INVESTIGATION OF VIRAL GENETIC AND BIOLOGIC DETERMINANTS OF HIV-1 SUBTYPE C PREDOMINANCE IN INDIA

by

Milka A. Rodriguez

BS, State University of New York at Stony Brook, 1998

Submitted to the Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This dissertation was presented

by

Milka A. Rodriguez

It was defended on

August 14, 2007

and approved by

Velpandi Ayyavoo, PhD, Professor, Department of Infectious Diseases and Microbiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA

Todd Reinhart, ScD, Associate Professor, Department of Infectious Diseases and Microbiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA

Ronald Montelaro, PhD, Professor, Department of Molecular Genetics and Biochemistry, School of Medicine, University of Pittsburgh, Pittsburgh, PA

Dissertation Advisor: Phalguni Gupta, PhD, Professor, Department of Infectious Diseases and Microbiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA

INVESTIGATION OF VIRAL GENETIC AND BIOLOGIC DETERMINANTS OF HIV-1 SUBTYPE C PREDOMINANCE IN INDIA

Milka A. Rodriguez, PhD

University of Pittsburgh, 2007

In India, HIV-1 subtype C has been the predominant subtype throughout the course of the HIV-1 epidemic, regardless of geographic region in the country. We hypothesize that the dominance of HIV-1 subtype C compared to other subtypes in India is due to enhanced replication fitness and/or enhanced transmission efficiency of this subtype across the mucosal surface over other subtypes present in India. The specific aims of this project are: (1) to compare the replication fitness between Indian HIV-1 subtype A and subtype C; (2) to evaluate the transmission efficiency of Indian HIV-1 subtype A and subtype C across the mucosa of cervical tissue; and (3) to determine the role of the LTR and *env* gene in replication fitness and transmission efficiency. Replication fitness was assessed using a dual infection growth competition assay. We observed that primary HIV-1 subtype C isolates had higher overall relative fitness and transmission efficiency than primary subtype A isolates in PBMC and in an ex vivo cervical tissue derived organ culture, respectively. Furthermore, a comparison of replicative fitness between a subtype A/subtype C half genome chimeric virus and parental subtype A virus indicates that the higher replication fitness and transmission efficiency of subtype C virus over subtype A virus from India is not due to the env gene alone. We have also characterized the genetic structure and functional characteristics of subtype A and subtype C LTRs from India. Despite their apparent variability, no significant difference was observed in the transcriptional activity between the LTRs of subtype A and subtype C. Therefore, the LTR region alone is not responsible for higher

replication fitness of subtype C over subtype A. The findings presented in this study are significant for public health because an understanding of the mechanism of the asymmetric distribution of HIV-1 subtypes in India is an important component in the development of strategies to control HIV-1 infection in this country.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

1.1 HIV VIRUSES AND THE GLOBAL EPIDEMIC

Human immunodeficiency virus (HIV), the causative agent of acquired immune deficiency syndrome (AIDS) in the human population, is a member of the genus *Lentivirus* in the family *Retroviridae*. HIV-1 is responsible for the current pandemic, having spread to more than 150 countries on six continents (1). The spread of HIV-2 has not been nearly as extensive as HIV-1 although infections have been reported in Europe, the United States and South America and accounts for a substantial number of infections in West Africa (1, 2). There are currently an estimated 40 million people worldwide living with HIV/AIDS (Figure 1) (3). Forty-nine percent are men, forty-five percent are women and approximately six percent are children under 15 years old (3). The major route of HIV transmission is through heterosexual contact, although injection drug use, men who have sex with men and mother-to-child transmission also constitute a considerable fraction of transmission groups in some countries (1, 2).

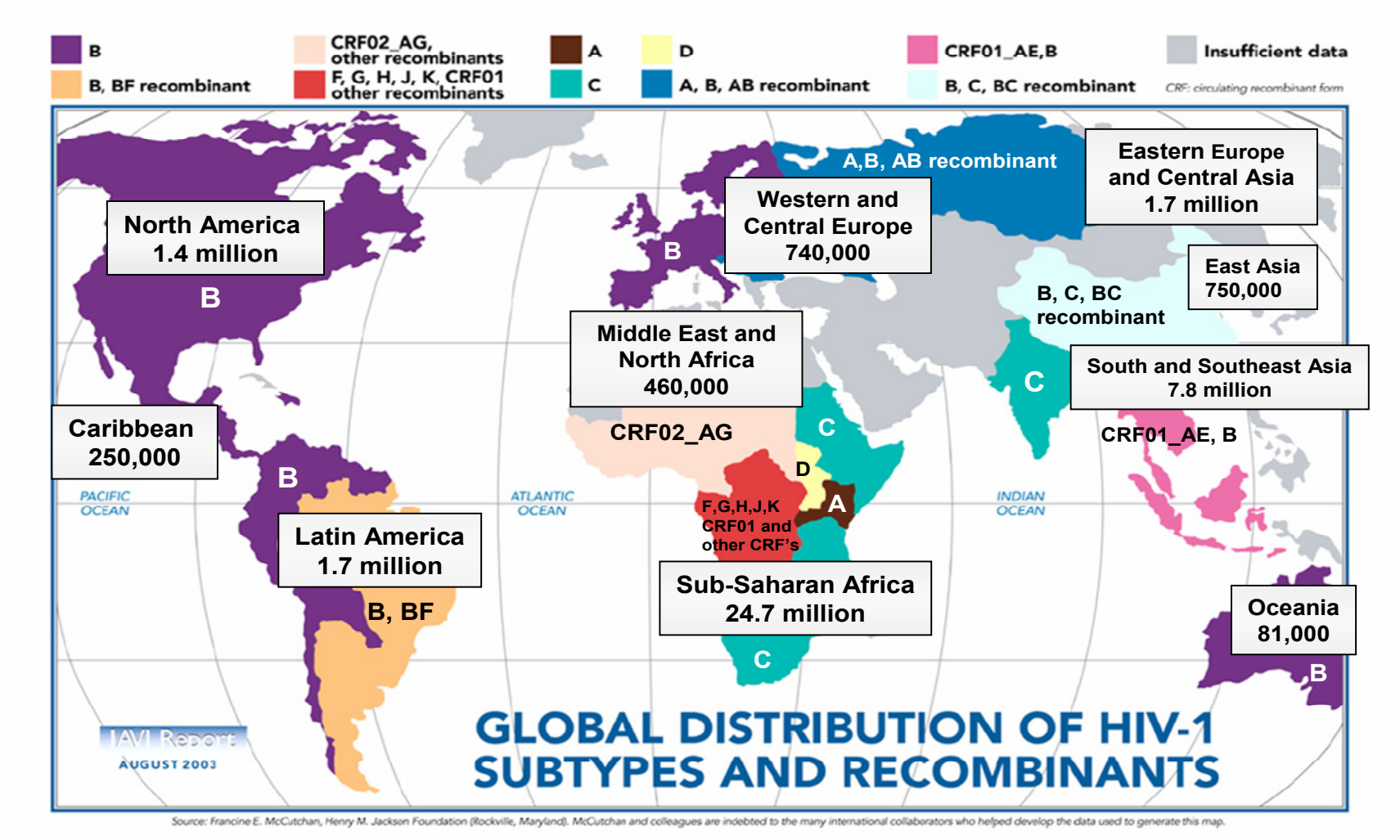

Figure 1. Global numbers of adults and children estimated to be living with HIV in 2006.

Adapted from publicly available data from UNAIDS and WHO. www.unaids.org. The colors depict regional patterns of HIV variation which are also indicated by letters in each region. Photo/ illustration is reprinted with permission from Francine E. McCutchan and Henry M. Jackson Foundation (Rockville, Maryland) ©2006 the International AIDS Vaccine Initiative (IAVI).

1.2 CLASSIFICATION OF HIV-1

HIV-1 was introduced into the human population through three separate cross-species transmissions from a *Pan troglodytes troglodytes* chimpanzee reservoir in central Africa (4). Each of these transmission events is represented by phylogenetic groups termed group M (main), group O (outlier) and group N (non-M, non-O) (Figure 2). Considerable diversification has evolved within these groups - and especially in group M - due to the error prone nature of the reverse transcriptase (RT) enzyme which can introduce approximately 1 mutation per genome, per round of replication. This coupled with high viral turnover by which 10^9 - 10^{10} virions are produced per day creates a swarm of genetically diverse but related populations termed quasispecies (5, 6). An additional factor contributing to viral evolution is recombination which can occur by way of alternate copying of viral RNA strands by RT. The crossover frequency ranges from 7 to 30 per genome per replication round (7).

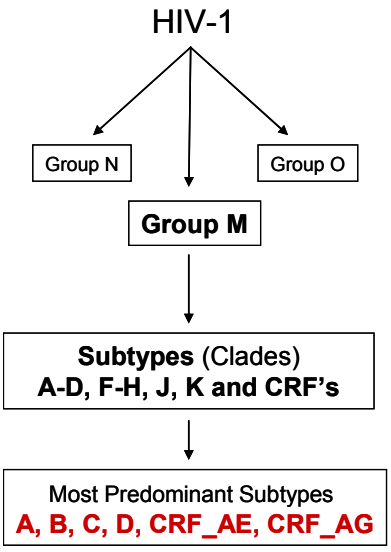

Figure 2. Classification of HIV-1 genetic forms

The high genetic diversity of HIV-1 has lead to further sub-classification of the viral groups. HIV-1 Group M, which is responsible for the global pandemic, is further subdivided into nine genetic groups called subtypes or clades, based on sequence diversity in the envelope gene (Figure 2). All Group M subtypes, namely A – D, F – H, J and K, are believed to have originated in Central Africa. At least three epidemiologically unlinked sequences are required to define a subtype (1). Intrasubtype amino acid diversity in the envelope gene ranges from 5-20% and intersubtype diversity ranges from 25-35%. There is also a growing percentage of the HIV-1 epidemic that is now comprised of circulating recombinant forms (CRFs). CRFs, of which at least 20 have now been classified, are the product of intersubtype recombination events.

1.3 GLOBAL DISTRIBUTION OF HIV-1 GENETIC FORMS

Group M viruses are responsible for the global pandemic whereas group O viruses are mainly found in Cameroon and some neighboring countries and group N viruses have been reported exclusively in Cameroon. Among the group M subtypes the most predominant are A, B, C and D and CRF01_AE and CRF02_AG (7). Subtype A is concentrated in East and West Africa, Eastern Europe and Central Asia and subtype B in the Americas, Western Europe and Australia. Subtype C which has become the most predominant subtype worldwide is concentrated in Southern and Eastern Africa and India. Subtype D is also prevalent in East and West Africa. CRF01_AE is the predominant subtype throughout Southeast Asia and CRF02_AG has emerged as the dominant subtype in West and West Central Africa (1, 7) (Figure 1). Geographically, there appears to be an asymmetric distribution of HIV-1 subtypes. Some explanations for this unequal spread of

HIV-1 subtypes include founder effects, social and behavioral practices, human genetic susceptibility and viral attributes such as transmission efficiency and replication fitness.

1.4 BIOLOGY OF HIV-1

This section describes the basic structure and genomic organization of HIV, the HIV life cycle and HIV pathogenesis.

1.4.1 HIV Virion Structure and Genomic Organization

The HIV virion is 100 to 120 nm in diameter (2) (Figure 3A). It is composed of two RNA strands that are surrounded by a conical shaped core which is composed of gag-p24 capsid protein. The viral RNA-dependent DNA polymerase (RT) and nucleocapsid proteins are closely associated with the genomic RNA within the core. Outside of the core lie accessory and structural proteins involved in early and late events in the replication cycle. The exact locations of the *vif* and *nef* proteins are unclear, although they are closely associated with the core. Vif promotes infectivity and aids in proviral DNA synthesis and maturation while *nef* has been reported to either increase or decrease virus replication. Vpr is found outside the core and has several functions involved in virus replication, regulation of cell cycling and viral transactivation. The matrix protein is located just below the viral membrane and functions in viral maturation and budding (8). The entire virus particle is surrounded by a lipid bilayer membrane which it obtains upon budding from an infected cell. Finally, embedded in the viral membrane is the envelope glycoprotein composed of a gp120 and gp41 subunit. These proteins form trimers, referred to as spikes, on the

surface of the virus. The gp120 subunit binds to the host cell CD4 receptor and a coreceptor to gain entry into a cell. The gp41 subunit is a transmembrane protein that is noncovalently bound to gp120. It is involved in conformational changes in the envelope protein necessary for membrane fusion and in envelope glycoprotein incorporation into maturing virions (2).

The HIV genome is approximately 10kb and contains 9 genes which code for 15 proteins (Figure 3B). At the 5' end of the genome lies the long terminal repeat (LTR) which is also duplicated at the 3' end. The LTR functions as the viral promoter containing several transcription factor binding sites that modulate virus replication. Following the 5' LTR are the *gag* and *pol* genes. The *gag* precursor is cleaved into MA, p17; CA, p24; p7, p6, p2 and p1 proteins. The *pol* precursor is cleaved into 3 enzymes – protease (PR), reverse transcriptase (RT) and integrase (IN). Protease functions in posttranslational processing of viral proteins. Reverse transcriptase generates cDNA from the viral RNA genome and also contains an RNase H domain that functions during reverse transcription. Following p*ol* are the accessory genes which code for the v*if, vpr* and v*pu (vpx* is the *vpr* homolog in HIV-2) proteins. These accessory proteins serve a variety of functions throughout the virus life cycle. The regulatory proteins Tat and Rev are generated from multiply spliced mRNA regions of the envelope gene. The e*nv* precursor gp160 is cleaved into gp120 and gp41 proteins. The viral protein Nef is partially encoded in the 3'LTR.

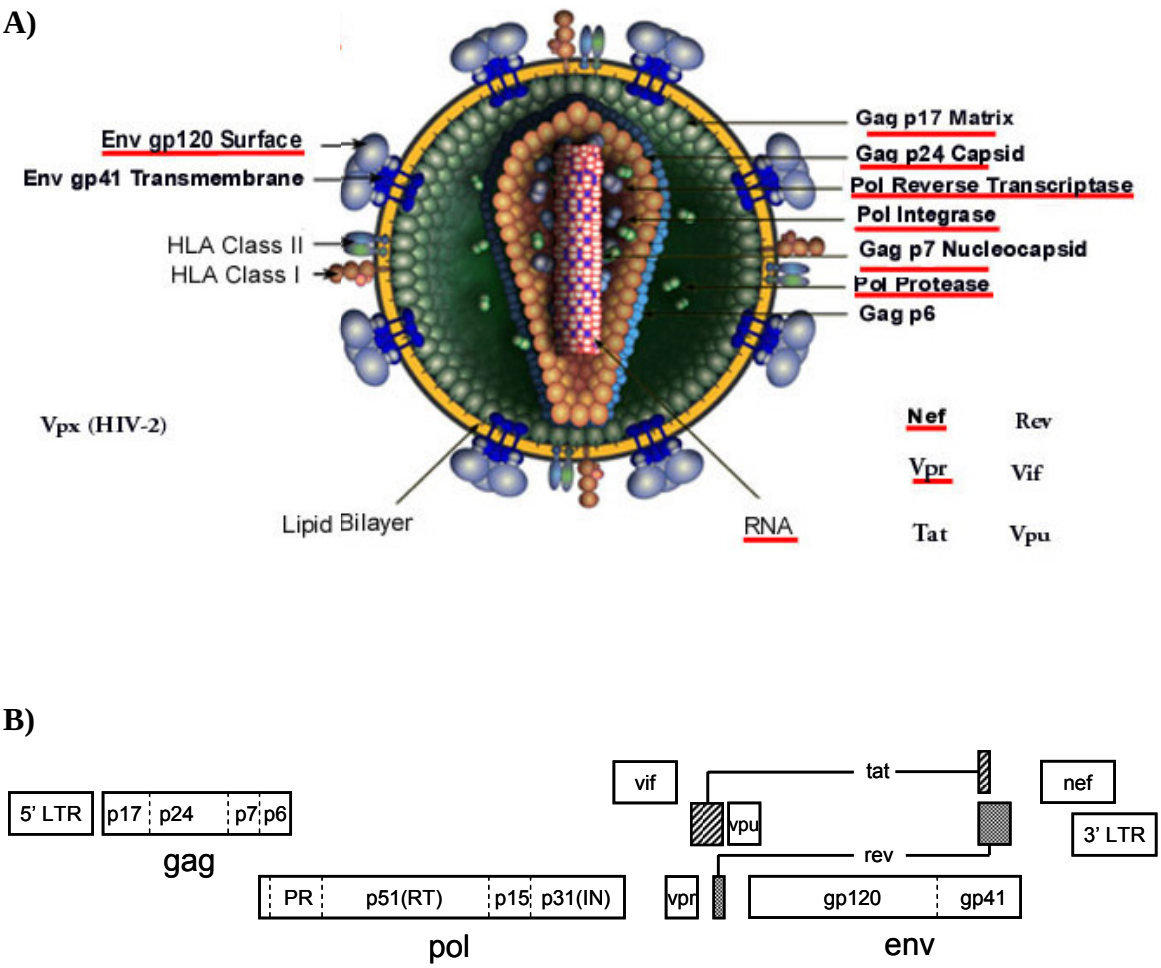

Figure 3. HIV-1 virion structure and genomic organization

A) Mature HIV-1 virion structure. Image was reproduced from a publicly available source from the Los Alamos National Laboratory HIV Database, Los Alamos, New Mexico. www.hiv.lanl..gov. B) HIV-1 genomic structure.

