USING GENOTYPIC AND PHENOTYPIC METHODS TO DETERMINE THE HIV CO-RECEPTOR PHENOTYPE IN THE CLINICAL SETTING

by

Andrew John Low

B.Eng., The University of Victoria, 2002

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Experimental Medicine)

THE UNIVERSITY OF BRITISH COLUMBIA

December 2007

ABSTRACT

Objective: The human immunodeficiency virus type 1 (HIV-1) currently infects over 30 million people worldwide. It uses one of two main co-receptors to infect cells. The primary objective of this thesis is to evaluate genotypic and phenotypic assays for co-receptor usage in the clinical setting and investigate approaches for improvement of these assays.

Methods: The concordance of recombinant co-receptor phenotyping assays and the predictive ability of genotype-based methods including the '11/25' rule, position specific scoring matrices (PSSMs), and support vector machines (SVMs) were evaluated in the clinical setting using patient-derived plasma samples. Samples and patient data were evaluated in cross-sectional analyses from a retrospective population-based cohort of HIV-infected individuals enrolled in the HIV/AIDS Drug Treatment Program in British Columbia, Canada.

Results: Current implementations of HIV V3 region-based predictors for HIV co-receptor usage tested on patient derived samples are inadequate in the clinical setting, primarily due to low sensitivities as a result of difficult to detect minority species. Recombinant phenotype assays also show discordances when tested against each other on the same set of patient derived samples, raising doubts if any of these assays can truly be considered a 'gold standard'. Significant associations between clinical progression, viral sequence-based predictors of co-receptor usage and the output of recombinant assays are observed, suggesting that sensitivity can be improved by incorporating CD4%

into genotype-based predictors. This is verified with a SVM model which showed a 17% increase in sensitivity when CD4% was incorporated into training and testing.

Conclusion: This work in this thesis has exposed the difficulty in determining the co-receptor phenotype in the clinical setting, primarily due to minority species. Although genotypic methods of screening for HIV co-receptor usage prior to the administration of CCR5 antagonists may reduce costs and increase turn-around time over phenotypic methods, they are currently inadequate for use in the clinical setting due to low sensitivities. Although the addition of clinical parameters such as CD4 count significantly increases the predictive ability of genotypic methods, the presence of low-levels of X4 virus continues to reduce the sensitivity of both genotypic and phenotypic methods.

TABLE OF CONTENTS

ABSTRACT ii

TABLE OF CONTENTS iv

LIST OF TABLES viii

LIST OF FIGURES ix

LIST OF ILLUSTRATIONS xi

ACKNOWLEDGEMENTS xii

CO-AUTHORSHIP STATEMENT xiii

CHAPTER 1

GENERAL INTRODUCTION AND THESIS OBJECTIVES 1

1.1 Thesis Objectives and Organization 1

1. 2 Global Impact of the Human Immunodeficiency Virus 3

1.3 Viral Pathogenesis 4

1.4 Virus Structure and Replication 6

1.5 Antiviral Therapy: Brief History and Current Standard of Care 10

1.6 Novel Therapeutic Options 13

1.7 List of Illustrations 14

1.7 References 24

CHAPTER 2

LITERATURE REVIEW: HIV CO-RECEPTOR PHENOTYPING IN THE CLINICAL SETTING 42

2.1 Introduction 42

2.2 Chemokine Receptors and Viral Entry 43

2.3 History of Co-receptor Assays 45

2.4 Co-receptor Phenotype over the Natural History of HIV Infection and Theories on the Co-receptor Switch 49

2.5 Genotypic markers of co-receptor usage 53

2.6 Relevance and Limitations of Co-receptor usage in the era of Co-receptor Antagonists 56

2.7 Summary 59

2.8 References 61

CHAPTER 3

CURRENT V3 GENOTYPING ALGORITHMS ARE INADEQUATE FOR PREDICTING X4 CO-RECEPTOR USAGE IN CLINICAL ISOLATES 87

3.1 INTRODUCTION 87

3.2 METHODS 90

3.3 RESULTS 94

3.4 DISCUSSION 97

3.5 REFERENCES 109

CHAPTER 4

IS THERE A GOLD STANDARD FOR DETERMINING HIV CO-RECEPTOR USE IN CLINICAL SAMPLES? THE DEGREE OF CORRELATION BETWEEN TWO PHENOTYPIC ASSAYS AND A BIOINFORMATIC MODEL 118

4.1 INTRODUCTION 118

4.2 METHODS 121

4.3 RESULTS 123

4.4 DISCUSSION 128

4.5 REFERENCES 138

CHAPTER 5

CD4-DEPENDENT CHARACTERISTICS OF CO-RECEPTOR USE AND HIV-1 V3 SEQUENCE IN A LARGE POPULATION OF THERAPY-NAÏVE INDIVIDUALS 143

5.1 INTRODUCTION 143

5.2 METHODS 146

5.3 RESULTS 150

5.4 DISCUSSION 156

5.5 REFERENCES 169

CHAPTER 6

PREDICTING HIV CO-RECEPTOR USAGE ON THE BASIS OF GENETIC AND CLINICAL COVARIATES 180

6.1 INTRODUCTION 180

6.2 METHODS 184

6.3 RESULTS 188

6.4 DISCUSSION 193

6.5 REFERENCES 204

CHAPTER 7

GENERAL DISCUSSION AND CONCLUSION 211

7.1 Introduction 211

7.2 Sensitivity of HIV Genotype-Based Co-receptor Prediction Algorithms 213

7.3 Comparison of HIV Recombinant Co-receptor Phenotype Assays: Is there a 'Gold Standard'? 215

7.4 Correlations of the output of a Recombinant Co-receptor Phenotype Assay with V3 Genotype and Clinical Immune Status 217

7.5 Using Clinical Parameters to Improve Prediction of HIV V3 Genotype-Based Predictive Algorithm 219

7.6 Concluding Discussion 220

7.7 References 222

APPENDIX 1: Certificates of Ethical Approval 228

LIST OF TABLES

Table 3.1: Initial Predictive performance of V3 genotyping algorithms. 102

Table 3.2: Clonal analysis to determine levels of undetected minority species. 103

Table 4.1: Potentially relevant differences in phenotypic assays. 131

Table 4.2: Summary of median baseline CD4 count and baseline viral load for samples categorized as either R5 or R5/X4 using the Trofile or "TRT" co-receptor usage assays. 132

Table 4.3: Characteristics of samples with discordant results from the Trofile or TRT entry co-receptor assays 133

Table 6.1: V3 mutations significantly associated with co-receptor usage in univariate analyses of the population-based data. 197

Table 6.2. Univariate association of individual parameters with co-receptor usage. 198

LIST OF FIGURES

Figure 3.1: Sensitivity of genotype-based predictors stratified by CXCR4 RLU. 104

Figure 3.2: Positive predictive value for X4-tropic virus stratified by CD4 count. 106

Figure 4.1: Receiver Operating Characteristic Curve of Trofile HIV co-receptor assay vs. TRT co-receptor assay. 134

Figure 4.2: Support Vector Machine predicted probability of CXCR4 co-receptor usage 135

Figure 4.3: Boxplot summary of predicted probability of CXCR4 co-receptor usage... 136

Figure 4: Concordances between phenotypic co-receptor assays and a SVM bioinformatic predictor. 137

Figure 5.1: Distribution of CCR5 and CXCR4 RLU for R5 and DM samples. 163

Figure 5.2: Relative Light Units (RLUs) from CD4/CCR5 and CD4/CXCR4 cells from samples identified as using the CCR5 or CXCR4 co-receptors. 164

Figure 5.3: Comparison of the percentage of sequence conservation and local net charge in samples with R5 and DM phenotype. 165

Figure 5.4: Proportions and prevalence of samples with specific patterns of V3 sequence. 166

Figure 5.5: Associations of V3 loop predictors of X4-tropic HIV with CXCR4 RLU in samples phenotyped as DM. 167

Figure 6.1. Predictive performance of the 11/25 rule and five statistical learning methods, assessed on clonal data. 199

Figure 6.2. V3 loop predictors of CXCR4 usage. 200

Figure 6.3. Predicting co-receptor usage on population-based data. 201

Figure 6.4. Feature ranking for the SVM trained on the clinical dataset..... 202

LIST OF ILLUSTRATIONS

Illustration 1.1. AIDS was the leading cause of death in those aged 25-44 in the USA by the year 1992 14

Illustration 1.2. Natural History of HIV Infection 15

Illustration 1.3. Structure of HIV 16

Illustration 1.4. HIV-1 genome, HXB2 strain 17

Illustration 1.5. HIV-1 life cycle 18

Illustration 1.6. Chemical structure of the NRTIs Zidovudine and Lamivudine 19

Illustration 1.7. Chemical structures of NNRTIs Nevirapine and Delavirdine 20

Illustration 1.8. Chemical structure of the PIs Lopinavir and Nelfinavir 21

Illustration 1.9. Chemical Structure of the integrase inhibitor Raltegravir 22

Illustration 1.10. Chemical structure of the CCR5-antagonists Maraviroc and Vicriviroc 23

ACKNOWLEDGEMENTS

I would like to extend my sincere thanks to my thesis supervisor, Dr. Richard Harrigan, as well as my thesis committee members, Dr. Julio Montaner, and Dr. Artem Cherkasov, for their continuing direction, support, encouragement and mentorship.

I would also like to thank my fellow students and coworkers who have provided invaluable expertise, advice, support, and encouragement throughout the years: Zabrina Brumme, Chanson Brumme, Dennison Chan, Celia Chui, Winnie Dong, Vikram Gill, Erin Hudson, Theresa Mo, Emilie Stevens, Conan Woods, and Brian Wynhoven.

I gratefully acknowledge the financial support of the British Columbia Centre for Excellence in HIV/AIDS for their support through graduate stipend.

CO-AUTHORSHIP STATEMENT

Each of chapters 3-6 consists of a manuscript that is published or accepted in an international, peer-reviewed journal. The candidate is either the lead author (chapters 3, 5), co-lead author (chapter 4), or second author (chapter 6) on all of these manuscripts.

This statement is to certify that the candidate was the major contributor to study identification and design, supervision of laboratory data collection, as well as statistical and data analysis. The candidate also wrote the manuscripts.

Primary co-authors of these manuscripts include the thesis supervisor (Dr. Richard Harrigan, thesis committee members (Dr. Julio Montaner), statistical analysts (Chanson Brumme), authors who provided data access and assistance with data analysis (Chanson Brumme, Dennison Chan), laboratory technical staff and students (Zabrina Brumme, Winnie Dong, Theresa Mo, and Peter Cheung) and external collaborators.

CHAPTER 1

GENERAL INTRODUCTION AND THESIS OBJECTIVES

1.1 Thesis Objectives and Organization

This thesis is divided into seven chapters, and is prepared according to the University of British Columbia manuscript-based format. This first chapter provides a general introduction to HIV pathogenesis, structure and treatment. Chapter two is a detailed literature review about determining the HIV co-receptor phenotype in the clinical setting. Chapters three through six address main theses objectives as outlined below. Chapter seven discusses and summarizes the overall study results, and comments upon the implications and relevance of these findings in the context of the current state of knowledge. References are found at the end of each chapter.

The general aim of this thesis is to evaluate the performance of genotype-based predictors of co-receptor usage, as well as recombinant co-receptor phenotype assays in the clinical setting with the purpose of investigating how they can be improved. All studies are population-based studies focused on a cross-sectional population of therapy naïve individuals. These studies evaluate the performance of currently available HIV envelope sequence-based algorithms for predicting HIV entry (chapter three) as well as the concordance between recombinant phenotype assays for determining HIV entry on patient-derived samples (chapter four). This is followed by an analysis of how known viral and clinical determinants of HIV entry correlate with the output of a recombinant phenotype entry assay (chapter five). The final chapter (chapter 6) includes an evaluation

of a modified envelope sequence-based algorithm which uses other clinical and viral determinants of HIV entry to improve performance on patient-derives samples. Chapters three through six each represent stand-alone manuscripts. Chapters three and four have already been published, while chapters five and six are currently in press in international, peer-reviewed journals. Journals in which the individual thesis chapters have been published include *Journal of Clinical Microbiology* and *AIDS.* Journals on which the individual thesis chapters are in press include *Antiviral Therapy* and *AIDS Research and Human Retroviruses*. The candidate is either the lead author (chapters three, and five), co-first author (chapter four) or second author (chapter six) on all manuscripts.

1. 2 Global Impact of the Human Immunodeficiency Virus

Between June 1, 1981, and September 15, 1982, a total of 593 cases of acquired immune deficiency syndrome (AIDS) were reported in the United States [1]. In the subsequent year, a novel pathogenic retrovirus, belonging to the group of viruses know as human T-cell leukemia viruses (HTLV), was isolated in patients with AIDS and some of those at risk for developing AIDS [2,3]. Initially named HTLV type III (HTLV-III), this human retrovirus, subsequently designated as the human immunodeficiency virus type 1 (HIV-1), was identified as a novel pathogenic retrovirus and the cause of AIDS [4-6]. Shortly afterwards, HIV-2 was isolated from Western Africa as an etiological agent of AIDS distinct from HIV-1 [7]. HIV would subsequently be responsible for a quickly growing pandemic. Between 1992 and 1996, AIDS was the leading cause of death in those aged 25-44 in the USA [8] (Illustration 1.1). By the end of 2006, a total of 39.5 million people (37.2 million adults and 2.3 million children under 15 years of age) worldwide were estimated to be living with HIV, with sub-Saharan Africa representing a disproportionately large number of infections (24.7 million) [9]. During 2006, an estimated 4.3 million people worldwide were newly infected with HIV, and approximately 2.9 million people died as a result of AIDS associated infections. Almost three quarters (72%) of all deaths due to AIDS in 2006 occurred in sub-Saharan Africa [9]. In North-America, approximately 1.4 million adults and children are currently living with HIV/AIDS, with 18,000 new infections annually by the end of December, 2006 [9]. In Canada, the prevalence of HIV/AIDS was estimated at 58,000 by the end of 2005. This represented an increase of 16% over the 2002 estimate of 50,000, and included an estimated 15,800 people (27%) who were unaware of their infection status [10].

1.3 Viral Pathogenesis

The hallmark of HIV-1 infection is the depletion of CD4+ T-lymphocytes, thus leading to immunodeficiency. In the absence of antiretroviral treatment, infection and destruction of CD4+ T-lymphocytes impairs almost all facets of the adaptive immune system, leaving infected individuals progressively more susceptible to opportunistic bacterial, viral, parasitic and/or fungal infections, and eventually results in death.

Worldwide, most new HIV infections are sexually transmitted [11]. During sexual transmission, the virus enters the body via the mucosal surfaces of the genital tract, and is believed to first come into contact with Langerhans cells present in the epidermis and mucosal epithelium, as well as macrophages present in the sub-mucosal tissues [12]. Freshly isolated Langerhans cells express CCR5, but not CXCR4, while macrophages express both CCR5 and CXCR4. It is believed that macrophages and dendritic cells contribute to continued pathogenesis by disseminating the virus throughout the body, where the virus then begins to productively infect CD4+ T-lymphocytes [13-16].

The natural history of HIV disease progression varies widely between infected individuals (Illustration 1.2). Rare individuals may remain uninfected despite multiple exposures [17,18] while others progress to AIDS and death within a relatively short period of time (months to years) following infection [19]. In one circumstance, an individual appears to have cleared the virus completely [20], although it is still possible that latent virus may re-emerge. Certain individuals with specific human genetic factors and/or attenuated viral strains may remain symptom-free and maintain CD4 cell counts

greater than 500-600 cells/mm^3 for greater than 7-10 years (one definition of long-term non-progression (LTNP)) [21-24], where the normal range for CD4 counts in seronegative (negative for HIV antibodies) individuals is between 500-1500 cells/mm^3. During the natural course of infection, a patient's plasma viral load (pVL), which measures the amount of HIV RNA in the peripheral plasma [25] increases sharply during primary or acute infection. This is usually accompanied by flu-like symptoms, named 'acute seroconversion illness', although other more severe symptoms may also manifest [26]. The acute stage of infection is also accompanied by a significant drop in the CD4+ T-lymphocyte count, or CD4 count (a surrogate marker of immunological function in HIV-infected individuals [27]). The acute state of infection is followed by a subsequent increase in the CD4 count, mirrored by a drastic drop in the plasma viral load, establishing a viral 'set point', and marking the beginning of the asymptomatic phase of HIV infection [28]. The early stage of infection is characterized by continuous viral replication and a gradual loss of CD4+ lymphocytes, with CD4 counts remaining above 500 cells/mm^3 in approximately 50% of patients [8]. When CD4 counts range between 200 and 500 cells/mm^3 and symptoms are usually still absent or mild, the disease has progressed to intermediate stage of infection. After CD4 counts drop below 200 cells/mm^3, and at least one AIDS-defining illness is evident, infection has progressed to late-stage disease [29]. Late-stage disease is typically accompanied by a resurgence in plasma viral load to levels similar to those seen in primary infection. The current definition of AIDS, as classified by the U.S. Centres for Disease Control, is the diagnosis of an AIDS-defining illness, a CD4+ T-lymphocyte count of less than 200 cells/mm^3 or a CD4+ percentage less than 14% [30].

1.4 Virus Structure and Replication

HIV-1 is approximately 110 nm in diameter, consisting of a protein core or 'capsid' surrounded by a lipid bilayer membrane or 'envelope'. The three-dimensional structure of this lipid bilayer is held in place by underlying 'matrix' proteins. Most major structural and regulatory proteins necessary for viral replication and infectivity are encoded within two single strands of positive-sense RNA, each approximately 9700 base-pairs long [31], encapsidated and stabilized by a nucleocapsid protein (Illustration 1.3). The HIV-1 genome is organized into three large open reading frames (*gag*, *pol*, and *env*), encoding the structural and non-structural components common to all replication-competent retroviruses [32,33] (Illustration 1.4). The HIV-1 genome also includes reading frames for regulatory proteins (*Tat* and *Rev*), and accessory proteins (*Nef*, *Vif*, *Vpr*, and *Vpu*) [32,33]. The *gag* gene encodes the Gag polyproteins precursor (pr55) which is cleaved by the viral protease to produce the matrix (p17), capsid (p24), and nucleocapsid (p7) proteins, thus comprising the major structural components of the virion. The HIV-1 *pol* gene is encoded in a different reading frame than *gag* and is initially transcribed as a larger *gag-pol* polyprotein precursor (pr160) resulting from approximately 5-10% of all transcriptional events as a result of a ribosomal frameshift [34]. The *pol* gene encodes the major functional enzymes necessary for converting the viral RNA into proviral DNA, integrating the proviral DNA into the host genome and maturing the newly produced virus. These include the reverse transcriptase enzyme and ribonuclease H (RNaseH) domain, required for reverse transcription of the viral RNA, as well as the HIV integrase, the enzyme responsible for integration of the HIV proviral DNA into the host genome. The HIV protease is the viral enzyme required for the

maturation of new virions. The *env* gene also encodes a polyprotein precursor (gp160), which is cleaved by cellular proteases to produce the envelope glycoproteins, gp120 and gp41, which form the surface and transmembrane components of the HIV envelope, respectively [35].

The HIV life cycle can be broken down into five main steps: binding and fusion of the free virion to the host cell; reverse transcription of the viral RNA into double stranded DNA (dsDNA); integration of the dsDNA into the host cell genome; transcription and assembly of viral proteins; and finally, budding and maturation of new virus. Binding and fusion are initiated when the CD4 binding domain of HIV gp120 (envelope) contacts a CD4 receptor on the surface of a CD4-expressing cell [36] (Illustration 1.5). After binding to the CD4 receptor, HIV gp120 undergoes a conformational change by exposing another binding site [37]. This binding requires a co-receptor on the cell surface to complete viral binding and fusion. These human chemokine co-receptors were identified as belonging to the seven transmembrane g-protein coupled receptor family, later named CXCR4 and CCR5 [38-43]. While the CXCR4 receptor is expressed primarily on activated CD4+ T-lymphocytes, the CCR5 receptor is expressed primarily on naïve and memory CD4+ T-lymphocytes and macrophages, which are the primary cellular targets of HIV infection. Rarely, other HIV co-receptors have also been identified to mediate infection of HIV-1 *in vitro*, but these are unlikely to play a significant role in infection *in vivo* [43].

Binding of the gp120 protein to the CD4 receptor and chemokine co-receptor induces a conformational change to the gp120 and gp41 envelope proteins resulting in the insertion of the hydrophobic gp41 fusion peptide (located at the N-terminus of gp41) into the host cell membrane. Virus-cell fusion is followed by the uptake of the viral core or capsid into the host cell cytoplasm. Once inside the cell, the viral core undergoes a poorly understood 'uncoating' process, whereby capsid protein molecules (p24) become partially dissociated and the viral genetic material, accessory molecules, and enzymes are released into the cytoplasm [32]. The HIV reverse transcriptase then converts the single-stranded RNA template into double-stranded DNA (dsDNA), and a preintegration complex (PIC) forms. The PIC consists of the dsDNA, reverse transcriptase, integrase, Vpr and matrix proteins (p24) [44], and allows the virion components to cross through the nuclear membrane, resulting in infection of non-dividing cells [45]. Once inside the nucleus, the retroviral DNA is integrated into the host cell DNA through the concerted action of the viral integrase and host cell DNA polymerase as well as other host cell factors. The viral DNA can integrate at virtually any sequence; however, integration is more efficient in actively transcribed regions where the DNA may be more accessible to the integrase [46]. Once integrated, the viral DNA is referred to as proviral DNA, and serves as a template for the synthesis of viral RNA. HIV RNA transcripts are initiated at the 5'Long-Terminal Repeat (LTR), located at the 5' end of the viral genome. During the early stages of infection, viral mRNAs encoding Tat, Rev and Nef accumulate. Tat increases the level of HIV mRNA transcription from the 5' LTR, while Rev mediates the expression of full-length mRNA transcripts. These include full-length genomic RNAs, (for inclusion into new viral particles), singly and doubly spliced mRNAs, and

multicistronic (single transcripts encoding multiple proteins) mRNAs which serve as templates for the synthesis of structural, enzymatic and accessory proteins [34,47]. In contrast to Gag and Gag-Pol, the Env polyproteins are translated on ribosomes associated with the endoplasmic reticulum, allowing them to be cleaved by host cell proteases. This is followed by glycosylation, producing glycosylated Env precursor proteins (gp160). These precursor proteins are then transported to the Golgi Apparatus where the sugar side chains are processed [48]. Finally, the gp160 precursor protein is cleaved into gp120 and gp41 by the host protease furin during transport to the cell surface [49].

After synthesis, the Gag (pr55) and Gag-Pol (pr160) polyproteins migrate to the host cell plasma membrane, targeted by the matrix domain (p24) of Gag [32], where the assembly of new viral particles takes place. The Gag proteins are also responsible for packaging several components of the virus particle, including viral RNA and the viral accessory protein Vpr. The Gag p6 protein, located at the 3' end of Gag, includes the Late (L) domain responsible for mediating the release of immature virions from the plasma membrane [50]. The gp120-gp41 complex binds to the budding Gag shell in the plasma membrane. After the release of immature progeny virions, a final maturation step begins as the two protease domains (a homodimer) of the Gag-Pol precursor proteins become active and cleave themselves, and subsequently cleave the remainder of the Gag and Gag-Pol polyproteins, releasing them into their various constituents [51].

1.5 Antiviral Therapy: Brief History and Current Standard of Care

Prior to the introduction of antiretroviral therapy, AIDS-associated mortality had steadily increased in North America and by 1994 eclipsed all other causes of death, becoming the leading cause of mortality for people aged 25-44 in the United States of America [8]. In the absence of antiretroviral treatment, the median time of survival for HIV-infected individuals was approximately 5 to 11 years, depending on age at serconversion [52], and the median survival time following an AIDS diagnosis was only 9.5 to 12 months [53,54]. Zidovudine (AZT) was introduced as the first nucleoside analogue reverse transcriptase inhibitor (NRTI) (Illustration 1.6). NRTIs inhibit reverse transcription by terminating synthesis of the HIV DNA. AZT was responsible for increasing the survival time for patients with AIDS [55,56]; however, in symptom-free individuals infected with HIV, AZT monotherapy was not associated with a significant difference in disease progression, although patients treated with AZT did have higher CD4 counts than patients treated with a placebo [57]. The subsequent development and licensure of the NRTIs Didanosine (ddI), zalcitabine (ddC) and lamivudine (3TC) allowed for the introduction of dual-combination therapies, and resulted in improved clinical responses and survival when compared to AZT monotherapy alone [58-64]. However, it soon became obvious that mono and dual combination therapy lost their effectiveness after a relatively short period of time, largely due to the selection of drug resistant HIV variants [65-70]. During the mid 1990's, additional drug classes were developed including the non-nucleotide reverse transcriptase inhibitors or NNRTIs (Illustration 1.7), which were capable of hindering the synthesis of HIV DNA through allosteric inhibition of the HIV reverse transcriptase enzyme, and the protease inhibitors

or PIs (Illustration 1.8), a drug class which prevents viral maturation by binding strongly to the active site of the HIV. These new classes were very effective at lowering plasma viral load concentrations in combination with NRTI regimens [71,72]. The safety and efficacy of these regimens allowed for the introduction, and subsequent universal recommendation of triple combination therapy. Triple combination therapy (also known as "Highly Active Antiretroviral Therapy", or HAART) resulted in dramatic and sustained reductions in HIV-associated mortality and morbidity [72-76]. In addition, the introduction of protease inhibitors 'boosted' with ritonavir have resulted in well tolerated regimens that have superior activity compared to non-boosted PI regimens [77]. Boosted PI regimens have further increased the efficacy of HAART and are currently the most-preferred method of administering protease inhibitors.

Currently, there are a total 8 nucleoside or nucleotide analogue reverse transcriptase inhibitors (NRTI) developed and approved by the U.S. Food and Drug Administration (FDA) and/or Health Canada. These include zidovudine, didanosine, zalcitabine, stavudine, lamivudine, abacavir, tenofovir, emtracitabine, and combinations thereof. Protease inhibitors represent the second class of HIV antiretroviral agents. They currently include ritonavir, nelfinavir, indinavir, amprenavir, saquinavir, lopinavir, atazanavir, tipranavir, darunavir, and fosamprenavir. The third class of antiretroviral agents, the NNRTIs, include nelfinavir, delavirdine, and efavirenz. Most first line regimens include a 'backbone' of two NRTIs in combination with either a boosted PI or NNRTI [78]. A fourth class of antiretroviral agents which inhibits virus-cell fusion by interacting with HIV gp41 is represented by the peptide enfuvirtide (T20). Although this

drug is more inconvenient to take, it has often been used successfully in regimens where the patient has become resistant to first or second line regimens [79,80]. Combination pills such as Atripla, which includes efavirenz, emtricitabine and tenofovir provide patients with a much easier regimen to follow, which encourages strong adherence, one of the strongest known predictors of acquiring antiretroviral drug resistance [81-83]. Regimens which encourage strong adherence are important because no therapy 'cures' HIV infection, therefore therapy is lifelong.

1.6 Novel Therapeutic Options

The development of novel antiretroviral agents which are effective against resistant strains as well as the development of novel drug classes is necessary in order to treat patients resistant to current regimens and ensure that future regimens remain available in case resistance develops. Some of these novel therapeutic classes include integrase inhibitors, such as raltegravir (Illustration 1.9) which suppress viral replication by inhibiting the integration step of the viral life cycle. Raltegravir is expected to be approved soon and in clinical trials, has demonstrated effective reductions in plasma viral load [84]. Another novel viral inhibition strategy has been to block viral entry via one of the chemokine co-receptors, CCR5 or CXCR4. Thus far, no CXCR4 antagonists have progressed beyond phase II of clinical trials; however, two CCR5 antagonists, mariviroc and vicriviroc (Illustration 1.10) have been shown to be effective at reducing plasma viral load in clinical trials [85,86], and maraviroc is expected to be approved shortly in Canada. Patients with a detectable CXCR4-using virus as determined by the Trofile Co-receptor Assay (Monogram Biosciences) showed no appreciable decrease in plasma viral load while undergoing CCR5-antagonist based therapy (mariviroc) [87], thus indicating a need for HIV co-receptor usage screening prior to co-receptor-based therapy.

Illustration 1.1: AIDS was the leading cause of death in those aged 25-44 in the USA by the year 1992

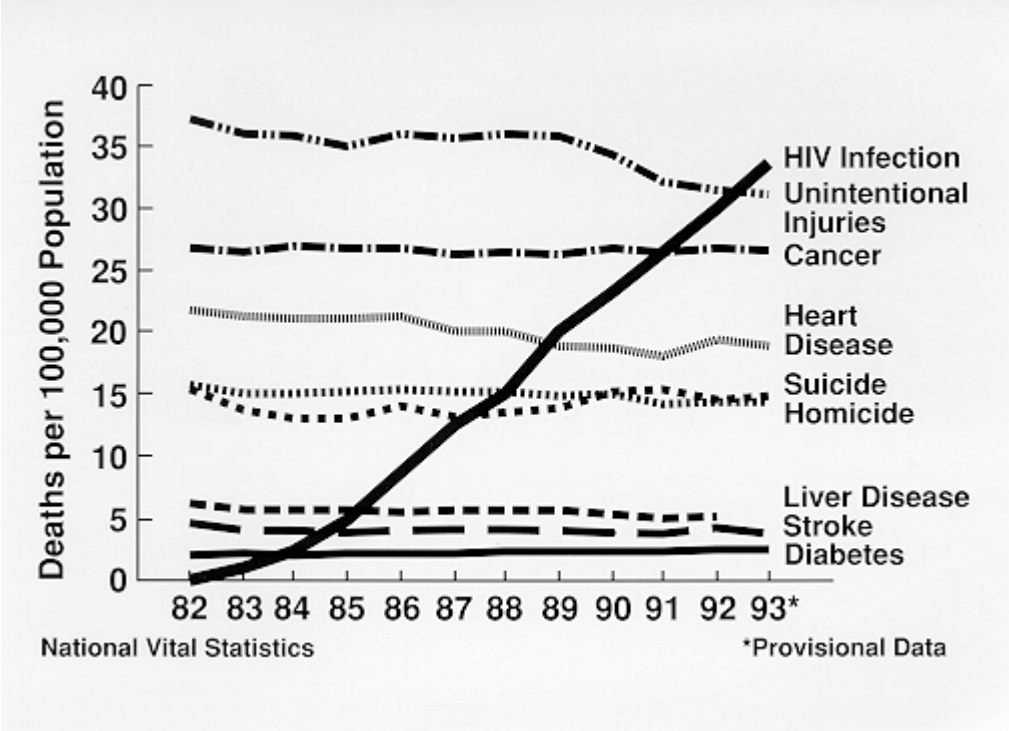
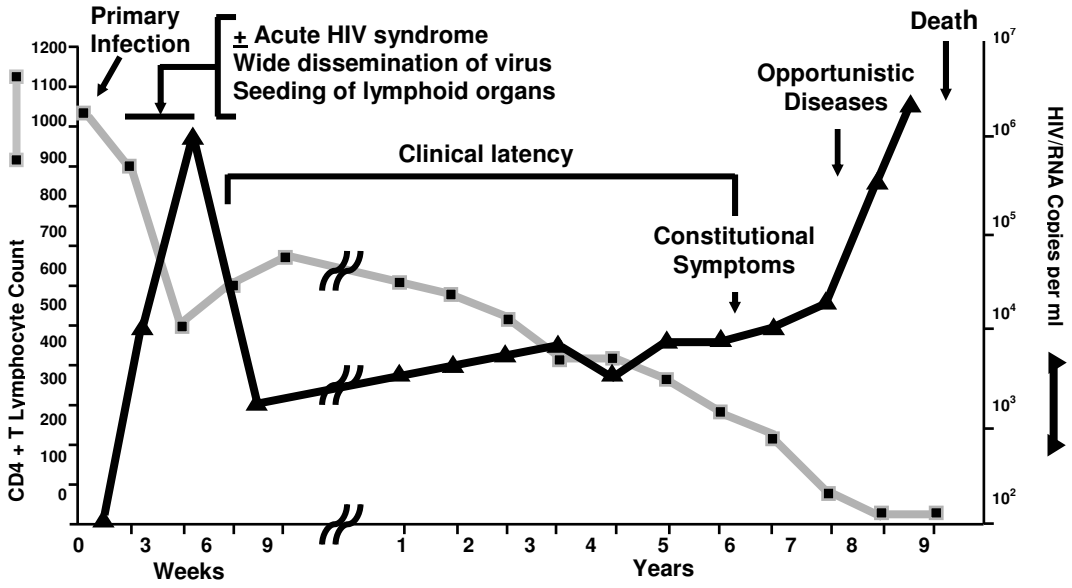


Illustration 1.2: Natural history of HIV infection



Adapted from Fauci T. et. al. Ann Intern Med. 1996 124(7):654-63

Illustration 1.3: Structure of HIV

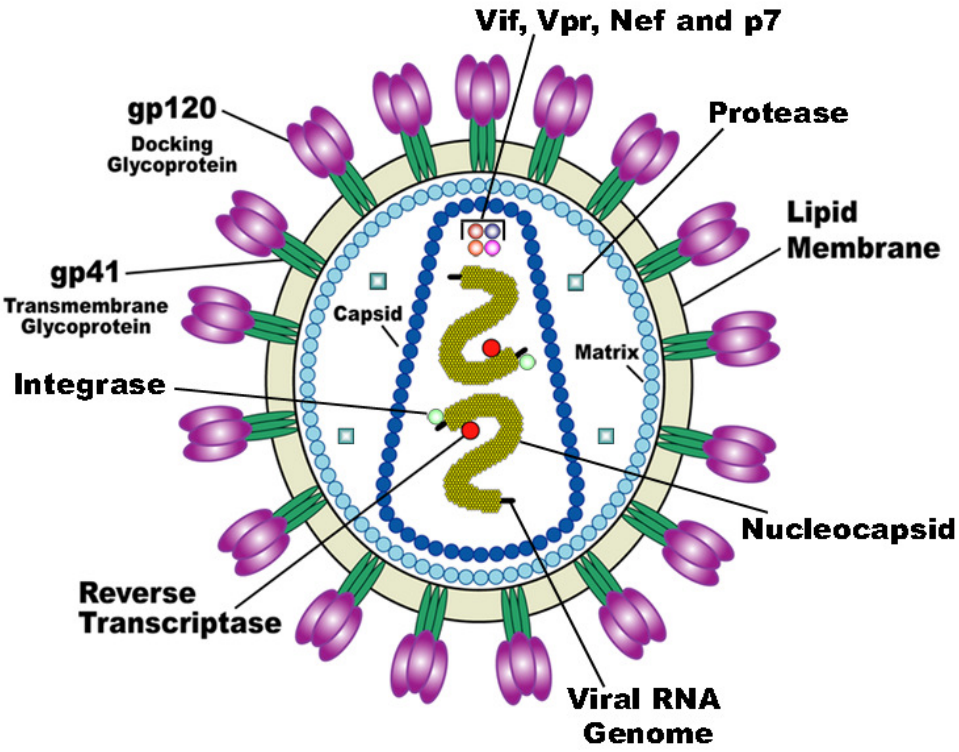
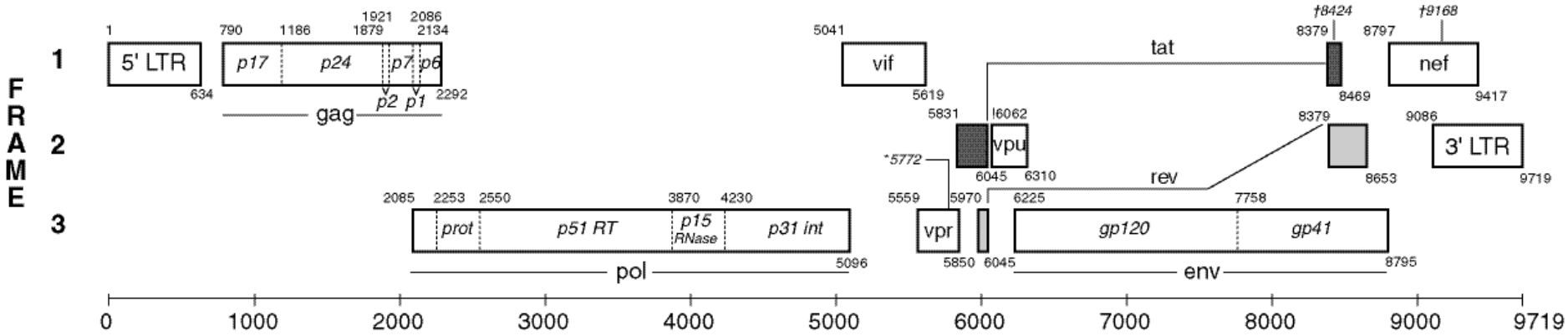


Illustration 1.4: HIV-1 genome, HXB2 strain



Open reading frames are shown as rectangles. The gene start, indicated by the small number in the upper left corner of each rectangle normally records the position of the a in the atg start codon for that gene while the number in the lower right records the last position of the stop codon. For pol, the start is taken to be the first t in the sequence ttttttag which forms part of the stem loop that potentiates ribosomal slippage on the RNA and a resulting -1 frameshift and the translation of the gag-pol polyprotein. The tat and rev spliced exons are shown as shaded rectangles. In HXB2, *5772 marks position of frameshift in the vpr gene caused by an "extra" t relative to most other subtype B viruses; !6062 indicates a defective acg start codon in vpu; †8424, and †9168 mark premature stop codons in tat and nef.

Adapted from Korber B. et. al. Numbering Positions in HIV Relative to HXB2CG, in the database compendium, *Human Retroviruses and AIDS*, 1998.

