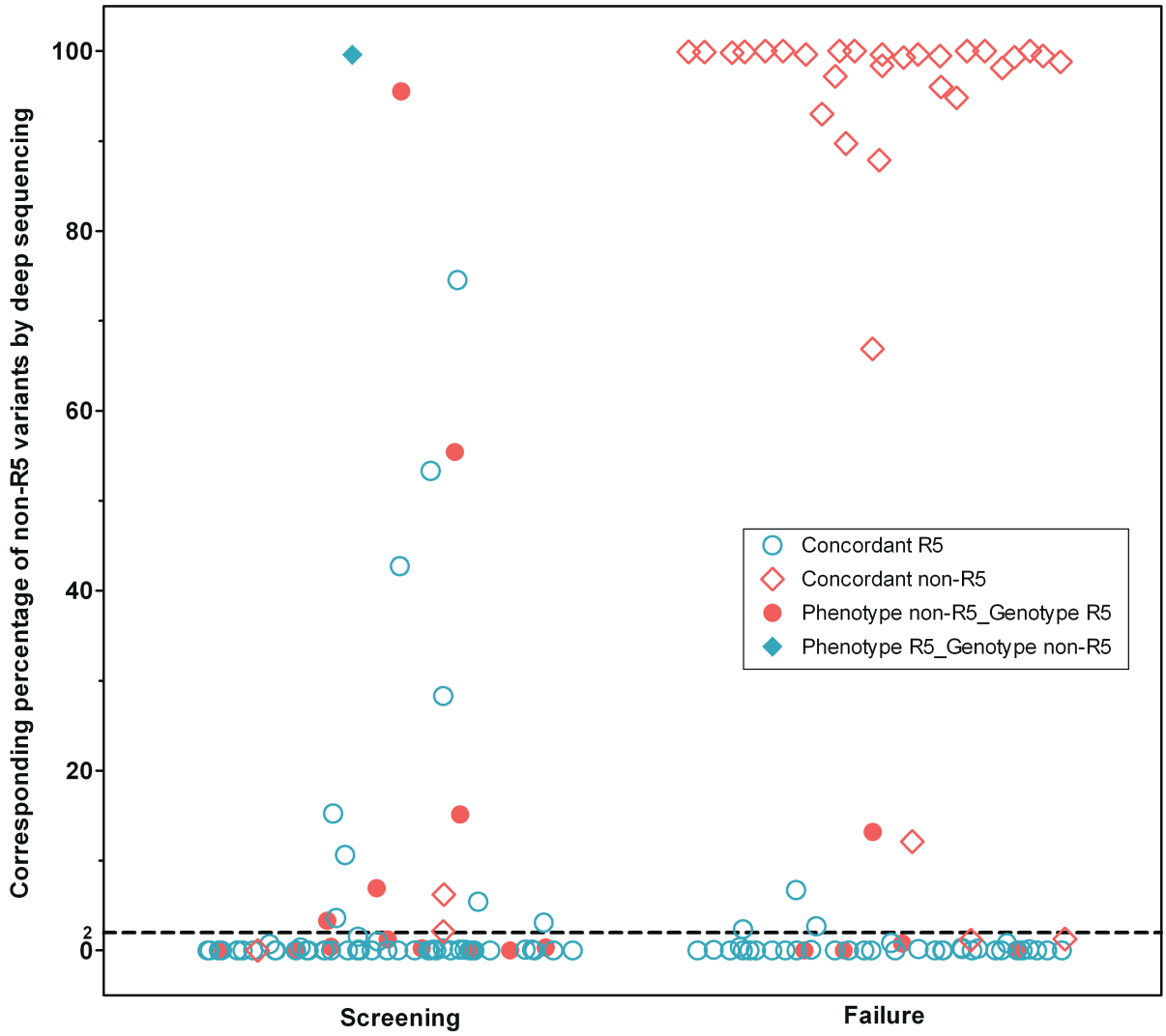


Figure 5.4: The Percentage of Non-R5 Variants in Deep Sequencing Results at Screening & Failure.
A scatterplot of the percent non-R5 variants for all patients with deep sequencing results at screening and failure. Points are marked by whether tropism results in the same time-point were concordant between phenotype and population-based genotype (see legend). The failure column illustrates how the majority of failure samples had very high or very low non-R5 prevalence. Phenotypes were performed by ESTA at screening and original Trofile assay (OTA) at failure. A dashed line at 2% non-R5 prevalence represents a cutoff for deep sequencing, above which a sample was classified as having non-R5 tropism

5.3.6 Phylogenetic Relationship between Screening & Failure Sequences

Phylogenetic trees were generated using the screening deep sequencing data and the failure population-based sequence. For many patients, a distinct minority subpopulation of non-R5 variants

was detected by deep sequencing at screening. This minority subpopulation often emerged following treatment, and was detected with standard population-based sequencing methods. The trees were inspected manually to assess the degree of the phylogenetic relationship between the failure V3 sequence and sequences detected by deep sequencing prior to maraviroc treatment.

Overall, 70% of patients (64/91) with non-R5 HIV at failure had a closely related non-R5 subpopulation present prior to treatment with maraviroc, confirming previous reports of the selection of pre-treatment non-R5 reservoirs by maraviroc [232,316]. These CXCR4-using subpopulations were present despite patients being pre-screened as having R5 HIV with OTA. A number of these patients were missed by population-based sequencing as well. A set of representative example trees is given in Figure 5.5-5.8 and Appendices V–XIV.

5.3.7 Comparison of Tropism Methods

The performance of population-based sequencing could be assessed at both screening and failure by comparing the results using deep sequencing as the "gold standard". ESTA results were available for comparison at screening, and OTA at failure. When the two genotypic tropism methods were compared at screening, population-based sequencing had 30% sensitivity (15/50 non-R5 samples) and 95% specificity (125/131 R5 samples) relative to deep sequencing. However, performance of population-based sequencing was dramatically better at failure. This method achieved 88% sensitivity (29/33 called non-R5) and 95% specificity (38/40 called R5) relative to deep sequencing, likely due to the higher proportions of non-R5 variants after maraviroc treatment.

Figure 5.5: Phylogenetic Tree from a Patient Who Had a Small Pre-Treatment X4 Population

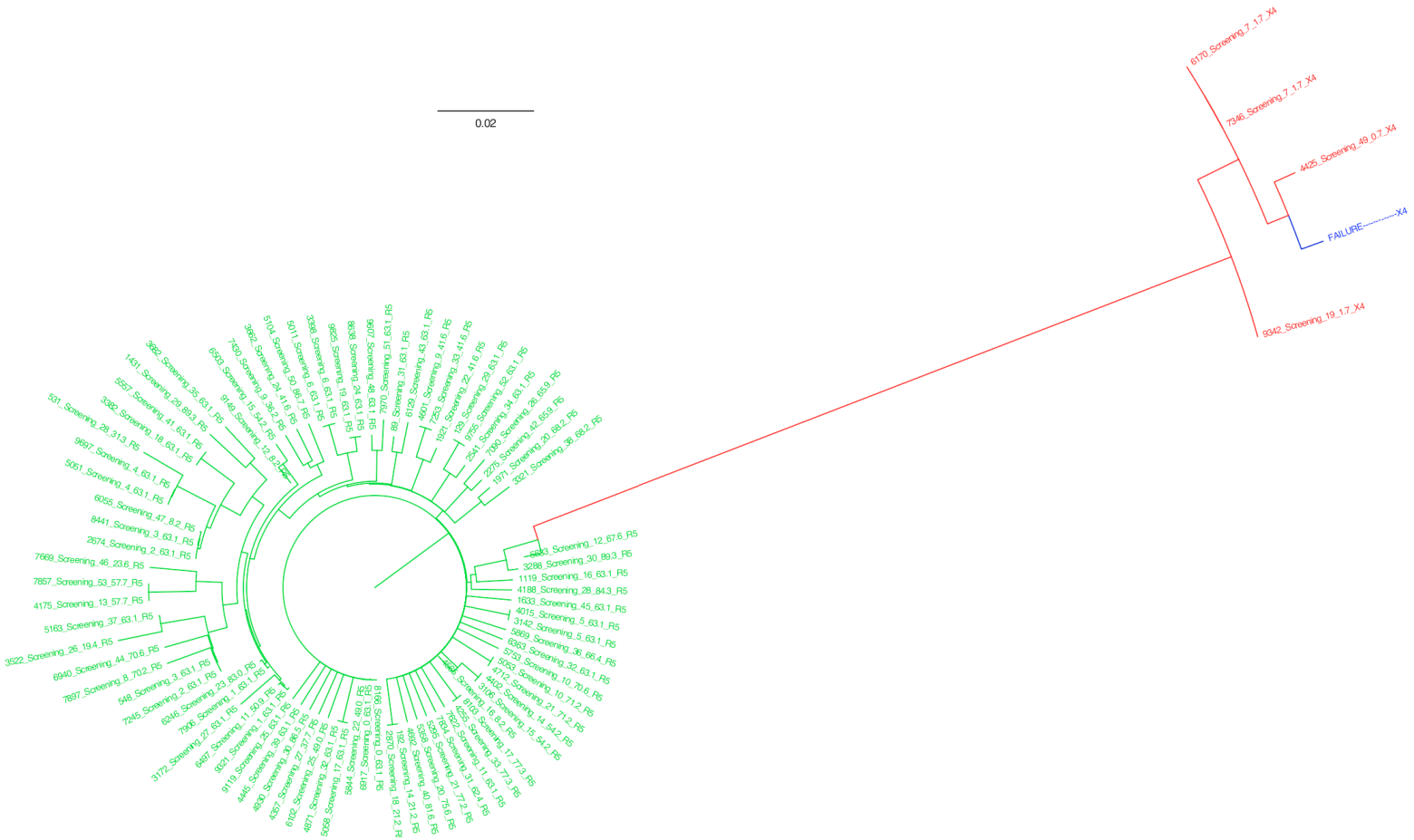


Figure 5.5: Phylogenetic Tree from a Patient with a Small Pre-Treatment X4 Population. Phylogenetic tree generated from the deep sequencing data at screening and the population-based genotype at failure. Screening R5 sequences are shown in green, X4 sequences are shown in red, and the failure sequence is shown in blue. Failure was due to a small pre-treatment X4 population detected by deep sequencing (1% X4) but not by phenotyping (ESTA R5). The failure sample was X4 by population-based genotype and X4 by the original Trofile assay.

Figure 5.6: Phylogenetic Tree from a Patient Who Had a Large Pre-Treatment X4 Population

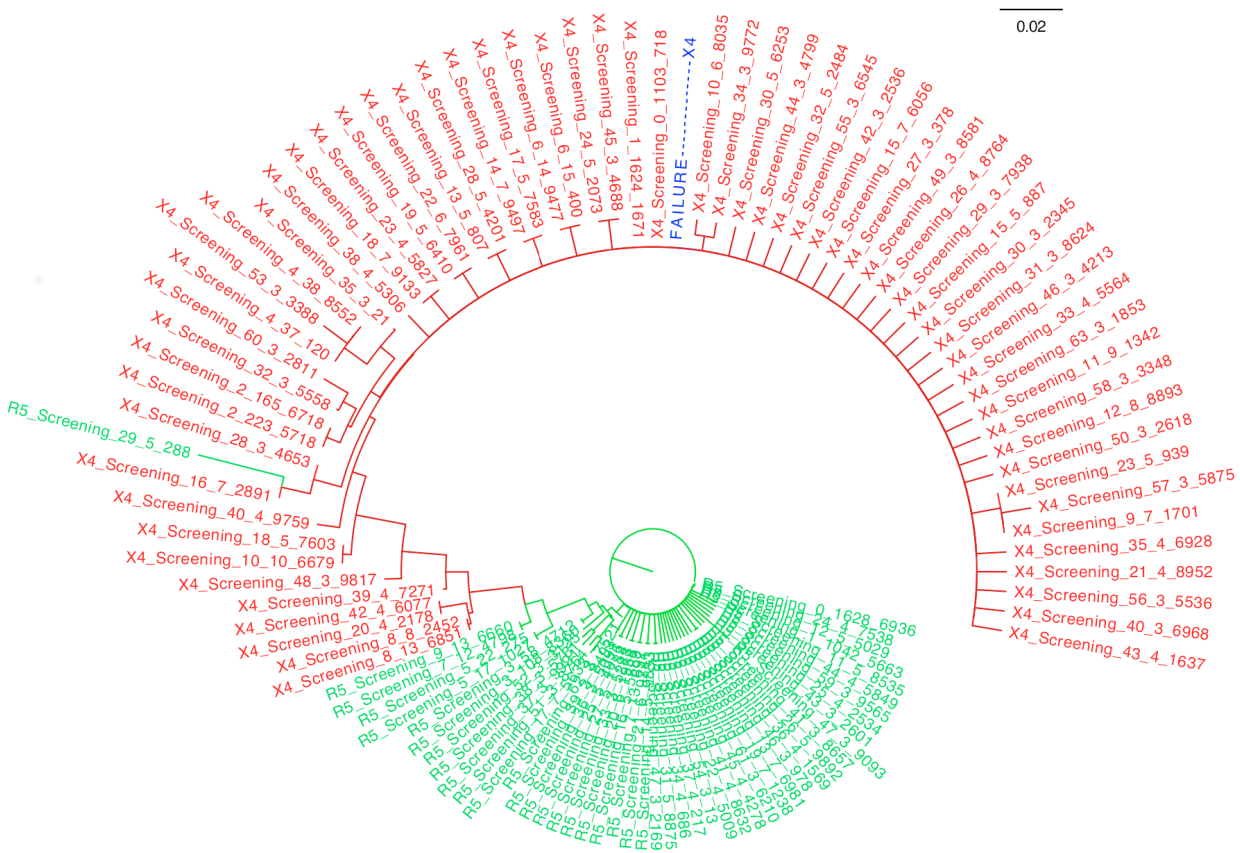


Figure 5.6: Phylogenetic Tree from a Patient with a Large Pre-Treatment X4 Population. Phylogenetic tree generated from the deep sequencing data at screening and the population-based genotype at failure. Screening R5 sequences are shown in green, X4 sequences are shown in red, and the failure sequence is shown in blue. Failure was due to a large pre-treatment X4 population detected by deep sequencing (53% X4) but not by phenotyping (R5 by ESTA). The failure sample was X4 by population-based genotype and Dual-Mixed by OTA.

Figure 5.7: Phylogenetic Tree from a Patient for Whom Deep Sequencing Failed to Detect X4 HIV

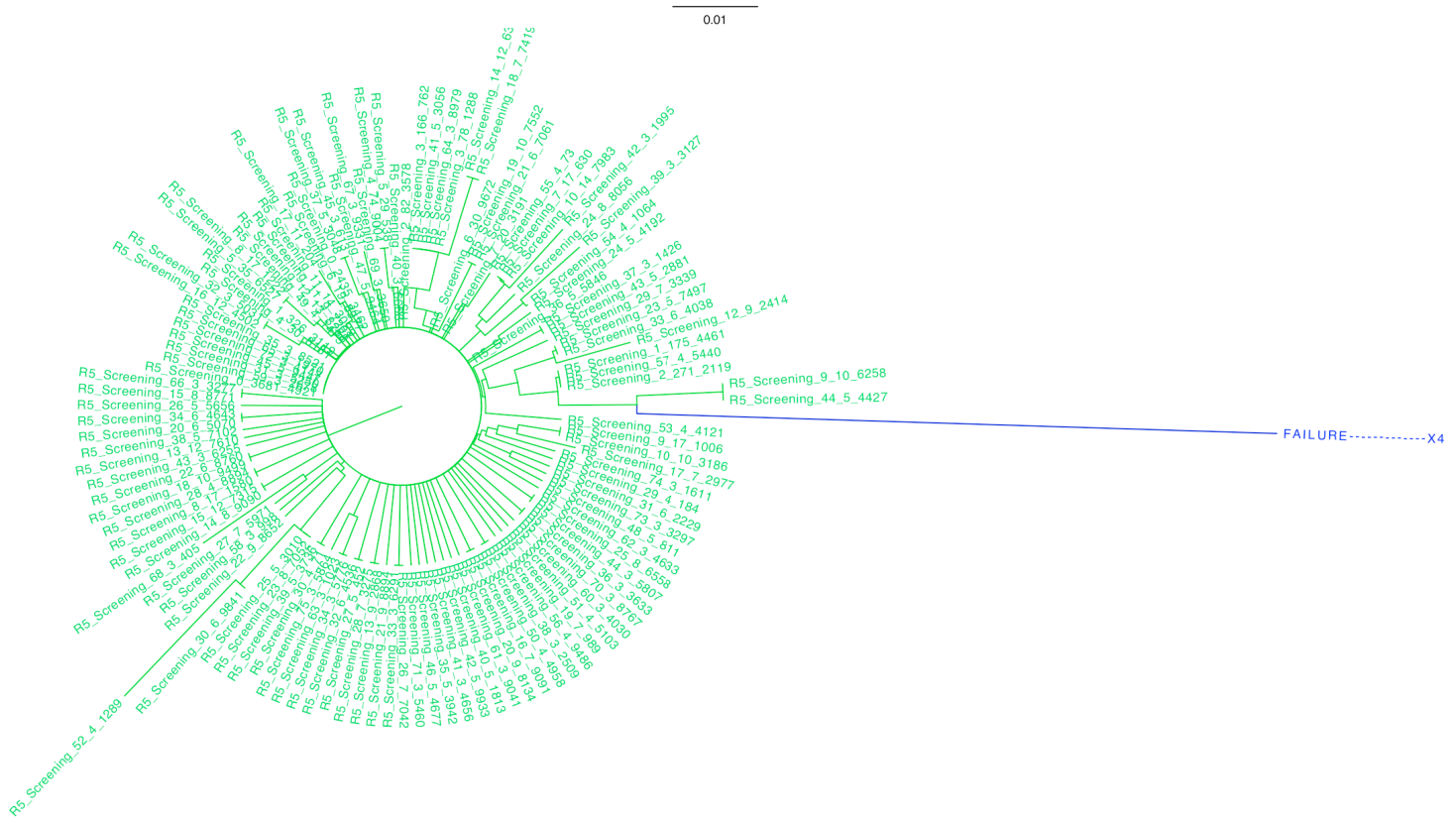


Figure 5.7: Phylogenetic Tree from a Patient for Whom Deep Sequencing Failed to Detect X4 HIV. Phylogenetic tree generated from the deep sequencing data at screening and the population-based genotype at failure. Screening R5 sequences are shown in green, X4 sequences are shown in red, and the failure sequence is shown in blue. Pre-treatment X4 sequences were not detected by deep sequencing (0% X4) and screening phenotype was R5 by ESTA. However, failure occurred with a non-R5 genotype and phenotype.

Figure 5.8: Phylogenetic Tree from a Patient Who Failed with R5 HIV

Figure 5.8: Phylogenetic Tree from a Patient Who Failed with R5 HIV. Phylogenetic tree generated from the deep sequencing data at screening and the population-based genotype at failure. Screening R5 sequences are shown in green, X4 sequences are shown in red, and the failure sequence is shown in blue. A patient who experienced failure with an R5 genotype and phenotype. One outlier branch has been truncated for display purposes. Both screening and failure time points were R5 by all genotypic or phenotypic tests performed

The genotypes were also compared to the phenotypes. At screening, population-based sequencing had 19% sensitivity (11/57) and 92% specificity (114/124) relative to ESTA. Deep sequencing had 53% sensitivity (30/57), and 84% specificity (104/124) relative to ESTA at screening. At failure, when population-based genotypes were compared to the OTA phenotypes in the same time-point the assays were 90% concordant (163/181 samples). Overall sensitivity of genotyping compared to phenotyping was 85% (89/105 non-R5), with 97% specificity (74/76 R5) in these failure samples. This performance was comparable to the performance of deep sequencing relative to OTA at failure: 83% sensitivity (30/36), 92% specificity (34/37).

5.3.8 Virologic Responses to Maraviroc

While all patients were R5 by OTA at screening, they could be stratified by their genotypic tropism results in their failure visit. Patients with a non-R5 genotype by population-based sequencing at failure had overall poorer virologic responses to maraviroc. At week 8, the median decline in plasma viral load (pVL) from baseline was 2.0 $\log_{10}$ in those with R5 HIV but 0.4 $\log_{10}$ in those with non-R5 HIV at failure ($p<0.001$). In contrast, the median change in pVL at week 8 for the maraviroc arms in the MOTIVATE trials overall was approximately 2.4 $\log_{10}$ copies/mL. This was larger than the viral load decreases for either group in the current study (Figure 5.9, $p<0.01$).

A total of 88% of patients with non-R5 population genotypes at failure (80/91) failed to achieve an undetectable viral load during the study, versus 77% (69/90) of those with R5 genotypes at failure. Protocol-defined virologic failure was documented for 69% of patients with non-R5 genotypes at failure (63/91), compared to 41% of patients with R5 at failure (37/90). Patients with R5 at failure had higher rates of virologic rebound compared to those with non-R5: 14% (13/90) versus 2% (2/91). They were also twice as likely to have never suppressed throughout the study but remain enrolled: 32% (29/90) versus 16% (15/91).

Figure 5.9: Virologic Responses Were Reduced among Patients with Non-R5 Genotype Results at Failure

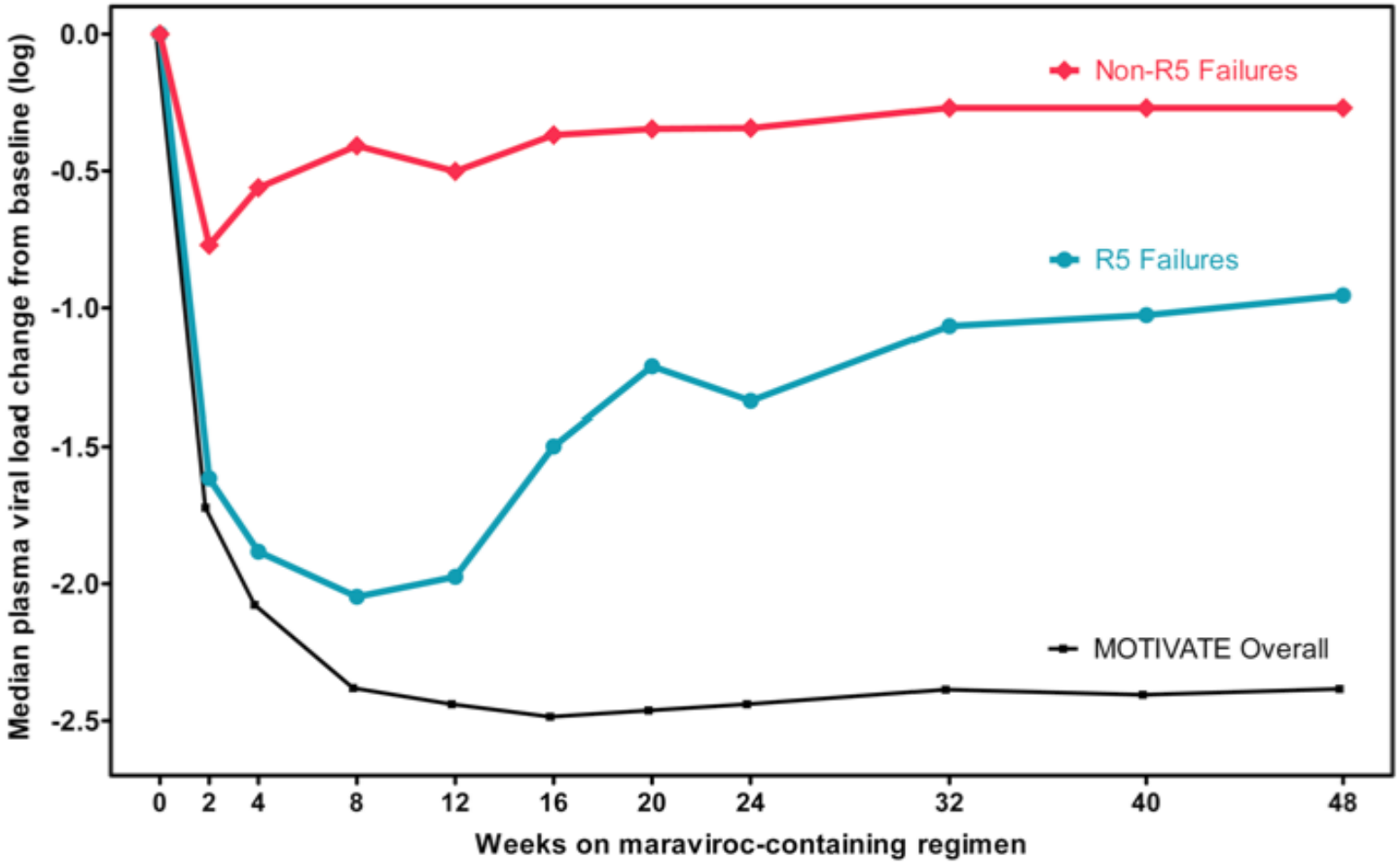


Figure 5.9: Virologic Responses Were Reduced among Patients with Non-R5 Genotype Results at Failure. The median change in plasma viral load from baseline among maraviroc recipients is shown. Patients are stratified according to whether their first available failure sample had an R5 (turquoise line) or non-R5 (red line) population genotype. For comparison, the median viral load change of the maraviroc arms in the MOTIVATE trials overall is also shown (black line).

5.3.9 Comparison to the Enhanced Sensitivity Trofile Assay

As stated above, sensitivities of population-based sequencing and deep sequencing were 19% and 53% relative to ESTA at screening, respectively, with concordance of 69% and 74%. Of those where re-screening by ESTA indicated pre-treatment non-R5 phenotypes, 50 of 57 (88%) patients were confirmed to have OTA non-R5 phenotypes at failure, similar to the results by genotyping (47 of 57 patients, 82%). Patients with pre-treatment R5 phenotypes by both OTA and ESTA were more likely to fail therapy with R5 phenotypes or genotypes (56% or 65%) than non-R5 (44% or 35%).

All follow-up results were tested with OTA, but it is important to note that this study indicates that very low minority non-R5 variants present only at the time of failure are not commonly associated with suboptimal maraviroc responses. Deep sequence analysis demonstrated that when phenotypic tropism changes occurred, they were generally accompanied by very high non-R5 prevalence (Figure 5.4). Accordingly, the current results are likely to be unaffected by the fact that the failure phenotypes were performed using OTA rather than ESTA. Furthermore, largely similar results were obtained when analyses were restricted only to patients with R5 by ESTA at screening (Figures 5.10-5.12).

5.4 Discussion & Conclusions

In this study, genotypic analysis indicated that failure on maraviroc followed two distinct pathways. Those patients who experienced an HIV tropism shift to non-R5 had a large decline in geno2pheno false-positive rate and accumulated V3 substitutions at multiple codons, and had a large increase in the prevalence of non-R5 variants to a median of 99% according to deep sequencing. Patients with R5 results at failure tended to have very similar geno2pheno values to their screening values, and accumulated few amino acid substitutions in V3 compared to the pre-treatment sequences.

Figure 5.10: Overall Change in Geno2pheno False-Positive Rate between Screening & Failure in a Subset of Patients with R5 Results at Screening by the Enhanced Sensitivity Trofile Assay

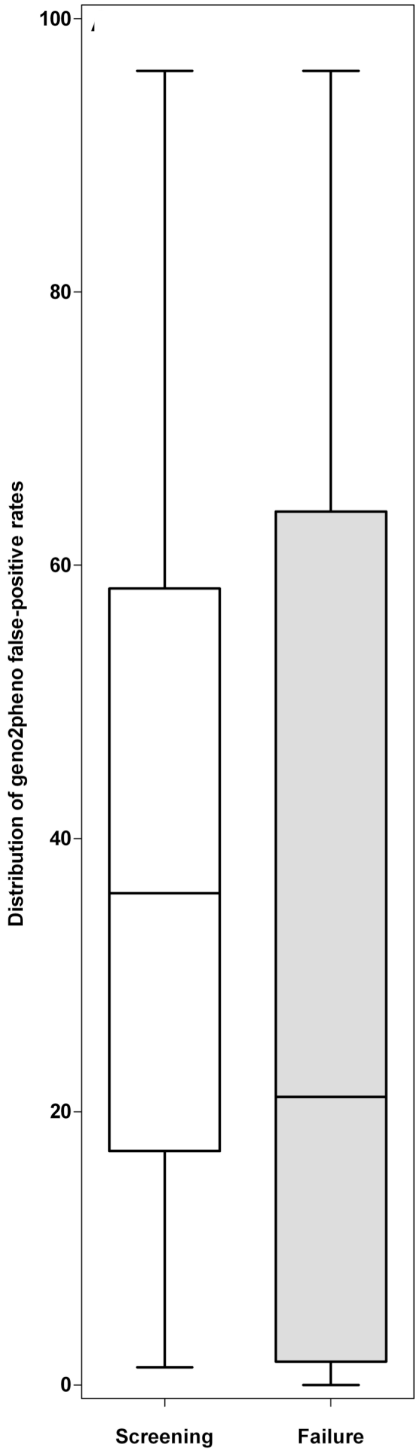


Figure 5.10: Overall Change in Geno2pheno False-Positive Rate between Screening & Failure in a Subset of Patients with R5 Results at Screening by the Enhanced Sensitivity Trofile Assay. The distribution of geno2pheno false-positive rate (FPR) values is shown for the screening (left) and failure sequences (right). Boxes indicate the interquartile range of the values, with the median value indicated by a solid horizontal line. Whiskers correspond to 1.5 times the interquartile range.

Figure 5.11: Individual Geno2pheno False-Positive Rate Values at Screening & Failure in a Subset of Patients with R5 Results at Screening by the Enhanced Sensitivity Trofile Assay

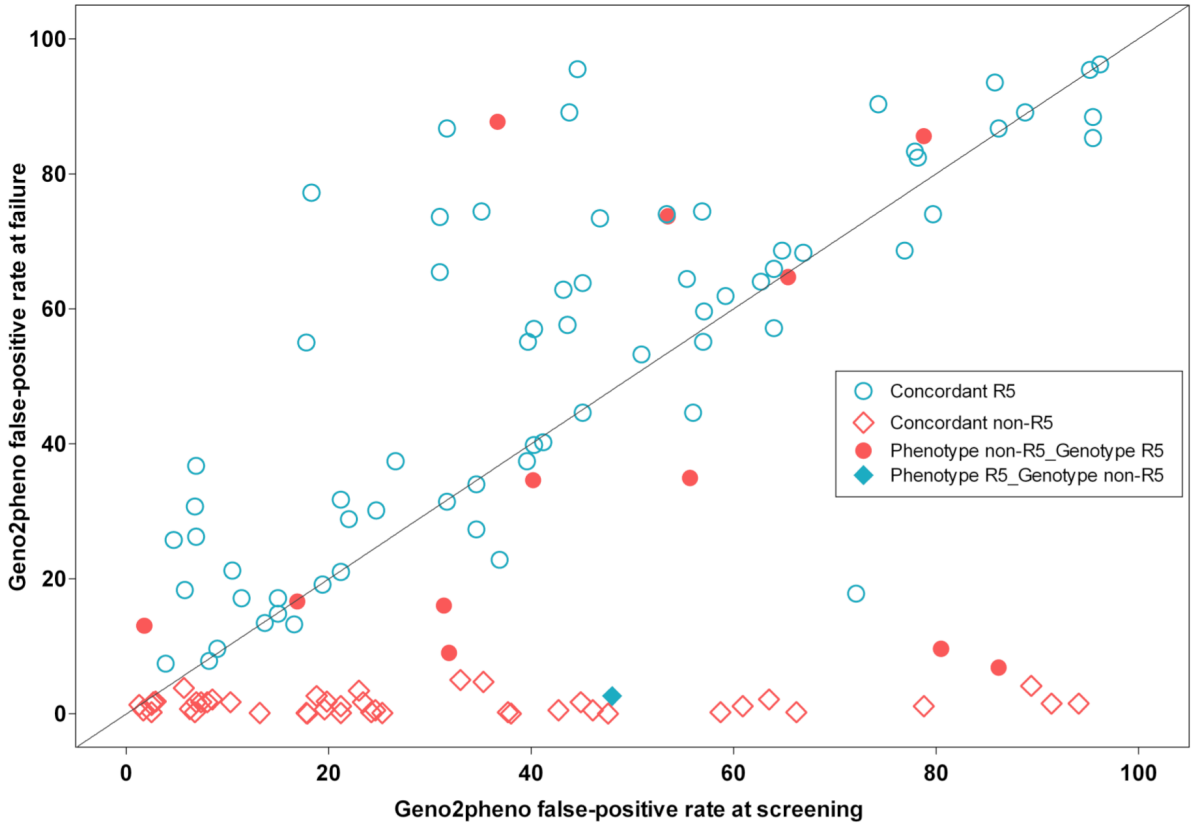


Figure 5.11: Individual Geno2pheno False Positive Rate Values between Screening & Failure in a Subset of Patients with R5 Results at Screening by the Enhanced Sensitivity Trofile Assay. A scatterplot of the geno2pheno false-positive rate (FPR) of all patients with coordinates at two time-points: screening value on the horizontal axis and failure value on the vertical axis. Points are marked by whether tropism results at failure were concordant between phenotype and genotype (see legend). The geno2pheno decreased by a large amount between screening and failure for patients in the non-R5 group, but changed very little for those in the R5 group.