1.4.2 The HIV Life Cycle

HIV enters a cell by binding of its surface gp120 protein to CD4 and a chemokine receptor (mainly CCR5 or CXCR4) on the host cell (2). This binding promotes a conformational change which induces fusion of the virus to the target cell membrane followed by subsequent release of the viral core into the cell. In the cytoplasm the viral RNA is reverse transcribed into cDNA by RT and a host of additional viral and cellular proteins (9). The viral cDNA is then transported to the nucleus where it can become integrated into the host chromosome. Transcription of the viral genome occurs in the nucleus of the host cell mediated by the viral LTR and several host cell transcription factors. Following the synthesis of viral proteins and shuttling of full-length copies of the viral RNA genome to the cytoplasm, the virion is assembled at the plasma membrane where budding occurs (10).

1.4.3 HIV pathogenesis

In 1983 HIV was first identified as the causative agent of AIDS. Acute infection by HIV is characterized by flu-like symptoms which present within 1 to 4 weeks of infection. These symptoms include sore throat, fever, muscle ache, swollen lymph nodes and rash (2). The asymptomatic or chronic phase in which there is low level virus replication and no apparent illness is unique for each individual and can range from 3 to more than 14 years (11, 12). The symptomatic phase generally presents within 10 years after infection and is characterized by high

viral loads, high genetic diversity, increased viral replication fitness, low $CD4^+$ T cells numbers and a variety of immune disorders (11, 13) (Figure 4).

Upon transmission to a new host, HIV targets $CCR5^+$ $CD4^+$ effector memory T cells (14, 15). These target cells are majorly present at mucosal sites such as the gut and cervicovaginal mucosa. This results in massive depletion of these immune cells at these mucosal sites within the first few weeks of acute infection (14-16). Mucosal depletion is soon followed by a state of chronic activation resulting in 1) increased numbers of activated and memory T cells, 2) increased production of proinflammatory cytokines and 3) increased turnover of immune cells (14). During the acute phase, viremia can be very high (10^6 to 10^7 copies/ml) in plasma and peripheral blood (2, 17). Over 10 million to 100 million infected $CD4^+$ cells die per day (2, 18). The $CD8^+$ T cell response is also high during acute infection and its ability to control viremia during the acute phase is usually the best predictor of a long-term asymptomatic period (2).

Virus infection is established once the infection has traveled from the initial site to local lymph nodes. Although a robust immune response is mounted soon after infection with the infiltration of $CD8^+$ T cells and production of proinflammatory cytokines and chemokines, it is generally too late and the infection is already well established (2). Seroconversion normally occurs 1 to 3 weeks after infection and viral set point is reached usually after 3 to 5 months.

At 3 to 6 months after primary infection, $CD4^+$ T cell numbers increase, however they never fully recover to pre-infection levels. This period begins the asymptomatic or chronic phase of infection. The decrease in viremia could be due to an immune response mediated by $CD8^+$ T cell killing of virus infected cells, substrate exhaustion due to the massive loss of target cells or a combination of both. During this time until the symptomatic phase begins (8 to 10 years in a typical progressor), there is persistent low level virus replication in the lymph nodes and

peripheral blood, increased viral genetic diversity and increased viral replicative fitness (Figure 4). Low level virus replication is maintained by an antiviral $CD8^+$ T cell response.

The symptomatic phase is characterized by increased viral load, a decline in $CD4^+$ T cells to less than 200 cells/ul and a deteriorated $CD8^+$ T cell response. In the lymph nodes there is massive breakdown of the tissue architecture and destruction of follicular dendritic cells. All of these events directly precede or coincide with the emergence of more pathogenic fast replicating viruses and CXCR4 coreceptor using viruses (in approximately 50% of cases) and AIDS related malignancies.

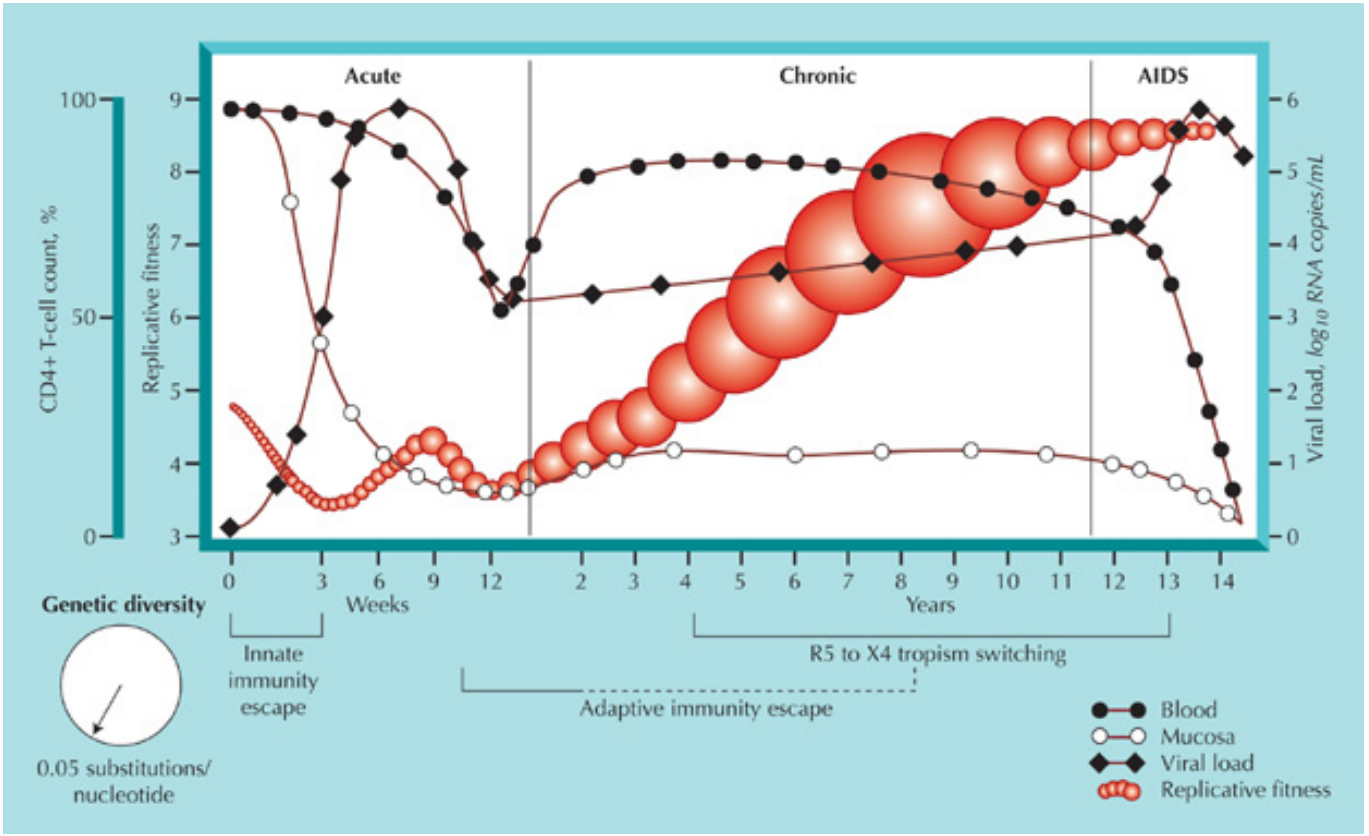

Figure 4. HIV-1 fitness and disease progression model.

Changes in replication fitness and genetic diversity, CD4+ T cell count in the blood and the mucosa and viral loads are shown during HIV disease progression. During the acute phase there is high viral load, rapid depletion of CD4+ T cells in the mucosa and an initial decrease in replication fitness followed by an increase in fitness upon viral escape of the innate immune response. During the chronic phase, HIV adaptive immunity escape occurs and replication fitness increases. All CD4+ T cells are depleted during the AIDS stage, and the viral load increases. Fitness remains high, but HIV diversity is low. Reprinted with permission Current Medicine Group, LLC from Henry KR, *et al*. 2007. The impact of viral and host elements on HIV fitness and disease progression. Curr HIV/AIDS Rep 4(1):36-41.

1.5 HIV-1 IN INDIA

1.5.1 India's HIV epidemic

According to the UNAIDS 2006 report, India – which has a population of about 1.1 billion, has approximately 5.7 million people living with HIV-1. This makes India the second largest HIV-1 infected population in the world next to South Africa. The highest prevalence rates are found in the industrialized south and west in the states of coastal Andhra Pradesh, the Nammakkal district of Tamil Nadu, Mumbai-Karnataka corridor, and the Nagpur area of Maharashtra; and in the northeastern tip of the country in the states of Nagaland and Manipur (3, 19) (Figure 5).

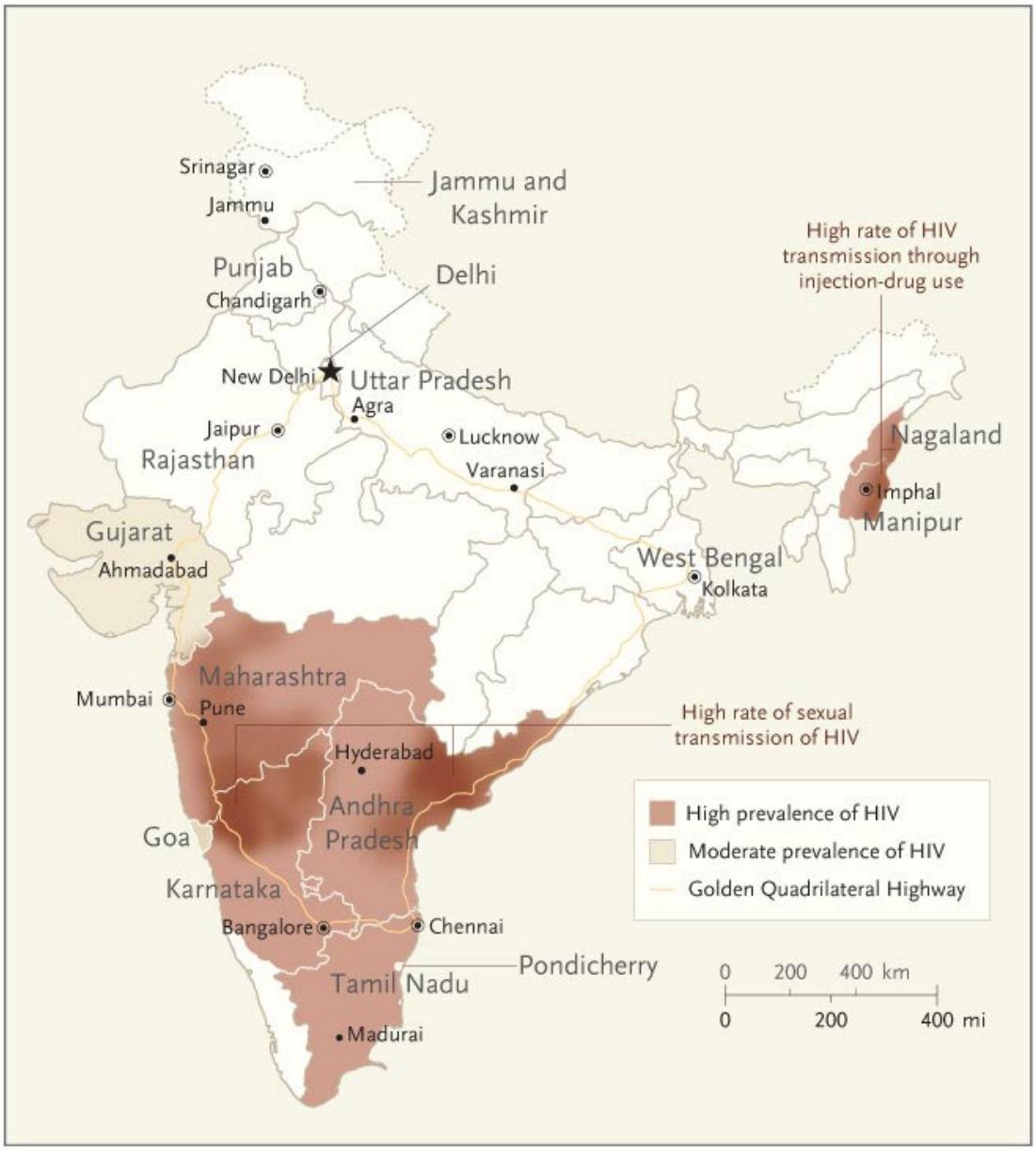

Figure 5. HIV-1 prevalence in India

Map of India showing HIV-1 prevalence and the Golden Quadrilateral Highway. Reprinted with permission from the Massachusetts Medical Society from Steinbrook, R. HIV in India--a complex epidemic. 2007. N Engl J Med 356(11):1089-93. Copyright © 2007 Massachusetts Medical Society. All rights reserved.

The Golden Quadrilateral Highway is an express highway that connects India's four largest cities. It is through this highway network (among other means) that migrant workers, truck drivers and traveling sex workers can contract HIV-1 and spread the virus to different parts of the country (19) (Figure 5). Although prostitution, homosexuality and injection drug use are all illegal in India, these are the major routes of HIV-1 transmission. Sexual contact accounts for roughly 85% of transmissions of the virus (19, 20) a growing proportion of which is from men who have sex with men (MSM) (21). Injection drug use follows as the next major contributor to the spread of HIV-1 in India (3, 19, 22). Additional routes of transmission are perinatally, via breast-feeding and through paid blood donations. Beginning in the early 1990's, intervention projects were started among small groups to bring awareness to sex workers about the modes of transmission of HIV and the importance of condom use, control of STD's and client negotiation skills (20). More recently large scale efforts have been made in the country to educate all sexually active groups about HIV (20, 22). The National AIDS Control Plan III is India's most recent response to their growing HIV epidemic (3). It will focus on scaling up investment on prevention activities among high risk populations, increasing antiretroviral treatment coverage and continued involvement and support of organizations and institutions in both public and private sectors. Although a significant step in the right direction, many obstacles still remain due to cultural, legal and medical factors making HIV prevention and treatment in India a huge challenge (3, 19).

1.5.2 HIV-1 subtypes circulating in India

Genetic analyses of HIV-1 circulating in different parts of India have shown that the predominant proportion of HIV-1 circulating in India is of subtype C origin with a small fraction made up of subtypes A and B (13, 23-28). Further genetic analysis has also shown that HIV-1 subtype C from India shares homology to subtype C strains from Zambia and South Africa, however most C3 (Indian) strains were more closely related to each other than to subtype C from other countries (27, 29, 30). A recent study compared subtype C sequences from India to subtype C sequences sampled from Botswana, Burundi, South Africa, Tanzania, and Zimbabwe. Overall, HIV-1 type C sequences from different parts of India were more closely related to each other (10%) than to subtype C sequences from Botswana, Burundi, South Africa, Tanzania and Zimbabwe (15-21%) (13) (Figure 6). These results indicate that subtype C sequences in India are distinct from subtype C sequences sampled from other countries. Additionally, several studies suggest a recent introduction of HIV-1 subtype C and HIV-2 into India from a similar set of founder strains or ancestor viruses (29, 31, 32) implying a more rapid spread of HIV-1 subtype C throughout the country.

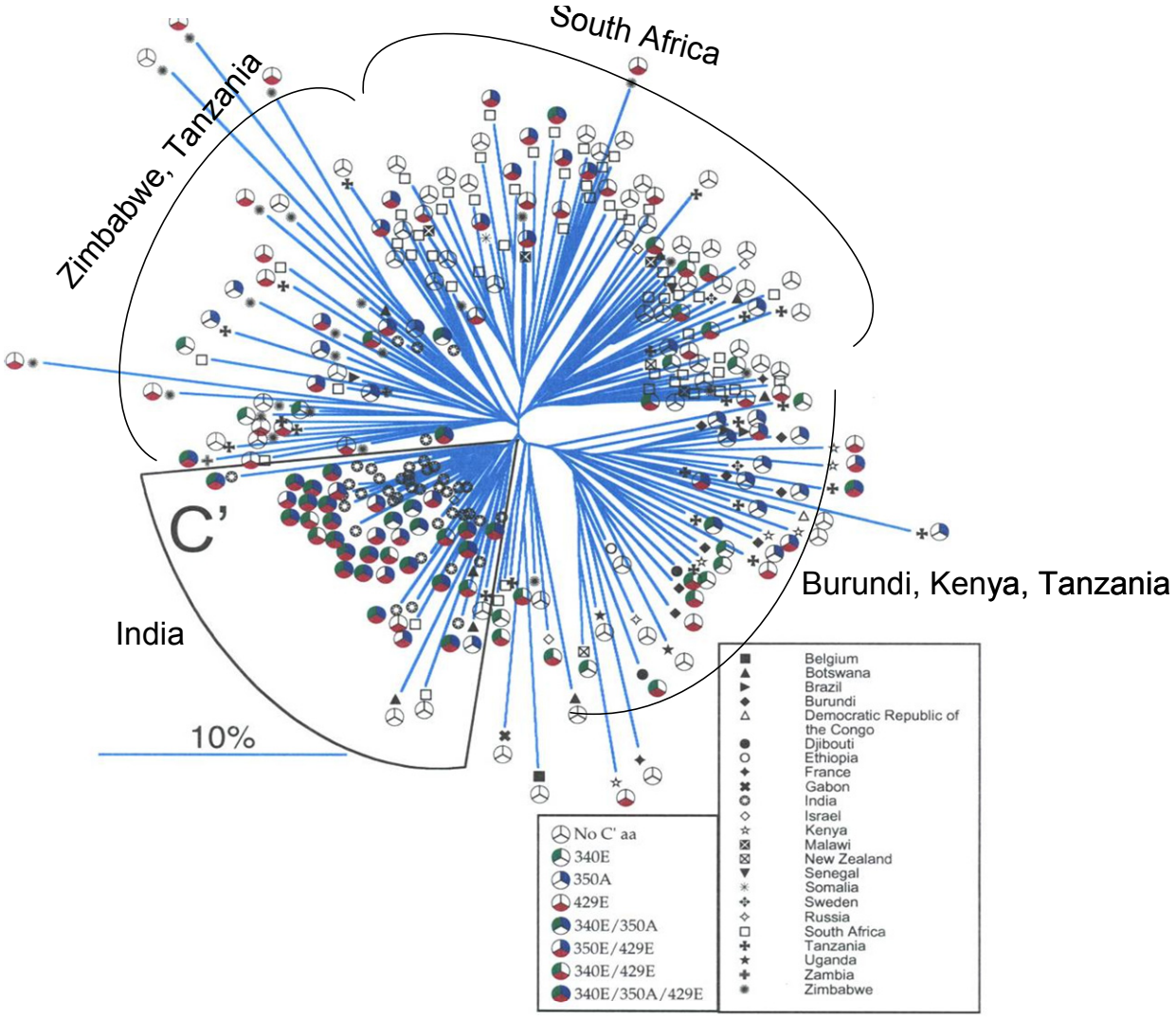

Figure 6. Phylogenetic relationships among HIV-1 subtype C *env* sequences sampled from different countries.