Illustration 1.5: HIV-1 life cycle

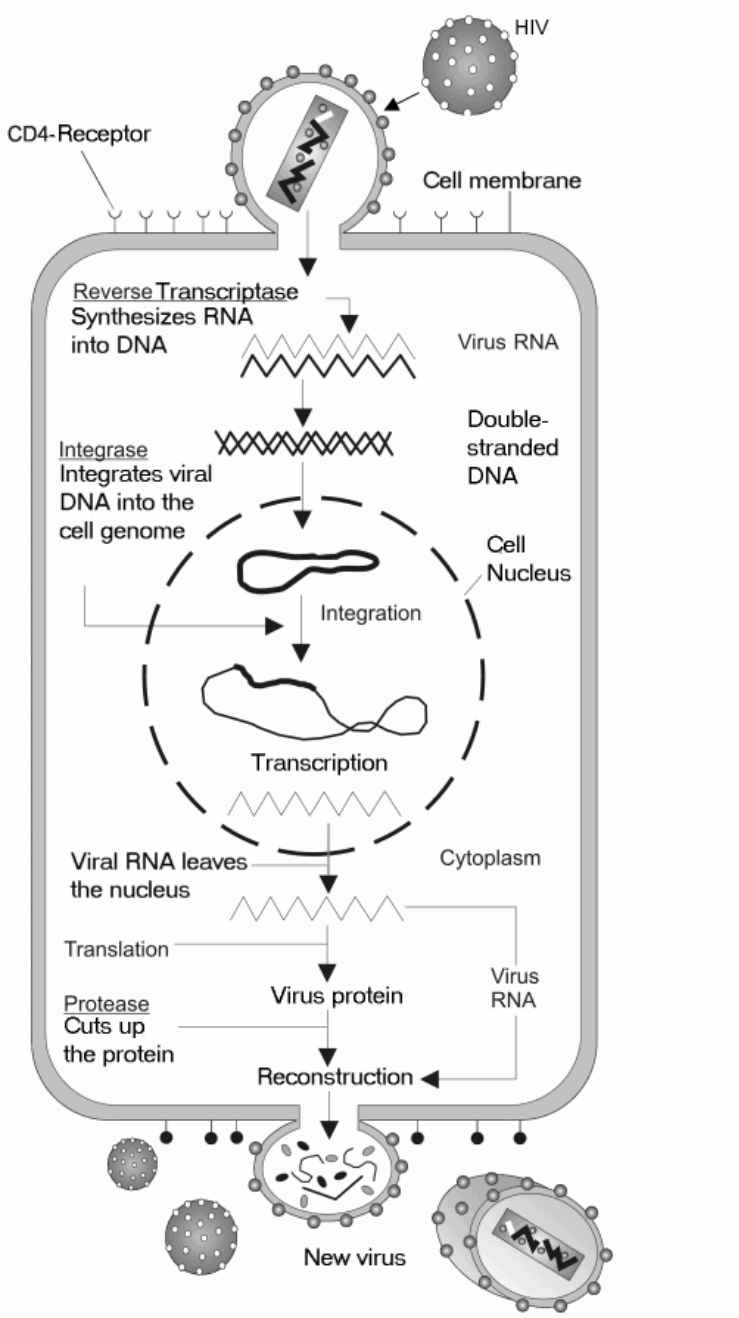
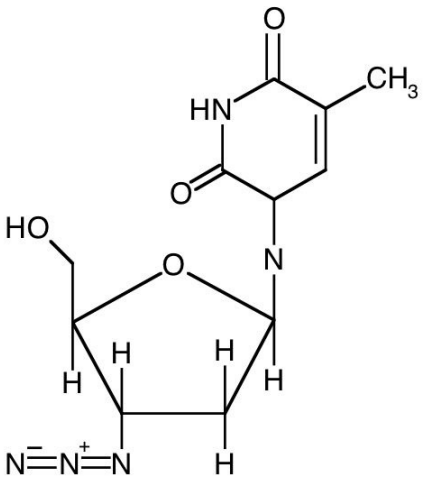
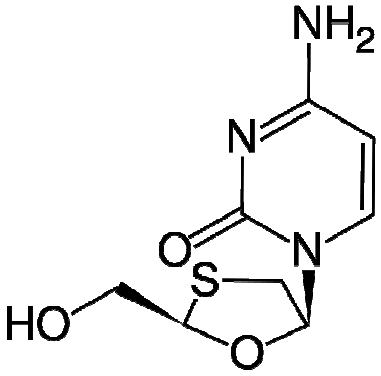


Illustration 1.6: Chemical structure of the NRTIs Zidovudine and Lamivudine

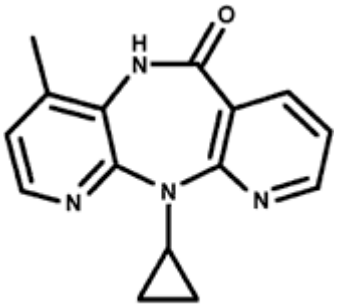


Zidovudine

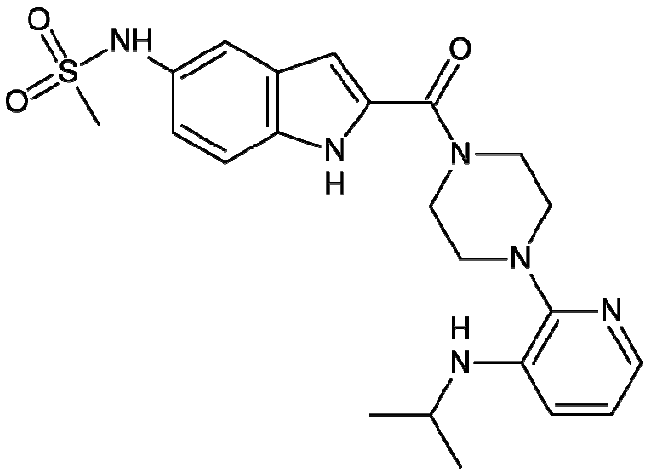


Lamivudine

Illustration 1.7: Chemical structures of NNRTIs Nevirapine and Delavirdine

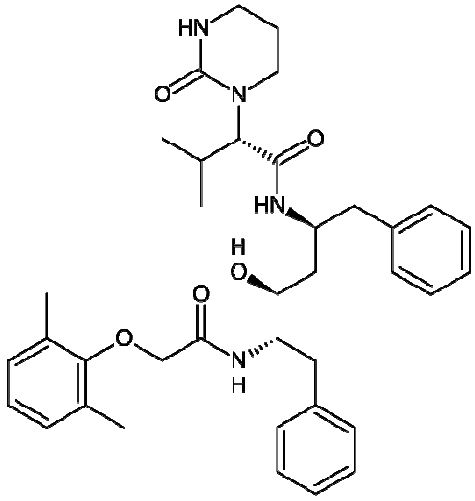


Nevirapine

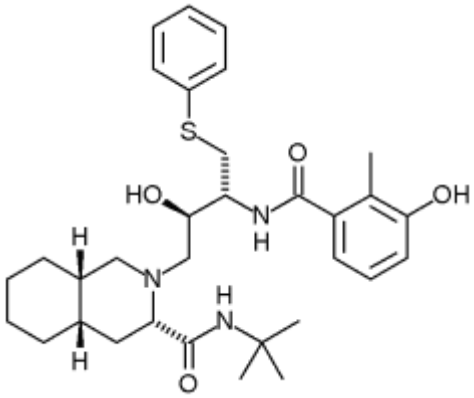


Delavirdine

Illustration 1.8: Chemical structure of the PIs Lopinavir and Nelfinavir

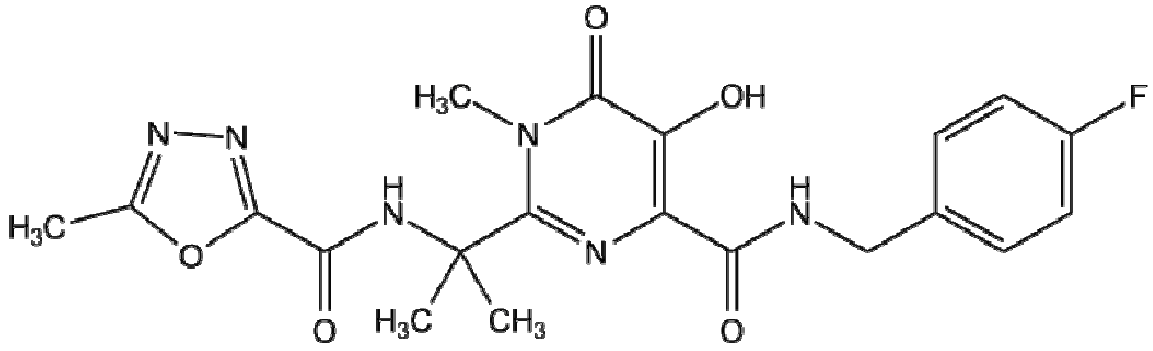


Lopinavir



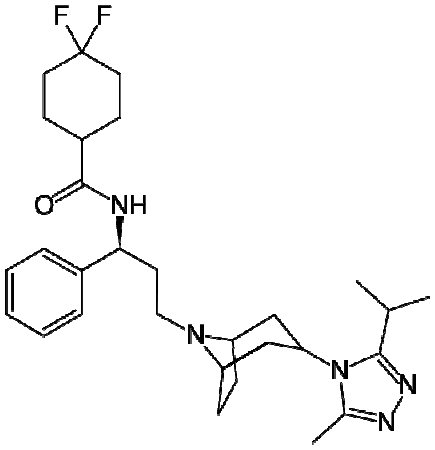
Nelfinavir

Illustration 1.9: Chemical Structure of the integrase inhibitor Raltegravir

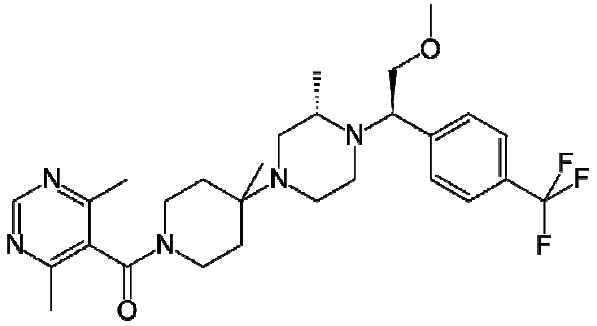


Raltegravir

Illustration 1.10: Chemical structure of the CCR5-antagonists Maraviroc and Vicriviroc



Maraviroc



Vicriviroc

1.7 References

1. Centers for Disease Control (CDC).Update on acquired immune deficiency syndrome (AIDS)-United States. *MMWR Morb Mortal Wkly Rep.* 1982; **31**(37):507-8, 513-4.

2. Gallo RC, Sarin PS, Gelmann EP, Robert-Guroff M, Richardson E, Kalyanaraman VS, et al. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science.* 1983; **220**(4599):865-867.

3. Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science.* 1983; **220**(4599):868-871.

4. Hahn BH, Shaw GM, Arya SK, Popovic M, Gallo RC,Wong-Staal F. Molecular cloning and characterization of the HTLV-III virus associated with AIDS. *Nature.* 1984; **312**(5990):166-169.

5. Alizon M, Sonigo P, Barre-Sinoussi F, Chermann JC, Tiollais P, Montagnier L, et al. Molecular cloning of lymphadenopathy-associated virus. *Nature.* 1984;

312(5996):757-760.

6. Broder S, Gallo RC. A pathogenic retrovirus (HTLV-III) linked to AIDS. *N Engl J Med.* 1984; **311**(20):1292-1297.

7. Clavel F, Guetard D, Brun-Vezinet F, Chamaret S, Rey MA, Santos-Ferreira MO, et al. Isolation of a new human retrovirus from west african patients with AIDS. *Science.* 1986; **233**(4761):343-346.

8. National Center for Injury Prevention and Control. WISQARS Leading Causes of Death Reports, 1981-1998 <http://webapp.cdc.gov/sasweb/ncipc/leadcaus9.html>. Accessed 08/21, 2007. Center for Disease Control and Prevention, Atlanta, Georgia, USA, July 10, 2007.

9. UNAIDS/WHO. 2006 AIDS epidemic update. WHO Library Cataloguing. AIDS Epidemic Update: special report on HIV/AIDS: December 2006

10. Boulos D, Yan P, Schanzer D, Remis RS,Archibald CP. Estimates of HIV prevalence and incidence in canada, 2005. *Can Commun Dis Rep.* 2006; **32**(15):165-174.

11. Quinn TC.Global burden of the HIV pandemic. *Lancet.* 1996; **348**(9020):99-106.

12. Miller CJ, Hu J. T cell-tropic simian immunodeficiency virus (SIV) and simian-human immunodeficiency viruses are readily transmitted by vaginal inoculation of rhesus macaques, and langerhans' cells of the female genital tract are infected with SIV. *J Infect Dis.* 1999; **179 Suppl 3**:S413-7.

13. Zhang Z, Schuler T, Zupancic M, Wietgrefe S, Staskus KA, Reimann KA, et al. Sexual transmission and propagation of SIV and HIV in resting and activated CD4+ T cells. *Science.* 1999; **286**(5443):1353-1357.

14. 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):383-387.

15. Kahn JO, Walker BD. Acute human immunodeficiency virus type 1 infection. *N Engl J Med.* 1998; **339**(1):33-39.

16. Ho DD, Rota TR,Hirsch MS. Infection of monocyte/macrophages by human T lymphotropic virus type III. *J Clin Invest.* 1986; **77**(5):1712-1715.

17. Kulkarni PS, Butera ST,Duerr AC. Resistance to HIV-1 infection: Lessons learned from studies of highly exposed persistently seronegative (HEPS) individuals. *AIDS Rev.* 2003; **5**(2):87-103.

18. McNicholl JM, Promadej N. Insights into the role of host genetic and T-cell factors in resistance to HIV transmission from studies of highly HIV-exposed thais. *Immunol Res.* 2004; **29**(1-3):161-174.

19. Phair J, Jacobson L, Detels R, Rinaldo C, Saah A, Schrager L, et al. Acquired immune deficiency syndrome occurring within 5 years of infection with human immunodeficiency virus type-1: The multicenter AIDS cohort study. *J Acquir Immune Defic Syndr.* 1992; **5**(5):490-496.

20. J. Zaunders, W. B. Dyer, M. Churchill, M. L. Munier, P. Cunningham, K. Suzuki, et al. Clearance of transfusion-acquired, attenuated nef-deleted HIV-1 infection by a long-term non-progressor with heterozygous CCR5 delta32 and HLA-B57 genotype [Abstract]. *4th IAS Conference on HIV Pathogenesis, Treatment, and Prevention. Sydney, Australia* July 22-25, 2007; MOPEA025.

21. Sheppard HW, Lang W, Ascher MS, Vittinghoff E,Winkelstein W. The characterization of non-progressors: Long-term HIV-1 infection with stable CD4+

T-cell levels. *AIDS.* 1993; **7**(9):1159-1166.

22. Ashton LJ, Carr A, Cunningham PH, Roggensack M, McLean K, Law M, et al. Predictors of progression in long-term nonprogressors. australian long-term nonprogressor study group. *AIDS Res Hum Retroviruses.* 1998; **14**(2):117-121.

23. Petrucci A, Dorrucci M, Alliegro MB, Pezzotti P, Rezza G, Sinicco A, et al. How many HIV-infected individuals may be defined as long-term nonprogressors? A report from the italian seroconversion study. italian seroconversion study group (ISS). *J Acquir Immune Defic Syndr Hum Retrovirol.* 1997; **14**(3):243-248.

24. Pantaleo G, Menzo S, Vaccarezza M, Graziosi C, Cohen OJ, Demarest JF, et al. Studies in subjects with long-term nonprogressive human immunodeficiency virus infection. *N Engl J Med.* 1995; **332**(4):209-216.

25. Mulder J, Resnick R, Saget B, Scheibel S, Herman S, Payne H, et al. A rapid and simple method for extracting human immunodeficiency virus type 1 RNA from plasma: Enhanced sensitivity. *J Clin Microbiol.* 1997; **35**(5):1278-1280.

26. Schacker T, Collier AC, Hughes J, Shea T,Corey L. Clinical and epidemiologic features of primary HIV infection. *Ann Intern Med.* 1996; **125**(4):257-264.

27. [Anonymous]. The prognostic value of cellular and serologic markers in infection with human immunodeficiency virus type 1. *Dis Markers.* 1990; **8**(3):164-165.

28. Jurriaans S, Van Gemen B, Weverling GJ, Van Strijp D, Nara P, Coutinho R, et al. The natural history of HIV-1 infection: Virus load and virus phenotype independent determinants of clinical course? *Virology.* 1994; **204**(1):223-233.

29. Gazzard B, Stein SE. :Factors Influencing the Natural History of HIV Infection. In: *Human Immunodeficiency Virus.* Richman D (editor) London, Atlanta, USA: International Medical Press; 2003. pp. 1-21.

30. [Anonymous]. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Recomm Rep.* 1992; **41**(RR-17):1-19.

31. Ratner L, Haseltine W, Patarca R, Livak KJ, Starcich B, Josephs SF, et al. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature.* 1985; **313**(6000):277-

284.

32. Freed EO.HIV-1 replication. *Somat Cell Mol Genet.* 2001; **26**(1-6):13-33.

33. Frankel AD, Young JA. HIV-1: Fifteen proteins and an RNA. *Annu Rev Biochem.* 1998; **67**:1-25.

34. Alonso A, Peterlin M. :HIV-1: control of gene expression by the viral regulatory proteins Tat and Rev. In: *HIV and the New Viruses.* Dalgleish AG, Weiss RA (eds). (editor) San Diego: Academic Press; 1999. pp. 1-28.

35. Moulard M, Decroly E. Maturation of HIV envelope glycoprotein precursors by cellular endoproteases. *Biochim Biophys Acta.* 2000; **1469**(3):121-132.

36. Berger EA, Murphy PM,Farber JM. Chemokine receptors as HIV-1 coreceptors: Roles in viral entry, tropism, and disease. *Annu Rev Immunol.* 1999; **17**:657-700.

37. Sattentau QJ, Moore JP. Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding. *J Exp Med.* 1991; **174**(2):407-415.

38. Cocchi F, DeVico AL, Garzino-Demo A, Arya SK, Gallo RC,Lusso P. Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells. *Science.* 1995; **270**(5243):1811-1815.

39. Oberlin E, Amara A, Bachelerie F, Bessia C, Virelizier JL, Arenzana-Seisdedos F, et al. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature.* 1996; **382**(6594):833-835.

40. Bleul CC, Farzan M, Choe H, Parolin C, Clark-Lewis I, Sodroski J, et al. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature.* 1996; **382**(6594):829-833.

41. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, et al. Identification of a major co-receptor for primary isolates of HIV-1. *Nature.* 1996; **381**(6584):661-666.

42. Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, et al. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature.* 1996; **381**(6584):667-673.

43. Doranz BJ, Rucker J, Yi Y, Smyth RJ, Samson M, Peiper SC, et al. A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell.* 1996; **85**(7):1149-1158.

44. Goldfarb DS.HIV-1 virology. simply marvelous nuclear transport. *Curr Biol.* 1995; **5**(6):570-573.

45. Haffar OK, Popov S, Dubrovsky L, Agostini I, Tang H, Pushkarsky T, et al. Two nuclear localization signals in the HIV-1 matrix protein regulate nuclear import of the HIV-1 pre-integration complex. *J Mol Biol.* 2000; **299**(2):359-368.

46. Farnet CM, Bushman FD. HIV cDNA integration: Molecular biology and inhibitor development. *AIDS.* 1996; **10 Suppl A**:S3-11.

47. Strebel K.Virus-host interactions: Role of HIV proteins vif, tat, and rev. *AIDS.* 2003; **17 Suppl 4**:S25-34.

48. Fenouillet E, Jones IM. The glycosylation of human immunodeficiency virus type 1 transmembrane glycoprotein (gp41) is important for the efficient intracellular transport of the envelope precursor gp160. *J Gen Virol.* 1995; **76 (Pt 6)**(Pt 6):1509-

1514.

49. 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):358-361.

50. Demirov DG, Orenstein JM,Freed EO. The late domain of human immunodeficiency virus type 1 p6 promotes virus release in a cell type-dependent manner. *J Virol.* 2002; **76**(1):105-117.

51. Debouck C, Gorniak JG, Strickler JE, Meek TD, Metcalf BW,Rosenberg M. Human immunodeficiency virus protease expressed in escherichia coli exhibits autoprocessing and specific maturation of the gag precursor. *Proc Natl Acad Sci U S A.* 1987; **84**(24):8903-8906.

52. Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: A collaborative re-analysis. collaborative group on AIDS incubation and HIV survival including the CASCADE EU concerted action. concerted action on SeroConversion to AIDS and death in europe. *Lancet.* 2000; **355**(9210):1131-1137.

53. Luo K, Law M, Kaldor JM, McDonald AM,Cooper DA. The role of initial AIDS-defining illness in survival following AIDS. *AIDS.* 1995; **9**(1):57-63.

54. Vella S, Giuliano M, Pezzotti P, Agresti MG, Tomino C, Floridia M, et al. Survival of zidovudine-treated patients with AIDS compared with that of contemporary untreated patients. italian zidovudine evaluation group. *JAMA.* 1992; **267**(9):1232-1236.

55. Lemp GF, Payne SF, Neal D, Temelso T,Rutherford GW. Survival trends for patients with AIDS. *JAMA.* 1990; **263**(3):402-406.

56. Cox PH, Martin MA, Styer CM,Beall GN. Outcomes of treatment with AZT of patients with AIDS and symptomatic HIV infection. *Nurse Pract.* 1990; **15**(5):36, 39-44.

57. Concorde coordinating committee. Concorde: MRC/ANRS randomised double-blind controlled trial of immediate and deferred zidovudine in symptom-free HIV infection. *Lancet.* 1994; **343**(8902):871-881.

58. Delta coordinating committee. Delta: A randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with

zidovudine alone in HIV-infected individuals. *Lancet.* 1996; **348**(9023):283-291.

59. Hammer SM, Katzenstein DA, Hughes MD, Gundacker H, Schooley RT, Haubrich RH, et al. A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. AIDS clinical trials group study 175 study team. *N Engl J Med.* 1996; **335**(15):1081-1090.

60. Eron JJ, Benoit SL, Jemsek J, MacArthur RD, Santana J, Quinn JB, et al. Treatment with lamivudine, zidovudine, or both in HIV-positive patients with 200 to 500 CD4+ cells per cubic millimeter. north american HIV working party. *N Engl J Med.* 1995; **333**(25):1662-1669.

61. Katlama C, Ingrand D, Loveday C, Clumeck N, Mallolas J, Staszewski S, et al. Safety and efficacy of lamivudine-zidovudine combination therapy in antiretroviral-naive patients. A randomized controlled comparison with zidovudine monotherapy. lamivudine european HIV working group. *JAMA.* 1996; **276**(2):118-125.

62. Schooley RT, Ramirez-Ronda C, Lange JM, Cooper DA, Lavelle J, Lefkowitz L, et al. Virologic and immunologic benefits of initial combination therapy with zidovudine and zalcitabine or didanosine compared with zidovudine monotherapy. wellcome resistance study collaborative group. *J Infect Dis.* 1996; **173**(6):1354-

1366.

63. Gatell JM, Leal M, Mallolas J, Vidal C, Pumarola T, Parra R, et al. A pilot case-control study of zidovudine compared with zidovudine plus didanosine in patients with advanced HIV-1 disease and no previous experience with antiretrovirals. *Antivir Ther.* 1996; **1**(2):105-112.

64. Collier AC, Coombs RW, Fischl MA, Skolnik PR, Northfelt D, Boutin P, et al. Combination therapy with zidovudine and didanosine compared with zidovudine alone in HIV-1 infection. *Ann Intern Med.* 1993; **119**(8):786-793.

65. Larder BA, Darby G,Richman DD. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. *Science.* 1989; **243**(4899):1731-1734.

66. Larder BA, Kemp SD. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science.* 1989; **246**(4934):1155-1158.

67. Larder BA, Kemp SD,Harrigan PR. Potential mechanism for sustained antiretroviral efficacy of AZT-3TC combination therapy. *Science.* 1995; **269**(5224):696-699.

68. D'Aquila RT, Johnson VA, Welles SL, Japour AJ, Kuritzkes DR, DeGruttola V, et al. Zidovudine resistance and HIV-1 disease progression during antiretroviral therapy. AIDS clinical trials group protocol 116B/117 team and the virology committee resistance working group. *Ann Intern Med.* 1995; **122**(6):401-408.

69. Shirasaka T, Yarchoan R, O'Brien MC, Husson RN, Anderson BD, Kojima E, et al. Changes in drug sensitivity of human immunodeficiency virus type 1 during therapy with azidothymidine, dideoxycytidine, and dideoxyinosine: An in vitro comparative study. *Proc Natl Acad Sci U S A.* 1993; **90**(2):562-566.

70. Shirasaka T, Kavlick MF, Ueno T, Gao WY, Kojima E, Alcaide ML, et al. Emergence of human immunodeficiency virus type 1 variants with resistance to multiple dideoxynucleosides in patients receiving therapy with dideoxynucleosides. *Proc Natl Acad Sci U S A.* 1995; **92**(6):2398-2402.

71. D'Aquila RT, Hughes MD, Johnson VA, Fischl MA, Sommadossi JP, Liou SH, et al. Nevirapine, zidovudine, and didanosine compared with zidovudine and didanosine in patients with HIV-1 infection. A randomized, double-blind, placebo-controlled trial. national institute of allergy and infectious diseases AIDS clinical trials group protocol 241 investigators. *Ann Intern Med.* 1996; **124**(12):1019-1030.

72. Hammer SM, Squires KE, Hughes MD, Grimes JM, Demeter LM, Currier JS, et al. 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. *N Engl J Med.* 1997; **337**(11):725-733.

73. Palella FJ,Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV outpatient study investigators. *N Engl J Med.* 1998; **338**(13):853-860.

74. Hogg RS, Yip B, Kully C, Craib KJ, O'Shaughnessy MV, Schechter MT, et al. Improved survival among HIV-infected patients after initiation of triple-drug antiretroviral regimens. *CMAJ.* 1999; **160**(5):659-665.

75. Mocroft A, Vella S, Benfield TL, Chiesi A, Miller V, Gargalianos P, et al. Changing patterns of mortality across europe in patients infected with HIV-1. EuroSIDA study group. *Lancet.* 1998; **352**(9142):1725-1730.

76. Detels R, Munoz A, McFarlane G, Kingsley LA, Margolick JB, Giorgi J, et al. Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. multicenter AIDS cohort study investigators. *JAMA.*

1998; **280**(17):1497-1503.

77. Walmsley S, Bernstein B, King M, Arribas J, Beall G, Ruane P, et al. Lopinavir-ritonavir versus nelfinavir for the initial treatment of HIV infection. *N Engl J Med.* 2002; **346**(26):2039-2046.

78. Hammer SM, Saag MS, Schechter M, Montaner JS, Schooley RT, Jacobsen DM, et al. Treatment for adult HIV infection: 2006 recommendations of the international AIDS society-USA panel. *JAMA.* 2006; **296**(7):827-843.

79. Lazzarin A, Clotet B, Cooper D, Reynes J, Arasteh K, Nelson M, et al. Efficacy of enfuvirtide in patients infected with drug-resistant HIV-1 in europe and australia. *N Engl J Med.* 2003; **348**(22):2186-2195.

80. Lalezari JP, Henry K, O'Hearn M, Montaner JS, Piliero PJ, Trottier B, et al. Enfuvirtide, an HIV-1 fusion inhibitor, for drug-resistant HIV infection in north and south america. *N Engl J Med.* 2003; **348**(22):2175-2185.

81. Low-Beer S, Yip B, O'Shaughnessy MV, Hogg RS,Montaner JS. Adherence to triple therapy and viral load response. *J Acquir Immune Defic Syndr.* 2000; **23**(4):360-361.

82. Bangsberg DR, Acosta EP, Gupta R, Guzman D, Riley ED, Harrigan PR, et al. Adherence-resistance relationships for protease and non-nucleoside reverse transcriptase inhibitors explained by virological fitness. *AIDS.* 2006; **20**(2):223-231.

83. Harrigan PR, Hogg RS, Dong WW, Yip B, Wynhoven B, Woodward J, et al. Predictors of HIV drug-resistance mutations in a large antiretroviral-naive cohort initiating triple antiretroviral therapy. *J Infect Dis.* 2005; **191**(3):339-347.

84. M. Markowitz, B. -Y Nguyen, E. Gotuzzo, F. Mendo, W. Ratanasuwan, C. Kovacs, et al. Rapid onset and durable antiretroviral effect of raltegravir (MK-0518), a novel HIV-1 integrase inhibitor, as part of combination ART in treatment HIV-1 infected patients: 48-week data [Abstract]. *4th IAS Conference on HIV Pathogenesis, Treatment, and Prevention. Sydney, Australia* July 22-25, 2007; TUAB104.

85. M. S. Saag, P. Ive, J. Heera, M. Tawadrous, E. DeJesus, N. Clumeck, et al. A multicenter, randomized, double-blind, comparative trial of a novel CCR5 antagonist, maraviroc versus efavirenz, both in combination with Combivir (zidovudine [ZDV]/lamivudine [3TC]), for the treatment of antiretroviral naive subjects infected with R5 HIV-1: week 48 results of the MERIT study [Abstract]. *4th IAS Conference on HIV Pathogenesis, Treatment, and Prevention. Sydney, Australiah* July 22-25, 2007; WESS104.

86. Strizki JM, Tremblay C, Xu S, Wojcik L, Wagner N, Gonsiorek W, et al. Discovery and characterization of vicriviroc (SCH 417690), a CCR5 antagonist with potent activity against human immunodeficiency virus type 1. *Antimicrob Agents Chemother.* 2005; **49**(12):4911-4919.

87. H. Mayer, Van der Ryst, E., M. Saag, B. Clotet, G. Fätkenheuer, N. Clumeck, et al. Safety and efficacy of maraviroc, a novel CCR5 antagonist, when used in combination with optimized background therapy for the treatment of antiretroviral-experiences subjects infected with dual/mixed-tropic HIV-1: 24-week results of a phase 2b exploratory trial.[Abstract]. *XVI International AIDS Conference, Toronto, Ontario, Canada* August 13-18, 2006; THLB0215.

CHAPTER 2

LITERATURE REVIEW:

HIV CO-RECEPTOR PHENOTYPING IN THE CLINICAL SETTING[1]

2.1 Introduction

HIV co-receptor phenotyping has recently become an area of great relevance, largely due to the introduction of co-receptor antagonists which should not be prescribed unless the co-receptor tropism of the patient is correctly determined. New recombinant assays for HIV co-receptor usage have been developed for use in the clinical setting and have also been essential to the assessment of the epidemiology and clinical correlates of HIV co-receptor usage in large observational cohort studies. In addition, the development and evaluation of genotype-based predictive methods, which use bioinformatic and machine learning approaches to predict the HIV co-receptor phenotype from the HIV envelope genotype in a rapid and cost-effective manner, are also emerging as an area of intense interest. Development and training of genotype-based predictors, however, require large clinically-derived datasets for which both the co-receptor phenotype and HIV envelope genotype are known.

This chapter will focus on the relevance of accurately determining the HIV co-receptor phenotype in the clinical setting, and briefly describe an overview of historic and current assays for determining the HIV co-receptor phenotype and their associations with disease progression. In addition, genotype-based predictors of these assays and their associations with HIV clinical outcomes will be discussed.

[1] A version of this chapter will be submitted for publication

2.2 Chemokine Receptors and Viral Entry

Binding of the CD4 receptor to the viral surface envelope protein, gp120, results in conformational changes in the envelope, however these changes alone are often insufficient to induce virus-cell fusion [1]. The requirement for a co-factor in conjunction with CD4 for HIV entry was demonstrated in studies showing that most human CD4-expressing cells, as well as non-human CD4-expressing cells are not permissive to HIV infection [2-9]. In late 1995, Lusso and Gallo indicated that CC-chemokines were involved in inhibiting HIV-1 replication [10]. Early in 1996, Berger discovered that the chemokine receptor CXCR4 (originally known as "fusin"), was a necessary co-factor required for HIV infection of CD4 expressing cells [11-13].Viral isolates able to use the CCR5 or CXCR4 co-receptors in addition to CD4 to gain entry into host cells are referred to as CCR5 or CXCR4-using, respectively. In addition, viral isolates able to use both co-receptors are referred to as "dual-tropic"; the dual-tropic phenotype could represent an evolutionary transitional state between solely CCR5-tropic and CXCR4-tropic virus [14].

The requirement of cellular co-receptors for viral entry was additionally demonstrated by the discovery that certain individuals possess a natural resistance to infection due to the presence of a rare 32 base-pair deletion in the CCR5 gene. This "CCR5Δ32 mutation" causes a frameshift mutation resulting in a premature stop codon, leading to the formation of a severely truncated protein which is not expressed on the cell surface [15,16]. Individuals homozygous for the CCR5Δ32 mutation are largely "resistant" to infection with CCR5-using strains of HIV [17,18], although others report

no correlation between CCR5 surface expression and co-receptor usage [19]. Individuals heterozygous for the CCR5Δ32 mutation are not protected against HIV infection, but do appear to have a decreased rate of disease progression once infected [17,18,20].

The levels or cell surface density of both HIV co-receptors as well as CD4 receptors appear to have a significant effect on viral fusion and infection [21]. Infection via the CCR5 co-receptor occurs more efficiently when cell surface concentrations of CD4 are increased, indicating that only low levels of CCR5 are necessary when cell surface levels of CD4 are high [21]. Conversely, at a lower CD4 concentration, larger amounts of CCR5 are necessary to achieve the same high efficiency of infection [21]. The levels of CD4 and co-receptor expression can vary from 2 to 5 fold between donors [22]. Although CXCR4 is generally expressed on a higher percentage of total lymphocytes compared with CCR5, the levels of CCR5 expression appear to be near 2.5 fold greater [22], suggesting that although there may be more cellular targets which express the CXCR4 co-receptor, infection via the CCR5 co-receptor may be more efficient.

2.3 History of Co-receptor Assays

The ability to induce the formation of multinucleated giant cells, or "syncytia", in peripheral blood mononuclear cells (PBMC) was discovered to be a stable transmissible phenotypic property exhibited by specific HIV strains in 1987 [23]. Indeed, the first *in vitro* assays developed to assess the HIV co-receptor phenotype was the SI ("syncytium inducing") assay. This assay was based on the ability of a viral isolate to induce syncytium formation in cultures of peripheral blood mononuclear cells [23-25]. Viral isolates able to form such syncytia were classified as "SI" (syncytium-inducing), whereas those unable to do so were classified as "NSI" (non-syncytium-inducing). Later it was noted that isolates capable of forming syncytium in PBMC culture, in general, displayed the ability to infect T-cell lines [25], while non-syncytium inducing isolates, in general, displayed a much greater capacity to infect monocyte-derived macrophages (MDM), leading to the nomenclature "T-tropic" and "M-tropic", respectively.

In order to reliably differentiate between SI or NSI HIV variants, a MT-2 cell line assay was developed to be a sensitive, specific and higher-throughput method [26,27]. The MT-2 cell assay isolates virus from HIV seropositive donors by cocultivation with stimulated PBMCs from seronegative donors [26-28]. Note that the ability to replicate in PBMCs is not exclusively linked to the SI phenotype in the MT-2 cell assay. Isolates must also be able to induce cytopathic changes and demonstrate increasing levels of HIV-1 antigen in cell culture in two consecutive samples [28]. The appearance of SI virus in the MT-2 cell assay in untreated individuals is followed by a steeper decline in

CD4 cells and a more rapid progression to AIDS in comparison to individuals with non-syncytium inducing virus (NSI) [24,29-34].

Although the MT-2 cell assay has been shown to be an accurate measure of HIV co-receptor usage, as well as a prognostic marker of disease progression, the length of time and labor required for the MT-2 assays limited its use as a functional, clinically relevant assay. A more recent innovation, therefore, was the development of more convenient and efficient phenotypic co-receptor usage assays, which classified viral isolates based on their ability to gain entry into CD4+ T-cells via either the CCR5 and/or CXCR4 co-receptors, leading to the nomenclature "CCR5-using" , "CXCR4-using", or "dual/mixed" tropism. Generally, CXCR4-using isolates are "SI" and CCR5-using isolates are "NSI", however, some disparity exists. One of the earliest phenotypic co-receptor usage assays used GHOST cells (derived from human osteosarcoma (HOS) cell lines) as an indicator cell line: these cells were engineered to express CCR5 and/or CXCR4, and carried an HIV-1 tat-inducible (HIV-2 LTR) green fluorescent protein (GFP) expression cassette [35,36]. However, the HOS cell lines expressed low levels of endogenous CXCR4 on their surface which can lead to false positive results for the CXCR4-using phenotype [37].

The most recently developed assays for determining the HIV co-receptor phenotype are recombinant co-receptor phenotype assays. These assays co-transfect all or part of a patient-derived HIV envelope, with a stock, envelope deficient virus engineered to express a reporter gene upon successful infection. The produced virions are

capable of only a single round of infection and are used to infect human glioma cell lines (U87, U373) engineered to express CD4 in addition to CXCR4 and/or CCR5. These cell lines have no endogenous CXCR4 or CCR5 expression and therefore have very low background infectivity levels [37]. This approach, however, is not without possible drawbacks. Interactions with the stock gag vector and the patient derived envelope may affect pseudotyping efficiency and variation in toxicity of different envelopes in the producer cells [37]. Also, the use of U87 cell lines engineered to express CD4 and either CXCR4 or CCR5 may not represent co-receptor densities *in vivo*.

There are currently two commercially available recombinant co-receptor phenotype assays. These are the Trofile Co-receptor assay by Monogram Biosciences [38], and the Tropism Recombinant Test assay by VIRalliance [39]. Another HIV co-receptor assay will also soon to be made available by Virco [40]. Infection of cell lines is detected using a reporter gene (β-galactosidase for the VIRalliance TRT and Luciferase for the Trofile and Virco assays). Results are then reported as R5, X4 or R5/X4 (dual/mixed), based on the dose-dependent decrease in reporter output after the addition of a strong co-receptor antagonist. These assays have been shown to be effective at detecting minority species of CXCR4-using virus at levels as low as 5-20% [41], although these results are likely achieved using idealized, highly CCR5-capable or CXCR4-capable strains. The Monogram Trofile assay was used exclusively as the assay for clinical trials of the new CCR5 antagonist, maraviroc [42-45]. However, the concordance between the Monogram Trofile assay and the TRT assay has been estimated at approximately 85% [46], and no results have yet been published on the repeatability of

these assays on identical, clinically derived samples, shedding some doubt on whether any of these assays can reliably be considered a ‘gold-standard’ in the clinical setting. Although all co-receptor assays discussed are well-validated, they are not necessarily comparable, and should prompt users of this data to be aware of the potential for bias and discordance between reported results.

2.4 Co-receptor Phenotype over the Natural History of HIV Infection and Theories on the Co-receptor Switch

Monocytotropic, non-syncytium-inducing isolates are detectable in over 95% of HIV infected individuals during the early asymptomatic phase, while isolates with syncytium-inducing capability emerge over the course of infection, and are detectable in approximately half of untreated HIV-infected individuals progressing to AIDS [24,25,34,47]. Seroconverter studies suggest that although transmission of SI strains may occur, these SI are likely outcompeted by their NSI counterparts shortly after the transmission event [48,49], suggesting that NSI variants have a selective advantage under conditions where immune function is not yet severely compromised [50]. Large scale population based studies also indicated that approximately only 18% of HIV variants isolated from the plasma of drug-naïve, chronically infected individuals have the ability to use the CXCR4 co-receptor, while virus from the remaining individuals use the CCR5 co-receptor exclusively [51-53]. This has also been found to be the case in recent seroconverters [54]. However, in the advanced stages of HIV infection, CXCR4-using variants may be detectable in up to 50% or more of patients [52].

Although the factors mediating the R5-to-X4 "phenotype switch" are incompletely understood, the consequences of this "switch" may be quite severe with respect to clinical prognosis. The emergence of SI (or, in other studies, CXCR4-using) virus is associated with a rapid CD4 cell count decline, accelerated disease progression, and reduced survival time in untreated individuals [24,31,33,55-57]. The mechanisms whereby X4 viruses contribute to accelerated disease progression and a poorer prognosis

in HIV-infected individuals are incompletely understood, however, evidence suggests that X4 viruses are able to directly infect and kill CD4+ T-cells and indirectly kill CD8+ T-cells via a mechanism mediated by the CXCR4 chemokine receptor [58]. In addition, X4 strains have been shown to block specific immune responses that are not blocked by R5 strains, thereby directly contributing to a decline in immune function [59]. Recent evidence suggests that the co-receptor switch is both a cause and effect of low CD4 count, due to an increase in the ratio of CXCR4 to CCR5 expressing T-cells as HIV infection progresses [60], however, it remains largely unresolved if X4 virus is a cause and/or emerges as a consequence of immune depletion [59-64].

The reconstitution of the immune system does not decrease the levels of CXCR4-using virus, and patients on successful antiretroviral therapy who have significant increases to their CD4 counts do continue to have detectable CXCR4-using virus [53]. These results suggest that the nadir (lowest recorded) CD4 count should be the one used when associating the presence or absence of CXCR4-using virus with CD4 count levels in therapy experienced individuals. In the context of artificial, drug-mediated selection pressure, the CCR5 antagonist maraviroc did select for pre-existing X4 variants which were previously undetectable at screening [42]. This reinforces the suggestion that even in samples with purely CCR5-tropic virus, CXCR4-capable virus may be circulating as minority species. Importantly, it indicates that even low levels of CXCR4-capable virus may be sufficient to render CCR5-antagonist based therapies ineffective.

The lack of CXCR4-using virus during the earlier stages of disease is likely a result of many contributing factors [65]. From a purely mathematical point of view, if the majority of the infected population is infected with CCR5 tropic virus, it would be expected that CCR5-using strains would be the most prevalent virus transmitted. Even if an individual is characterized as having a CXCR4-using virus, this virus will likely be present as a minority species, limiting the probability that transmitted virus will be CXCR4 tropic. There are also many biological factors involved as well, including preferential expression of CCR5 on Langerhans cells [66], and preferential trapping and inactivation of X4-virus by mucin and other innate antiviral proteins [67-70]. Macrophages, which have been shown to provide a more persistent viral reservoir in the face of a cytotoxic lymphocyte response [71], provide a greater pool of susceptible targets for R5 virus than X4 virus [65,71]. Dendritic cells have also been observed to show preferential binding for R5 virus over X4 virus, thereby increasing the opportunity for R5 virus transmission through the lymph nodes [72,73]. Recent studies have identified a novel, conserved gp120-neutralizing epitope that is cryptic in R5 variants, but is constitutively accessible to neutralizing antibodies in X4 (or R5X4) variants [65,74]. Any of these mechanisms alone have been postulated to provide only low levels of resistance to CXCR4-using virus, however, the synergistic operation of these and other imperfect barriers to CXCR4-tropic has been proposed as a very likely mechanism for the high prevalence of CCR5-tropic virus in all stages of HIV-infection [65]. Even in HIV-infected patients with CCR5-using virus only, CXCR4-tropic virus may be circulating as a rarely encountered minority species and/or be seeded in different lymphatic compartments such as in the gut-associated lymphatic tissue (GALT), which is

known to harbor cells with a preponderance of both CCR5 and CXCR4 co-receptors [75], and is an important early target of HIV replication and a site for severe CD4+ T-cell depletion [68,76].