Figure 5.12: Individual Changes in Geno2pheno False-Positive Rate Values between Screening & Failure in a Subset of Patients with R5 Results at Screening by the Enhanced Sensitivity Trofile Assay

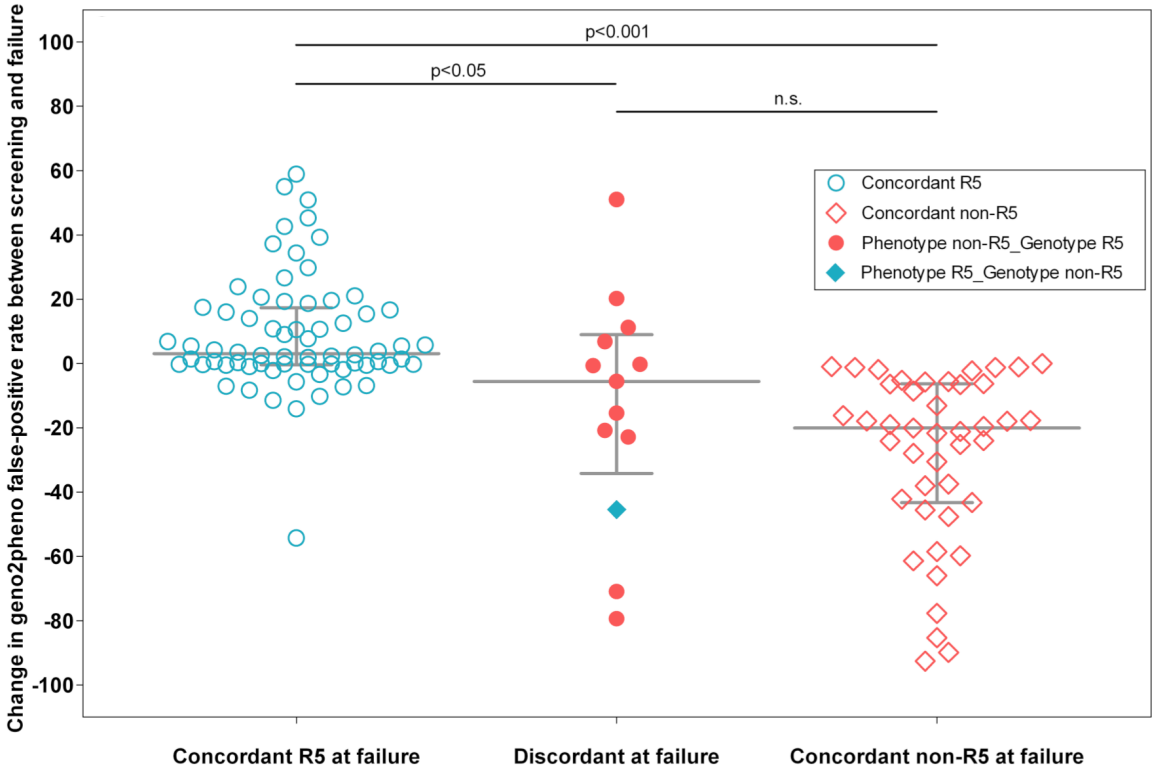


Figure 5.12: Individual Changes in Geno2pheno False-Positive Rate Values between Screening & Failure in a Subset of Patients with R5 Results at Screening by the Enhanced Sensitivity Trofile Assay. A scatterplot of the false-positive rate change between screening and failure. Horizontal lines denote the median values, with error bars indicating the interquartile ranges.

In 70% of patients, deep sequencing was able to detect a pre-treatment non-R5 subpopulation which emerged at failure. This study also demonstrated that standard population-based sequencing is capable of identifying on-treatment tropism changes accompanying maraviroc failure, and does so with high sensitivity (85%) relative to phenotyping. The sensitivity reported here for population-based sequencing is much higher than previously reported results [325]. The sensitivity of population-based sequencing more than tripled after patients began treatment with maraviroc (88% sensitivity on-treatment versus 24% at screening when compared to deep sequencing). Other studies have typically reported much lower sensitivities for population-based sequencing [187,190,229,325,371].

The high sensitivity reported for the failure samples in the current study is likely attributable to the selective effect of maraviroc treatment on patient HIV. For those who failed maraviroc-based therapy with non-R5 HIV, the average percentage of non-R5 variants rose to 99% according to deep sequencing. This likely increased the ability of population-based sequencing to give a non-R5 result. Under these circumstances, population-based methods performed better than usual since non-R5 prevalence is usually masked by a predominantly R5 viral population, thus limiting sensitivity. This rapid emergence of high prevalence non-R5 variants has also been reported for cases of treatment failure with other CCR5 antagonists [317,372].

Maraviroc recipients with non-R5 HIV infection at failure had poorer virologic responses to the medication than those whose virus did not change tropism — even within this population of patients who failed maraviroc-based therapy. This is likely due to the additional loss of maraviroc activity in patients who fail therapy with non-R5 HIV, whereas failure with R5 HIV may have been due to a number of reasons, including maraviroc resistance, resistance to the other background antiretroviral agents, and/or poor adherence. However, response to antiretroviral therapy in general may also be impacted by the presence of non-R5 HIV infection [155,173].

Some limitations of this study should be acknowledged. It may be difficult to extend these findings generally to all maraviroc-treated populations as this study was conducted in a selected population pre-screened for R5 HIV by OTA. However, the proportion of non-R5 OTA results at failure in the current study (57%) was quite reflective of the proportion seen in the MOTIVATE trials overall (58%) [289], so these results are likely generalizable to the larger MOTIVATE study population. While ESTA results were available at screening, the follow-up phenotypes were performed using OTA. However, additional sensitivity of ESTA over OTA (reported detection limit of 0.3% versus 10% non-R5 HIV [365]) likely had very little affect on the above results, since genotypic analyses indicated that phenotypic tropism changes on maraviroc were associated with extremely high non-R5 prevalence well above the 10% detection limit of the original Trofile assay.

Despite its high sensitivity, deep sequencing could not identify pre-existing non-R5 populations in 30% of patients with non-R5 HIV at failure. Thus, this approach may still lack sufficient sensitivity, or the sampling volume may have been insufficient for detection of minority variants. Alternatively, non-R5 variants may evolve from R5 populations more rapidly during maraviroc treatment, or may have emerged between screening and enrollment, as previously reported in 8% of MOTIVATE participants [289]. Indeed, among the studied population, 25 patients had switched to non-R5 phenotypes at maraviroc initiation (14%). It is also possible that non-R5 variants may also emerge from compartments other than blood plasma.

This study was limited in its ability to better characterize maraviroc resistance. To date, a reduced maximal percentage inhibition (MPI) in a phenotypic assay assessing susceptibility of the patient virus to maraviroc is the only consistent characteristic of maraviroc resistance [373]; no signature mutations have been observed for maraviroc [374,375] or other CCR5 antagonists [376]. Similar to these prior findings, no consistent patterns of mutations were noted that were associated with virologic failure while maintaining an R5 population. This may be due to a number of factors, such as insufficient time on the medication to induce resistance-associated mutations, the possibility that mutations may emerge outside the V3 loop, and/or the possibility that maraviroc resistance mutations are patient-specific and difficult to generalize. Furthermore, only a small number of patients who experienced failure on maraviroc with R5 viruses have actually been shown to have maraviroc-resistant HIV in phenotypic assays [374]. In the remaining patients, other factors such as adherence or resistance to the other agents in their regimens may be involved.

The aforementioned analyses indicate that maraviroc treatment dichotomized V3 sequences and inferred coreceptor usage. Genotypic tropism analyses demonstrated large decreases in geno2pheno values, and large increases in the percentage of non-R5 variants. Patients with non-R5 HIV at failure experienced suboptimal virologic responses to maraviroc in this study, likely driven by their non-R5 status. However, genotypic analysis in patients with failure R5 results was not informative, with little change in geno2pheno values, and few amino acid substitutions that might be attributed to maraviroc resistance. In contrast, the results in the non-R5 population were unambiguous and

striking, suggesting that both deep and population-based sequencing approaches are useful monitoring tools for patients receiving maraviroc.

While this and all previous chapters have assessed deep sequencing as applied to patients with previous treatment experience, Chapter 6 will extend this application to patients beginning antiretroviral therapy for the first time. This treatment-naïve population may be especially relevant since CCR5-using HIV-1 tends to be most prevalent early in infection [145,146]. It may therefore be a preferred population in which to prescribe maraviroc. Additionally, the next chapter aims to more thoroughly assess deep sequencing in comparison to a phenotypic assay with higher sensitivity than the original Trofile assay The following chapter also represents a completely independent validation of next-generation sequencing methods since they were originally optimized in the treatment-experienced population described previously.

Chapter 6: Deep V3 Sequencing for HIV-1 Tropism in Treatment-Naïve Patients: A Reanalysis of the MERIT Trial of Maraviroc

6.1 Background & Introduction

Human Immunodeficiency Virus type 1 (HIV-1) infects cells using the CD4 receptor and a coreceptor. The chemokine receptor CCR5 is a necessary coreceptor for strains of HIV called R5 [140], which predominate in antiretroviral-naïve individuals [139,155,173]. The CCR5 coreceptor is also the target of the HIV entry inhibitor, maraviroc. Maraviroc inhibits the ability of HIV to interact with and infect cells via CCR5 [108]. The use of an alternative coreceptor emerges in approximately half of clade-B-infected individuals [100]. Therefore, a tropism test is performed prior to maraviroc administration to exclude patients whose viral population (or some subpopulation of it) is non-R5 and unlikely to respond to maraviroc [67].

A number of genotypic HIV tropism approaches have been developed to provide alternatives to phenotypic tropism assays such as the Monogram Biosciences Trofile assay [377] and Enhanced Sensitivity Trofile Assay (ESTA) [195]. Commonly, genotypic approaches use the sequence of the third variable (V3) region of the HIV *gp120* gene, since the V3 loop itself interacts with the HIV coreceptor [122] and mutations encoded by V3 are associated with measurable changes in HIV-tropism [123,324]. Tropism is then inferred using a bioinformatic algorithm such as geno2pheno [197,204].

While population-based genotypic tropism assays can infer the coreceptor usage of a patient's most common HIV quasispecies, these tests may miss non-R5 variants comprising low-level minorities within a predominantly R5 population [229]. The ability to detect minority non-R5 variants is important because these subpopulations may undergo selection by maraviroc treatment and lead to virologic failure [232,291,316].

There have been four large clinical trials of maraviroc to date [288–291]. The Maraviroc versus Efavirenz Regimens as Initial Therapy (MERIT) trial assessed two doses of maraviroc (plus lamivudine/zidovudine) in antiretroviral-naïve patients, with a comparator arm of efavirenz (plus lamivudine/zidovudine) [290]. The trial consisted only of patients with R5 HIV at screening by the original Trofile assay. The maraviroc once-daily (QD) arm was discontinued early after failing to meet pre-specified efficacy criteria.

Although superior to placebo in the MOTIVATE trials, maraviroc was inferior to efavirenz in the primary analysis of the MERIT trial using the original screening population. However, when patients in MERIT were retrospectively re-screened using the higher-sensitivity ESTA, with exclusion of those now identified with non-R5 HIV, maraviroc twice daily (BID) was non-inferior to efavirenz for the primary study endpoint [290,292].

Deep sequencing refers to the application of next-generation sequencing technology such as the Genome Sequencer FLX (GS-FLX) [240]. The GS-FLX can simultaneously sequence and quantify thousands of individual variants within a viral population, allowing an in-depth quantification of the proportion of non-R5 variants in a given sample [232,352], and therefore the proportion unlikely to respond to maraviroc [348]. The aim of this study was to assess whether the high-sensitivity of deep sequencing could also have been a successful screening tool for the treatment-naïve patients in the MERIT trial.

6.2 Materials & Methods

6.2.1 Samples & MERIT Trial Design

A total of 859 plasma screening samples from the MERIT trial were examined. All samples were R5 by the original Trofile assay. Most patients entered either the maraviroc BID arm (N=347) or the efavirenz arm (N=346). The trial's primary endpoint was the proportion of patients with a viral load <50 HIV RNA copies/mL at week 48. A third arm consisting of maraviroc QD was also partially

enrolled. Screening samples from those initially assigned to the maraviroc QD arm (N=166) were also tested.

6.2.2 V3 Amplification Method

HIV RNA was extracted from 500 μL of each of the 859 stored screening plasma samples using automated extraction methods with a NucliSENS easyMAG (bioMérieux). One-step RT-PCR was performed in triplicate using 4 μL of sample extract per amplification. A second-round PCR amplification was then performed using customized primers to allow multiplexing (48 samples per sequencing run). PCR amplifications were then quantified. Each amplification was combined in equal proportions with the others to a concentration of 2×10^{12} DNA molecules per sample. This combined set of PCR products then underwent emulsion PCR and deep sequencing with a GS-FLX. A detailed methodology has been published [333,348] and is detailed in previous chapters.

In addition, a second round PCR amplification was also performed using the same triplicate RT-PCR template. These PCR products underwent individual standard, population-based sequencing on an ABI 3730 XL DNA analyzer according to previously described methods [325,333,348,366].

6.2.3 Bioinformatic Analyses

The false positive rate (FPR) cutoff for geno2pheno tropism assignments had previously been optimized and validated in the maraviroc treatment-experienced trials, as had the cutoff for the percentage of non-R5 variants needed for a sample to be classified as non-R5 [335–337,348]. A sample was considered R5 if fewer than 2% of the variants detected using deep sequencing fell below a geno2pheno FPR of 3.5 [335]. Population-based V3 sequencing used a geno2pheno FPR cutoff of 5.75 [336].

6.2.4 Ethics Statement

Written, informed consent was obtained from all individuals, including consent to allow other tropism testing to be performed on their samples. The University of British Columbia-Providence Health Care Research Ethics Board reviewed the research project and granted ethical approval. All data were analyzed anonymously.

6.2.5 Data Analysis

The maraviroc BID arm was the primary dataset for assessing the utility of deep sequencing. The efavirenz arm served as a comparator. The maraviroc QD arm was also examined as a complementary analysis. Unless otherwise stated, any reference to maraviroc should be taken as a reference to maraviroc BID.

Virologic outcomes examined included the viral load change from baseline, the percentage of patients with virologic suppression, and a time to a change in a patient's Trofile result from R5 to non-R5 (i.e., a tropism "switch"). Where data were missing, the last observation was carried forward, except in the case of the percentage of patients with a pVL <50 copies, where missing values were imputed to be >50 ("failures"). Deep sequencing was also compared to the performance of both ESTA and standard population-based sequencing in the same dataset.

Differences between tropism groups (R5 versus non-R5) were tested for statistical significance using three tests. The Mann-Whitney test tested for statistically significant differences between median measurements, such as median pVL declines. The Fisher's exact test examined differences in the proportions of patients, such as differences in virologic suppression or clade. The log-rank test examined differences in the Kaplan-Meier curves for tropism changes. No statistical comparisons between the populations defined by Trofile and deep sequencing were performed because these populations were not independent.

6.3 Results

6.3.1 Patient Characteristics

Baseline characteristics of patients stratified by deep sequencing tropism result at screening are largely similar to the original MERIT population [290]. Those found to have non-R5 HIV by deep sequencing were more likely to be white, MSM, infected with clade-B HIV, and have lower CD4 counts than those found to have R5 HIV by deep sequencing, though these differences were relatively minor (Table 6.1).

6.3.2 Identification of Non-R5 Screening Samples Using Deep Sequencing

Deep sequencing generated a mean of 5002 sequences per sample (median: 4529; Inter-quartile Range [IQR]: 3715 – 6024). Sequence depth did not have a discernable impact on deep sequencing's sensitivity or ability to predict virologic outcomes (data not shown). Overall, re-screening MERIT patients using deep sequencing classified an additional 10% of maraviroc BID recipients (35/347) as being unlikely to respond to their regimens due to the presence of ≥2% non-R5 virus prior to treatment. Similarly, 13% (22/166) in the maraviroc QD arm, and 9% (30/346) in the efavirenz arm were classified as having non-R5 HIV by deep sequencing.

Samples screened non-R5 by deep sequencing had non-R5 variants at a median proportion of 20.9% (IQR: 5.4 – 44.1). Samples screened R5 had a median of 0% non-R5 HIV (IQR: 0 – 0). Seventy-four percent of patients (511/693) treated with maraviroc BID or efavirenz had no detectable non-R5 variants at screening by deep sequencing. Additionally, 60% of all non-R5 samples had >10% non-R5 variants by this method, despite having been already pre-screened with the Trofile assay, which has a reported cut-off of 10% non-R5 virus [334]. There were a total of 94 maraviroc recipients with detectable non-R5 virus by deep sequencing. When the non-R5 prevalence was extended to the absolute amount of non-R5 at screening, these patients had a median non-R5 viral load of 2.9 $\log_{10}$ copies/mL (IQR: 2.2 – 3.5).

Table 6.1: Baseline Patient Characteristics in the MERIT trial

	Combined MVC BID and EFV arms (N=693)	Deep Sequencing Non-R5 (N=65)	Deep Sequencing R5 (N=628)	Statistical significance p value
Age — median (range)	36 (18 – 77)	39 (21 – 68)	36 (18 – 77)	n.s. 0.09
Male sex - no. (%)	503 (73)	53 (82)	450 (72)	n.s. 0.11
Race or ethnicity — no. (%)	White - 394 (57); Black - 238 (34), Asian, other – 61 (9)	White - 45 (69); Black - 13 (20); Asian, other – 7 (11)	White - 349 (56); Black - 225 (36); Asian, other – 54 (9)	0.04 (white versus non-white)
Clade — no. (%)	B 414 (60); C – 205 (30); Other - 74 (11)	B – 48 (74); C – 10 (15); Other - 7 (11)	B - 366 (58); C - 195 (31); Other - 67 (11)	0.02 (B versus non-B)
Mode of Transmission — no. (%)	Het - 328 (47); MSM - 292 (42); IDU - 48 (7); Other - 25 (4)	Het - 23 (36); MSM - 37 (58); IDU - 0 (0); Other - 4 (6)	Het- 305 (49); MSM - 254 (40); IDU - 48 (8); Other - 21 (3)	0.01 (MSM versus non-MSM)
Median baseline pVL — log_{10} copies/mL (IQR)	5.0 (4.5 – 5.3)	5.0 (4.5 – 5.3)	4.9 (4.4 – 5.2)	n.s. 0.54
Median baseline CD4 count — cells/mm3 (IQR)	251 (183 – 323)	236 (135 – 300)	252 (185 – 327)	0.03

Table 6.1: Baseline Patient Characteristics in the MERIT Trial. Baseline patient characteristics for the study population. MERIT participants are shown for the group as a whole (N=693), as well as for the groups who were classified as R5 or non-R5 by deep sequencing. The difference in baseline characteristics between the R5 and non-R5 groups were also tested for their statistical significance. Het — heterosexual; MSM — Men who have sex with men; IDU — Injection Drug Use; pVL — plasma viral load; IQR — interquartile range; n.s. — not significant.

6.3.3 Viral Load Decline from Baseline

Overall viral load declines from baseline through 96 weeks are shown for both arms in Figures 6.1 and 6.2, with patients grouped according to their deep sequencing result. Patients with R5 results had similar responses in both the maraviroc and efavirenz arms (Figure 6.1), but patients with non-R5 results had expectedly poorer responses in the maraviroc arm compared to the efavirenz arm (Figure 6.2).

Maraviroc recipients with R5 HIV by deep sequencing showed a median 2.7 $\log_{10}$ decline in pVL from baseline to week 8 (IQR: 2.3 – 3.1), while the non-R5 group had a smaller decline: 2.3 $\log_{10}$ (IQR: 1.9 – 2.6), p<0.0001. The efavirenz arm had similar virologic responses as the R5-infected maraviroc recipients, regardless of tropism assessment by deep sequencing: 2.8 $\log_{10}$ (IQR: 2.4 – 3.2) for R5, and 2.9 $\log_{10}$ (IQR: 2.5 – 3.2) for non-R5, p=0.56.

6.3.4 Virologic Suppression

The larger pVL changes observed when patients were classified using the deep sequencing method was also reflected in the percentage of patients who achieved an undetectable viral load at 48 weeks. Where deep sequencing had indicated R5 HIV at screening, a total of 67% of maraviroc recipients (208/312) had a pVL <50 HIV RNA copies/mL at week 48 (i.e., virologic suppression). In contrast, only 46% of non-R5-infected maraviroc recipients (16/35) achieved week 48 virologic suppression, p=0.02 (Figure 6.3). In terms of non-R5 viral load, the percentage of maraviroc recipients with week 48 virologic suppression was: 68% (173/255) of those with <1 $\log_{10}$ non-R5 copies/mL; 77% (10/13) with 1-2 $\log_{10}$; 56% (22/39) with 2-3 $\log_{10}$; 52% (14/27) with 3-4 $\log_{10}$; and 38% (5/13) of those with >4 $\log_{10}$ non-R5 copies/mL. In the efavirenz arm, suppression was 69% (219/316) in those with R5 HIV — similar to the maraviroc arm. This was 70% (21/30) in efavirenz recipients with non-R5 HIV by deep sequencing, p=n.s.. The percentages of patients achieving virologic suppression for both arms are shown in Figure 6.3, with data to week 96.

Figure 6.1: Median Decline in Plasma Viral Load from Baseline in Patients Screened with R5 HIV by Deep Sequencing Who Received Either Maraviroc Twice-Daily or Efavirenz

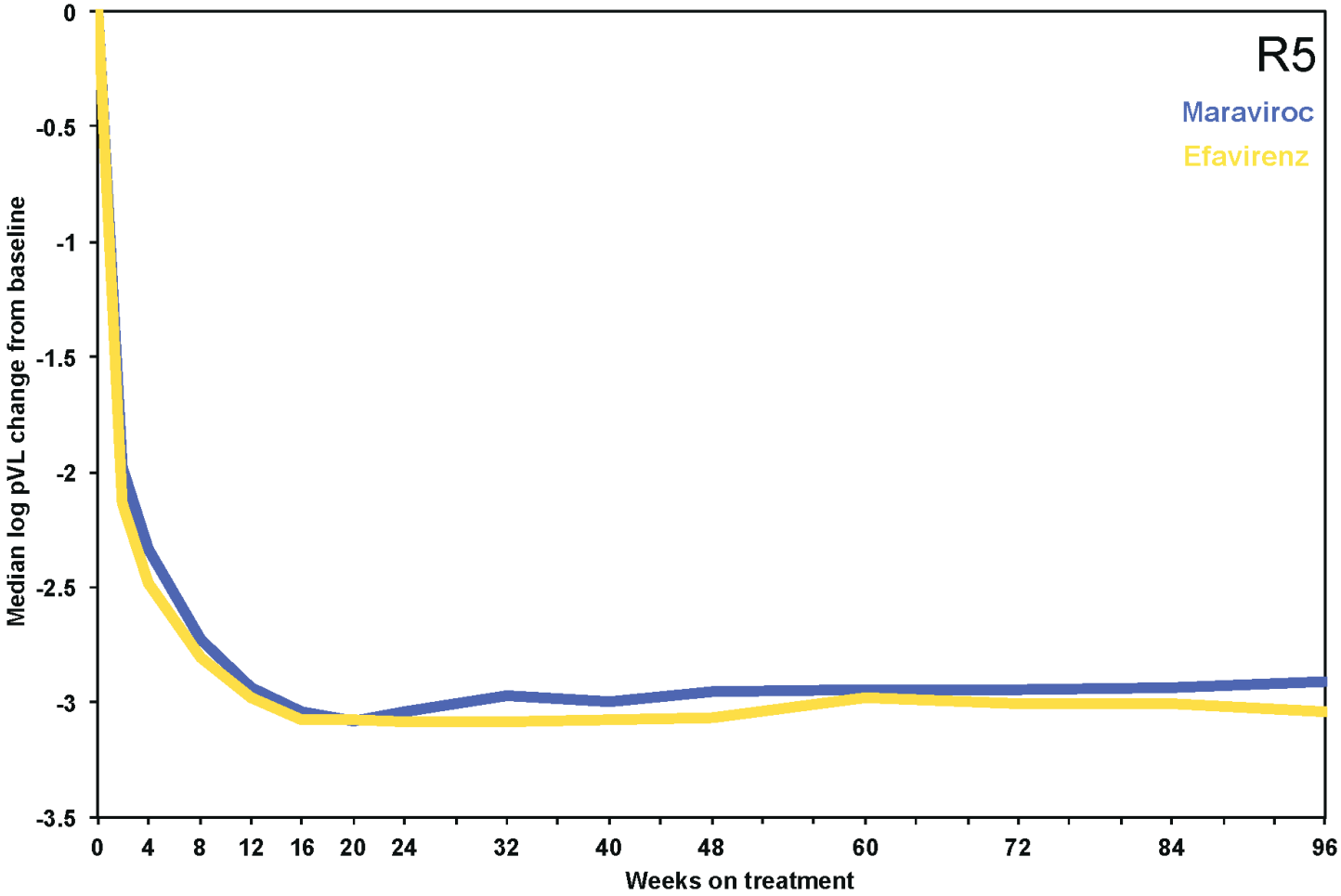


Figure 6.1: Median Decline in Plasma Viral Load from Baseline in Patients Screened with R5 HIV by Deep Sequencing Who Received Either Maraviroc Twice-Daily or Efavirenz. The $\log_{10}$-transformed decline in plasma viral load for R5 HIV-infected patients. The blue line indicates patients receiving maraviroc BID (N=312), and the yellow line indicates those receiving efavirenz (N=316). With screening by deep sequencing, both groups had a median pVL decline from baseline of approximately 3 $\log_{10}$ HIV RNA copies/mL, which was sustained to week 96.

Figure 6.2: Median Decline in Plasma Viral Load from Baseline in Patients Screened with Non-R5 HIV by Deep Sequencing Who Received Either Maraviroc Twice-Daily or Efavirenz

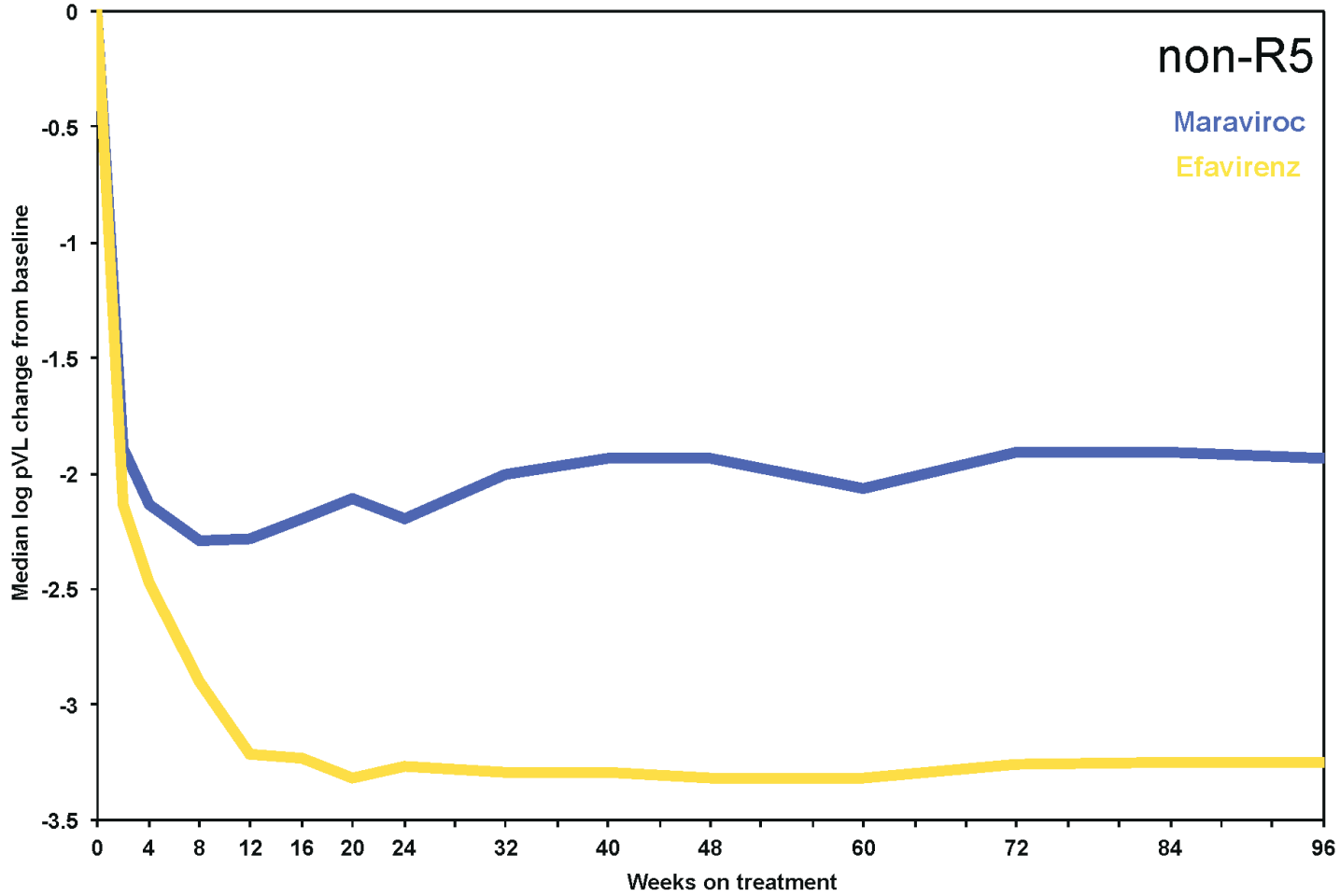


Figure 6.2: Median Decline in Plasma Viral Load from Baseline in Patients Screened with Non-R5 HIV by Deep Sequencing Who Received Either Maraviroc Twice-Daily or Efavirenz. The log_{10}-transformed decline in plasma viral load for non-R5 HIV-infected patients. The blue line indicates patients receiving maraviroc BID (N=35), and the yellow line indicates those receiving efavirenz (N=30). With screening by deep sequencing, those found to have non-R5 HIV had lower pVL declines from baseline when treated with maraviroc BID versus efavirenz.

Figure 6.3: Percentage of Maraviroc Twice-Daily & Efavirenz Recipients with Plasma Viral Loads Less than 50 Copies/mL with R5 HIV at Screening by Deep Sequencing

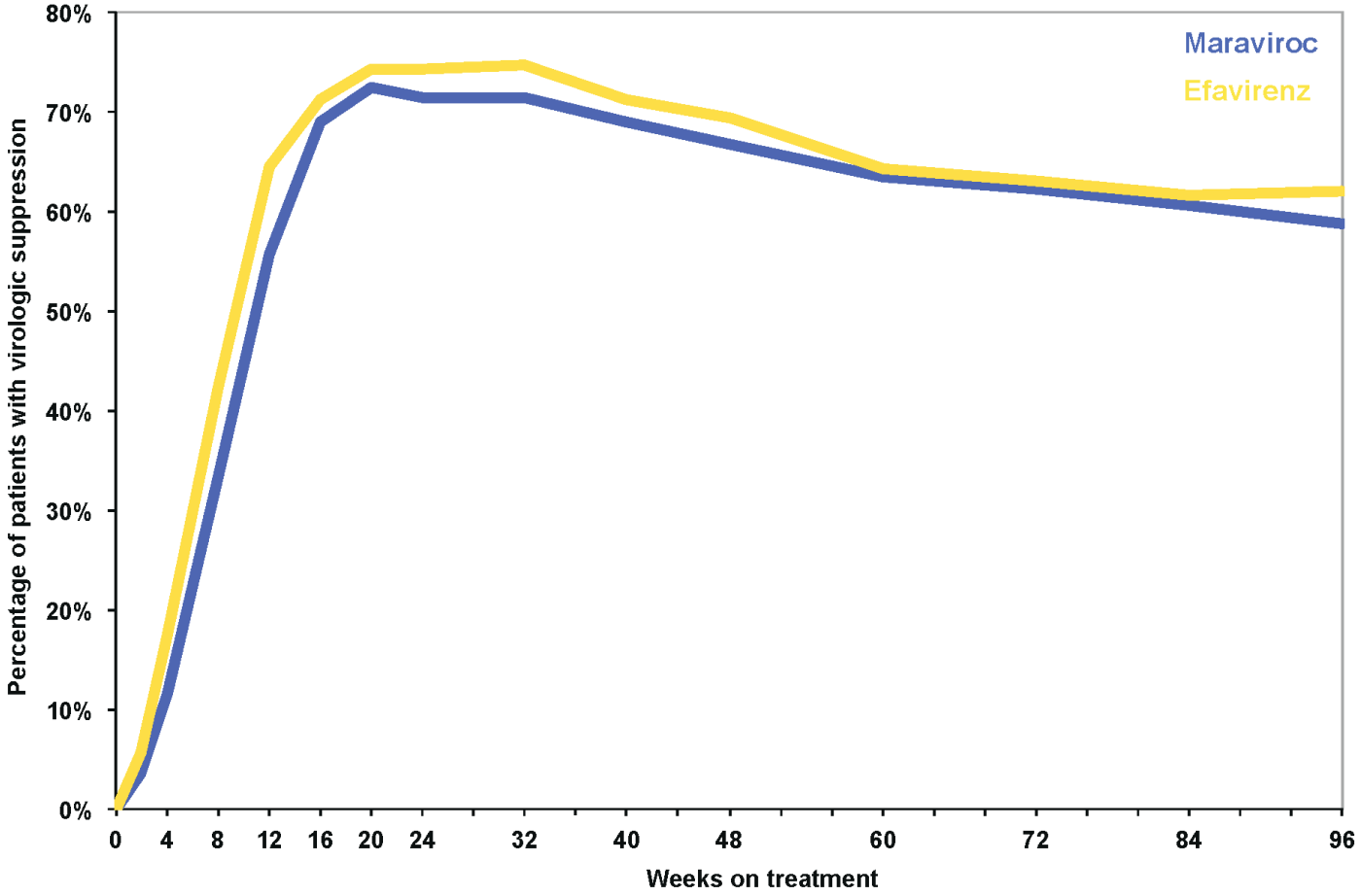


Figure 6.3: Percentage of Maraviroc Twice-Daily & Efavirenz Recipients with Plasma Viral Loads Less than 50 Copies/mL with R5 HIV at Screening by Deep Sequencing. The black line indicates R5 HIV-infected patients receiving maraviroc BID (N=312), and the dashed-dotted line indicates those receiving efavirenz (N=316). Similar percentages of patients had virologic suppression at week 48 in the two treatment arms when patients with non-R5 HIV at screening by deep sequencing were excluded. The percentage of patients with virologic suppression at week 48 is indicated for both the maraviroc (MVC) and efavirenz (EFV) arms.