Neighbor joining analysis using 192 sequences encoding the V3-V4 region. The presence of C_{IN} signature amino acids are indicated by colored circles. Clustering of select countries are indicated. Reprinted with permission from the American Society for Microbiology from Shankarappa, *et al*. 2001. Human immunodeficiency virus type 1 env sequences from Calcutta in eastern India: identification of features that distinguish subtype C sequences in India from other subtype C sequences. J Virol 75(21):10479-87.

Reports of subtype A and B viruses in India date back to 1991 and 1994, respectively (28). Since then these subtypes have been reported at a much lower frequency than subtype C although they have been transmitted via the same primary routes, i.e. sexual contact and injection drug use (13, 33, 34). This taken together with the fact that HIV-1 isolated from different parts of India at different times are closely related (13, 35, 36) suggest that the preponderance of subtype C viruses over other subtypes is most probably not due to continual introductions of HIV-1 subtype C into the country or to recent immigration or representative of a cluster of isolated individuals.

Over the course of the HIV-1 epidemic in India, and regardless of geographical area, subtype C has consistently accounted for 90-95% of infections with subtypes A, B and others accounting for the remaining 5-10% (27, 28, 31, 37, 38). It is possible that the disproportionate distribution of HIV-1 subtypes in India may be due to one or a combination of factors including founder effect, replication fitness and transmission efficiency.

1.6 CONCEPTS IN VIRAL FITNESS

This section introduces evolutionary and fitness theories, defines fitness as it relates to RNA viruses, describes assays used to measure fitness and describes in vitro replication fitness as it relates to this study.

1.6.1 Concepts in Viral Fitness

Distinctive features of RNA viruses such as HIV include high genetic variation as a result of mutations caused by error-prone replication machinery and a lack of repair and proofreading mechanisms, high yields of virions and short replication times (39, 40). Due to these unique traits, RNA viruses have frequently been used to test both evolutionary and fitness theories (39-45). Among the principles tested with RNA viruses are the Red Queen hypothesis and the competitive exclusion principle (using vesicular stomatitis virus, VSV) and Muller's ratchet (using VSV followed by foot-and-mouth-disease, FMDV). The Red Queen hypothesis states that viral quasispecies that are cocultured in the same environment tend to gain fitness with each viral passage, although one eventually overgrows the other (42). The competitive exclusion principle states that when two species coexist in the same environment with limited resources, one will always outgrow the other (40, 46, 47). Finally, changes in environment such as viral passage in tissue culture, plaque-to-plaque transfers or transmission events which alter or decrease population size, may create genetic bottlenecks which can decrease gains in fitness or cause an overall fitness loss. In such cases the Muller's ratchet hypothesis states that decreases in average

fitness due to an irreversible gain of deleterious mutations in limited populations cannot be compensated for and will overwhelm the appearance of mutations improving fitness (40, 41, 46, 47). Initially HIV was not used as a tool to test these fitness theories. However, more recently there have been increasing data using HIV variants in growth competition assays to study in vitro replication fitness dynamics. Although several implications have been made, the relationship between in vivo replication capacity (including disease progression, transmission and spread) and in vitro fitness remains unclear.

1.6.2 Fitness: Definition and Influential Factors

Fitness is an evolutionary term used to describe the ability of an organism to reproduce and adapt to its particular environment (39). For RNA viruses, fitness can be estimated by the relative ability to produce stable infectious progeny in a given environment (39). HIV fitness can be affected by any combination of host or viral factors. For example, the first obstacle encountered by the virus is upon sexual transmission in which a severe bottleneck selects for those variants that use the CCR5 coreceptor (11). This event reduces the population size from donor to recipient (48) resulting in a more homogeneous population with a narrower genetic distribution of quasispecies as compared to the donor (49). Thus, as stated for Mueller's ratchet evolutionary theory, continued bottleneck events can result in an overall fitness loss to the viral population. Additional host factors affecting viral fitness include mutations or polymorphisms in the CCR5 coreceptor that could affect transmission or disease progression in the host and selective pressure imposed by HIV-specific CTLs which could select for escape variants with reduced fitness (11). Viral factors affecting fitness include the high mutation rate of HIV which greatly contributes to the adaptability of the virus which can result in overall fitness gains (11). In addition, differences

in *in vitro* replication fitness have been shown to be due to efficiency of binding related to the gp120 region of the envelope gene. The *nef* gene is another viral factor that may influence fitness since it has been reported that some long term nonprogressors harbored HIV strains having *nef* deletions (50).

1.6.3 Growth Competition Assays as a Tool to Measure Fitness

Ex vivo fitness assays are valuable because they focus solely on replication efficiency and eliminate selective pressures by the human host on the infecting virus - which may vary from one host to another (47). These assays can discern small differences in fitness as opposed to side-by-side growth kinetic assays that may only distinguish gross changes in replicative fitness (49). Furthermore, dual infection assays provide the internal control that cannot be ensured in separate monoinfections and relative fitness can be directly compared since one clone will eventually outgrow the other (39, 44, 47). Ex vivo (in vitro) fitness has been shown to correlate with disease progression and therefore can be a useful predictor of disease progression (51).

1.6.4 Relationship Of in vitro Replication Fitness to HIV-1 Disease Progression and Global Distribution

It is not fully understood how in vitro replication fitness and transmission efficiency relate to the in vivo situation or the current asymmetric distribution of HIV-1 subtypes globally, but as new data continues to support this hypothesis it further strengthens the idea that genetic variability within subtypes may be correlated with in vivo fitness and HIV population dynamics. The seminal study examining this issue came in 2000 where using an in vitro dual infection growth

competition assay, Quinones-Mateu and colleagues showed that there was a correlation between *ex vivo* (*in vitro*) fitness and disease progression. In this study, HIV-1 isolates from progressive patients were shown to be significantly more fit than HIV-1 isolates from long-term survivors – in several cases independent of viral phenotype. It was also shown that a subtype C primary isolate from Brazil was outcompeted and thus less fit than both a Rwandan subtype A primary isolate and a laboratory adapted subtype B strain (51). In a more recent study, the *ex vivo* fitness of CCR5-tropic HIV-1 isolates of subtypes B and C were compared by performing pair-wise competitions (52). All subtypes C isolates were outcompeted by subtype B isolates in all cell types except Langerhans cells. In these cells, subtype C demonstrated competitive replication efficiency against subtype B indicating that transmission between these two subtypes may be similar. These findings suggest that slower disease progression to AIDS (reduced viral fitness) coupled with efficient transmission may have contributed to the current subtype C global distribution.

Recently, an order of relative fitness was assigned to HIV viruses: HIV-1 group M > HIV-2 >> HIV-1 group O (53). In this study in vitro growth competition assays were used as a tool to evaluate fitness differences among HIV groups and types and suggests that the lower replicative capacity of group O and HIV-2 compared to group M may have led to decreased transmission and distribution of these groups in the human population. In this scenario, a drastically lower replicative capacity may be impairing transmission of the virus due to low viral loads, resulting in reduced spread of the less fit variant. Another recent study found that CRF02_AG isolates had higher in vitro replication capacities in PBMCs than subtype A and G viruses from the same geographic region. The higher fitness was independent of coreceptor tropism and irrespective of high or low $CD4^+$ T cell counts (54, 55). Over the last 10 years

CRF02_AG has become the predominant subtype in West and West Central Africa. While this current state may be due to founder effect, these findings along with others previously mentioned suggest that altered replicative capacity acquired by intersubtype recombination events could have contributed to the disproportionate spread of CRF02_AG.

However, since heterosexual transmission in India is more closely related to subtype C viruses, it is possible that HIV-1 subtype C viruses may posses enhanced fitness properties ideal for mucosal entry and establishment of infection as compared to other subtypes present in India. This argument is further supported by the fact that the length of time to progression to AIDS appears to be shortest in women infected with subtype C as compared to those infected with subtype A, D or G (56). In fact, women infected with subtype C appear to be 15 times more likely to develop AIDS than those infected with subtype A (56, 57) and women infected with subtype C virus have been associated with increased vaginal shedding compared to infection with subtypes A or D (58).

1.7 SEXUAL TRANSMISSION OF HIV-1

This section describes the mechanisms by which HIV-1 is heterosexually transmitted and describes relevant findings on the relationship between HIV-1 subtypes and sexual transmission.

1.7.1 Sexual Transmission of HIV-1

Heterosexual intercourse is the primary route of HIV-1 transmission and accounts for 80% of new infections worldwide (59). According to the UNAIDS 2006 report, 17.7 million women are infected with HIV. This accounts for nearly half of infected individuals living with HIV. The majority of these women live in Sub-Saharan Africa and Asia where the epidemic is rampant. The probability of transmission for each heterosexual encounter is low (about 0.001) and is related to dose, i.e. the amount of virus present in genital fluids (60-62) and disease stage (2). Strains that use the CCR5 coreceptor are preferentially transmitted over strains that use CXCR4. This is likely due to the host induced bottleneck in which epithelial cells selectively captures and transfers R5 viruses (61, 63). This is supported by the fact that individuals that posses a deletion in CCR5 are generally resistant to infection (64, 65) and those that contain a mutation in the CCR5 allele progress to AIDS more slowly than those that have wild type CCR5.

In vivo, the lower female genital tract is comprised of different anatomical regions including the vaginal mucosa, the ectocervix and the endocervix (66). The ectocervix and vaginal mucosa contain several layers of non-keratinized stratified squamous epithelial cells forming a barrier to block infection. The squamocolumnar junction or cervical transformation zone is the region in

which the multilayered ectocervix transitions to a single layer of columnar epithelium that is the endocervix (62, 66). The topology of the endocervix likely makes it more vulnerable to HIV-1 infection (67). The target cells for HIV virus in the underlying cervicovaginal submucosa are 'resting' and activated $CD4^+$ T cells, macrophages and dendritic cells (68) (Figure 7). It is still unclear how HIV manages to infiltrate an intact mucosal surface since when intact, the epithelial layer should serve as a suitable barrier to block infection due to vaginal mucus, which can trap and dilute virus; low pH and antimicrobial peptides like cytokines and chemokines, which can block infection (69). However STD's, inflammation or trauma or microscopic breaks in the epithelial surface could allow entry of the virus and exposure of susceptible cells in the submucosa. There are four potential mechanisms by which HIV can penetrate the mucosal surface: 1) transcytosis through epithelial cells or specialized M cells; 2) epithelial transmigration of infected donor cells; 3) association with dendritic cells that can both capture virus in the lumen via C-type lectins (61) or themselves become infected and transfer the virus to susceptible cells in the lamina propia and lymph nodes; and 4) through physical breaks in the epithelial barrier (2, 61, 62) (Figure 7).

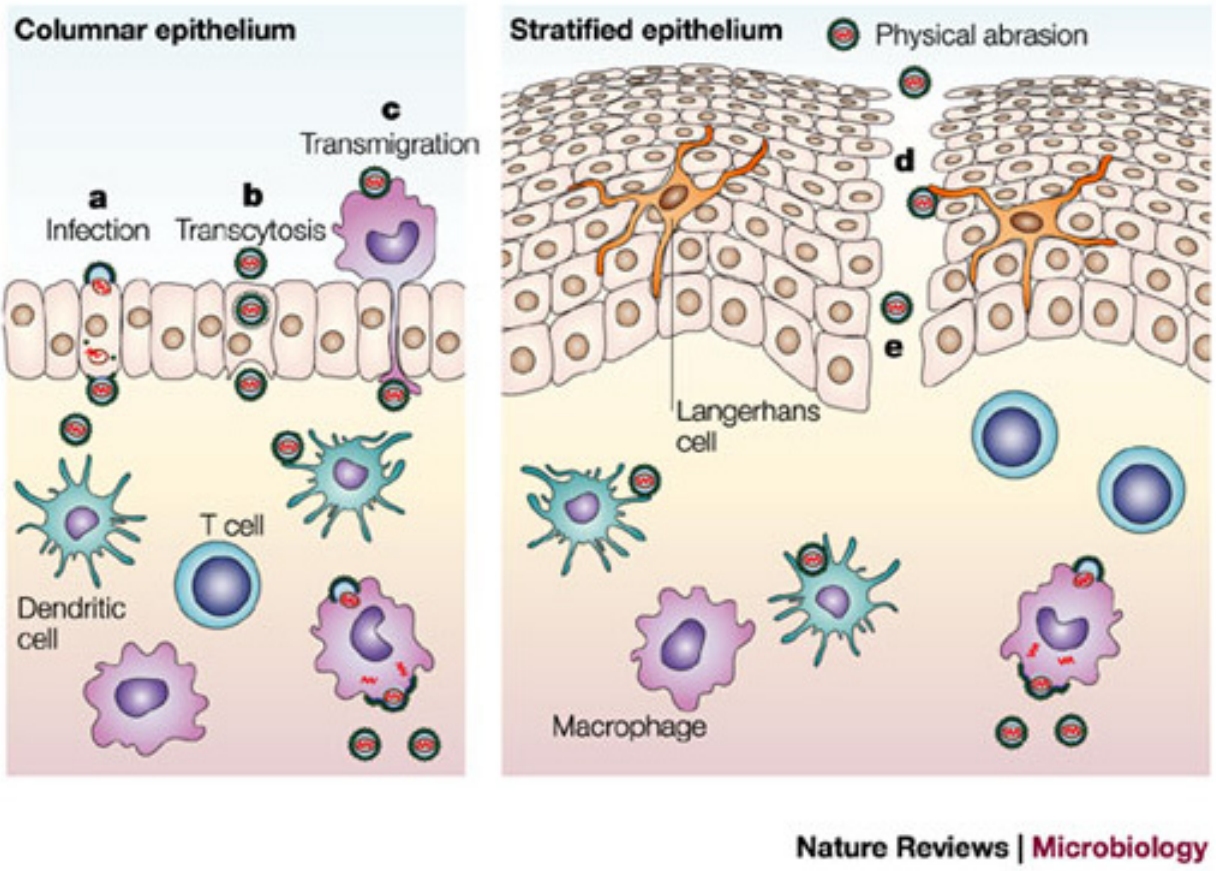

Figure 7. Potential mechanisms for HIV-1 transmission across the mucosal epithelium

a) Direct infection of epithelial cells. b) Transcytosis through epithelial cells and or M cells. c) Epithelial transmigration of infected donor cells. d) Uptake by intraepithelial Langerhans cells. e) Crossing of the epithelial barrier through physical abrasions. Reprinted by permission from Macmillan Publishers Ltd. Nat Rev Microbiol. Shattock RJ and Moore JP. Inhibiting sexual transmission of HIV-1 infection. 2003;1(1):25-34, copyright 2003.

Once the virus has successfully penetrated the mucosal epithelial barrier, it now has access to underlying susceptible cells. Using a modified organ culture system, Hladik *et al*. (70) observed that intraepithelial memory T cells expressing CD4 and CCR5 became infected within two hours of virus exposure. HIV was also found associated with Langerhans cells, however no interaction was observed at this early timepoint between LCs and T cells. This suggests that HIV infection of T cells is not dependent on LC-mediated viral uptake or enhancement of infection in trans (70, 71). Upon emigration from the epithelium several interactions between LCs, T cells and DC's have been suggested for the propagation of infection. First, LCs may be involved in transfer of virus to T cells across an "infectious synapse". Second, LCs could transfer virus to DC's in the stroma and finally, stromal DC's could acquire virus from infected T cells (71). Using a cervical tissue derived organ culture, our laboratory has shown that activated memory CD4+T cells are the first cell types that become infected within 6 hours of infection. Infection of macrophages and DCs were detected 24-96 hr after infection (72).