2.5 Genotypic markers of co-receptor usage

HIV-1 co-receptor tropism is determined by the *env* gene [77-81], which encodes the gp120 (surface) and gp41 (transmembrane) envelope glycoproteins. The gp120 surface protein is organized into five hypervariable domains, designated V1-V5, interspersed with conserved regions [40,82]. The third hypervariable loop, or V3 loop of the gp120 surface protein, contains the most predictive genotypic determinants of SI or NSI tropism [83,84], although it has been demonstrated in molecular clones that the V1-V2 region may also be involved in determining SI capacity in MT-2 cells [85]. One of the most commonly observed quantitative genotypic changes associated with a phenotypic switch from NSI to SI viral strains is an increase in the positive charge of the V3 loop [83,86]. Specifically, a point mutation analysis indicated that positively charged residues (lysine or arginine) at either positions 306 or 320 in the proviral clone HXB2 (V3 positions 11 or 25) were causally involved in determining SI capacity [83]. A polymerase chain reaction (PCR) based assay was later demonstrated to be a sensitive and reliable method for quickly determining the HIV-1 SI phenotype based on the V3 genotype [87]. This method, termed the '11/25' rule, has been cited as a sensitive genotypic method of determining SI or NSI HIV-1. The '11/25' rule has high sensitivity and specificity (>90%) for determining the SI/NSI phenotype from clonal data [88], but its sensitivity for predicting the R5/X4 phenotype is reduced to <60% on clonal data and <40% on population-based, clinical data at the same specificity [87-90]. In addition, the '11/25' rule has also been shown to be an independent predictor of HIV disease progression during therapy [91], while the X4 phenotype (as determined by the Monogram Trofile assay) has not [92].

It has been demonstrated through an exchange between X4 and R5 V3 domains in envelope constructs that the principal region in gp120 that is generally both necessary and sufficient to mediate CXCR4-dependent infection of T-lymphocytes is the V3 region [93,94]. However, productive infection of primary macrophages via the CXCR4 co-receptor requires CXCR4-associated mutations in the V3 domain in combination with complex determinants in gp120 outside the V3 loop [95,96]. Different mutations or substitutions in the extracellular domains of CXCR4 or CCR5 appear to have very different effects in different HIV-1 strains suggesting that not all gp120 envelope proteins have the same requirements for binding and infection with their co-receptors [97]. These results indicate the difficulty of mapping multiple pathways of CXCR4 or CCR5 tropism based on the V3 loop and its context within the sequence of surrounding envelope protein, and also illustrates why methods such as the simple '11/25' rule do not work in all situations. More complex methods which use the V3 loop genotype to predict HIV co-receptor tropism have resulted in increased sensitivity over the '11/25' rule (at the same specificity). These include regression methods [98], a position specific scoring matrix (PSSM) [99], as well as machine learning methods, including neural networks (NN) [90], and support vector machines (SVM) [100,101]. These methods use 'training' data, or sequence information with known co-receptor affinity to produce weightings or associations based on the presence of certain amino acids at codons in the V3 loop. However, sequence information alone is often not enough for clinically derived samples. New methods for predicting CXCR4-usage improve their predictive ability by using other genotypic information, such as the presence or absence of insertions and deletions in the V3 loop as well as immunological status as measured by CD4 count of CD4%

[102]. In addition, the predicted V3-loop structure has been shown to improve the predictive ability of genotype-based predictors [103].

2.6 Relevance and Limitations of Co-receptor usage in the era of Co-receptor Antagonists

The anticipated licensing of CCR5 antagonists as treatments for HIV infection [43,45,104-107] have greatly increased the necessity for assays to accurately determine HIV co-receptor usage. Specifically, clinical trials with CCR5 antagonists have indicated that patients with detectable CXCR4-using virus are unlikely to show a significant decrease in viral load in response to CCR5 antagonists [44], thus necessitating the determination viral co-receptor usage prior to initiation of co-receptor antagonist-based regimens. However, the inability to accurately detect CXCR4-using virus present as a minority species presents challenges for both phenotype and genotype-based methods of determining co-receptor phenotype in patient-derived samples [46]. Genotype-based methods performed on population based samples, or 'bulk' sequence data, will often result in sequence ambiguities or 'mixtures' in the genotype. Often low level mixtures at levels below 10-25% [108-111] are not detected and therefore, CXCR4-using virus, which may often exist as a minority species, can be overlooked. Recombinant phenotyping methods are also subject to the limitation of detecting minority species in populational samples at levels between 10-20% [112].

Initially, there was concern that resistance to co-receptor inhibitors may involve a 'switch' to predominantly CXCR4-using virus, thus resulting in accelerated disease progression as observed with the appearance of CXCR4-using virus during the natural history of the disease [106]. Initially, this hypothesis was supported by some i*n vitro* evidence, where X4-variants were selected for during culture with a chemically modified

analog of the CCR5 analog RANTES [113]. I*n vivo* support includes observations indicating that individuals with naturally occurring CCR5Δ32 mutations [18,114-116] , and therefore reduced levels of CCR5 expression, are more likely to harbor CXCR4-using virus than those with CCR5 wild-type alleles [89,117,118], although a longitudinal analysis of co-receptor expression five years after seroconversion did not indicate an association between the emergence of CXCR4-capable virus and levels of co-receptor expression on target cells [119].

Despite the concern that CXCR4 virus could emerge *de novo* in response to the administration of CCR5-antagonists, recent evidence suggests that resistance to CCR5-antagonists occurs via the selection of virus that is often still capable of utilizing inhibitor-bound CCR5. The appearance of this virus involves the ordered accumulation of mutations in the viral envelope, both in the V3 and elsewhere in gp120 [120]. Further evidence for resistance to CCR5 antagonists to occur via continued usage of the CCR5 co-receptor includes recent studies which have described resistance to AD101 (a precursor of the CCR5 antagonist Vicriviroc) to be conferred by four amino acid substitutions in the V3 region, allowing the virus to continue to enter primary T-cells via CCR5 [121-123]. It has also been suggested that these mutations do not come at a cost to fitness, resulting in a virus that will most likely remain resistant to these agents indefinitely [124]; however, this study is based upon evidence obtained from the use of a single CCR5-antagonist only. Together, these studies suggest that when CXCR4-using virus does emerge, it does not appear to be the result of a viral 'switch' from a CCR5-tropic clone under the selection pressure of CCR5 antagonists. Rather, CXCR4-virus

most likely emerges as an outgrowth of a pretreatment CXCR4-using reservoir [125]; however, when therapy is discontinued, the virus sampled in the peripheral plasma appears to revert back to CCR5-using virus [125]. Whether the CXCR4-using virus which emerges as a result of exposure to CCR5-antagonists results in more advanced disease progression, is still unknown, and remains an important question in the context of CCR5-antagonist therapy.

2.7 Summary

The HIV envelope faces multiple evolutionary pressures *in vivo*. It must escape from neutralizing antibodies, as well as bind and infect cells with differing levels and types of target cell receptors; both between individuals during transmission and within individuals during disease progression. This high level of variability makes sensitive and repeatable assays for co-receptor usage based on the HIV envelope of patient derived samples very challenging. In addition, it is still unknown if free floating virions in the peripheral plasma represent the most ideal viral samples with which to determine tropism. Increased sensitivity may be obtained from sampling different cellular or physical compartments, such as macrophage associated virus, or T-cell associated virus in the GALT.

The introduction of CCR5 antagonists have resulted in increased demand for efficient, quick and accurate assays of HIV co-receptor usage. Previous assays based on MT-2 cells, although clinically useful for understanding the effect of syncytium-inducing virus on disease progression have been mostly replaced by recombinant assays, which are designed to determine if virus can enter cells expressing very specific types and levels of CD4 and CCR5 or CXCR4. However, low levels of CXCR4-capable minority species present in patient-derived samples, and the high rate of variability in the HIV envelope, have created challenges for making repeatable co-receptor phenotype assays. A greater understanding about the significance of the magnitude of reporter output used to indicate successful infection, such as the relative light units (RLU) in the Trofile assay or

the colorimetric measure of β-galactosidase in the TRT assay, may improve our interpretation of the results of these assays.

Genotype-based predictors could be used to complement the use of current co-receptor phenotype assays or, if their sensitivity can be significantly increased, replace them. Importantly, it is likely that response to CCR5-antagonist therapy could be directly predicted from the HIV envelope genotype (or co-receptor phenotype), and therefore using genotypic data to predict response to CCR5 antagonist-based therapies may be much more relevant than predicting the HIV co-receptor usage itself. However, whether the predictive power of using the V3-loop alone is sufficient or must be increased by using other clinically relevant determinants as well as genotypic determinants within gp120 and gp41 in combination with the V3 loop has yet to be completely determined, and is a major component of this thesis.

2.8 References

1. Sattentau QJ, Moore JP. Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding. *J Exp Med.* 1991; **174**(2):407-415.

2. Ashorn PA, Berger EA, Moss B. Human immunodeficiency virus envelope glycoprotein/CD4-mediated fusion of nonprimate cells with human cells. *J Virol.* 1990; **64**(5):2149-2156.

3. Broder CC, Berger EA. Fusogenic selectivity of the envelope glycoprotein is a major determinant of human immunodeficiency virus type 1 tropism for CD4+ T-cell lines vs. primary macrophages. *Proc Natl Acad Sci U S A.* 1995; **92**(19):9004-9008.

4. Broder CC, Dimitrov DS, Blumenthal R, Berger EA. The block to HIV-1 envelope glycoprotein-mediated membrane fusion in animal cells expressing human CD4 can be overcome by a human cell component(s). *Virology.* 1993; **193**(1):483-491.

5. Chesebro B, Buller R, Portis J, Wehrly K. Failure of human immunodeficiency virus entry and infection in CD4-positive human brain and skin cells. *J Virol.* 1990; **64**(1):215-221.

6. Clapham PR, Blanc D, Weiss RA. Specific cell surface requirements for the infection of CD4-positive cells by human immunodeficiency virus types 1 and 2 and by simian immunodeficiency virus. *Virology.* 1991; **181**(2):703-715.

7. Dragic T, Charneau P, Clavel F, Alizon M. Complementation of murine cells for human immunodeficiency virus envelope/CD4-mediated fusion in human/murine heterokaryons. *J Virol.* 1992; **66**(8):4794-4802.

8. Dragic T, Picard L, Alizon M. Proteinase-resistant factors in human erythrocyte membranes mediate CD4-dependent fusion with cells expressing human immunodeficiency virus type 1 envelope glycoproteins. *J Virol.* 1995; **69**(2):1013-1018.

9. Harrington RD, Geballe AP. Cofactor requirement for human immunodeficiency virus type 1 entry into a CD4-expressing human cell line. *J Virol.* 1993; **67**(10):5939-5947.

10. Cocchi F, DeVico AL, Garzino-Demo A, Arya SK, Gallo RC, Lusso P. Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells. *Science.* 1995; **270**(5243):1811-1815.

11. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science.* 1996; **272**(5263):872-877.

12. Bleul CC, Farzan M, Choe H, Parolin C, Clark-Lewis I, Sodroski J, et al. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature.* 1996; **382**(6594):829-833.

13. Oberlin E, Amara A, Bachelerie F, Bessia C, Virelizier JL, Arenzana-Seisdedos F, et al. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature.* 1996; **382**(6594):833-835.

14. Collman R, Balliet JW, Gregory SA, Friedman H, Kolson DL, Nathanson N, et al. An infectious molecular clone of an unusual macrophage-tropic and highly cytopathic strain of human immunodeficiency virus type 1. *J Virol.* 1992; **66**(12):7517-7521.

15. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, et al. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature.* 1996; **382**(6593):722-725.

16. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell.* 1996; **86**(3):367-377.

17. Huang Y, Paxton WA, Wolinsky SM, Neumann AU, Zhang L, He T, et al. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. *Nat Med.* 1996; **2**(11):1240-1243.

18. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, Allikmets R, et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. hemophilia growth and development study, multicenter AIDS cohort study, multicenter hemophilia cohort study, san francisco city cohort, ALIVE study. *Science.* 1996; **273**(5283):1856-1862.

19. de Roda Husman AM, Koot M, Cornelissen M, Keet IP, Brouwer M, Broersen SM, et al. Association between CCR5 genotype and the clinical course of HIV-1 infection. *Ann Intern Med.* 1997; **127**(10):882-890.

20. Stewart GJ, Ashton LJ, Biti RA, Ffrench RA, Bennetts BH, Newcombe NR, et al. Increased frequency of CCR-5 delta 32 heterozygotes among long-term non-progressors with HIV-1 infection. the Australian long-term non-progressor study

group. *AIDS.* 1997; **11**(15):1833-1838.

21. Platt EJ, Wehrly K, Kuhmann SE, Chesebro B, Kabat D. Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1. *J Virol.* 1998; **72**(4):2855-2864.

22. Lee B, Sharron M, Montaner LJ, Weissman D, Doms RW. Quantification of CD4, CCR5, and CXCR4 levels on lymphocyte subsets, dendritic cells, and differentially conditioned monocyte-derived macrophages. *Proc Natl Acad Sci U S A.* 1999; **96**(9):5215-5220.

23. Tersmette M, de Goede RE, Al BJ, Winkel IN, Gruters RA, Cuypers HT, et al. Differential syncytium-inducing capacity of human immunodeficiency virus isolates: Frequent detection of syncytium-inducing isolates in patients with acquired immunodeficiency syndrome (AIDS) and AIDS-related complex. *J Virol.* 1988; **62**(6):2026-2032.

24. Tersmette M, Gruters RA, de Wolf F, de Goede RE, Lange JM, Schellekens PT, et al. Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency syndrome: Studies on sequential HIV isolates. *J Virol.* 1989; **63**(5):2118-2125.

25. Schuitemaker H, Kootstra NA, de Goede RE, de Wolf F, Miedema F, Tersmette M. Monocytotropic human immunodeficiency virus type 1 (HIV-1) variants detectable in all stages of HIV-1 infection lack T-cell line tropism and syncytium-inducing ability in primary T-cell culture. *J Virol.* 1991; **65**(1):356-363.

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

27. Boucher CA, Lange JM, Miedema FF, Weverling GJ, Koot M, Mulder JW, et al. HIV-1 biological phenotype and the development of zidovudine resistance in relation to disease progression in asymptomatic individuals during treatment. *AIDS.* 1992; **6**(11):1259-1264.

28. Karlsson A, Parsmyr K, Sandstrom E, Fenyo EM, Albert J. MT-2 cell tropism as prognostic marker for disease progression in human immunodeficiency virus type 1 infection. *J Clin Microbiol.* 1994; **32**(2):364-370.

29. Tersmette M, Gruters RA, de Wolf F, de Goede RE, Lange JM, Schellekens PT, et al. Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency syndrome: Studies on sequential

HIV isolates. *J Virol.* 1989; **63**(5):2118-2125.

30. Tersmette M, Miedema F. Interactions between HIV and the host immune system in the pathogenesis of AIDS. *AIDS.* 1990; **4 Suppl 1**:S57-66.

31. Richman DD, Bozzette SA. The impact of the syncytium-inducing phenotype of human immunodeficiency virus on disease progression. *J Infect Dis.* 1994; **169**(5):968-974.

32. Maas JJ, Gange SJ, Schuitemaker H, Coutinho RA, van Leeuwen R, Margolick JB. Strong association between failure of T cell homeostasis and the syncytium-inducing phenotype among HIV-1-infected men in the amsterdam cohort study. *AIDS.* 2000; **14**(9):1155-1161.

33. Koot M, Keet IP, Vos AH, de Goede RE, Roos MT, Coutinho RA, et al. Prognostic value of HIV-1 syncytium-inducing phenotype for rate of CD4+ cell depletion and progression to AIDS. *Ann Intern Med.* 1993; **118**(9):681-688.

34. Schuitemaker H, Koot M, Kootstra NA, Dercksen MW, de Goede RE, van Steenwijk RP, et al. Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: Progression of disease is associated with a

shift from monocytotropic to T-cell-tropic virus population. *J Virol.* 1992; **66**(3):1354-1360.

35. Berkowitz RD, Alexander S, Bare C, Linquist-Stepps V, Bogan M, Moreno ME, et al. CCR5- and CXCR4-utilizing strains of human immunodeficiency virus type 1 exhibit differential tropism and pathogenesis in vivo. *J Virol.* 1998; **72**(12):10108-10117.

36. Trkola A, Ketas T, Kewalramani VN, Endorf F, Binley JM, Katinger H, et al. Neutralization sensitivity of human immunodeficiency virus type 1 primary isolates to antibodies and CD4-based reagents is independent of coreceptor usage. *J Virol.* 1998; **72**(3):1876-1885.

37. Coakley E, Petropoulos CJ, Whitcomb JM. Assessing chemokine co-receptor usage in HIV. *Curr Opin Infect Dis.* 2005; **18**(1):9-15.

38. Whitcomb JM, Huang W, Fransen S, Limoli K, Toma J, Wrin T, et al. Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism. *Antimicrob Agents Chemother.* 2007; **51**(2):566-575.

39. Trouplin V, Salvatori F, Cappello F, Obry V, Brelot A, Heveker N, et al. Determination of coreceptor usage of human immunodeficiency virus type 1 from patient plasma samples by using a recombinant phenotypic assay. *J Virol.* 2001; **75**(1):251-259.

40. Ping LH, Nelson JA, Hoffman IF, Schock J, Lamers SL, Goodman M, et al. Characterization of V3 sequence heterogeneity in subtype C human immunodeficiency virus type 1 isolates from Malawi: Underrepresentation of X4 variants. *J Virol.* 1999; **73**(8):6271-6281.

41. Whitcomb JM, Huang W, Fransen S, Limoli K, Toma J, Wrin T, et al. Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism. *Antimicrob Agents Chemother.* 2007; **51**(2):566-575.

42. M. Lewis, P. Simpson, S. Francen, W. Huang, J. Whitcomb and M. Mosley. CXCR4-using virus detected in patients receiving maraviroc in the Phase III studies MOTIVATE 1 and 2 originates from a pre-existing minority of CXCR4-using virus [Abstract]. *16th International HIV Drug Resistance Workshop, Barbados, West Indies* June 12-16, 2007; 56.

43. Lalezari J, Goodrich J, DeJesus E, Lampiris H, Gulick R, Saag M, et al. Efficacy and Safety of Maraviroc plus Optimized Background Therapy in Viremic ART-experienced Patients Infected with CCR5-tropic HIV-1: 24-Week Results of a Phase 2b/3 Study in the US and Canada [Abstract]. *14th Conference on Retroviruses and Opportunistic Infections, Los Angeles, California, USA* February 25-28, 2007; 104bLB.

44. Mayer H, Van der Ryst E, Saag M, Clotet B, Fätkenheuer G, Clumeck N, et al. Safety and efficacy of maraviroc, a novel CCR5 antagonist, when used in combination with optimized background therapy for the treatment of antiretroviral-experiences subjects infected with dual/mixed-tropic HIV-1: 24-week results of a phase 2b exploratory trial.[Abstract]. *XVI International AIDS Conference, Toronto, Ontario, Canada* August 13-18, 2006; THLB0215.

45. Nelson M, Fätkenheuer G, Konourina I, Lazzarin A, Clumeck N, Horban A, et al. Efficacy and Safety of Maraviroc plus Optimized Background Therapy in Viremic, ART-experienced Patients Infected with CCR5-tropic HIV-1 in Europe, Australia, and North America: 24-Week Results [Abstract]. *14th Conference on Retroviruses and Opportunistic Infections, Los Angeles, California, USA* February 25-28, 2007; 104aLB.

46. Skrabal K, Low AJ, Dong W, Sing T, Cheung PK, Mammano F, et al. Determining human immunodeficiency virus coreceptor use in a clinical setting: Degree of correlation between two phenotypic assays and a bioinformatic model. *J Clin Microbiol.* 2007; **45**(2):279-284.

47. Roos MT, Lange JM, de Goede RE, Coutinho RA, Schellekens PT, Miedema F, et al. Viral phenotype and immune response in primary human immunodeficiency virus type 1 infection. *J Infect Dis.* 1992; **165**(3):427-432.

48. Groenink M, Fouchier RA, de Goede RE, de Wolf F, Gruters RA, Cuypers HT, et al. Phenotypic heterogeneity in a panel of infectious molecular human immunodeficiency virus type 1 clones derived from a single individual. *J Virol.* 1991; **65**(4):1968-1975.

49. Baur A, Schwarz N, Ellinger S, Korn K, Harrer T, Mang K, et al. Continuous clearance of HIV in a vertically infected child. *Lancet.* 1989; **2**(8670):1045.

50. Katz JD, Nishanian P, Mitsuyasu R, Bonavida B. Antibody-dependent cellular cytotoxicity (ADCC)-mediated destruction of human immunodeficiency virus (HIV)-coated CD4+ T lymphocytes by acquired immunodeficiency syndrome (AIDS) effector cells. *J Clin Immunol.* 1988; **8**(6):453-458.

51. Brumme ZL, Goodrich J, Mayer HB, Brumme CJ, Henrick BM, Wynhoven B, et al. Molecular and clinical epidemiology of CXCR4-using HIV-1 in a large population of antiretroviral-naive individuals. *J Infect Dis.* 2005; **192**(3):466-474.

52. Moyle GJ, Wildfire A, Mandalia S, Mayer H, Goodrich J, Whitcomb J, et al. Epidemiology and predictive factors for chemokine receptor use in HIV-1 infection. *J Infect Dis.* 2005; **191**(6):866-872.

53. Hunt PW, Harrigan PR, Huang W, Bates M, Williamson DW, McCune JM, et al. Prevalence of CXCR4 tropism among antiretroviral-treated HIV-1-infected patients with detectable viremia. *J Infect Dis.* 2006; **194**(7):926-930.

54. de Mendoza C, Rodriguez C, Garcia F, Eiros JM, Ruiz L, Caballero E, et al. Prevalence of X4 tropic viruses in patients recently infected with HIV-1 and lack of association with transmission of drug resistance. *J Antimicrob Chemother.* 2007; **59**(4):698-704.

55. Zhu T, Mo H, Wang N, Nam DS, Cao Y, Koup RA, et al. Genotypic and phenotypic characterization of HIV-1 patients with primary infection. *Science.* 1993; **261**(5125):1179-1181.

56. Blaak H, van't Wout AB, Brouwer M, Hooibrink B, Hovenkamp E, Schuitemaker H. In vivo HIV-1 infection of CD45RA(+)CD4(+) T cells is established primarily by syncytium-inducing variants and correlates with the rate of CD4(+) T cell decline. *Proc Natl Acad Sci U S A.* 2000; **97**(3):1269-1274.

57. M. Goetz, R. Leduc, J. Kostman, A. Labriola, Y. Lie, J. Weidler, et al. Relationship between HIV co-receptor tropism and disease progression in persons with untreated chronic HIV infection [Abstract]. *4th IAS Conference on HIV Pathogenesis, Treatment, and Prevention. Sydney, Australia* July 22-25, 2007; TUPEB092.

58. Herbein G, Mahlknecht U, Batliwalla F, Gregersen P, Pappas T, Butler J, et al. Apoptosis of CD8+ T cells is mediated by macrophages through interaction of HIV gp120 with chemokine receptor CXCR4. *Nature.* 1998; **395**(6698):189-194.

59. Glushakova S, Grivel JC, Fitzgerald W, Sylwester A, Zimmerberg J,Margolis LB. Evidence for the HIV-1 phenotype switch as a causal factor in acquired immunodeficiency. *Nat Med.* 1998; **4**(3):346-349.

60. Ribeiro RM, Hazenberg MD, Perelson AS,Davenport MP. Naive and memory cell turnover as drivers of CCR5-to-CXCR4 tropism switch in human immunodeficiency virus type 1: Implications for therapy. *J Virol.* 2006; **80**(2):802-809.

61. 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. *Proc Natl Acad Sci U S A.* 1999; **96**(2):663-668.

62. Kreisberg JF, Kwa D, Schramm B, Trautner V, Connor R, Schuitemaker H, et al. Cytopathicity of human immunodeficiency virus type 1 primary isolates depends on coreceptor usage and not patient disease status. *J Virol.* 2001; **75**(18):8842-8847.

63. Harouse JM, Buckner C, Gettie A, Fuller R, Bohm R, Blanchard J, et al. CD8+ T cell-mediated CXC chemokine receptor 4-simian/human immunodeficiency virus suppression in dually infected rhesus macaques. *Proc Natl Acad Sci U S A.* 2003; **100**(19):10977-10982.

64. Davenport MP, Zaunders JJ, Hazenberg MD, Schuitemaker H,van Rij RP. Cell turnover and cell tropism in HIV-1 infection. *Trends Microbiol.* 2002; **10**(6):275-278.

65. Margolis L, Shattock R. Selective transmission of CCR5-utilizing HIV-1: The 'gatekeeper' problem resolved? *Nat Rev Microbiol.* 2006; **4**(4):312-317.

66. Patterson BK, Landay A, Siegel JN, Flener Z, Pessis D, Chaviano A, et al. Susceptibility to human immunodeficiency virus-1 infection of human foreskin and cervical tissue grown in explant culture. *Am J Pathol.* 2002; **161**(3):867-873.

67. Bergey EJ, Cho MI, Blumberg BM, Hammarskjold ML, Rekosh D, Epstein LG, et al. Interaction of HIV-1 and human salivary mucins. *J Acquir Immune Defic Syndr.* 1994; **7**(10):995-1002.

68. Moulard M, Lortat-Jacob H, Mondor I, Roca G, Wyatt R, Sodroski J, et al. Selective interactions of polyanions with basic surfaces on human immunodeficiency virus type 1 gp120. *J Virol.* 2000; **74**(4):1948-1960.

69. Vives RR, Imberty A, Sattentau QJ,Lortat-Jacob H. Heparan sulfate targets the HIV-1 envelope glycoprotein gp120 coreceptor binding site. *J Biol Chem.* 2005; **280**(22):21353-21357.

70. Quinones-Mateu ME, Lederman MM, Feng Z, Chakraborty B, Weber J, Rangel HR, et al. Human epithelial beta-defensins 2 and 3 inhibit HIV-1 replication. *AIDS.* 2003; **17**(16):F39-48.

71. Schutten M, van Baalen CA, Guillon C, Huisman RC, Boers PH, Sintnicolaas K, et al. Macrophage tropism of human immunodeficiency virus type 1 facilitates in vivo escape from cytotoxic T-lymphocyte pressure. *J Virol.* 2001; **75**(6):2706-2709.

72. Granelli-Piperno A, Delgado E, Finkel V, Paxton W,Steinman RM. Immature dendritic cells selectively replicate macrophagetropic (M-tropic) human immunodeficiency virus type 1, while mature cells efficiently transmit both M- and T-tropic virus to T cells. *J Virol.* 1998; **72**(4):2733-2737.

73. Turville SG, Santos JJ, Frank I, Cameron PU, Wilkinson J, Miranda-Saksena M, et al. Immunodeficiency virus uptake, turnover, and 2-phase transfer in human dendritic cells. *Blood.* 2004; **103**(6):2170-2179.

74. Lusso P, Earl PL, Sironi F, Santoro F, Ripamonti C, Scarlatti G, et al. Cryptic nature of a conserved, CD4-inducible V3 loop neutralization epitope in the native envelope glycoprotein oligomer of CCR5-restricted, but not CXCR4-using, primary human immunodeficiency virus type 1 strains. *J Virol.* 2005; **79**(11):6957-6968.

75. Mowat AM, Viney JL. The anatomical basis of intestinal immunity. *Immunol Rev.* 1997; **156**:145-166.

76. Schneider T, Jahn HU, Schmidt W, Riecken EO, Zeitz M,Ullrich R. Loss of CD4 T lymphocytes in patients infected with human immunodeficiency virus type 1 is more pronounced in the duodenal mucosa than in the peripheral blood. berlin Diarrhea/Wasting syndrome study group. *Gut.* 1995; **37**(4):524-529.

77. Cheng-Mayer C, Quiroga M, Tung JW, Dina D,Levy JA. Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation. *J Virol.* 1990; **64**(9):4390-4398.

78. de Jong JJ, Goudsmit J, Keulen W, Klaver B, Krone W, Tersmette M, et al. Human immunodeficiency virus type 1 clones chimeric for the envelope V3 domain differ in syncytium formation and replication capacity. *J Virol.* 1992; **66**(2):757-765.

79. O'Brien WA, Koyanagi Y, Namazie A, Zhao JQ, Diagne A, Idler K, et al. HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain. *Nature.* 1990; **348**(6296):69-73.

80. 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):167-169.

81. Westervelt P, Trowbridge DB, Epstein LG, Blumberg BM, Li Y, Hahn BH, et al. Macrophage tropism determinants of human immunodeficiency virus type 1 in vivo. *J Virol.* 1992; **66**(4):2577-2582.

82. Lamers SL, Sleasman JW, She JX, Barrie KA, Pomeroy SM, Barrett DJ, et al. Persistence of multiple maternal genotypes of human immunodeficiency virus type I in infants infected by vertical transmission. *J Clin Invest.* 1994; **93**(1):380-390.

83. 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. *J Virol.* 1992; **66**(11):6777-6780.

84. Freed EO, Myers DJ,Risser R. Identification of the principal neutralizing determinant of human immunodeficiency virus type 1 as a fusion domain. *J Virol.* 1991; **65**(1):190-194.

85. Todd BJ, Kedar P,Pope JH. Syncytium induction in primary CD4+ T-cell lines from normal donors by human immunodeficiency virus type 1 isolates with non-syncytium-inducing genotype and phenotype in MT-2 cells. *J Virol.* 1995; **69**(11):7099-7105.

86. Moore JP, McCutchan FE, Poon SW, Mascola J, Liu J, Cao Y, et al. Exploration of antigenic variation in gp120 from clades A through F of human immunodeficiency virus type 1 by using monoclonal antibodies. *J Virol.* 1994; **68**(12):8350-8364.

87. Fouchier RA, Brouwer M, Broersen SM,Schuitemaker H. Simple determination of human immunodeficiency virus type 1 syncytium-inducing V3 genotype by PCR. *J Clin Microbiol.* 1995; **33**(4):906-911.

88. Jensen MA, van 't Wout AB. Predicting HIV-1 coreceptor usage with sequence analysis. *AIDS Rev.* 2003; **5**(2):104-112.

89. Brumme ZL, Goodrich J, Mayer HB, Brumme CJ, Henrick BM, Wynhoven B, et al. Molecular and clinical epidemiology of CXCR4-using HIV-1 in a large population of antiretroviral-naive individuals. *J Infect Dis.* 2005; **192**(3):466-474.

90. 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.

91. Brumme ZL, Dong WW, Yip B, Wynhoven B, Hoffman NG, Swanstrom R, et al. Clinical and immunological impact of HIV envelope V3 sequence variation after

starting initial triple antiretroviral therapy. *AIDS.* 2004; **18**(4):F1-9.

92. Edgar RC.MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; **32**(5):1792-1797.

93. 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):167-169.

94. Chesebro B, Wehrly K, Nishio J,Perryman S. Mapping of independent V3 envelope determinants of human immunodeficiency virus type 1 macrophage tropism and syncytium formation in lymphocytes. *J Virol.* 1996; **70**(12):9055-9059.

95. Lee MK, Heaton J,Cho MW. Identification of determinants of interaction between CXCR4 and gp120 of a dual-tropic HIV-1DH12 isolate. *Virology.* 1999; **257**(2):290-296.

96. Smyth RJ, Yi Y, Singh A,Collman RG. Determinants of entry cofactor utilization and tropism in a dualtropic human immunodeficiency virus type 1 primary isolate. *J Virol.* 1998; **72**(5):4478-4484.

97. Labrosse B, Treboute C, Brelot A,Alizon M. Cooperation of the V1/V2 and V3 domains of human immunodeficiency virus type 1 gp120 for interaction with the CXCR4 receptor. *J Virol.* 2001; **75**(12):5457-5464.

98. 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):2937-2939.

99. Jensen MA, Li FS, van 't Wout AB, Nickle DC, Shriner D, He HX, et al. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences. *J Virol.* 2003; **77**(24):13376-13388.

100. Pillai S, Good B, Richman D,Corbeil J. A new perspective on V3 phenotype prediction. *AIDS Res Hum Retroviruses.* 2003; **19**(2):145-149.

101. T. Sing, T. Lengauer, N. Beerenwinkel, R. Kaiser, M. Daumer, A. Low, et al. Immunological and virological parameters improve genotypic prediction of HIV-1 coreceptor usage [Abstract]. *1st Int. Workshop on Targeting HIV Entry, Bethesda, Maryland, USA*. December 2-3, 2005; Abstract 30.

102. T. Sing, R. Harrigan, N. Beerenwinkel, M. Daumer, R. Kaiser, A. Low, et al. Immunologic markers improve genotypic prediction of HIV-1 co-receptor usage [Abstract]. *Int. Workshop on Targeting HIV Entry* 2005; **1:**.

103. Sander O, Sing T, Sommer I, Low AJ, Cheung PK, Harrigan PR, et al. Structural descriptors of gp120 V3 loop for the prediction of HIV-1 coreceptor usage. *PLoS Comput Biol.* 2007; **3**(3):e58.

104. Maeda K, Nakata H, Ogata H, Koh Y, Miyakawa T,Mitsuya H. The current status of, and challenges in, the development of CCR5 inhibitors as therapeutics for HIV-1 infection. *Curr Opin Pharmacol.* 2004; **4**(5):447-452.

105. Seibert C, Sakmar TP. Small-molecule antagonists of CCR5 and CXCR4: A promising new class of anti-HIV-1 drugs. *Curr Pharm Des.* 2004; **10**(17):2041-2062.

106. Westby M, van der Ryst E. CCR5 antagonists: Host-targeted antivirals for the treatment of HIV infection. *Antivir Chem Chemother.* 2005; **16**(6):339-354.

107. Strizki JM, Tremblay C, Xu S, Wojcik L, Wagner N, Gonsiorek W, et al. Discovery and characterization of vicriviroc (SCH 417690), a CCR5 antagonist with potent

activity against human immunodeficiency virus type 1. *Antimicrob Agents Chemother.* 2005; **49**(12):4911-4919.

108. Larder BA, Kohli A, Kellam P, Kemp SD, Kronick M,Henfrey RD. Quantitative detection of HIV-1 drug resistance mutations by automated DNA sequencing. *Nature.* 1993; **365**(6447):671-673.

109. Leitner T, Halapi E, Scarlatti G, Rossi P, Albert J, Fenyo EM, et al. Analysis of heterogeneous viral populations by direct DNA sequencing. *BioTechniques.* 1993; **15**(1):120-127.

110. 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. *J Clin Microbiol.* 1999; **37**(7):2291-2296.

111. Van Laethem K, Van Vaerenbergh K, Schmit JC, Sprecher S, Hermans P, De Vroey V, et al. Phenotypic assays and sequencing are less sensitive than point mutation assays for detection of resistance in mixed HIV-1 genotypic populations. *J Acquir Immune Defic Syndr.* 1999; **22**(2):107-118.

112. J. L. Labernardiere, S. Lebel-Binay, J. L. Faudo, A. Holguin, V. Soriano and A. Cheret. Tropism determination and performance of Phenoscript™ HIV-1 entry inhibitor assay [Abstract]. *Antiviral Therapy* 2005; **9:**S141.

113. Mosier DE, Picchio GR, Gulizia RJ, Sabbe R, Poignard P, Picard L, et al. Highly potent RANTES analogues either prevent CCR5-using human immunodeficiency virus type 1 infection in vivo or rapidly select for CXCR4-using variants. *J Virol.* 1999; **73**(5):3544-3550.

114. Eugen-Olsen J, Iversen AK, Garred P, Koppelhus U, Pedersen C, Benfield TL, et al. Heterozygosity for a deletion in the CKR-5 gene leads to prolonged AIDS-free survival and slower CD4 T-cell decline in a cohort of HIV-seropositive individuals. *AIDS.* 1997; **11**(3):305-310.

115. Wu L, Paxton WA, Kassam N, Ruffing N, Rottman JB, Sullivan N, et al. CCR5 levels and expression pattern correlate with infectability by macrophage-tropic HIV-1, in vitro. *J Exp Med.* 1997; **185**(9):1681-1691.

116. Benkirane M, Jin DY, Chun RF, Koup RA,Jeang KT. Mechanism of transdominant inhibition of CCR5-mediated HIV-1 infection by ccr5delta32. *J Biol Chem.* 1997; **272**(49):30603-30606.

117. D'Aquila RT, Sutton L, Savara A, Hughes MD,Johnson VA. CCR5/delta(ccr5) heterozygosity: A selective pressure for the syncytium-inducing human immunodeficiency virus type 1 phenotype. NIAID AIDS clinical trials group protocol 241 virology team. *J Infect Dis.* 1998; **177**(6):1549-1553.

118. Zhang L, He T, Huang Y, Chen Z, Guo Y, Wu S, et al. Chemokine coreceptor usage by diverse primary isolates of human immunodeficiency virus type 1. *J Virol.* 1998; **72**(11):9307-9312.

119. van Rij RP, Hazenberg MD, van Benthem BH, Otto SA, Prins M, Miedema F, et al. Early viral load and CD4+ T cell count, but not percentage of CCR5+ or CXCR4+ CD4+ T cells, are associated with R5-to-X4 HIV type 1 virus evolution. *AIDS Res Hum Retroviruses.* 2003; **19**(5):389-398.

120. Westby M, Smith-Burchnell C, Mori J, Lewis M, Mosley M, Stockdale M, et al. Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry. *J Virol.* 2007; **81**(5):2359-2371.

121. Kuhmann SE, Pugach P, Kunstman KJ, Taylor J, Stanfield RL, Snyder A, et al. Genetic and phenotypic analyses of human immunodeficiency virus type 1 escape

from a small-molecule CCR5 inhibitor. *J Virol.* 2004; **78**(6):2790-2807.

122. Trkola A, Kuhmann SE, Strizki JM, Maxwell E, Ketas T, Morgan T, et al. HIV-1 escape from a small molecule, CCR5-specific entry inhibitor does not involve CXCR4 use. *Proc Natl Acad Sci U S A.* 2002; **99**(1):395-400.

123. Pugach P, Marozsan AJ, Ketas TJ, Landes EL, Moore JP,Kuhmann SE. HIV-1 clones resistant to a small molecule CCR5 inhibitor use the inhibitor-bound form of CCR5 for entry. *Virology.* 2007; **361**(1):212-228.

124. Anastassopoulou CG, Marozsan AJ, Matet A, Snyder AD, Arts EJ, Kuhmann SE, et al. Escape of HIV-1 from a small molecule CCR5 inhibitor is not associated with a fitness loss. *PLoS Pathog.* 2007; **3**(6):e79.

125. Westby M, Lewis M, Whitcomb J, Youle M, Pozniak AL, James IT, et al. 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. *J Virol.* 2006; **80**(10):4909-4920.

CHAPTER 3

CURRENT V3 GENOTYPING ALGORITHMS ARE INADEQUATE FOR PREDICTING X4 CO-RECEPTOR USAGE IN CLINICAL ISOLATES[2]

3.1 INTRODUCTION

HIV gains entry into CD4-expressing cells using the CXCR4 and/or CCR5 co-receptors [1-5]. The capacity to use CXCR4, measured by the ability of the virus to form syncytia in CXCR4 expressing MT-2 cells [6] has been associated with rapid CD4 decline, accelerated disease progression, and reduced survival time in untreated individuals [7-9]. Accurately determining co-receptor tropism is of current clinical concern, especially in the context of screening patients prior to CCR5 antagonist-based therapies, as patients with detectable CXCR4-using virus do not show a significant virological response when administered CCR5 antagonists [10-12].