6.3.5 Non-Inferiority Analysis

In the original MERIT study, maraviroc was found to be inferior to efavirenz. This analysis was based on comparing the differences in the percentage of patients achieving virologic suppression <50 copies/mL in the maraviroc and efavirenz arms, when these patients were screened by the original Trofile assay. The criterion was that the lower bound of the 97.5% confidence interval for the difference between arms must not fall below -10 for the maraviroc arm, but this criterion was not met. In contrast, when patients in the current study were tested using deep sequencing, the lower bound of the 97.5% confidence interval for the difference between arms was -8.67; less than the pre-specified minimum value of -10 for determining non-inferiority of maraviroc at week 48 (Table 6.2).

Table 6.2: Non-Inferiority Analysis between the Maraviroc & Efavirenz Arms

	Number and proportion of virologic successes at 48 weeks								
	Maraviroc BID arm			Efavirenz arm			Stratified		
Assay result	n	N	%	n	N	%	Raw diff (MVC - EFV)	Adj Diff	97.5% LCB
DeepSeq R5	210	312	67.31	217	316	68.67	-1.36	-1.48	-8.67
DeepSeq non-R5	17	35	48.57	21	30	70.00	-21.43	-42.19	-60.71
ESTA R5	205	300	68.33	196	290	67.59	0.75	0.17	-7.21
ESTA non-R5	22	47	46.81	42	56	75.00	-28.19	-31.15	-48.87
Trofile R5	227	347	65.42	238	346	68.79	-3.37	-3.73	-10.61

Table 6.2: Non-Inferiority Analysis between the Maraviroc & Efavirenz Arms. Non-inferiority analysis comparing the maraviroc BID arm to the efavirenz arm. The lower confidence bound of the difference between drug arms for R5 screened participants was smaller than -10%, indicating non-inferiority between arms. MVC — maraviroc; EFV — efavirenz; Adj Diff — adjusted difference; LCB — lower confidence bound; ESTA — enhanced sensitivity Trofile assay; DeepSeq — Deep Sequencing

This analysis also confirms the poor virologic response among maraviroc recipients screened with non-R5 HIV by deep sequencing, compared to those who received efavirenz. Together, these analyses suggest that had patients been screened with deep sequencing rather than the original Trofile assay, the maraviroc BID arm would have likely been found to be non-inferior to the efavirenz arm.

6.3.6 Changes in HIV Tropism

Maraviroc administration unmasks and can select non-R5 virus that was present prior to maraviroc administration [316]. Maraviroc-recipients with non-R5 HIV by deep sequencing were more likely to change phenotypic tropism over the course of the study compared to those with R5 HIV by deep sequencing ($p<0.0001$). Of those with non-R5 HIV, 43% (15/35) changed their Trofile result from R5 to non-R5 between screening and 96 weeks, versus only 7% (23/312) of the deep sequencing R5 group (Figure 6.4).

In the non-R5 group, patients switched tropism a mean of 5 weeks after beginning treatment. This was earlier than the 17 weeks seen in the R5 group. Maraviroc recipients who changed tropism also had a higher proportion of non-R5 variants present pre-treatment according to deep sequencing, with a median of 0.8% non-R5 variants (IQR: 0.0 – 7.4%) versus 0% (IQR: 0 – 0%) for those who did not change tropism. Deep sequencing was able to detect at least low levels (>0%) of non-R5 HIV in a majority, 61%, of maraviroc recipients who switched tropism, versus 23% of those who did not switch tropism.

6.3.7 Effects of HIV Subtype

For all patients analysed in the current study, 60% had HIV-1 clade-B, 29% had clade-C, and 11% had other clades of HIV. Non-R5 tropism seemed to be overrepresented amongst clade-B-infected individuals, with 74% of the deep sequencing non-R5 group consisting of clade-B-infected patients, higher than the overall clade-B composition of 60%, $p=0.02$. Conversely, clade-C was underrepresented amongst non-R5-infected patients, at 15%, $p=0.001$.

Figure 6.4: Time to Change in Tropism for Maraviroc Twice-Daily Recipients

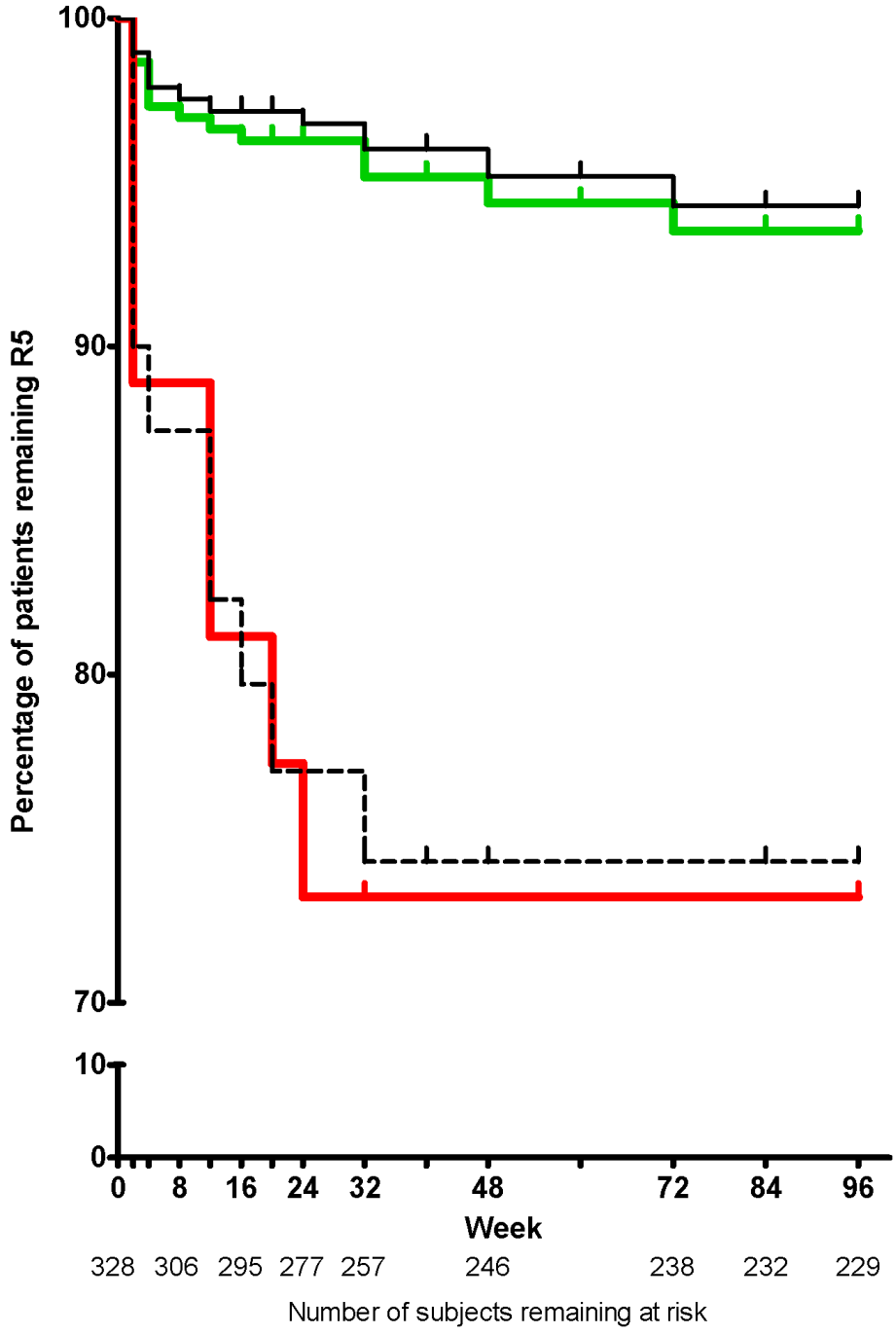


Figure 6.4: Time to Change in Tropism for Maraviroc Twice-Daily Recipients.

This analysis examined the likelihood of a change from an original Trofile assay result of R5 to non-R5 over the course of the study. The upper green line indicates patients screened with R5 HIV by deep sequencing (N=312). The lower red line indicates those screened with non-R5 HIV by deep sequencing (N=35). The upper solid black and lower dotted black lines indicate the ESTA-R5 (N=300) and ESTA-non-R5 (N=47) groups. Patients screened with non-R5 HIV by either assay were more likely to change Trofile results to non-R5 during the study. The numbers of patients remaining at risk for a change in their Trofile result are shown below the week numbers. The scale of the Y axis has been modified for display purposes.

Global concordance in the entire study population between deep sequencing and ESTA was 79% in the clade B-infected population and 87% in the non-clade B-infected population. Importantly, both deep sequencing and ESTA had similar performance in predicting virologic outcome in both clade-B non-clade-B-infected patients (Tables 6.3 & 6.4).

Table 6.3: Virologic Outcomes in Maraviroc Recipients Infected with Subtype B HIV-1, Stratified by Tropism Assessment by Deep Sequencing & the Enhanced Sensitivity Trofile Assay

Assay	Tropism Call	Week 8 pVL decline, log_{10} scale Median (IQR)	Week 48 virologic suppression %, n/N
Deep sequencing	R5	2.7 (2.4 – 3.1)	72% (132/183)
	Non-R5	2.3 (2.0 – 2.6)	46% (12/26)
ESTA	R5	2.7 (2.3 – 3.0)	73% (128/176)
	Non-R5	2.6 (2.1 – 2.9)	48% (16/33)

Table 6.4: Virologic Outcomes in Maraviroc Recipients Infected with Non-Subtype B HIV-1, Stratified by Tropism Assessment with Deep Sequencing & the Enhanced Sensitivity Trofile Assay

Assay	Tropism Call	Week 8 pVL decline, log_{10} scale Median (IQR)	Week 48 virologic suppression %, n/N
Deep sequencing	R5	2.8 (2.2 – 3.1)	59% (76/129)
	Non-R5	1.9 (1.9 – 2.4)	44% (4/9)
ESTA	R5	2.8 (2.2 – 3.1)	60% (75/124)
	Non-R5	2.2 (1.7 – 2.9)	36% (5/14)

6.3.8 Comparison of Deep Sequencing to the Enhanced Sensitivity Trofile Assay & Population-Based Sequencing

Maraviroc recipients rescreened with R5- and non-R5 HIV using ESTA had week 8 pVL declines of 2.7 (IQR: 2.3 – 3.1) and 2.4 (IQR: 1.9 – 3.0), respectively, similar to the deep sequencing results. Likewise, the percentage of patients who achieved virologic suppression on maraviroc was similar regardless of the assay used to determine tropism (Figures 6.5 & 6.6).

When compared with each other, deep sequencing and ESTA had a global concordance of 82%. Perhaps surprisingly, only 22 samples of the 693 total (3%) were called non-R5 by both methods, or only 15% of the 146 samples called non-R5 by either method. Consequently, both assays had low sensitivity relative to the other. Deep sequencing had 21% sensitivity and 93% specificity using ESTA as a reference; ESTA had 34% sensitivity and 87% specificity using deep sequencing as a reference. Despite this, the groups called R5 and non-R5 by either method had similar virologic outcomes regardless of the assay. Retrospective screening by ESTA identified 14% of maraviroc recipients (47/347) as having non-R5 HIV. This was 16% (56/346) in the efavirenz arm. Overall, the ESTA-non-R5 group had a median of 0% non-R5 HIV (IQR: 0 – 0.8%; mean = 7.4%), according to deep sequencing using the geno2pheno algorithm; the ESTA-R5 group also had a median of 0% (IQR: 0 – 0%; mean = 2.2%).

Deep sequencing was also compared to population-based sequencing where available. Population-based sequencing was concordant with deep sequencing in 93% of cases (638/688), and gave 54% sensitivity relative to deep sequencing. Samples identified by population-based sequencing as non-R5 had a median of 9.1% non-R5 variants in their deep sequencing result (IQR: 0.7 – 41.0%; mean = 26.3%). Virologic responses of patients grouped by discordance of deep sequencing with ESTA or population-based sequencing are shown in Table 6.5.

Figure 6.5: Median Decline in Plasma Viral Load from Baseline in Maraviroc Twice-Daily Recipients with Screening by Deep Sequencing & the Enhanced Sensitivity Trofile Assay

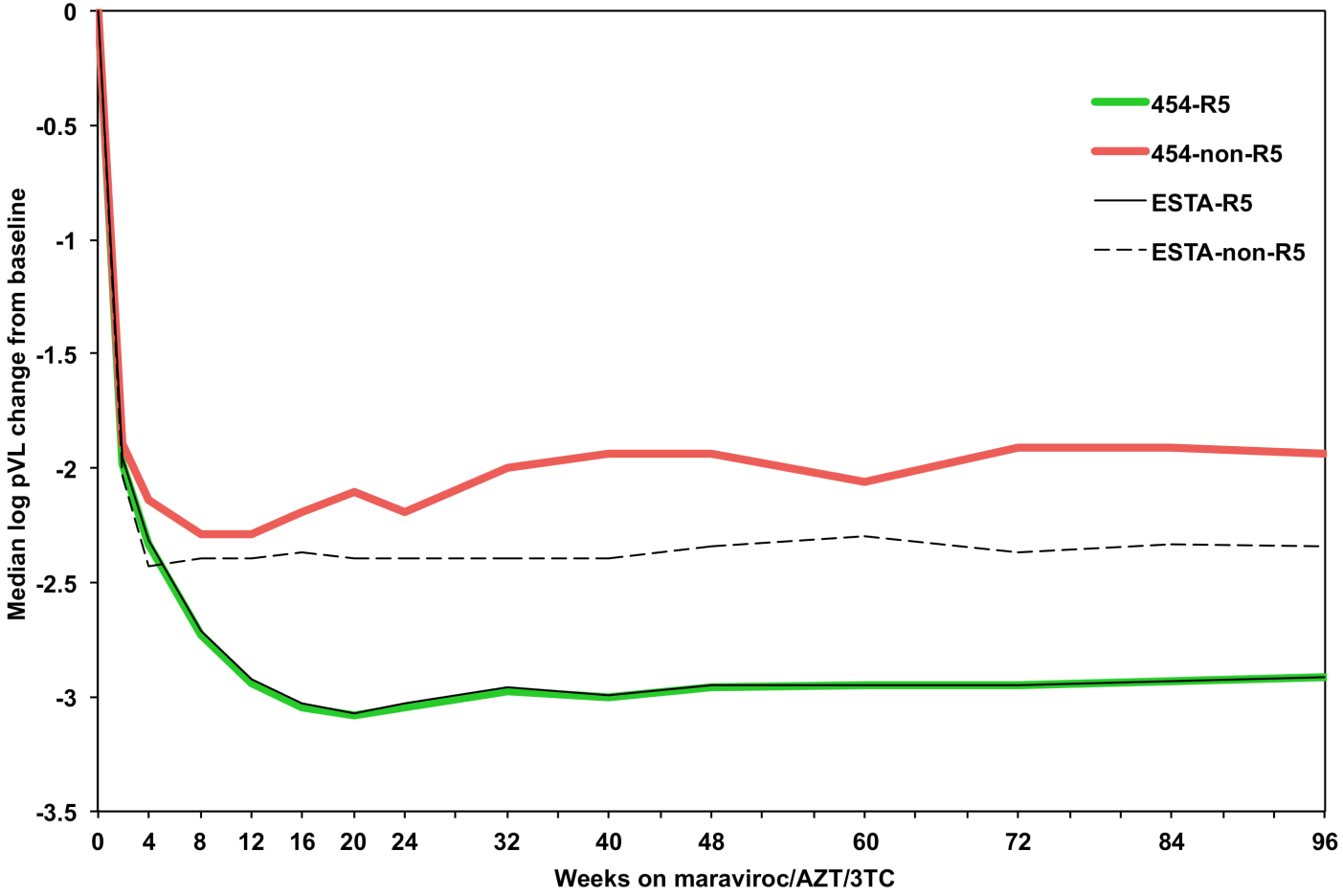


Figure 6.5: Median Decline in Plasma Viral Load from Baseline in Maraviroc Twice-Daily Recipients with Screening by Deep Sequencing & the Enhanced Sensitivity Trofile Assay. The green line indicates maraviroc BID recipients screened with R5 HIV by deep sequencing (N=312). The red line indicates those screened with non-R5 HIV by deep sequencing (N=35). The solid black and dashed black lines indicate the ESTA-R5 (N=300) and ESTA-non-R5 (N=47) groups, respectively. Deep sequencing and ESTA performed similarly in terms of distinguishing between virologic responders and non-responders on maraviroc.

Figure 6.6: Percentage of Maraviroc Twice-Daily Recipients with Plasma Viral Loads below 50 Copies/mL with Screening by Deep Sequencing & the Enhanced Sensitivity Trofile Assay

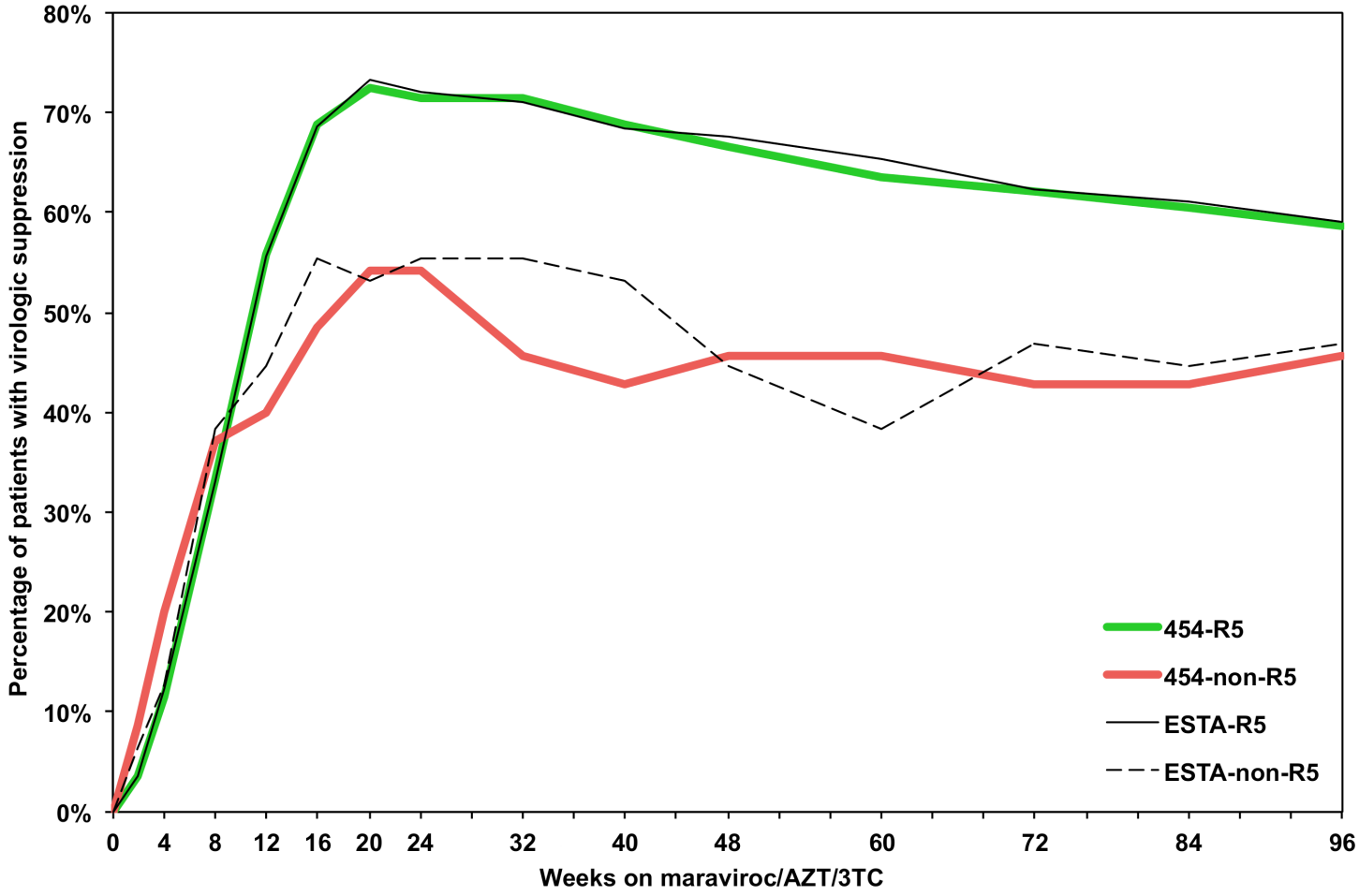


Figure 6.6: Percentage of Maraviroc Twice-Daily Recipients with Plasma Viral Loads below 50 Copies/mL with Screening by Deep Sequencing & the Enhanced Sensitivity Trofile Assay. The green line indicates maraviroc BID recipients screened with R5 HIV by deep sequencing (N=312). The red line indicates those screened with non-R5 HIV by deep sequencing (N=35). The solid black and dashed black lines indicate the ESTA-R5 (N=300) and ESTA-non-R5 (N=47) groups

Table 6.5: Overall Virologic Responses of Maraviroc Recipients Grouped by Discordance between Deep Sequencing & the Enhanced Sensitivity Trofile Assay or Population-Based Sequencing

Concordance or Discordance DeepSeq Result versus Other assay result	Median week 8 log_{10} pVL change from baseline, (IQR) Deep Sequencing versus:		Percentage of patients with week 48 virologic suppression, (n) Deep Sequencing versus:	
	ESTA	Population-based sequencing	ESTA	Population-based sequencing
R5/ R5	2.7 (2.3 – 3.1)	2.7 (2.3 – 3.1)	68% (188/276)	67% (202/301)
R5/ Non-R5	2.6 (2.2 – 3.1)	2.8 (2.1 – 3.1)	56% (20/36)	50% (4/8)
Non-R5/ R5	2.4 (2.2 – 2.7)	2.4 (1.9 – 2.7)	63% (15/24)	47% (8/17)
Non-R5/ Non-R5	1.9 (1.3 – 2.1)	2.1 (1.9 – 2.3)	9% (1/11)	44% (8/18)

Table 6.5: Overall Virologic Responses of Maraviroc Recipients Grouped by Discordance between Deep Sequencing & the Enhanced Sensitivity Trofile Assay or Population-Based Sequencing. pVL — plasma viral load; IQR — interquartile range; ESTA — enhanced sensitivity Trofile assay; DeepSeq — Deep Sequencing

In addition, virologic responses were compared with classification by deep sequencing versus ESTA or population-based sequencing. Overall, where screening assays differed, there was no clear indication as to which assay was the "gold standard". Indeed, deep sequencing, ESTA, and population-based sequencing all performed quite similarly in terms of predicting virologic response to maraviroc in this population. There was a possible trend towards slightly superior predictions by deep sequencing (Figures 6.7 & 6.8).

Figure 6.7: Declines in Plasma Viral Load from Baseline in Patients with Concordant & Discordant Results between Tropism Assays

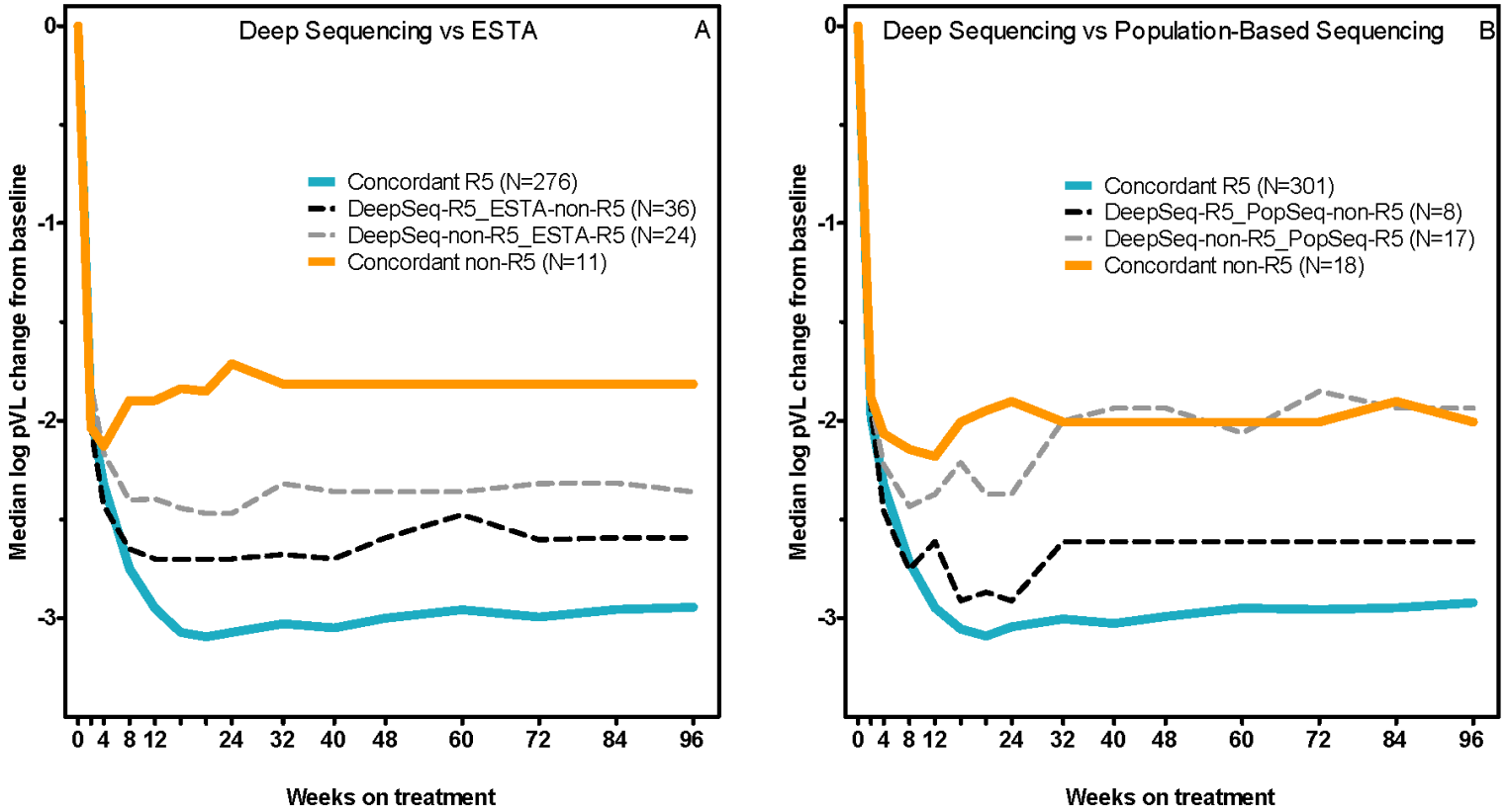


Figure 6.7: Declines in Plasma Viral Load from Baseline in Patients with Concordant & Discordant Results between Tropism Assays. Left: Panel A shows the plasma viral load decline on maraviroc for patient groups where deep sequencing and ESTA gave the same (concordant) and different (discordant) tropism calls. The solid blue and solid orange lines indicate the concordant R5 and non-R5 groups, respectively. The black dashed line indicates the group called R5 by deep sequencing but non-R5 by ESTA. The grey dashed line indicates the group called non-R5 by deep sequencing but R5 by ESTA. Right: Panel B shows the median plasma viral load change from baseline in maraviroc recipients. Patients are grouped where deep sequencing and population-based sequencing gave the same (concordant) or different (discordant) tropism calls. The solid blue and solid orange lines indicate the concordant R5 and non-R5 groups, respectively. The black dashed line indicates the group called R5 by deep sequencing but non-R5 by population-based sequencing. The grey dashed line indicates the group called non-R5 by deep sequencing but R5 by population-based sequencing.

Figure 6.8: Proportion of Patients with Plasma Viral Loads below 50 Copies/mL in Groups with Concordant & Discordant Results between Tropism Assays

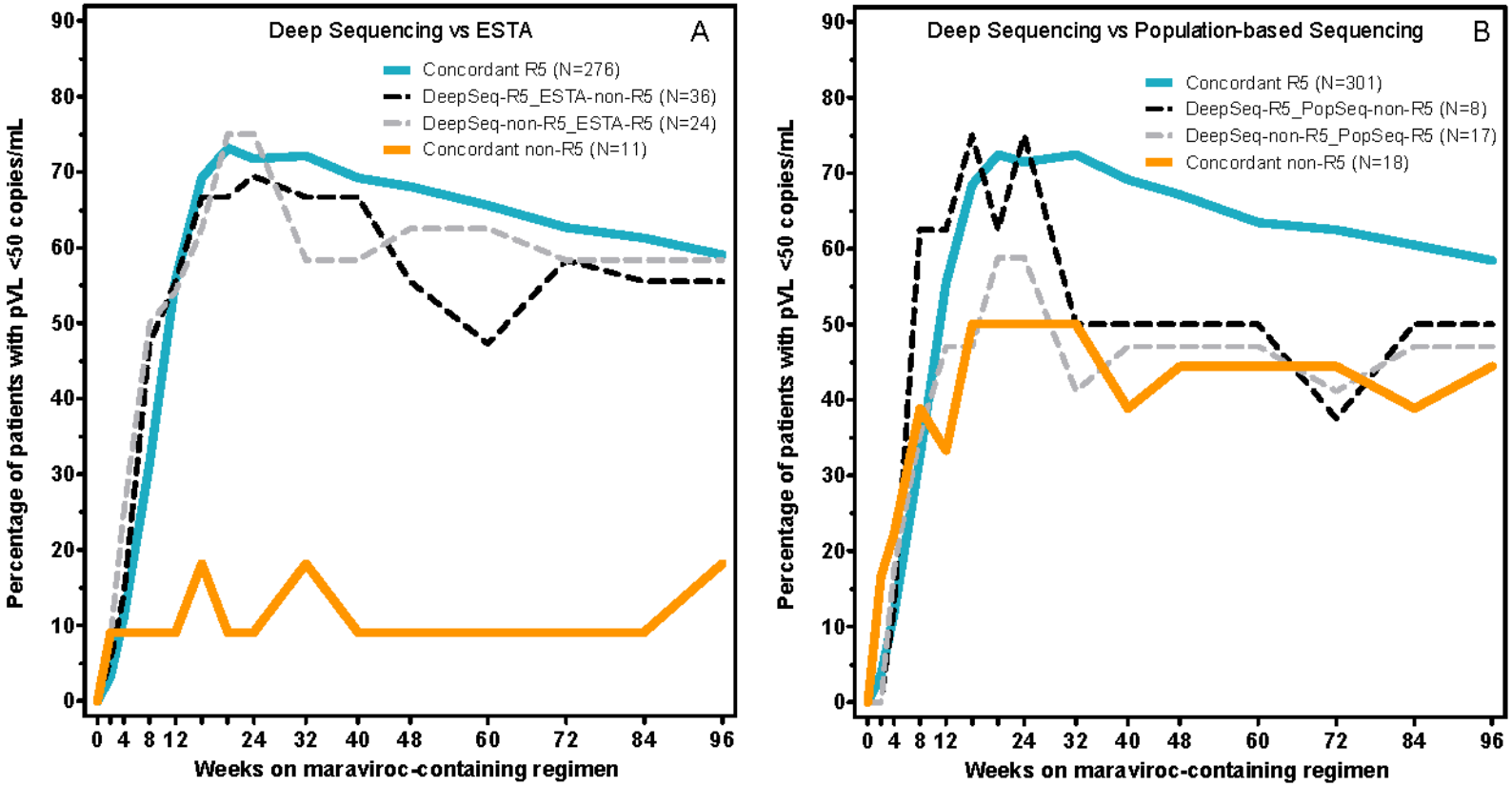


Figure 6.8: Proportion of Patients with Plasma Viral Loads below 50 Copies/mL in Groups with Concordant & Discordant Results between Tropism Assays. Left: Panel A shows the percentage of patients on maraviroc who had viral loads below 50 copies/mL. Patients are grouped where deep sequencing and ESTA gave the same (concordant) or different (discordant) tropism calls. The solid blue and solid orange lines indicate the concordant R5 and non-R5 groups, respectively. The black dashed line indicates the group called R5 by deep sequencing but non-R5 by ESTA. The dashed line indicates the group called non-R5 by deep sequencing but R5 by ESTA. Right: Panel B shows the percentage of patients on maraviroc who had viral loads below 50 copies/mL. Patients are grouped where deep sequencing and population-based sequencing gave the same (concordant) or different (discordant) tropism calls. The solid blue and solid orange lines indicate the concordant R5 and non-R5 groups, respectively. The dashed black line indicates the group called R5 by deep sequencing but non-R5 by population-based sequencing. The dashed grey line indicates the group called non-R5 by deep sequencing but R5 by population-based sequencing.

6.3.9 Maraviroc Once-Daily Arm

The patients who were randomized into the maraviroc QD arm were also examined with deep sequencing (N=166). This dataset served as an independent validation of the deep V3 sequencing method. The maraviroc QD arm was originally discontinued partway through the MERIT study due to a protocol-defined lack of demonstrated non-inferiority to efavirenz. Maraviroc QD recipients were then allowed to switch to maraviroc BID for the remainder of the study. The performance of deep sequencing as a screening tool for tropism was assessed in this population (Figure 6.9). Analyses were performed where responses were censored or uncensored after patients switched to maraviroc BID. The week 8 $\log_{10}$ pVL declines from baseline were similar between the maraviroc QD and BID arms in the uncensored analysis. The median decline of those screened as having R5 HIV (N=144) was 2.8 $\log_{10}$ (IQR: 2.4 – 3.1) versus 2.6 $\log_{10}$ (IQR: 1.3 – 3.0) for those with non-R5 HIV (N=22). Note that 26 patients in the R5 group (18%) and 6 in the non-R5 group (27%) had discontinued therapy or switched to maraviroc BID by week 8. Viral load declines from baseline for the uncensored groups are shown in Figure 6.9. The R5-group censored for those remaining on maraviroc QD is also shown and gave similar results (Figure 6.9).