Using a rhesus macaque simian immunodeficiency virus (SIV) model, Miller *et al*. (73) observed that in the first three days post-intravaginal infection there was limited penetration of the mucosal barrier despite having used very high titer virus inoculum. They also observed that within these first 72 hours, the infection was confined to a small focus of founder populations of infected cells and that there was rapid dissemination to distal sites from this portal of entry. These founder populations served as a continuous source of virus from which distal sites and lymphatic tissue compartments could establish productive infection. Productive systemic infection was observed within six to ten days post infection. These studies using ex vivo organ cultures and the rhesus macaque model have been instrumental in our understanding of target

cells involved in HIV transmission and the extent and timing of HIV-1 infection and dissemination. Furthermore, we can draw relevant conclusions from the macaque model since target cells, physiology and immunology of the macaque female genital tract have all been shown to be similar to humans (73).

1.7.2 Role of Subtypes in HIV- Transmission

The association between HIV subtype and mode and efficiency of transmission has been under investigation since early on in the global HIV epidemic. Some studies have reported an association between subtype B and homosexual practices (74, 75). Whereas subtype E was identified in the majority of cases where there was heterosexual contact (76). Additionally a South African study found a highly significant correlation between subtypes B and C in homosexual and heterosexual contacts, respectively (75). The majority of infections by HIV are through heterosexual intercourse. The subtypes of HIV-1 that appear to spread most successfully by heterosexual transmission may have biologic features that contribute to efficient heterosexual transmission. Recently an order of relative fitness was assigned to HIV viruses based on in vitro growth competition assays in PBMC and a dendritic cell – T-cell transmission model system (77). The order (group M > HIV-2 >> group O) is consistent with the current global subtype distribution and suggests a correlation between in vitro and in vivo fitness, replication capacity and transmission.

There is still debate over which cells are the initial targets of HIV upon heterosexual exposure. Langerhans cells (52, 78-80), dendritic cells (81), macrophages (82) and activated CD4+ T-cells (70, 72, 83) have all been implicated. Nevertheless, Langerhans cells have been the primary cell used in HIV-1 transmission model systems. Langerhans cells are immature

phagocytic dendritic cells found under most surface pluristratified epithelia such as the female cervix and vagina (84). These cells have been proposed to be primary target cells following sexual exposure to HIV (52, 74, 78, 79). Recent studies have shown that although it is outcompeted by other subtypes in PBMC, subtype C competes efficiently in Langerhans cells (52, 77). Soto-Ramirez *et al*. (74) reported that HIV-1 subtypes E and C strains from Thai and Indian heterosexuals, respectively, replicated in Langerhans cells more efficiently than subtype B strains from US homosexual men (74). Therefore Langerhans cells in the foreskin of the penis and the cervicovaginal epithelium could be very important for transmission and dissemination during sexual transmission of HIV-1. Recently primary Langerhans cells isolated from cervical tissues were shown to be refractory to HIV-1 infection by way of Langerin which may act as a natural barrier to HIV-1 infection (166). More studies are necessary to determine the role of Langerhans cells in HIV-1 infection and to determine the degree of subtype-specific replication in these cells. The efficiency of replication in Langerhans cells may imply more efficient transmission across the epithelial surface. Subtype-specific replication efficiency differences in Langerhans cells could result in subtype-specific differences in transmission efficiency and dissemination throughout a population.

1.7.3 Models for the Study of HIV-1 Transmission

The development of model systems to study HIV-1 transmission has been very useful to investigate the early molecular and cellular events involved in this process. For example, primary and transformed genital epithelial cells have been used to investigate viral (85) and cellular (66) transmigration, respectively. To identify initial targets of HIV-1 infection, tissue sections and isolated cell cultures (86), epidermal sheets (87) and ex vivo organ culture models (70, 88-90)

have been used. A more comprehensive picture of the virological and immunological events surrounding HIV-1 transmission has been given by the animal model. Rhesus macaques can efficiently transmit SIV or SIV/HIV chimeras by vaginal inoculation and endure a natural course of infection similar to what is observed in humans. Using animals, investigators have been able to examine early and late events after intravaginal exposure and better understand viral propagation and systemic dissemination (59, 91). While animal models hold the most promise, they are generally expensive to use on a routine basis and could potentially yield misleading information regarding heterosexual transmission of HIV-1 in humans (92).

As discussed in section 1.7.2, a limited number of studies have investigated the differences in replication fitness or transmission efficiency of HIV-1 subtypes in cells from the vaginal mucosa. This study is novel in that it uses tissues rather than a monolayer of cells to determine replication fitness differences. Our ex vivo organ culture model most closely represents the in vivo situation in humans since it provides the natural tissue architecture of the female genital organ, including epithelial cells, submucosa and immune cells, such as T cells, macrophages and Langerhans cells (88).

1.8 ROLE OF LTR AND ENVELOPE GENE IN HIV-1 REPLICATION FITNESS

The long terminal repeat (LTR) of HIV regulates the expression of the viral genome through its interaction with both cellular and viral transcription factors (93). The HIV-1 LTR is approximately 640 base pairs long and is divided into three structural regions designated, U3, R and U5 (Figure 8). The U3 region is further subdivided into the modulatory, enhancer and core/promoter functional regions (93, 94). The U3 region is of particular interest because it

contains binding sites for several key transcription factors such as NF-κB, Sp-1, NF-AT, USF and TCF1-α, which regulate HIV transcription (94, 95).

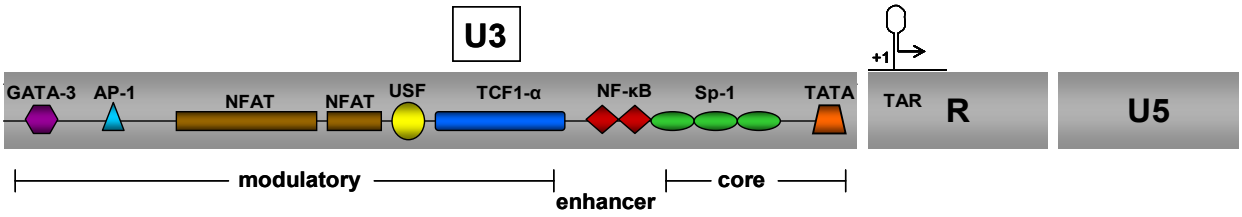

Figure 8. Schematic diagram of the structure of the HIV-1 LTR.

The binding sites for select cellular transcription factors are shown. The transcription start site (+1) is indicated.

There is a growing body of data indicating subtype specific differences of the HIV-1 LTR at both the molecular and functional levels (95-99). Comparisons of HIV-1 subtypes show differences in the number and organization of transcription factor binding sites within the promoter region (100). The subtype C LTR in particular has been shown to contain three NF-κB binding sites instead of two that are carried by the majority of subtypes (101-103). This additional NF-κB enhancer site in subtype C has been correlated with increased promoter activity and increased viral replication (96, 98, 104, 105). A study by van Opijnen *et al*. , showed that differences in viral replication and fitness rates were determined by the HIV-1 subtype-specific LTR and was dependent on both the host cell type and activation status of the cell (98). In another study by Centlivre *et al*., Chinese rhesus macaques were co-infected with subtype B, C and E short terminal repeat cloned virus which consisted of the HIV-1 subtype-specific LTR core promoter/enhancer region in a neutral SIVmac239 backbone. This study showed that viral replication was HIV-1 LTR subtype-dependent and suggests that subtype C is particularly adapted to sustain viral replication in primary viremia. This conclusion was based on the observation that there was a strong predominance of the subtype C chimera during primary

infection followed by the dominance of subtype B in all tissues later in infection. Taken together these data suggest that HIV-1 LTR subtype variability may influence virus replication in a host cell specific manner and at different stages of disease. Therefore the higher replicative fitness of HIV-1 subtype C early in infection may generate higher viral loads during the asymptomatic period compared to other subtypes which could promote viral transmission. It is therefore relevant to investigate whether these same conclusions can be made regarding HIV-1 subtype C in India. This may help to explain the predominance of this subtype in India today and throughout the HIV epidemic.

The envelope gene (*env*) of HIV-1 is responsible for viral infectivity, co-receptor usage, and pathogenesis (in vitro). For example, changes in amino acids within the V3 loop of gp120 are correlated with coreceptor usage and disease progression. Basic amino acids at positions 11 and 25 in the V3 region are correlated with CXCR4 coreceptor usage, rapid disease progression and syncytium formation in vitro. Viruses with this genotype usually emerge late in the chronic/symptomatic stage of disease and correspond with a rapid decline in $CD4^+$ T cells and progression to AIDS. Whereas an absence of basic amino acids at these same positions, are correlated with CCR5 coreceptor usage and slow disease progression. Viruses of this genotype are principally found after initial infection and in the asymptomatic phase of disease.

High mutation frequency caused by the viral reverse transcriptase coupled with selective pressures of the host environment have caused the envelope gene to exhibit the highest variability (5-25% nucleotide diversity) within the genome among HIV-1 subtypes. Considering the important functions of *env* together with its high sequence diversity has led investigators to speculate that this gene may be one of the factors involved in determining replication capacity and fitness of the virus. Several studies have proposed that viral entry mediated by *env* is the

major determinant of fitness (52, 106, 107). Rangel *et al*. (106), and Marozsan *et al*. (107), both used chimeric *env* viruses to show that HIV-1 fitness was e*nv* dependent. Furthermore, Ball *et al*. showed that in dual competition experiments between HIV-1 isolates of different subtypes, the more fit variant was determined within 8-24 hours and this observed difference in fitness occurred at the level of entry rather than other steps in the retroviral life cycle such as reverse transcription, integration or viral mRNA transcription. Based on these findings, Marozsan *et al*, (107) extended these studies and determined that the ability of one isolate to outcompete another was due to the efficiency of host cell entry (specifically binding and fusion) mediated by the gp120 coding region of e*nv*. These data suggest that HIV-1 subtype fitness differences occur at the level of viral entry and are due to diversity within the envelope gene.

Therefore Indian subtype C may be more replication fit than Indian subtype A due to subtype specific variability within the LTR or to a higher rate of infectivity mediated by *env*. This fitness advantage of subtype C over subtype A may help to explain the asymmetric distribution of subtypes of HIV-1 in India.

2.0 HYPOTHESIS AND SPECIFIC AIMS

Rationale

The dominance and disproportionate spread of HIV-1 subtype C compared to other subtypes in India suggests that subtype C may possess a biological growth advantage making it more replication efficient and/or more efficiently transmitted across the mucosal surface. Ex vivo fitness assays are valuable tools for exploring such biological differences because they focus solely on replication efficiency and eliminate selective pressures by the human host on the infecting virus-which may vary from one host to another (47). Furthermore, ex vivo fitness has been shown to correlate with disease progression and thus be a useful predictor of disease progression (51). There is a growing amount of data which suggest that fitness differences among HIV-1 subtypes are dependent on the envelope gene. Chimeric constructs of HIV-1 subtype envelope genes in neutral backbones have demonstrated differences in replicative capacity, fusion and competitive binding, all of which were shown to be dependent on *env*. The LTR of HIV-1 has also been proposed as the region responsible for determining differences in subtype fitness. Studies have reported that variations in LTR transcription factor binding (TFBS) sites among HIV-1 subtypes resulted in altered transcription efficiency in the presence and absence of stimuli and in different cell types.

Hypothesis

The asymmetric distribution of HIV-1 subtype C in India is due to enhanced replication fitness and/or enhanced transmission efficiency of this subtype over other subtypes present in India.

Specific Aims

The overall objective of this proposal is to elucidate the mechanism of asymmetric distribution of HIV-1 subtypes in India. Specific aims of this study are:

1) To construct and characterize an infectious molecular clone of HIV-1 subtype A of Indian origin.
2) To evaluate the replication fitness of HIV-1 subtypes A and C from India using an in vitro growth competition assay
3) To compare the transmission efficiency of HIV-1 subtypes A and C from India across the mucosa of cervicovaginal tissue.
4) To evaluate the roles of the LTR and e*nv* in replication fitness and transmission efficiency across cervicovaginal tissue.

3.0 CHAPTER ONE. CONSTRUCTION AND CHARACTERIZATION OF AN INFECTIOUS MOLECULAR CLONE OF HIV-1 SUBTYPE A OF INDIAN ORIGIN

3.1 PREFACE

This chapter is adapted from a published study (Milka A. Rodriguez[a] , Yue Chen[a] , Jodi K. Craigo[b], Ramdas Chatterjee[c] , Deena Ratner[a] , Masashi Tatsumi[d] , Pratima Roy[e] , Dhruba Neogi[e] and Phalguni Gupta[a]. 2005. Virology. 345:328-336.). Work described in this chapter is in fulfillment of specific aim 1.

[a]Department of Infectious Diseases and Microbiology, Graduate School of Public Health, University of Pittsburgh, PA 15261, USA

[b]Department of Molecular Genetics and Biochemistry, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15261

[c]Department of Virology, Chittaranjan Cancer Research Institute, Calcutta, India

[d]Department of Veterinary Science, National Institute of Infectious Diseases, Toyama, Shinjuku-ku, Tokyo, Japan, 162-8640

[e]Department of Virology, School of Tropical Medicine, Calcutta, India

3.2 ABSTRACT

India has the second highest number of HIV-1 infected people next to South Africa. The predominant proportion of HIV-1 circulating in India is of subtype C origin, with a small fraction made up of subtypes A and B. In this report, we describe the construction and characterization of the first full length infectious molecular clone p1579A-1 HIV-1, from an HIV-1 subtype A infected person from India, using long PCR and successive ligation of the amplimers. Phylogenetic analysis of the sequence of the entire proviral DNA and LTR confirmed p1579A to be an HIV-1 subtype A. Analysis of the envelope gene of p1579A-1 showed a conserved GPGQ motif and the absence of basic amino acids at positions 11 and 25 suggesting CCR5 co-receptor usage. Analysis of *env* N-linked glycosylation sites revealed fewer sites in the V1 region of envelope compared to other subtype A. Transcription factor binding site analysis of the LTR sequences identified conserved as well as unique transcription factor binding sites (TFBS) in p1579A-1. This infectious clone of HIV-1 can be useful to study the molecular mechanism of dominance of subtype C in India.

3.3 INTRODUCTION

HIV-1 was first reported in Tamil Nadu, India in 1986 among female commercial sex workers (108, 109). Since then, HIV-1 has spread rapidly throughout the country and has been reported in nearly every major city. Today, India has the second largest number of HIV-1 infected individuals in the world next to South Africa (3). Genetic analyses of HIV-1 sequences circulating in different parts of India have shown that the predominant proportion of HIV-1 circulating in India is of subtype C origin with a small fraction of subtypes A and B (13, 23-28). Furthermore, phenotypic and genotypic analysis of *env* sequences of HIV-1 in India indicate that they are CCR5 tropic with non-syncytium inducing phenotype even when they are isolated from late stage of infection (110). We have recently analyzed subtype C sequences from India and then compared this set of sequences to subtype C sequences sampled from Botswana, Burundi, South Africa, Tanzania, and Zimbabwe. Overall, HIV-1 type C sequences from different parts of India were more closely related to each other (10%) than to subtype C sequences from Botswana, Burundi, South Africa, Tanzania and Zimbabwe (15-21%) (13). These results indicate that subtype C sequences in India are distinct from subtype C sequences sampled from other countries.

Over the course of the HIV-1 epidemic in India, and regardless of geographical area, subtype C has consistently accounted for 90-95% of infections with subtypes A, B and others accounting for the remaining 5-10% (27, 28, 31, 37, 38). This together with the fact HIV-1 isolated at different parts of India at different times are closely related (13) suggests that, the preponderance of subtype C viruses over other subtypes is most probably not due to new and continual introductions of HIV-1 subtype C into the country or to recent immigration or

representative of a cluster of isolated individuals. It is possible that the disproportionate distribution of HIV-1 subtypes in India may be due to one or a combination of factors including founder effect, replication fitness and transmission efficiency. In order to elucidate the molecular dynamics of HIV-1 subtype distribution in India it is important to have an infectious molecular clone, such as subtype A, from India. Reports of subtype A viruses in India date back to 1991 (28) and since then they have been periodically reported in various studies (13, 33, 34). To date there are no sequence data of the complete genome of an HIV-1 subtype A nor does a replication competent infectious molecular clone of subtype A exist. Here we report the construction of a full-length infectious molecular clone of subtype A HIV-1 from India and delineation of its complete genomic sequence.

3.4 MATERIALS AND METHODS

3.4.1 Virus Cultures

PBMC from 19 HIV-1 positive Indian patients were cultured in the presence of phytohemagglutin (PHA)-stimulated, CD8-depleted normal donor PBMC in RPMI 1640 media supplemented with IL-2 as described previously (111, 112). Production of virus was monitored by HIV-1 p24 production in the culture supernatant. Virus was harvested every 5 days for 40 days. 10-20 million cultured cells were pelleted and stored at -80°C for DNA isolation and subtyping.

3.4.2 DNA Isolation and HIV-1 Env Subtyping by HMA

High molecular weight chromosomal DNA was purified from cultured cell pellets using PUREGENE DNA purification kit. The DNA was then subjected to a heteroduplex mobility assay (HMA), as described previously (31). ED5/ED12 primers were used to PCR amplify the V1-V5 region of the envelope gene followed by a nested PCR amplification of the C2-V5 region using primers DR7/DR8. The cycling conditions for both primer pairs were as follows: 3 cycles of 94°C for 1 minute, 55°C for 1 minute, 72°C for 1 minute. Followed by 32 cycles of 94°C for 15 seconds, 55°C for 45 seconds, 72°C for 1 minute. Final extension at 72°C for 5 minutes. DR7/DR8 PCR products from each of the reference plasmids and the unknown HIV-1 were then combined, denatured at 99°C for 5 minutes and then reannealed on wet ice for 10 minutes.