HIV co-receptor use in clinical samples is most commonly determined using a recombinant phenotype assay [13,14]. Although highly accurate on clonal samples, when these assays [13,14] were compared using clinically derived samples, they showed an 85.1% concordance with each other [15], with discordances likely due to low level minority species [15]. Bioinformatic predictors based on viral genotype may be able to

[2] A version of this chapter has been published as:
Low AJ, Dong W, Chan D, Sing T, Swanstrom R, Jensen M, et. al. Current V3 genotyping algorithms are inadequate for predicting X4 co-receptor usage in clinical isolates. *AIDS*. 2007;**21**(14):F17-F24.

predict co-receptor usage in a more cost-effective and timely manner. The '11/25' charge rule [16], which is based on the presence of positively charged amino acids at positions 11 and/or 25 of the third hypervariable loop (V3 loop) of the envelope glycoprotein gp120, is a simple genotypic predictor of syncytium-inducing (SI) HIV and has been shown to be associated with reduced CD4 response and decreased survival time following the initiation of HAART [17]. Although the '11/25' rule displays >90% sensitivity and specificity for predicting the SI phenotype on clonal HIV sequences [18], this sensitivity is reduced to <60% for predicting the X4/R5 phenotype on clonal data and is further reduced when tested on 'bulk' (population-based) sequences from clinically derived samples [18-21]. Other published bioinformatic methods, such as support vector machines (SVM) [22,23], neural networks (NN) [21], and position specific scoring matrices (PSSM) [24] have also demonstrated relatively high sensitivities for predicting SI HIV (SVM=75%; NN=90%; PSSM=93%), and with a modest decrease in sensitivity, CXCR4-using HIV (SVM=67%; NN=75%; PSSM=62%) when tested on sequences derived from clonal samples [18].

The use of CCR5 antagonists requires screening for co-receptor usage prior to their use. The requirement for screening could result in publicly available genotype-based predictors developed for HIV clones being used for clinical screening purposes. However, the heterogeneous nature of the HIV viral population sampled in the peripheral blood of HIV-infected patients creates difficulty in determining co-receptor usage with both genotype and phenotype based approaches. Genotypic methods trained on clonal data may not be able to accurately interpret sequence ambiguities present in 'bulk'

population-based sequence data, and minority species may be lost during PCR amplification for both genotype and phenotype based approaches.

In this study, we compared the sensitivities and specificities of 5 publicly available co-receptor predictors, which use V3 genotype, including the '11/25' rule [16], two PSSM's ($PSSM_{SI/NSI}$, and $PSSM_{X4/R5}$ [24], a NN, and two SVM's ($SVM_{genomiac}$ [22], and $SVM_{geno2pheno}$ [23]) for predicting CXCR4 co-receptor phenotype (Monogram 'Trofile' assay).

3.2 METHODS

Study subjects: the British Columbia HOMER cohort

The present study represents a baseline cross-sectional analysis of the 'HOMER' [25] cohort, where V3 genotypes and co-receptor phenotypes were determined prior to the initiation of HAART [17,19]. Plasma viral loads, CD4 counts as well as genotypic and phenotypic parameters represent the latest pre-therapy measurements collected in the 180 days prior to HAART initiation [25].

Determination of baseline HIV co-receptor phenotype

The Trofile assay, performed at Monogram Biosciences [13] has previously been used to assess co-receptor phenotype in the latest pre-therapy plasma sample for each subject and the clinical predictors of the co-receptor phenotype were described [19]. Briefly, a RT-PCR product spanning the entire gp160 is digested, purified and ligated into an *E. Coli* expression vector, and gene libraries are constructed. Recombinant viruses are harvested after 48 hours and assessed for their ability to infect U87-CD4 cells expressing CCR5 or CXCR4, as determined by a luciferase read-out measured in Relative Light Units (RLU) on each of the cell lines. The Trofile assay classifies samples as CCR5-using, CXCR4-using or "DM" (indicating dual and/or mixed-tropic virus) based on confirmation of decreased RLU values by 50% or more upon addition of specific co-receptor inhibitors. Phenotypic data from 977 isolates were available, with 799, 177 and one characterized as R5, R5/X4 (DM) and X4-only, respectively, with a higher prevalence of X4-capable virus being detected in individuals with lower CD4 cell

counts [19]. Here, isolates phenotyped as "DM" or "X4" were combined and designated as "DM" for the remainder of analyses; CCR5-only using virus is referred to as R5.

Determination of corresponding baseline envelope V3 sequence

Aliquots of the same baseline plasma samples were used to determine 'bulk' population HIV V3 envelope sequence as described [17]. Matched baseline co-receptor phenotypes and V3 genotypes were available for 953 of 1188 (80.2%) HOMER subjects. Sequences with more than 7 amino acid mixtures in the V3 were excluded due to combinatorial factors (alignment and submission to online algorithms was either not possible or was prohibitively time-consuming), leaving 920 samples with matched baseline co-receptor phenotype and genotype (GenBank accession EF 637088-EF638007). V3 sequences with nucleotide mixtures were translated into all possible amino acid permutations, resulting in 5512 unique V3 amino acid sequences, as well as all possible nucleotide permutations, resulting in 11,447 unique V3 nucleotide sequences, referred to as the amino acid and nucleotide datasets, respectively. Sequence alignments were performed using MUSCLE [26], followed by visual inspection.

Bioinformatic Predictions

After alignment, sequences with positively charged amino acids at codons 11 and/or 25 of the V3 loop, associated with an HIV syncytium-inducing (SI) phenotype, were classified as having an '11/25' genotype. Data from the NN [21] had been previously collected [17], with no changes to the algorithm made since. The $PSSM_{X4/R5}$,

which is the PSSM trained with X4 and R5 data, and PSSM $_{SI/NSI,}$ which is the PSSM trained on SI and NSI data, available at: http://ubik.microbiol.washington.edu/computing/pssm/ (February 2007) [24], use unaligned sequence data from the amino acid dataset as an input, and output categorical (R5 or X4) scores as well as a continuous output variable, which will be referred to as the PSSM score. The $SVM_{genomiac}$, available at: http://genomiac2.ucsd.edu:8080/wetcat/v3.html (February 2007) [22], outputs a categorical score (CCR5 or CXCR4) and uses the amino acid dataset aligned to the following standard, "**CTRPNNNT-RK*I*I--GPG*AFY*-TG*I-IGDIRQAHC**", where (*) indicates any amino acid or gap. The $SVM_{geno2pheno}$ available at: http://coreceptor.bioinf.mpi-sb.mpg.de/cgi-bin/coreceptor.pl (February 2007) [27] also outputs a categorical score (R5 or X4) and uses the unaligned nucleotide dataset as an input. For this analysis, the default false positive rate of 0.1 was chosen to yield a specificity of approximately 90%. Another more recent version of this SVM was also developed to include clinical data (CD4%, number of sequence ambiguities, host CCR5Δ32 heterozygosity, and presence of insertions/deletions) as input variables [28]. This model was compared to the previous SVM to determine the impact clinical markers have on co-receptor prediction. For all predictors, a sample was scored as being X4 if ≥25% of all expanded sequence permutations were classified as X4 [17].

Clonal analysis for low level detection of minority X4-capable variants

Cloning of amplified PCR product was performed using the Invitrogen TOPO TA cloning kit (catalog # K4550-40, Burlington, Ontario, Canada), containing the PCR 2.1-

TOPO vector with chemically competent TOP10F' one shot cells, according to package instructions. Clones (N=48) were picked for 8 samples chosen to address whether 'bulk' sequencing methods contained undetected, low level minority variants. Sequencing was performed on the population sample as well as all viable clones using standard automated sequencing techniques [17].

3.3 RESULTS

V3 genotypes of known co-receptor tropism phenotype (N=920) were submitted to six different V3 genotype-based algorithms. The initial sensitivity and specificity of the '11/25' rule (31% sensitivity/93% specificity), NN (44%/88%), $PSSM_{(sinsi)}$ (34%/96%), $PSSM_{(x4/r5)}$ (24%/97%), $SVM_{genomiac}$ (22%/90%) and $SVM_{geno2pheno}$ (50%/89%) demonstrates that although specificities for all genotype-based methods were high, the sensitivities for all methods when tested on clinically derived samples were relatively low (Table 3.1). The PSSMs provide a raw score (PSSM score) in addition to the categorical predicted phenotype which, after averaging over all permutations, provided a single PSSM score for each sample. This PSSM score could be optimized with receiver-operating characteristic (ROC) curves by reducing the specificity to 90%, resulting in a sensitivity of 43.7% (cutoff = -5.93) for the $PSSM_{SI/NSI}$ and 43.7% (cutoff = -8.12) for the $PSSM_{x4/r5}$. Adjusted sensitivities and specificities represented by a more aggressive approach for the '11/25' rule (31% sensitivity/92% specificity), $PSSM_{(sinsi)}$ (38%/95%), $PSSM_{(x4r5)}$ (27%/96%), $SVM_{genomiac}$ (24%/88%) and $SVM_{geno2pheno}$ (50%/89%) were calculated by categorizing a sample as X4 if any of its sequence permutations were scored as X4 (included in Table 3.1). This method could be applied to both SVM and PSSM methods, but not the NN. In addition, the PSSM methods were further optimized by assigning each sample the highest PSSM score (most likely to be X4) of all its permutations, instead of the average PSSM score. Using ROC curve analysis at a specificity of 90% on this data resulted in a sensitivity of 51.7% (cutoff = -4.81) for the $PSSM_{SI/NSI}$ and 49.0% (cutoff = -7.18) for the $PSSM_{X4/R5}$. Overall, the total concordance of the bioinformatic methods with each other was 61.1% (35.8% in X4

samples only and 66.1% in R5 samples only). When limited to the methods with the greatest sensitivity (the PSSMs and $SVM_{geno2pheno}$) the concordance was 84.6% (X4 samples 78.1%; R5 samples = 85.8%).

In the phenotype assay, luciferase produced in CCR5 or CXCR4 expressing U87 cells is measured in relative light units (RLU). To examine how sensitivity changed with Trofile Assay output, samples were grouped into 9 $\log_{10}$(CXCR4 RLU) strata (Figure 3.1). All optimized predictive methods showed a significant increase in sensitivity with increasing CXCR4 RLU ($p<<0.05$).

The dependence of predictor sensitivity on CXCR4 RLU suggests that population or 'bulk' V3 genotype may be missing low level minority species. To test this hypothesis, 8 patient samples, with plasma viral loads (pVl) above 4 log (to reduce the likelihood of PCR resampling) were chosen for cloning and sequencing. Two samples were chosen with matching R5 phenotype and genotype ('11/25'-) while three samples with matching X4-capable (DM) phenotype and genotype ('11/25'+) acted as controls. The final three samples were chosen with discordant DM phenotype and '11/25'- genotype. A total of 48 clones were picked for each sample and after sequencing, each clone was categorized as R5 or X4 based on the '11/25' rule. All samples which were concordant between phenotype and genotype showed no clones with discordant genotypes from the 'bulk' genotype (Table 3.2). Samples 1 and 2 were both R5 in the phenotype assay and 'bulk' genotype ('11/25' rule) and all clones were genotyped as being R5. Samples 3, 4 and 5 were DM in the phenotype assay and X4 in the 'bulk'

genotype. All clones for samples 3 and 4 were genotyped as being X4, while sample 5 was 23.5% X4 (26 R5 clones; 8 X4 clones). However, samples discordant between the genotype and phenotype showed varying levels of an X4 genotype population in the clonal samples when compared to the 'bulk' sequence indicating a minority species which was undetectable in the 'bulk' genotype. Samples 6, 7 and 8 were DM in the phenotype assay and R5 in the 'bulk' genotype. Sample 6 was 22.2% X4 (35 R5 clones; 10 X4 clones), sample 7 was 21.4% X4 (33 R5 clones; 9 X4 clones) and sample 8 was 8.9% X4 (41 R5 clones; 4 X4 clones).

The positive predictive value (PPV) of the different genotype-to-phenotype predictors increased with decreasing CD4 strata (p=0.03-0.01) (Figure 3.2). Sensitivity did not vary significantly across CD4 strata (data not shown). This dependence of PPV on CD4 indicates a potential benefit for including clinical information such as CD4 count or CD4 percent (CD4%) into co-receptor predictors [27].

3.4 DISCUSSION

Although implementations of HIV V3-loop based co-receptor predictors perform relatively well in clonal samples [18], their performance had not yet been tested on patient-derived samples. Our data suggests these methods, available in February 2007, are too insensitive to be implemented in a clinical setting, due to the presence of low level mixtures in standard 'bulk' genotyping, and require improvement. Adjusting the cutoffs of the raw score output of the PSSM's, [24] and using a more sensitive approach, where any X4 permutation results in an X4 sample to increase sensitivity are two effective examples which highlight the importance of optimizing these methods for testing on clinical samples

The $SVM_{genomiac}$ [22] had the lowest sensitivity of all the predictors, which is in contrast to the $SVM_{geno2pheno}$ [23] (one the best performing predictors), although both predictors were based on the SVM machine-learning model. The $SVM_{genomiac}$ was trained on a dataset comprised of multiple HIV clades, while all other methods were trained on datasets which were composed primarily of clade-B virus. The clinical test set from BC is composed of 97.5% clade-B virus, thus biasing the results in favor of those methods trained on primarily clade-B sequences. Additionally, the PSSM and the $SVM_{geno2pheno}$ were trained on the same set of clonal samples, and both methods performed nearly equally well at all CXCR4 RLU and CD4 strata, despite the different approaches used.

In another study, predictions were performed using the $SVM_{geno2pheno}$ model, but with additional clinical parameters included ($\log_{10}$(CD4%); host Δ32 heterozygosity;

number of ambiguous amino acid V3 positions; and a variable indicating the presence of insertions or deletions in the V3 sequence) [28]. Sensitivity of the predictor with clinical parameters was approximately 10% greater than the purely sequence based approach at a specificity of 90% [28]. Although it has been previously demonstrated that there is no significant association between CXCR4 RLU and CD4 count [29], the association of CD4 count with PPV and the increase in sensitivity, achieved by integrating clinical information into training and prediction, indicate the importance of integrating clinical parameters for co-receptor prediction. However, it is not yet clear whether low CD4 are indicative of a more favorable immune environment for X4-capable virus or if CD4 count is an independent predictor of X4-capable virus. The log-linear dependence of CXCR4 RLU sensitivity for all methods indicates that samples which produce higher levels of CXCR4 RLU are more likely to be genotypically characterized as X4. Lower CXCR4 RLU signals may indicate a minority species at levels undetectable to 'bulk' sequencing techniques and/or a genotypically unique X4-tropic viral species present in vivo, which predictors trained on clonal data cannot detect.

Results of the clonal analyses performed here indicate that 'bulk' sequencing techniques were unable to reliably detect minority species present below approximately 22%, which is in agreement with previous studies that estimate the limit of detection for minority species by direct sequencing of bulk PCR products to be between 10-25% [30-33]. Unlike genotype-based HIV antiretroviral resistance testing, low level minority species of X4-using virus appear to be the rule rather than the exception. This is likely a reason for much of the discordance and lack of sensitivity observed, and also indicates

that an area for improvement would be increasing the sensitivity of detection for minority variants. Techniques which may improve the detection of minority variants include adopting more aggressive base-calling techniques or examining the area under the curve of the chromatograms as a variable when nucleotide mixtures are observed. New technologies capable of more sensitive base-calling or detection of minority DNA species [34] may result in greatly increased sensitivities for genotype-based predictors. Alternatively, methods capable of separating genotypic variants, such as the heteroduplex tracking assays [35] combined with sequence analysis may provide another approach for the detection of minority variants.

In this large clinically based study, all V3 genotype and co-receptor phenotype results were derived on the same set of samples using well defined methodologies, and all predictive methods were tested on the same data. However, this study is limited by the fact that associations between HIV co-receptor usage and clinical parameters are based on cross-sectional data from a population of therapy-naïve individuals initiating their first antiretroviral treatment during the period 1996 to 1999, and may not be representative of the HIV-infected, population in general, nor of individuals treated with co-receptor antagonists. It should be emphasized that these results were limited to therapy naïve individuals and therefore, cannot necessarily be extrapolated to therapy experienced individuals. The methods investigated in this study limited their training and test set to only the V3 loop of the gp120 surface glycoprotein. Although the V3-loop region contains the majority of mutations predictive of co-receptor usage, expanding the genotypic sequence sampled to regions outside the V3 may result in increased

sensitivities. In addition, the Monogram Trofile assay, or any other co-receptor assay, cannot yet be considered a 'gold-standard'. This was illustrated in a comparison of the Trofile assay to another well-validated phenotype co-receptor assay, which yielded approximately 85% concordance [15]. Also, in clinical trials, maraviroc selected for pre-existing X4 variants which were undetected by the Trofile assay [36]. Therefore, although this assay is in the unique position of being the sole test used in screening for co-receptor clinical trials, other phenotype assays are being developed [37,38], and there is still room for improvement.

In conclusion, the current sensitivities of genotype-based co-receptor predictors between 30-50% are inadequate for the prediction of CXCR4 co-receptor usage in a clinical setting although the exact level of desired sensitivity is not known, sensitivities upwards of 85% would be equivalent to the concordance of co-receptor phenotype assays with one another [15] and should be sufficient. Although, even if a genotype-based method approached or exceeded the sensitivity of the phenotype assay, it would still need to be validated before it could be used as a widespread clinical tool. Importantly, the results obtained here indicate that the type of predictive model used is not nearly as important as the reliability and effectiveness of the parameters chosen to represent the phenotype. Genotype-based co-receptor predictors do have the potential to significantly decrease turn-around time and cost in comparison to current phenotype assay approaches, and although current implementations of genotype-based predictors are inadequate, there are identifiable methods to improve their sensitivity. In the interim, however, these

approaches should not be used alone as screening tools for clinical use of co-receptor antagonists.

Table 3.1: Initial Predictive performance of V3 genotyping algorithms.

Method	Sensitivity	Specificity	$Sensitivity_{Adjusted}$	$Specificity_{Adjusted}$
11/25 rule	30.5%	93.4%	30.5%	92.3%
Neural Network	44.4%	87.5%	-	-
$PSSM_{SI/NSI}$	33.8% (43.7%)*	95.3% (90%)	37.7% (51.7%)*	94.9% (90%)
$PSSM_{X4/R5}$	24.5% (43.7%)*	96.9% (90%)	26.5% (49.0%)*	96.3% (90%)
$SVM_{genomiac}$	21.8%	89.6%	23.8%	88.4%
$SVM_{geno2pheno}$	44.7%	90.6%	50.3%	89.2%

Adjusted values classified a sample as X4 if ANY permutation was classified as X4
*Values in brackets were obtained by adjusting the specificity of the PSSM score to 90%
Table 3.1 lists the sensitivity and specificity of all methods. Adjusted methods were obtained by calling a sample X4 if any permutation of a sequence containing amino acid mixtures was X4. This data was not available to be modified for the neural network method.

Table 3.2: Clonal analysis to determine levels of undetected minority species.

Sample ID	Phenotype[1]	CXCR4 RLU (log_{10})	CCR5 RLU (log_{10})	'Bulk' Genotype[2]	Number 11/25- Clones[2]	Number 11/25+ Clones[2]	Percent 11/25+ (%)
1	R5	1.95	6.12	R5	47	0	0.0
2	R5	2.08	6.26	R5	41	0	0.0
3	R5/X4	5.18	5.32	X4	0	46	100.0
4	R5/X4	5.45	5.71	X4	0	44	100.0
5	R5/X4	4.00	5.86	X4	26	8	23.5
6	R5/X4	4.78	5.83	R5	35	10	22.2
7	R5/X4	3.93	5.64	R5	33	9	21.4
8	R5/X4	2.95	4.75	R5	41	4	8.9

[1] Using the Monogram Trofile Assay

[2] Using the '11/25' rule

To determine if, and at what proportion, 'bulk' sequencing of the V3 is missing minority genotypic variants, samples were cloned (48 attempted clones each), sequenced and then compared against the 'bulk' sequence (using the '11/25' rule). Included is the phenotype, the RLU from the CXCR4 and CCR5 expressing cell lines, the bulk genotype ('11/25' rule), the number of '11/25'- and '11/25'+ clones, as well as the percentage of '11/25'+ clones present.

Figure 3.1: Sensitivity of genotype-based predictors stratified by CXCR4 RLU.

Sensitivity, defined as the proportion of all positives (DM by Phenotype Assay) detected by the predictive method, is calculated for all predictors grouped within 11 CXCR4 RLU strata: <2.6 (N=11), 2.6-3.0 (N=22), 3.0-3.4 (N=23), 3.4-3.8 (N= 13), 3.8-4.2 (N=21), 4.2-4.6 (N=26), 4.6-5.0 (N=15), 5.0-5.4 (N=10), ≥5.4 (N=10). Linear regressions are derived from the sensitivity of each stratum, where each stratum is represented by incremental integer values. Symbols and linear regression coefficients for the methods used are:
'11/25' rule (■ with a dashed, black trendline; R^2=0.77), $PSSM_{SI/NSI}$ (♦ with a solid black trendline; R^2=0.94), $PSSM_{X4/R5}$ (● with a solid grey trendline; R^2=0.90);
B) Neural Network (■ with a dashed, black trendline; R^2=0.85), $SVM_{genomiac}$ (♦ with a solid black trendline; R^2=0.52), $SVM_{geno2pheno}$ (● with a solid grey trendline; R^2=0.83).

Figure 3.1A

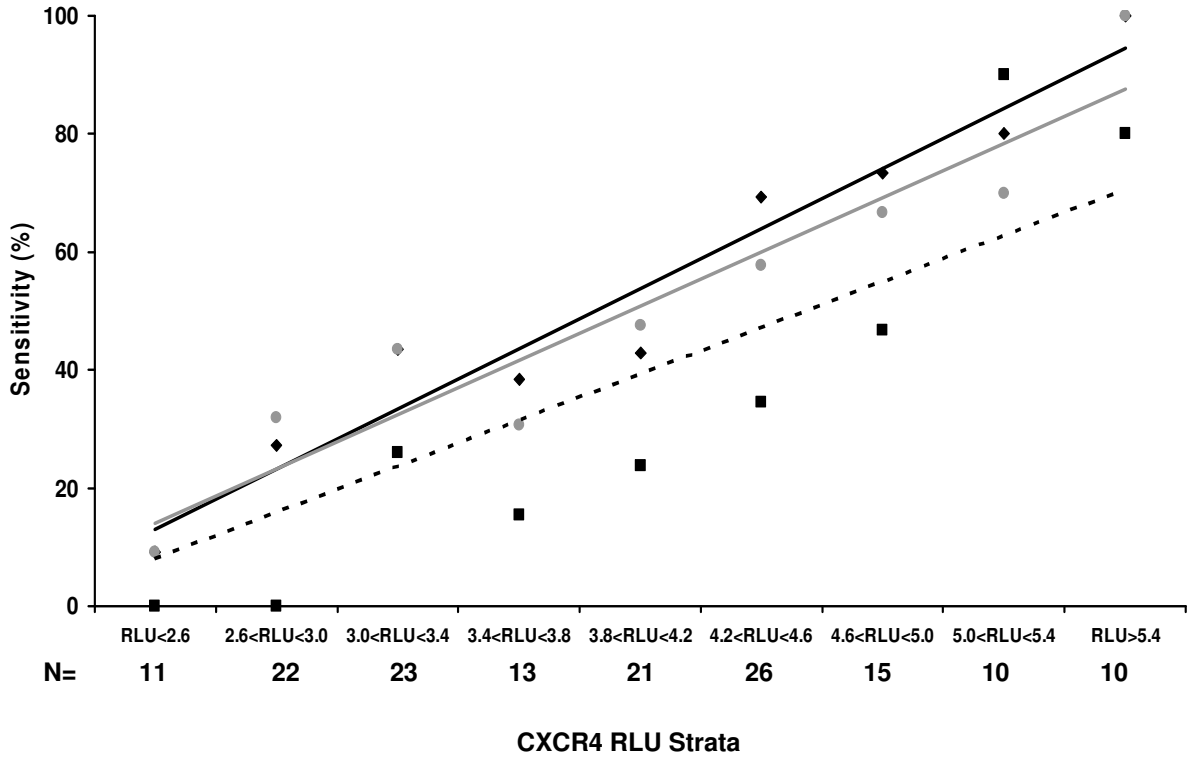


Figure 3.1B

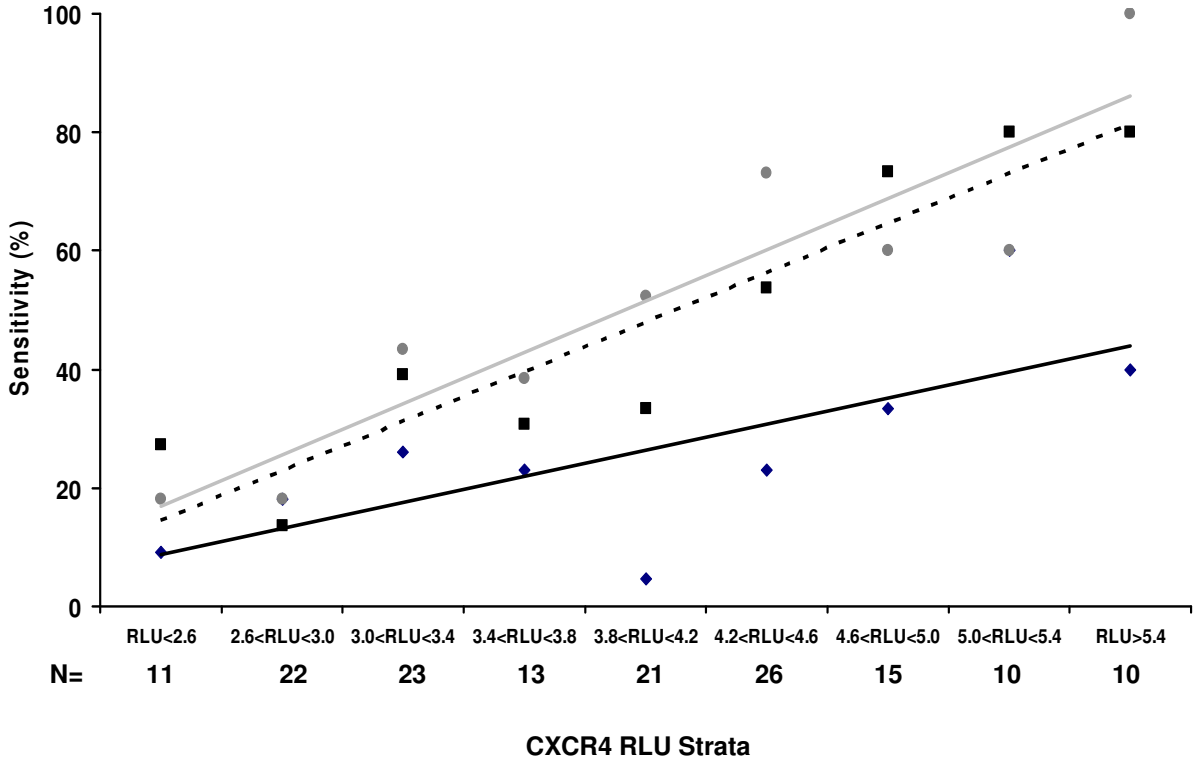

100
80
60
40
20
0
Sensitivity (%)
RLU<2.6
2.6<RLU<3.0
3.0<RLU<3.4
3.4<RLU<3.8
3.8<RLU<4.2
4.2<RLU<4.6
4.6<RLU<5.0
5.0<RLU<5.4
RLU>5.4
N=
11
22
23
13
21
26
15
10
10
CXCR4 RLU Strata

Figure 3.2: Positive predictive value for X4-tropic virus stratified by CD4 count.

Positive predictive value, defined as the proportion of all predicted positives which are true positives, is calculated for all predictors grouped within 7 CD4 strata: <25 (N=69), 25-49 (N=56), 50-99 (N=75), 100-199 (N=163), 200-349 (N=238), 350-499 (N=182), >500 (N=137). Linear regressions are derived from the sensitivity of each stratum, where each stratum is represented by incremental integer values. Symbols and linear regression coefficients for the methods used are:
A) '11/25' rule (■ with a dashed, black trendline; R^2=0.71), $PSSM_{SI/NSI}$ (♦ with a solid black trendline; R^2=0.65), $PSSM_{X4/R5}$ (● with a solid grey trendline; R^2=0.80);
B) Neural Network (■ with a black trendline made a series of dots; R^2=0.80), $SVM_{genomiac}$ (♦ with a solid black trendline; R^2=0.62), $SVM_{geno2pheno}$ (○ with a solid grey trendline; R^2=0.91).

Figure 3.2A

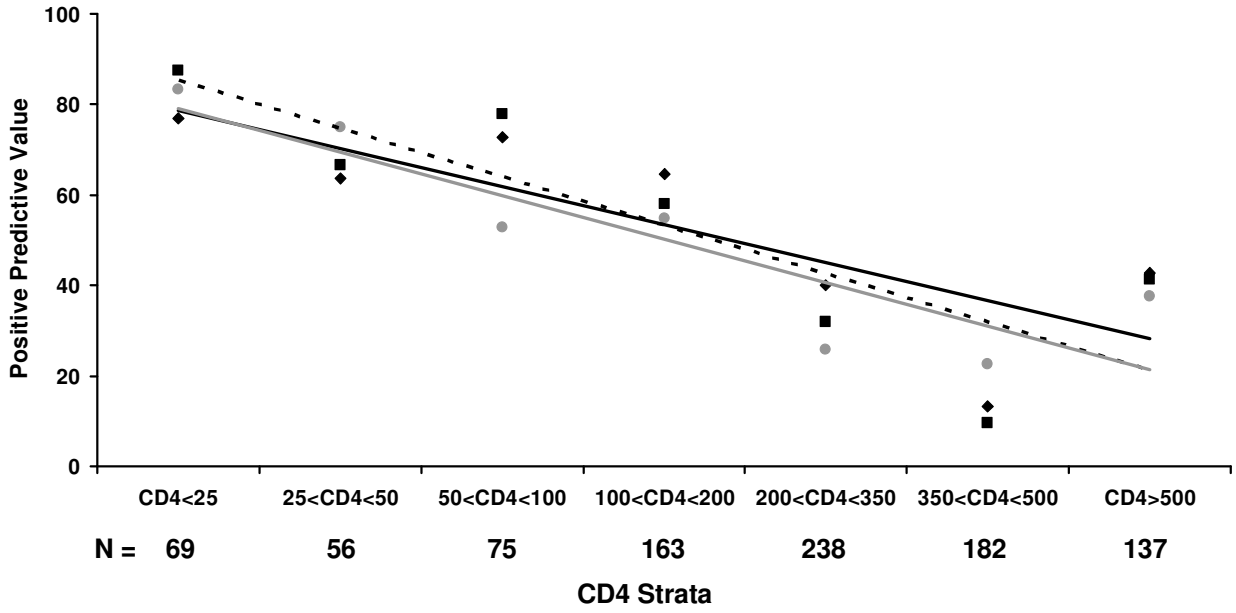
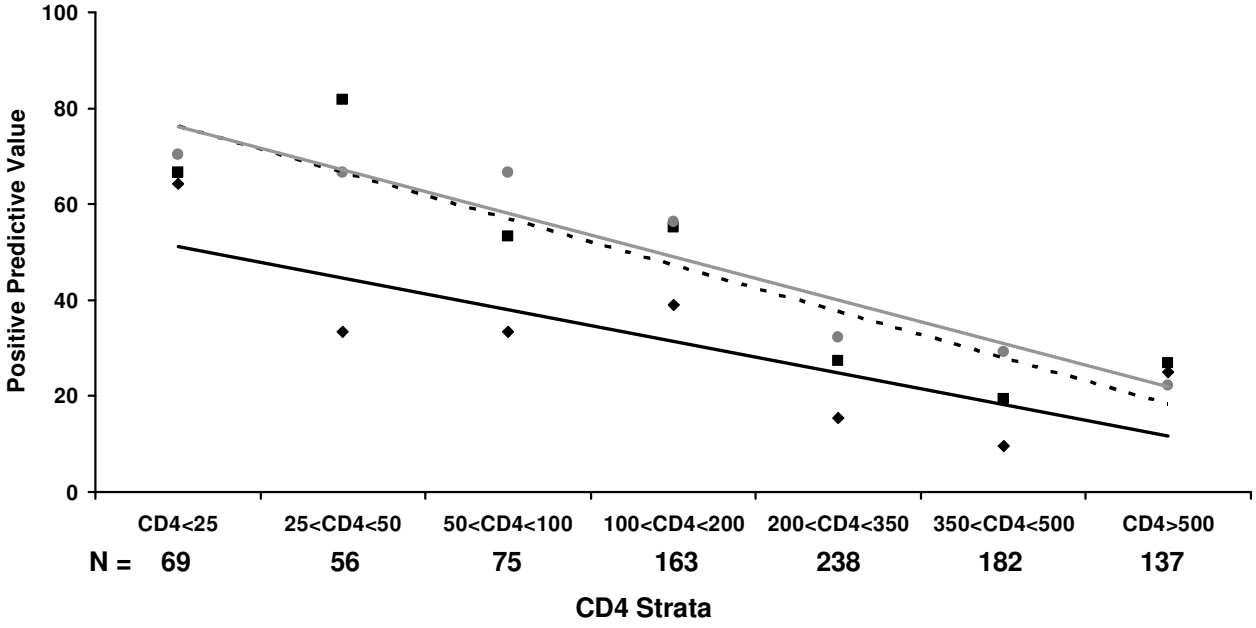


Figure 3.2B


Positive Predictive Value
100
80
60
40
20
0
CD4<25
25<CD4<50
50<CD4<100
100<CD4<200
200<CD4<350
350<CD4<500
CD4>500
N = 69 56 75 163 238 182 137
CD4 Strata

ACKNOWLEDGEMENTS

Ethical approval for this study was obtained from the institutional ethics board (Providence HealthCare/University of British Columbia).

We would like to thank James Goodrich and Howie Mayer from Pfizer Inc., as well as Brian Wynhoven, Theresa Mo and Conan Woods from the BC Centre for Excellence in HIV/AIDS for their contributions to this study. This study was supported through an Rx&D research grant from the Canadian Institutes for Health Research (CIHR) and by funding from Pfizer, Inc. Richard Harrigan has previously consulted for Monogram, Virco, and other companies in the pharmaceutical and diagnostic industries.

3.5 REFERENCES

1. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, et al. Identification of a major co-receptor for primary isolates of HIV-1. *Nature.* 1996; **381**(6584):661-666.

2. Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, et al. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature.* 1996; **381**(6584):667-673.

3. Feng Y, Broder CC, Kennedy PE,Berger EA. HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science.* 1996; **272**(5263):872-877.

4. Moore JP, Kitchen SG, Pugach P,Zack JA. The CCR5 and CXCR4 coreceptors--central to understanding the transmission and pathogenesis of human immunodeficiency virus type 1 infection. *AIDS Res Hum Retroviruses.* 2004; **20**(1):111-126.

5. Zhang L, He T, Huang Y, Chen Z, Guo Y, Wu S, et al. Chemokine coreceptor usage by diverse primary isolates of human immunodeficiency virus type 1. *J Virol.* 1998;

72(11):9307-9312.

6. Schuitemaker H, Kootstra NA, de Goede RE, de Wolf F, Miedema F,Tersmette M. Monocytotropic human immunodeficiency virus type 1 (HIV-1) variants detectable in all stages of HIV-1 infection lack T-cell line tropism and syncytium-inducing ability in primary T-cell culture. *J Virol.* 1991; **65**(1):356-363.

7. Schuitemaker H, Koot M, Kootstra NA, Dercksen MW, de Goede RE, van Steenwijk RP, et al. Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: Progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus population. *J Virol.* 1992; **66**(3):1354-1360.

8. Tersmette M, Gruters RA, de Wolf F, de Goede RE, Lange JM, Schellekens PT, et al. Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency syndrome: Studies on sequential HIV isolates. *J Virol.* 1989; **63**(5):2118-2125.

9. Koot M, Keet IP, Vos AH, de Goede RE, Roos MT, Coutinho RA, et al. Prognostic value of HIV-1 syncytium-inducing phenotype for rate of CD4+ cell depletion and progression to AIDS. *Ann Intern Med.* 1993; **118**(9):681-688.

10. H. Mayer, Van der Ryst, E., M. Saag, B. Clotet, G. Fätkenheuer, N. Clumeck, et al. Safety and efficacy of maraviroc, a novel CCR5 antagonist, when used in combination with optimized background therapy for the treatment of antiretroviral-experiences subjects infected with dual/mixed-tropic HIV-1: 24-week results of a phase 2b exploratory trial.[Abstract]. *XVI International AIDS Conference, Toronto, Ontario, Canada* August 13-18, 2006; THLB0215.

11. M. Nelson, G. Fätkenheuer, I. Konourina, A. Lazzarin, N. Clumeck, A. Horban, et al. Efficacy and Safety of Maraviroc plus Optimized Background Therapy in Viremic, ART-experienced Patients Infected with CCR5-tropic HIV-1 in Europe, Australia, and North America: 24-Week Results [Abstract]. *14th Conference on Retroviruses and Opportunistic Infections, Los Angeles, California, USA* February 25-28, 2007; 104aLB.

12. J. Lalezari, J. Goodrich, E. DeJesus, H. Lampiris, R. Gulick, M. Saag, et al. Efficacy and Safety of Maraviroc plus Optimized Background Therapy in Viremic ART-experienced Patients Infected with CCR5-tropic HIV-1: 24-Week Results of a Phase 2b/3 Study in the US and Canada [Abstract]. *14th Conference on Retroviruses and Opportunistic Infections, Los Angeles, California, USA* February 25-28, 2007; 104bLB.

13. Whitcomb JM, Huang W, Fransen S, Limoli K, Toma J, Wrin T, et al. Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism. *Antimicrob Agents Chemother.* 2007; **51**(2):566-575.

14. Trouplin V, Salvatori F, Cappello F, Obry V, Brelot A, Heveker N, et al. Determination of coreceptor usage of human immunodeficiency virus type 1 from patient plasma samples by using a recombinant phenotypic assay. *J Virol.* 2001; **75**(1):251-259.

15. Skrabal K, Low AJ, Dong W, Sing T, Cheung PK, Mammano F, et al. Determining human immunodeficiency virus coreceptor use in a clinical setting: Degree of correlation between two phenotypic assays and a bioinformatic model. *J Clin Microbiol.* 2007; **45**(2):279-284.

16. 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. *J Virol.* 1992; **66**(11):6777-6780.

17. Brumme ZL, Dong WW, Yip B, Wynhoven B, Hoffman NG, Swanstrom R, et al. Clinical and immunological impact of HIV envelope V3 sequence variation after

starting initial triple antiretroviral therapy. *AIDS.* 2004; **18**(4):F1-9.

18. Jensen MA, van 't Wout AB. Predicting HIV-1 coreceptor usage with sequence analysis. *AIDS Rev.* 2003; **5**(2):104-112.

19. Brumme ZL, Goodrich J, Mayer HB, Brumme CJ, Henrick BM, Wynhoven B, et al. Molecular and clinical epidemiology of CXCR4-using HIV-1 in a large population of antiretroviral-naive individuals. *J Infect Dis.* 2005; **192**(3):466-474.

20. Fouchier RA, Brouwer M, Broersen SM,Schuitemaker H. Simple determination of human immunodeficiency virus type 1 syncytium-inducing V3 genotype by PCR. *J Clin Microbiol.* 1995; **33**(4):906-911.

21. 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.

22. Pillai S, Good B, Richman D,Corbeil J. A new perspective on V3 phenotype prediction. *AIDS Res Hum Retroviruses.* 2003; **19**(2):145-149.

23. T. Sing, T. Lengauer, N. Beerenwinkel, R. Kaiser, M. Daumer, A. Low, et al. Immunological and virological parameters improve genotypic prediction of HIV-1 coreceptor usage [Abstract]. *BioSapiens-viRgil Workshop on Bioinformatics for Infectious Diseases, Bonn, Germany.* September 21-23, 2005; 22.

24. Jensen MA, Li FS, van 't Wout AB, Nickle DC, Shriner D, He HX, et al. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences. *J Virol.* 2003; **77**(24):13376-13388.

25. Hogg RS, Yip B, Chan KJ, Wood E, Craib KJ, O'Shaughnessy MV, et al. Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy. *JAMA.* 2001; **286**(20):2568-2577.