6.4 Discussion & Conclusions

This study represents the first large clinical comparison of two highly sensitive HIV tropism assays: deep sequencing and ESTA, and the largest application of deep sequencing in antiretroviral-naïve patients to date. Retrospective screening by deep sequencing, with removal of patients classified with non-R5 HIV, led to similar rates of week 48 virologic suppression between the maraviroc BID and efavirenz arms. Maraviroc recipients screened with R5 HIV by this approach had larger on-treatment pVL declines, were more likely to achieve virologic suppression, and were less likely to change tropism than those screened with non-R5 virus.

Figure 6.9: Median Decline in Plasma Viral Load from Baseline in Maraviroc Once-Daily Recipients

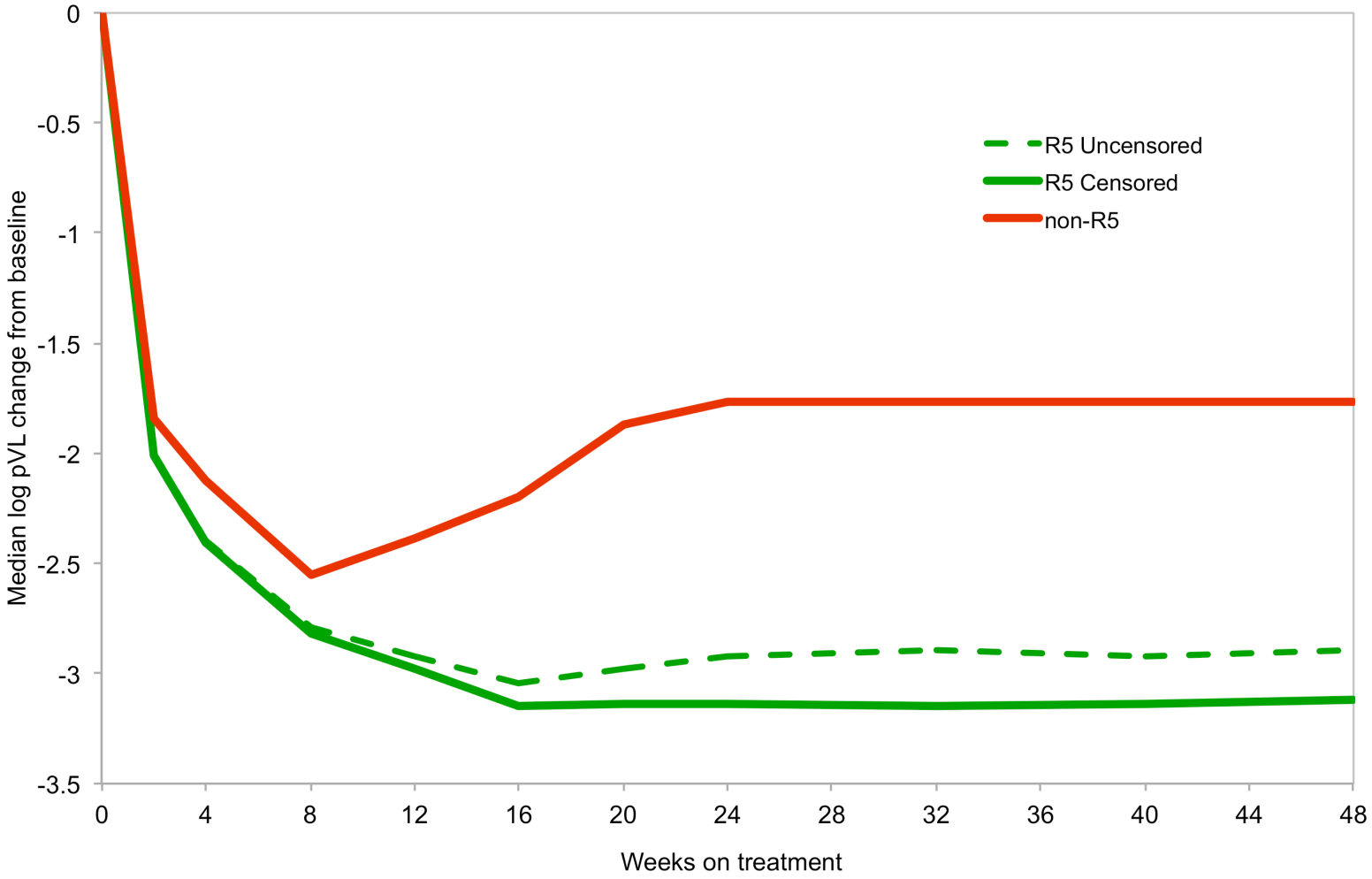


Figure 6.9: Median Decline in Plasma Viral Load from Baseline in Maraviroc Once-Daily Recipients. The viral load changes for patients in the discontinued maraviroc once-daily (QD) arm. The red line indicates those for whom deep sequencing indicated non-R5 HIV at screening. The solid green line indicates those patients classified as having R5 HIV at screening, with patients censored when they were switched to maraviroc twice-daily dosing. The dashed green line indicates the same patients as the R5 group, but with continued follow-up after switching to twice-daily from once-daily dosing.

Deep sequencing also had similar performance to ESTA, which is widely used in the clinic. Virologic responses were similar between groups that had discordant results by either assay, suggesting that neither assay is significantly more "correct" than the other. Interestingly, the decline in viral load from baseline was greater than 2 $\log_{10}$ copies/mL even in the maraviroc-treated non-R5 group. This is likely due to the activity of the background zidovudine-lamivudine, and perhaps some residual activity of maraviroc.

The additional clinical utility of deep sequencing over standard population-based sequencing was not clearly demonstrated in this study, despite a possible trend in a previous study in treatment-experienced patients [348]. In fact, concordance was over 90% between the methods in the current study, though the sensitivity of population-based sequencing remained low relative to deep sequencing. A common critique of bioinformatic algorithms for HIV tropism is that most are trained primarily on clade B sequences. However, the deep sequencing genotypic assay presented here performed similarly to the phenotypic ESTA assay in MERIT, including in non-clade-B-infected patients, lending confidence to the utility of this approach in such populations (see also [378]).

Some limitations of this study and the use of deep sequencing should be acknowledged. The MERIT trial itself only included patients pre-screened as having R5 HIV by the original Trofile assay, so an analysis of maraviroc treatment in an antiretroviral-naïve population infected with non-R5 virus by the Trofile assay was not possible, though an analysis of deep sequencing in a non-R5 treatment-experienced trial has been published [348] and is presented in Chapter 3. The pre-screening of these patients may also have diminished the ability to demonstrate improved tropism prediction of any assay over any other given the small number of patients re-screened as having non-R5 HIV. The analysis of the maraviroc QD arm should also be examined with caution given the small number of patients continuing once-daily maraviroc treatment. Finally, the deep sequencing method itself is costly in both time and capital, which currently limits its utility in clinical settings.

Overall, deep sequencing is a useful tool for distinguishing between probable responders and non-responders to maraviroc. This high sensitivity method performed similarly to ESTA, which is

currently the most commonly used clinical phenotypic tropism assay. Had deep sequencing been used to screen patients, maraviroc would have likely been found to be non-inferior to efavirenz in the MERIT trial.

The next and final chapter of this thesis will summarize the findings of the previous chapters, and remark on conclusions and future implications that these results may have within the HIV treatment field.

Chapter 7: General Discussion & Conclusions

7.1 Thesis Summary & Overall Conclusions

This thesis described a number of aspects of HIV tropism. Next-generation sequencing was used to better evaluate coreceptor usage and its clinical relevance in patients treated with medication that specifically antagonizes the HIV-1 coreceptor, CCR5. This approach was able to accurately determine HIV tropism in a number of patient populations, including an observational cohort, and a total of four large-scale, randomized clinical trials of the CCR5 antagonist maraviroc. Deep sequencing either met or surpassed the performance of a number of alternative tropism assays, including population-based sequencing, the original Trofile assay, and the Enhanced Sensitivity Trofile assay.

This approach could both accurately quantify the coreceptor usage of patients' viral populations as well as predict the virologic responses that those patients would experience while receiving treatment with maraviroc. Next-generation sequencing could be performed using both the plasma and peripheral blood mononuclear cell compartments to test either HIV RNA or DNA. Patients with CXCR4-using HIV as assessed by next-generation sequencing had smaller viral load declines on maraviroc, were less likely to achieve a suppressed viral load, and were more likely to switch phenotypic tropism results to non-R5 while on maraviroc-containing regimens. Furthermore, in a majority of cases, the pre-treatment assessments were phylogenetically related to the viral outgrowth populations following maraviroc administration, indicating that the CXCR4-using variants detected prior to treatment with maraviroc did in fact give rise to those which would later emerge during treatment failure.

A major advantage of the use of next-generation sequencing in the treatment of HIV is the flexibility of the technique. Not only can it be used to assess viral coreceptor usage, it can also simultaneously quantify drug resistance to all known classes of antiretrovirals. Increasingly, there are opportunities to sequence the entire genome of HIV on a single next-generation sequencing platform. Moreover, the genome or exome of the host may also be sequenced in parallel with the virus, enabling clinicians to

gain insight into potential genetic factors that may influence response to therapy or risk for adverse events. This flexible and high-output approach is now standard-of-care for HIV tropism determination in many settings including Canada and much of Europe. It will likely also become increasingly used to quantify other types of drug resistance in order to further personalize and tailor HIV treatment to the individual patient.

7.2 Specific Conclusions of the Thesis

7.2.1 Conclusions for Chapter Two

Chapter 2 provided the foundation for all of the studies detailed in the subsequent chapters It established that genotypic tropism testing was viable for predicting phenotypic assay results. This study also established that higher sensitivities for detecting CXCR4-using HIV could be achieved through the use of triplicate amplification methods and deep sequencing. The bioinformatic algorithms used to infer coreceptor usage were observed to vary in their outputs with phenotypic tropism results, such that it was feasible to assign them classification cutoffs. Furthermore, both the plasma RNA-based and peripheral blood mononuclear cell DNA-based amplifications gave useful results. The methodologies were performed in patients from the observational cohort HOMER, and were validated in a small subset of patients from the MOTIVATE clinical trials of maraviroc.

7.2.2 Conclusions for Chapter Three

Chapter 3 examined an extensive dataset of patients enrolling in the phase III clinical trials of maraviroc in treatment-experienced participants. Using 75% of the dataset, cutoffs for the bioinformatic algorithms and the allowable percentage of CXCR4-using variants were established by optimizing against phenotypic tropism assay results and virologic responses to maraviroc between baseline and week eight. For deep sequencing, the primary cutoffs established for defining X4 HIV were a PSSM cutoff of $\geq$-4.75 and a geno2pheno cutoff of $\leq$3.5. A sample was defined as non-R5 if 2% or more of the viral sequences were classified as CXCR4-using by either of the optimized algorithms.

These cutoffs were confirmed in the remaining quarter of patients in the dataset. Compared to the sensitivity of X4 detection with population-based sequencing in Chapter 2 (69%-75%), the sensitivity of next-generation sequencing relative to the phenotypic Trofile assay was 84%, with comparable specificity. More importantly, next-generation sequencing was able to predict virologic responses to maraviroc in all three of the clinical trials examined. Deep sequencing predicted viral load declines, rates of virologic suppression, and rates of tropism changes while on maraviroc. It was superior at predicting these responses than either the original Trofile assay or population-based sequencing. Results were reproducible when the samples were processed at an independent laboratory, and were comparable to the Enhanced Sensitivity Trofile assay, where these results were available.

There has been past criticism that the true performance of next-generation sequencing for HIV tropism cannot be determined from the maraviroc trials described in Chapter 3 [234,354]. This has been levelled at the fact that patients were pre-screened using the original Trofile assay, and thus are not representative of a general treatment population. Importantly, however, this thesis demonstrates that very similar results were obtained when analyses were restricted to the unbiased sample set which is described in Section 3.3.8 and Figure 3.14. The patients selected for this unbiased data set were all treated regardless of their tropism status. Thus, the composition of this dataset was not impacted by screening. Even in this unbiased dataset, deep sequencing was an excellent predictor of virologic outcomes on maraviroc and was comparable to the Enhanced Sensitivity Trofile Assay. Furthermore, similar results were obtained by an independent laboratory which replicated these methods on the same dataset.

7.2.3 Conclusions for Chapter Four

The majority of the work presented in Chapters 2 and 3 focussed on using plasma HIV RNA from viraemic individuals. Chapter 4 differs from but expands on the previous chapters because it presents results obtained from HIV DNA derived from peripheral blood mononuclear cells rather than free virions. Approximately 2% of the overall lymphocyte population is present in the systemic circulation

at any one time [61], and these can be obtained during a blood draw. HIV DNA may be integrated in the cells' chromosomes, or may otherwise be associated with them (e.g., episomal DNA or recently reverse-transcribed viral genomes). Since the HIV DNA is used as template for new viral progeny, it represents useful alternative material for assessing HIV tropism. A tropism test is strongly recommended for patients about to start maraviroc-containing regimens [67]. However, in order to perform an RNA-based tropism test, plasma viraemia must be high enough to have sufficient material. For patients with lower or suppressed viral loads, only the cellular compartment may have enough material available for testing.

Chapter 4 evaluates this DNA-based tropism method and offers a comparison with RNA-based methods. Broadly, these methods performed very similarly, and DNA-based tropism testing was shown to be a good indicator of successful treatment with maraviroc. In a subset of patients however, it appeared that where DNA- and RNA-based methods gave discordant results, that the RNA-based results tended to be slightly more indicative of therapy success. A major success of this study was that it directly helped to inform the design of an international clinical trial. The MARCH study is a randomized clinical trial designed to evaluate the utility of DNA-based tropism testing in virologically-suppressed individuals on antiretroviral therapy [344,346].

7.2.4 Conclusions for Chapter Five

Chapter 5 examines the longitudinal changes in HIV coreceptor usage as assessed by genotypic tropism assays and next-generation sequencing. The study finds that a majority of patients who experience a change in tropism on maraviroc harboured CXCR4-using HIV prior to treatment and that this can be detected with next-generation sequencing. Furthermore, the study confirms that the viral population present at treatment failure is phylogenetically related to and is likely derived from the pre-treatment CXCR4-using population. This is important because it provides evidence that the pre-treatment viral species are not merely artifacts of the next-generation sequencing methodology, but represent the actual viral populations which undergo selection by maraviroc.

The experiments performed in Chapter 5 also revealed an intriguing aspect of maraviroc failure. There were two distinct patterns of failure. Tropism changes were accompanied by a number of mutations in the V3 loop and a massive increase in the proportion of non-R5 variants. In contrast, where virologic failure was not associated with a tropism change, there were few, if any, changes in V3 and the proportion of non-R5 variants remained low at both screening and failure. These changes were so distinct and drastic that while X4 variants could only be detected at low levels by deep sequencing at screening, they could be easily detected at failure by standard population-based sequencing due to their massive increase in prevalence.

7.2.5 Conclusions for Chapter Six

All previous chapters described the application of deep sequencing to treatment-experienced populations. The final chapter applies this technology in a completely different population of treatment-naïve patients entering the MERIT trial of maraviroc. Additionally, all previous analyses were performed on patients with primarily subtype B HIV-1. Chapter 6 includes patients with other subtypes of HIV-1. Interestingly, although deep sequencing was optimized on subtype B-infected, treatment-experienced patients, it was still an excellent discriminator of responses to maraviroc in treatment-naïve patients, many of whom were infected with non-subtype B HIV. Thus, the methodologies appear robust to very different bioinformatic approaches, treatment populations, and HIV subtypes.

Next-generation sequencing had similar performance to ESTA in this population. Also, a retrospective non-inferiority analysis of the trial indicated that had deep sequencing been used to screen patients entering MERIT, the maraviroc arm would have likely been found to be non-inferior to the efavirenz arm of the trial. Where results were discordant between assays, again next-generation sequencing tended to give better results relative to ESTA or population-based sequencing. Furthermore, when the once-daily maraviroc arm was retrospectively re-tested using next-generation sequencing, the virologic responses in this arm were similar to those of the twice-daily arm. Hence,

better tropism screening of the maraviroc once-daily arm may have preventing its early termination by the study's data, safety and monitoring board.

7.2.6 Summary of Specific Conclusions

Overall, this thesis establishes that next-generation sequencing is a useful tool in the clinical treatment of HIV infection. It can be used to accurately assess HIV tropism from both the plasma RNA and cellular DNA compartments. The coreceptor usage assessments by next-generation sequencing are similar, if not superior to population-based sequencing and two versions of a phenotypic tropism assay. The data obtained with this high-sensitivity genotyping approach can be used to predict several independent outcomes on therapy, including viral load declines and suppression, as well as phenotypic changes in HIV tropism over time. In most cases, the variants detected by next-generation sequencing are those which experience drug selection pressure, and contribute to treatment failure. Deep sequencing is therefore a highly detailed, clinically relevant, and predictive application of next-generation sequencing, and provides valuable insight into the treatment of HIV infection.

7.3 Future Directions & Applications

There are a number of applications and research avenues which can build upon the results described in this thesis. Importantly, the studies described in this thesis have been key in implementing next-generation sequencing into HIV treatment. In fact, since beginning these studies, the use of next-generation sequencing to determine HIV-1 tropism has now become the standard-of-care for HIV therapy, and is recommended by expert guidelines panels including the United States Department of Health and Human Services [276,277,379].

The bulk of the above detailed studies were performed using an earlier example of a next-generation sequencing platform, the Roche/454 Life Sciences Genome Sequencer FLX. As cited in Chapter 1, there are a number of other platforms which can generate similar results, but differ in their chemistry

[244,246–248,321]. As the field progresses, these platforms tend towards having increasingly long sequence read-lengths, higher read depths, lower error rates, lower costs per base, and faster turn-around times. Developments in single molecule and real-time sequencing herald exciting and unforeseeable advancements in next-generation sequencing [253,380]. Thus, there will surely be an increasing number of applications of next-generation sequencing in a vast array of fields, including HIV. Future implementation of next-generation sequencing must include transition from older platforms such as the GS-FLX to newer ones. This must also include validation that the new methods are comparable to the previously adopted ones.

There are already a number of interesting applications of next-generation sequencing in HIV infection other than for determining coreceptor usage. For example, next-generation sequencing has been used for detecting drug resistant HIV or hepatitis C virus; for investigating viral coinfections; for studying the development of HIV neutralizing antibodies; and for more accurately determining linked HIV transmissions [234,381–384]. Human genetic testing to personalize medical treatment is also possible with next-generation sequencing [254], and the above applications combined with others have high potential to lead to better treatment for people living with HIV.

Although this thesis specifically focuses on the application of next-generation sequencing to determining HIV tropism, the approach itself is readily transferrable to other types of antiretroviral drug resistance and viral evolution in general. The thesis candidate has also used deep sequencing in such a context. For example, deep sequencing can be used to detect drug resistance mutations in protease, reverse transcriptase, and integrase [237,385,386]. It can be used to monitor longitudinal evolution of HIV after transmission of drug resistance, and this may have implications on the fitness costs of various mutations [385]. Deep sequencing can also be used to monitor viral evolution upon resumption of antiretroviral therapy, though does not appear to have much utility to detect possibly archived historic mutations [386]. Finally, deep sequencing can be used to probe evolution within genes not currently associated with drug resistance, and the candidate has performed experiments to evaluate co-evolving sites within HIV-1 *nef* [387].

In addition to providing insight into treatment with maraviroc, the candidate has also used deep V3 sequencing to assess viral tropism in untreated individuals [161,164]. This can provide insight into viral evolution and coreceptor switching. Deep V3 sequencing can also potentially be used to investigate the hypothesis that CXCR4-using HIV may be protective against the development of breast cancer via interaction with CXCR4-expressing tumour tissue [388,389]. The approach is also useful in treatment with other CCR5 antagonists, including vicriviroc and cenicriviroc [110,111,317,390].

Pre-exposure prophylaxis (PrEP) or post-exposure prophylaxis (PEP) regimens containing maraviroc have been investigated for the prevention of HIV transmission [391,392]. Microbicide formulations and vaginal rings containing maraviroc have also been considered as biomedical prevention technologies [393–395]. Success with these prevention strategies will likely be contingent upon the exposing virus being CCR5-using, which appears to be common for the majority of HIV transmissions [149]. Cases in which transmission still occurs could be examined using deep V3 sequencing, and may provide additional insight. For instance, it could be used to determine whether a donor harboured CXCR4-using virus which was transmitted to the recipient. Furthermore, next-generation sequencing could be used to retrospectively examine a recipient's viral population to determine whether they were infected with a CXCR4-using virus, and also to examine how this virus evolves under maraviroc pressure during early infection prior to diagnosis. Finally, the absence of detection of any CXCR4-using virus by deep sequencing could indicate that non-adherence was the primary cause of transmission under maraviroc PrEP or PEP, rather than exposure to non-R5 virus.

A number of novel approaches to actually cure HIV infection have involved the *CCR5* gene. The first well-documented cure of HIV was published in 2009, with a follow-up confirmation study two years later [352,396]. These studies described an HIV-infected patient who underwent stem-cell transplantation for treatment of acute myeloid leukemia. The donor material was remarkable in that it was specifically selected from an individual who was homozygous for the *CCR5 Δ32* allele, which confers near total resistance to HIV infection. Following successful transplantation and cessation of antiretroviral therapy, HIV RNA and DNA remained persistently undetectable in the patient several years after the initial procedure, and there was significant immunological recovery [396]. Deep

sequencing was performed on a sample from the patient prior to achieving an undetectable viral load, and the proportion of non-R5 variants was 2.9% [352], a value which is provocatively close to the cutoff established in this thesis. It is possible that if such a procedure were attempted in the future on a patient with a higher proportion of non-R5 variants, that a cure might not be achieved in this context. Hence it may be prudent to perform deep sequencing in these cases in order to confirm the patient's R5 HIV status prior to transplantation. Recent attempts at replicating this cure have indicated that the defective *CCR5* allele seems to be a crucial part of achieving long-term remission/cure, rather than the stem cell transplant itself [397].

Partially based on the findings of the stem cell transplant cure, there have been efforts at mimicking the stem-cell cure using gene therapy which targets *CCR5* [351,398–400]. Generally, these approaches have involved direct modification of the genomes of stem cells drawn from HIV-infected patients. Using enzymatic proteins such as zinc-finger nucleases, the *CCR5* gene can be disrupted, and the modified cells can be infused back into the patient. Cells where *CCR5* has been disrupted will express lower concentrations of CCR5 on their surfaces, and will be less susceptible to HIV infection. Again, however, there may be a requirement for a very low proportion of non-R5 HIV variants in these patients in order for the procedures to work. Therefore, deep V3 sequencing may also be used in these situations to identify the best candidates.

Next-generation sequencing has revolutionized the ability to describe and evaluate the viral populations within HIV-infected individuals. Pre-existing minority variants can be selected by suboptimal antiretroviral therapy, and lead to treatment failure. However, deep sequencing can accurately detect such variants, and this information can be used to predict whether a patient will respond to their regimen. Not only can next-generation sequencing lead to personalized treatment for HIV, it may also act as a component in strategies aimed at curing HIV infection entirely.

Bibliography

1. UNAIDS. *Global report: UNAIDS report on the global AIDS epidemic 2013*.; 2013.

2. Gottlieb M, Schanker H, Fan P, Saxon A, Weisman J, Pozalski I. Pneumocystis Pneumonia -- Los Angeles. *MMWR Morbidity and mortality weekly report*. 1981;30:250–2.

3. Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vézinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*. 1983;220(4599):868–71.

4. Hu DJ, Dondero TJ, Rayfield MA, George JR, Schochetman G, Jaffe HW, Luo CC, Kalish ML, Weniger BG, Pau CP, Schable CA, Curran JW. The emerging genetic diversity of HIV. The importance of global surveillance for diagnostics, research, and prevention. *JAMA: the journal of the American Medical Association*. 1996;275(3):210–6.

5. Campbell-Yesufu OT, Gandhi RT. Update on human immunodeficiency virus (HIV)-2 infection. *Clinical infectious diseases*. 2011;52(6):780–7.

6. Varmus H. Retroviruses. *Science*. 1988;240(4858):1427–1435.

7. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, Gao F, Hahn BH, Kalish ML, Kuiken C, Learn GH, Leitner T, McCutchan F, Osmanov S, Peeters M, Pieniazek D, Salminen M, Sharp PM, Wolinsky S, et al. HIV-1 nomenclature proposal. *Science*. 2000;288(5463):55–6.

8. Taylor BS, Sobieszczyk ME, McCutchan FE, Hammer SM. The challenge of HIV-1 subtype diversity. *The New England journal of medicine*. 2008;358(15):1590–1602.

9. Hemelaar J, Gouws E, Ghys PD, Osmanov S, others. Global trends in molecular epidemiology of HIV-1 during 2000–2007. *AIDS*. 2011;25(5):679.

10. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, Kalengayi M, Marck E Van, Muyembe J, Kabongo JM, Gilbert MTP, Wolinsky SM. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature*. 2008;455:661–665.

11. Heeney JL, Dalgleish AG, Weiss R a. Origins of HIV and the evolution of resistance to AIDS. *Science*. 2006;313(5786):462–6.

12. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, Bibollet-Ruche F, Chen Y, Wain L V, Liegeois F, Loul S, Ngole EM, Bienvenue Y, Delaporte E, Brookfield JFY, Sharp PM, Shaw GM, Peeters M, Hahn BH. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science*. 2006;313(5786):523–6.

13. Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. *Cold Spring Harbor perspectives in medicine*. 2011;1(1):a006841.

14. Lemey P, Pybus OG, Wang B, Saksena NK, Salemi M, Vandamme A-M. Tracing the origin and history of the HIV-2 epidemic. *Proceedings of the National Academy of Sciences of the United States of America*. 2003;100(11):6588–92.

15. Van Heuverswyn F, Li Y, Neel C, Bailes E, Keele BF, Liu W, Loul S, Butel C, Liegeois F, Bienvenue Y, Ngolle EM, Sharp PM, Shaw GM, Delaporte E, Hahn BH, Peeters M. Human immunodeficiency viruses: SIV infection in wild gorillas. *Nature*. 2006;444(7116):164.

16. Thomson MM, Pérez-Alvarez L, Nájera R. Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. *The Lancet infectious diseases*. 2002;2(8):461–71.

17. Plantier J-C, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemée V, Damond F, Robertson DL, Simon F. A new human immunodeficiency virus derived from gorillas. *Nature medicine*. 2009;15(8):871–2.

18. Richman DD, Kinchington D, Fitch N, Griffin S, Persaud D, Siliciano R, Stein SEG, Gazzard BG, Powderly WG, Tebas P, Race E, Rizzardi GP, Pantaleo G, Razavi B, Friedland G, Pawinski R, Govender M, Govender T. *Human Immunodeficiency Virus*. 1st ed. (Richman DD, ed.). London, , UK: International Medical Press; 2003.

19. Frankel A, Young J. HIV-1: Fifteen Proteins and an RNA. *Annual review of biochemistry*. 1998;67:1–25.

20. Wong RW, Balachandran A, Ostrowski MA, Cochrane A. Digoxin suppresses HIV-1 replication by altering viral RNA processing. *PLoS pathogens*. 2013;9(3):e1003241.

21. Clapham PR. HIV and chemokines: ligands sharing cell-surface receptors. *Trends in cell biology*. 1997;7(7):264–8.

22. Engelman A, Cherepanov P. The structural biology of HIV-1: mechanistic and therapeutic insights. *Nature reviews Microbiology*. 2012;10(4):279–90.

23. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*. 1998;393(6686):648–59.

24. Rizzuto CD, Wyatt R, Hernández-Ramos N, Sun Y, Kwong PD, Hendrickson WA, Sodroski J. A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. *Science*. 1998;280(5371):1949–53.

25. Xiang S-H, Finzi A, Pacheco B, Alexander K, Yuan W, Rizzuto C, Huang C-C, Kwong PD, Sodroski J. A V3 loop-dependent gp120 element disrupted by CD4 binding stabilizes the

human immunodeficiency virus envelope glycoprotein trimer. *Journal of virology*. 2010;84(7):3147–61.

26. Lapham CK, Ouyang J, Chandrasekhar B, Nguyen NY, Dimitrov DS, Golding H. Evidence for cell-surface association between fusin and the CD4-gp120 complex in human cell lines. *Science*. 1996;274(5287):602–5.

27. Quiñones-Mateu ME, Lederman MM, Feng Z, Chakraborty B, Weber J, Rangel HR, Marotta ML, Mirza M, Jiang B, Kiser P, Medvik K, Sieg SF, Weinberg A. Human epithelial beta-defensins 2 and 3 inhibit HIV-1 replication. *AIDS*. 2003;17(16):F39–48.

28. Coffey MJ, Woffendin C, Phare SM, Strieter RM, Markovitz DM. RANTES inhibits HIV-1 replication in human peripheral blood monocytes and alveolar macrophages. *The American journal of physiology*. 1997;272(5 Pt 1):L1025–9.

29. Weiss CD. HIV-1 gp41: mediator of fusion and target for inhibition. *AIDS reviews*. 2003;5(4):214–21.

30. Chan DC, Fass D, Berger JM, Kim PS. Core structure of gp41 from the HIV envelope glycoprotein. *Cell*. 1997;89(2):263–73.

31. Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC. Atomic structure of the ectodomain from HIV-1 gp41. *Nature*. 1997;387(6631):426–30.

32. Pornillos O, Ganser-Pornillos BK, Yeager M. Atomic-level modelling of the HIV capsid. *Nature*. 2011;469(7330):424–7.

33. Schaller T, Ocwieja KE, Rasaiyaah J, Price AJ, Brady TL, Roth SL, Hué S, Fletcher AJ, Lee K, KewalRamani VN, Noursadeghi M, Jenner RG, James LC, Bushman FD, Towers GJ. HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS pathogens*. 2011;7(12):e1002439.

34. Forshey BM, von Schwedler U, Sundquist WI, Aiken C. Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *Journal of virology*. 2002;76(11):5667–77.

35. Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science*. 1998;282(5394):1669–75.

36. Albrecht D, Altfeld M, Behrens G, Bredeek UF, Buhk T, Dickinson L, Lyall H, Meemken L. *HIV Medicine*. 15th ed. (Hoffmann C, Rockstroh JK, Kamps BS, eds.). Flying Publisher; 2007.

37. Preston BD, Poiesz BJ, Loeb LA. Fidelity of HIV-1 reverse transcriptase. *Science*. 1988;242(4882):1168–71.

38. Guo F, Cen S, Niu M, Saadatmand J, Kleiman L. Inhibition of tRNA(Lys)-primed reverse transcription by human APOBEC3G during human immunodeficiency virus type 1 replication. *Journal of virology*. 2006;80(23):11710–22.

39. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature*. 2003;424(6944):99–103.

40. Sheehy AM, Gaddis NC, Malim MH. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nature medicine*. 2003;9(11):1404–7.

41. Strebel K, Luban J, Jeang K-T. Human cellular restriction factors that target HIV-1 replication. *BMC medicine*. 2009;7:48.

42. Fassati A, Görlich D, Harrison I, Zaytseva L, Mingot J-M. Nuclear import of HIV-1 intracellular reverse transcription complexes is mediated by importin 7. *The EMBO journal*. 2003;22(14):3675–85.

43. Bukrinsky M, Adzhubei A. Viral protein R of HIV-1. *Reviews in medical virology*. 1999;9(1):39–49.

44. Maertens GN, Hare S, Cherepanov P. The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature*. 2010;468(7321):326–9.

45. Chun T-W, Fauci AS. HIV reservoirs: pathogenesis and obstacles to viral eradication and cure. *AIDS*. 2012;26(10):1261–8.

46. Saad JS, Miller J, Tai J, Kim A, Ghanam RH, Summers MF. Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proceedings of the National Academy of Sciences of the United States of America*. 2006;103(30):11364–9.

47. Malim MH, Emerman M. HIV-1 accessory proteins--ensuring viral survival in a hostile environment. *Cell host & microbe*. 2008;3(6):388–98.

48. Hallenberger S, Bosch V, Angliker H, Shaw E, Klenk HD, Garten W. Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. *Nature*. 1992;360(6402):358–61.

49. Moulard M, Hallenberger S, Garten W, Klenk HD. Processing and routage of HIV glycoproteins by furin to the cell surface. *Virus research*. 1999;60(1):55–65.

50. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature*. 1988;331(6153):280–3.

51. Scarlata S, Carter C. Role of HIV-1 Gag domains in viral assembly. *Biochimica et biophysica acta*. 2003;1614(1):62–72.

52. Jouvenet N, Simon SM, Bieniasz PD. Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106(45):19114–9.

53. Hurley JH, Hanson PI. Membrane budding and scission by the ESCRT machinery: it's all in the neck. *Nature reviews Molecular cell biology*. 2010;11(8):556–66.

54. Dubé M, Roy BB, Guiot-Guillain P, Binette J, Mercier J, Chiasson A, Cohen E a. Antagonism of tetherin restriction of HIV-1 release by Vpu involves binding and sequestration of the restriction factor in a perinuclear compartment. *PLoS pathogens*. 2010;6(4):e1000856.

55. Fitzgerald PMD, Springer JP. Structure and function of retroviral proteases. *Annual review of biophysics and biophysical chemistry*. 1991;20(1):299–320.

56. Huff JR. HIV protease: a novel chemotherapeutic target for AIDS. *Journal of medicinal chemistry*. 1991;34(8):2305–14.

57. Sundquist WI, Kräusslich H-G. HIV-1 assembly, budding, and maturation. *Cold Spring Harbor perspectives in medicine*. 2012;2(7):a006924.

58. Hladik F, McElrath MJ. Setting the stage: host invasion by HIV. *Nature reviews Immunology*. 2008;8(6):447–57.

59. Tebit DM, Ndembi N, Weinberg A, Quiñones-Mateu ME. Mucosal transmission of human immunodeficiency virus. *Current HIV research*. 2012;10(1):3–8.