Heteroduplexes were resolved on a 5% nondenaturing polyacrylamide gel at 250V for 3 hours. The gel was then stained with ethidium bromide and visualized in a phosphoimager.

3.4.3 Construction of the Full Length Infectious Molecular Clone

The entire 1579A genome was amplified in three fragments. Long PCR was performed using the high fidelity AccuPrime Pfx DNA polymerase (Invitrogen) and 1ug of chromosomal DNA from HIV-1 1579A infected cells. Three overlapping sub-genomic fragments of 2.5kb, 5'LTR-*gag*, 6.2kb, *gag*-*env*, and 2.7kb, *env*-LTR3', were generated using PCR primers shown in Table 1. The cycling conditions were as follows: 1 cycle of 95°C for 2 minutes. Followed by 10 cycles of 94°C for 10 seconds, 55°C for 30 seconds and 66°C for 2 minutes. Followed by 25 cycles of 94°C for 15 seconds, 55°C for 30 seconds and 66°C for 2 minutes. Final extension at 66°C for 7 minutes. Extension times were increased by 5 seconds every additional cycle. Annealing temperatures for 2.5kb, 6.2kb and 2.7kb were 50°C, 55°C and 57°C, respectively. PCR products were gel purified using Mo Bio Laboratories DNA purification kit and then cloned into pCR-Blunt II TOPO vectors (Invitrogen) resulting in plasmids A, B and C, respectively. The plasmid A carrying the 2.5kb HIV-1 DNA fragment and the plasmid B carrying the 6.2kb HIV-1 DNA fragment was digested with unique restriction enzymes. The 4.4kb fragment from the plasmid A was ligated to the 6.7kb fragment from the plasmid B to generate a plasmid D that contained a 7.6kb fragment encompassing the 5'LTR-env region of the 1579A HIV-1 genome. The plasmid C carrying the 2.7kb HIV-1 DNA fragment and plasmid D were then digested with unique restriction enzymes. The 9.6kb fragment from the plasmid D was ligated to the 3.9kb fragment

from the plasmid C to generate a plasmid that contained the complete full length HIV-1 proviral genome of 1579A and was designated p1579A.

Table 1. Primers Used for Amplification of HIV-1 Subtype A Genome

Fragment Length	Position	Primer	Sequence (5'→ 3')
2.5kb	**5'LTR-pol**	**U3XhoI**	**ATT ACT CGA GTG GAT GGG TTA ATT**
		HIV-1Pol1	**ACT GGT ACA GTC TCA ATA GGA CTA ATT G**
6.2kb	**gag-env**	**SK145**	**AGT GGG GGG ACA TCA AGC AGC CAT GCA AAT**
		DR8	**GGG ACA ATT GGA GAA GTG**
2.7kb	**env-3'LTR**	**DR7**	**CAA CTG CTG TTA AAT GGC AGT CTA GC**
		New LTR 3' NotI	**ATA AGC GGC CGC CCA CTG CTA GAG ATT T**

3.4.4 Transfection, replication kinetics, co-receptor usage and MT2 Assay

HEK293T cells (ATCC) were maintained in DMEM supplemented with 15% FBS, 1% L-glutamine, 1% penicillin-streptomycin, 1ug/ml puromycin and 300ug/ml G418. 293T cells that were grown to 90-95% confluency were transfected with 14ug of p1579A-1 DNA or 35ug of control plasmid pIndie C1 with Lipofectamine 2000 reagent (Invitrogen). Transfected HEK293T cells were cultured in 15% D-MEM media for 48hrs. After 48hrs, supernatant from transfection cultures was passed through a 0.45micron filter and added to 5 million PHA stimulated CD8 depleted PBMC treated with 5ug/ml polybrene for 1 hour in RPMI media supplemented with IL-2. Cultures were maintained at 37°C. Cultures were tested for the presence of p24 antigen every 2-3 days and fed with 2.5 million PHA stimulated and polybrene treated CD8 depleted PBMC every 7 days.

To determine the co-receptor usage of p1579A-1 virus U87.CD4 cells were infected as described (113). Briefly, U87.CD4 cells expressing either the CXCR4 or CCR5 chemokine receptors were seeded in duplicate in a 24-well microtiter plate at a concentration of 250,000 cells/well in a final volume of 500ul 15%D-MEM. The cells were incubated at 37°C until they became 90% confluent. The media was removed from confluent cells and the cell monolayer was washed with 500ul PBS. Approximately 0.5 ml (30 pg p24 equivalent) of p1579A-1 or the control HIV-1 BAL or IIIB was added to duplicate wells of U87.CD4.CCR5 and CXCR4 cells. One set of wells for each co-receptor was left uninfected as a negative control. An additional 0.5 ml of DMEM containing 15% FCS was added to each plate and the plates were incubated at 37°C for 24hrs. After 24hrs, the supernatant was removed. The cells were rinsed with PBS and

incubated in 1 ml fresh DMEM supplemented with 15% FCS. HIV-1 p24 production in the supernatant was monitored every 2 days.

To determine the replication kinetics of p1579A-1 cloned virus 10 million PHA-stimulated CD8 depleted PBMC treated with 5ug/ml polybrene for 1 hour, were infected with approximately 30ng equivalent HIV-1 p24 virus of either cloned virus or the parental isolate. The infections were carried out for 2 hours, with shaking every 30 minutes at 37°C. After the 2 hours, the virus supernatant was removed and the cells were washed and cultured at 37°C in RPMI 1640 medium containing 20% FCS and 5% IL-2. Virus growth was monitored every 2 days by p24 antigen production in the culture supernatant.

MT2 assay to measure syncytia inducing activity of HIV-1 was performed as described previously (114).

3.4.5 Sequencing, phylogenetic analysis and molecular characterization

The entire viral genome was sequenced using an automated sequencer. Sequences were manually joined and contiguous sequence fragments were assembled using the Vector NTI suite software program (Informax, Oxford, UK). DNA and protein alignments were constructed using the CLUSTAL X alignment program of Vector NTI. Consensus sequences and subtype reference sequences used for alignments were obtained from the Los Alamos HIV-1 database.

Phylogenetic analyses of the entire genome sequence was constructed by the neighbor-joining method of Jukes Cantor corrected distances (protein, total mean character corrected distances) with the optimality criterion set to distance as measured in PAUP. Statistical significance of branchings and various clustering were assessed by bootstrap re-sampling of

1000 pseudoreplicates on the complete data set. The trees were edited for publication using Treeview68K version 1.5. Phylogenetic analyses of LTR sequences were conducted through the construction of maximum-likelihood trees as measured in PAUP. Statistical significance of branchings and various clustering were assessed by bootstrap re-sampling of 1000 pseudoreplicates on the complete data set. The trees were edited for publication using Treeview68K version 1.5.

N-glycosylation site analysis was performed at the Los Alamos HIV-1 database. Transcription factor binding site predictions were performed using MatInspector.

GenBank Accession Number: **DQ083238**

3.5 RESULTS

3.5.1 DNA isolation and subtype determination

From the molecular epidemiological data it is expected that very few subjects will be infected with subtype A HIV-1. Therefore, we have performed virus isolation from a number of HIV-1 infected subjects and determined their subtype specificity. The infected subjects for this study were recruited from Calcutta, the major eastern city of India. A summary of patient characteristics is shown in Table 1.

Table 2. Characteristics of HIV-1 infected subjects from India

Patient ID#	Age	Sex	Clinical Symptoms	CD4 Count	Subtype
2145	33	F	NAD	678/ul	C
1577	18	F	NAD	814	C
2161	53	M	asymptomatic	427	C
1590	16	F	weakness, anorexia	736	C
2024	25	F	NAD	350	C
1310	25	F	abdomen pain, diarrhea	343	C
1581	21	F	fever, cough, weakness	377	C
1580	18	F	NAD	810	C
1540	20	F	total body pain, menstrual problems	640	C
1579	16	F	Sore throat, weight loss	477	A
1443	34	F	weakness	702	C
2167	25	F	abdomen pain, leucorrhoea	255	C
1585	21	F	weakness	331	C
2163	25	F	weakness, anorexia	275	C
2177	18	F	cold & cough, weakness	193	C

NAD – nothing abnormal detected

HIV-1 was isolated from these patients by coculturing their PBMC with PHA-stimulated CD8 depleted normal donor PBMC as described previously (111, 112). Production of virus was monitored by HIV-1 p24 production in the culture supernatant. Following thirty to forty days of cultivation, chromosomal DNA was extracted from the infected cell pellets. The DNA was then used to determine the subtypes of HIV-1 isolates using a heteroduplex mobility assay (HMA) kit obtained through the NIH AIDS Repository (31, 93, 115). In this assay, a nested PCR reaction was employed which in the first round amplified a ~1.25kb long amplimer encompassing the V1-V5 region of the HIV-1 envelope gene from proviral DNA. The 1.25kb DNA was then used as a template in the second round of PCR to amplify a ~0.7kb C2-V5 envelope gene fragment. Similar nested PCR primer pairs were also used for amplification of the C2-V5 region from reference plasmids carrying HIV-1 subtype A, B and C env sequences. Nested PCR products from each of the reference plasmids and the unknown were then combined, denatured and reannealed. Subtypes were assigned based on the mobility of heteroduplexes on a nondenaturing polyacrylamide gel. Among the 19 HIV-1 positive Indian isolates tested, one, 1579, was found to be a subtype A (Table 1). Subtypes were further verified by sequencing the 0.7kb PCR products followed by its blast analysis at the Los Alamos HIV-1 Database (data not shown). Proviral DNA from this isolate (1579A) was then used to construct the subtype A infectious molecular clone.

3.5.2 Construction of subtype A infectious molecular clone

To construct the full length molecular clone, long PCR was performed using a high fidelity polymerase and chromosomal DNA from HIV-1 1579A infected cells. Three overlapping sub-genomic fragments of 2.5kb: 5'LTR-pol, 6.2kb: gag-env, and 2.7kb: env-3'LTR were generated.

These three fragments were cloned into pCR-Blunt II-Topo cloning vectors (Invitrogen) resulting in plasmids A, B and C, respectively (Figure 9). Plasmids A and B were digested with unique restriction enzymes AhdI and FseI and the resulting fragments were ligated to generate plasmid D containing a 7.6kb fragment of 1579A genome. Plasmids C and D were then digested with unique restriction enzymes DraIII and SnaBI and the resulting fragments were ligated to generate the full length infectious molecular clone, p1579A. Our cloning strategy exploited the pUC origin of replication and kanamycin resistance gene within the cloning vector such that the fragments being ligated together contained either one or the other of these features. This technique significantly lowered the background of our ligation products and facilitated cloning of these large fragments.

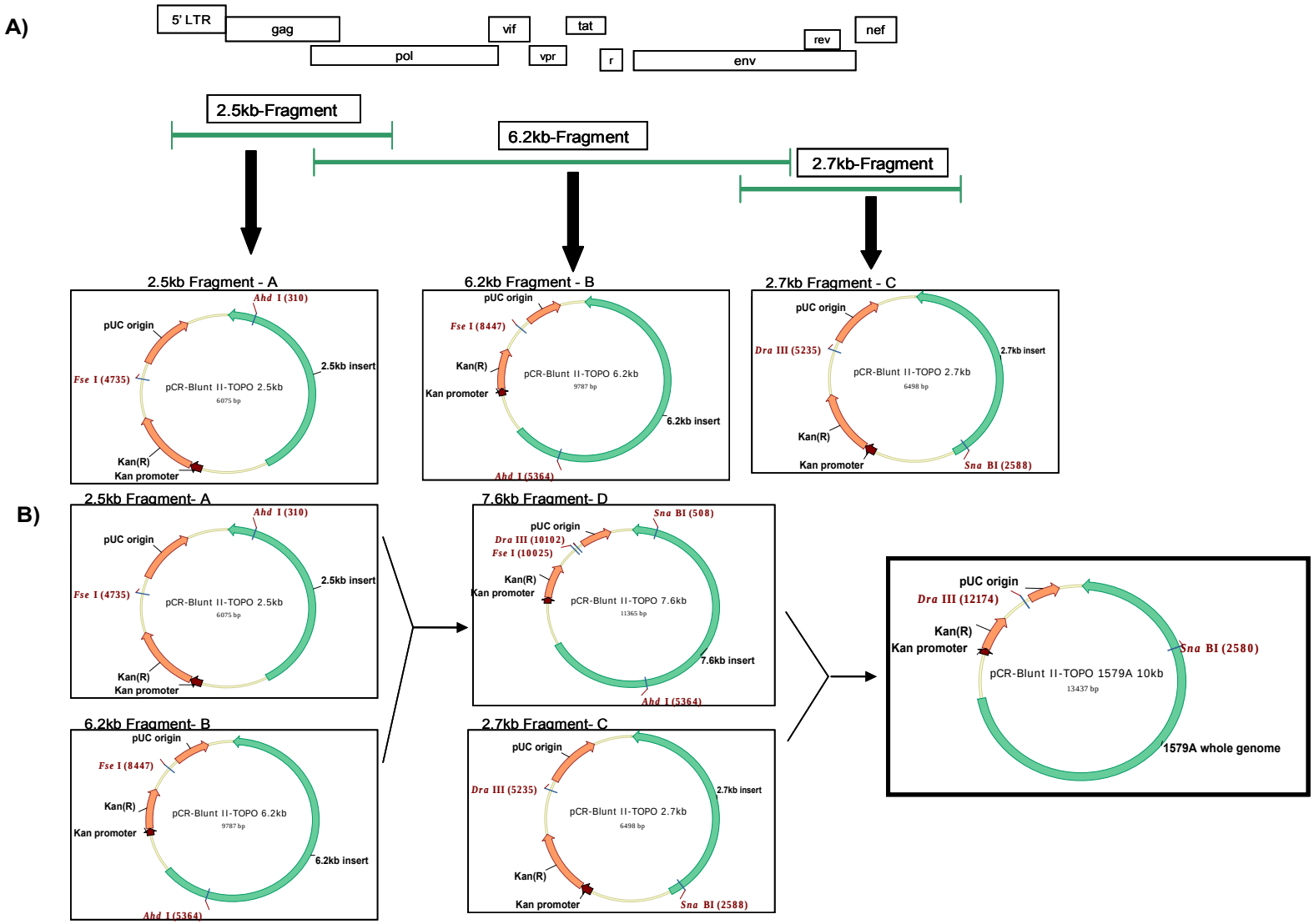

Figure 9. Construction of an HIV-1 subtype A infectious molecular clone.

(A) PCR and cloning of three overlapping subgenomic fragments of 1579A genome. (B) Cloning strategy. Each of the plasmids was cut with unique restriction enzymes and the products ligated together to recover a complete full-length clone of 1579A.

3.5.3 Validation of molecular clone

To validate the presence of the entire HIV-1 genome in p1579A clones, PCR was performed on three separate clones using *gag, pol and env* gene specific primers. Each of the three clones produced amplimers of correct size, 0.9kb, 1kb, and 0.7kb, corresponding to HIV-1 *gag, pol and env* genes, respectively (Figure 10A). The intactness of these three clones gained further support by the analysis of the digestion pattern with NotI restriction enzyme which cleaves outside of the HIV-1 subtype A genome on either side. An approximate 10kb band corresponding to full length HIV-1 DNA and a 3.8kb band corresponding to the plasmid vector was observed for each clone that was digested (Figure 10B). One of these three clones, p1579A-1, was used in subsequent biologic and genetic studies.