26. Edgar RC.MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; **32**(5):1792-1797.

27. T. Sing, R. Harrigan, N. Beerenwinkel, M. Daumer, R. Kaiser, A. Low, et al. Immunologic markers improve genotypic prediction of HIV-1 co-receptor usage [Abstract]. *Int. Workshop on Targeting HIV Entry.* December 2-3, 2005; 30

28. T. Sing, P. R. Harrigan, N. Beerenwinkel, M. Daumer, R. Kaiser, A. Low, et al. Immunologic markers improve genotypic prediction of HIV-1 coreceptor usage [Abstract]. *Keystone symposium on genome sequence variation and the inherited basis of common disease and complex traits, Big Sky, Montana, USA* Jan 8-13, 2006; 30.

29. A. J. Low, C. J. Brumme, Z. L. Brumme, R. S. Hogg, J. S. Montaner and P. R. Harrigan.
Immunologic, Virologic, and Genotypic Correlates of Phenotypic CCR5/CXCR4 Co-receptor Usage Assay Readouts in a Large Panel of Clinically Derived Isolates [Abstract]. *14th Conference on Retroviruses and Opportunistic Infections, Los Angeles, California, USA* February 25-28, 2007; 177.

30. Larder BA, Kohli A, Kellam P, Kemp SD, Kronick M,Henfrey RD. Quantitative detection of HIV-1 drug resistance mutations by automated DNA sequencing. *Nature.* 1993; **365**(6447):671-673.

31. Leitner T, Halapi E, Scarlatti G, Rossi P, Albert J, Fenyo EM, et al. Analysis of heterogeneous viral populations by direct DNA sequencing. *BioTechniques.* 1993; **15**(1):120-127.

32. 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. *J Clin Microbiol.* 1999; **37**(7):2291-2296.

33. Van Laethem K, Van Vaerenbergh K, Schmit JC, Sprecher S, Hermans P, De Vroey V, et al. Phenotypic assays and sequencing are less sensitive than point mutation assays for detection of resistance in mixed HIV-1 genotypic populations. *J Acquir Immune Defic Syndr.* 1999; **22**(2):107-118.

34. M. Lewis, I. James, M. Braverman, B. Desany, T. Jarvie, M. Penny, et al. Evaluation of an Ultra-deep Sequencing Method to Identify Minority Sequence Variants in the HIV-1 *env* Gene from Clinical Samples [Abstract]. *14th Conference on Retroviruses and Opportunistic Infections, Los Angeles, California, USA* February 25-28, 2007; 680.

35. Nelson JA, Baribaud F, Edwards T,Swanstrom R. Patterns of changes in human immunodeficiency virus type 1 V3 sequence populations late in infection. *J Virol.* 2000; **74**(18):8494-8501.

36. Lewis M, Simpson P, Francen S, Huang W, Whitcomb J, Mosley M, et. al. CXCR4-using virus detected in patients receiving maraviroc in the Phase III studies MOTIVATE 1 and 2 originates from a pre-existing minority of CXCR4-using virus. *16th International HIV Drug Resistance Workshop, Barbados, West Indies*. June 12-16, 2007; 56

37. Roulet V, Faudon J-L, Rochas S, Labernardiere J-L, Mammano F, Level-Binay S, et. al. A sensitive assay for the determination of viral tropism suitable for clinical management of HIV subtype B and non-B infected individuals treated with coreceptor antagonists : PHENOSCRIPT® ENV [Abstract]. *16th International HIV Drug Resistance Workshop, Barbados, West Indies*. June 12-16, 2007; 147

38. Vandenbroucke I, Van Eygen V, Rondelez E, Van Baelen K, Styver LJ. Detection limit of X4 minority virus in abundant R5-tropic virus populations in standard genotypic and phenotypic assays [Abstract]. *16th International HIV Drug Resistance Workshop, Barbados, West Indies*. June 12-16, 2007; 139

CHAPTER 4

IS THERE A GOLD STANDARD FOR DETERMINING HIV CO-RECEPTOR USE IN CLINICAL SAMPLES? THE DEGREE OF CORRELATION BETWEEN TWO PHENOTYPIC ASSAYS AND A BIOINFORMATIC MODEL[3]

4.1 Introduction

HIV-1 can be characterized according to the host chemokine co-receptor used to gain entry into CD4-expressing cells. The phenotypic designations are R5 (for CCR5-using variants), X4 (for CXCR4-using variants) or R5/X4 (defined here either as a mixture of both R5 and X4 variants or as a dual-tropic variant able to enter cells using both CXCR4 and CCR5 co-receptors). Most population-based phenotype assays cannot readily distinguish between virus populations that are truly dual-tropic and those that are comprised of mixtures of viruses with different co-receptor phenotypes [1]. The fact that viral populations within an individual may contain heterogeneous mixtures composed of any combination of the three classes of virus complicates the accurate and reliable assessment of tropism in clinical samples.

Recently, with the development of co-receptor antagonists capable of blocking either the CCR5 [2,3] or the CXCR4 [3,4] co-receptors, the need for accurate HIV

[3] A version of this chapter has been published as:
Skrabal K, Low AJ, Dong W, Sing T, Cheung PK, Mammano F, et. al. Is there a gold standard for determining HIV co-receptor use in clinical samples? The degree of correlation between two phenotypic assays and a bioinformatic model. *J. Clin. Microbiol.* 2007;**45**(2):279-84.

phenotyping assays has increased. This is especially important in preliminary tests to screen patients for clinical trials of these agents and for monitoring the evolution of co-receptor usage under the pressure of CCR5 and/or CXCR4 antagonists [5].

Many, but not all of the genetic determinants of co-receptor usage reside in the HIV envelope, in particular, the V3 loop [6]. Consequently, this region is highly useful for the training and development of computerized genotypic predictors of co-receptor usage, which have the potential to make rapid screening and testing of HIV infected patients faster and more cost-effective. The simplest genotypic approach, known as the "11/25 rule" classifies a virus as X4 if positively charged amino acids (lysine or arginine) are present at positions 11 and/or 25 of the V3 loop [7,8]. Evaluation of bioinformatics-based genotypic predictors such as neural networks [9], position specific scoring matrices (PSSM) [10] and support vector machines indicate that the co-receptor phenotypes may be accurately determined with genotypic predictors [11] when tested on clonally derived sequence data. However, both phenotypic assays and genotypic prediction methods are challenged when applied to heterogeneous sequence data obtained from actual clinical samples. If testing for HIV co-receptor usage becomes necessary for treatment with co-receptor inhibitors, it would be expected to be performed on clinically derived samples. Under these circumstances, a high level of reliability and repeatability should be expected.

Two previously validated recombinant phenotypic assays for HIV co-receptor use (the Trofile HIV Co-receptor Tropism Assay –formerly PhenoSense HIV-Entry Assay-

by Monogram Biosciences and the Tropism Recombinant Test (TRT) assay, performed in the lab of Fabrizio Mammano at INSERM (to be available through Eurofins Viralliance Inc) were employed here, as well as genotypic analysis of the HIV V3 loop on 74 clinically derived HIV-1 samples in order to evaluate concordance between phenotypic assays and determine the ability of an SVM based genotypic method to predict co-receptor usage.

4.2 Methods

Sample Selection

Samples represented a non-random selection of 93 baseline samples from the well characterized HOMER cohort [12], a large cohort of antiretroviral-naïve individuals initiating triple combination therapy in British Columbia. Baseline plasma samples from this cohort have been previously genotyped for HIV envelope V3 loop sequence [13]. Samples were chosen for analysis on the basis of predicted phenotype using the 11/25 charge rule [6,7,8,14], a neural network method [9], and results from a position-specific scoring matrix (PSSM) analysis [10], in order to over-represent expected X4 variants to an approximate 50% prevalence. Note that the entire HOMER cohort was later phenotyped for HIV co-receptor usage; these results have been subsequently published [15].

Assay Methods

The two phenotypic assays [16,17] use virus stocks pseudotyped with envelope sequences derived from patient plasma samples to infect cell lines engineered to express CCR5 or CXCR4. Some relevant differences in the phenotypic assays are highlighted in Table 4.1. The Trofile assay confirms co-receptor usage by verifying a decrease in RLU after the addition of a CCR5 or CXCR4 antagonist. The TRT assay did not currently perform this step at the time of these analyses. Note that the TRT assay was performed on PCR products amplified by our lab in British Columbia, and the Trofile assay was performed on the corresponding baseline plasma samples. More detailed information

regarding the two phenotypic assays can be found in [1,16,18] (Trofile) and [1,17] (TRT).

SVM Prediction

Population based HIV envelope V3 loop sequences (~105 bp) of corresponding baseline plasma samples were obtained using previously described, standard automated sequence analysis [13]. Genotypic predictions of the HIV co-receptor usage were performed using a support vector machine (SVM) analysis of the sequenced HIV V3 loop (www.geno2pheno.org). Robust alignments of the envelope V3 region were obtained using the multiple alignment software MUSCLE [19]. SVMs were trained using all genotype-phenotype data from the HOMER cohort, determined using the Trofile assay, excluding the 74 samples compared here. Codon mixtures in sequence data are considered ambiguous with respect to sequence identity and were treated using an "X4-sensitive" strategy during training and a "combinatorial" strategy for prediction, as described in [11]. Briefly, in the combinatorial prediction strategy, non-ambiguous V3 sequences are generated by random sampling from the ambiguous genotype and a numeric score, indicating the probability of CXCR4 usage, is predicted for each of the non-ambiguous sequences. The final score is taken as the 0.75 percentile of the individual scores. A cutoff of 0.4 on this score was chosen to predict X4 variants. This cutoff was determined by maximizing the area under the ROC curve for all matched genotype-phenotype samples in the HOMER cohort, excluding the 74 samples analyzed here.

4.3 Results

Phenotypic Assay Comparison

Of the 93 samples tested, results were obtained for 78 (84.9%) samples with the Trofile co-receptor tropism assay and results were obtained for 90 (96.8%) samples with the TRT co-receptor usage assay. Note that no comparison can be made with respect to the number of results obtained by each assay, as the TRT assay was performed on PCR products amplified by our lab in British Columbia, while the Trofile assay was performed on the corresponding baseline plasma sample. Results from the two phenotypic assays were obtained for 74 samples (79.6%) and were in general, but not complete agreement.

Overall, an 85.1% concordance was observed between the two assays. Of the 74 samples, 42 (56.8%) were identified as R5 by both assays and 21 (28.4%) were identified as R5/X4 by both assays. An additional 8 (10.8%) were identified as R5 by the Trofile assay but as R5/X4 by the TRT assay and the remaining 3 (4.1%) were identified as R5/X4 by the Trofile assay but as R5 by the TRT assay. The Trofile assay characterized one sample as being pure X4. This sample was characterized as being R5/X4 by the TRT assay, but for the purposes of our analysis was classified as being R5/X4 by the Trofile assay and therefore treated as concordant.

We then investigated clinical parameters of samples with concordant and discordant results (Table 4.2). Low CD4 cell count has been shown to be predictive of CXCR4 using HIV [20], and as expected, the samples identified as R5 in both assays showed significantly higher median CD4 cell counts ($p<0.01$) than those samples

identified as R5/X4 in both assays, although plasma viral loads were not significantly different ($p>0.1$). Median R5 RLU (Trofile assay) and R5 OD (TRT assay) values were consistently high in both assays, regardless of whether these were identified as R5 or R5/X4, implying a high efficiency of replication in the R5 cell lines. Discordant results were observed where R5/X4 was detected in the Trofile assay and not the TRT assay, and vice versa (shown individually in Table 4.3). When samples were grouped as "discordant" and "concordant", we observed no statistically significant difference in viral load or CD4 count. ($p>0.05$). Also, when discordant groups were compared against each other, we observed no statistically significant difference in CD4 or viral load ($p>>0.5$) (Wilcoxon rank sum test), although the small sample size of the discordant groups should be noted.

The median X4 signal measured in RLU (Trofile assay) or OD_{sample}/OD_{Ctrl} (TRT assay) of the discordant samples was close to the lower quartile of the X4 signal for samples concordantly phenotyped as being R5/X4 by both assays. Therefore, a possible explanation for discordances is misclassification as a result of using OD or RLU "cut-off" values for categorizing samples as R5 or R5/X4. Receiver Operator Characteristic (ROC) curves examine the effects of all possible cut-off values (for X4 RLU [Trofile Entry assay] or X4 OD [TRT assay]) on sensitivity (ability to identify true positives) and specificity (ability to identify false positives). Both assays appeared to use cut-off values such that their sensitivity and specificity were optimized by placing them in the upper left quadrant of the ROC graph (Figure 4.1). In comparing the area under the ROC curve (AUC), the X4 RLUs resulted in an AUC of 0.88 when using the TRT assay as the gold

standard and the X4 ODs resulted in an AUC of 0.87 when using the Trofile assay as the gold standard. There appeared to be no significant difference between the two assays when using one assay as a reference and varying the cutoffs used for the other assay and vice-versa ($p>0.1$), suggesting that the use of different cut-off values in either assay would not improve the predictive performance of one phenotypic assay with respect to the other. Of interest, 6 of the 11 discordant samples (54.5%) had positively charged amino acids at codons 11 and/or 25, (Table 4.3) implying a high probability that many of these discordant samples harbored X4 variants which gave only marginally detectable X4 signal strength.

Genotypic Prediction

Using the HIV envelope V3 sequences, genotypic predictions of co-receptor usage were made using a Support Vector Machine (SVM) trained on separate V3 genotype-phenotype (Trofile assay) data, as described above, excluding the 74 samples tested here. In the dataset of 74 clinically derived isolates, amino acid mixtures at ≥1 codon of the HIV V3 loop were observed in 56 (75.7%) of sequences. According to the "combinatorial" prediction strategy, mixture-containing sequences were expanded into all possible permutations (median of 2 permutations per isolate, range 1 to 128), giving rise to multiple values for each sample, each falling between 0 (low likelihood of being X4) and 1 (high likelihood of being X4) (Figure 4.2).

Samples were grouped according to the results obtained from both phenotypic co-receptor assays (concordant **R5**/R5, concordant **R5X4**/R5X4, discordant **R5X4**/R5, and

discordant **R5**/R5X4, where the Trofile result is indicated first and in bold lettering). When comparing between groups, the maximum X4 score per sample was used. With regards to the Trofile assay, 16 of the 24 samples phenotyped as R5/X4 were predicted to be R5/X4 by the SVM, resulting in a sensitivity of 67%. A total of 48 of the 50 samples phenotyped as R5 were predicted of being R5 by the SVM, yielding a specificity of 96%. Similarly, for the TRT assay, 16 of the 28 samples phenotyped as R5/X4 were predicted to be R5/X4 by the SVM, yielding a sensitivity of 55% and 43 of the 45 samples phenotyped as R4 were predicted of being R5, yielding a specificity of 96%. In comparison, the "11/25" rule, had a sensitivity and specificity for predicting X4 usage in the Trofile assay of 54% and 44%, respectively, and 52% and 42% for the TRT assay. The median X4 scores of the samples classified as R5/X4 by both phenotypic co-receptor assays was 0.68. This score was significantly higher (Figure 4.3) than the median 0.08 X4 score of the samples classified as R5 by both phenotypic co-receptor assays ($p<<0.001$). It also exceeded the median X4 scores for the samples classified as R5 by the Trofile assay and R5/X4 by the TRT assay (0.11; $p<0.01$) and vice versa (0.10; $p<0.01$). There was no significant difference in X4 score between the two groups with discordant phenotypic co-receptor classifications ($p=0.99$). Two outliers in the **R5**/R5 and one outlier in the **R5**/R5X4 categories may be identified in Figure 4.3. These outliers have a very high probability of harboring X4 as determined by the genotypic predictor, but are categorized as R5 according to one or both of the phenotypic co-receptor assays. Of note, samples classified as R5/X4 virus by both phenotypic assays were more heterogeneous (contained more amino acid mixtures in the V3 loop) than samples classified exclusively as R5 [median 1.5 amino acid mixtures/sequence, IQR[0.75-4.25],

for R5/X4 samples vs. median 1.0 amino acid mixture/sequence, IQR[1.0-2.0], for R5 samples.

Results of the two phenotypic co-receptor assays were compared against the V3 sequence based genotypic Support Vector Machine (SVM) predictions for the 74 patient-derived samples (19). The Trofile assay and SVM genotypic predictor resulted in concordance for 64 (86.5%) samples, where the TRT assay and SVM resulted in concordance for 59 (79.7%) samples (Figure 4.4). Both the SVM and the Trofile co-receptor assay had a higher prevalence of classifying samples as R5 than the TRT assay.

4.4 Discussion

This study compares two well-characterized recombinant phenotypic co-receptor assays on a set of clinically derived plasma samples selected to contain roughly equal numbers of R5 and X4-containing HIV-1 isolates. The overall concordance of 85.1% between the assays indicates that although both are in good agreement with each other, there remains considerable disagreement. Discordant samples appear to show no significant difference between viral loads or CD4 counts, although concordant R5/R5 samples had significantly higher CD4 counts when compared to samples identified by any assay as harboring X4 variants, and concordant X4/X4 samples had much higher X4 signals as measured by RLU (Trofile assay) or OD (TRT assay). The SVM and Trofile co-receptor assay had higher concordance and a greater prevalence of classifying samples as R5 than the TRT assay. This may be a result of the SVM being trained on sequence data phenotyped by the Trofile co-receptor assay. The small number of discordances (N=11) makes the determination of associations in observed clinical or assay parameters between the two discordant groups difficult. A more thorough analysis of discordances would require a larger sample size.

Some potentially relevant differences between the two assays may account for many of the observed differences in co-receptor phenotype observed here. These differences include the size of the viral RNA fragment excised and amplified, the cells used to express the co-receptors, and the reporter gene. While the Trofile HIV assay uses the entire gp160 gene fragment, approximately 2500 base pairs, the TRT assay uses a 900-bp-long viral fragment which spans the V1-V3 region, including 150-bp-long

extensions on each side to allow for homologous recombination during transfection [17]. One hypothesis that may require further investigation is whether amplification of a larger DNA fragment leads to a greater loss of minority species during PCR amplification. Other potential reasons for discordance between assay results have been reviewed by Coakley, et al. [1] including the interaction between the viral vector gag proteins and the patient-derived-envelope proteins, which may affect pseudotyping efficiency and the usage of reporter cell lines. The use of different target cell lines could be another possible source of discrepancy between the assays, as the tropism assays made use of human malignant glioma cell lines (U87, U373) engineered to express CD4 and various chemokine co-receptors on their surface. Previous studies have shown that co-receptor utilization may be influenced by the level of receptor expression and the ratio of the receptors relative to each other [21]. However, results from at least one study suggest that co-receptor levels do not impact the ability to measure co-receptor usage [22].
Clinical samples contain a heterogeneous viral population. Studies into the detection of minority species within inter-individual viral populations have shown that the limit of detection is in the range of 10-20% for the phenotypic co-receptor assays [1], while the limit of detection of conventional genotypic assays (consisting of RT-PCR followed by population ('bulk') sequencing) is also approximately 25%-30%. This is largely because low level signals cannot be detected by conventional automated DNA sequencing, a factor that limits both the phenotypic co-receptor assays and genotypic predictors (SVM).

Developing a reliable genotypic predictor of co-receptor usage requires a well-validated phenotypic co-receptor assay to use as a reference. Although the "11/25" rule

for determining tropism, which categorizes variants as syncytium-inducing (SI) if they display positively charged residues at codons 11 and/or 25 within the V3 loop, and non-SI if they do not, is reasonably specific, it lacks sensitivity (<60%) for predicting co-receptor usage [10]. SVM analysis of clinically derived sequences appears to correlate well with results of phenotypic co-receptor assays and has higher sensitivity and specificity when compared against other bioinformatics models of determining co-receptor usage from clonal data [11]. However, the inability to consistently detect minority species in clinically derived sequence data presents difficulties both in using genotypic predictors and in training these predictors with sequence data phenotyped using assays which are also unable to consistently detect minority species.

In summary, the two phenotypic co-receptor assays tested here generated largely concordant results. However, in cases where there are differences in assay results, it is not clear which of these assays is "correct". Discordances do not show significant trends that could be attributed to differences between the two phenotypic co-receptor assays or to patient characteristics. It is likely that the presence of heterogeneous viral populations in clinically derived isolates poses difficulties of detection for both phenotypic and genotypic/bioinformatic approaches of determining co-receptor use. For the purposes of determining co-receptor usage in a clinical setting, an understanding of the limitations of the test used is essential. Although genotypic predictive models have the potential to reduce the cost and time of determining co-receptor usage of patient-derived isolates, validation of genotypic methods against a verified reference would be required before they can be adopted for clinical use.

Table 4.1: Potentially relevant differences in phenotypic assays.

Parameter	Value	
	Trofile	TRT
Insert	Entire gp160, ~2500 bp RT-PCR product	~900bp RT-PCR product spanning V1-V3
Vector	pCXAS-envelope expression vector plus HIV genomic-Luc vector	pNL43-ΔV
Construction	Restriction enzymes	Recombination
Producer cells	293	293T
Target cells	U87-CD4-CCR5 and U87-CD4-CXCR4	U373-CD4-CCR5 and U373-CD4-CXCR4 with HIV-1 LTR-lacZ cassette
Reporter gene	Luciferase	
Detection Assay	RLU	OD using a colorimetric assay

Table 4.2: Summary of median baseline CD4 count and baseline viral load for samples categorized as either R5 or R5/X4 using the Trofile or "TRT" co-receptor usage assays.

TF - TRT[a]	Baseline no. of CD4 cells/mm^3 (IQR)[c]	Log (VL Base) (IQR)	TF RLU		TRT	
			R5 signal (log) (IQR)	X4 signal (log) (IQR)	R5 signal OD_{Sample}/OD_{Ctrl} (IQR)	X4 signal OD_{Sample}/OD_{Ctrl} (IQR)
R5 - R5 (N=42)	325 (217-457)	5.0 (4.5-5.4)	5.66 (5.3-5.9)	1.87 (1.8-1.9)	17.1 (15.3-18.6)	1.1 (1.1-1.2)
R5/X4 - R5/X4 (N=21)	100 (30-170)	5.1 (4.6-5.6)	5.27 (4.9-5.6)	4.52 (3.6-5.2)	18.4 (15.0-19.6)	11.2 (2.2-19.2)
R5 - R5/X4 (N=8)	165 (100-342)	5.5 (5.2-5.7)	5.49 (5.4-5.7)	1.88 (1.8-2.0)	19.3 (18.4-20.4)	2.0 (1.6-3.3)
R5/X4 - R5 (N=3)	100 (80-275)	5.0 (4.8-5.6)	5.69 (5.5-5.8)	3.89 (3.2-3.9)	14.4 (13.8-17.6)	1.1 (1.1-1.2)

[a] TF, Trofile HIV co-receptor tropism assay; TRT, TRT co-receptor assay.
[b] VL base, baseline viral load in $\log_{10}$ HIV RNA copies/ml.
[c] IQR, interquartile range

Table 4.3: Characteristics of samples with discordant results from the Trofile or TRT entry co-receptor assays

TF - TRT[a]	Sample Identifier	Baseline no. of CD4 cells/mm^3	Log(VL Base)[b]	TRT		TF		Amino Acid observed at codon:	
				R5 signal OD_{Sample}/OD_{Ctrl}	X4 signal OD_{Sample}/OD_{Ctrl}	RLU R5 signal (log)	RLU X4 signal (log)	11	25
R5/X4 - R5	1	450	4.6	14.4	1.1	5.31	3.89	G	E/G
	2	60	5	21.3	1.3	5.88	2.48	S	R
	3	100	5.9	13.2	1.1	5.69	3.94	G	R
R5 - R5/X4	4	150	5.5	20.0	1.6	5.71	1.86	R/S	R
	5	320	4.2	21.9	2.2	5.40	1.89	G	K
	6	110	5.9	18.4	1.5	5.57	2.19	G	A/E
	7	410	5.9	19.1	1.5	6.22	1.99	G	K
	8	70	5.3	16.6	20.5	5.76	1.87	G/S	D
	9	20	5.5	18.3	1.8	3.80	1.73	G	D
	10	530	5.1	21.4	6.5	5.31	1.83	G	E
	11	180	5.7	19.4	16.4	5.38	2.25	D/G	R

[a] TF, Trofile HIV co-receptor tropism assay; TRT, TRT co-receptor assay.
[b] VL base, baseline viral load $\log_{10}$ HIV RNA copies/ml.

Figure 4.1: Receiver Operating Characteristic Curve of Trofile HIV co-receptor assay vs. TRT co-receptor assay.

The grey line uses the Trofile co-receptor results as the reference, while varying the cut-off for the X4 OD_{Sample}/X4 OD_{Ctrl} data obtained from the TRT assay. The black line uses the TRT co-receptor results as the reference, while varying the cut-off for the X4 RLU data obtained from the Trofile assay.

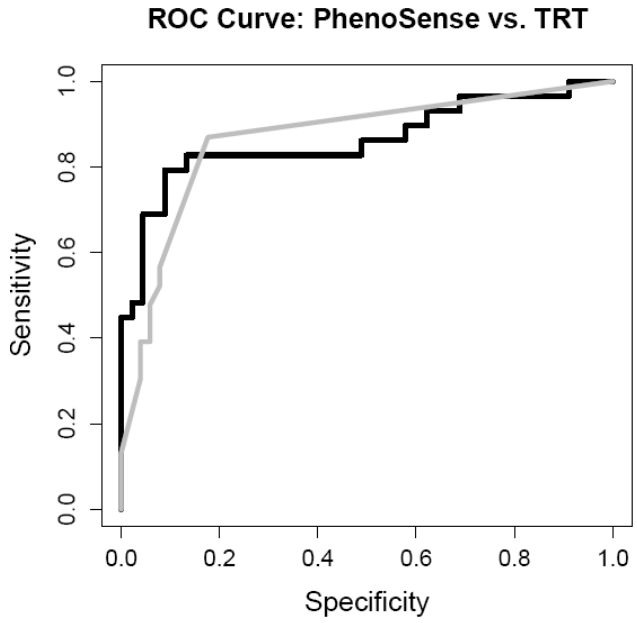


Figure 4.2: Support Vector Machine predicted probability of CXCR4 co-receptor usage

Each position on the x-axis indicates a sample from a single patient. Sequences containing amino acid mixtures were expanded into all possible permutations per sample and each permutation assigned a score based on the probability of harboring X4-containing virus (where 0= low probability and 1= high probability). Sample permutations were tested and scored individually, and the probability of being X4 is indicated on the y-axis. Each individual sample permutation is represented by a single symbol, generating a series of vertically-aligned symbols for each clinically-derived sample. Samples were grouped along the x-axis and labeled according to the results obtained from both phenotypic assays as "R5" or "R5/X4" and ordered by "**Trofile assay/TRT assay**". Solid triangles represent samples phenotyped as R5/X4 by both assays. Open squares represent samples phenotyped as R5/X4 by the Trofile assay and R5 by the TRT assay. Solid diamonds represent samples phenotyped as R5 by the Trofile assay and R5/X4 by the TRT assay. Open circles represent samples phenotyped as R5 by both assays. The dashed horizontal line represents the SVM cutoff, above which samples will be classified as being R5/X4

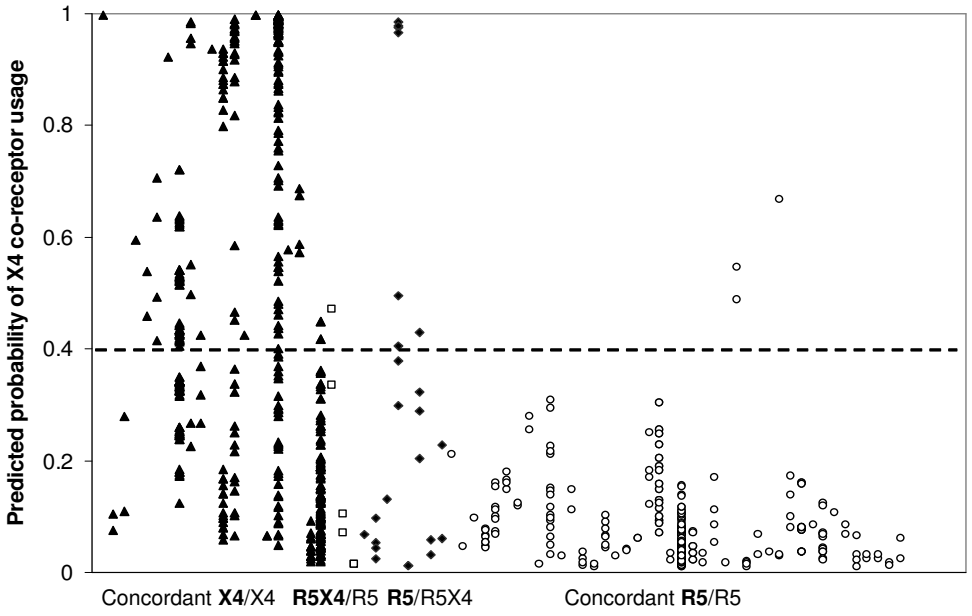


Figure 4.3: Boxplot summary of predicted probability of CXCR4 co-receptor usage

Mixtures were separated into all possible permutations per sample. Samples were grouped and labeled according to the results obtained from both phenotypic assays as "R5" or "R5/X4" and ordered by "**Trofile assay/TRT assay**". The maximum score observed for each isolate was used to create the boxplot. Boxes represent the interquartile range and the solid line inside the box represents the median of the maxima of the group of samples. Boxplot whiskers represent the range of the data in each category and outliers are represented as open circles.

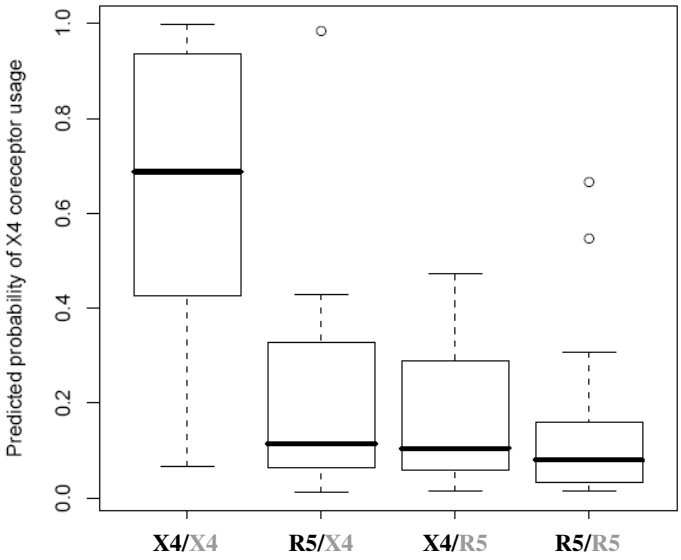
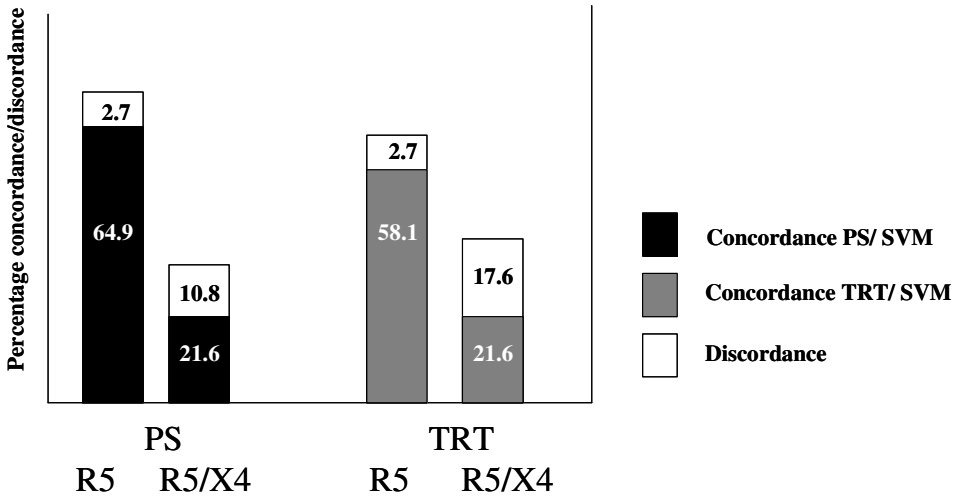


Figure 4: Concordances between phenotypic co-receptor assays and a SVM bioinformatic predictor.

Black bars represent the concordances for the Trofile entry co-receptor assay and support vector machine (SVM), and grey bars represent concordances for the "TRT" co-receptor assay and SVM. Percent discordances are represented with white bars. R5 and X4 concordances and discordances with the SVM model by the Trofile assay (TF) and the TRT assay (TRT).



4.5 References

1. Coakley E, Petropoulos CJ, Whitcomb JM. Assessing chemokine co-receptor usage in HIV. *Curr. Opin. Infect. Dis*. 2005;**18**:9-15.

2. Maeda K, Nakata H, Ogata H, Koh Y, Miyakawa T, Mitsuya H. The current status of, and challenges in, the development of CCR5 inhibitors as therapeutics for HIV-1 infection. *Curr opinion Pharmacol*. 2004*;* **4**:447-52.

3. Seibert C, Sakmar TP. Small-molecule antagonists of CCR5 and CXCR4: A promising new class of anti-HIV-1 drugs. *Curr Pharm Des.* 2004;**10**:2041-62.

4. Tanaka Y, Okuma K, Tanaka R, Kumakura S, Shimoyamada A, Hirose K, et. al. Development of Novel Orally Bioavailable CXCR4 Antagonists, KRH-3955 and KRH-3140: Binding Specificity, Pharmacokinetics and Anti-HIV-1 Activity *in vivo* and *in vitro*. [Abstarct]. *13th Conference on Retroviruses and Opportunistic Infections, Denver, Colorado, USA*. February 6-8, 2006; Abstract 49LB.

5. Kuhmann SE, Moore JP. The HIV-1 Phenotypic Variants: Deadly and Deadlier. *J. Viral Entry.* 2005*;***1**(1):4-16.

6. Fouchier RA, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema F, et. al. Phenotype-associated sequence variation of the third variable domain of the human immunodificiency virus type 1 gp120 molecule. *J Virol.* 1992;**66**:3183-3187.

7. 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 syncitium-inducing phenotype: analysis by single amino acid substitution. *J. Virol.* 1992;**66**:6777-6780.

8. Fouchier RA, Brouwer M, Broersen SM, Schuitemaker H. Simple determination of human immunodificiency virus type 1 syncytium-inducing V3 genotype by PCR. *J Clin Microbiol.* 1995;**33**:906-911.

9. 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**:51–62.

10. Jensen MA, van't Wout AB, Nickle D, Shriner D, He HX, McLaughlin S, et. al. Predicting HIV-1 co-receptor usage using sequence analysis. *AIDS Rev.* 2003;**5**:104–112.

11. Sing T, Harrigan R, Beerenwinkel N, Daumer M, Kaiser R, Low A, et. al. Immunologic markers improve genotypic prediction of HIV-1 coreceptor usage.

[Abstract]. *1st Int. Workshop on Targeting HIV Entry, Bethesda, Maryland, USA*. December 2-3, 2005; Abstract 30.

12. Hogg RS, Yip B, Chan KJ, Wood E, Craib KJ, O'Shaughnessy MV, et. al. Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy. *JAMA*. 2001;**286**:2568-77.

13. Brumme ZL, Dong WWY, Yip B, Wynhoven B, Hoffman NG, Swanstrom R, et. al. Clinical and immunological impact of HIV envelope V3 sequence variation after starting initial antiretroviral therapy. *AIDS*. 2004;**18**:F1-F9.

14. Hoffman NG, Seilier-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. *J Virol.* 2002;**76**:3852-3864.

15. Brumme ZL, Goodrich J, Mayer HB, Brumme CJ, Henrick BM, Wynhoven B, et. al.. Molecular and Clinical Epidemiology of CXCR4-Using HIV-1 in a Large Population of Antiretroviral-Naïve Individuals. *J Infect Dis.* 2005;**192**:466-74.

16. Huang W, Wrin T, Yap J, Fransen S, Beauchaine J, Reddy M, et. al.. A rapid, multi-functional HIV-1 entry assay for measuring drug susceptibility, co-receptor tropism and antibody neutralization. [Abstract]. *Program and Abstracts of the 41st Annual*

Interscience Conference on Antimicrobial Agents and Chemotherapy (ICAAC). December 16-19, 2001; Abstract 1968.

17. Trouplin V, Salvatori F, Cappello F, Obry V, Brelot A, Heveker N, et. al. Determination of Coreceptor Usage of Human Immunodeficiency Virus Type 1 from Patient Plasma Samples by Using a Recombinant Phenotypic Assay. *J Virol.* 2000;**75**:251-259.

18. Capon, Daniel, Petropoulos, Christos J. Compositions and methods for determining anti-viral drug susceptibility and resistance and anti-viral drug screening. *ViroLogic Inc.,USA*. November 17, 1998;US Patent 5,837,464.

19. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research.* 2004;**32**(5):1792-7.

20. Zhu T, Mo H, Wang N, Nam DS, Cao Y, Koup RA, et. al. Genotypic and phenotypic characterization of HIV-1 patients with primary infection. *Science.* 1993;**261**:1179-81.

21. Reeves JD, Gallo SA, Ahmad N, Miamidian JL, Harvey PE, Sharron M, et. al. Sensitivity of HIV-1 to entry inhibitors correlates with envelope/coreceptor affinity, receptor density, and fusion kinetics. *Proc Natl Acad Sci USA.* 2002;**99**:16249-16254

22. Princen K, Hatse S, Vermeire K, Bridger GJ, Skerlj RT, De Clercq E, et. al. The antiviral activity of the CXCR4 antagonist AMD 3100 is independent of the cytokine-induced CXCR4/HIV coreceptor expression level. *AIDS Res Hum Retroviruses.* 2003;**19**:1135-1139.

CHAPTER 5

CD4-DEPENDENT CHARACTERISTICS OF CO-RECEPTOR USE AND HIV-1 V3 SEQUENCE IN A LARGE POPULATION OF THERAPY-NAÏVE INDIVIDUALS[4]

5.1 INTRODUCTION

Human Immunodeficiency Virus type 1 (HIV-1) uses host cell membrane chemokine receptors, in combination with CD4, to gain entry into host cells [1-3]. The most important of these are CCR5 [1,2,4] and CXCR4 [5]. HIV entry via one or both of these co-receptors is generally associated with two different viral phenotypes identified shortly after the development of HIV culture techniques; namely the ability to form syncytia in the MT-2 cell line, leading to the nomenclature of syncytium inducing (SI) and non-syncytium inducing (NSI) [6,7]. While NSI variants are generally detectable over the entire course of HIV infection, SI variants emerge in approximately 40-50% of subtype B infected persons [6-8]. In addition, dual-tropic variants capable of using both major co-receptors may also arise over the course of disease [9]. The emergence of SI variants is associated with a rapid decline in $CD4^+$ T-lymphocyte counts, accelerated disease progression and reduced survival time in untreated individuals [10-13], but is not necessary for $CD4^+$ cell decline or disease progression to occur. The factors mediating

[4] A version of this chapter is in press as:
Low AJ, Marchant D, Brumme CJ, Brumme ZL, Dong W, Hogg RS, et. al. CD4-Dependent Characteristics of Co-receptor use and HIV-1 V3 Sequence in a Large Population of Therapy Naïve Individuals. *AIDS Res and Hum. Retro.* (In Press).

the NSI to SI phenotype "switch" over the natural course of HIV infection are not fully understood, and it is not clear whether it is a cause and/or a consequence of progressive immune dysfunction [14]. As MT-2 cells express CXCR4 but not CCR5, “SI” tropism as determined by the MT-2 assay is an indirect measurement of ability to use the CXCR4 co-receptor; for this reason , “SI” and “NSI” designations are considered largely equivalent to “X4” (CXCR4-using) and “R5” (CCR5-using) phenotypes, respectively, although there are exceptions [15]. Recently, a renewed focus on co-receptor use has emerged, in part, as a result of development of new classes of antiretroviral agents which specifically inhibit HIV binding to either CCR5 or CXCR4. A number of different phenotypic co-receptor usage assays have been developed, generally based on the development of cell lines engineered to express either CCR5 or CXCR4 [16,17].

The third hypervariable region of the HIV envelope glycoprotein gp120 (“V3 loop”) is known to be a key (but not exclusive) determinant of HIV co-receptor usage [18]. Genetic determinants include the presence of positively charged amino acids at codons 11 and/or 25 of the V3 (termed the “11/25” genotype [19]), the presence of mutations within the region of codons 6 through 8, indicative of the loss of an N-linked glycosylation site at position 6 [20], and increased V3 loop heterogeneity [21]. Currently, a variety of bioinformatics methods are being used to predict co-receptor usage by examining specific amino acid changes the V3; these include Support Vector Machines [22,23] Neural Networks [24] and Position Specific Scoring Matrices [25].