60. Stevenson M. HIV-1 pathogenesis. *Nature medicine*. 2003;9(7):853–60.

61. Blankson JN, Persaud D, Siliciano RF. The challenge of viral reservoirs in HIV-1 infection. *Annual review of medicine*. 2002;53(6):557–93.

62. Romagnani S. The Th1/Th2 paradigm. *Immunology today*. 1997;18(6):263–6.

63. Fahey JL, Taylor JM, Detels R, Hofmann B, Melmed R, Nishanian P, Giorgi J V. The prognostic value of cellular and serologic markers in infection with human immunodeficiency virus type 1. *The New England journal of medicine*. 1990;322(3):166–72.

64. Puhan M a, Van Natta ML, Palella FJ, Addessi A, Meinert C. Excess mortality in patients with AIDS in the era of highly active antiretroviral therapy: temporal changes and risk factors. *Clinical infectious diseases*. 2010;51(8):947–56.

65. The Opportunistic Infections Project Team of the Collaboration of Observational HIV Epidemiological Research in Europe (COHERE) in EuroCoord. CD4 Cell Count and the Risk of AIDS or Death in HIV-Infected Adults on Combination Antiretroviral Therapy with a Suppressed Viral Load: A Longitudinal Cohort Study from COHERE. *PLoS medicine*. 2012;9(3):e1001194.

66. Mellors JW, Rinaldo CR, Gupta P, White RM, Todd JA, Kingsley LA. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science*. 1996;272(5265):1167–70.

67. Thompson MA, Aberg JA, Hoy JF, Telenti A, Benson C, Cahn P, Eron JJ, Günthard HF, Hammer SM, Reiss P, Richman DD, Rizzardini G, Thomas DL, Jacobsen DM, Volberding PA. Antiretroviral treatment of adult HIV infection: 2012 recommendations of the International Antiviral Society-USA panel. *JAMA: the journal of the American Medical Association*. 2012;308(4):387–402.

68. Castro KG, Ward JW, Slutsker L, Buehler JW, Jaffe HW, Berkelman RL, Curran JW. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Morbidity and mortality weekly report*. 1992;41(RR-17):1–19.

69. Doitsh G, Galloway NLK, Geng X, Yang Z, Monroe KM, Zepeda O, Hunt PW, Hatano H, Sowinski S, Muñoz-Arias I, Greene WC. Cell death by pyroptosis drives CD4 T-cell depletion in HIV-1 infection. *Nature*. 2014;505(7484):509–14.

70. Monroe KM, Yang Z, Johnson JR, Geng X, Doitsh G, Krogan NJ, Greene WC. IFI16 DNA Sensor Is Required for Death of Lymphoid CD4 T Cells Abortively Infected with HIV. *Science*. 2013;11(11):997–1004.

71. McCune JM. The dynamics of CD4+ T-cell depletion in HIV disease. *Nature*. 2001;410(6831):974–9.

72. Grossman Z, Meier-Schellersheim M, Sousa AE, Victorino RMM, Paul WE. CD4+ T-cell depletion in HIV infection: are we closer to understanding the cause? *Nature medicine*. 2002;8(4):319–23.

73. Douek DC, Brenchley JM, Betts MR, Ambrozak DR, Hill BJ, Okamoto Y, Casazza JP, Kuruppu J, Kunstman K, Wolinsky S, Grossman Z, Dybul M, Oxenius A, Price DA, Connors M, Koup RA. HIV preferentially infects HIV-specific CD4+ T cells. *Nature*. 2002;417(6884):95–8.

74. Embretson J, Zupancic M, Ribas JL, Burke A, Racz P, Tenner-Racz K, Haase AT. Massive covert infection of helper T lymphocytes and macrophages by HIV during the incubation period of AIDS. *Nature*. 1993;362(6418):359–62.

75. Cameron PU, Freudenthal PS, Barker JM, Gezelter S, Inaba K, Steinman RM. Dendritic cells exposed to human immunodeficiency virus type-1 transmit a vigorous cytopathic infection to CD4+ T cells. *Science*. 1992;257(5068):383–7.

76. Trono D, Van Lint C, Rouzioux C, Verdin E, Barre-Sinoussi F, Chun T-W, Chomont N. HIV Persistence and the Prospect of Long-Term Drug-Free Remissions for HIV-Infected Individuals. *Science*. 2010;329(5988):174–180.

77. Richman DD, Margolis DM, Delaney M, Greene WC, Hazuda D, Pomerantz RJ. The Challenge of Finding a Cure for HIV Infection: Can We Do Better Than HAART ? *Science*. 2009; (March):1304–1307.

78. Han Y, Wind-Rotolo M, Yang H-C, Siliciano JD, Siliciano RF. Experimental approaches to the study of HIV-1 latency. *Nature reviews Microbiology*. 2007;5(2):95–106.

79. Simon V, Ho DD. HIV-1 dynamics in vivo: implications for therapy. *Nature reviews Microbiology*. 2003;1(3):181–90.

80. Chomont N, El-Far M, Ancuta P, Trautmann L, Procopio FA, Yassine-Diab B, Boucher G, Boulassel M-R, Ghattas G, Brenchley JM, Schacker TW, Hill BJ, Douek DC, Routy J-P, Haddad EK, Sékaly R-P. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nature medicine*. 2009;15(8):893–900.

81. Hatano H, Hayes TL, Dahl V, Sinclair E, Lee T-H, Hoh R, Lampiris H, Hunt PW, Palmer S, McCune JM, Martin JN, Busch MP, Shacklett BL, Deeks SG. A Randomized, Controlled Trial of Raltegravir Intensification in Antiretroviral-treated, HIV-infected Patients with a Suboptimal CD4+ T Cell Response. *The Journal of infectious diseases*. 2011;203(7):960–8.

82. Pope M, Haase AT. Transmission, acute HIV-1 infection and the quest for strategies to prevent infection. *Nature medicine*. 2003;9(7):847–52.

83. Hughes JP, Baeten JM, Lingappa JR, Magaret AS, Wald A, de Bruyn G, Kiarie J, Inambao M, Kilembe W, Farquhar C, Celum C. Determinants of Per-Coital-Act HIV-1 Infectivity Among African HIV-1-Serodiscordant Couples. *The Journal of infectious diseases*. 2012;205(3):358–365.

84. Quinn TCC, Wawer MJJ, Sewankambo N, Serwadda D, Li C, Wabwire-Mangen F, Meehan MOO, Lutalo T, Gray RHH, Rakai Project Study Group. Viral Load and Heterosexual Transmission of Human Immunodeficiency Virus Type 1. *New England Journal of Medicine*. 2000;342(13):921–929.

85. Gray RH, Wawer MJ, Brookmeyer R, Sewankambo NK, Serwadda D, Wabwire-Mangen F, Lutalo T, Li X, vanCott T, Quinn TC. Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1-discordant couples in Rakai, Uganda. *Lancet*. 2001;357(9263):1149–53.

86. McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF. The immune response during acute HIV-1 infection: clues for vaccine development. *Nature reviews Immunology*. 2010;10(1):11–23.

87. Pedersen C, Lindhardt BO, Jensen BL, Lauritzen E, Gerstoft J, Dickmeiss E, Gaub J, Scheibel E, Karlsmark T. Clinical course of primary HIV infection: consequences for subsequent course of infection. *BMJ*. 1989;299(6692):154–7.

88. Mindel A, Tenant-Flowers M. ABC of AIDS: Natural history and management of early HIV infection. *BMJ*. 2001;322(7297):1290–3.

89. Altfeld M, Fadda L, Frleta D, Bhardwaj N. DCs and NK cells: critical effectors in the immune response to HIV-1. *Nature Reviews Immunology*. 2011;11(3):176–186.

90. Kawai T, Akira S. Innate immune recognition of viral infection. *Nature immunology*. 2006;7(2):131–7.

91. Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, Peddada L, Heldebrant C, Smith R, Conrad A, Kleinman SH, Busch MP. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. *AIDS*. 2003;17(13):1871–9.

92. Moir S, Fauci AS. B cells in HIV infection and disease. *Nature reviews Immunology*. 2009;9(4):235–45.

93. McMichael A, Rowland-Jones SL. Cellular immune responses to HIV. *Nature*. 2001;410(6831):980–7.

94. Brumme ZL, John M, Carlson JM, Brumme CJ, Chan D, Brockman MA, Swenson LC, Tao I, Szeto S, Rosato P, Sela J, Kadie CM, Frahm N, Brander C, Haas DW, Riddler SA, Haubrich R, Walker BD, Harrigan PR, et al. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. *PloS one*. 2009;4(8):e6687.

95. Yang OO, Kalams SA, Trocha A, Cao H, Luster A, Johnson RP, Walker BD. Suppression of human immunodeficiency virus type 1 replication by CD8+ cells: evidence for HLA class I-

restricted triggering of cytolytic and noncytolytic mechanisms. *Journal of virology*. 1997;71(4):3120–8.

96. Shankar P, Xu Z, Lieberman J. Viral-specific cytotoxic T lymphocytes lyse human immunodeficiency virus-infected primary T lymphocytes by the granule exocytosis pathway. *Blood*. 1999;94(9):3084–93.

97. Goulder PJR, Watkins DI. Impact of MHC class I diversity on immune control of immunodeficiency virus replication. *Nature reviews Immunology*. 2008;8(8):619–30.

98. Collins KL, Chen BK, Kalams SA, Walker BD, Baltimore D. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. *Nature*. 1998;391(6665):397–401.

99. Raport CJ, Gosling J, Schweickart VL, Gray PW, Charo IF. Molecular cloning and functional characterization of a novel human CC chemokine receptor (CCR5) for RANTES, MIP-1beta, and MIP-1alpha. *The Journal of biological chemistry*. 1996;271(29):17161–6.

100. Berger EA, Murphy PM, Farber JM. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annual review of immunology*. 1999;17:657–700.

101. Clapham PR, McKnight A. HIV-1 receptors and cell tropism. *British medical bulletin*. 2001;58:43–59.

102. De Clercq E. Recent advances on the use of the CXCR4 antagonist plerixafor (AMD3100, Mozobil™) and potential of other CXCR4 antagonists as stem cell mobilizers. *Pharmacology & therapeutics*. 2010;128(3):509–18.

103. Kucia M, Jankowski K, Reca R, Wysoczynski M, Bandura L, Allendorf DJ, Zhang J, Ratajczak J, Ratajczak MZ. CXCR4-SDF-1 signalling, locomotion, chemotaxis and adhesion. *Journal of molecular histology*. 2004;35(3):233–45.

104. Wu Y, Yoder A. Chemokine coreceptor signaling in HIV-1 infection and pathogenesis. *PLoS pathogens*. 2009;5(12):e1000520.

105. Bleul CC, Wu L, Hoxie JA, Springer TA, Mackay CR. The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America*. 1997;94(5):1925–30.

106. D'Souza MP, Harden VA. Chemokines and HIV-1 second receptors. Confluence of two fields generates optimism in AIDS research. *Nature medicine*. 1996;2(12):1293–300.

107. Glass WG, McDermott DH, Lim JK, Lekhong S, Yu SF, Frank W a, Pape J, Cheshier RC, Murphy PM. CCR5 deficiency increases risk of symptomatic West Nile virus infection. *The Journal of experimental medicine*. 2006;203(1):35–40.

108. Dorr P, Westby M, Dobbs S, Griffin P, Irvine B, Macartney M, Mori J, Rickett G, Smith-Burchnell C, Napier C, Webster R, Armour D, Price D, Stammen B, Wood A, Perros M. Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity. *Antimicrobial agents and chemotherapy*. 2005;49(11):4721–32.

109. De Clercq E. The bicyclam AMD3100 story. *Nature reviews Drug discovery*. 2003;2(7):581–7.

110. Gulick RM, Su Z, Flexner C, Hughes MD, Skolnik PR, Wilkin TJ, Gross R, Krambrink A, Coakley E, Greaves WL, Zolopa A, Reichman R, Godfrey C, Hirsch M, Kuritzkes DR. Phase 2 study of the safety and efficacy of vicriviroc, a CCR5 inhibitor, in HIV-1-Infected, treatment-experienced patients: AIDS clinical trials group 5211. *The Journal of infectious diseases*. 2007;196(2):304–12.

111. Lalezari J, Gathe J, Brinson C, Thompson M, Cohen C, Dejesus E, Galindez J, Ernst JA, Martin DE, Palleja SM. Safety, Efficacy, and Pharmacokinetics of TBR-652, a CCR5/CCR2 Antagonist, in HIV-1 – Infected, CCR5 Antagonist–Naive Subjects. *Journal of acquired immune deficiency syndromes*. 2011;57(2):118–125.

112. DiPersio JF, Stadtmauer EA, Nademanee A, Micallef INM, Stiff PJ, Kaufman JL, Maziarz RT, Hosing C, Früehauf S, Horwitz M, Cooper D, Bridger G, Calandra G. Plerixafor and G-CSF

versus placebo and G-CSF to mobilize hematopoietic stem cells for autologous stem cell transplantation in patients with multiple myeloma. *Blood*. 2009;113(23):5720–6.

113. DiPersio JF, Uy GL, Yasothan U, Kirkpatrick P. Plerixafor. *Nature reviews Drug discovery*. 2009;8(2):105–6.

114. Pantophlet R, Burton DR. GP120: target for neutralizing HIV-1 antibodies. *Annual review of immunology*. 2006;24:739–69.

115. Wyatt R. The HIV-1 Envelope Glycoproteins: Fusogens, Antigens, and Immunogens. *Science*. 1998;280(5371):1884–1888.

116. Briz V, Poveda E, Soriano V. HIV entry inhibitors: mechanisms of action and resistance pathways. *The Journal of antimicrobial chemotherapy*. 2006;57(4):619–27.

117. Madani N, Perdigoto AL, Srinivasan K, Cox JM, Chruma JJ, LaLonde J, Head M, Smith AB, Sodroski JG. Localized changes in the gp120 envelope glycoprotein confer resistance to human immunodeficiency virus entry inhibitors BMS-806 and #155. *Journal of virology*. 2004;78(7):3742–52.

118. Xiang S, Kwong PD, Gupta R, Rizzuto CD, Casper DJ, Wyatt R, Wang L, Hendrickson WA, Doyle ML, Sodroski J. Mutagenic stabilization and/or disruption of a CD4-bound state reveals distinct conformations of the human immunodeficiency virus type 1 gp120 envelope glycoprotein. *Journal of virology*. 2002;76(19):9888–99.

119. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S. Molecular architecture of native HIV-1 gp120 trimers. *Nature*. 2008;455(7209):109–13.

120. White TA, Bartesaghi A, Borgnia MJ, Meyerson JR, de la Cruz MJ V, Bess JW, Nandwani R, Hoxie J a, Lifson JD, Milne JLS, Subramaniam S. Molecular Architectures of Trimeric SIV and HIV-1 Envelope Glycoproteins on Intact Viruses: Strain-Dependent Variation in Quaternary Structure. *PLoS pathogens*. 2010;6(12):e1001249.

121. Azimi I, Matthias LJ, Center RJ, Wong JWH, Hogg PJ. Disulfide bond that constrains the HIV-1 gp120 V3 domain is cleaved by thioredoxin. *The Journal of biological chemistry*. 2010;285(51):40072–80.

122. Huang C, Tang M, Zhang M-Y, Majeed S, Montabana E, Stanfield RL, Dimitrov DS, Korber B, Sodroski J, Wilson IA, Wyatt R, Kwong PD. Structure of a V3-containing HIV-1 gp120 core. *Science*. 2005;310(5750):1025–8.

123. Shioda T, Levy JA, Cheng-Mayer C. Macrophage and T cell-line tropisms of HIV-1 are determined by specific regions of the envelope gp120 gene. *Nature*. 1991;349(6305):167–9.

124. Hwang SS, Boyle TJ, Lyerly HK, Cullen BR. Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. *Science*. 1991;253(5015):71–4.

125. Yuan W, Craig S, Si Z, Farzan M, Sodroski J. CD4-induced T-20 binding to human immunodeficiency virus type 1 gp120 blocks interaction with the CXCR4 coreceptor. *Journal of virology*. 2004;78(10):5448–57.

126. Cormier E, Dragic T. The crown and stem of the V3 loop play distinct roles in human immunodeficiency virus type 1 envelope glycoprotein interactions with the CCR5 coreceptor. *Journal of virology*. 2002;76(17):8953–8957.

127. Huang C-C, Lam SN, Acharya P, Tang M, Xiang S-H, Hussan SS-U, Stanfield RL, Robinson J, Sodroski J, Wilson IA, Wyatt R, Bewley CA, Kwong PD. Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. *Science*. 2007;317(5846):1930–4.

128. Basmaciogullari S, Babcock GJ, Van Ryk D, Wojtowicz W, Sodroski J. Identification of conserved and variable structures in the human immunodeficiency virus gp120 glycoprotein of importance for CXCR4 binding. *Journal of virology*. 2002;76(21):10791.

129. Rangel HR, Weber J, Chakraborty B, Gutierrez A, Marotta ML, Mirza M, Kiser P, Martinez MA, Este JA, Quiñones-Mateu ME. Role of the human immunodeficiency virus type 1 envelope gene in viral fitness. *Journal of virology*. 2003;77(16):9069–73.

130. Lobritz MA, Marozsan AJ, Troyer RM, Arts EJ. Natural variation in the V3 crown of human immunodeficiency virus type 1 affects replicative fitness and entry inhibitor sensitivity. *Journal of virology*. 2007;81(15):8258–69.

131. Pastore C, Nedellec R, Ramos A, Pontow S, Ratner L, Mosier DE. Human immunodeficiency virus type 1 coreceptor switching: V1/V2 gain-of-fitness mutations compensate for V3 loss-of-fitness mutations. *Journal of virology*. 2006;80(2):750–8.

132. Cardozo T, Kimura T, Philpott S, Weiser B, Burger H, Zolla-Pazner S. Structural basis for coreceptor selectivity by the HIV type 1 V3 loop. *AIDS research and human retroviruses*. 2007;23(3):415–26.

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

134. Tan Q, Zhu Y, Li J, Chen Z, Han GW, Kufareva I, Li T, Ma L, Fenalti G, Li J, Zhang W, Xie X, Yang H, Jiang H, Cherezov V, Liu H, Stevens RC, Zhao Q, Wu B. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science*. 2013;341(6152):1387–90.

135. Kajumo F, Thompson DA, Guo Y, Dragic T. Entry of R5X4 and X4 human immunodeficiency virus type 1 strains is mediated by negatively charged and tyrosine residues in the amino-terminal domain and the second extracellular loop of CXCR4. *Virology*. 2000;271(2):240–7.

136. Zanetti G, Briggs J a G, Grünewald K, Sattentau QJ, Fuller SD. Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ. *PLoS pathogens*. 2006;2(8):e83.

137. Gottlieb MS, Schroff R, Schanker HM, Weisman JD, Fan PT, Wolf RA, Saxon A. Pneumocystis carinii Pneumonia and Mucosal Candidiasis in Previously Healthy Homosexual Men - Evidence of a New Acquired Cellular Immunodeficiency. *New England Journal of Medicine*. 1981;305(24):1425–1431.

138. Klatzmann D, Barré-Sinoussi F, Nugeyre MT, Danquet C, Vilmer E, Griscelli C, Brun-Veziret F, Rouzioux C, Gluckman JC, Chermann JC. Selective tropism of lymphadenopathy associated virus (LAV) for helper-inducer T lymphocytes. *Science*. 1984;225(4657):59–63.

139. Miedema F, Meyaard L, Koot M, Klein MR, Roos MT, Groenink M, Fouchier RA, Van't Wout A, Tersmette M, Schellekens PT, H S. Changing virus-host interactions in the course of HIV-1 infection. *Immunological reviews*. 1994;140(140):35–72.

140. Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, Cayanan C, Maddon PJ, Koup RA, Moore JP, Paxton WA. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature*. 1996;381(6584):667–73.

141. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, Di Marzio P, Marmon S, Sutton RE, Hill CM, Davis CB, Peiper SC, Schall TJ, Littman DR, Landau NR. Identification of a major co-receptor for primary isolates of HIV-1. *Nature*. 1996;381(6584):661–6.

142. Feng Y, Broder CC, Kennedy PE, Berger E. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*. 1996;272(5263):872–7.

143. Zhang YJ, Moore JP. Will multiple coreceptors need to be targeted by inhibitors of human immunodeficiency virus type 1 entry? *Journal of virology*. 1999;73(4):3443–8.

144. Pollakis G, Paxton W a. Use of (alternative) coreceptors for HIV entry. *Current opinion in HIV and AIDS*. 2012;7(5):440–9.

145. Schuitemaker H, van 't Wout AB, Lusso P. Clinical significance of HIV-1 coreceptor usage. *Journal of translational medicine*. 2011;9(Suppl 1):S5.

146. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, Sun C, Grayson T, Wang S, Li H, Wei X, Jiang C, Kirchherr JL, Gao F, Anderson JA, Ping L-H, Swanstrom R, Tomaras GD, Blattner WA, et al. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proceedings of the National Academy of Sciences of the United States of America*. 2008;105(21):7552–7.

147. Salazar-Gonzalez JF, Bailes E, Pham KT, Salazar MG, Guffey MB, Keele BF, Derdeyn CA, Farmer P, Hunter E, Allen S, Manigart O, Mulenga J, Anderson JA, Swanstrom R, Haynes BF, Athreya GS, Korber BTM, Sharp PM, Shaw GM, et al. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *Journal of virology*. 2008;82(8):3952–70.

148. van't Wout AB, Kootstra NA, Mulder-Kampinga GA, Albrecht-van Lent N, Scherpbier HJ, Veenstra J, Boer K, Coutinho RA, Miedema F, Schuitemaker H. Macrophage-tropic variants initiate human immunodeficiency virus type 1 infection after sexual, parenteral, and vertical transmission. *The Journal of clinical investigation*. 1994;94(5):2060–7.

149. Margolis L, Shattock R. Selective transmission of CCR5-utilizing HIV-1: the "gatekeeper" problem resolved? *Nature reviews Microbiology*. 2006;4(4):312–7.

150. Soto-Ramirez LE, Renjifo B, McLane MF, Marlink R, O'Hara C, Sutthent R, Wasi C, Vithayasai P, Vithayasai V, Apichartpiyakul C, Auewarakul P, Peña Cruz V, Chui DS, Osathanondh R, Mayer K, Lee TH, Essex M. HIV-1 Langerhans' cell tropism associated with heterosexual transmission of HIV. *Science*. 1996;271(5253):1291–3.

151. Chalmet K, Dauwe K, Foquet L, Baatz F, Seguin-Devaux C, Van Der Gucht B, Vogelaers D, Vandekerckhove L, Plum J, Verhofstede C. Presence of CXCR4-using HIV-1 in patients with recently diagnosed infection: correlates and evidence for transmission. *The Journal of infectious diseases*. 2012;205(2):174–84.

152. De Mendoza C, Rodriguez C, García F, Eiros JM, Ruíz L, Caballero E, Aguilera A, Leiva P, Colomina J, Gutierrez F, del Romero J, Aguero J, Soriano V. Prevalence of X4 tropic viruses in

patients recently infected with HIV-1 and lack of association with transmission of drug resistance. *The Journal of antimicrobial chemotherapy*. 2007;59(4):698–704.

153. Hedskog C, Mild M, Albert J. Transmission of the X4 phenotype of HIV-1: is there evidence against the "random transmission" hypothesis? *The Journal of infectious diseases*. 2012;205(2):163–5.

154. Martinson JJ, Chapman NH, Rees DC, Liu YT, Clegg JB. Global distribution of the CCR5 gene 32-basepair deletion. *Nature genetics*. 1997;16(1):100–103.

155. Brumme ZL, Goodrich J, Mayer HB, Brumme CJ, Henrick BM, Wynhoven B, Asselin JJ, Cheung PK, Hogg RS, Montaner JSG, Harrigan PR. Molecular and clinical epidemiology of CXCR4-using HIV-1 in a large population of antiretroviral-naive individuals. *The Journal of infectious diseases*. 2005;192(3):466–74.

156. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, Saragosti S, Lapoumeroulie C, Cognaux J, Forceille C, Muyldermans G, Verhofstede C, Burtonboy G, Georges M, Imai T, Rana S, Yi Y, Smyth RJ, Collman RG, et al. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature*. 1996;382(6593):722–5.

157. Liu R, Paxton W a, Choe S, Ceradini D, Martin SR, Horuk R, MacDonald ME, Stuhlmann H, Koup R a, Landau NR. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell*. 1996;86(3):367–77.

158. O'Brien TR, Winkler C, Dean M, Nelson JA, Carrington M, Michael NL, White GC. HIV-1 infection in a man homozygous for CCR5 delta 32. *Lancet*. 1997;349(9060):1219.

159. Michael NL, Nelson JA, KewalRamani VN, Chang G, O'Brien SJ, Mascola JR, Volsky B, Louder M, White GC, Littman DR, Swanstrom R, O'Brien TR. Exclusive and persistent use of the entry coreceptor CXCR4 by human immunodeficiency virus type 1 from a subject homozygous for CCR5 delta32. *Journal of virology*. 1998;72(7):6040–7.

160. Michael NL, Chang G, Louie LG, Mascola JR, Dondero D, Birx DL, Sheppard HW. The role of viral phenotype and CCR-5 gene defects in HIV-1 transmission and disease progression. *Nature medicine*. 1997;3(3):338–40.

161. Bunnik EM, Swenson LC, Edo-Matas D, Huang W, Dong W, Frantzell A, Petropoulos CJ, Coakley E, Schuitemaker H, Harrigan PR, van 't Wout AB. Detection of Inferred CCR5- and CXCR4-Using HIV-1 Variants and Evolutionary Intermediates Using Ultra-Deep Pyrosequencing. Trkola A, ed. *PLoS Pathogens*. 2011;7(6):e1002106.

162. Pastore C, Ramos A, Mosier DE. Intrinsic obstacles to human immunodeficiency virus type 1 coreceptor switching. *Journal of virology*. 2004;78(14):7565–74.

163. Alizon S, Boldin B. Within-host viral evolution in a heterogeneous environment: insights into the HIV co-receptor switch. *Journal of evolutionary biology*. 2010;23(2005):2625–2635.

164. Poon AFY, Swenson LC, Bunnik EM, Edo-Matas D, Schuitemaker H, van 't Wout AB, Harrigan PR. Reconstructing the Dynamics of HIV Evolution within Hosts from Serial Deep Sequence Data. Fraser C, ed. *PLoS Computational Biology*. 2012;8(11):e1002753.

165. Verhofstede C, Vandekerckhove L, Eygen V Van, Demecheleer E, Vandenbroucke I, Winters B, Plum J, Vogelaers D, Stuyver L. CXCR4-using HIV type 1 variants are more commonly found in peripheral blood mononuclear cell DNA than in plasma RNA. *Journal of acquired immune deficiency syndromes*. 2009;50(2):126–36.

166. Bachis A, Cruz MI, Mocchetti I. M-tropic HIV envelope protein gp120 exhibits a different neuropathological profile than T-tropic gp120 in rat striatum. *The European journal of neuroscience*. 2010;32(4):570–8.

167. Koot M, Keet IPM, Vos AHV, de Goede REY, Roos MTL, Coutinho RA, Miedema F, Schellekens PTA, Tersmette M. Prognostic value of HIV-1 syncytium-inducing phenotype for rate of CD4+ cell depletion and progression to AIDS. *Annals of internal medicine*. 1993;118(9):681.

168. Poveda E, Briz V, Quiñones-Mateu M, Soriano V. HIV tropism: diagnostic tools and implications for disease progression and treatment with entry inhibitors. *AIDS*. 2006;20(10):1359–67.

169. Penn ML, Grivel JC, Schramm B, Goldsmith MA, Margolis L. CXCR4 utilization is sufficient to trigger CD4+ T cell depletion in HIV-1-infected human lymphoid tissue. *Proceedings of the National Academy of Sciences of the United States of America*. 1999;96(2):663–8.

170. Shepherd JC, Jacobson LP, Qiao W, Jamieson BD, Phair JP, Piazza P, Quinn TC, Margolick JB. Emergence and persistence of CXCR4-tropic HIV-1 in a population of men from the multicenter AIDS cohort study. *The Journal of infectious diseases*. 2008;198(8):1104–12.

171. Tersmette M, Lange JM, de Goede RE, de Wolf F, Eeftink-Schattenkerk JK, Schellekens PT, Coutinho RA, Huisman JG, Goudsmit J, Miedema F. Association between biological properties of human immunodeficiency virus variants and risk for AIDS and AIDS mortality. *Lancet*. 1989;1(8645):983–5.

172. Richman DD, Bozzette SA. The impact of the syncytium-inducing phenotype of human immunodeficiency virus on disease progression. *The Journal of infectious diseases*. 1994;169(5):968–74.

173. Brumme ZL, Dong WWY, Yip B, Wynhoven B, Hoffman NG, Swanstrom R, Jensen MA, Mullins JI, Hogg RS, Montaner JSG, Harrigan PR. Clinical and immunological impact of HIV envelope V3 sequence variation after starting initial triple antiretroviral therapy. *AIDS*. 2004;18(4):F1–9.

174. Lin NH, Kuritzkes DR. Tropism testing in the clinical management of HIV-1 infection. *Current opinion in HIV and AIDS*. 2009;4(6):481–7.

175. Low AJ, Swenson LC, Harrigan PR. HIV coreceptor phenotyping in the clinical setting. *AIDS reviews*. 2008;10(3):143–51.

176. Durant J, Clevenbergh P, Halfon P, Delgiudice P, Porsin S, Simonet P, Montagne N, Boucher C, Schapiro JM, Dellamonica P. Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial. *Lancet*. 1999;353(9171):2195–9.

177. Winters B, Montaner J, Harrigan PR, Gazzard B, Pozniak A, Miller MD, Emery S, van Leth F, Robinson P, Baxter JD, Perez-Elias M, Castor D, Hammer S, Rinehart A, Vermeiren H, Van Craenenbroeck E, Bacheler L. Determination of clinically relevant cutoffs for HIV-1 phenotypic resistance estimates through a combined analysis of clinical trial and cohort data. *Journal of acquired immune deficiency syndromes*. 2008;48(1):26–34.

178. Shafer RW, Dupnik K, Winters MA, Eshleman SH. A Guide to HIV-1 Reverse Transcriptase and Protease Sequencing for Drug Resistance Studies. *HIV sequence compendium*. 2001;2001:1–51.

179. Winters B, Van Craenenbroeck E, Van Der Borght K, Lecocq P, Villacian J, Bacheler L. Clinical cut-offs for HIV-1 phenotypic resistance estimates: update based on recent pivotal clinical trial data and a revised approach to viral mixtures. *Journal of Virological Methods*. 2009;162:101–108.

180. Schuurman R, Demeter L, Reichelderfer P, Tijnagel J, de Groot T, Boucher C. Worldwide Evaluation of DNA Sequencing Approaches for Identification of Drug Resistance Mutations in the Human Immunodeficiency Virus Type 1 Reverse Transcriptase. *Journal of clinical microbiology*. 1999;37(7):2291–2296.

181. Harrigan PR, Côté HC. Clinical utility of testing human immunodeficiency virus for drug resistance. *Clinical infectious diseases*. 2000;30 Suppl 2:S117–22.

182. Hirsch MS, Conway B, D'Aquila RT, Johnson VA, Brun-Vézinet F, Clotet B, Demeter LM, Hammer SM, Jacobsen DM, Kuritzkes DR, Loveday C, Mellors JW, Vella S, Richman DD. Antiretroviral drug resistance testing in adults with HIV infection: implications for clinical management. International AIDS Society--USA Panel. *JAMA: the journal of the American Medical Association*. 1998;279(24):1984–91.

183. Clavel F, Hance AJ. HIV drug resistance. *The New England journal of medicine*. 2004;350(10):1023–35.

184. DeGruttola V, Dix L, D'Aquila R, Holder D, Phillips A, Ait-Khaled M, Baxter J, Clevenbergh P, Hammer S, Harrigan R, Katzenstein D, Lanier R, Miller M, Para M, Yerly S, Zolopa A, Murray J, Patick A, Miller V, et al. The relation between baseline HIV drug resistance and response to antiretroviral therapy: re-analysis of retrospective and prospective studies using a standardized data analysis plan. *Antiviral therapy*. 2000;5(1):41–8.

185. Koot M, Vos AH, Keet RP, de Goede RE, Dercksen MW, Terpstra FG, Coutinho RA, Miedema F, Tersmette M. HIV-1 biological phenotype in long-term infected individuals evaluated with an MT-2 cocultivation assay. *AIDS*. 1992;6(1):49–54.

186. Björndal A, Deng H, Jansson M, Fiore JR, Colognesi C, Karlsson A, Albert J, Scarlatti G, Littman DR, Fenyö EM. Coreceptor usage of primary human immunodeficiency virus type 1 isolates varies according to biological phenotype. *Journal of virology*. 1997;71(10):7478–87.

187. De Mendoza C, Van Baelen K, Poveda E, Rondelez E, Zahonero N, Stuyver L, Garrido C, Villacian J, Soriano V. Performance of a population-based HIV-1 tropism phenotypic assay and correlation with V3 genotypic prediction tools in recent HIV-1 seroconverters. *Journal of acquired immune deficiency syndromes*. 2008;48(3):241–4.

188. Raymond S, Delobel P, Mavigner M, Cazabat M, Souyris C, Encinas S, Bruel P, Sandres-Sauné K, Marchou B, Massip P, Izopet J. Development and performance of a new recombinant virus phenotypic entry assay to determine HIV-1 coreceptor usage. *Journal of clinical virology*. 2010;47(2):126–30.

189. Saliou A, Delobel P, Dubois M, Nicot F, Raymond S, Calvez V, Masquelier B, Izopet J. Concordance between two phenotypic assays and ultradeep pyrosequencing for determining HIV-1 tropism. *Antimicrobial agents and chemotherapy*. 2011;55(6):2831–6.