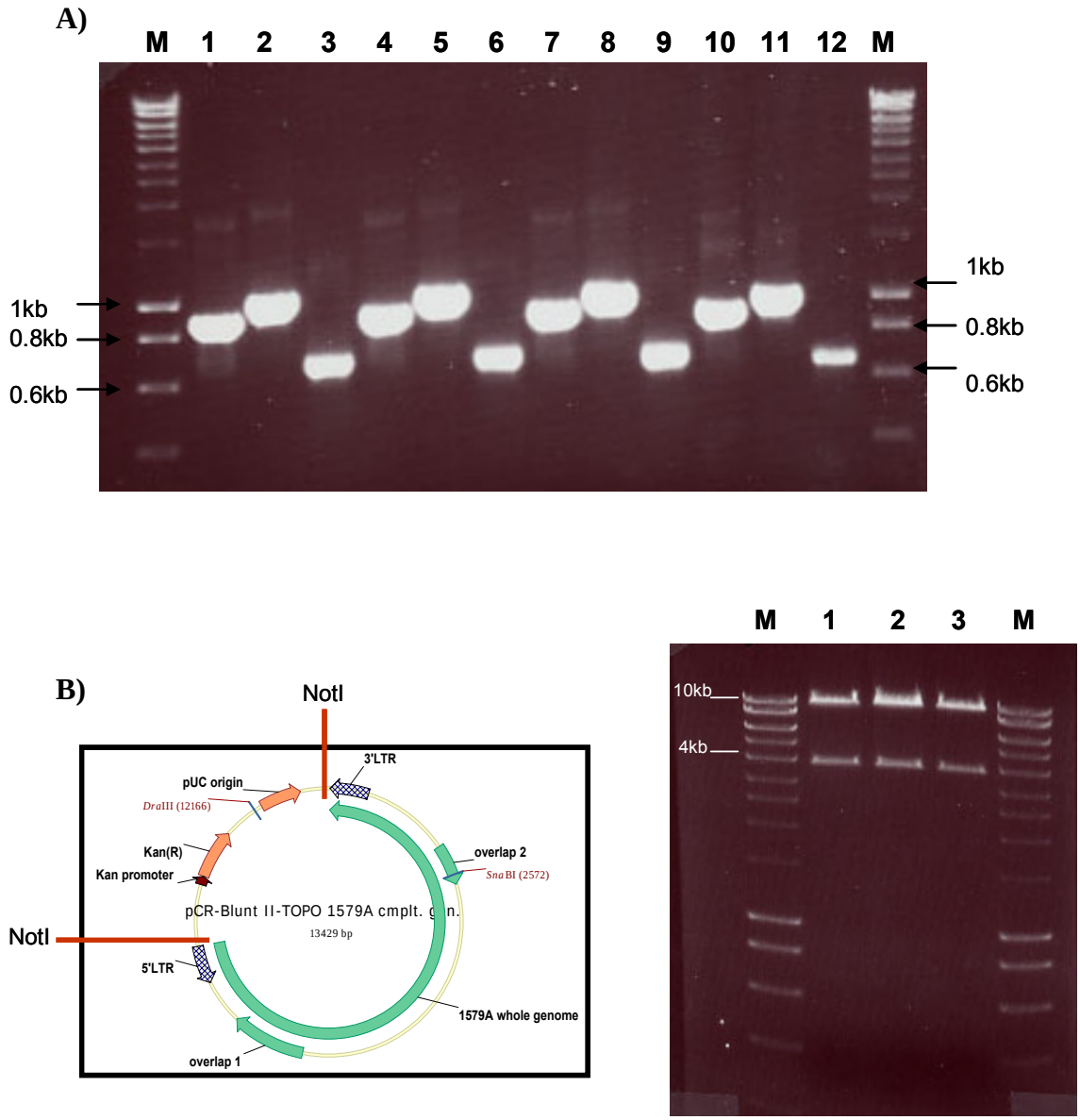

Figure 10. Validation of 1579A whole genome clone

A) Validation of whole genome clone using HIV-1 PCR primers for *gag, pol and env* genes. M=marker. Lanes 1, 2, 3; *gag, pol, env* (respectively) PCR products for clone #1. Lanes 4, 5, 6; *gag, pol, env* (respectively) PCR products for clone #2. Lanes 7, 8, 9; *gag, pol, env* (respectively) PCR products for clone #3. Lanes 10, 11, 12; *gag, pol, env* (respectively) PCR products for positive control of chromosomal DNA from 1579-infected PBMC. B) NotI restriction digest of whole genome clones. Schematic indicates NotI positions in plasmid. Gel image indicates 1579 full length genome (~10kb) and plasmid vector (3.8kb). M=marker, Lanes 1, 2, 3 correspond to clones #1, #2 and #3.

3.5.4 Infectivity of p1579A-1

The infectivity of p1579A-1 was examined by transfection of human endothelial kidney (HEK) 293T cells. Forty eight hours after transfection, filtered culture supernatant was used to infect CD8-depleted PBMC (116). Virus growth was monitored by measuring HIV-1 p24 antigen in culture supernatant. An infectious molecular clone of HIV-1 subtype C of Indian origin, pIndieC1, (117) was used as a positive control for transfection. Results shown in Figure 11A indicate that p1579A-1 is able to produce replication competent HIV-1.

3.5.5 Molecular characterization of p1579A-1

To characterize the growth properties of HIV-1 p1579A-1, CD8 depleted PBMC were infected with p1579A-1 virus (equivalent to 21ng of p24) collected from transfected culture supernatant and an equivalent p24 value of the parental HIV-1 subtype A isolate. Kinetics of viral replication was determined by measuring HIV-1 p24 antigen in culture supernatant. Figure 11B shows that p1579A-1 cloned virus has similar growth properties as the parental isolate. However, the parental isolate produced slightly higher p24 values than p1579A-1.

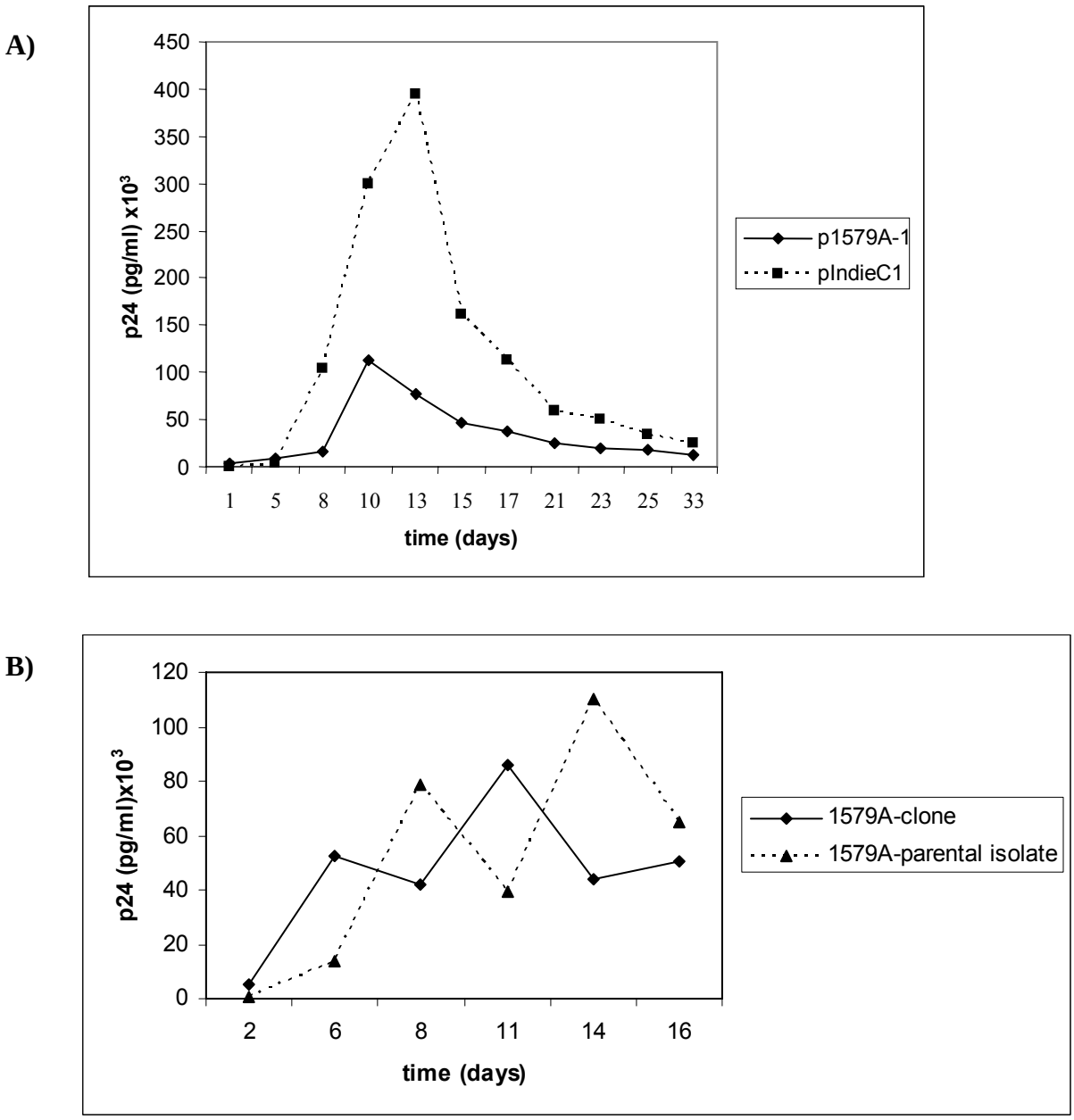

Figure 11. Transfection–infection of p1579A-1 DNA.

(A) HEK293T cells were first transfected with p1579A-1 or the control pIndieC1 DNA. Two days after transfection, culture supernatant was then used to infect CD8 depleted PBMC. Production of HIV-1 p24 antigen on each day sampled is shown for p1579A-1 and the control pIndieC1. (B) Replication kinetics of p1579A-1. CD8 depleted PBMC were infected with 21ng of HIV-1 p24 equivalent virus supernatant of p1579A-1 or the parental isolate, 1579A. Virus growth was monitored by measuring HIV-1 p24 antigen in the culture supernatant every 2 days.

To determine the co-receptor usage of p1579A-1 cloned virus, U87.CD4 cells expressing either the CXCR4 or CCR5 co-receptors were infected with virus containing supernatant equivalent to 30ng of HIV-1 p24 antigen. Virus growth was monitored by measuring p24 antigen in culture supernatant. p1579A-1 HIV-1 replicated more efficiently in U87.CD4 cells expressing CCR5 than CXCR4, indicating that p1579A virus is an R5 tropic virus (Table 3). Consistent with CCR5 co-receptor usage, p1579A-1 HIV-1 was found to be non-syncytia inducing (NSI), due to its inability to form syncytia in MT2 cells (data not shown).

Table 3. Coreceptor usage of p1579A-1 cloned virus

Coreceptor Usage-U87.CD4.CCR5/CXCR4						
	1579A[a]		**Bal[b]**		**IIIB[c]**	
	CCR5	CXCR4[d]	CCR5	CXCR4[e]	CCR5[f]	CXCR4
D2	183	276	2,276	601	338	>200,000
D4	705	210	99,275	650	425	>200,000
D6	3,480	220	ND	ND	ND	ND
D8	12,770	280	ND	ND	ND	ND
D10	33,625	300	ND	ND	ND	ND

[a,b,c]HIV-1 p24 antigen production (pg/ml) in U87.CD4.CCR5 or U87.CD4.CXCR4 expressing cells.

[d,e,f]An increase in p24 value of at least 3-fold over the day 2 value is considered positive HIV replication.

ND—not determined**.**

The complete sequence of the proviral DNA of p1579A-1 was determined by primer walking along the entire genome. The sequences were verified using the HIV-1 Sequence Locator tool at the Los Alamos HIV-1 sequence database and then manually joined. The full length genome of p1579A-1 is 9699 base pairs long. All reading frames of this clone were found to be open. The entire genome was subjected to HIV-1 recombination analysis at the Los Alamos Sequence Database and was not found to be an HIV-1 recombinant. The complete genome of p1579A-1 and the long terminal repeat (LTR) were aligned with HIV-1 group M subtype reference sequences obtained from the Los Alamos HIV-1 database. Phylogenetic analysis of both the complete genome and the LTR showed that 1579A clusters within the HIV-1 subtype A lineage (Figures 12 and 13).

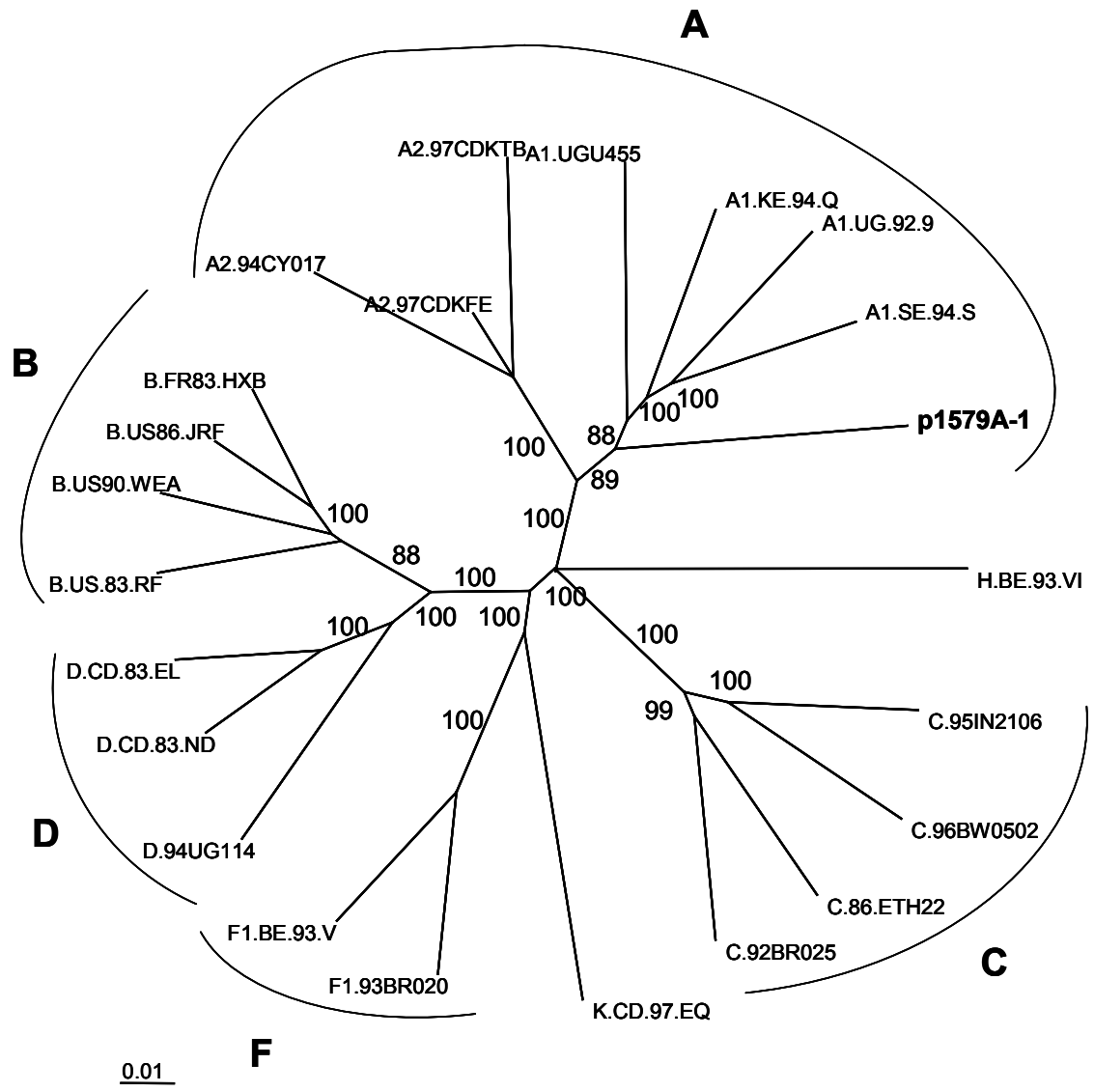

Figure 12. Phylogenetic analysis of p1579A-1.

Neighbor-joining Jukes Cantor unrooted tree showing the phylogenetic relationship of the complete genomic sequence of p1579A-1 to HIV-1 group M reference subtypes. The scale bar at the bottom represents percent genetic divergence.

The envelope gene is important for viral attachment, infectivity and coreceptor usage. We analyzed the envelope gene by aligning the predicted amino acid sequence of p1579A-1 with those of the consensus sequence of reference subtypes A1 and A2 and IndieC1. The characteristic GPGQ motif at the crown of the V3 loop was conserved in all sequences (Figure 14) (43). Consistent with earlier findings of CCR5 coreceptor usage and an NSI phenotype, there was an absence of basic amino acids at positions 11 and 25 in 1579A (Figure 14). This was also true for IndieC1, as reported earlier (117). An analysis of potential N-linked glycosylation sites did not show any major differences between 1579A and reference subtype A1 or A2 obtained from the Los Alamos HIV database or Indie C1. The number of predicted potential sites ranged from 27-31; with IndieC1 having the most, 31 and 1579A having 29. Variations in the locations and number of sites were seen predominantly in the variable regions V1, V2, V4 and V5. Only one potential site was predicted for 1579A between positions 131-150 of the V1 region whereas at least 2 sites were predicted for other subtype A's and 3 were predicted for IndieC1 (Figure 14). Lack of potential N-linked glycosylation sites in this region is probably due to a 14 amino acid deletion in this region. Absence of glycans in this region could have an effect on CD4-gp120 interaction and viral entry (118, 119).

The long terminal repeat (LTR) is known to contain essential elements that control HIV-1 transcription and hence replication. As a first step to explore biological differences between subtypes from India, we compared the LTR of p1579A-1 with the LTR of IndieC1 by performing a transcription factor binding site (TFBS) prediction analysis. Several differences were observed in the number of TFBS between subtypes A and C (Figure 15). Most notably, both p1579A-1 and IndieC1 were found to contain two NF-kappa B binding sites. This is in contrast to HIV-1 subtype C of African origin which has been shown to carry three functional

NF-kappa B enhancers. This is not unusual, since Indian subtype C has been shown to be genetically distinct from other HIV-1 subtype C around the world (13). Furthermore, IndieC1 and not p1579A-1 has the binding sites for GATA-binding factor 3 (GATA) and Octamer-binding factor 1 (Oct1). Finally, p1579A-1 but not IndieC1 has the binding site for GC Box elements which have been reported to be a component of Vpr-mediated LTR activation (120).