Here we examine associations between co-receptor phenotype using the Monogram Biosciences Trofile Tropism Assay [17] raw phenotypic assay readout data, and genotypic variation in the HIV V3 loop in a large cohort previously characterized for both V3 sequence [26] and phenotypic co-receptor usage [27]. Furthermore we investigated how these relationships varied with host immune status (as measured by CD4 count).

5.2 METHODS

Study subjects: the British Columbia HOMER cohort

In British Columbia, Canada, antiretrovirals are provided free of charge through a centralized drug treatment program according to guidelines established by the B.C. Therapeutic Guidelines Committee [28]. Routine clinical monitoring of patients takes place at approximately 3-month intervals at which time plasma viral load testing (Roche Amplicor Assay) and CD4 counts are performed. The HAART Observational Medical Evaluation and Research (HOMER) cohort includes all HIV-positive, antiretroviral naïve adults who started triple antiretroviral therapy (consisting of two nucleoside reverse transcriptase inhibitors [NRTIs] and either a protease inhibitor [PI] or a non-nucleoside reverse transcriptase inhibitor [NNRTI]) through the B.C. Drug Treatment Program between August 1996 and September 1999 (N=1188). This cohort has been the focus of a number of population-based studies and has been described in detail previously [29]. The present study represents a cross-sectional analysis of the HOMER cohort at baseline (prior to initiation of HAART). Reported plasma viral loads, CD4 counts as well as genotypic and phenotypic parameters represent the latest pre-therapy measurements collected in the 180 days prior to therapy initiation.

Determination of baseline HIV co-receptor phenotype

The "Trofile" co-receptor entry assay, performed by Monogram Biosciences had been previously used to assess co-receptor phenotype in the latest pre-therapy plasma sample for each subject [27]. Briefly, a 2.5-kb RT-PCR product spanning the entire

gp160 is digested, purified and ligated into an *E. Coli* expression vector, and gene libraries are constructed. A replication-defective retroviral vector (pHIVluc) containing a luciferase expression cassette inserted within the *env* gene is then used to co-transfect human embryonic kidney cell cultures with the sample plasmid DNA. Recombinant viruses are harvested after 48 hours and assessed for their ability to infect U87-CD4 cells expressing CCR5 or CXCR4, as determined by a luciferase read-out measured in Relative Light Units (RLU) on each of the two cell lines. The Trofile assay classifies isolates as CCR5-using, CXCR4-using or "DM" (indicating dual and/or mixed-tropic virus) based on confirmation of decreased RLU values upon addition of specific co-receptor inhibitors. The RLU values reported from the Trofile assay are normalized to a well characterized, dual-tropic reference strain (92HT594). Specifically, RLUs from 92HT594-infected $CXCR4^+/CD4^+$/U87 or $CCR5^+/CD4^+$/U87 cells were adjusted to 1,000,000 counts to give two independent scaling factors (i.e., to normalize counts on $CXCR4^+$ and $CCR5^+$ cells, respectively) [17]. Phenotypic data from a total of 977 isolates were available, with 799, 177 and one characterized as R5, R5/X4 (DM) and X4 only, respectively, with a higher prevalence of X4 virus being detected in individuals with lower CD4 cell counts [27]. In this manuscript, isolates phenotyped as "DM" or "X4" were combined and designated as "DM" for the remainder of analyses while CCR5-using virus is referred to as R5. This approach was used because only a single sample was phenotyped as being purely X4.

Determination of corresponding baseline envelope V3 sequence

Aliquots of the same baseline plasma samples were used to determine bulk "population" HIV V3 envelope sequence as described previously [26]. Matched baseline co-receptor phenotypes and V3 genotypes were available for 953 of 1188 (80.2%) HOMER subjects. In the case where nucleotide mixtures were present in V3, sequences were translated into all possible amino acid permutations, resulting in a total of 1388 unique V3 amino acid sequences in the entire dataset. Alignment of this dataset to a standard consensus B V3 sequence was made using the multiple sequence alignment software tool, MUSCLE [30], followed by visual inspection. Sequences displaying positively charged amino acids at codons 11 and/or 25 of the V3 loop, associated with an HIV syncytium-inducing (SI) phenotype, were classified as having an "11/25 genotype". A subject was classified as harboring "11/25" isolates if any of its expanded sequence permutations were classified as "11/25" positive.

Investigation of relationship between V3 amino acid variation and co-receptor usage

To determine the degree of amino acid variability across the V3 loop, a consensus sequence of the aligned dataset was created using the Los Alamos Consensus Maker program, available at http://www.hiv.lanl.gov/content/hivdb/CONSENSUS/CONSENSUS_TOOL/ConsensusTools.html. Note that for this and the following analyses, all expanded V3 sequences from a given patient isolate were assigned the co-receptor phenotype of the sample (for example, in a subject phenotyped as DM, whose V3 sequence was expanded into 4

possible permutations based on the presence of mixtures at two codons, all 4 resulting individual sequence permutations were classified as "DM". A codon-by-codon analysis was undertaken in which 2x2 contingency tables were constructed using co-receptor usage (R5 vs. DM) and amino acid (consensus vs. non consensus) as categories. Statistical significance was evaluated using the chi-squared test, with corrections made for multiple comparisons (Bonferroni method).

Investigation of relationship between V3 charge and co-receptor usage

Local charges were determined for all amino acid sequences using an exponential, five residue, sliding window. The charge of the residue at the midpoint of the sliding window was calculated, based on Coulomb's Law, by adding the charge of the centre residue to the two nearest residues, weighted by 70% plus the two distal charges, weighted by 10%. Charge values of -1, 0 or +1 were used for all amino acid residues depending on the residue specific charge at a biological pH of 7.5. Statistical significance was calculated using the Wilcoxon rank sum test.

5.3 RESULTS

In vitro replication signal strength -RLUs- in CCR5 and CXCR4 target cells

Luciferase is produced in the target cells upon a successful single round of infection and is measured upon cell lysis in Relative Light Units (RLU). CCR5 RLU in R5-tropic samples are roughly log-normally distributed (Figure 5.1). The range of CXCR4 RLU in DM samples was much greater and appears bimodal, with a trough at approximately 3.8log CXCR4 RLU. In order to investigate the relationship between CD4 count and output of the Monogram Trofile Assay, median CCR5 and CXCR4 RLU values from N= 977 clinical isolates phenotyped as R5 or DM were compared with CD4 count (Figure 5.2).

CCR5 RLU signal strength was robust in both R5 as well as DM-phenotyped isolates (median 5.6 log, inter-quartile range (IQR)=[5.3-5.9] vs. 5.6 log, IQR=[5.2-5.9] RLU in R5 vs. DM subjects, respectively). In contrast, median CXCR4 RLU signal in DM samples was much lower (43-fold) than the corresponding median CCR5 RLU in these samples (median 3.9 log, IQR=[3.1-4.5], N = 178). As expected, the median CXCR4 RLU signal was consistently low (median 73 RLU) in samples designated R5, presumably corresponding to assay background levels.

In univariate log-linear regression analyses of R5 samples, CCR5 RLU in R5 isolates did not vary with pVL (p=0.13), and showed a significant, but small increase with decreasing CD4 (β(slope)=$-4.3\text{x}10^{-4}$ $_{\text{log RLU/log pVL,}}$ [95% Confidence Interval (CI) -

$8x10^{-4}$:$4x10^{-5}$], p<<0.001; N=799). However, in DM samples, CCR5 RLU decreased with decreasing pVL ($\beta=0.27_{\text{log RLU/log pVL,}}$ [CI -0.12:0.67], p<0.05; N=178), but did not vary significantly with CD4 (p=0.17), while CXCR4 RLU increased with decreasing pVL ($\beta=-0.49_{\text{log RLU/ 100CD4}}$, [-1.03:0.05], p<0.001) and did not significantly vary with CD4 (p=0.82). In multivariate log-linear regression analyses, CCR5 RLU in R5 isolates increased with decreasing baseline CD4 count ($\beta=-0.04_{\text{log RLU/100CD4}}$, [95%CI -0.06:-0.02], p<0.001), but did not vary with baseline plasma viral load (pVL) (p=0.74). In contrast, the CCR5 RLU in DM samples did not increase, but instead tended to decrease with decreasing CD4 count ($\beta=0.04_{\text{log RLU/100CD4}}$, [95%CI 0.00:0.10], p=0.067) and decreasing pVL ($\beta=0.30_{\text{log RLU/log pVL}}$, [95%CI -0.29:0.84], p=0.001), while CXCR4 RLU increased with decreasing pVL (β=-0.50, [95%CI -0.78:-0.25], p=0.0001) but did not vary with CD4 (p=0.78).

Associations of Population Based V3 Sequence Patterns and Co-receptor Usage

The V3 sequence variability in R5 and DM samples versus consensus was determined in sequences generated from R5 vs. DM phenotyped isolates, respectively (Figure 5.3A). V3 sequence variability of DM samples was generally greater than that of R5 samples, with significant differences in variability observed at positions 11, 13, 25, 32 and 34 (p<<0.001). In addition, V3 loop codons 6 through 8 (possibly linked to N-linked glycosylation [31]) were conserved in almost 100% of R5 samples, but showed 20-30% variability in the DM samples (p<<0.001). When the criteria was narrowed to mutations away from consensus at positions 6-8, a total of 20 of 168 (11.9%) DM

samples had a mutation in these residues in comparison to 5 of 785 (0.6%) R5 samples ($p<<0.001$). Due to the strong association of a higher net and local charge with the DM phenotype, the mean local charge was analyzed between samples phenotyped as R5 and as DM for all CD4 strata, revealing that the net local charge in the V3 was similar across R5 and the DM samples, except in the region of codons 23 to 26 (Figure 5.3B). These codons tended to have higher charge in samples harbouring X4 variants than R5 samples ($p<<0.001$), an observation most likely explained by the substantial proportion of DM samples exhibiting positively-charged amino acids around, but not limited to, codon 25 of the V3 loop. A positively charged amino acid (Lysine or Arginine) at position 25 was observed in 31 (18.4%) X4 samples and 52 (6.6%) R5 samples ($p<<0.001$). Although the prevalence of arginine residues at position 25 was greater than lysine residues in DM samples (20 vs 14), the prevalence of arginine was less than lysine at position 25 in R5 samples (25 vs 32), resulting in a greater likelihood of a sample being DM if an arginine was detected at position 25 (44.4%) than a lysine (30.4%; $p<0.01$). Codon 11 also tended to have a more positive charge in samples phenotyped as DM than R5 ($p<<0.001$), but the magnitude of the difference was less pronounced. A total of 29 (17.2%) X4 samples had a positively charged amino acid at position 11, compared to 2 (0.2%) R5 samples ($p<<0.001$). In all cases, the positively charged amino acid was identified as arginine. Similarly, the detection of nucleic acid mixtures in the V3 region, presumably reflecting the co-existence of multiple viral strains at relatively high prevalence, was associated with an increased likelihood of detection of CXCR4-using virus in the Trofile assay. A total of 414 of the 785 R5 samples (53 %) contained at least one amino acid mixture in the V3 loop, compared with 129 of 168 (77 %) DM samples ($p<0.001$). The median

number of these ambiguous codons was 1.0 in the R5 samples, with an interquartile range (IQR) of 0-2, compared with a median of 2.0 in the DM samples (IQR=1.0–4.0, $p<0.001$). The distribution of mutations associated with X4-capable virus was compared in DM samples between the two CXCR4 RLU strata indicated in the observed bimodal distribution in Figure 1 (CXCR4 RLU<3.8log and CXCR4 RLU≥3.8log). Samples in the lower strata were less like to have positively charged amino acids at positions 25 (4.0% vs 29.0%, N=31; $p<<0.001$), as well slightly less likely to have positively charged amino acids at position 11 as (16.7% vs 20.4%, N=29; $p=0.2$) or a non-consensus amino acid between positions 6-8 (13.7% vs 16.1%, N=20; $p=0.4$), than those in the higher strata.

The HIV Genotype-Phenotype Relationship is Dependent on Host Immune Status

We examined changes in the positive predictive value (PPV) of genotypic determinants of the DM phenotype (positively charged amino acids at codons 11 and/or 25, lack of conservation in codons 6-8, and V3 loop heterogeneity) as a function of 7 CD4 strata (defined as <25, 25-49, 50-99, 100-199, 200-349, 350-499, 500+). The PPV is defined here as the proportion of samples predicted to be DM, based on genotypic determinants, which actually are DM according to the phenotypic assay. In addition, we also examined the change in the proportion of samples with one of the above genotypic determinants which are not actually DM as a function of CD4 strata for comparison. As expected, almost all samples (N=29 of 31, 93.5 %) with a positively charged amino acid residue at position 11 were identified as DM, regardless of the baseline CD4 cell count ($p<<0.01$) (Figure 5.4A). However, the PPV of the presence of a positively charged

amino acid at codon 25 increased significantly with decreasing CD4 count ($p<<0.001$; Cochran-Armitage Trend Test) (Figure 5.4B). Not only was the detection of this polymorphism (N = 83 phenotyped samples) strongly dependent upon the individual's CD4 cell count, but also almost all samples (N=12 of 14, 85.7%) with a positive charge at codon 25 of the V3 loop and a CD4 count below 100 cells/mm^3 were phenotyped as DM, compared to 19 of 69 (27.5%) with this mutation where the CD4 count was above 100 cells/mm^3 ($p<0.001$). This provides evidence of a CD4-dependent genotype to phenotype relationship and is not merely a reflection of DM at low CD4.

The PPV for the DM phenotype from the presence of one or more non-consensus amino acids within the potentially N-linked glycosylation region of codons 6 through 8 of the V3 loop (Figure 5.4C) showed a trend of increasing with decreasing CD4 count ($p=0.2$). A total of 10 of 11 (90.9%) samples from individuals with CD4 cell counts below 100 cells/mm^3 with mutations at these positions were identified as DM, compared with 10 of 14 samples (71.4%), with the same set of mutations, in the CD4 stratum greater than 100 ($p=0.48$).

A significantly larger proportion of sequences containing amino acid mixtures were observed among DM samples with CD4 counts below 100 cells/mm^3 (N= 68 of 81, 83.9%) compared to those with CD4 counts >100 cells/mm^3 (N=61 of 87 (70%), $p=0.05$). The PPV for the DM phenotype from the presence of one or more amino acid mixtures in the V3 loop increased significantly ($p<<0.001$) with decreasing CD4 (Figure 5.4D).

V3 Loop Genotypic also Predicts Magnitude of CXCR4 RLU Signal in In Vitro Phenotype Test

Genotypic predictors of X4 coreceptor usage were also associated with raw assay output values (RLU) among samples phenotyped as DM (Figure 5). Those samples with the 11/25 genotype exhibited significantly higher median CXCR4 RLU (median [IQR] =4.5 log [3.9-5.2]; N=57), than samples with the non-11/25 genotype (3.5 log [2.8-4.2]; N=111 $p<<0.001$). Interestingly, when codons 11 and 25 were analyzed independently, a positive charge at codon 25 alone was found to be associated with a higher median CXCR4 RLU (4.7 log [4.4-5.1]; N=28, $p<<0.001$), than a positive charge at codon 11 alone (4.2 log [3.5-5.2]; N=26, $p=0.004$). This effect appeared to be additive, as the CXCR4 RLU of samples with positively charged residues at both positions 11 and 25 was even greater (5.9 log [5.6-5.8]; N=3, $p<<0.001$).

Samples phenotyped as DM containing a mutation away from consensus within codons 6 through 8 of the V3 loop, potentially indicative of a loss of an N-linked glycosylation site, had a median CXCR4 RLU of 4.5 log [4.1-5.2] (N=20), while samples phenotyped as DM without mutations in this region had a median CXCR4 RLU of 3.9 log [3.1-4.4] (N=148, $p=0.02$). CXCR4 RLU in samples containing any V3 loop mixtures were higher (4.1 log [3.3-4.5]; N=129) than samples without mixtures (3.5[2.7-4.8] log; N=39, $p=0.08$).

5.4 DISCUSSION

We have previously investigated the clinical predictors of HIV co-receptor tropism [27], in this cohort. This manuscript examines the Monogram Trofile assay output (co-receptor phenotype and RLU on the CXCR4 and CCR5 cell lines). Here we examine how the raw RLU values vary with patient and virus characteristics. Previously, the standard Monogram protease (PR)/reverse transcriptase (RT) resistance assay [32] has been used to determine viral replicative capacity using an analogous procedure [33]. The RLU's represent two distinct envelope replicative capacity assays (Env-RC_{CCR5} and Env-RC_{CXCR4}). The range of the Env-RC_{CXCR4} is much broader than the range of the Env-RC_{CCR5} and both are much greater than the PR/RT-RC assay, suggesting that envelope variation is more critical to viral replicative capacity than variance in the protease or reverse transcriptase. In addition the observation of two distinct peaks in the CXCR4-using samples may indicate the existence of two subpopulations within the DM population. The observation that there is a relationship between Env-RC and CD4 cell count also suggests a potential relationship between RLU and infectivity using a single cycle entry assay, indicating that it may be reflect a biologically significant parameter.

Consistent with results from other cross-sectional studies [34-37] we observed a strong correlation between the presence of CXCR4-using HIV and CD4 cell count. What has not been previously reported is the significant, near 8-fold increase in CCR5 RLU with decreasing CD4 count in samples phenotyped as being R5 only; this remained significant in multivariate analysis, adjusted for plasma viral load. Although, this observation would be expected as increased fitness of R5 virus has been known to occur

in late stages of disease progression [38-40]. This may be an indication of increasing infectivity of the R5-tropic virus in patients who experience progressive immune dysfunction, but do not show any sign of a circulating X4-tropic strain and is consistent with the suggestion that viral replicative capacity increases with decreasing immune surveillance [37] and that disease progression occurs even in the absence of an X4 switch. What was unexpected is the opposing decrease in CCR5 RLU with CD4 count in samples with a detectable X4-using virus, presumably as a result of a decrease in the proportion of R5-only virus present.

Changes in immune function as measured by CD4 count not only affect X4 prevalence [3], but also affect the HIV V3 genotype. This variability in immune pressure may therefore alter the genotype to phenotype relationship over the course of disease progression, influencing the ability to predict co-receptor phenotype from V3 genotype. The presence of positively charged amino acids at V3 positions 11 and 25 are well known to be associated with CXCR4 co-receptor phenotype [19], and have been shown to be correlated with increased risk of clinical progression in the absence of antiretroviral therapy [18,19,41] as well as poorer response to HAART [27]. However, the CD4 dependence and previously discussed CXCR4 RLU and CCR5 RLU associations are less well characterized. Our observations suggest that positions 11 and 25, as well other potential codons within the V3 loop, including a loss of a glycosylation site at positions 6-8 and V3 heterogeneity, contribute independently to increases in magnitude of the median CXCR4 RLU as well as the absolute tropism category. We found specific V3 mutations (such as 11R and 13Y – data not shown) are consistently associated with DM

virus, regardless of CD4 count. However, the opposite is true for samples with mutations at codon 25, where almost all samples derived from patients with low CD4 count are DM, but for patients with high CD4 count the vast majority are not DM. One possible explanation is that the context of the gp-120, within which the V3 appears, may be influenced by host immune status, consistent with Figure 4. This is supported by the fact that 'swapping' V3 sequences between X4 and R5 strains usually, but not invariably changes in tropism [42, 43]. This independence of positions 11 and dependence of position 25 may indicate that the traditional '11/25 rule' may in fact be dissociated, and position 11 and 25 may be analyzed independently, depending upon CD4 count. Also, the predictive value for CXCR4-usage with arginine compared to lysine at position 25 is in contrast to other studies which have indicated that lysine is more predictive of X4-usage [44, 45]. One possible reason for this disparity may be the different study populations used or the fact that the other study was performed in clones instead of in patient derived, population-based samples. The CD4 dependence of positively charged residues at position 25 and corresponding CXCR4 RLU in DM samples may reflect a decrease in immune surveillance and increasing viral infectivity. Indeed, CXCR4-using virus had significantly greater CXCR4 RLU if their V3 loop genotype indicated positive amino acids at positions 11 and/or 25 (11/25 genotype), as compared to samples phenotyped as DM without the 11/25 genotype. This is consistent with the finding of previous studies, linking V3 genotype to viral ex-vivo fitness [21] using a position specific scoring matrix (PSSM) [25] approach.

The presence of the N-linked glycosylation site at position 6 of the V3 loop has been associated with high levels of resistance to antibody-mediated neutralization [20], while the absence of this V3 loop glycan has been found to increase the efficiency with which HIV is able to enter target cells via the chemokine co-receptors CCR5 and CXCR4 [20] and is associated with increased syncytium formation [46]. Our results are consistent with these data, indicating a much higher conservation of the N-linked glycosylation site in R5 isolates compared to DM samples, as well as a much greater tendency for CXCR4-using isolates to have mutations within this glycosylation site at lower CD4 strata. In addition to this, deviations from consensus within codons 6 through 8, potentially predictive of a loss of the N-linked glycan at this site, correlated with increased viral efficiency [20], were associated with greater CXCR4 RLU.

Furthermore, the number of amino acid mixtures observed in the V3 loop, indicative of increased envelope heterogeneity, and potentially indicative of increased viral diversity, is shown to be significantly associated with the presence of HIV samples capable of using the CXCR4 co-receptor. This association strongly suggests that many of the samples which are typically referred to as “dual and/or mixed tropic” are likely made of mixtures of both R5 and X4 viruses. However, systematic sequence and phenotype analyses of clones of these samples were not attempted in order to confirm this. In addition, increased V3 loop heterogeneity, a genotypic predictor for X4-using HIV samples, was also associated with increased CXCR4 RLU, consistent with data linking V3 heterogeneity in X4 isolates with improved ex vivo fitness [21].

It should be emphasised that the population-based sequencing methods and the phenotype assay are able to detect only the circulating viral species which make up a significant proportion of the total virus present. Limitations of this study include the fact that associations between HIV co-receptor usage and clinical parameters are based on cross-sectional data from a population of therapy-naïve individuals initiating their first antiretroviral treatment during the period 1996 to 1999, and may not be representative of the HIV-infected, antiretroviral naïve population in general, nor of individuals treated with HAART or co-receptor antagonists.

Another potential limitation is that observations based on RLU results obtained from the Trofile assay are limited most directly by the lack of understanding of all the factors contributing to expressed RLU. RLU are highly complicated and dependent on a variety of cellular expression mechanisms as well as viral envelope and envelope/expression vector protein interactions. In addition, the fact that the R5 signal in X4-capable (DM) samples is approximately 50 times greater (1.6 log) than the X4 signal in the same samples is of interest, and may be partially due to the fact that X4 variants are a minority species in the total virus population, although this likely does not explain this discrepancy completely. The cell lines in which these assays are performed are based on U87 CD4 cells in which the CXCR4 co-receptor is stably expressed [47], however, it is known that chemokine receptors can adopt alternate conformations [48], and/or be subject to alternate post-translational modifications [49]. It is also not clear whether co-receptor expression in these cells is entirely representative of co-receptor expression in vivo. The ability of X4 virus to infect CXCR4-expressing cells may be partly limited by

the interaction between the viral vector gag proteins and the patient-derived envelope proteins [50]. Furthermore, as the CCR5 and CXCR4 results are derived from infection of independent cell lines direct comparisons of RLU values between the two cell lines should be made with caution. Note though, that some CXCR4 signals as high as 10^6 RLU were observed, confirming a very large dynamic range of CXCR4 RLU signal can be detected in this assay system.

These data support the correlation between decreasing immune function, as measured by CD4 count, with genotypic changes, indicative of increased infectivity. Due to the cross-sectional nature of this study, we cannot comment on the causality of this relationship. However, our results clearly demonstrate the tightly linked association between the immune system and viral genotype as measured by CD4 count and HIV V3 loop mutations and RLU. This association may be especially important if co-receptor usage is inferred using genotypic predictors based on HIV V3 loop sequence (available online; http://ubik.microbiol.washington.edu/computing/pssm/ [25]; http://coreceptor.bioinf.mpi-sb.mpg.de/cgi-bin/coreceptor.pl [23]; http://genomiac2.ucsd.edu:8080/wetcat/v3.html [22]) to screen for the suitability of co-receptor antagonist therapy or other purposes, and illustrates the need to integrate clinical information such as CD4 count into the training and testing of future genotype based co-receptor predictors. In addition, the strong observed association between increasing CCR5 RLU for R5-tropic samples at decreasing CD4 and the converse, decreasing CCR5 RLU for DM samples suggests that clinical information regarding immune status cannot

be discounted when interpreting phenotype data, and importantly, understanding the association of HIV phenotype from V3 genotype.

Figure 5.1: Distribution of CCR5 and CXCR4 RLU for R5 and DM samples.

The grey line represents the percentage of the median CCR5 RLU for all samples characterized as using the CCR5 co-receptor exclusively (N=799), while the black line represents the percentage of the median CXCR4 RLU for all samples characterized as using the CXCR4 co-receptor (N=178). Bin sizes were increased by a factor of 1.7, starting at 0.01. Frequency of occurrence for each bin was normalized by the total N. Trend lines were created by manually.

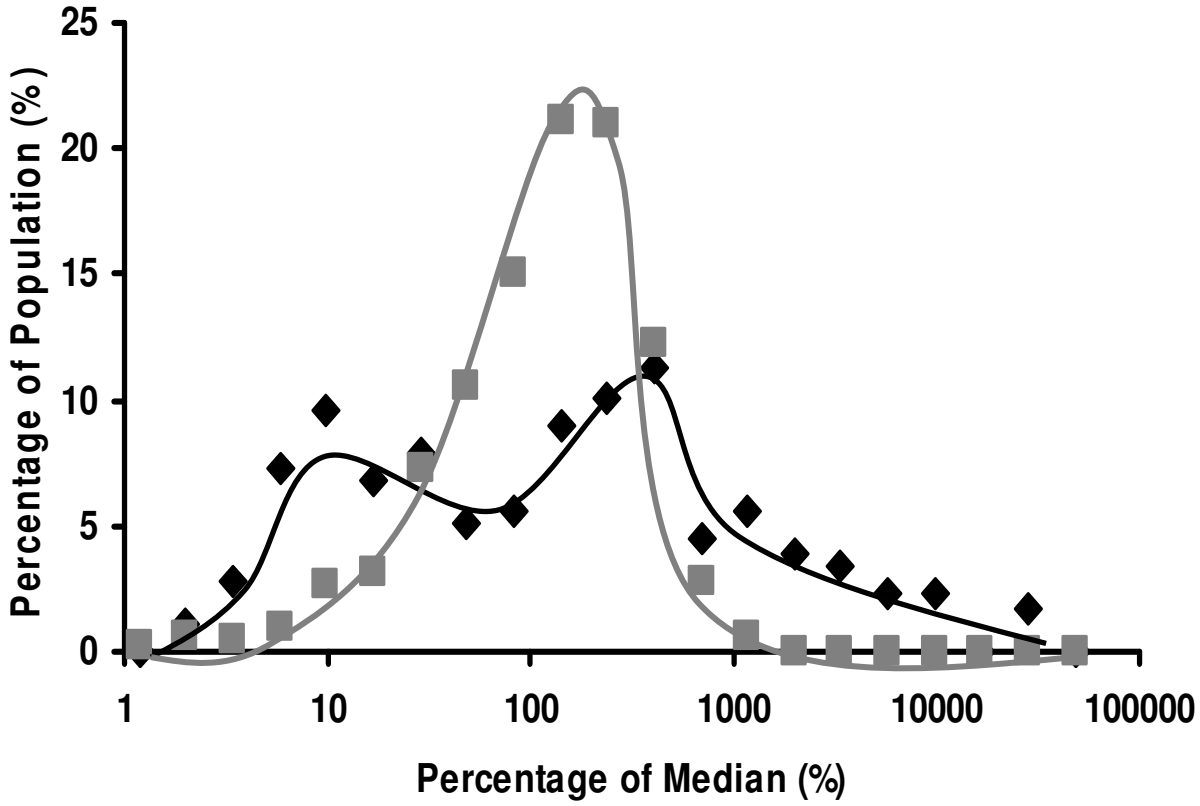


Figure 5.2: Relative Light Units (RLUs) from CD4/CCR5 and CD4/CXCR4 cells from samples identified as using the CCR5 or CXCR4 co-receptors.

The median and interquartile ranges of CCR5 and CXCR4 RLU values (log scale) from the Monogram Biosciences (Virologic) Trofile Tropism assay of samples from antiretroviral naïve individuals identified as using the CCR5 co-receptor ("R5"; N = 799) or the CXCR4 co-receptor with or without CCR5 co-receptor ("DM"; N = 178) are indicated as a function of the CD4 cell count. CD4 counts are grouped into 7 strata: < 25 (N = 79); 25 - 49 (N = 45); 50 - 99 (N = 82); 100 - 199 (N = 167); 200 - 349 (N = 259); 350 - 499 (N = 193); ≥ 500 (N = 152). Linear regressions were derived from raw, rather than stratified data. Interquartile ranges are depicted by the vertical error bars. Symbols used are: CCR5 RLU from R5 virus (◆); CCR5 RLU from DM virus (▲); CXCR-4 RLU from DM virus (■); and CXCR4 RLU from R5 virus (■).

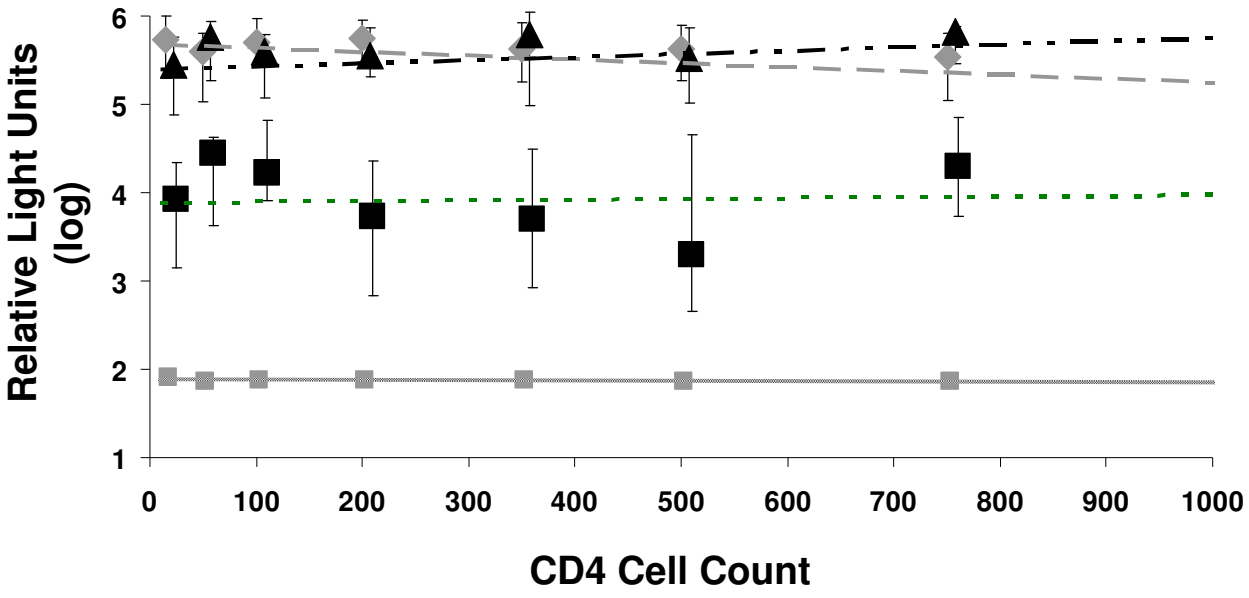


Figure 5.3: Comparison of the percentage of sequence conservation and local net charge in samples with R5 and DM phenotype.

(A) The level of sequence conservation at each codon of the HIV-1 V3 loop for R5 or DM samples where phenotype and genotype are available (N = 953). (B) A comparison of the inferred local amino acid charges of DM and R5 viruses derived from clinical isolates of the HOMER cohort where phenotype and genotype are available (N = 953).

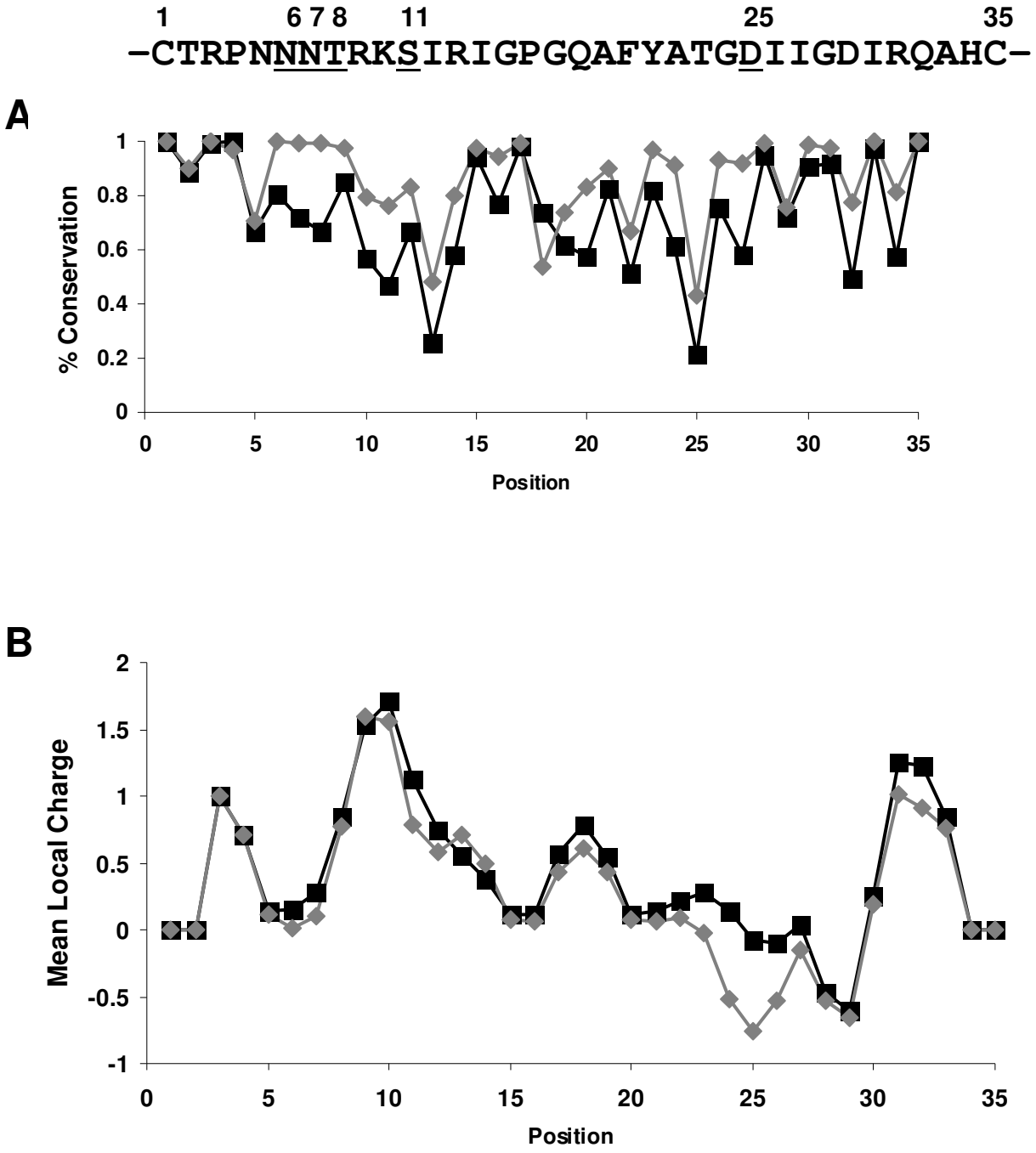


Figure 5.4: Proportions and prevalence of samples with specific patterns of V3 sequence.

In each panel, the proportion of samples with the mutation that are actually DM (left hand axis; black diamonds), as well as the proportion of samples of samples without the mutation that are DM (left hand axis; open triangles).in each CD4 group, as in Figure 2, are shown. The number of samples with a specific mutation pattern are also shown (right hand axis; grey symbols). Data are indicated for those with a positively charged amino acid at codon 11 (A), a positive charged amino acid at codon 25 (B), any mutation within V3 loop codons 6 to 8 (C) or the presence of one or more amino acid mixtures in the V3 loop (D). Note that these categories were not necessarily mutually exclusive.

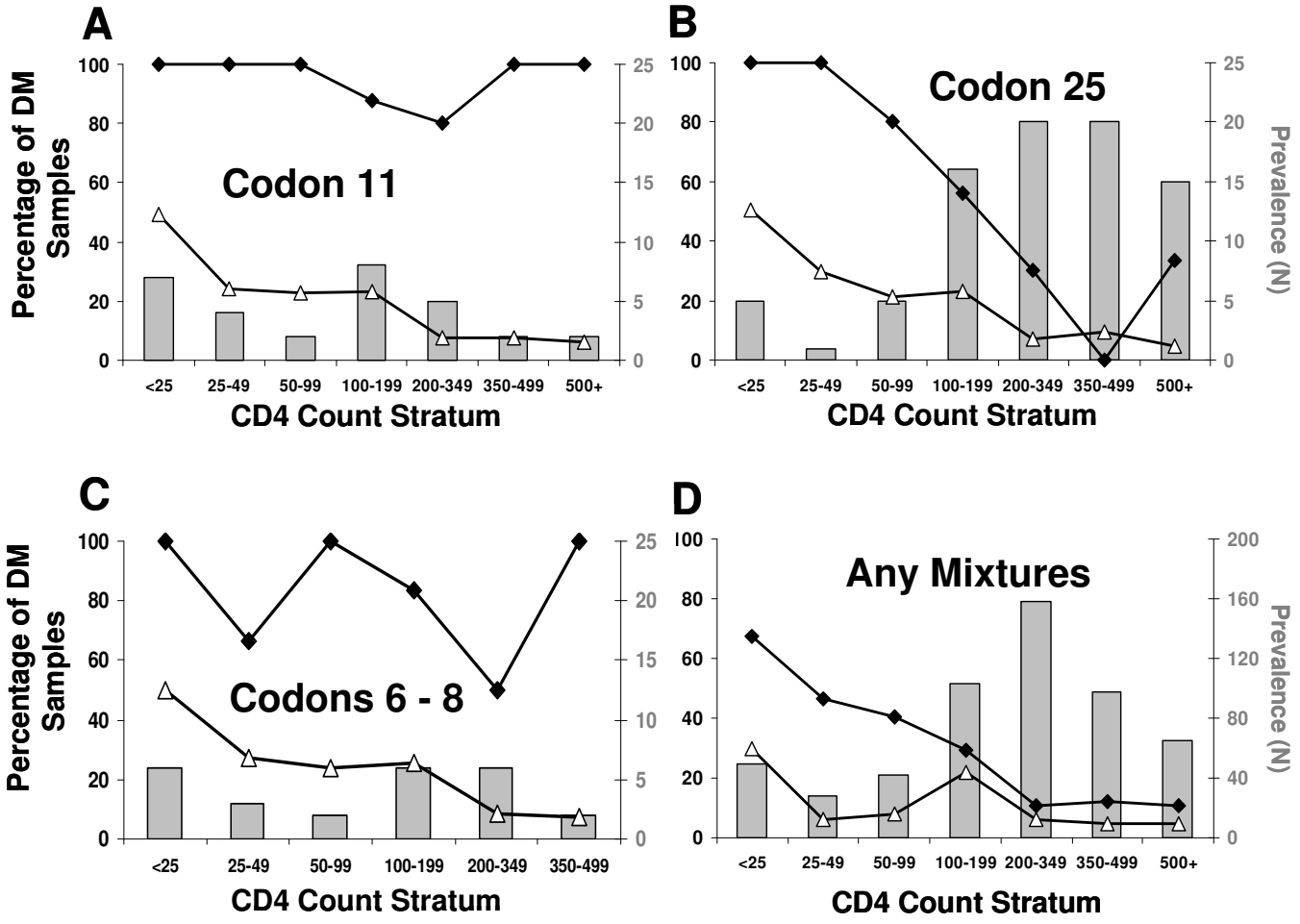
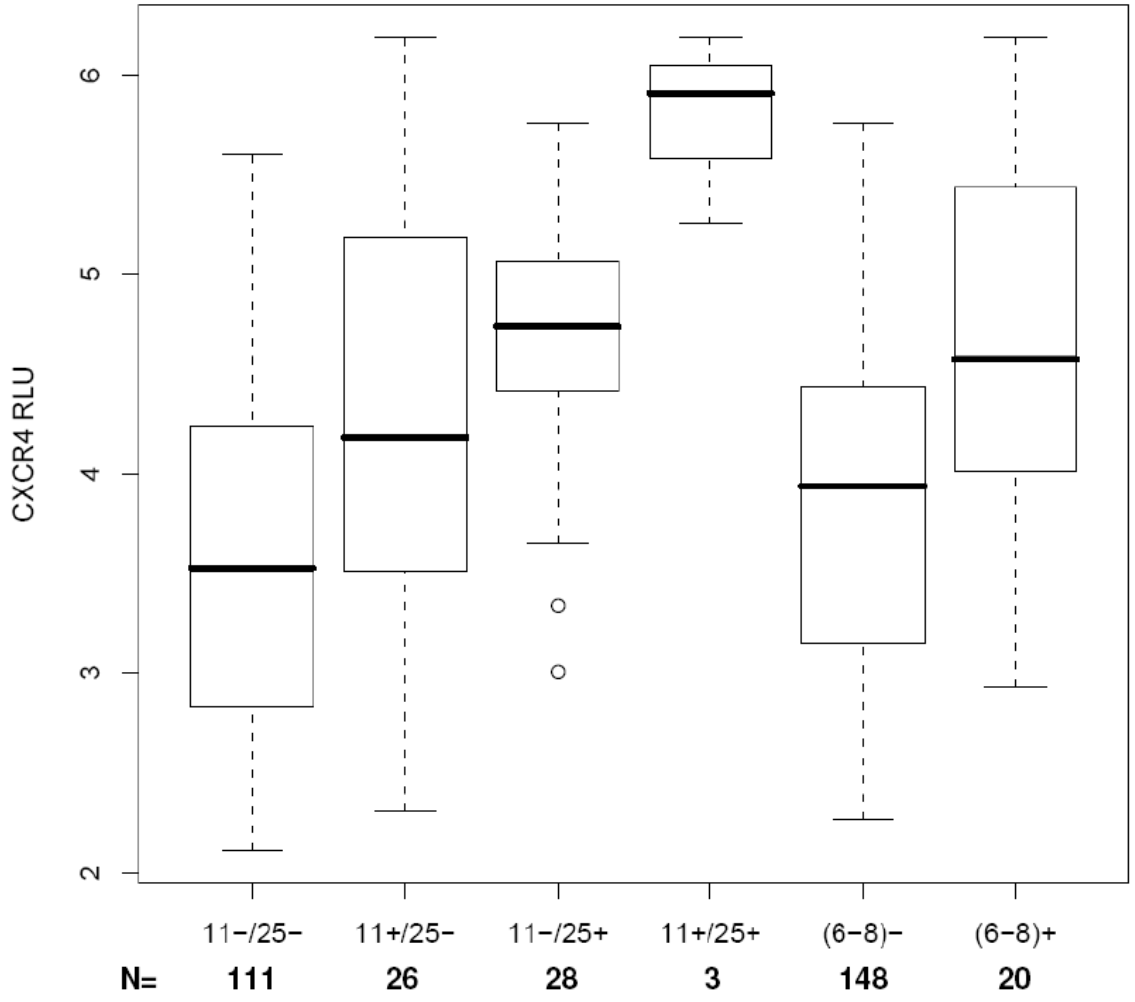


Figure 5.5: Associations of V3 loop predictors of X4-tropic HIV with CXCR4 RLU in samples phenotyped as DM.