190. Skrabal K, Low AJ, Dong W, Sing T, Cheung PK, Mammano F, Harrigan PR. Determining human immunodeficiency virus coreceptor use in a clinical setting: degree of correlation between two phenotypic assays and a bioinformatic model. *Journal of clinical microbiology*. 2007;45(2):279–84.

191. Raymond S, Delobel P, Izopet J. Phenotyping methods for determining HIV tropism and applications in clinical settings. *Current opinion in HIV and AIDS*. 2012;7(5):463–9.

192. Gonzalez-Serna A, Leal M, Genebat M, Abad MA, Garcia-Perganeda A, Ferrando-Martinez S, Ruiz-Mateos E. TROCAI (tropism coreceptor assay information): a new phenotypic tropism test and its correlation with Trofile enhanced sensitivity and genotypic approaches. *Journal of clinical microbiology*. 2010;48(12):4453–8.

193. Weber J, Vazquez AC, Winner D, Gibson RM, Rhea AM, Rose JD, Wylie D, Henry K, Wright A, King K, Archer J, Poveda E, Soriano V, Robertson DL, Olivo PD, Arts EJ, Quiñones-Mateu ME. Sensitive cell-based assay for determination of human immunodeficiency virus type 1 coreceptor tropism. *Journal of clinical microbiology*. 2013;51(5):1517–27.

194. Whitcomb JM, Huang W, Fransen S, Limoli K, Toma J, Wrin T, Chappey C, Kiss LDB, Paxinos EE, Petropoulos CJ. Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism. *Antimicrobial agents and chemotherapy*. 2007;51(2):566–75.

195. Reeves JD, Coakley E, Petropoulos CJ, Whitcomb JM. An enhanced sensitivity Trofile HIV coreceptor tropism assay for selecting patients for therapy with entry inhibitors targeting CCR5: a review of analytical and clinical studies. *Journal of Viral Entry*. 2009;3:94–102.

196. Woods CK, Brumme CJ, Liu TF, Chui CKS, Chu AL, Wynhoven B, Hall T a, Trevino C, Shafer RW, Harrigan PR. Automating HIV drug resistance genotyping with RECall, a freely accessible sequence analysis tool. *Journal of clinical microbiology*. 2012;50(6):1936–42.

197. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of HIV coreceptor usage. *Nature biotechnology*. 2007;25(12):1407–10.

198. Garrido C, Roulet V, Chueca N, Poveda E, Aguilera A, Skrabal K, Zahonero N, Carlos S, García F, Faudon JL, Soriano V, de Mendoza C. Evaluation of eight different bioinformatics tools to predict viral tropism in different human immunodeficiency virus type 1 subtypes. *Journal of clinical microbiology*. 2008;46(3):887–91.

199. Pillai S, Good B, Richman D, Corbeil J. A new perspective on V3 phenotype prediction. *AIDS research and human retroviruses*. 2003;19(2):145–9.

200. Swenson LC, Boehme R, Thielen A, McGovern RA, Harrigan PR. Genotypic determination of HIV-1 tropism in the clinical setting. *HIV Therapy*. 2010;4(3):293–303.

201. Lee GQ, Cheung PK, Swenson LC, Harrigan PR. HIV tropism and CCR5 antagonists. *Hot Topics in HIV and Other Retroviruses*. 2012;3:1–21.

202. Evans MC, Paquet AC, Huang W, Napolitano L, Frantzell A, Toma J, Stawiski EW, Goetz MB, Petropoulos CJ, Whitcomb J, Coakley E, Haddad M. A case-based reasoning system for genotypic prediction of HIV-1 co-receptor tropism. *Journal of bioinformatics and computational biology*. 2013;11(4):1350006.

203. Jensen M, Li F, van't Wout A, Nickle D. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop. *Journal of Virology*. 2003;77(24):13376–13388.

204. Sing T, Low AJ, Beerenwinkel N, Sander O, Cheung PK, Domingues FS, Büch J, Däumer M, Kaiser R, Lengauer T, Harrigan PR. Predicting HIV coreceptor usage on the basis of genetic and clinical covariates. *Antiviral therapy*. 2007;12(7):1097–106.

205. Jensen MA, Coetzer M, van't Wout AB, Morris L, Mullins JI. A reliable phenotype predictor for human immunodeficiency virus type 1 subtype C based on envelope V3 sequences. *Journal of virology*. 2006;80(10):4698.

206. Shankarappa R, Margolick JB, Gange SJ, Rodrigo A, Upchurch D, Farzadegan H, Gupta P, Rinaldo CR, Learn GH, He X, Huang XL, Mullins JI. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *Journal of virology*. 1999;73(12):10489–502.

207. Curlin ME, Zioni R, Hawes SE, Liu Y, Deng W, Gottlieb GS, Zhu T, Mullins JI. HIV-1 Envelope Subregion Length Variation during Disease Progression. Trkola A, ed. *PLoS Pathogens*. 2010;6(12):e1001228.

208. Weiser B, Philpott S, Klimkait T, Burger H, Kitchen C, Bürgisser P, Gorgievski M, Perrin L, Piffaretti J-C, Ledergerber B. HIV-1 coreceptor usage and CXCR4-specific viral load predict clinical disease progression during combination antiretroviral therapy. *AIDS*. 2008;22(4):469–79.

209. Nelson JAE, Fiscus SA, Swanstrom R. Evolutionary variants of the human immunodeficiency virus type 1 V3 region characterized by using a heteroduplex tracking assay. *Journal of virology*. 1997;71(11):8750–8758.

210. McLaughlin S, Swenson LC, Hu S, Hughes P, Harrigan PR, Coombs RW, Frenkel LM. Development of a novel codon-specific polymerase chain reaction for the detection of CXCR4-utilizing HIV type 1 subtype B. *AIDS research and human retroviruses*. 2013;29(5):814–25.

211. Abbate I, Rozera G, Tommasi C, Bruselles A, Bartolini B, Chillemi G, Nicastri E, Narciso P, Ippolito G, Capobianchi MR. Analysis of co-receptor usage of circulating viral and proviral HIV genome quasispecies by ultra-deep pyrosequencing in patients who are candidates for CCR5 antagonist treatment. *Clinical microbiology and infection*. 2011;17(5):725–31.

212. Uzgiris AJ, Pandit S, Shields A, Zhang N, Chapman D, Valdez H, Heera J, Kaiser R, Thielen A, Turczyn J. Performance Characteristics of a DNA Sequencing-Based HIV-1 Coreceptor Tropism Assay for Use in a Prospective Double-Blind Randomized Phase 3 Study (A4001095, MODERN). In: *American Association for Clinical Chemistry Annual Meeting*. Los Angeles, CA; 2012:Abstract E–97.

213. Svicher V, Alteri C, Montano M, D'Arrigo R, Andreoni M, Angarano G, Antinori A, Antonelli G, Allice T, Bagnarelli P, Baldanti F, Bertoli A, Borderi M, Boeri E, Bon I, Bruzzone B, Callegaro AP, Capobianchi MR, Carosi G, et al. Performance of genotypic tropism testing on proviral DNA in clinical practice: results from the DIVA study group. *New Microbiologica*. 2012;35(1):17–25.

214. Soulié C, Fourati S, Lambert-Niclot S, Malet I, Wirden M, Tubiana R, Valantin M-A, Katlama C, Calvez V, Marcelin A-G. Factors associated with proviral DNA HIV-1 tropism in antiretroviral therapy-treated patients with fully suppressed plasma HIV viral load: implications for the clinical use of CCR5 antagonists. *Journal of antimicrobial chemotherapy*. 2010;65(4):749–51.

215. Soulie C, Calvez V, Marcelin A-G. Coreceptor usage in different reservoirs. *Current opinion in HIV and AIDS*. 2012;7(5):450–455.

216. Monno L, Saracino A, Scudeller L, Punzi G, Brindicci G, Altamura M, Lagioia A, Ladisa N, Angarano G. Impact of mutations outside the V3 region on coreceptor tropism phenotypically assessed in patients infected with HIV-1 subtype B. *Antimicrobial agents and chemotherapy*. 2011;55(11):5078–84.

217. Thielen A, Sichtig N, Kaiser R, Lam J, Harrigan PR, Lengauer T. Improved Prediction of HIV-1 Coreceptor Usage with Sequence Information from the Second Hypervariable Loop of gp120. *The Journal of infectious diseases*. 2010;202(9):1435–1443.

218. Thielen A, Lengauer T, Swenson LC, Dong WWY, McGovern RA, Lewis M, James I, Heera J, Valdez H, Harrigan PR. Mutations in gp41 are correlated with coreceptor tropism but do not improve prediction methods substantially. *Antiviral Therapy*. 2011;16:319–328.

219. Huang W, Toma J, Fransen S, Stawiski E, Reeves JD, Whitcomb JM, Parkin N, Petropoulos CJ. Coreceptor tropism can be influenced by amino acid substitutions in the gp41 transmembrane subunit of human immunodeficiency virus type 1 envelope protein. *Journal of virology*. 2008;82(11):5584–93.

220. Hoffman NG, Seillier-Moiseiwitsch F, Ahn J, Walker JM, Swanstrom R. Variability in the human immunodeficiency virus type 1 gp120 Env protein linked to phenotype-associated changes in the V3 loop. *Journal of virology*. 2002;76(8):3852–64.

221. Esbjörnsson J, Månsson F, Martínez-Arias W, Vincic E, Biague AJ, da Silva ZJ, Fenyö EM, Norrgren H, Medstrand P. Frequent CXCR4 tropism of HIV-1 subtype A and CRF02_AG during late-stage disease - indication of an evolving epidemic in West Africa. *Retrovirology*. 2010;7:23.

222. Huang W, Eshleman SH, Toma J, Fransen S, Stawiski E, Paxinos EE, Whitcomb JM, Young AM, Donnell D, Mmiro F, Musoke P, Guay LA, Jackson JB, Parkin NT, Petropoulos CJ. Coreceptor tropism in human immunodeficiency virus type 1 subtype D: high prevalence of CXCR4 tropism and heterogeneous composition of viral populations. *Journal of virology*. 2007;81(15):7885–93.

223. Gonzalez-Serna A, McGovern RA, Harrigan PR, Vidal F, Poon AFY, Ferrando-Martinez S, Abad MA, Genebat M, Leal M, Ruiz-Mateos E. Correlation of the virological response to short-term maraviroc monotherapy with standard and deep-sequencing-based genotypic tropism prediction methods. *Antimicrobial agents and chemotherapy*. 2012;56(3):1202–7.

224. Genebat M, Ruiz-Mateos E, León JA, González-Serna A, Pulido I, Rivas I, Ferrando-Martínez S, Sánchez B, Muñoz-Fernández MA, Leal M. Correlation between the Trofile test and

virological response to a short-term maraviroc exposure in HIV-infected patients. *The Journal of antimicrobial chemotherapy*. 2009;64(4):845–9.

225. Gonzalez-Serna A, Romero-Sánchez MC, Ferrando-Martinez S, Genebat M, Vidal F, Muñoz-Fernández MÁ, Abad MA, Leal M, Ruiz-Mateos E. HIV-1 tropism evolution after short-term maraviroc monotherapy in HIV-1-infected patients. *Antimicrobial agents and chemotherapy*. 2012;56(7):3981–3.

226. Genebat M, Pulido I, Romero-Sánchez MC, González-Serna A, Ferrando-Martínez S, Machmach K, Pacheco YM, Muñoz-Fernández MÁ, Ruiz-Mateos E, Leal M. Patients on a combined antiretroviral therapy after maraviroc clinical test show no immunovirological impairment. *Antiviral research*. 2012;95(3):207–11.

227. Chueca N, Garrido C, Álvarez M. Improvement in the determination of HIV-1 tropism using the V3 gene sequence and a combination of bioinformatic tools. *Journal of medical virology*. 2009;81:763–767.

228. Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H, Rock D, Falloon J, Davey RT, Dewar RL, Metcalf JA, Hammer S, Mellors JW, Coffin JM. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *Journal of clinical microbiology*. 2005;43(1):406–13.

229. Low AJ, Dong W, Chan D, Sing T, Swanstrom R, Jensen M, Pillai S, Good B, Harrigan PR. Current V3 genotyping algorithms are inadequate for predicting X4 co-receptor usage in clinical isolates. *AIDS*. 2007;21(14):F17–24.

230. Raymond S, Delobel P, Mavigner M, Cazabat M, Souyris C, Sandres-Sauné K, Cuzin L, Marchou B, Massip P, Izopet J. Correlation between genotypic predictions based on V3 sequences and phenotypic determination of HIV-1 tropism. *AIDS*. 2008;22(14):F11–6.

231. Prosperi MCF, Bracciale L, Fabbiani M, Di Giambenedetto S, Razzolini F, Meini G, Colafigli M, Marzocchetti A, Cauda R, Zazzi M, De Luca A. Comparative determination of HIV-1 co-

receptor tropism by Enhanced Sensitivity Trofile, gp120 V3-loop RNA and DNA genotyping. *Retrovirology*. 2010;7:56.

232. Archer J, Braverman MS, Taillon BE, Desany B, James I, Harrigan PR, Lewis M, Robertson DL. Detection of low-frequency pretherapy chemokine (CXC motif) receptor 4 (CXCR4)-using HIV-1 with ultra-deep pyrosequencing. *AIDS*. 2009;23(10):1209–18.

233. Li JZ, Paredes R, Ribaudo HJ, Svarovskaia ES, Metzner KJ, Kozal MJ, Hullsiek KH, Balduin M, Jakobsen MR, Geretti AM, Thiebaut R, Ostergaard L, Masquelier B, Johnson JA, Miller MD, Kuritzkes DR. Low-frequency HIV-1 drug resistance mutations and risk of NNRTI-based antiretroviral treatment failure: a systematic review and pooled analysis. *JAMA: the journal of the American Medical Association*. 2011;305(13):1327–35.

234. Li JZ, Kuritzkes DR. Clinical Implications of HIV-1 Minority Variants. *Clinical infectious diseases*. 2013;56(11):1667–74.

235. Vandenbroucke I, Van Marck H, Verhasselt P, Thys K, Mostmans W, Dumont S, Van Eygen V, Coen K, Tuefferd M, Aerssens J. Minor variant detection in amplicons using 454 massive parallel pyrosequencing: experiences and considerations for successful applications. *BioTechniques*. 2011;51(3):167–77.

236. Lee GQ, Harrigan PR, Dong W, Poon AFY, Heera J, Demarest J, Rinehart A, Chapman D, Valdez H, Portsmouth S. Comparison of population and 454 “deep” sequence analysis for HIV type 1 tropism versus the original trofile assay in non-B subtypes. *AIDS research and human retroviruses*. 2013;29(6):979–84.

237. Lee GQ, Swenson LC, Poon AFY, Martin JN, Hatano H, Deeks SG, Harrigan PR. Prolonged and Substantial Discordance in Prevalence of Raltegravir-Resistant HIV-1 in Plasma versus PBMC Samples Revealed by 454 “Deep” Sequencing. Wainberg M, ed. *PLoS ONE*. 2012;7(9):e46181.

238. Archer J, Rambaut A, Taillon BE, Harrigan PR, Lewis M, Robertson DL. The Evolutionary Analysis of Emerging Low Frequency HIV-1 CXCR4 Using Variants through Time—An Ultra-Deep Approach. Kosakovsky Pond SL, ed. *PLoS Computational Biology*. 2010;6(12):e1001022.

239. Fischer W, Ganusov V V, Giorgi EE, Hraber PT, Keele BF, Leitner T, Han CS, Gleasner CD, Green L, Lo C-C, Nag A, Wallstrom TC, Wang S, McMichael AJ, Haynes BF, Hahn BH, Perelson AS, Borrow P, Shaw GM, et al. Transmission of Single HIV-1 Genomes and Dynamics of Early Immune Escape Revealed by Ultra-Deep Sequencing. *PloS one*. 2010;5(8).

240. Droege M, Hill B. The Genome Sequencer FLX System--longer reads, more applications, straight forward bioinformatics and more complete data sets. *Journal of biotechnology*. 2008;136(1-2):3–10.

241. McGovern RA, Symons J, Poon AFY, Harrigan PR, van Lelyveld SFL, Hoepelman AIM, van Ham PM, Dong W, Wensing AMJ, Nijhuis M. Maraviroc treatment in non-R5-HIV-1-infected patients results in the selection of extreme CXCR4-using variants with limited effect on the total viral setpoint. *The Journal of antimicrobial chemotherapy*. 2013;68(9):2007–14.

242. Mardis ER. Next-generation DNA sequencing methods. *Annual review of genomics and human genetics*. 2008;9:387–402.

243. Fakhrai-Rad H, Pourmand N, Ronaghi M. Pyrosequencing: an accurate detection platform for single nucleotide polymorphisms. *Human mutation*. 2002;19(5):479–85.

244. Metzker ML. Sequencing technologies - the next generation. *Nature reviews Genetics*. 2010;11(1):31–46.

245. Wang C, Mitsuya Y, Gharizadeh B, Ronaghi M, Shafer RW. Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. *Genome research*. 2007;17(8):1195–201.

246. Glenn TC. Field guide to next-generation DNA sequencers. *Molecular Ecology Resources*. 2011;11(5):759–769.

247. Shendure J, Ji H. Next-generation DNA sequencing. *Nature biotechnology*. 2008;26(10):1135–45.

248. Archer J, Weber J, Henry K, Winner D, Gibson R, Lee L, Paxinos E, Arts EJ, Robertson DL, Mimms L, Quiñones-Mateu ME. Use of Four Next-Generation Sequencing Platforms to Determine HIV-1 Coreceptor Tropism. *PloS one*. 2012;7(11):e49602.

249. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben L a, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes X V, Godwin BC, He W, Helgesen S, Ho CH, Ho CH, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005;437(7057):376–80.

250. Zhou X, Ren L, Li Y, Zhang M, Yu Y, Yu J. The next-generation sequencing technology: a technology review and future perspective. *Science China Life sciences*. 2010;53(1):44–57.

251. Fedurco M, Romieu A, Williams S, Lawrence I, Turcatti G. BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic acids research*. 2006;34(3):e22.

252. Dressman D, Yan H, Traverso G, Kinzler KW, Vogelstein B. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proceedings of the National Academy of Sciences of the United States of America*. 2003;100(15):8817–22.

253. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, et al. Real-time DNA sequencing from single polymerase molecules. *Science*. 2009;323(5910):133–8.

254. Frese KS, Katus HA, Meder B. Next-Generation Sequencing: From Understanding Biology to Personalized Medicine. *Biology*. 2013;2:378–398.

255. Peckham HE, Mclaughlin SF, Ni JN, Rhodes MD, Malek JA, Mckernan KJ, Blanchard AP. SOLiD™ Sequencing and 2-Base Encoding. *SOLiD™ Sequencing Chemistry, Applied Biosystems*. (2624). Available at: http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_057810.pdf. Accessed December 12, 2013.

256. Applied Biosystems. *Specification Sheet - 5500 Series Genetic Analysis Systems*.; 2011.

257. Shafer RW. Low-abundance drug-resistant HIV-1 variants: finding significance in an era of abundant diagnostic and therapeutic options. *The Journal of infectious diseases*. 2009;199(5):610–2.

258. Gianella S, Richman DD. Minority variants of drug-resistant HIV. *The Journal of infectious diseases*. 2010;202(5):657–66.

259. Weiss RA. HIV and AIDS in relation to other pandemics. *EMBO reports*. 2003;4:S10–S14.

260. Hogg RS, Schechter MT, Montaner JS, Goldstone I, Craib K, O'Shaughnessy M V. Impact of HIV infection and AIDS on death rates in British Columbia and Canada. *CMAJ: Canadian Medical Association journal*. 1994;150(5):711–7.

261. McMichael AJ, Hanke T. HIV vaccines 1983-2003. *Nature medicine*. 2003;9(7):874–80.

262. Pomerantz RJ, Horn DL. Twenty years of therapy for HIV-1 infection. *Nature medicine*. 2003;9(7):867–73.

263. Richman DD. HIV chemotherapy. *Nature*. 2001;410(6831):995–1001.

264. Concorde Coordinating Committee. Concorde: MRC/ANRS randomised double-blind controlled trial of immediate and deferred zidovudine in symptom-free HIV infection. Concorde Coordinating Committee. *Lancet*. 1994;343(8902):871–81.

265. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*. 1995;373(6510):123–6.

266. Fauci AS. HIV and AIDS: 20 years of science. *Nature medicine*. 2003;9(7):839–43.

267. Sepkowitz KA. AIDS--the first 20 years. *The New England journal of medicine*. 2001;344(23):1764–72.

268. Hammer SM, Squires KE, Hughes MD, Grimes JM, Demeter LM, Currier JS, Eron JJ, Feinberg JE, Balfour HH, Deyton LR, Chodakewitz JA, Fischl MA. A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *The New England journal of medicine*. 1997;337(11):725–33.

269. Montaner JS, Reiss P, Cooper D, Vella S, Harris M, Conway B, Wainberg MA, Smith D, Robinson P, Hall D, Myers M, Lange JM. A randomized, double-blind trial comparing combinations of nevirapine, didanosine, and zidovudine for HIV-infected patients: the INCAS Trial. Italy, The Netherlands, Canada and Australia Study. *JAMA: the journal of the American Medical Association*. 1998;279(12):930–7.

270. Mellors JW, Kingsley LA, Rinaldo CR, Todd JA, Hoo BS, Kokka RP, Gupta P. Quantitation of HIV-1 RNA in plasma predicts outcome after seroconversion. *Annals of internal medicine*. 1995;122(8):573–9.

271. Weber J, Vazquez AC, Winner D, Rose JD, Wylie D, Rhea AM, Henry K, Pappas J, Wright A, Mohamed N, Gibson R, Rodriguez B, Soriano V, King K, Arts EJ, Olivo PD, Quiñones-Mateu ME. Novel method for simultaneous quantification of phenotypic resistance to maturation, protease, reverse transcriptase, and integrase HIV inhibitors based on

3'Gag(p2/p7/p1/p6)/PR/RT/INT-recombinant viruses: a useful tool in the multitarget era of antiretro. *Antimicrobial agents and chemotherapy*. 2011;55(8):3729–42.

272. Liu TF, Shafer RW. Web resources for HIV type 1 genotypic-resistance test interpretation. *Clinical infectious diseases*. 2006;42(11):1608–18.

273. Baxter JD, Mayers DL, Wentworth DN, Neaton JD, Hoover ML, Winters MA, Mannheimer SB, Thompson MA, Abrams DI, Brizz BJ, Ioannidis JP, Merigan TC. A randomized study of antiretroviral management based on plasma genotypic antiretroviral resistance testing in patients failing therapy. CPCRA 046 Study Team for the Terry Beirn Community Programs for Clinical Research on AIDS. *AIDS*. 2000;14(9):F83–93.

274. Hirsch MS, Günthard HF, Schapiro JM, Brun-Vézinet F, Clotet B, Hammer SM, Johnson V a, Kuritzkes DR, Mellors JW, Pillay D, Yeni PG, Jacobsen DM, Richman DD. Antiretroviral drug resistance testing in adult HIV-1 infection: 2008 recommendations of an International AIDS Society-USA panel. *Clinical infectious diseases*. 2008;47(2):266–85.

275. Van Baelen K, Salzwedel K, Rondelez E, Van Eygen V, De Vos S, Verheyen A, Steegen K, Verlinden Y, Allaway GP, Stuyver LJ. Susceptibility of human immunodeficiency virus type 1 to the maturation inhibitor bevirimat is modulated by baseline polymorphisms in Gag spacer peptide 1. *Antimicrobial agents and chemotherapy*. 2009;53(5):2185–8.

276. Vandekerckhove L, Wensing A, Kaiser R, Brun-Vézinet F, Clotet B, De Luca A, Dressler S, Garcia F, Geretti A, Klimkait T, Korn K, Masquelier B, Perno C, Schapiro J, Soriano V, Sönnerborg A, Vandamme A-M, Verhofstede C, Walter H, et al. European guidelines on the clinical management of HIV-1 tropism testing. *The Lancet infectious diseases*. 2011;11(5):394–407.

277. The Therapeutic Guidelines Committee of the BC Centre for Excellence in HIV/AIDS. Therapeutic Guidelines: Antiretroviral Treatment (ARV) of Adult HIV Infection. 2013; (February):1–48.

278. Williams I, Churchill D, Anderson J, Boffito M, Bower M, Cairns G, Cwynarski K, Edwards S, Fidler S, Fisher M, Freedman A, Geretti AM, Gilleece Y, Horne R, Johnson M, Khoo S, Leen C, Marshall N, Nelson M, et al. British HIV Association guidelines for the treatment of HIV-1-positive adults with antiretroviral therapy 2012 British HIV Association guidelines for the treatment of HIV-1-positive adults with antiretroviral therapy 2012. *HIV Medicine*. 2012;13(July):1–85.

279. European AIDS Clinical Society, EACS. European AIDS Clinical Society Guidelines: Clinical Management and Treatment of HIV-infected Adults in Europe Version 7.0. 2013;(October):1–82.

280. Dybul M, Fauci AS, Bartlett JG, Kaplan JE, Pau AK. Guidelines for using antiretroviral agents among HIV-infected adults and adolescents. *Annals of internal medicine*. 2002;137(5 Pt 2):381–433.

281. Hunt PW, Harrigan PR, Huang W, Bates M, Williamson DW, McCune JM, Price RW, Spudich SS, Lampiris H, Hoh R, Leigler T, Martin JN, Deeks SG. Prevalence of CXCR4 tropism among antiretroviral-treated HIV-1-infected patients with detectable viremia. *The Journal of infectious diseases*. 2006;194(7):926–30.

282. Seclén E, Soriano V, González MM, Martín-Carbonero L, Gellermann H, Distel M, Kadus W, Poveda E. Impact of Baseline HIV-1 Tropism on Viral Response and CD4 Cell Count Gains in HIV-Infected Patients Receiving First-line Antiretroviral Therapy. *The Journal of infectious diseases*. 2011;204(1):139–44.

283. Waters LJ, Scourfield AT, Marcano M, Gazzard BG, Bower M, Nelson M, Stebbing J. The Evolution of Coreceptor Tropism in HIV-infected Patients Interrupting Suppressive Antiretroviral Therapy. *Clinical infectious diseases*. 2011;52(5):671–3.

284. Kuritzkes D, Kar S, Kirkpatrick P. Maraviroc. *Nature Reviews Drug Discovery*. 2008;7(1):15–16.

285. Marier J-F, Trinh M, Pheng LH, Palleja SM, Martin DE. Pharmacokinetics and pharmacodynamics of TBR-652, a novel CCR5 antagonist, in HIV-1-infected, antiretroviral treatment-experienced, CCR5 antagonist-naïve patients. *Antimicrobial agents and chemotherapy*. 2011;55(6):2768–74.

286. Carter NJ, Keating GM. Maraviroc. *Drugs*. 2007;67(15):2277–88.

287. Fätkenheuer G, Pozniak AL, Johnson M a, Plettenberg A, Staszewski S, Hoepelman AIM, Saag MS, Goebel FD, Rockstroh JK, Dezube BJ, Jenkins TM, Medhurst C, Sullivan JF, Ridgway C, Abel S, James IT, Youle M, van der Ryst E. Efficacy of short-term monotherapy with maraviroc, a new CCR5 antagonist, in patients infected with HIV-1. *Nature medicine*. 2005;11(11):1170–2.

288. Gulick RM, Lalezari J, Goodrich J, Clumeck N, DeJesus E, Horban A, Nadler J, Clotet B, Karlsson A, Wohlfeiler M, Montana JB, McHale M, Sullivan J, Ridgway C, Felstead S, Dunne MW, van der Ryst E, Mayer H. Maraviroc for Previously Treated Patients with R5 HIV-1 Infection. *New England Journal of Medicine*. 2008;359(14):1429–1441.

289. Fätkenheuer G, Nelson M, Lazzarin A, Konourina I, Hoepelman AIM, Lampiris H, Hirschel B, Tebas P, Raffi F, Trottier B, Bellos N, Saag M, Cooper DA, Westby M, Tawadrous M, Sullivan JF, Ridgway C, Dunne MW, Felstead S, et al. Subgroup analyses of maraviroc in previously treated R5 HIV-1 infection. *The New England journal of Medicine*. 2008;359(14):1442–55.

290. Cooper D, Heera J, Goodrich J, Tawadrous M, Saag M, Dejesus E, Clumeck N, Walmsley S, Ting N, Coakley E, Reeves JD, Reyes-Teran G, Westby M, Van Der Ryst E, Ive P, Mohapi L, Mingrone H, Horban A, Hackman F, et al. Maraviroc versus efavirenz, both in combination with zidovudine-lamivudine, for the treatment of antiretroviral-naive subjects with CCR5-tropic HIV-1 infection. *The Journal of infectious diseases*. 2010;201(6):803–13.

291. Saag M, Goodrich J, Fätkenheuer G, Clotet B, Clumeck N, Sullivan J, Westby M, van der Ryst E, Mayer H. A double-blind, placebo-controlled trial of maraviroc in treatment-experienced patients infected with non-R5 HIV-1. *The Journal of infectious diseases*. 2009;199(11):1638–47.

292. Saag M, Heera J, Goodrich J, Dejesus E, Clumeck N, Cooper D, Walmsley S, Ting N, Coakley E, Reeves J, Westby M, van der Ryst E, Mayer H. Reanalysis of the MERIT Study with the Enhanced Trofile Assay (MERIT-ES). In: *48th Annual ICAAC/IDSA 46th Annual Meeting*.Vol 88. Washington DC; 2008:Poster H–1232.

293. Eigen M. Viral quasispecies. *Scientific American*. 1993;269(1):42–9.

294. Nijhuis M, Deeks S, Boucher C. Implications of antiretroviral resistance on viral fitness. *Current opinion in infectious diseases*. 2001;14(1):23–8.

295. Gianella S, Delport W, Pacold ME, Young J a, Choi JY, Little SJ, Richman DD, Kosakovsky Pond SL, Smith DM. Detection of minority resistance during early HIV-1 infection: natural variation and spurious detection rather than transmission and evolution of multiple viral variants. *Journal of virology*. 2011;85(16):8359–67.

296. Johnson JA, Li J-F, Wei X, Lipscomb J, Irlbeck D, Craig C, Smith A, Bennett DE, Monsour M, Sandstrom P, Lanier ER, Heneine W. Minority HIV-1 drug resistance mutations are present in antiretroviral treatment-naïve populations and associate with reduced treatment efficacy. *PLoS medicine*. 2008;5(7):e158.

297. Peuchant O, Thiébaut R, Capdepont S, Lavignolle-Aurillac V, Neau D, Morlat P, Dabis F, Fleury H, Masquelier B. Transmission of HIV-1 minority-resistant variants and response to first-line antiretroviral therapy. *AIDS*. 2008;22(12):1417–23.

298. Metzner KJ, Giulieri SG, Knoepfel S a, Rauch P, Burgisser P, Yerly S, Günthard HF, Cavassini M. Minority quasispecies of drug-resistant HIV-1 that lead to early therapy failure in treatment-naive and -adherent patients. *Clinical infectious diseases*. 2009;48(2):239–47.

299. Metzner KJ, Rauch P, Walter H, Boesecke C, Zöllner B, Jessen H, Schewe K, Fenske S, Gellermann H, Stellbrink H-J. Detection of minor populations of drug-resistant HIV-1 in acute seroconverters. *AIDS*. 2005;19(16):1819–25.

300. Charpentier C, Laureillard D, Piketty C, Tisserand P, Batisse D, Karmochkine M, Si-Mohamed A, Weiss L. High frequency of integrase Q148R minority variants in HIV-infected patients naive of integrase inhibitors. *AIDS*. 2010;24(6):867–73.

301. Doualla-Bell F, Avalos A, Brenner B, Gaolathe T, Mine M, Gaseitsiwe S, Oliveira M, Moisi D, Ndwapi N, Moffat H, Essex M, Wainberg M a. High prevalence of the K65R mutation in human immunodeficiency virus type 1 subtype C isolates from infected patients in Botswana treated with didanosine-based regimens. *Antimicrobial agents and chemotherapy*. 2006;50(12):4182–5.

302. Kozal MJ, Chiarella J, St John EP, Moreno EA, Simen BB, Arnold TE, Lataillade M. Prevalence of low-level HIV-1 variants with reverse transcriptase mutation K65R and the effect of antiretroviral drug exposure on variant levels. *Antiviral therapy*. 2011;16(6):925–9.

303. Varghese V, Shahriar R, Rhee S-Y, Liu T, Simen BB, Egholm M, Hanczaruk B, Blake LA, Gharizadeh B, Babrzadeh F, Bachmann MH, Fessel WJ, Shafer RW. Minority variants associated with transmitted and acquired HIV-1 nonnucleoside reverse transcriptase inhibitor resistance: implications for the use of second-generation nonnucleoside reverse transcriptase inhibitors. *Journal of acquired immune deficiency syndromes*. 2009;52(3):309–15.

304. Paredes R, Lalama CM, Ribaudo HJ, Schackman BR, Shikuma C, Giguel F, Meyer W a, Johnson V a, Fiscus S a, D'Aquila RT, Gulick RM, Kuritzkes DR. Pre-existing minority drug-resistant HIV-1 variants, adherence, and risk of antiretroviral treatment failure. *The Journal of infectious diseases*. 2010;201(5):662–71.

305. Dykes C, Najjar J, Bosch RJ, Wantman M, Furtado M, Hart S, Hammer SM, Demeter LM. Detection of drug-resistant minority variants of HIV-1 during virologic failure of indinavir, lamivudine, and zidovudine. *The Journal of infectious diseases*. 2004;189(6):1091–6.

306. Roquebert B, Malet I, Wirden M, Tubiana R, Valantin M-A, Simon A, Katlama C, Peytavin G, Calvez V, Marcelin A-G. Role of HIV-1 minority populations on resistance mutational pattern evolution and susceptibility to protease inhibitors. *AIDS*. 2006;20(2):287–9.