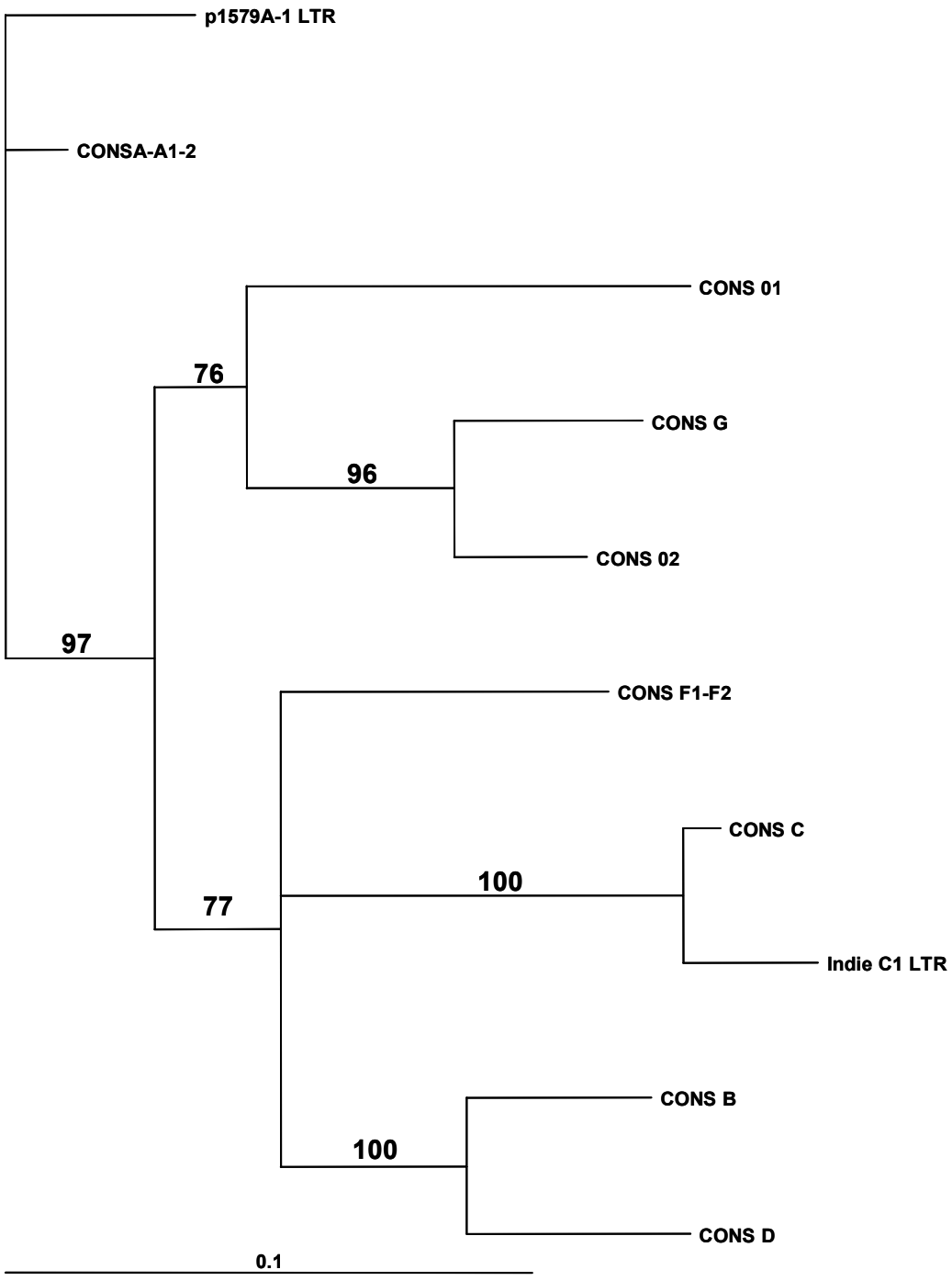

Figure 13. Phylogentic analysis of LTR

Maximum likelihood unrooted tree showing the phylogenetic relationship of p1579A-1 long terminal repeat (LTR) sequence to the consensus (CONS) sequence of HIV-1 group M LTR. All HIV-1 group M reference sequences were obtained from the Los Alamos HIV Database (www.hiv.lanl.gov). The scale bar at the bottom represents percent genetic divergence.

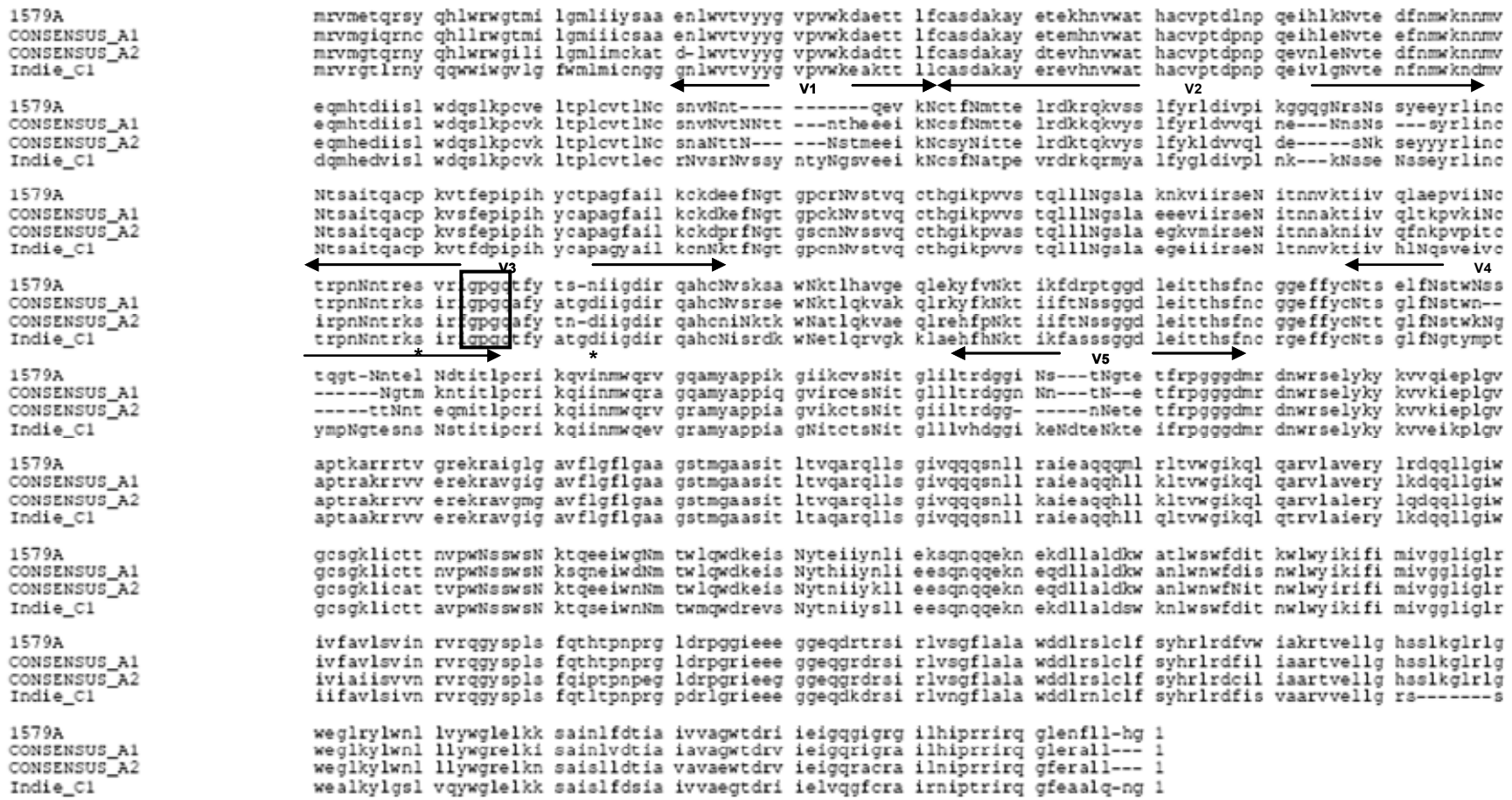

Figure 14. Molecular characterization of HIV-1 p1579A-1 envelope gene.

Amino acid alignment of p1579A-1, reference subtype A1 and A2 (subgroups of HIV-1 group M, subtype A) consensus sequences obtained from Los Alamos HIV-1 database and IndieC1 envelope gene. N=predicted N-linked glycosylation site. The GPGQ motif is surrounded by solid box. Positions 11 and 25 are represented by an asterisk (*).

(A)

Transcription Factor Binding Site Predictions		1579A	Concensus A	Indie C1
AHRARNT	(Aryl hydrocarbon receptor/Arnt heterodimers)	1	0	0
AML3	(Runt-related transcription factor 2/CBFA19core binding factor, runt domain, alpha subunit 1))	1	1	0
AP4	(Activator protein 4)	1	1	0
CMYB	(c-Myb,important in hematopoesis,cellular equivalent to avian myoblastosis virus oncogene v-myb)	1	1	1
CP2	(LBP-1c(leader-binding protein-1c), LSF(late SV40 factor), CP2, SEF(SAA3 enhanced factor)	1	1	1
CREL	(c-Rel)	0	0	1
EVI1	(Ecotropic viral integration site 1 encoded factor)	1	1	1
GATA	(GATA-binding factor 3)	0	0	1
GC	(GC box elements)	1	1	0
MEIS1	(Binding site for monomeric Meis1 homeodomain protein)	2	2	1
NFAT	(Nuclear Factor of activated T-cells)	0	0	1
NFKAPPAB	(NF-kappa B)	2	2	2
OCT1	(Octamer-binding factor 1)	0	1	1
PAX6	(PAX6 paired domain and homeodomain are required for binding to this site)	2	1	0
SMAD3	(Smad3 transcription factor involved in TGF-beta signaling)	0	0	1
SP1	(Stimulating protein 1 SP1, ubiquitous zinc finger transcription factor)	0	0	1
VMAF	(v-Maf)	1	0	1

(B)

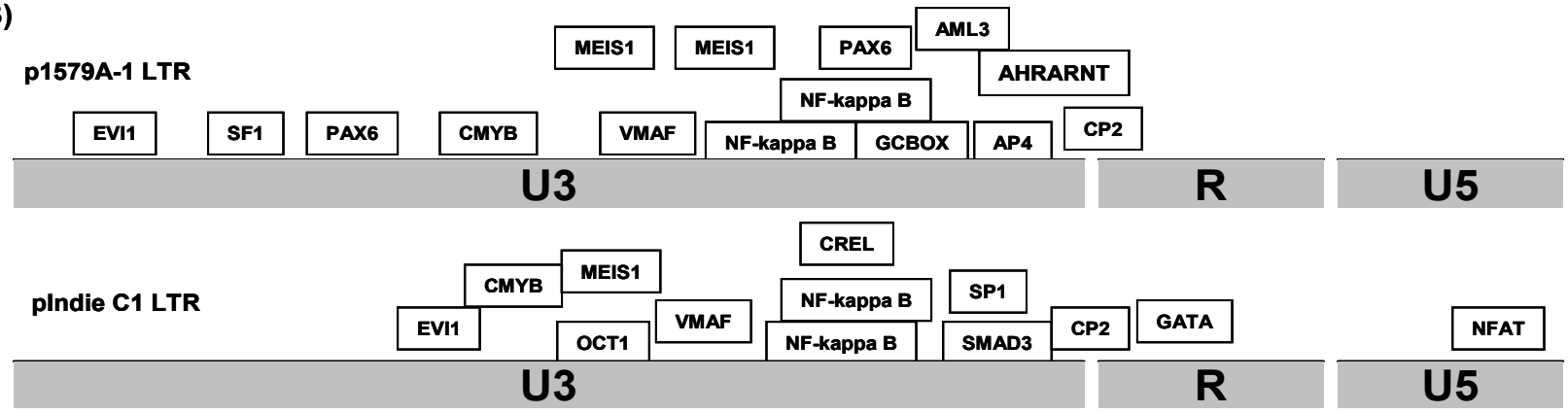

Figure 15. Analysis of HIV-1 long terminal repeat (LTR).

(A) Chart listing predicted transcription factor binding sites. Values in each column represent the number of sites predicted for each transcription factor binding site. (B) Schematic representation of predicted TFBS within the LTRs of 1579A (upper) and IndieC1 (lower).

3.6 DISCUSSION

Out of 19 HIV-1 Indian isolates only one was found to be a subtype A. This finding is consistent with the low frequency (5-10%) of subtype A prevalence in India. From this HIV-1 subtype A isolate the molecular clone p1579A-1 was constructed. This clone was found to be replication competent in CD8-depleted PBMC and CCR5 tropic as determined by genotypic and phenotypic analysis. It is important to note that p1579A-1 was constructed based on unique restriction sites pre-existing within the genome and was not subject to changes in the sequence in order to generate these sites. Furthermore the complete cloned sequence was replication competent. It was unnecessary to provide genes in trans, reconstruct or substitute any parts in order to recover infectious virus which, depending on the location of the substitution, could alter the replication properties of the virus and not be representative of the parental isolate. Therefore p1579A-1 represents biologically relevant virus isolated from an HIV-1 infected individual.

Sequence analysis of p1579A indicated it to be of the HIV-1 subtype A lineage. The branching pattern also indicates that p1579A-1 is more closely related to the sub-subtype A2 than A3. This may indicate an introduction of subtype A into India from the eastern region of Africa however more Indian subtype A analysis would be necessary to confirm this relationship.

GATA 3 is a member of a family of transcriptional activating proteins and is expressed in T lymphocytes. In vitro studies have shown enhanced HIV-1 LTR-directed transcription in the presence of this protein (121). Octamer binding proteins (Oct) have been reported to both positively and negatively regulate the expression of a variety of genes (122, 123). In vitro studies have shown Oct-1 to either repress LTR-directed transcription of HIV-1 in fibroblasts (122) or to

have no effect on HIV-1 transcription in primary CD4 T cells (123). It can be hypothesized that the presence or absence of these binding sites in the LTR of subtypes C and A may be responsible for differences in replication efficiency, resulting in the higher prevalence of subtype C over other subtypes of HIV-1 in India. Further studies are necessary to evaluate this hypothesis.

In conclusion, p1579A-1 represents the first infectious molecular clone of HIV-1 subtype A which replicates in primary CD4 lymphocytes. Availability of a molecular clone of subtype A HIV-1 will be useful for the study of replication dynamics and evolution of subtypes in India. Furthermore, its full genomic sequence will be important for characterizing other HIV-1 subtype A around the world.

Acknowledgements: We thank Ming Ding, Kathy Kulka and Mary White for technical assistance. This work was supported by AIDS-FIRCA grant R03 TH00971 (PG) and AI51661 (SH).

Yue Chen provided technical assistance with construction of the clone.

Jodi Craigo performed phylogenetic analysis and constructed phylogenetic trees.

Ramadas Chatterjee, Pratima Roy, Dhruba Neogi provided PBMC from HIV-1 infected patients from India.

Deena Ratner propagated, expanded and isolated HIV-1 from PBMC obtained from HIV-1 infected patients from India.

4.0 CHAPTER TWO. GENETIC AND FUNCTIONAL CHARACTERIZATION OF THE LTR OF HIV-1 SUBTYPES A AND C CIRCULATING IN INDIA

4.1 PREFACE

This chapter is adapted from a manuscript accepted for publication in AIDS Research and Human Retroviruses. (Milka A. Rodriguez[a] , Chengli Shen[a], Deena Ratner[a], Ramesh S. Paranjape[b], Smita S. Kulkarni[b], Ramdas Chatterjee[c] , and Phalguni Gupta[a]. Genetic and Functional Characterization of the LTR of HIV-1 Subtypes A and C Circulating in India. Reprinted with permission from Mary Ann Leibert. Work described in this chapter is in partial fulfillment of specific aim 4.

[a]Department of Infectious Diseases and Microbiology, Graduate School of Public Health, University of Pittsburgh, PA 15261, USA

[b]National AIDS Research Institute, Pune, India

[c]Department of Virology, Chittaranjan Cancer Research Institute, Calcutta, India

4.2 ABSTRACT

Genetic analysis of HIV-1 sequences circulating in different parts of India have shown that the predominant proportion of HIV-1 subtypes circulating in India is type C and a small fraction are subtypes A, B, E and CRFs. We sequenced the HIV-1 long terminal repeat (LTR) promoter region of seven subtype C and five subtype A isolates obtained from two major cities in India. Sequence analysis of the complete promoter and TAR regions revealed conserved subtype-specific variability in several major binding sites. Three NF-κB sites were present in all subtype C isolates and two isolates contained an insertion in the most frequent naturally occurring length polymorphism (MFNLP). The transcriptional activity of one these isolates may have been hindered due to this insertion. Despite the apparent variability between the LTRs we did not observe any significant difference in the transcriptional activity between subtype C and subtype A. To our knowledge, this is the first study characterizing the genetic structure and functional attributes of subtype A LTRs from India.

4.3 INTRODUCTION

In the current global epidemic, subtype C accounts for more than 50% of all infections and predominates in South and Eastern Africa, India and China (7). With a population of 1.1 billion, India has the second largest number of HIV infected individuals in the world next to South Africa. The estimated HIV prevalence among people 15-49 years old is 0.5%-1.5% (19). Greater than 90% of the infections in India are made up of subtype C (124). The asymmetric distribution of HIV-1 subtypes in India suggests that there may be a biological basis for such dissemination. Genetic diversity of the virus may likely play a key role.

The long terminal repeat (LTR) of HIV regulates the expression of the viral genome through its interaction with both cellular and viral transcription factors (93). The HIV-1 LTR is approximately 640 base pairs long and is divided into 3 structural regions designated, U3, R and U5. The U3 region is further subdivided into the modulatory, enhancer and core/promoter functional regions (93, 94). The U3 region is of particular interest because it contains binding sites for several key transcription factors such as NF-κB, Sp-1, NF-AT, USF and TCF1-α, which regulate HIV transcription (94, 95). There is a growing body of data indicating subtype specific differences of the HIV-1 LTR at both the genetic and functional levels (95-99). The subtype C LTR in particular has been shown to contain three NF-κB binding sites instead of two that are carried by the majority of subtypes (101-103). The additional NF-κB enhancer site in subtype C has been correlated with increased promoter activity and increased viral replication (96, 98, 104, 105). The purpose of this study is to characterize the genetic structure of the LTR of HIV-1 subtype C and subtype A of Indian origin and to compare their functional domains.