Associations between CXCR4 RLU for samples phenotyped as DM and the absence and presence of genotypic predictors are shown. 11-/25-, 11+/25-, 11-/25+ and 11+/25+ represent the absence or presence, respectively, of positively charged amino acids at both positions 11 and 25. (6-8)- and (6-8)+ represent the absence or presence, respectively, of mutations in the region of codons 6 through 8 in the V3 loop, potentially indicative of a loss of N-linked glycosylation at position 6. Medians are represented by the solid lines. Inter-quartile ranges are represented by the box boundaries and minima and maxima are represented by the whiskers.



Acknowledgements

Ethical approval for this study was obtained from the institutional ethics board (Providence HealthCare/University of British Columbia).

This study was supported through an Rx&D research grant from the Canadian Institutes for Health Research (CIHR) and by funding from Pfizer, Inc. Z. Brumme is supported by doctoral research awards from the Michael Smith Foundation for Health Research (MSFHR) and the CIHR. R. Hogg is supported by the MSFHR through a Senior Scholar Award. Richard Harrigan has previously consulted for Monogram, Virco, and other companies in the pharmaceutical and diagnostic industries.

5.5 REFERENCES

1. Alkhatib G, Combadiere C, Broder CC, Feng Y, Kennedy PE, Murphy PM, et al. CC CKR5: A RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science.* 1996; **272**(5270):1955-1958.

2. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, et al. Identification of a major co-receptor for primary isolates of HIV-1. *Nature.* 1996; **381**(6584):661-666.

3. Berger EA, Doms RW, Fenyo EM, Korber BT, Littman DR, Moore JP, et al. A new classification for HIV-1. *Nature.* 1998; **391**(6664):240.

4. Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, et al. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature.* 1996; **381**(6584):667-673.

5. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science.* 1996; **272**(5263):872-877.

6. Schuitemaker H, Kootstra NA, de Goede RE, de Wolf F, Miedema F, Tersmette M. Monocytotropic human immunodeficiency virus type 1 (HIV-1) variants detectable in all stages of HIV-1 infection lack T-cell line tropism and syncytium-inducing ability in primary T-cell culture. *J Virol.* 1991; **65**(1):356-363.

7. Connor RI, Ho DD. Human immunodeficiency virus type 1 variants with increased replicative capacity develop during the asymptomatic stage before disease progression. *J Virol.* 1994; **68**(7):4400-4408.

8. Berger EA, Murphy PM,Farber JM. Chemokine receptors as HIV-1 coreceptors: Roles in viral entry, tropism, and disease. *Annu Rev Immunol.* 1999; **17**:657-700.

9. Glushakova S, Yi Y, Grivel JC, Singh A, Schols D, De Clercq E, et al. Preferential coreceptor utilization and cytopathicity by dual-tropic HIV-1 in human lymphoid tissue ex vivo. *J Clin Invest.* 1999; **104**(5):R7-R11.

10. Tersmette M, Gruters RA, de Wolf F, de Goede RE, Lange JM, Schellekens PT, et al. Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency syndrome: Studies on sequential HIV isolates. *J Virol.* 1989; **63**(5):2118-2125.

11. Schuitemaker H, Koot M, Kootstra NA, Dercksen MW, de Goede RE, van Steenwijk RP, et al. Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: Progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus population. *J Virol.* 1992; **66**(3):1354-1360.

12. Koot M, Keet IP, Vos AH, de Goede RE, Roos MT, Coutinho RA, et al. Prognostic value of HIV-1 syncytium-inducing phenotype for rate of CD4+ cell depletion and progression to AIDS. *Ann Intern Med.* 1993; **118**(9):681-688.

13. Richman DD, Bozzette SA. The impact of the syncytium-inducing phenotype of human immunodeficiency virus on disease progression. *J Infect Dis.* 1994; **169**(5):968-974.

14. Shaheen F, Collman RG. Co-receptor antagonists as HIV-1 entry inhibitors. *Curr Opin Infect Dis.* 2004; **17**(1):7-16.

15. Dong XN, Chen X, Chen Y, Ablimit A, Ye Z, Wu Y, et al. Short communication: HIV type 1 phenotype, tropism, and sequence patterns: Association and preference. *AIDS Res Hum Retroviruses.* 2005; **21**(3):234-238.

16. Trouplin V, Salvatori F, Cappello F, Obry V, Brelot A, Heveker N, et al. Determination of coreceptor usage of human immunodeficiency virus type 1 from patient plasma samples by using a recombinant phenotypic assay. *J Virol.* 2001; **75**(1):251-259.

17. Whitcomb JM, Huang W, Fransen S, Limoli K, Toma J, Wrin T, et. al. Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism. *Antimicrob Agents Chemother.* 2007;**51**(2):566-75

18. 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. *J Virol.* 1992; **66**(11):6777-6780.

19. Fouchier RA, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema F, et al. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J Virol.* 1992; **66**(5):3183-3187.

20. Malenbaum SE, Yang D, Cavacini L, Posner M, Robinson J, Cheng-Mayer C. The N-terminal V3 loop glycan modulates the interaction of clade A and B human immunodeficiency virus type 1 envelopes with CD4 and chemokine receptors. *J Virol.* 2000; **74**(23):11008-11016.

21. Troyer RM, Collins KR, Abraha A, Fraundorf E, Moore DM, Krizan RW, et al. Changes in human immunodeficiency virus type 1 fitness and genetic diversity during disease progression. *J Virol.* 2005; **79**(14):9006-9018.

22. Pillai S, Good B, Richman D,Corbeil J. A new perspective on V3 phenotype prediction. *AIDS Res Hum Retroviruses.* 2003; **19**(2):145-149.

23. T. Sing, R. Harrigan, N. Beerenwinkel, M. Daumer, R. Kaiser, A. Low, et al. Immunologic markers improve genotypic prediction of HIV-1 co-receptor usage [Abstract]. *Int. Workshop on Targeting HIV Entry* 2005; **1:**.

24. 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.

25. Jensen MA, van 't Wout AB. Predicting HIV-1 coreceptor usage with sequence analysis. *AIDS Rev.* 2003; **5**(2):104-112.

26. Brumme ZL, Dong WW, Yip B, Wynhoven B, Hoffman NG, Swanstrom R, et al. Clinical and immunological impact of HIV envelope V3 sequence variation after starting initial triple antiretroviral therapy. *AIDS.* 2004; **18**(4):F1-9.

27. Brumme ZL, Goodrich J, Mayer HB, Brumme CJ, Henrick BM, Wynhoven B, et al. Molecular and clinical epidemiology of CXCR4-using HIV-1 in a large population of antiretroviral-naive individuals. *J Infect Dis.* 2005; **192**(3):466-474.

28. Wood E, Montaner JS, Tyndall MW, Schechter MT, O'Shaughnessy MV,Hogg RS. Prevalence and correlates of untreated human immunodeficiency virus type 1 infection among persons who have died in the era of modern antiretroviral therapy. *J Infect Dis.* 2003; **188**(8):1164-1170.

29. Hogg RS, Yip B, Chan KJ, Wood E, Craib KJ, O'Shaughnessy MV, et al. Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy. *JAMA.* 2001; **286**(20):2568-2577.

30. Edgar RC.MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; **32**(5):1792-1797.

31. Pollakis G, Kang S, Kliphuis A, Chalaby MI, Goudsmit J, Paxton WA. N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization. *J Biol Chem.* 2001; **276**(16):13433-13441.

32. Petropoulos CJ, Parkin NT, Limoli KL, Lie YS, Wrin T, Huang W, et al. A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. *Antimicrob Agents Chemother.* 2000; **44**(4):920-928.

33. Deeks SG, Wrin T, Liegler T, Hoh R, Hayden M, Barbour JD, et al. Virologic and immunologic consequences of discontinuing combination antiretroviral-drug therapy in HIV-infected patients with detectable viremia. *N Engl J Med.* 2001; **344**(7):472-480.

34. Bozzette SA, McCutchan JA, Spector SA, Wright B, Richman DD. A cross-sectional comparison of persons with syncytium- and non-syncytium-inducing human immunodeficiency virus. *J Infect Dis.* 1993; **168**(6):1374-1379.

35. Furrer H, Wendland T, Minder C, Christen A, von Overbeck J, Grunow R, et al. Association of syncytium-inducing phenotype of HIV-1 with CD4 cell count, viral load and sociodemographic characteristics. *AIDS.* 1998; **12**(11):1341-1346.

36. G. J. Moyle, A. Wildfire, S. Mandalia, H. Mayer, J. l Goodrich, J. Whitcomb, et al. Prevalence and predictive factors for CCR5 and CXCR4 co-receptor usage in a large cohort of HIV positive individuals [Abstract]. *International Conference on Antimicrobial Agents and Chemotherapy (ICAAC)* 2004; **44**

37. J. Demarest, T. Bonny, C. Vavro, C. Labranche, K. Kitrinos, C. McDanal, et al. HIV-1 co-receptor tropism in treatment naive and experienced subjects [Abstract]. *International Conference on Antimicrobial Agents and Chemotherapy (ICAAC)* 2004; **44**.

38. Scoggins RM, Taylor JR, Patrie J, van't Wout AB, Schuitemaker H, Camerini D. Pathogenesis of primary R5 human immunodeficiency virus type 1 clones in SCID-hu mice. *J Virol.* 2000; **74**(7):3205-3216.

39. van't Wout AB, Ran LJ, Kuiken CL, Kootstra NA, Pals ST, Schuitemaker H. Analysis of the temporal relationship between human immunodeficiency virus type 1 quasispecies in sequential blood samples and various organs obtained at autopsy. *J*

Virol. 1998; **72**(1):488-496.

40. Olivieri K, Scoggins RM, Bor YC, Matthews A, Mark D, Taylor JR, et al. The envelope gene is a cytopathic determinant of CCR5 tropic HIV-1. *Virology.* 2007; **358**(1):23-38.

41. Toulson AR, Harrigan R, Heath K, Yip B, Brumme ZL, Harris M, et al. Treatment interruption of highly active antiretroviral therapy in patients with nadir CD4 cell counts >200 cells/mm3. *J Infect Dis.* 2005; **192**(10):1787-1793.

42. Nabatov A, 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. *J. Virol*. 2004;**78**(1):524-30

43. Bagnerelli P, Fiorelli L, Vecchi M, Monachetti A, Menzo S, Clementi M. Analysis of the functional relationship between V3 loop and gp120 context with regard to human immunodeficiency virus coreceptor usage using naturally selected sequences and different viral backbones. *Virology*. 2003;**307**(2):328-40

44. 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.

45. Delobel P, Nugeyre MT, Cazabat M, Pasquier C, Marchou B, Massip P, et. al. Population-Based Sequencing of the V3 Region of env for Predicting the Coreceptor Usage of Human Immunodeficiency Virus Type 1 Quasispecies. *J. Clin. Microbiol.* 2007; **45**(5):1572-80.

46. Etemad-Moghadam B, Sun Y, Nicholson EK, Fernandes M, Liou K, Gomila R, et al. Envelope glycoprotein determinants of increased fusogenicity in a pathogenic simian-human immunodeficiency virus (SHIV-KB9) passaged in vivo. *J Virol.* 2000; **74**(9):4433-4440.

47. Westermark B, Ponten J, Hugosson R. Determinants for the establishment of permanent tissue culture lines from human gliomas. *Acta Pathol Microbiol Scand [A].* 1973; **81**(6):791-805.

48. Marmor M, Hertzmark K, Thomas SM, Halkitis PN, Vogler M. Resistance to HIV infection. *J Urban Health.* 2006; **83**(1):5-17.

49. Chabot DJ, Chen H, Dimitrov DS, Broder CC. N-linked glycosylation of CXCR4 masks coreceptor function for CCR5-dependent human immunodeficiency virus type 1 isolates. *J Virol.* 2000; **74**(9):4404-4413.

50. Coakley E, Petropoulos CJ,Whitcomb JM. Assessing chemokine co-receptor usage in HIV. *Curr Opin Infect Dis.* 2005; **18**(1):9-15.

CHAPTER 6

PREDICTING HIV CO-RECEPTOR USAGE ON THE BASIS OF GENETIC AND CLINICAL COVARIATES[5]

6.1 INTRODUCTION

HIV-1 enters target cells through a multi-stage interaction of the viral envelope protein gp120 with the CD4 host cell receptor and a cellular co-receptor, usually CCR5 or CXCR4 (reviewed in [1]). Individual virions are able to use one or the other or both co-receptors (R5/X4 phenotype). *In vivo*, R5-only virus is generally present over the entire course of infection [2], and *X4-capable* (X4 and/or R5/X4) variants are detected in approximately 50% of infected individuals at end-stage disease. The reason for this co-receptor switch remains unclear, but both *in vitro* studies and experiments in animal models suggest that the emergence of X4 virus is strongly associated with $CD4^+$ cell depletion and thus may be an important determinant of pathogenesis [3]. The question of whether X4 virus is a cause or emerges as a result of $CD4^+$ cell depletion (or both) as well as the evolutionary reasons for the development of these variants, remain largely unresolved [4]. The capacity of the virus to use CXCR4 lies at least partially in a change of several amino acids in the third hypervariable (V3) loop of HIV gp120 [5], although sequence changes outside the V3 region also contribute to co-receptor use [6].

[5] A version of this chapter is published as:
Sing T, Low AJ, Beerenwinkel N, Sander O, Cheung PK, Domingues FS, et. al. Predicting HIV Co-receptor Usage on the Basis of Genetic and Clinical Covariates. *Antiviral Ther.*2007;12(7):1097-106

In recent years, substantial attention has been devoted to HIV co-receptors due to their potential as drug targets, and antagonists of the CCR5 co-receptor are now in advanced clinical studies [7, 8]. Notably, co-receptor antagonists represent the first class of anti-HIV drugs targeting a host protein, rather than a viral protein. Unfortunately, the static nature of this target (as opposed to the rapidly adapting viral proteins targeted by classical drugs), does not prevent the emergence of resistant mutants. Specifically, resistance to co-receptor antagonists can include increased viral binding affinity to the co-receptor, changes in CCR5 binding, or the emergence of newly produced or pre-existing X4 variants [9]. Given the link between emerging X4 virus and disease progression, the need for careful monitoring of viral co-receptor usage for screening and treatment with CCR5 antagonists is apparent.

Phenotypic assays for monitoring viral drug resistance or co-receptor usage are commercially available, but are relatively expensive and have a relatively slow turnaround. Approaches based on the viral genotype promise potential alternatives for routine clinical usage. The complex relationship between viral genotype and phenotype and/or response to therapy has led to the development of sophisticated interpretation algorithms which have been successfully implemented for HIV drug resistance testing. These algorithms are now widely used to support treatment with the antiretroviral drug classes of protease (PIs) and reverse transcriptase inhibitors (RTIs). Genotypic approaches for monitoring co-receptor usage aim at detecting X4-capable virus with high sensitivity, while minimizing the number of false positives, i.e. R5-only variants that are incorrectly predicted as X4-capable. To date, the most popular genotypic predictor of

X4-capable virus is the simple “11/25” rule, based on the presence of arginine or lysine at positions 11 and/or 25 of the third hypervariable (V3) region of the envelope protein gp120 [5, 10-12]. However, previous studies have found the overall reliability of sequence motif-based methods for phenotype inference, especially for co-receptor usage prediction, to be limited [13]. Moreover, it has been suggested that for many V3 backgrounds, basic changes at 11 or 25 are neither necessary nor sufficient for a phenotype switch [14]. As a consequence, several alternatives based on statistical learning methods have been developed (reviewed in [15]), including linear regression [16], artificial neural networks [13], decision trees [17], support vector machines [17], position-specific scoring matrices [14], and mixtures of localized rules [18].

Here, we provide a comparison of these methods in a joint cross-validation setting, using a large clonal dataset of HIV V3 genotype-phenotype pairs. The raw output of all methods is a numeric score, optimized for class prediction (R5-only versus X4-capable) via a pre-defined and method-dependent cutoff. In contrast to the scenario reflected by clonal data, the viral population *in vivo* is a swarm of genetically and phenotypically heterogeneous variants, often termed a "quasispecies" [19]; therefore, approaches which give satisfactory results on clonal data may not be satisfactory on clinically derived data. In order to obtain a representative sample of this quasispecies, a substantial number of clones would have to be phenotyped and/or genotyped, which is not presently feasible in routine clinical practice. Instead, both the genotype and the phenotype are obtained using bulk or “population-based” approaches. To address these challenges of population-based data, we describe a strategy for dealing with sequence

ambiguities and integrating clinically derived data into the prediction of co-receptor usage.

6.2 METHODS

Sequence Alignment

Clonal and clinical samples were aligned with the multiple alignment package MUSCLE [20], using standard parameters, followed by visual inspection. No manual alignment correction was necessary.

Statistical analysis

Throughout, the association between discrete quantities was assessed using Fisher's exact test, while differences between the medians of numerical quantities in specific groups were assessed using Wilcoxon's rank sum test. The association of specific mutations with co-receptor usage and the covariation among pairs of mutations were also assessed using Fisher's exact test. In both cases, correction for multiple testing was performed using the Benjamini-Hochberg method at a false discovery rate of 5%. The R package *covaRius* [21] was used for analyzing mutation covariation, and is available from the authors upon request.

Statistical learning

The unifying principle underlying all statistical learning methods is that they rely on a training set of data in order to fit a model. In our case, the training data consists of V3 sequences (and augmented with clinical parameters in the second part of the study) for which the corresponding phenotypic co-receptor usage has been determined. The trained model can then be used to predict co-receptor usage from the V3 loop sequence

(and clinical data if applicable) from other samples. Support vector machines were trained using the package *libsvm*. Decision trees were trained with *Q5.0*, and artificial neural networks were trained using the backpropagation algorithm implemented in the R package *nnet*. PSSMs were implemented in R according to the description in [14], and mixtures of localized rules were implemented in Perl as described in [15]. For SVMs, a linear kernel was used, as standard nonlinear kernels did not improve prediction. Adjustable parameters were optimized in 10x10-fold cross-validation using the area under the ROC curve as the objective function (SVM cost: 2^{-5}; neural network weight decay 2^{3}). For SVMs, probabilistic output was obtained using Platt's method [22], as implemented in *libsvm*. As described in detail in the Results, different SVMs were trained separately on a clonal dataset and on data from population-based sequencing, to assess potential differences in predictive performance that can be expected from the two different sources.

Model evaluation

Throughout, predictive models were inferred and tested using 10x10-fold cross-validation. The clonal dataset consists of 1,100 samples from 332 patients. To reduce epidemiological bias, a maximum of one R5-only and one X4-capable sample from each patient was included in each cross-validation run. For many patients, multiple sequences were available. Those that are used in a particular training or test run were determined by random sampling in order to avoid overfitting to a single patient. This resulted in 100 (10 replicates of 10-fold cross-validation) training/test sets where seventy of these training sets contained 240 R5-only sequences, and 30 of them contained 241 R5-only sequences,

while 80 training sets contained 70 X4-capable sequences and the other 20 contained 71 X4-capable sequences. Predictive performance was analyzed using the R evaluation package *ROCR* [23], focusing on the trade-off between sensitivity and specificity which can be controlled by choosing a prediction cutoff for turning the continuous scores into actual class predictions. The area under the ROC curve was taken as a cutoff-independent class separation criterion. Averaged ROC curves were estimated from the 10x10 individual cross-validation curves using vertical or threshold averaging, as indicated in Figures 1 and 3. The calibration error [23] of the clinical SVM was assessed by pooling the scores obtained from ten individual folds within a replicate (using the ROCR default bandwidth of 100 samples), followed by vertical averaging over the ten replicates.

Feature ranking from SVM models

The weights of individual features in SVM models were obtained by exploiting the bilinearity of the scalar product that defines the linear kernel, as described in [21]. Briefly, the SVM decision function can be written as $f(x)=\sum y_i \alpha_i k(x_i,x) + b$, with (x_i, y_i) denoting training sequences and their true co-receptor usage, where x denotes a novel sequence whose co-receptor usage is to be predicted, and α_i and b are the support vector weights and a general offset, respectively. In the case of a linear kernel $k(x,y)=<x,y>$, the decision function can be rewritten as a model which is linear in the features (mutations and clinical parameters), allowing for direct assessment of the model weights: $f(x)=<\sum y_i \alpha_i x_i, x>+b$.

Treatment of sequence ambiguities in SVMs

The standard amino acid indicator representation used for SVMs consists of 20 variables (of value zero or one) for each of the 35 main V3 positions. Values of one indicate the presence of a particular amino acid at a particular position, whereas values of zero indicate its absence. In contrast to clonal samples, population-based clinical samples require a strategy for handling sequence ambiguities. The strategy used in this study ensured maximal sensitivity in X4 detection. This means that during model training, when the true class of samples is known, in X4-capable samples only X4-associated mutations are retained at ambiguous positions, and the dual approach is performed for R5-only samples. When a model is used for prediction, the sequence of all samples is modified as if they were X4-capable. This strategy corresponds to a "worst-case" scenario and improved predictive performance (data not shown) compared to a naive approach in which all mutations in an ambiguous sequence are considered in the indicator representation. V3 sequence positions that are not represented within the standard 35 amino acids due to insertions or deletions were ignored, except in the calculation of net charge or ambiguous positions.

Molecular graphics

The importance of specific sequence positions was superimposed on the three-dimensional structure of a V3 loop in the context of the gp120 core (Protein Data Bank identifier 2B4C) using the molecular viewer Pymol.

6.3 RESULTS

Reliability of prediction methods, evaluated on matched clonal pairs of V3 genotype and co-receptor usage phenotype

We compare the predictive performance of the "11/25" rule with alternatives based on statistical learning [13, 14, 17, 18] using ten replicates of 10-fold cross-validation. The comparison is based on 1,110 clonal genotype-phenotype pairs obtained from the Los Alamos HIV Sequence Database and from selected publications (the dataset is available for download at www.geno2pheno.org). The samples originate from 332 patients, with 769 R5, 131 R5/X4, and 210 X4 phenotypes. R5/X4 and X4 variants were pooled into a single class (*X4-capable*), as opposed to variants that are limited to using CCR5 (*R5-only*). 156 samples (14%) had insertions or deletions relative to the subtype B V3 consensus sequence, - CTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHC - the reference for position numbering in this study. It should be noted that the V3 region of the HXB2 [24] sequence has two insertions and one deletion relative to this consensus sequence.

The "11/25" rule has a mean sensitivity of 59.5% in detecting X4-capable variants and a mean specificity of 92.5% on the clonal isolates, in good agreement with previous studies. We compared this with a variety of other prediction methods in the framework of receiver operating characteristic (ROC) analysis [23] to analyze the sensitivity/specificity trade-off across the range of all possible cutoffs (Figure 6.1). These analyses, and all previous studies, are based on genetically and phenotypically homogeneous clonal samples. The ROC curve [23] shows the trade-off between

sensitivity and specificity by varying the score cutoff for all compared predictive methods. In our analysis, decision trees, neural networks, mixtures of localized rules or simple modifications of the '11/25' rule led to only minor improvements in predictive performance over the '11/25' rule when a method-specific cutoff corresponding to the '11/25' specificity of 92.5% was chosen. In contrast, classifiers based on position-specific scoring matrices (PSSMs) or support vector machines (SVMs) significantly outperformed the '11/25' rule, increasing sensitivity by 12.4% and 16.9% respectively at the '11/25' rule specificity. Using Wilcoxon's ranks sum test, the differences in sensitivity between the SVM and PSSM ($p=0.03$) or 11/25 ($p<10^{-12}$) were significant at this specificity. SVMs and PSSMs also showed significantly higher areas under the ROC curve (AUC) than the three other methods (0.91 and 0.90 respectively), indicating an overall improvement in the ability to distinguish X4-capable samples from R5-only samples.

Reliability of prediction methods on population-based genotype-phenotype pairs obtained from clinical practice

Clinically derived samples often contain mixtures of co-existing viral variants. Sequence ambiguities that remain undetected either by genotyping or phenotyping may differ from the exact genotype-phenotype match seen in clonal samples. The extent to which the population-based data confound the genotype-phenotype relation is evaluated here on a large clinical dataset. In order to assess the predictive performance of these methods on clinically derived isolates, we analyzed plasma samples from 952 antiretroviral-naïve patient samples with matched V3 genotype and co-receptor

phenotype. V3 genotype was previously determined using population-based 'bulk' sequencing techniques and co-receptor phenotype was obtained using the Trofile Co-receptor Assay (Monogram Biosciences) [25]. Only one sample was phenotyped as pure X4 with the co-receptor assay, suggesting that *in vivo*, X4 virus is only rarely present without concomitant R5 variants.

A total of 41 mutations at 19 V3 positions were significantly associated with co-receptor usage (25 predicting X4-capability, 16 predicting R5-only variants) in univariate analysis, using Fisher's exact test and correcting for multiple testing with the Benjamini-Hochberg method at a false discovery rate of 0.05 (Table 6.1). In Figure 6.2, the 14 V3 positions with X4-associated mutations are highlighted on the only V3 loop structure in context of the gp120 core available to date (Protein Data Bank identifier 2B4C). Most of the significant mutations are located close to the tip of the loop, where the two strands of the β-hairpin are in close spatial proximity. The well-known 11R mutation ($p=1.58x10^{-20}$) was present in 31 samples (often in mixtures also containing glycine or serine), 29 of which (94%) were phenotyped as X4-capable. However, other mutations were also strong X4 determinants. Mutation 13Y ($p<10^{-8}$) occurred in 21 samples, 16 of which were phenotyped as X4-capable. In all cases in which samples with 11R or 13Y were phenotyped as R5, sequence ambiguities were present at these positions, possibly indicating cases of genotype-phenotype mismatch. To assess the evolutionary role of 13Y and other mutations, we analyzed mutation covariation between all pairs of the 25 X4-associated V3 mutations, again using Fisher's exact test with the Benjamini-Hochberg method at a false discovery rate of 0.05. In this analysis, mutation 13Y was not

significantly associated with any mutation in a prominent cluster comprising 9S, 11R, 13S, and 24R, which had been partly described in a previous covariation analysis based on mutual information [12]. This suggests that 13Y (and possibly other V3 mutations) are involved in alternative evolutionary pathways from R5-only to X4-capable genotypes.

After this univariate analysis, we evaluated the predictive performance of the multivariable models on the population-based data, in comparison to the performance on the clonal data. We focused on the 11/25 rule (as the classical approach) and the SVM (as the best-performing prediction method on the clonal data). The specificity of the 11/25 rule remained close to that observed for the clonal data (93.5%), but the sensitivity decreased to 30.5% from 59.9% as observed in the clonal dataset. SVM-based prediction was evaluated using 10x10-fold cross-validation. Averaged over the 100 test sets from the cross-validation data, the mean sensitivity of the 11/25 rule was even lower (25.9%) than on the dataset as a whole, at a mean specificity of 93.9%. SVM-based prediction again outperformed the 11/25 rule, but, as compared to the clonal data, also decreased substantially in sensitivity to 39.8% at the 11/25 specificity of 93.5%.

Improving predictive performance for population-based data by integrating clinical markers

To identify potentially useful markers of co-receptor usage, we examined clinical parameters, including plasma viral load (VL), CD4 and CD8 cell counts, and the percentage of $CD4^+$ T-cells (CD4%) at the time of sampling for all of the 976 HOMER

patients [25]. Moreover, patients were tested for heterozygosity at the CCR5 Δ32 allele, a 32-basepair deletion resulting in non-functional CCR5 co-receptors [26]. The univariate association between clinical features and co-receptor usage is summarized in Table 6.2. To assess the predictive benefit of these features when combined with a purely sequence-based prediction approach, we evaluated different feature subsets in combination with a SVM-based classifier. The best feature combination relied on four additional features ($\log_{10}$(CD4%); host Δ32 heterozygosity; number of ambiguous amino acid V3 positions; and a variable indicating the presence of insertions or deletions in the V3 sequence). The improvements in sensitivity over the 11/25 rule and the purely sequence-based SVM were substantial when the clinical parameters are considered (Figure 6.3). The sensitivity of 63% at the "11/25" specificity corresponds to a 2.4-fold improvement in detecting X4-capable samples relative to the "11/25" rule. A SVM-based feature ranking showed that 221 of the 704 variables had non-zero weights in the combined model. In particular, 98 variables contributed to increased CXCR4 propensity, and 123 variables to increased CCR5 propensity, with the 30 variables for both showing an initially sharp decline in feature ranking before flattening out (Figure 6.4). The top 5-ranking variables were CD4 percent, the presence of mutation 13Y, the presence of mutation 11R, the number of ambiguous V3 positions, and the presence of mutation 24G. The probabilistic SVM output was generally well-calibrated (approximately 75% of the samples with X4 probability of 75% were indeed X4-capable; data not shown).

6.4 DISCUSSION

Genotypic methods for determining HIV co-receptor use from HIV V3 sequence can complement phenotypic assessments. This may allow large scale routine clinical assessments of co-receptor usage, which will be essential for the implementation of co-receptor antagonists. This study aimed to identify how close we are to reaching this promise and to provide directions for further improvement.

SVMs provide substantial improvement over the classical "11/25" rule as well as a modest improvement over the PSSM method. Still, training and testing SVMs on our clinically derived, population based genotype-phenotype dataset reveals a substantial decrease in sensitivity, compared to clonal data. The extent of this decrease contrasts with previous experience with the prediction of phenotypic resistance to protease or reverse transcriptase inhibitors. In that case, satisfactory results can be obtained even when using population-based data, and from as few as 500 genotype-phenotype pairs [27].

We hypothesize that the transition from clonal to clinically-derived data has a more profound impact on predicting co-receptor usage than on predicting phenotypic resistance, due to the different nature of the reported phenotype. In drug resistance, the phenotype is reported as a continuous quantity — the fold-change in 50% inhibitory concentration as compared to a reference strain — and the presence of mixed populations will only cause the reported phenotype to be more variable around the mean value in repeated experiments, as compared to measurements on homogeneous populations. In

contrast, co-receptor usage phenotype is a categorical quantity, and instead of only increasing measurement variability, mixed populations can lead to a complete mismatch between genotype and phenotype. Furthermore, X4 virus is usually present as a mixture or a minority species whereas resistance mutations are often the predominant viral species. The mismatch can be due to X4-capable variants detected in the phenotype, but undetected in the genotype, usually because they are a minority species. As an example, some of the population-based samples share identical nucleotide sequences of the sequenced regions, but are associated with different phenotypes. Mutations outside of the V3 region might also be relevant for predicting the coreceptor phenotype. On the other hand, X4-capable variants may remain undetected in the phenotype assay; when 74 of these clinical samples were tested with a second recombinant phenotype assay an assay agreement of only 85.1% was observed [28].

The intricacies of population-based data can be compensated for by leveraging additional information which is routinely available in the clinical setting, but not with clonal data. Indeed, we have shown that the very source of these intricacies, namely the genetic diversity of the viral population as measured by the number of ambiguous sequence positions, can be one of the most important predictors of co-receptor usage both in univariate analyses and in SVM-based models. Remarkably, the association between increasing viral genetic diversity and increased CXCR4 propensity observed here on cross-sectional, population-based data supports the model of *env* evolution postulated in the longitudinal study [29] on the basis of multiple clones per time point. Beyond the viral genomic information, we show that the host genome (CCR5 Δ32 heterozygosity),

and more significantly, host immunological status (as measured by CD4%) contain information relevant for predicting the potential presence of undetected X4-capable variants. While these quantities may or may not be causally related to co-receptor usage, the association with these parameters clearly helps distinguish environments that are more typical of X4-capable variants. Alternatively, the actual relationship between the V3-genotype and viral co-receptor phenotype may itself be dependent on CD4 count or CD4%, due to the overall virological context in which the V3 genotype is considered . These additional data on viral diversity, host factors, and host immunological status contain non-redundant information, as evidenced by the substantial improvements in predictive performance compared to purely sequence-based prediction.

The most important limitation of the present and previous studies is the exclusive consideration of the V3 region. There is accumulating evidence that other parts of the "bridging sheet" that connects the gp120 inner and outer domain are critically involved in co-receptor selectivity [30] and using key mutation in these regions could substantially increase the sensitivity of any genotype-based prediction algorithm.

By assembling a large dataset on co-receptor usage for population-based samples, we demonstrate that performance evaluations of clonal data — the basis for all prediction studies to date — provide positively biased estimates of the predictive reliability of genotypic methods for predicting HIV co-receptor use in clinical practice. In the latter scenario, the prediction problem is complicated by the use of population-based genotyping and phenotyping, and the presence of genetically and phenotypically

heterogeneous — possibly even undetected — viral subpopulations. A multitude of V3 mutations appear to be associated with co-receptor usage, in addition to the classical mutations at positions 11 and 25. The role of these mutations should be confirmed *in vitro* using mutagenesis studies in a variety of genetic backgrounds. Our results show consistent advantages to statistical learning methods over the 11/25 rule, but also highlight that the choice of adequate covariates is as important as the choice of the learning method. In particular, additional information on viral diversity (number of ambiguous V3 sequence positions), host immunological status (CD4%), and host genetic factors (CCR5Δ32 haplotype) can indicate whether a given host ecosystem is typical for the presence of X4-capable variants.

Table 6.1: V3 mutations significantly associated with co-receptor usage in univariate analyses of the population-based data.

N6 97.0/0.14	**N7** 96.7/0.16	**T8** 96.9/0.16	**R9** 93.0/0.50	**S11** 56.4/1.29	**I12** 94.0/0.38	**H13** 70.5/1.91	**R18** 75.7/1.45	**A19** 86.5/0.85	**F20** 84.4/1.13
N (-3.8) 97.0/98.1/92.1	N (-6.0) 96.7/98.2/89.9 **Y** (-3.0) 0.4/0.0/2.2	**I** (-4.0) 1.4/0.4/4.5 T (-3.6) 96.9/98.0/92.1	R (-7.9) 93.0/95.5/82.0 **S** (-4.2) 2.8/1.6/7.9 **K** (-3.7) 0.5/0.0/2.8	**R** (-19.8) 3.2/0.3/16.3	I (-3.5) 94.0/95.4/87.6	**Y** (-8.8) 2.3/0.6/9.6 **S** (-4.6) 4.6/3.1/11.2 **T** (-2.8) 5.0/3.9/10.1 **R** (-2.8) 3.8/2.8/8.4	S (-2.2) 8.0/9.3/2.2	**V** (-6.0) 3.9/2.3/11.2 A (-3.2) 86.5/88.4/78.1	F (-6.2) 84.4/87.4/71.3 **V** (-5.2) 1.8/0.8/6.7

Y21 86.9/0.93	**A22** 71.5/1.15	**T23** 93.2/0.59	**G24** 84.7/1.07	**E25** 44.8/2.50	**I26** 92.4/0.39	**I27** 88.8/0.84	**I30** 97.1/0.13	**Q32** 82.9/1.17
H (-5.2) 1.8/0.8/6.7 **I** (-3.8) 0.9/0.3/3.9	A (-4.5) 71.5/74.5/58.4 **T** (-4.4) 27.0/24.2/39.9	T (-4.6) 93.2/95.0/85.4 **A** (-2.7) 1.8/1.1/5.1 **R** (-2.7) 1.2/0.6/3.9	**R** (-11.4) 2.3/0.4/10.7 G (-7.9) 84.7/88.1/69.7 **E** (-3.9) 5.7/4.3/12.4 **S** (-3.0) 0.6/0.1/2.8	D (-4.8) 34.4/37.4/20.8 **R** (-4.6) 4.6/3.1/11.2 **Q** (-3.8) 7.0/5.4/14.0 **N** (-2.6) 2.6/1.8/6.2	I (-2.6) 92.4/93.7/86.5	**V** (-4.2) 7.1/5.4/14.6 I (-3.7) 88.8/90.7/80.3 **A** (-3.0) 0.6/0.1/2.8	I (-2.7) 97.1/98.0/93.3	Q (-4.0) 82.9/85.2/72.5 **K** (-3.6) 13.5/11.5/22.5

* Header row shows the most frequent residue at the V3 position with relative frequency/entropy below. Underlined mutations predict X4-capable virus, while the remaining predict R5-only virus, with relative frequency of All/R5-only/X4 samples below. P-value is indicated in brackets as $\log_{10}$.

Table 6.2. Univariate association of individual parameters with co-receptor usage.