307. Le T, Chiarella J, Simen BB, Hanczaruk B, Egholm M, Landry ML, Dieckhaus K, Rosen MI, Kozal MJ. Low-abundance HIV drug-resistant viral variants in treatment-experienced persons correlate with historical antiretroviral use. *PloS one*. 2009;4(6):e6079.

308. Lecossier D, Shulman NS, Morand-Joubert L, Shafer RW, Joly V, Zolopa AR, Clavel F, Hance AJ. Detection of minority populations of HIV-1 expressing the K103N resistance mutation in patients failing nevirapine. *Journal of acquired immune deficiency syndromes*. 2005;38(1):37–42.

309. Flys TS, Chen S, Jones DC, Hoover DR, Church JD, Fiscus SA, Mwatha A, Guay LA, Mmiro F, Musoke P, Kumwenda N, Taha TE, Jackson JB, Eshleman SH. Quantitative analysis of HIV-1 variants with the K103N resistance mutation after single-dose nevirapine in women with HIV-1 subtypes A, C, and D. *Journal of acquired immune deficiency syndromes*. 2006;42(5):610–3.

310. Coovadia A, Hunt G, Abrams EJ, Sherman G, Meyers T, Barry G, Malan E, Marais B, Stehlau R, Ledwaba J, Hammer SM, Morris L, Kuhn L. Persistent minority K103N mutations among women exposed to single-dose nevirapine and virologic response to nonnucleoside reverse-transcriptase inhibitor-based therapy. *Clinical infectious diseases*. 2009;48(4):462–72.

311. Gorry PR, Zhang C, Wu S, Kunstman K, Trachtenberg E, Phair J, Wolinsky S, Gabuzda D. Persistence of dual-tropic HIV-1 in an individual homozygous for the CCR5 Delta 32 allele. *Lancet*. 2002;359(9320):1832–4.

312. Grivel J-C, Shattock RJ, Margolis LB. Selective transmission of R5 HIV-1 variants: where is the gatekeeper? *Journal of translational medicine*. 2011;9 Suppl 1(Suppl 1):S6.

313. Hoffmann C. The epidemiology of HIV coreceptor tropism. *European journal of medical research*. 2007;12(9):385–90.

314. Low AJ, Marchant D, Brumme CJ, Brumme ZL, Dong W, Sing T, Hogg RS, Montaner JSG, Gill V, Cheung PK, Harrigan PR. CD4-dependent characteristics of coreceptor use and HIV type 1 V3 sequence in a large population of therapy-naive individuals. *AIDS research and human retroviruses*. 2008;24(2):219–28.

315. Irlbeck DM, Amrine-Madsen H, Kitrinos KM, Labranche CC, Demarest JF. Chemokine (C-C motif) receptor 5-using envelopes predominate in dual/mixed-tropic HIV from the plasma of drug-naive individuals. *AIDS*. 2008;22(12):1425–31.

316. Westby M, Lewis M, Whitcomb J, Youle M, Pozniak AL, James IT, Jenkins TM, Perros M, Van Der Ryst E. Emergence of CXCR4-using human immunodeficiency virus type 1 (HIV-1) variants in a minority of HIV-1-infected patients following treatment with the CCR5 antagonist maraviroc is from a pretreatment CXCR4-using virus reservoir. *Journal of virology*. 2006;80(10):4909–4920.

317. Tsibris AMN, Korber B, Arnaout R, Russ C, Lo C-C, Leitner T, Gaschen B, Theiler J, Paredes R, Su Z, Hughes MD, Gulick RM, Greaves W, Coakley E, Flexner C, Nusbaum C, Kuritzkes DR. Quantitative deep sequencing reveals dynamic HIV-1 escape and large population shifts during CCR5 antagonist therapy in vivo. *PloS one*. 2009;4(5):e5683.

318. Su Z, Gulick RM, Krambrink A, Coakley E, Hughes MD, Han D, Flexner C, Wilkin TJ, Skolnik PR, Greaves WL, Kuritzkes DR, Reeves JD. Response to vicriviroc in treatment-experienced subjects, as determined by an enhanced-sensitivity coreceptor tropism assay: reanalysis of AIDS clinical trials group A5211. *The Journal of infectious diseases*. 2009;200(11):1724–8.

319. Briggs DR, Tuttle DL, Sleasman JW, Goodenow MM. Envelope V3 amino acid sequence predicts HIV-1 phenotype (co-receptor usage and tropism for macrophages). *AIDS*. 2000;14(18):2937–9.

320. Hogg RS, Yip B, Chan KJ, Wood E, Craib KJ, O'Shaughnessy M V, Montaner JS. Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy. *JAMA: the journal of the American Medical Association*. 2001;286(20):2568–77.

321. Swenson LC, Däumer M, Paredes R. Next-generation sequencing to assess HIV tropism. *Current opinion in HIV and AIDS*. 2012;7(5):478–485.

322. Huang X, Zhang J. Methods for comparing a DNA sequence with a protein sequence. *Computer applications in the biosciences : CABIOS*. 1996;12(6):497–506.

323. Simmonds P, Zhang LQ, McOmish F, Balfe P, Ludlam CA, Brown AJ. Discontinuous sequence change of human immunodeficiency virus (HIV) type 1 env sequences in plasma viral and lymphocyte-associated proviral populations in vivo: implications for models of HIV pathogenesis. *Journal of virology*. 1991;65(11):6266–76.

324. De Jong JJ, De Ronde A, Keulen W, Tersmette M, Goudsmit J. Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution. *Journal of virology*. 1992;66(11):6777–80.

325. McGovern RA, Thielen A, Mo T, Dong W, Woods CK, Chapman D, Lewis M, James I, Heera J, Valdez H, Harrigan PR. Population-based V3 genotypic tropism assay: a retrospective analysis using screening samples from the A4001029 and MOTIVATE studies. *AIDS*. 2010;24(16):2517–25.

326. Soriano V, Geretti AM, Perno C-F, Fätkenheuer G, Pillay D, Reynes J, Tambussi G, Calvez V, Alcamí J, Rockstroh J. Optimal use of maraviroc in clinical practice. *AIDS*. 2008;22(17):2231–40.

327. Esté JA, Telenti A. HIV entry inhibitors. *Lancet*. 2007;370(9581):81–8.

328. Rose JD, Rhea AM, Weber J, Quiñones-Mateu ME. Current tests to evaluate HIV-1 coreceptor tropism. *Current opinion in HIV and AIDS*. 2009;4(2):136–42.

329. Sierra S, Kaiser R, Thielen A. Genotypic coreceptor analysis. *European journal of medical research*. 2007;12:453–462.

330. Fouchier RA, Brouwer M, Broersen SM, Schuitemaker H. Simple determination of human immunodeficiency virus type 1 syncytium-inducing V3 genotype by PCR. *Journal of clinical microbiology*. 1995;33(4):906–11.

331. Chan SY, Speck RF, Power C, Gaffen SL, Chesebro B, Goldsmith MA. V3 recombinants indicate a central role for CCR5 as a coreceptor in tissue infection by human immunodeficiency virus type 1. *Journal of virology*. 1999;73(3):2350–8.

332. Bushman FD, Hoffmann C, Ronen K, Malani N, Minkah N, Rose HM, Tebas P, Wang GP. Massively parallel pyrosequencing in HIV research. *AIDS*. 2008;22(12):1411–5.

333. Swenson LC, Moores A, Low AJ, Thielen A, Dong W, Woods C, Jensen MA, Wynhoven B, Chan D, Glascock C, Harrigan PR. Improved detection of CXCR4-using HIV by V3 genotyping: application of population-based and "deep" sequencing to plasma RNA and proviral DNA. *Journal of acquired immune deficiency syndromes*. 2010;54(5):506–10.

334. Trinh L, Han D, Huang W, Wrin T, Larson J, Kiss L, Coakley E, Petropoulos C, Parkin N, Whitcomb J, Reeves J. Technical Validation of an Enhanced Sensitivity Trofile HIV Co-receptor Tropism Assay for Selecting Patients for Therapy with Entry Inhibitors Targeting CCR5. In: *XVII HIV Drug Resistance Workshop*. Sitges, Spain; 2008.

335. Swenson LC, McGovern RA, Dong W, Mo T, Lee GQ, Thielen A, Jensen M, Heera J, Lewis M, James I, Chapman D, Valdez H, Harrigan PR. Optimization of Clinically Relevant Cutoffs for Determining HIV Co-Receptor Use by Population and Deep Sequencing Methods. In: *Infectious Diseases Society of America 47th Annual Meeting*. Philadelphia, Pennsylvania: IDSA.

336. McGovern RA, Dong W, Mo T, Woods C, Zhong X, Thielen A, Jensen M, Heera J, Ellery S, Lewis M, James I, Biswas P, Chapman D, Valdez H, Harrigan R. Optimization of Clinically Relevant Cutpoints for the Determination of HIV Co-Receptor Usage to Predict Maraviroc Responses in Treatment Experienced (TE) Patients Using Population V3 Genotyping. In: *European AIDS Clinical Society*. Cologne, Germany; 2009:Poster PE3.4/8.

337. Swenson LC, Dong W, Mo T, Woods C, Zhong X, Thielen A, Jensen M, Biswas P, Ellery S, Lewis M, James I, Chapman D, Valdez H, Harrigan R. "Deep" Sequencing to Identify Treatment-Experienced Patients Who Respond to Maraviroc (MVC). In: *European AIDS Clinical Society*. Cologne, Germany; 2009:Poster PE3.3/2.

338. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical chemistry*. 1993;39(4):561–77.

339. Valdez H, Lewis M, Delogne C, Simpson P, Chapman D, Mcfadyen L, Coakley E, Ryst E Van Der, Westby M. Weighted OBT susceptibility score (wOBTSS) is a stronger predictor of virologic response at 48 weeks than baseline tropism result in MOTIVATE 1 and 2. In: *48th Annual ICAAC/IDSA 46th Meeting*. Washington DC; 2008:Poster H–1221.

340. Kagan RM, Johnson EP, Siaw M, Biswas P, Chapman DS, Su Z, Platt JL, Pesano RL. A Genotypic Test for HIV-1 Tropism Combining Sanger Sequencing with Ultradeep Sequencing Predicts Virologic Response in Treatment-Experienced Patients. Polis MA, ed. *PLoS ONE*. 2012;7(9):e46334.

341. Brumme C, Wilkin T, Su Z, Schapiro J, Kagan R, Chapman D, Heera J, Valdez H, Harrigan PR. Relative performance of ESTA, Trofile, 454 deep sequencing and reflex testing for HIV tropism in the MOTIVATE screening population of therapy experienced patients. In: *18th Conference on Retroviruses and Opportunistic Infections*. Boston, MA, USA: CROI; 2011:Abstract V–165.

342. Nabatov AA, Pollakis G, Linnemann T, Kliphius A, Chalaby MIM, Paxton WA. Intrapatient alterations in the human immunodeficiency virus type 1 gp120 V1V2 and V3 regions differentially modulate coreceptor usage, virus inhibition by CC/CXC chemokines, soluble CD4, and the b12 and 2G12 monoclonal antibodies. *Journal of virology*. 2004;78(1):524–30.

343. Cho MW, Lee MK, Carney MC, Berson JF, Doms RW, Martin MA. Identification of determinants on a dualtropic human immunodeficiency virus type 1 envelope glycoprotein that confer usage of CXCR4. *Journal of virology*. 1998;72(3):2509–15.

344. Kirby Institute, ViiV Healthcare, Pfizer. Maraviroc Switch Collaborative Study (MARCH). *ClinicalTrials.gov*. 2012. Available at: http://clinicaltrials.gov/ct2/show/NCT01384682. Accessed December 12, 2013.

345. Gill VS, Lima VD, Zhang W, Wynhoven B, Yip B, Hogg RS, Montaner JSG, Harrigan PR. Improved virological outcomes in British Columbia concomitant with decreasing incidence of HIV type 1 drug resistance detection. *Clinical infectious diseases*. 2010;50(1):98–105.

346. Tu E, Swenson LC, Land S, Pett S, Emery S, Marks K, Kelleher AD, Kaye S, Kaiser R, Schuelter E, Harrigan R. Results of external quality assessment for proviral DNA testing of HIV tropism in the Maraviroc Switch collaborative study. *Journal of clinical microbiology*. 2013;51(7):2063–71.

347. Verhofstede C, Brudney D, Reynaerts J, Vaira D, Fransen K, De Bel A, Seguin-Devaux C, De Wit S, Vandekerckhove L, Geretti A-M. Concordance between HIV-1 genotypic coreceptor tropism predictions based on plasma RNA and proviral DNA. *HIV medicine*. 2011;12(9):544–552.

348. Swenson LC, Mo T, Dong WWY, Zhong X, Woods CK, Jensen MA, Thielen A, Chapman D, Lewis M, James I, Heera J, Valdez H, Harrigan PR. Deep sequencing to infer HIV-1 co-receptor usage: application to three clinical trials of maraviroc in treatment-experienced patients. *The Journal of infectious diseases*. 2011;203(2):237–45.

349. Swenson LC, Mo T, Dong WWY, Zhong X, Woods CK, Thielen A, Jensen M, Knapp DJHF, Chapman D, Portsmouth S, Lewis M, James I, Heera J, Valdez H, Harrigan PR. Deep V3 Sequencing for HIV Type 1 Tropism in Treatment-Naive Patients: A Reanalysis of the MERIT Trial of Maraviroc. *Clinical infectious diseases*. 2011;53(7):732–42.

350. Lalezari J, Mitsuyasu R, Deeks S, Wang S, Lee G, Holmes M, Gregory P, Giedlin M, Tang W, Ando D. Successful and Persistent Engraftment of ZFN-M-R5-D Autologous CD4 T Cells (SB-728-T) in Aviremic HIV-infected Subjects on HAART. In: *18th Conference on Retroviruses and Opportunistic Infections*. Boston, MA, USA; :Abstract 46.

351. Ando D, WW T, D S. HAART Treatment Interruption following Adoptive Transfer of Zinc Finger Nuclease (ZFN) Modified Autologous CD4 T-cells (SB-728-T) to HIV-infected Subjects Demonstrates Durable Engraftment and Suppression of Viral Load. In: *51st Interscience Conference on Antimicrobial Agents and Chemotherapy*.Vol 3. Chicago, IL: ICAAC; 2011:H2–794a.

352. Hütter G, Nowak D, Mossner M, Ganepola S, Müssig A, Allers K, Schneider T, Hofmann J, Kücherer C, Blau O, Blau IW, Hofmann WK, Thiel E. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *The New England journal of medicine*. 2009;360(7):692–8.

353. El-Sadr W, Lundgren JD, Neaton J, Gordin F, Abrams D, Arduino R, Babiker A, Burman W, Clumeck N, Cohen C, others. CD4+ Count-Guided Interruption of Antiretroviral Treatment. *New England Journal of Medicine*. 2006;355(22):2283–2296.

354. Kuritzkes DR. Genotypic tests for determining coreceptor usage of HIV-1. *The Journal of infectious diseases*. 2011;203(2):146–8.

355. Thompson MA, Aberg JA, Cahn P, Montaner JSG, Rizzardini G, Telenti A, Gatell JM, Günthard HF, Hammer SM, Hirsch MS, Jacobsen DM, Reiss P, Richman DD, Volberding PA, Yeni P, Schooley RT. Antiretroviral treatment of adult HIV infection: 2010 recommendations of the International AIDS Society-USA panel. *JAMA: the journal of the American Medical Association*. 2010;304(3):321–33.

356. Vitiello P, Brudney D, MacCartney M, Garcia A, Smith C, Marshall N, Johnson M, Geretti AM. Responses to switching to maraviroc-based antiretroviral therapy in treated patients with suppressed plasma HIV-1-RNA load. *Intervirology*. 2012;55(2):172–8.

357. Paar C, Geit M, Stekel H, Berg J. Genotypic prediction of human immunodeficiency virus type 1 tropism by use of plasma and peripheral blood mononuclear cells in the routine clinical laboratory. *Journal of clinical microbiology*. 2011;49(7):2697–9.

358. Parisi SG, Andreoni C, Sarmati L, Boldrin C, Buonomini AR, Andreis S, Scaggiante R, Cruciani M, Bosco O, Manfrin V, D'Ettorre G, Mengoli C, Vullo V, Palù G, Andreoni M. HIV coreceptor tropism in paired plasma, peripheral blood mononuclear cell, and cerebrospinal fluid isolates from antiretroviral-naïve subjects. *Journal of clinical microbiology*. 2011;49(4):1441–5.

359. Lee GQ, Dong W, Mo T, Woods C, Kanters S, Shen A. Limited Evolution of Inferred HIV-1 Tropism While Viremia Is Undetectable During Standard HAART Therapy. In: *Infectious Diseases Society of America*. Vancouver, BC, Canada: IDSA; 2011:Abstract 3646.

360. Baldanti F, Paolucci S, Gulminetti R, Maserati R, Migliorino G, Pan A, Maggiolo F, Comolli G, Chiesa A, Gerna G. Higher levels of HIV DNA in memory and naive CD4(+) T cell subsets of viremic compared to non-viremic patients after 18 and 24 months of HAART. *Antiviral research*. 2001;50(3):197–206.

361. Jabara CB, Jones CD, Roach J, Anderson J a, Swanstrom R. Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(50):20166–71.

362. Andreoni M, Parisi SG, Sarmati L, Nicastri E, Ercoli L, Mancino G, Sotgiu G, Mannazzu M, Trevenzoli M, Tridente G, Concia E, Aceti A. Cellular proviral HIV-DNA decline and viral isolation in naïve subjects with <5000 copies/ml of HIV-RNA and >500 x 10(6)/l CD4 cells treated with highly active antiretroviral therapy. *AIDS*. 2000;14(1):23–9.

363. Sarmati L, Parisi SG, Nicastri E, D'Ettorre G, Palmisano L, Andreotti M, Andreoni C, Giuliano M, Gatti F, Boldrin C, Palù G, Vullo V, Vella S, Andreoni M. Association between cellular human immunodeficiency virus DNA level and immunological parameters in patients with undetectable plasma viremia level during highly active antiretroviral therapy. *Journal of clinical microbiology*. 2005;43(12):6183–5.

364. Schapiro JM, Boucher CA, Kuritzkes DR, van de Vijver DA, Llibre JM, Lewis M, Simpson P, Delogne C, McFadyen L, Chapman D, Perros M, Valdez H, Ryst E van der, Westby M. Baseline CD4+ T-cell counts and weighted background susceptibility scores strongly predict response to maraviroc regimens in treatment-experienced patients. *Antiviral therapy*. 2011;16:395–404.

365. Wilkin TJ, Goetz MB, Leduc R, Skowron G, Su Z, Chan ES, Heera J, Chapman D, Spritzler J, Reeves JD, Gulick RM, Coakley E. Reanalysis of coreceptor tropism in HIV-1-infected adults using a phenotypic assay with enhanced sensitivity. *Clinical infectious diseases*. 2011;52(7):925–8.

366. McGovern RA, Thielen A, Portsmouth S, Mo T, Dong W, Woods CK, Zhong X, Brumme CJ, Chapman D, Lewis M, James I, Heera J, Valdez H, Harrigan PR. Population-based sequencing of the V3-loop can predict the virological response to maraviroc in treatment-naive patients of the MERIT trial. *Journal of acquired immune deficiency syndromes*. 2012;61(3):279–86.

367. Raymond S, Saliou A, Nicot F, Delobel P, Dubois M, Cazabat M, Sandres-Sauné K, Marchou B, Massip P, Izopet J. Frequency of CXCR4-using viruses in primary HIV-1 infections using ultra-deep pyrosequencing. *AIDS*. 2011;25(13):1668–1673.

368. Tilton JC, Wilen CB, Didigu CA, Sinha R, Harrison JE, Agrawal-Gamse C, Henning EA, Bushman FD, Martin JN, Deeks SG, Doms RW. A maraviroc-resistant HIV-1 with narrow cross-resistance to other CCR5 antagonists depends on both N-terminal and extracellular loop domains of drug-bound CCR5. *Journal of virology*. 2010;84(20):10863–76.

369. Roche M, Salimi H, Duncan R, Wilkinson BL, Chikere K, Moore MS, Webb NE, Zappi H, Sterjovski J, Flynn JK, Ellett A, Gray LR, Lee B, Jubb B, Westby M, Ramsland PA, Lewin SR, Payne RJ, Churchill MJ, et al. A common mechanism of clinical HIV-1 resistance to the CCR5 antagonist maraviroc despite divergent resistance levels and lack of common gp120 resistance mutations. *Retrovirology*. 2013;10:43.

370. Lalezari J, Goodrich J, DeJesus E, Lampiris H, Gulick R, Saag M, Ridgway C, McHale M, Ryst E van der, Mayer H. Efficacy and Safety of Maraviroc plus Optimized Background Therapy In Viremic, ART- Experienced Patients Infected With CCR5-Tropic HIV-1: 24-Week Results of Phase 2b/3 Studies. In: *Program and abstracts of the 14th Conference on Retroviruses and Opportunistic Infections, Los Angeles*.; :Abstract 104bLB.

371. Obermeier M, Symons J, Wensing AMJ. HIV population genotypic tropism testing and its clinical significance. *Current opinion in HIV and AIDS*. 2012;7(5):470–477.

372. Demarest JF, Amrine-Madsen H, Irlbeck DM, Kitrinos KM. Virologic failure in first-line human immunodeficiency virus therapy with a CCR5 entry inhibitor, aplaviroc, plus a fixed-dose combination of lamivudine-zidovudine: nucleoside reverse transcriptase inhibitor

resistance regardless of envelope tropism. *Antimicrobial agents and chemotherapy*. 2009;53(3):1116–23.

373. Westby M, Smith-Burchnell C, Mori J, Lewis M, Mosley M, Stockdale M, Dorr P, Ciaramella G, Perros M. Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry. *Journal of virology*. 2007;81(5):2359–71.

374. Moore JP, Kuritzkes DR. A pièce de resistance: how HIV-1 escapes small molecule CCR5 inhibitors. *Current Opinion in HIV and AIDS*. 2009;4(2):118–124.

375. Delobel P, Cazabat M, Saliou A, Loiseau C, Coassin L, Raymond S, Requena M, Marchou B, Massip P, Izopet J. Primary resistance of CCR5-tropic HIV-1 to maraviroc cannot be predicted by the V3 sequence. *The Journal of antimicrobial chemotherapy*. 2013;68(11):2506–14.

376. McNicholas P, Vilchez RA, Greaves W, Kumar S, Onyebuchi C, Black T, Strizki JM. Detection of HIV-1 CXCR4 tropism and resistance in treatment experienced subjects receiving CCR5 antagonist-Vicriviroc. *Journal of clinical virology*. 2012;55(2):134–9.

377. Coakley E, Reeves JD, Huang W, Mangas-Ruiz M, Maurer I, Harskamp AM, Gupta S, Lie Y, Petropoulos CJ, Schuitemaker H, van 't Wout AB. Comparison of human immunodeficiency virus type 1 tropism profiles in clinical samples by the Trofile and MT-2 assays. *Antimicrobial agents and chemotherapy*. 2009;53(11):4686–93.

378. Walter H, Wolf E, Braun P, Daumer M, Thielen A, Noah C, Sichtig N, Muller H, Stumer M, Lengauer T, Kaiser R, Obermeier M. Performance of genotypic coreceptor measurement using geno2pheno[coreceptor] in B- and non-B HIV subtypes in a large cohort of therapy-experienced patients in Germany. In: *7th European HIV Drug Resistance Workshop*. Stockholm, Sweden; 2009.

379. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-infected adults and adolescents. Department of Health and Human Services. 2013;5(2):1–267.

380. Olasagasti F, Lieberman KR, Benner S, Cherf GM, Dahl JM, Deamer DW, Akeson M. Replication of individual DNA molecules under electronic control using a protein nanopore. *Nature Nanotechnology*. 2010;5(11):798–806.

381. Bartolini B, Giombini E, Zaccaro P, Selleri M, Rozera G, Abbate I, Comandini UV, Ippolito G, Solmone M, Capobianchi MR. Extent of HCV NS3 protease variability and resistance-associated mutations assessed by next generation sequencing in HCV monoinfected and HIV/HCV coinfected patients. *Virus Research*. 2013;177(2):205–208.

382. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, Hakim JG, Kumwenda J, Grinsztejn B, Pilotto JHS, Godbole S V, Mehendale S, Chariyalertsak S, Santos BR, Mayer KH, Hoffman IF, Eshleman SH, Piwowar-Manning E, Wang L, et al. Prevention of HIV-1 infection with early antiretroviral therapy. *The New England journal of medicine*. 2011;365(6):493–505.

383. Zhu J, Ofek G, Yang Y, Zhang B, Louder MK, Lu G, Mckee K, Pancera M, Skinner J, Zhang Z, Parks R, Eudailey J, Lloyd KE, Blinn J, Alam SM, Haynes BF, Simek M, Burton DR, Koff WC, et al. Mining the antibodyome for HIV-1-neutralizing antibodies with next-generation sequencing and phylogenetic pairing of heavy/light chains. *Proceedings of the National Academy of Sciences of the United States of America*. 2013;110(16):6470–6475.

384. Svarovskaia ES, Martin R, McHutchison JG, Miller MD, Mo H. Abundant drug-resistant NS3 mutants detected by deep sequencing in hepatitis C virus-infected patients undergoing NS3 protease inhibitor monotherapy. *Journal of clinical microbiology*. 2012;50(10):3267–74.

385. Swenson LC, Lee GQ, Mo T, Woods C, Harrigan PR. "Deep" Sequencing to Quantify in vivo Viral Fitness after Transmission of Resistant HIV. In: *49th Interscience Conference on Antimicrobial Agents and Chemotherapy.* San Francisco, CA; 2009.

386. Swenson LC, Hudson E, Mo T, Woods C, Archer J, Robertson D, Montaner J, Harrigan PR. Measurement of Low-level Drug-resistance Mutations Using Deep Sequencing Technology upon Restarting HAART. In: *Conference on Retroviruses and Opportunistic Infections*. Montreal, QC; 2009.

387. Poon AF, Swenson LC, Dong WW, Deng W, Kosakovsky Pond SL, Brumme ZL, Mullins JI, Richman DD, Harrigan PR, Frost SDW. Phylogenetic analysis of population-based and deep sequencing data to identify coevolving sites in the nef gene of HIV-1. *Molecular biology and evolution*. 2010;27(4):819–832.

388. Hessol N, Napolitano L, Smith D, Lie Y, Levine A, Young M, Cohen M, Minkoff H, Anastos K, D'Souza G, Greenblatt RM, Goedert JJ. HIV tropism and decreased risk of breast cancer. *PloS one*. 2010;5(12):e14349.

389. Velasco-Velazquez M, Jiao X, De La Fuente M, Pestell TG, Ertel A, Lisanti MP, Pestell RG. CCR5 antagonist blocks metastasis of basal breast cancer cells. *Cancer research*. 2012.

390. Kagan RM, Johnson EP, Siaw MF, Van Baelen B, Ogden R, Platt JL, Pesano RL, Lefebvre E. Comparison of genotypic and phenotypic HIV type 1 tropism assay: results from the screening samples of Cenicriviroc Study 202, a randomized phase II trial in treatment-naive subjects. *AIDS research and human retroviruses*. 2014;30(2):151–9.

391. Cohen MS, Muessig KE, Smith MK, Powers KA, Kashuba ADM. Antiviral agents and HIV prevention: controversies, conflicts, and consensus. *AIDS*. 2012;26(13):1585–98.

392. Méchai F, Quertainmont Y, Sahali S, Delfraissy J-F, Ghosn J. Post-exposure prophylaxis with a maraviroc-containing regimen after occupational exposure to a multi-resistant HIV-infected source person. *Journal of medical virology*. 2008;80(1):9–10.

393. Neff CP, Kurisu T, Ndolo T, Fox K, Akkina R. A topical microbicide gel formulation of CCR5 antagonist maraviroc prevents HIV-1 vaginal transmission in humanized RAG-hu mice. *PloS one*. 2011;6(6):e20209.

394. Malcolm RK, Veazey RS, Geer L, Lowry D, Fetherston SM, Murphy DJ, Boyd P, Major I, Shattock RJ, Klasse PJ, Doyle LA, Rasmussen KK, Goldman L, Ketas TJ, Moore JP. Sustained release of the CCR5 inhibitors CMPD167 and maraviroc from vaginal rings in rhesus macaques. *Antimicrobial agents and chemotherapy*. 2012;56(5):2251–8.

395. Fetherston SM, Boyd P, McCoy CF, McBride MC, Edwards K-L, Ampofo S, Malcolm RK. A silicone elastomer vaginal ring for HIV prevention containing two microbicides with different mechanisms of action. *European journal of pharmaceutical sciences*. 2012;48(3):406–415.

396. Allers K, Hütter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, Schneider T. Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation. *Blood*. 2011;117(10):2791–9.

397. Check Hayden E. Hopes of HIV cure in "Boston patients" dashed. *Nature*. 2013;785:6–8.

398. Wilen CB, Wang J, Tilton JC, Miller JC, Kim KA, Rebar EJ, Sherrill-Mix SA, Patro SC, Secreto AJ, Jordan APO, Lee G, Kahn J, Aye PP, Bunnell BA, Lackner AA, Hoxie JA, Danet-Desnoyers GA, Bushman FD, Riley JL, et al. Engineering HIV-resistant human CD4+ T cells with CXCR4-specific zinc-finger nucleases. *PLoS pathogens*. 2011;7(4):e1002020.

399. Voit RA, McMahon MA, Sawyer SL, Porteus MH. Generation of an HIV resistant T-cell line by targeted "stacking" of restriction factors. *Molecular therapy*. 2013;21(4):786–95.

400. Nazari R, Joshi S. CCR5 as target for HIV-1 gene therapy. *Current gene therapy*. 2008;8(4):264–72.

Appendices

Appendix I: Primers for Chapter 2

Population-based sequencing primers:

Standard, population-based sequencing from HIV RNA

First-round PCR primer, forward: 5' GAGCCAATTCCCATACATTATTGT 3′

First-round PCR primer, reverse: 5' TAAGTCTCTCAAGCGGTGGTAGCTGAA 3'

Second-round PCR primer, forward: 5' TGTGCCCCAGCTGGTTTTGCGAT 3'

Second-round PCR primer, reverse: 5' GGATCTGTCTCTGTCTCTCTCTCCA 3'

Standard, population-based sequencing from HIV DNA

First-round PCR primer, forward: 5' GAGCCAATTCCCATACATTATTGT 3′

First-round PCR primer, reverse: 5' TGTGCCCCAGCTGGTTTTGCGAT 3'

Second-round PCR primer, forward: 5' AGCACAGTACAATGTACACATGG 3′

Second-round PCR primer, reverse: 5' GAAAAATTCCCTTCCACAATTAAA 3′

Population-based sequencing primers

Sequencing primer, forward: 5' AATGTCAGYACAGTACAATGTACAC 3′

Sequencing primer, reverse: 5' GAAAAATTCCCTTCCACAATTAAA 3′

Deep sequencing primers:

Deep sequencing from HIV RNA

Second-round PCR primer, forward: 5′AATGCCAAAACCATAATAGTACA 3′

Second-round PCR primer, reverse: 5'GAAAAATTCCCTTCCACAATTAAA 3′

Deep sequencing from HIV DNA

Second-round PCR primer, forward: 5′ AATGCCAAAACCATAATAGTACA 3′

Second-round PCR primer, reverse: 5' GAAAAATTCCCTTCCACAATTAAA 3′

Deep sequencing fusion primer

5' GCCTCCCTCGCGCCATCAG 3′

Deep sequencing barcode tags:

(**A**) ACGAGTGCGT; (**B**) ACGCTCGACA; (**C**) AGACGCACTC; (**D**) AGCACTGTAG; (**E**) ATCAGACACG; (**F**) CGTGTCTCTA; (**G**) CTCGCGTGTC; (**H**) TAGTATCAGC; (**I**) TCTCTATGCG; (**J**) TGATACGTCT; (**K**) TACTGAGCTA; (**L**) ATATCGCGAG.

Example deep sequencing complete forward primer, tag A:

(fusion, then *barcode A*, then **PCR primer**)

5'GCCTCCCTCGCGCCATCAG*ACGAGTGCGT***AATGCCAAAACCATAATAGTACA**3′

Appendix II: Thermal Cycler Protocols for Chapter 2

RT-PCR:

30'@52°C; 2'@94°C; 40 cycles of (15"@94°C, 30"@55°C, 1'30"@68°C); 5'@68°C

2nd round PCR for standard sequencing:

2'@94°C; 35 cycles of (15"@94°C, 30"@55°C, 1'@72°C); 7'@72°C

2nd round PCR for deep sequencing:

2'@94°C; 35 cycles of (15"@94°C, 30"@55°C, 50"@72°C); 5'@72°C

Appendix III: Primers for Chapter 3

Primers for deep sequencing

First-round PCR primer, forward: 5' GAGCCAATTCCCATACATTATTGT 3'

First-round PCR primer, reverse: 5' GCCCATAGTGCTTCCTGCTGCTCCCAAG AACC 3'

Second-round PCR primer, forward: 5' AATGCCAAAACCATAATAGTACA 3'

Second-round PCR primer, reverse: 5' GAAAAATTCCCTTCCACAATTAAA 3'

Fusion primer

5' GCCTCCCTCGCGCCATCAG 3'

Deep sequencing barcode tags:

(**A**) ACGAGTGCGT; (**B**) ACGCTCGACA; (**C**) AGACGCACTC; (**D**) AGCACTGTAG; (**E**) ATCAGACACG; (**F**) CGTGTCTCTA; (**G**) CTCGCGTGTC; (**H**) TAGTATCAGC; (**I**) TCTCTATGCG; (**J**) TGATACGTCT; (**K**) TACTGAGCTA; (**L**) ATATCGCGAG.

Example of complete deep sequencing forward primer, tag A:

(fusion primer, then *barcode A*, then **PCR primer**)

5'GCCTCCCTCGCGCCATCAG*ACGAGTGCGT***AATGCCAAAACCATAATAGTACA**3'

Primers for population-based genotyping

First-round PCR primer, forward: 5' GAGCCAATTCCCATACATTATTGT 3'

First-round PCR primer, reverse: 5' GCCCATAGTGCTTCCTGCTGCTCCCAAG AACC 3'

Second-round PCR primer, forward: 5' TGTGCCCCAGCTGGTTTTGCGAT 3'

Second-round PCR primer, reverse: 5' TATAATTCACTTCTCCAATTGTCC 3'

Sequencing primer, forward: 5' AATGTCAGYACAGTACAATGTACAC 3'

Sequencing primer, reverse: 5' GAAAAATTCCCTTCCACAATTAAA 3'

Appendix IV: Description of Deep Sequencing Data Processing Pipeline

The output from the Genome Sequencer FLX platform was processed using a custom pipeline of Ruby and Python scripts. Identical nucleotide reads were merged into "variants" and the number of reads per variant was recorded. Variants were sorted by multiplexing tag and primer, tolerating a maximum of three nucleotide mismatches from the known primer sequence. No differences were tolerated in the tag sequences.

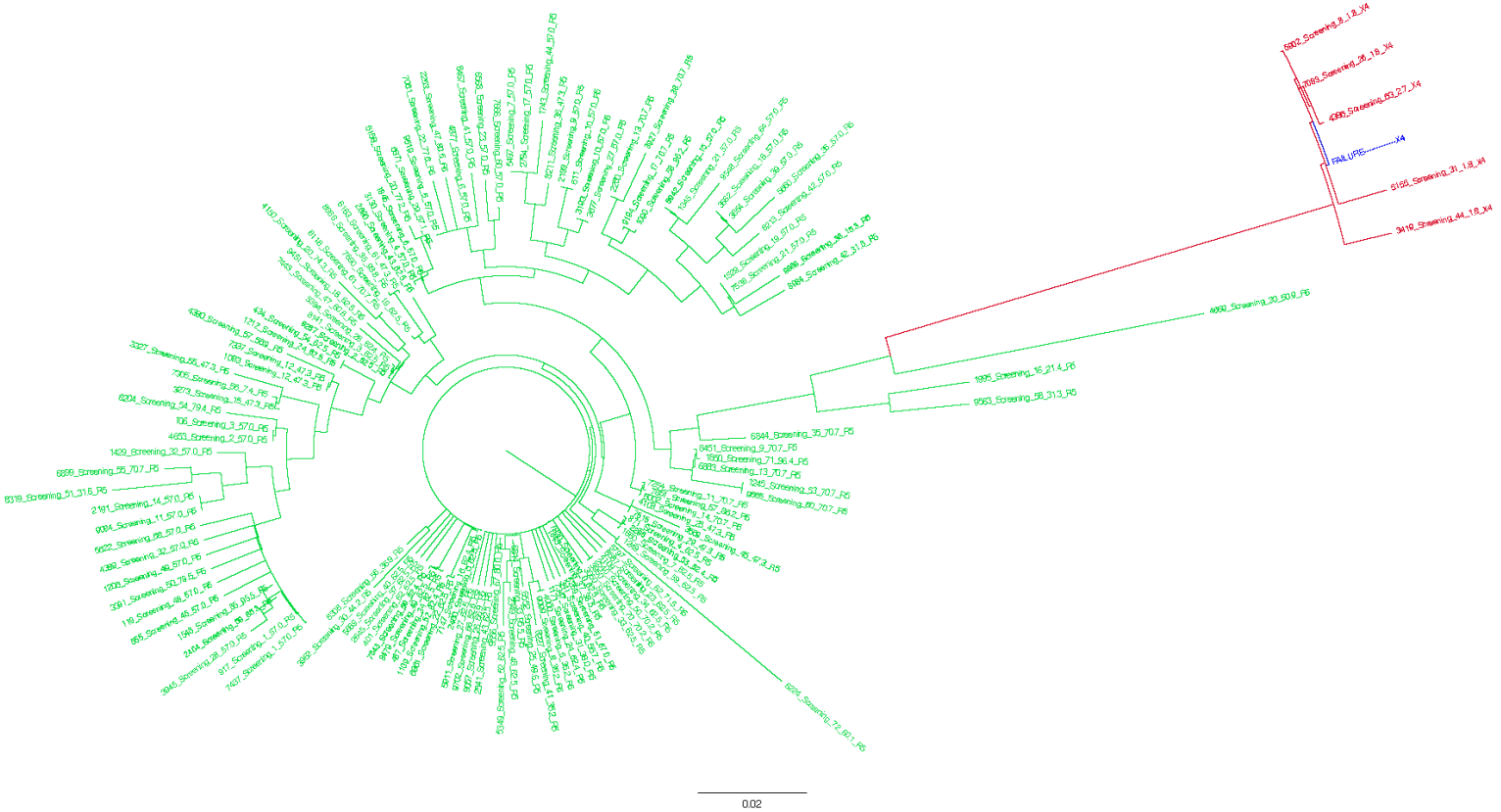
A sample-specific consensus sequence was generated from the three most abundant variants with a given tag and primer combination. Subsequently, all variants in the tag and primer-defined set were aligned pairwise against this consensus sequence, trimmed to the env V3 region, and screened for insertions and deletions that induced shifts in the reading frame.

Geno2pheno scores were calculated from the V3 sequences by aligning the amino acid translation against the geno2pheno reference sequence and scoring the residues at each reference position according to the geno2pheno support vector machine classifier. Any V3 protein sequences that (1) contained a stop codon; (2) contained residues other than cysteine on the 5' or 3' termini; or (3) comprised fewer than 33 or more than 40 residues did not receive a score. Samples with fewer than 750 reads were discarded and resequenced.

Appendix V: Phylogenetic Tree from a Maraviroc Recipient with a Small Pre-Treatment X4 Population Related to the Failure Sequence.

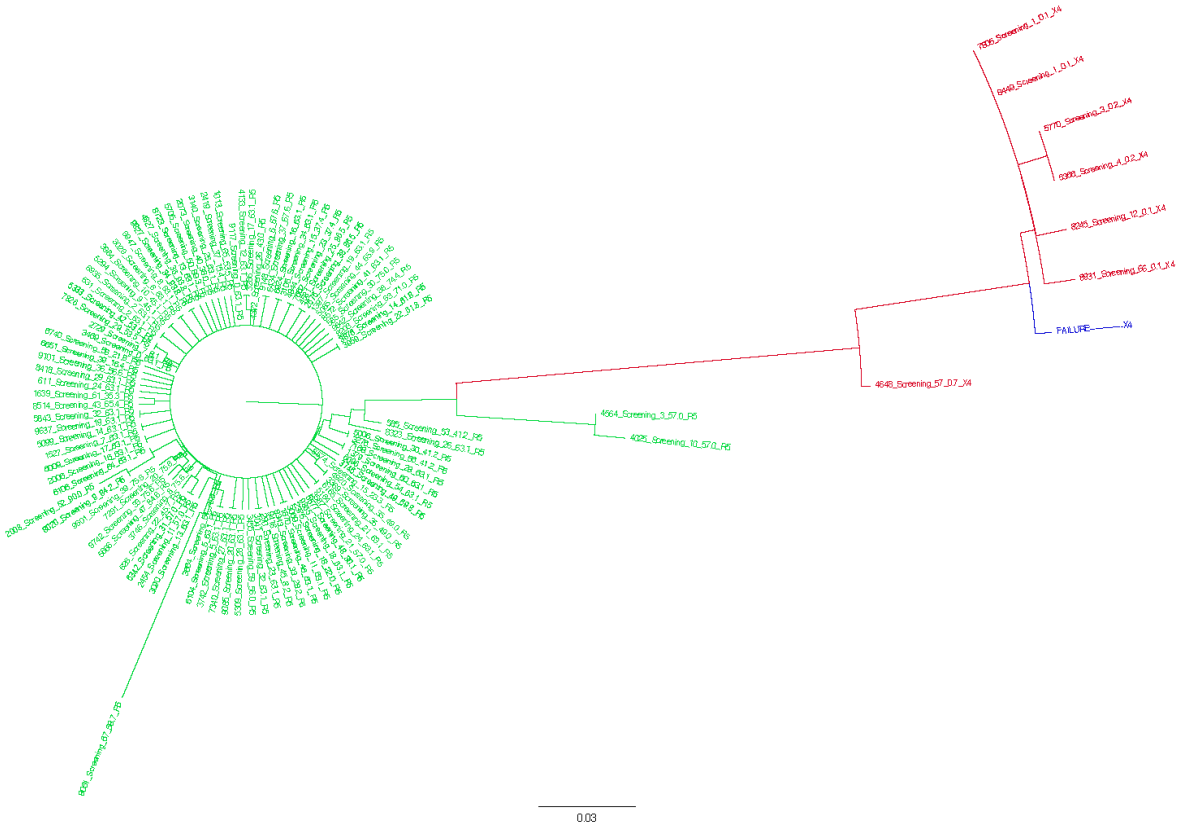
Appendix V: Phylogenetic Tree from a Maraviroc Recipient with a Small Pre-Treatment X4 Population Related to the Failure Sequence
Screening sample: 0.1% X4 by deep sequencing, and R5 by ESTA
Failure sample: X4 by population-based genotype and Dual/Mixed by the original Trofile assay.
X4 sequences are shown in red, R5 sequences are shown in green, and the failure sequence is shown in blue.

Appendix VI: Phylogenetic Tree from a Maraviroc Recipient with a Small Pre-Treatment X4 Population Related to the Failure Sequence



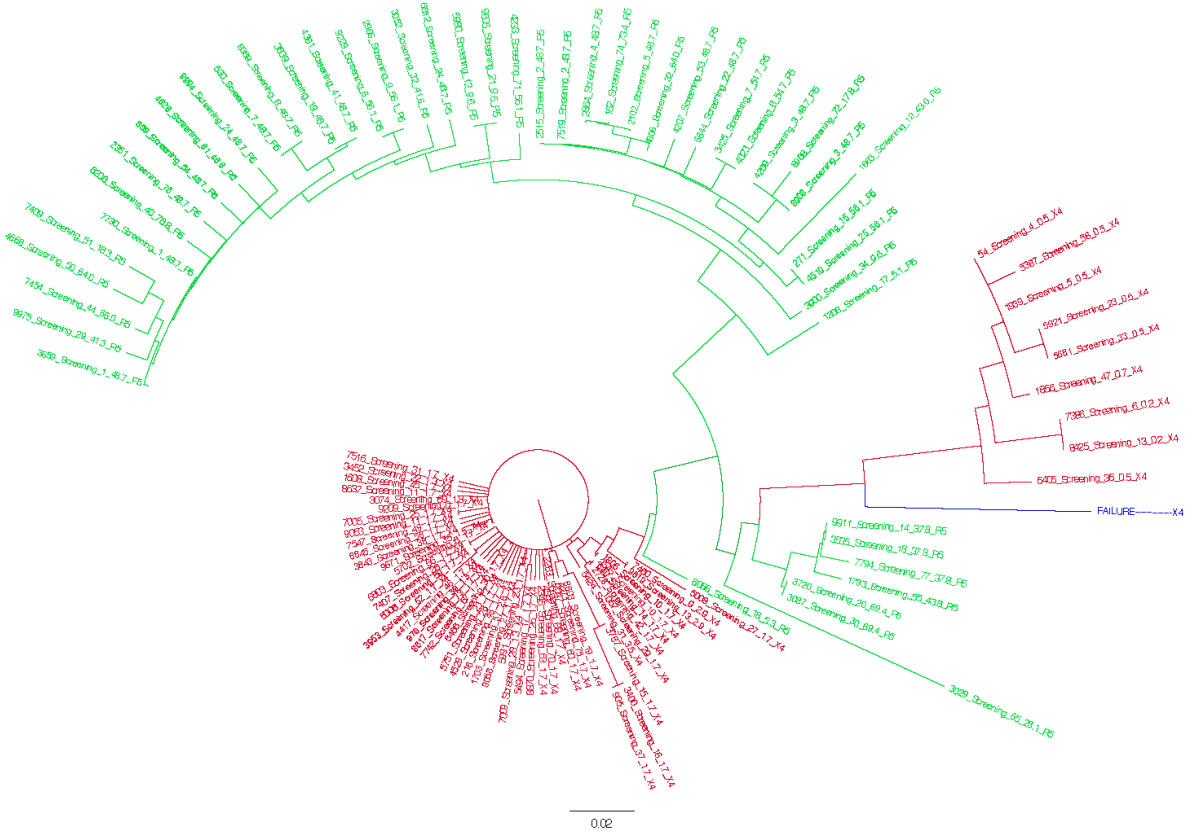
Appendix VI: Phylogenetic Tree from a Maraviroc Recipient with a Small Pre-Treatment X4 Population Related to the Failure Sequence
Screening sample: 2% X4 by deep sequencing, and R5 by ESTA
Failure sample: X4 by population-based genotype, and X4 by original Trofile assay.
R5 sequences shown in green, X4 sequences shown in red, failure sequence shown in blue.

Appendix VII: Phylogenetic Tree from a Maraviroc Recipient with a Small Pre-Treatment X4 Population Related to the Failure Sequence



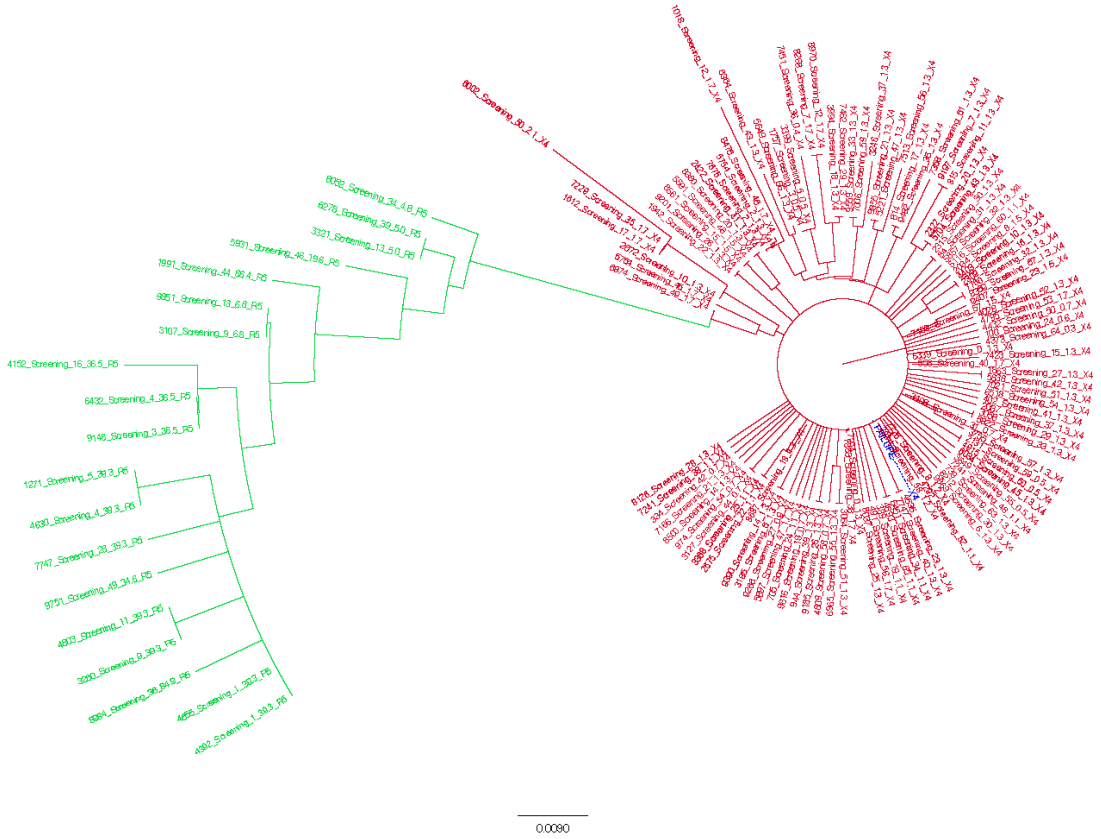
Appendix VII: Phylogenetic Tree from a Maraviroc Recipient with a Small Pre-Treatment X4 Population Related to the Failure Sequence
Screening sample: 5% X4 by deep sequencing, and Dual/Mixed by ESTA
Failure sample: X4 by population-based genotype, and X4 by original Trofile assay
R5 sequences shown in green, X4 sequences shown in red, failure sequence shown in blue.

Appendix VIII: Phylogenetic Tree from a Maraviroc Recipient with a Large Pre-Treatment X4 Population Related to the Failure Sequence



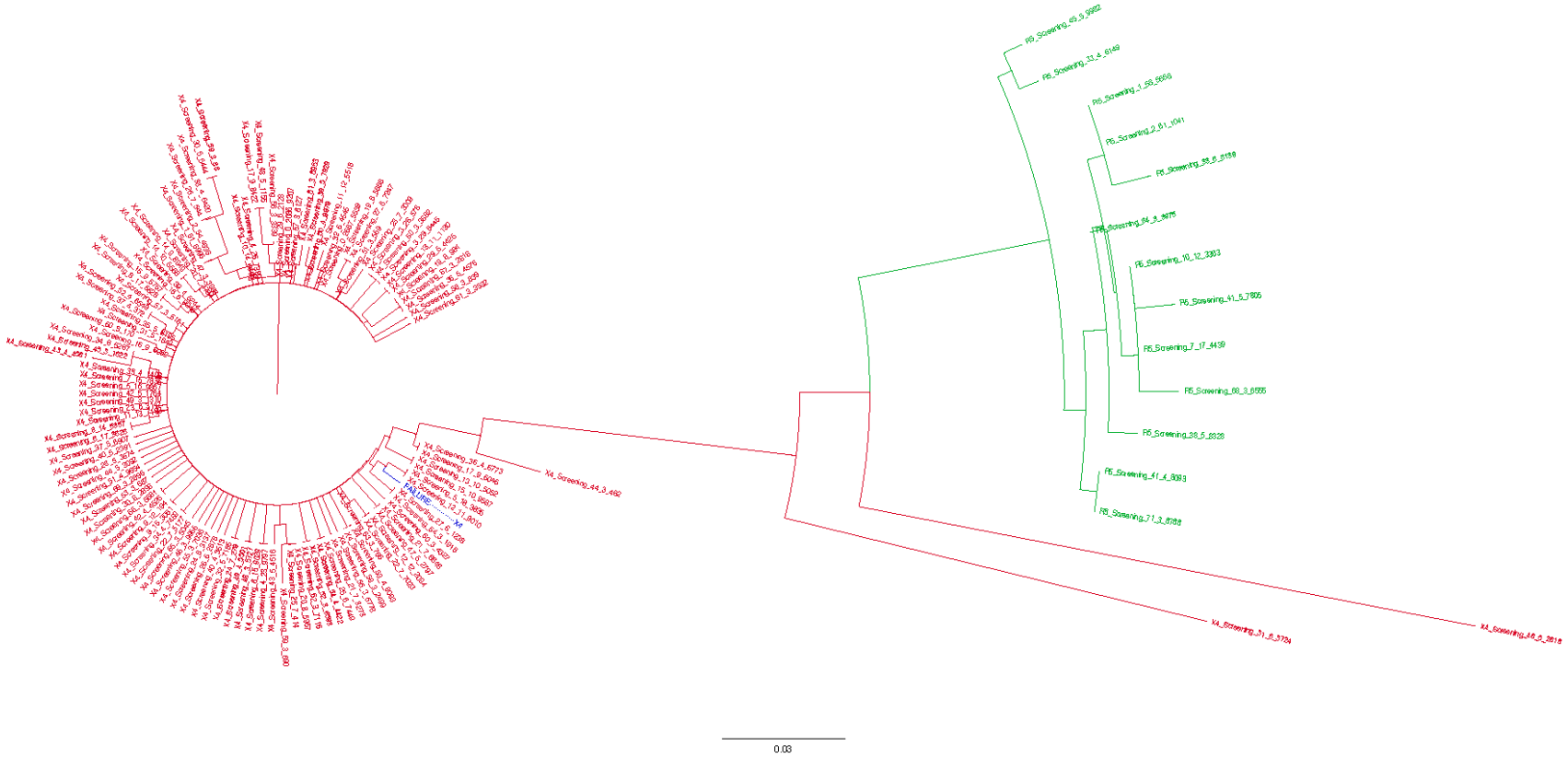
Appendix VIII: Phylogenetic Tree from a Maraviroc Recipient with a Large Pre-Treatment X4 Population Related to the Failure Sequence
Screening sample: 56% X4 by deep sequencing, and R5 by ESTA
Failure sample: X4 by population-based genotype, and Dual/Mixed by original Trofile assay.
X4 sequences are shown in red, R5 sequences are shown in green, and the failure sequence is shown in blue.

Appendix IX: Phylogenetic Tree from a Maraviroc Recipient with a Large Pre-Treatment X4 Population Related to the Failure Sequence



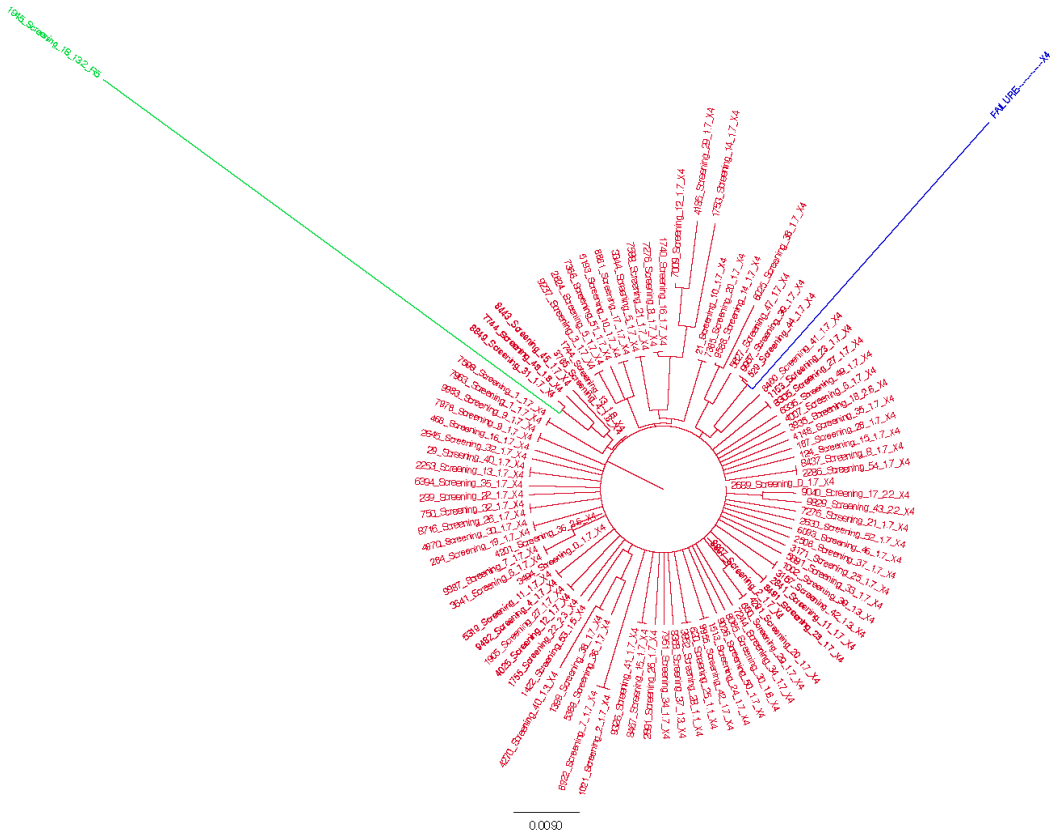
Appendix IX: Phylogenetic Tree from a Maraviroc Recipient with a Large Pre-Treatment X4 Population Related to the Failure Sequence
Screening sample: 92% X4 by deep sequencing, and R5 by ESTA
Failure sample: X4 by population-based genotype, and Dual/Mixed by original Trofile assay.
R5 sequences shown in green, X4 sequences shown in red, failure sequence shown in blue.

Appendix X: Phylogenetic Tree from a Maraviroc Recipient with a Large Pre-Treatment X4 Population Related to the Failure Sequence



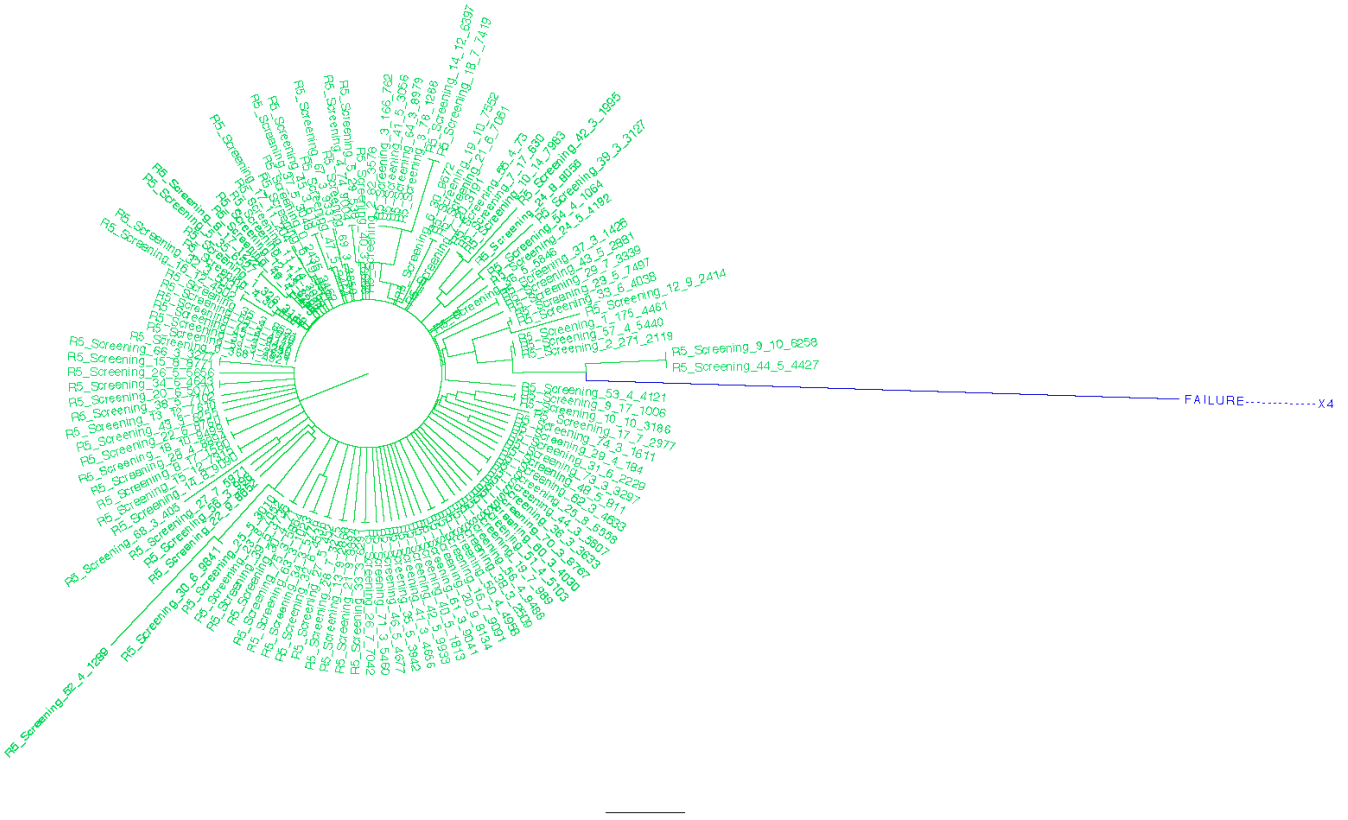
Appendix X: Phylogenetic Tree from a Maraviroc Recipient with a Large Pre-Treatment X4 Population Related to the Failure Sequence
Screening sample: 97% X4 by deep sequencing and R5 by ESTA
Failure sample: X4 by population-based genotype, and Dual/Mixed by original Trofile assay.
X4 sequences are shown in red, R5 sequences are shown in green, and the failure sequence is shown in blue.

Appendix XI: Phylogenetic Tree from a Maraviroc Recipient with a Large Pre-Treatment X4 Population Related to the Failure Sequence



Appendix XI: Phylogenetic Tree from a Maraviroc Recipient with a Large Pre-Treatment X4 Population Related to the Failure Sequence
Screening sample: 99.9% X4 by deep sequencing, and Dual/Mixed by ESTA
Failure sample: X4 by population-based genotype, and Dual/Mixed by original Trofile assay
R5 sequences shown in green, X4 sequences shown in red, and failure sequence shown in blue.

Appendix XII: Phylogenetic Tree from a Maraviroc Recipient for Whom Deep Sequencing Failed to Detect a Pre-Treatment X4 Population Despite Failure with X4 HIV

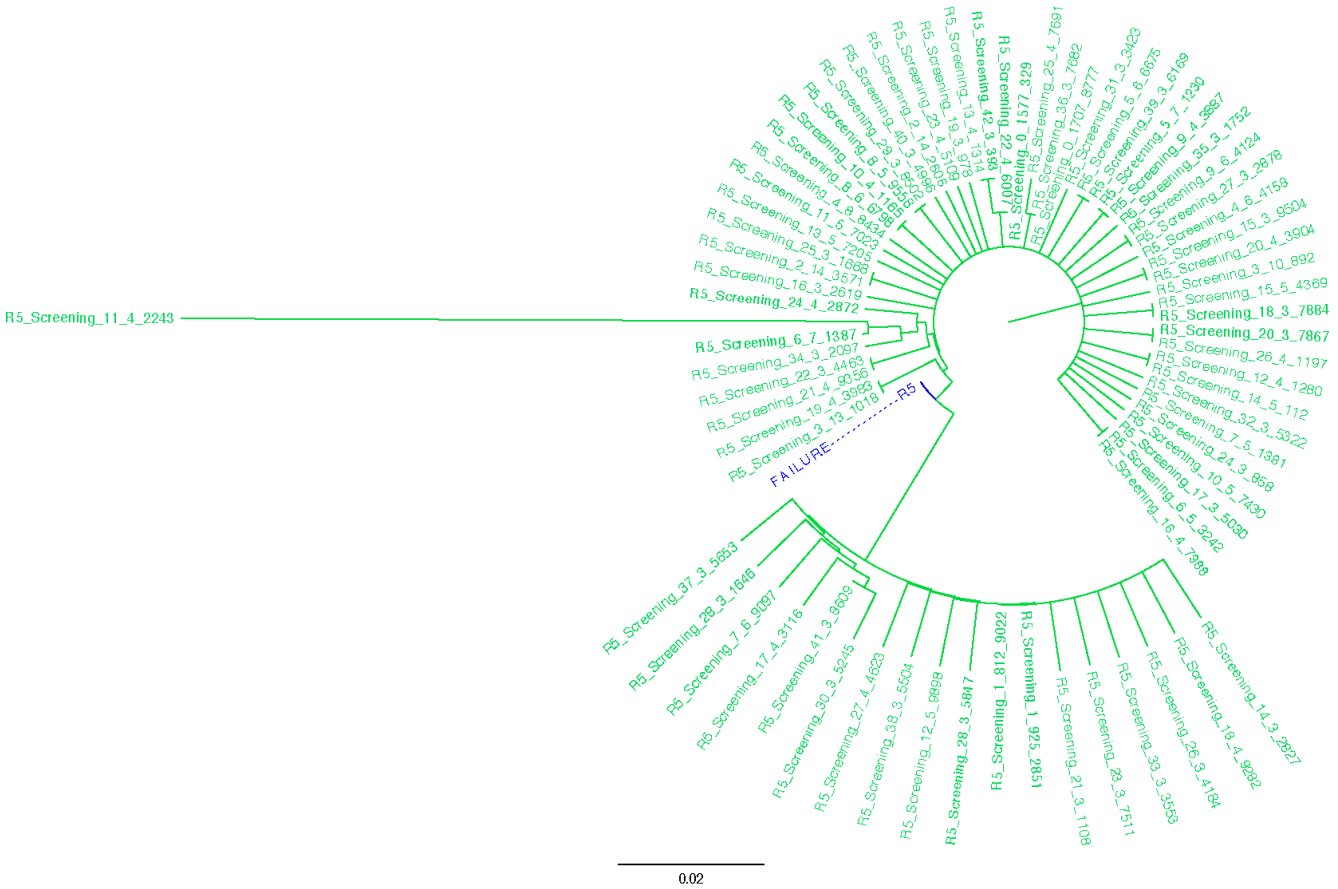


Appendix XII: Phylogenetic Tree from a Maraviroc Recipient for Whom Deep Sequencing Failed to Detect a Pre-Treatment X4 Population Despite Failure with X4 HIV
Screening sample: 0% X4 by deep sequencing, and R5 by ESTA
Failure sample: X4 by population-based genotype, and Dual/Mixed by original Trofile assay
R5 sequences shown in green, failure sequence shown in blue.

Appendix XIII: Phylogenetic Tree from a Maraviroc Recipient Who Failed with R5 HIV

Appendix XIII: Phylogenetic Tree from a Maraviroc Recipient Who Failed with R5 HIV
Screening sample: 0% X4 by deep sequencing, and Dual/Mixed by ESTA
Failure sample: R5 by population-based genotype, and R5 by original Trofile assay
R5 sequences shown in green, failure sequence shown in blue.

Appendix XIV: Phylogenetic Tree from a Maraviroc Recipient Who Failed with R5 HIV



Appendix XIV: Phylogenetic Tree from a Maraviroc Recipient Who Failed with R5 HIV
Screening sample: 0% X4 by deep sequencing, and R5 by ESTA
Failure sample: R5 by population-based genotype, and R5 by original Trofile assay.
R5 sequences shown in green, failure sequence shown in blue.