Parameter	N	R5 (N)	R5-only [mean; median (IQR)]	X4 (N)	X4-capable [mean; median (IQR)]	P-value (Wilcoxon)
Viral load (x10^3 copies/mL)	976	799	220.7; 120 (48.5-310)	177	279.5; 175 (84-415)	0.0008
CD4+ T-cell count (cells/mm^3)	976	799	318.6; 290 (155-430)	177	175.3; 110 (30-260)	$7.8x10^{-21}$
CD8+ T-cell count (cells/mm^3)	790	656	933.2; 840 (580-1140)	134	945.0; 790 (470-1210)	0.5157
CD4 percentage	790	656	20.12; 19.00 (12.00-27.00)	134	10.29; 8.00 (3.00-15.00)	<<0.0001
# ambiguous V3 positions	952	785	1.09; 1.00 (0.00-2.00)	167	2.64; 2.00 (1.00-4.00)	<<0.0001
V3 net charge	952	785	5.02; 5.00 (4.00-6.00)	167	5.78; 6.00 (5.00-7.00)	<<0.0001

Parameter	N	R5 (N)	R5-only (N; %)	X4 (N)	X4-capable (N; %)	P-value (Fisher's exact)
Insertions or deletions	952	785	64; 8.15%	167	24; 14.37%	0.0177
CCR5 wt/Δ32 genotype	964	789	94; 11.91%	175	34; 9.43%	0.0132

Figure 6.1. Predictive performance of the 11/25 rule and five statistical learning methods, assessed on clonal data.

Evaluation was performed using ten replicates of 10-fold cross-validation, followed by threshold averaging of the ROC curves. The dotted vertical line indicates the specificity of the 11/25 rule (92.4%). While decision trees (DT), mixtures of localized rules (MLR), and artificial neural networks (ANN) did not improve substantially over the 11/25 rule in our analysis, PSSMs and SVMs did significantly improve sensitivity by 12.4% and 16.9%, respectively. One standard error indicates the sensitivity spread of 11/25 rule, PSSM, and SVM at the specificity of 92.4%.

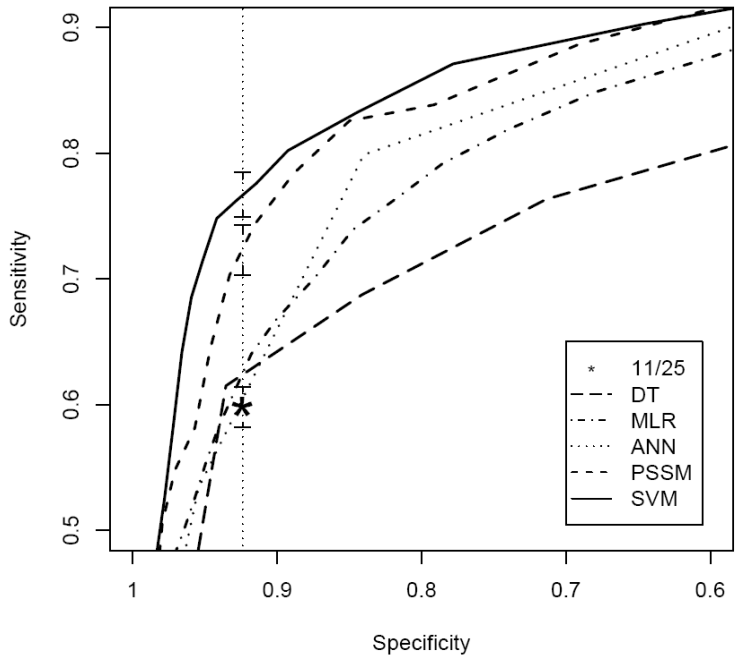


Figure 6.2. V3 loop predictors of CXCR4 usage.

The 35 amino acid V3 loop is shown extending from gp120. Amino acid positions are numbered from the N-terminus and colored according to the most relevant substitution (log_{10} p-value) from consensus for predicting co-receptor phenotype. Predictive ability varies from red (most significant) to blue (least significant). Positions without any mutation significantly associated with X4 virus are shown in grey.

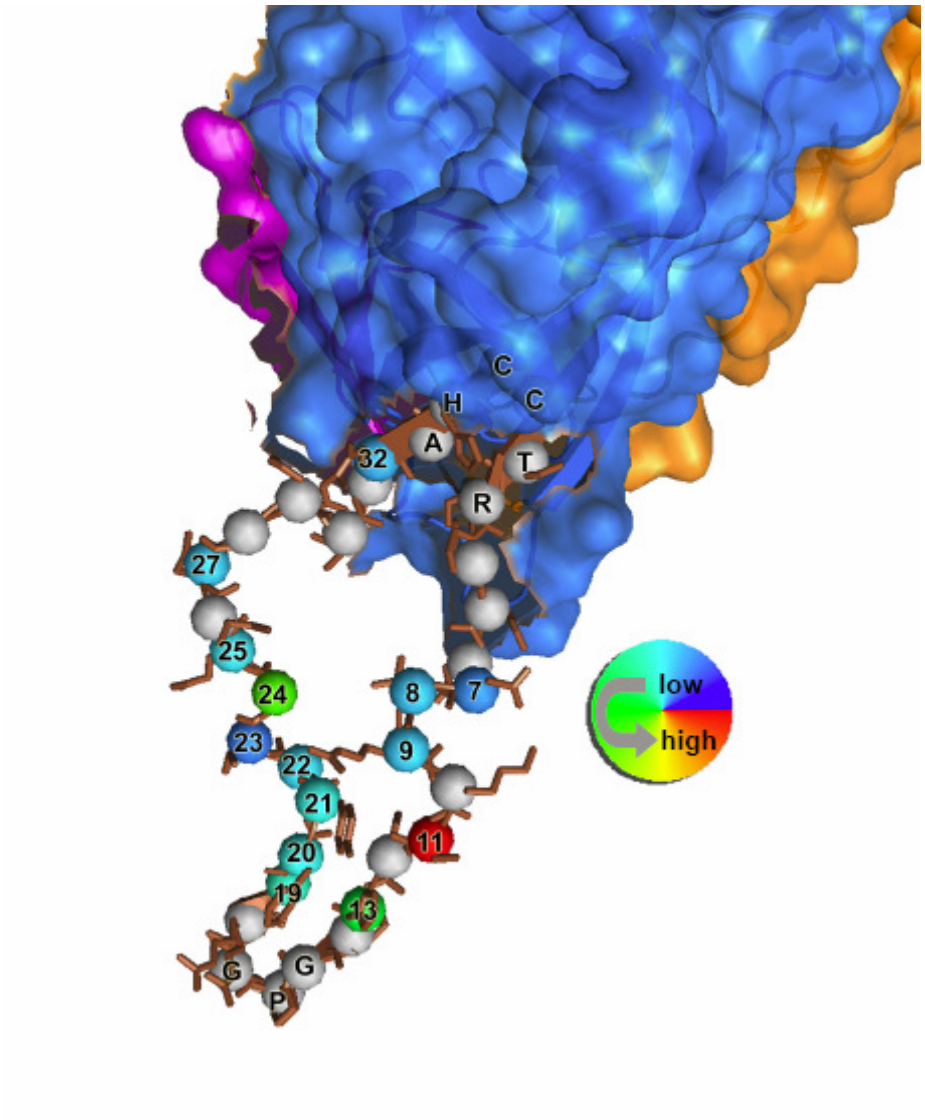


Figure 6.3. Predicting co-receptor usage on population-based data.

ROC curves are obtained by threshold averaging from ten replicates of 10-fold cross-validation. Bars indicate one standard error in horizontal and vertical direction. The 11/25 rule performs better when it is applied also in the presence of ambiguous positions 11 or 25 (*11/25[b]*), as compared to the requirement of unambiguous positions in these places (*11/25[a]*). Compared to the clonal dataset (Fig. 1), a substantial performance decrease is observed not only for the 11/25 rule, but also for the SVM with amino acid indicator representation (*SVM*). The inclusion of additional features (CD4%, number of sequence ambiguities, host CCR5 Δ32 heterozygosity, presence of insertions/deletions), leads to considerable improvements in predictive performance (SVM and CLINICAL).

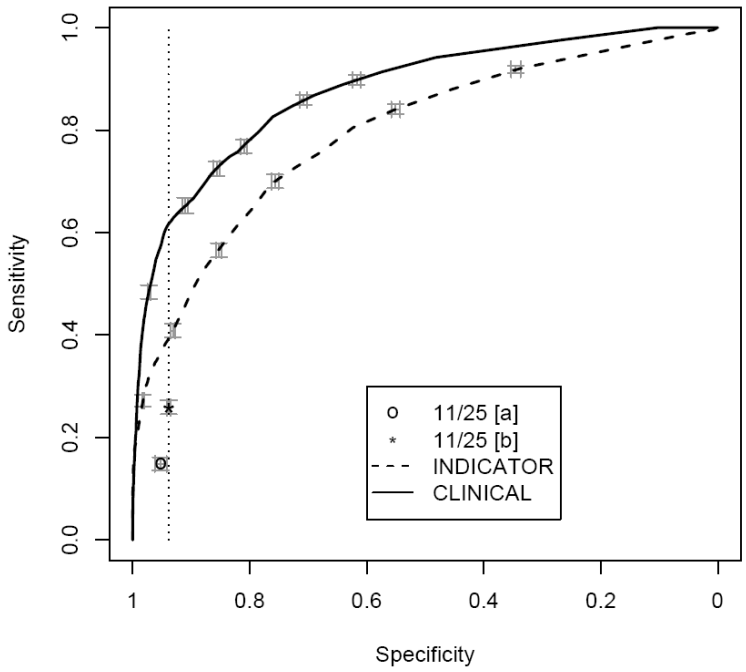
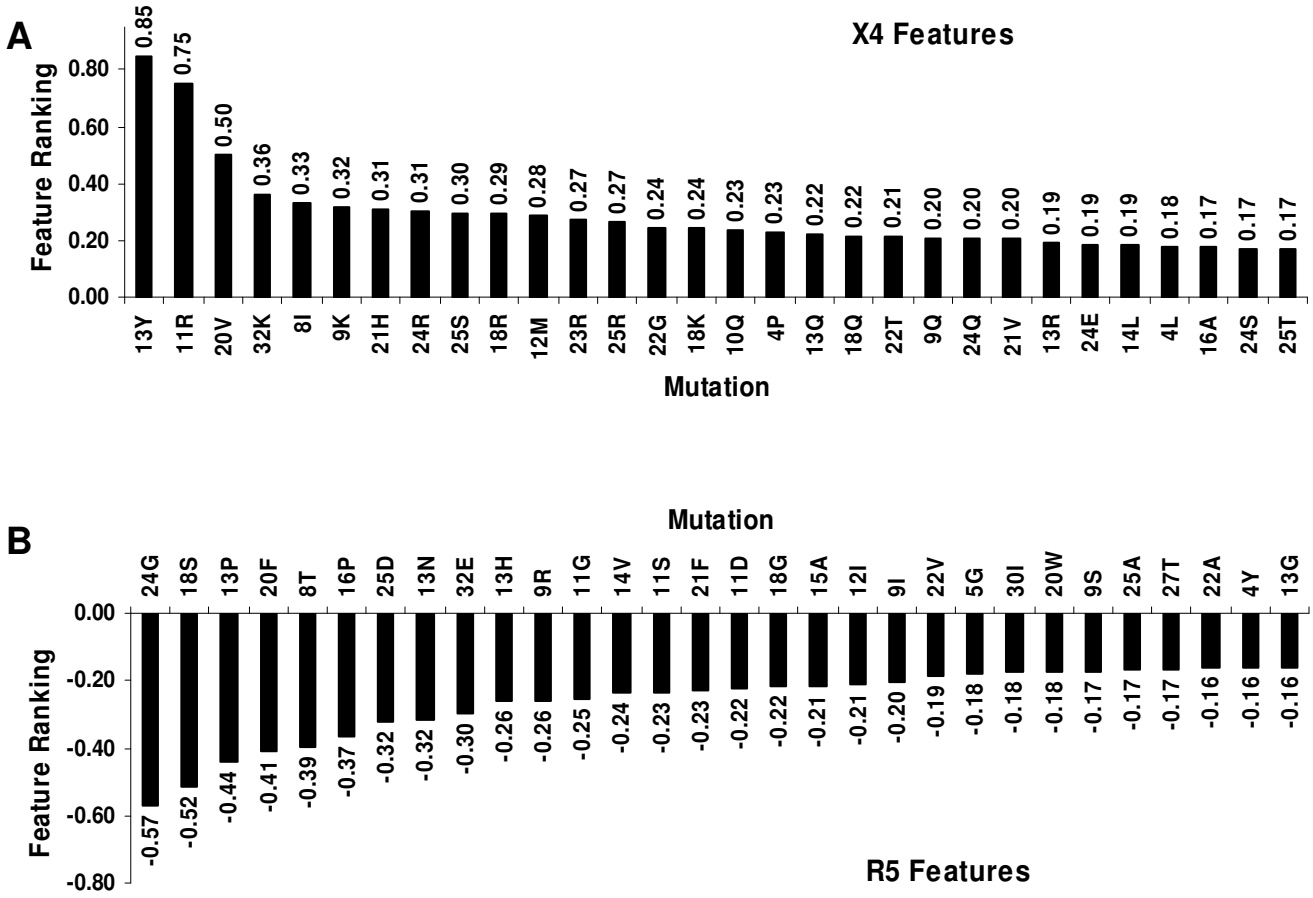


Figure 6.4. Feature ranking for the SVM trained on the clinical dataset.

The feature set consists of 704 variables: for each of the 35 V3 positions, 20 variables indicate the presence of specific amino acids and four additional variables represent CD4%, number of sequence ambiguities, host CCR5 Δ32 heterozygosity, and presence of insertions/deletions, respectively. The labels along the horizontal axis indicate the V3 position and the most prevalent amino acid at each position. Bar height represents the weight of a mutation. Only the 30 most important mutations for predicting X4-tropism (A) or R5-tropism (B) are illustrated. The weights of the four additional variables are as follows – *CD4%*: -1.40, *#sequence ambiguities*: 0.57, *Δ32*: 0.30*, indel*: -0.06.



Acknowledgements

Research at Max Planck Institute is conducted within the 6th framework EU program *EuResist* (www.euresist.org; IST-2004-027173) and the EU Network of Excellence *BioSapiens* (www.biosapiens.info; LSHG-CT-2003-503265). N.B. is funded by the Deutsche Forschungsgemeinschaft under grant no. BE 3217/1-1. Partial funding was provided by the Canadian Institute of Health Research and Pfizer Ltd.

6.5 REFERENCES

1. Sirois S, Sing T, Chou K C. HIV-1 gp120 V3 loop for structure-based drug design. Curr *Protein Pept Sci*. 2005; **6**:413-22.

2. Schuitemaker H, Kootstra NA, de Goede RE, de Wolf F, Miedema F, Tersmette M. Monocytotropic human immunodeficiency virus type 1 (HIV-1) variants detectable in all stages of HIV-1 infection lack T-cell line tropism and syncytium-inducing ability in primary T-cell culture. *J Virol.* 1991; **65**:356-63.

3. Regoes RR, Bonhoeffer S. The HIV coreceptor switch: a population dynamical P*erspective Trends Microbiol.* 2005; **13**:269-77.

4. Penn ML, Grivel JC, Schramm B, Goldsmithe MA, Margolis L. CXCR4 utilization is sufficient to trigger CD4+ T cell depletion in HIV-1-infected human lymphoid tissue. *Proc Natl Acad Sci USA*. 1999; **96**:663–8.

5. Fouchier RA, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema F, et. al. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J. Virol*. 1992; **66**:3183-7.

6. Groenink M, Andeweg AC, Fouchier RA, Broersen S, van der Jagt RC, Schuitemaker H, et. al. Phenotype-associated env gene variation among eight related human immunodeficiency virus type 1 clones: evidence for in vivo recombination and determinants of cytotropism outside the V3 domain. *J. Virol.* 1992; **66**:6175-80

7. Dorr P, Westby M, Dobbs S, Griffin P, Irvine B, Macartney M, et. al. 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. *Antimicrob Agents Chemother*. 2005; **49**:4721-32

8. Westby M, Van der Ryst E. CCR5 antagonists: host-targeted antivirals for the treatment of HIV infection. *Antivir Chem Chemother*. 2005; **16**:339-54.

9. Moore JP, Kitchen SG, Pugach P, Zack JA. The CCR5 and CXCR4 coreceptors--central to understanding the transmission and pathogenesis of human immunodeficiency virus type 1 infection. *AIDS Res Hum Retroviruses.* 2004; **20**(1):111-26.

10. De Jong JJ, Goudsmit J, Keulen W, Klaver B, Krone W, Tersmette M, et. al. Human immunodeficiency virus type 1 clones chimeric for the envelope V3 domain differ in syncytium formation and replication capacity. *J Virol*. 1992; **66**:757-65.

11. Fouchier RA, Brouwer M, Broersen SM, Schuitemaker H. Simple determination of human immunodeficiency virus type 1 syncytium-inducing V3 genotype by PCR. *J Clin Microbiol*. 1995; **33**:906-11.

12. Korber BT, Farber RM, Wolpert DH, Lapedes AS. Covariation of mutations in the V3 loop of human immunodeficiency virus type 1 envelope protein: an information theoretic analysis. *Proc Natl Acad Sci U S A*. 1993; **90**:7176-80.

13. 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**:51-62.

14. Jensen MA, Li FS, van 't Wout AB, Nickle DC, Shriner D, He HX, et. al. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences. *J Virol*. 2003; **77**:13376-88.

15. Jensen MA, van 't Wout AB. Predicting HIV-1 coreceptor usage with sequence analysis. *IDS Rev*. 2003; **5**:104-12.

16. 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**:2937-9.

17. Pillai S, Good B, Richman D, Corbeil J. A new perspective on V3 phenotype prediction. *AIDS Res Hum Retroviruses*. 2003; **19**:145-9.

18. Sing T, Beerenwinkel N, Lengauer T. Learning mixtures of localized rules by maximizing the area under the ROC curve. [Abstract]. *1st Int. Workshop on ROC Analysis in Artificial Intelligence*, Valencia, Spain. 2004; pp. 89-96.

19. Domingo E, Holland JJ, Biebricher CK, Eigen M. Quasispecies: The concept and the word. Molecular Evolution of the Viruses. In: Molecular Basis of Virus Evolution. A. Gibbs, C. Calisher and F. García-Arenal (editors). Cambridge University Press, Cambridge, UK. 1995; pp. 171–180.

20. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004; **32**:1792-7.

21. Sing T, Svicher V, Beerenwinkel N, Ceccherini-Silberstein F, Däumer M, Kaiser R, et. al. Characterization of Novel HIV Drug Resistance Mutations Using Clustering, Multidimensional Scaling, and SVM-Based Feature Ranking. In: *Knowledge Discovery in Databases: PKDD Vol. 3721.* Jorge A, Torgo L, Brazdil P, Camacho R, Gama J (editors). Porto, Portugal:Springer. 2005; pp.285-296.

22. Platt JC. Probabilities for SV Machines. In: *Advances in Large Margin Classifiers.* Smola A, Bartlett P, Schölkopf B, Schuurmans D (editors). Cambridge, MA:MIT Press. 2000; pp.61-74

23. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: Visualizing classifier performance in R. *Bioinformatics*. 2005; **21**:3940-1

24. Korber B, Foley BT, Kuiken C, Pillai SK, Sodroski JG . Numbering Positions in HIV Relative to HXB2. In: The Human Retroviruses and AIDS 1998 Compendium. Korber B. (editor). Los Alamos, NM:Los Alamos National Laboratory. 1998; pp.102-111.

25. Brumme ZL, Goodrich J, Mayer HB, Brumme CJ, Henrick BM, Wynhoven B, et. al. Molecular and clinical epidemiology of CXCR4-using HIV-1 in a large population of antiretroviral-naive individuals. *J Infect Dis*. 2005; **192**: 466-74.

26. Brumme ZL, Henrick BM, Brumme CJ, Hogg RS, Montaner JS, Harrigan PR. Short communication. Association of the CCR5delta32 mutation with clinical response and >5-year survival following initiation of first triple antiretroviral regimen. *Antivir Ther*. 2005; **10**:849-53.

27. Beerenwinkel N, Schmidt B, Walter H, Kaiser R, Lengauer T, Hoffmann D, et. al. Diversity and complexity of HIV-1 drug resistance: a bioinformatics approach to predicting phenotype from genotype. *Proc Natl Acad Sci U S A*. 2002 ;**99**:8271-6.

28. Skrabal K, Low A, Dong W, Mammano F, Sing T, Cheung P, et. al.. Is there a gold standard for determining HIV co-receptor use in clinical samples? The degree of correlation between two phenotypic assays and a bioinformatic model. *J Clin Microbiol*. 2007; **45**(2):279-84.

29. Shankarappa R, Margolick JB, Gange SJ, Rodrigo AG, Upchurch D, Farzadegan H, et. al.. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J Virol.* 1999; **73**:10489-502.

30. 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. *J Virol.* 2006; **80**:750-8.

CHAPTER 7

GENERAL DISCUSSION AND CONCLUSION

7.1 Introduction

Highly Active Antiretroviral Therapy (HAART) for HIV infection has made a substantial impact on AIDS mortality [1-5], but faces pressing challenges because of the life-long nature of therapy and the virus' capacity to mutate rapidly. The introduction of CCR5 antagonists [6,7] and integrase inhibitors [8], which will be available soon, have improved the ability to treat HIV infection by targeting novel mechanism for inhibiting HIV replication even in patients with virus resistant to current regimens. However, CCR5-antagonists pose a unique problem. CCR5-antagonists are not effective at reducing the plasma viral load in individuals with detectable circulating CXCR4-using virus [9], making these drugs unsuitable for patients with CXCR4-using virus. It is therefore essential to screen for HIV co-receptor usage in patients prior to considering them for CCR5-antagonist based therapies. Currently, screening is performed by recombinant co-receptor phenotyping assays, however, using the HIV envelope sequence to predict co-receptor usage may greatly reduce the cost and time associated with screening patients.

As described earlier, the general aim of this thesis is to evaluate the performance of genotype-based predictors of co-receptor usage as well as the recombinant phenotype assays on which they are tested in the clinical setting, with the purpose of investigating

how they can be improved. This chapter summarizes the major conclusions of each of the individual research papers, and discusses these conclusions in context of new studies that have been published since the undertaking of the studies presented here. This will be followed by a discussion of the broader implications of this work, as well as the potential application of these results in modern clinical practice.

7.2 Sensitivity of HIV Genotype-Based Co-receptor Prediction Algorithms

Bioinformatic, V3-loop-based approaches of predicting the genotype, such as support vector machines [10,11], neural networks and position specific scoring matrices [12] have already been proposed as useful methods for quickly and efficiently determining viral co-receptor phenotype in clonal isolates[13,14]. However, these studies failed to assess the sensitivity and specificity of these methods in clinical samples, which is of particular relevance if they are to be used to screen patients for CCR5 antagonists. It is for this reason that we undertook this study, which is the largest evaluation of the performance of HIV V3-genotype-based predictors on clinically-derived samples to date. This work highlights the lack of sensitivity inherent in these methods, due, in part to the difficulties associated with detection of minority variants.

It was surprising to note that the $SVM_{genomiac}$ displayed poorer sensitivity than the simple '11/25' rule, which predicts a sample as being X4-capable if positions 11 and/or 25 of the V3 loop are positively charged amino acids. However, it soon became apparent that this was likely a result of training with viral sequences originating from mixed clade datasets. This observation, combined with previously published data suggesting that V3 sequence-based PSSMs need to be specifically trained for virus originating from different clades [15], indicates the importance of determining viral clade before the use of genotype-based predictors.

Although genotype-based predictors of HIV co-receptor phenotype are currently inadequate for clinical applications, specific strategies have been proposed which could

enhance sensitivity of these methods. Increasing the limit of detection of minority variants will likely yield the most significant improvements in genotype-based co-receptor phenotype prediction. Technologies such as heteroduplex tracking assays [16], or new sequencing methods such as 454 'deep sequencing' [17] technologies are two examples which may improve the ability to detect minority X4 variants, and therefore increase sensitivity. Incorporating envelope sequences outside of the V3 region, and/or using clinical information such as CD4 count is also likely to result in increased sensitivity. Immediately evident is the observation that the type of bioinformatic algorithm used for training and prediction (PSSM, SVM etc) does not appear to influence the predictive ability of genotype-based methods as greatly the quality and appropriateness of the training set. The importance of training and testing on clinically derived data has not yet been evaluated, however, it is plausible that predictors, such as those evaluated in chapter three, which use training data from clonally derived sequences may not perform as well when tested on population based data, particularly due to undetected minority species. Until genotype-based predictors become as sensitive as co-receptor phenotype assays, or unless they can demonstrate increased effectiveness at predicting patient outcome with CCR5-antagonists, co-receptor phenotype assays will likely remain the 'gold-standard' for the screening of CCR5-antagonist based therapies.

7.3 Comparison of HIV Recombinant Co-receptor Phenotype Assays: Is there a 'Gold Standard'?

There are currently two commercially available co-receptor phenotype assays, including the Trofile co-receptor assay by Monogram Biosciences [18], and the Tropism Recombinant Test (TRT) assay by VIRAlliance [19]. A new co-receptor tropism assay by Virco will also be made available soon [20]. All assays have been well validated with regards to accuracy, reproducibility, and sensitivity of detection for minority variants using well defined clonal isolates [18-20]. However, these assays have yet to be validated against each other with patient-derived samples, where CXCR4-using virus may exist as a dual and/or mixed minority variant.

In order to examine the accuracy of these assays in clinical samples, we compared the results of the Trofile and TRT assays in a panel of 74 clinically-derived samples, and found an overall concordance of 85.1% between methods. Determining a pattern behind discordances would ultimately require a much larger sample size, however, our data suggested that most discordances occurred due to the inherent inconsistencies in sampling the extremely variable viral populations which occur in clinical samples. CXCR4-capable samples likely exist at varying levels and with varying affinities for the CXCR4 co-receptor. Even if an assay is 100% sensitive at determining if a viral isolate is X4-capable, there is no guarantee that a minority X4 variant will be amplified during RT-PCR, efficiently packaged into a pseudotyped virion, and ultimately detected using a phenotypic test. Another concern is that the levels of receptors and co-receptors expressed in the cell lines used in these assays may exceed those found on natural target

cells [21]. Since chemokine expression varies between individuals [22], it could be expected that the level of viral co-receptor affinity would vary, as a function of intra-individual evolution, correspondingly. Repeated tests on identical patient-derived plasma samples would significantly help in determining the reproducibility of these assays and would go a long way in characterizing their reliability in the clinical setting. However, despite these uncertainties, current recombinant HIV co-receptor assays are significantly more sensitive than current methods of predicting co-receptor usage from the HIV envelope. These recombinant assays are, at present, the most widely used methods of determining co-receptor usage and will likely remain the method of choice for screening prior to the use of CCR5-antagonist based therapies until there are improvements in genotyping technology.

7.4 Correlations of the output of a Recombinant Co-receptor Phenotype Assay with V3 Genotype and Clinical Immune Status

The use of recombinant co-receptor phenotype assays in clinical practice is relatively recent, and thus there has been little research focused on how these assays correlate with previously known clinical and genotypic predictors of HIV co-receptor usage. The published manuscript forming the basis of chapter five is the largest analysis to date comparing the results of a HIV recombinant co-receptor phenotype assay with clinical and genotypic markers of co-receptor usage. The association of relative light units (RLU) with CD4 count, which is a significant predictor of CXCR4-usage, as well as previously determined genotypic predictors of co-receptor usage present in envelope V3 loop suggests that the raw output of these recombinant assays may include relevant information regarding both HIV tropism as well as disease progression. A similar study performed with longitudinal samples may reveal tighter associations. In addition, the observed bi-modal distribution of CXCR4 RLU in patients with DM virus indicates the potential existence of two subtypes of CXCR4-capable virus. X4 virus has been hypothesized to evolve through a transitional period where its ability to bind to, and infect CXCR4-expresing CD4+ cells is at an intermediary level compared to truly X4 and truly R5 virus [23]. Thus, the bimodal distribution may represent both the transitional DM virus and a more strongly CXCR4-tropic, purely X4 virus. However, another study focused directly at answering this hypothesis will be required before this can be verified.

Since the magnitude of observed CXCR4 RLU appears to be associated with genotypic predictors of X4-tropic virus, it is possible that they may also provide a continuous variable which could function as a predictor of clinical progression. Perhaps most importantly, they may function as a more useful predictor of response to CCR5-antagonists than the binary tropism calls made by these assays. If CXCR4 RLU could indeed be used to predict clinical progression and/or response to CCR5-antagonists, this would be extremely useful in the clinical setting. We hope that the results of this study will lead to studies that may further dissect the relationship between the continuous outputs of recombinant co-receptor phenotype assays, genotypic outputs and clinical progression.

7.5 Using Clinical Parameters to Improve Prediction of HIV V3 Genotype-Based Predictive Algorithm

Although a number of HIV V3 genotype-based co-receptor phenotype predictors have been developed, their sensitivity on patient-derived samples is poor, thus limiting their usefulness in a clinical setting. One of the goals of this thesis work, however, was to identify methods which could improve the sensitivity of genotype-based predictors, while maintaining high specificities (>90%). The study in chapter six clearly demonstrates how the sensitivity of clonally trained HIV genotype-based predictors of co-receptor usage suffers when they are tested on patient-derived samples, likely due to the presence of minority variants. This study also demonstrated that leveraging additional information readily available from population genotyping methods, as well as other clinical (CD4 and CD4%) and host genome (CCR5Δ32 heterozygosity) information can improve sensitivity by up to 17%, depending on the predictor used. Host immune status, as measured by CD4%, provided the greatest increase in sensitivity. This data, combined with the CD4-dependence of predicting CXCR4-using virus with positively charged residues at position 25 of the V3 loop strongly suggests that the host immune system can provide useful information regarding the context within which the V3 region can be interpreted. In practical terms, the data in our study indicates that the V3 loop cannot be interpreted in isolation, which becomes especially important for future V3 loop-based genotypic predictors of HIV co-receptor usage.

7.6 Concluding Discussion

The introduction of CCR5 antagonists such as maraviroc increases the options available for constructing antiretroviral regimens; however it is coupled with the caveat that patients should be tested for HIV co-receptor tropism prior to initiating therapy. Failure to do so increases the risk that they may be integrating an ineffective drug into their regimen, thereby reducing the effectiveness of viral suppression as well as greatly increasing their risk for developing antiretroviral resistance. In addition, if administering CCR5-antagonists to patients with low or undetectable levels of CXCR4-using virus results in the emergence of high levels of CXCR4-using virus, this could potentially result in quicker CD4 cell count decline or quicker progression to AIDS; however, this last point is still unproven.

Although new antiretroviral treatments active against resistant strains and/or targeting different steps of the viral life cycle have been developed, the cost of such new treatments remains a concern. This becomes especially apparent with the introduction of CCR5-antagonists, which will likely require costly HIV co-receptor phenotyping assays before they can be used. Using the HIV envelope sequence to predict co-receptor usage would be a relatively faster and cheaper method than using recombinant co-receptor phenotyping assays which are currently used for screening. However, these predictive methods are currently inadequate for screening purposes due to low sensitivities, although even the recombinant assays display discordances when compared to against each other in patient-derived samples (as seen in chapter four). Gaining a better understanding of the output of these assays and correlating them with clinical progression

and known genotype-based predictors of co-receptor usage will provide some indication on how both genotype-based, and phenotypic assays for determining HIV co-receptor usage can be improved. Using a surrogate marker of the clinical status of a patient (such as CD4 count or CD4% as seen in chapter six) can also significantly increase the sensitivity of genotype-based predictors. However, it is still unclear if the predictive power of using the V3-loop alone is sufficient or must be increased by using other clinically relevant determinants as well as genotypic determinants within gp120 and gp41 in combination with the V3 loop. Future studies should focus on determining if HIV envelope based methods of predicting HIV co-receptor phenotype can also be applied to predicting the outcome of CCR5-antagonist based therapies, a result which is more valuable than predicting the co-receptor phenotype alone.

7.7 References

1. Hammer SM, Squires KE, Hughes MD, Grimes JM, Demeter LM, Currier JS, et al. 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. *N Engl J Med.* 1997; **337**(11):725-733.

2. Palella FJ,Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV outpatient study investigators. *N Engl J Med.* 1998; **338**(13):853-860.

3. Hogg RS, Yip B, Kully C, Craib KJ, O'Shaughnessy MV, Schechter MT, et al. Improved survival among HIV-infected patients after initiation of triple-drug antiretroviral regimens. *CMAJ.* 1999; **160**(5):659-665.

4. Mocroft A, Vella S, Benfield TL, Chiesi A, Miller V, Gargalianos P, et al. Changing patterns of mortality across europe in patients infected with HIV-1. EuroSIDA study group. *Lancet.* 1998; **352**(9142):1725-1730.

5. Detels R, Munoz A, McFarlane G, Kingsley LA, Margolick JB, Giorgi J, et al. Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. multicenter AIDS cohort study investigators. *JAMA*. 1998; **280**(17):1497-1503.

6. M. S. Saag, P. Ive, J. Heera, M. Tawadrous, E. DeJesus, N. Clumeck, et al. **A** multicenter, randomized, double-blind, comparative trial of a novel CCR5 antagonist, maraviroc versus efavirenz, both in combination with Combivir (zidovudine [ZDV]/lamivudine [3TC]), for the treatment of antiretroviral naive subjects infected with R5 HIV-1: week 48 results of the MERIT study [Abstract]. *4th IAS Conference on HIV Pathogenesis, Treatment, and Prevention. Sydney, Australiah* July 22-25, 2007; WESS104.

7. Strizki JM, Tremblay C, Xu S, Wojcik L, Wagner N, Gonsiorek W, et al. Discovery and characterization of vicriviroc (SCH 417690), a CCR5 antagonist with potent activity against human immunodeficiency virus type 1. *Antimicrob Agents Chemother.* 2005; **49**(12):4911-4919.

8. M. Markowitz, B. -Y Nguyen, E. Gotuzzo, F. Mendo, W. Ratanasuwan, C. Kovacs, et al. Rapid onset and durable antiretroviral effect of raltegravir (MK-0518), a novel HIV-1 integrase inhibitor, as part of combination ART in treatment HIV-1 infected patients: 48-week data [Abstract]. *4th IAS Conference on HIV Pathogenesis,*

Treatment, and Prevention. Sydney, Australia July 22-25, 2007; TUAB104.

9. H. Mayer, Van der Ryst, E., M. Saag, B. Clotet, G. Fätkenheuer, N. Clumeck, et al. Safety and efficacy of maraviroc, a novel CCR5 antagonist, when used in combination with optimized background therapy for the treatment of antiretroviral-experiences subjects infected with dual/mixed-tropic HIV-1: 24-week results of a phase 2b exploratory trial.[Abstract]. *XVI International AIDS Conference, Toronto, Ontario, Canada* August 13-18, 2006; THLB0215.

10. Pillai S, Good B, Richman D,Corbeil J. A new perspective on V3 phenotype prediction. *AIDS Res Hum Retroviruses.* 2003; **19**(2):145-149.

11. T. Sing, T. Lengauer, N. Beerenwinkel, R. Kaiser, M. Daumer, A. Low, et al. Immunological and virological parameters improve genotypic prediction of HIV-1 coreceptor usage [Abstract]. *1st Int. Workshop on Targeting HIV Entry, Bethesda, Maryland, USA*. December 2-3, 2005; Abstract 30.

12. Jensen MA, Li FS, van 't Wout AB, Nickle DC, Shriner D, He HX, et al. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences. *J Virol.* 2003; **77**(24):13376-13388.

13. V. Briz, E. Poveda, V. Roulet, M. Gonzalez, J. -L Faudon, K. Skrabal, et al. Correlation between a phenotypic assay and three bioinformatic tools for determining HIV co-receptor use [Abstract]. *4th IAS Conference on HIV Pathogenesis, Treatment, and Prevention. Sydney, Australia* July 22-25, 2007; TUPEB027.

14. C. de Mendoza, K. Van Baelen, E. Poveda, E. Rondelez, N. Zahonero, L. Stuyver, et al. Prevalence of X4 or X4/R5 dual-mixed viruses in recent HIV-1 seroconverters using a population-based HIV-1 tropism phenotypic assay [Abstract]. *4th IAS Conference on HIV Pathogenesis, Treatment, and Prevention. Sydney, Australia* July 22-25, 2007; WEPEA004.

15. 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. *J Virol.* 2006; **80**(10):4698-4704.

16. Nelson JA, Baribaud F, Edwards T,Swanstrom R. Patterns of changes in human immunodeficiency virus type 1 V3 sequence populations late in infection. *J Virol.* 2000; **74**(18):8494-8501.

17. M. Lewis, I. James, M. Braverman, B. Desany, T. Jarvie, M. Penny, et al. Evaluation of an Ultra-deep Sequencing Method to Identify Minority Sequence Variants in the

HIV-1 *env* Gene from Clinical Samples [Abstract]. *14th Conference on Retroviruses and Opportunistic Infections, Los Angeles, California, USA* February 25-28, 2007; Abstract 680.

18. Whitcomb JM, Huang W, Fransen S, Limoli K, Toma J, Wrin T, et al. Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism. *Antimicrob Agents Chemother.* 2007; **51**(2):566-575.

19. Trouplin V, Salvatori F, Cappello F, Obry V, Brelot A, Heveker N, et al. Determination of coreceptor usage of human immunodeficiency virus type 1 from patient plasma samples by using a recombinant phenotypic assay. *J Virol.* 2001; **75**(1):251-259.

20. Ping LH, Nelson JA, Hoffman IF, Schock J, Lamers SL, Goodman M, et al. Characterization of V3 sequence heterogeneity in subtype C human immunodeficiency virus type 1 isolates from malawi: Underrepresentation of X4 variants. *J Virol.* 1999; **73**(8):6271-6281.

21. Coakley E, Petropoulos CJ,Whitcomb JM. Assessing chemokine co-receptor usage in HIV. *Curr Opin Infect Dis.* 2005; **18**(1):9-15.

22. Lee B, Sharron M, Montaner LJ, Weissman D,Doms RW. Quantification of CD4, CCR5, and CXCR4 levels on lymphocyte subsets, dendritic cells, and differentially conditioned monocyte-derived macrophages. *Proc Natl Acad Sci U S A.* 1999; **96**(9):5215-5220.

23. Jensen MA, van 't Wout AB. Predicting HIV-1 coreceptor usage with sequence analysis. *AIDS Rev.* 2003; **5**(2):104-112.

APPENDIX

Certificates of Ethical Approval

PROVIDENCE HEALTH CARE
Research Institute

UBC-Providence Health Care
Research Institute
Office of Research Services
11th Floor Hornby Site - SPH
c/o 1081 Burrard St.
Vancouver, BC V6Z 1Y6
Tel: (604) 806-8567
Fax: (604) 806-8568

ETHICS CERTIFICATE OF EXPEDITED APPROVAL

PRINCIPAL INVESTIGATOR:	DEPARTMENT:	UBC-PHC REB NUMBER:
P. Richard Harrigan	Medicine - Infectious Disease	H07-00987

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:

Institution	Site
Providence Health Care	St. Paul's Hospital

Other locations where the research will be conducted:
Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill ; The Whitehead Institute

COINVESTIGATOR(S):

Eric Lander
David Heckerman
Satish Pillai
Ronald Swanstrom
Andrew J. Low
Benjamin Good
Thomas Lengauer

SPONSORING AGENCIES:

British Columbia Centre for Excellence in HIV/AIDS

PROJECT TITLE:
Determination of HIV CXCR4-using minority species using more sensitive or complementary genotypic methods

THE CURRENT UBC-PHC REB APPROVAL FOR THIS STUDY EXPIRES: June 21, 2008

The UBC-PHC Research Ethics Board Chair or Associate Chair, has reviewed the above described research project, including associated documentation noted below, and finds the research project acceptable on ethical grounds for research involving human subjects and hereby grants approval.

DOCUMENTS INCLUDED IN THIS APPROVAL:		APPROVAL DATE: June 21, 2007
Document Name	**Version**	**Date**
Protocol: HIV CXCR4-tropic minority species detection protocol	1	May 28, 2007

CERTIFICATION:

1. The membership of the UBC-PHC REB complies with the membership requirements for research ethics boards defined in Part C Division 5 of the Food and Drug Regulations of Canada.
2. The UBC-PHC REB carries out its functions in a manner fully consistent with Good Clinical Practices.
3. The UBC-PHC REB has reviewed and approved the research project named on this Certificate of Approval including any associated consent form and taken the action noted above. This research project is to be

conducted by the principal investigator named above at the specified research site(s). This review of the UBC-PHC REB have been documented in writing.

Approval of the UBC-PHC Research Ethics Board or Associate Chair, verified by the above signature of one of the following:

Dr. I. Fedoroff,
Chair

Dr. J. Kernahan,
Associate Chair