4.4 MATERIALS AND METHODS

4.4.1 Viruses, DNA isolation and HIV subtyping

HIV-1 subtype C isolates and the subtype A isolate 1579A were obtained from the peripheral blood mononuclear cells (PBMC) of HIV-1 positive Indian subjects from Calcutta, India by co-culture with PHA-stimulated CD8-depleted normal donor PBMC as described previously (124). DNA isolation and HIV-1 subtyping was performed as described previously (124). HIV-1 subtype A isolates A3, A6, A11 and A81 were obtained from Dr. Paranjape of the National AIDS Research Institute, Pune, India and subtyped by his laboratory. HIV-1 subtype C, IndieC1, was obtained from the molecular clone pIndieC1 (117).

4.4.2 Construction of LTR-luciferase reporter plasmids

The 5' LTR region was amplified from chromosomal DNA of cultured cells using a nested PCR reaction. Outer PCR primers were IN-C LTR XhoI-F or IN-A XhoI-F (5'-TACTCGAGTGGAAGGGTT AATTTACTCT-3' or 5'-ATCTCGAGACTGGATGGG TTAATTTACT-3', respectively) and SK101 (gag) (5'-GCTATGTCAGTTCCCCTTGGTTCTC-3'). Nested PCR primers were IN-C LTR XhoI-F or IN-A XhoI-F (sequence above) and IN-C LTR Hind III-R or IN-A LTR Hind III-R (5'-ATA AGCTTACTGACTAAAAGGGTCTGAG-3' or 5'-ATAAGCTTTGCTAGAGATTTTTACACCAACTAG-3', respectively). Cycling conditions for both outer and nested PCR were the following: 1 cycle at 95°C for 2 min, 35 cycles at 95°C for 15 sec, 55°C for 30 sec, and 68°C for 1 min, with a final extension at 68°C for 3 min. All PCR reactions were performed with high fidelity AccuPrime Pfx DNA polymerase

(Invitrogen). The gel purified nested PCR product was cloned into pCR-Blunt II TOPO vectors (Invitrogen). Forward and reverse sequencing was performed using M13F/R primers. The LTR fragment was then subcloned into pBlue LTR-luciferase plasmid (NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID) using XhoI and HindIII restriction enzymes.

4.4.3 Transcription factor binding site analysis and phylogenetic analysis

Sequences were aligned and manually edited using Vector NTI suite (Informax, Oxford, UK). HIV-1 group M subtype reference sequences from the Los Alamos HIV database were used in the construction of phylogenetic trees. Phlyogenetic analysis was performed using MEGA. The distance matrix was generated by the Kimura two-parameter model whereas the tree topology was determined using the neighbor-joining method. Statistical significance of branchings and various clustering were assessed by bootstrap re-sampling of 1000 pseudoreplicates on the complete data set.

4.4.4 Transfections

Jurkat or Jurkat-tat T cells were transiently co-transfected with LTR luciferase reporter plasmids and the internal control plasmid pGL4.74 which expresses Renilla luciferase (Promega). Transfections were performed using Lipofectamine 2000 reagent (Invitrogen) and 0.5ug of reporter plasmid DNA and 0.05ug of control plasmid. Forty-eight hours post-transfection, the cells were lysed and assayed for luciferase activity using a dual luciferase reporter assay system (Promega).

GenBank Accession Numbers:

EF592596, EF592597, EF592598, EF592599, EF592600, EF592601, EF592602, EF592603, EF592604, EF592605.

4.5 RESULTS

The LTR region from seven HIV-1 subtype C and five HIV-1 subtype A isolates from India were used in this study. A summary of patient characteristics is shown in Table 4. Figure 16 shows an unrooted phylogenetic tree of Indian subtype C and subtype A LTRs and HIV-1 group M LTR sequences. All of the Indian subtype C LTR clustered with subtype C. Five out of seven clustered within Indian subtype C. Two out of seven, C34 and C298, branched outside of the Indian subtype C lineage. These LTRs may be representative of separate introductions of subtype C into India. All of the Indian subtype A LTRs clustered within subtype A and formed their own sub-cluster with a moderate bootstrap value of 70. The 1579A isolate, although obtained from the eastern region of India, appears to be related to the subtype A isolates obtained from the southwestern region of the country.

Table 4. Characteristics of HIV-1 infected subjects from India

Isolate ID#	Age	Sex	City of Isolation	Clinical Symptoms	CD4 Count	Subtype	Coreceptor Usage	Phenotype
pIndieC1	ND	ND	ND	ND	N/A	C	R5	NSI
C31	25	F	Calcutta	abdomen pain, diarrhea	343	C	R5	NSI
C34	N/A	N/A	Calcutta	N/A	N/A	C	R5	NSI
C59	16	F	Calcutta	weakness, anorexia	736	C	R5	NSI
C267	25	F	Calcutta	abdomen pain, leuchorrhoea	255	C	R5	NSI
C293	24	F	Calcutta	enlarged lymph nodes, skin rashes	298	C	R5	NSI
C298	24	F	Calcutta	N/A	233	C	R5	NSI
A3	23	M	Pune	asymptomatic	N/A	A	R5	NSI
A6	30	M	Pune	asymptomatic	N/A	A	R5	NSI
A11	23	F	Pune	asymptomatic	905	A	R5	NSI
A15	16	F	Calcutta	soar throat, weight loss	477	A	R5	NSI
A81	45	M	Pune	asymptomatic	N/A	A	R5	NSI

ND or N/A = information not available
NAD = nothing abnormal detected

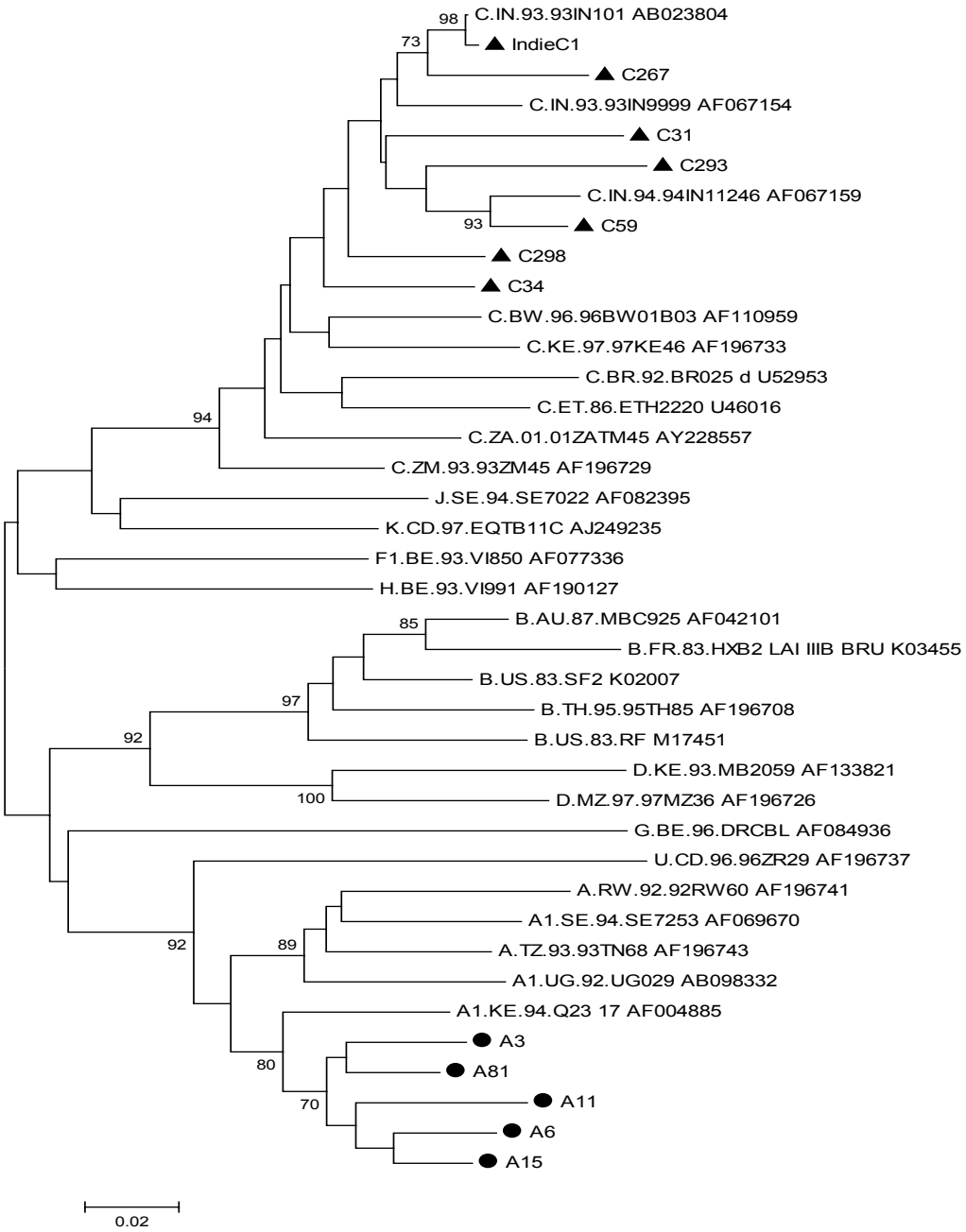

Figure 16. Phlylogenetic analysis of LTR sequences.

Neighbor-joining Kimura two-parameter unrooted tree showing the phylogenetic relationship of HIV-1 subtype A and subtype C LTRs to other HIV-1 group M LTRs. Triangle = subtype C LTR patient isolates. Circle = subtype A LTR patient isolates. The scale bar at the bottom represents percent genetic diversity.

The LTRs were analyzed for potential transcription factor binding sites (TFBS) using TESS. Several subtype specific variations were observed between subtype A and C LTRs (Figure 17). The NF-κB enhancer motifs were highly conserved and were in agreement with previously published reports that subtype C contained an additional NF-κB motif. GATA-3 is a member of a family of transcriptional activating proteins expressed in T lymphocytes. The GATA-3 site was highly conserved among both subtypes. In the subtype A LTR 4 out of 5 contained an A and 1 contained a T at position -449 of the GATA-3 binding site while in the subtype C LTR, 4 out of 7 contained a G, 2 contained a C and 1 contained an A at that same position. The significance of these changes and effects on transcriptional activity are unclear. The NF-AT sites I and II were the most variable among all of the predicted TFBS and contained several subtype-specific sequence variations. The upstream regulatory factor (USF), along with NF-AT, is one of several positive regulatory regions contained within the negative regulatory element (NRE) of the HIV-1 LTR (94, 101). We found several subtype specific changes within the USF and subtype specific signature sequences within the core NRE (positions -174 to -163 of HXB2). The core NRE sequence within the USF for subtype C (CRCAGACACNB) was only moderately conserved. However, the core NRE was highly conserved among subtype A (CTAAAACACAG) - with 2 isolates each having 1 base change compared to the consensus sequence of subtype A isolates analyzed in this study. T cell specific factor (TCF-1 alpha) was highly conserved with subtype specific variations. The most frequent naturally occurring length polymorphism (MFNLP) within the LTR region is located upstream of the NF-κB site at position -120 of HXB2 and can be from 15 to 34 bp long (94, 125). We observed two subtype C LTRs which contained insertions in this region. Isolate C31 contained a 6bp insertion while isolate C59

contained a 20bp insertion. Interestingly, we observed that the LTR from isolate C59 also had the lowest luciferase activity among all of the LTRs (Figure 18). This reduced activity may be due the binding of a nuclear factor which is identical to RBF-2 which has a negative effect on transcriptional activity (125). Sp1 sites, which are important for LTR activation, have been reported to be highly conserved across subtypes (95). Analysis of Indian subtype A and C LTRs showed that the Sp1 sites were also highly conserved with subtype-specific sequences. The TATA box was strictly conserved among all subtypes.

The TAR motif located within the R region of the HIV-1 LTR is an RNA transcriptional control recognition sequence that forms a hairpin stem-loop structure to which the viral Tat protein binds. Binding of Tat to the trinucleotide bulge of the TAR hairpin triggers a series of events that allows increased processivity of transcriptional complexes resulting in the efficient production of full-length transcripts and a marked increase of viral gene expression (88). We found that LTRs from all subtype A and one subtype C (C59) contained the C24T change (94). We also observed that there was a deletion at position 25 within the TAR bulge region of all 5 of the subtype A LTRs.

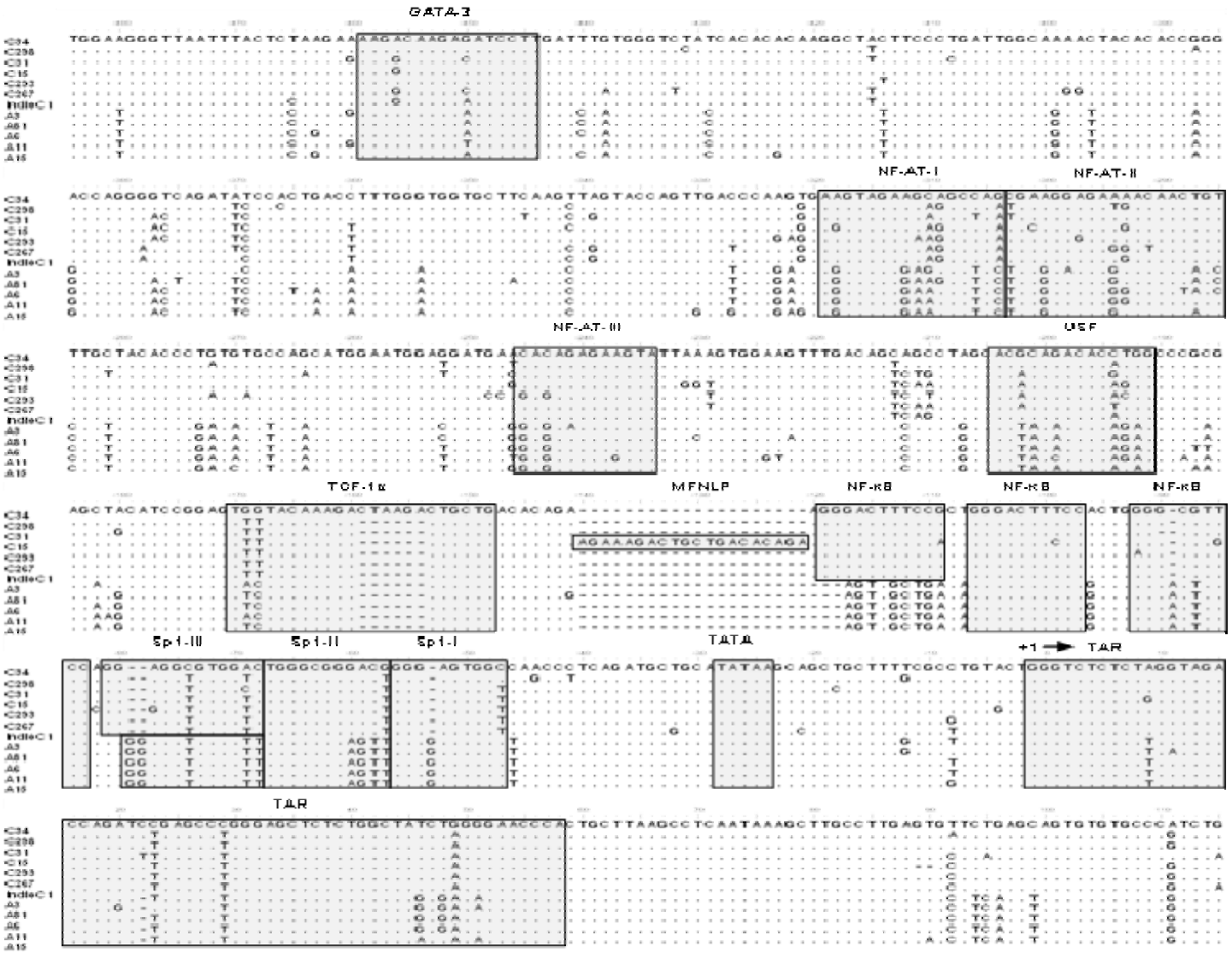

Figure 17. Alignment of subtype C and subtype A LTR's.

Transcription factor binding sites (TFBS) and regulatory regions are shaded. Identical sequence indicated by dots; insertions/deletions indicated by dashes**.**

To evaluate and compare the transcriptional activity of subtypes A and C of Indian origin a luciferase reporter assay was used. We observed no significant difference in the LTR transcriptional activity between subtypes C and A either in the absence ($p=0.10$) or presence ($p=0.59$) of HIV-1 Tat protein (Figure 18). As mentioned earlier, isolate C59 had the lowest transcriptional activity among all of the isolates tested. We speculate that this may be due to the 20bp insertion within the MFNLP. We also compared the transcriptional activity of the LTR of subtype C and subtype A after stimulation with increasing dosages of TNF-alpha (10ng-810ng) or PMA (10ng-30ng) and observed no significant difference ($p=0.10$) between the two groups (Figure 19).

.

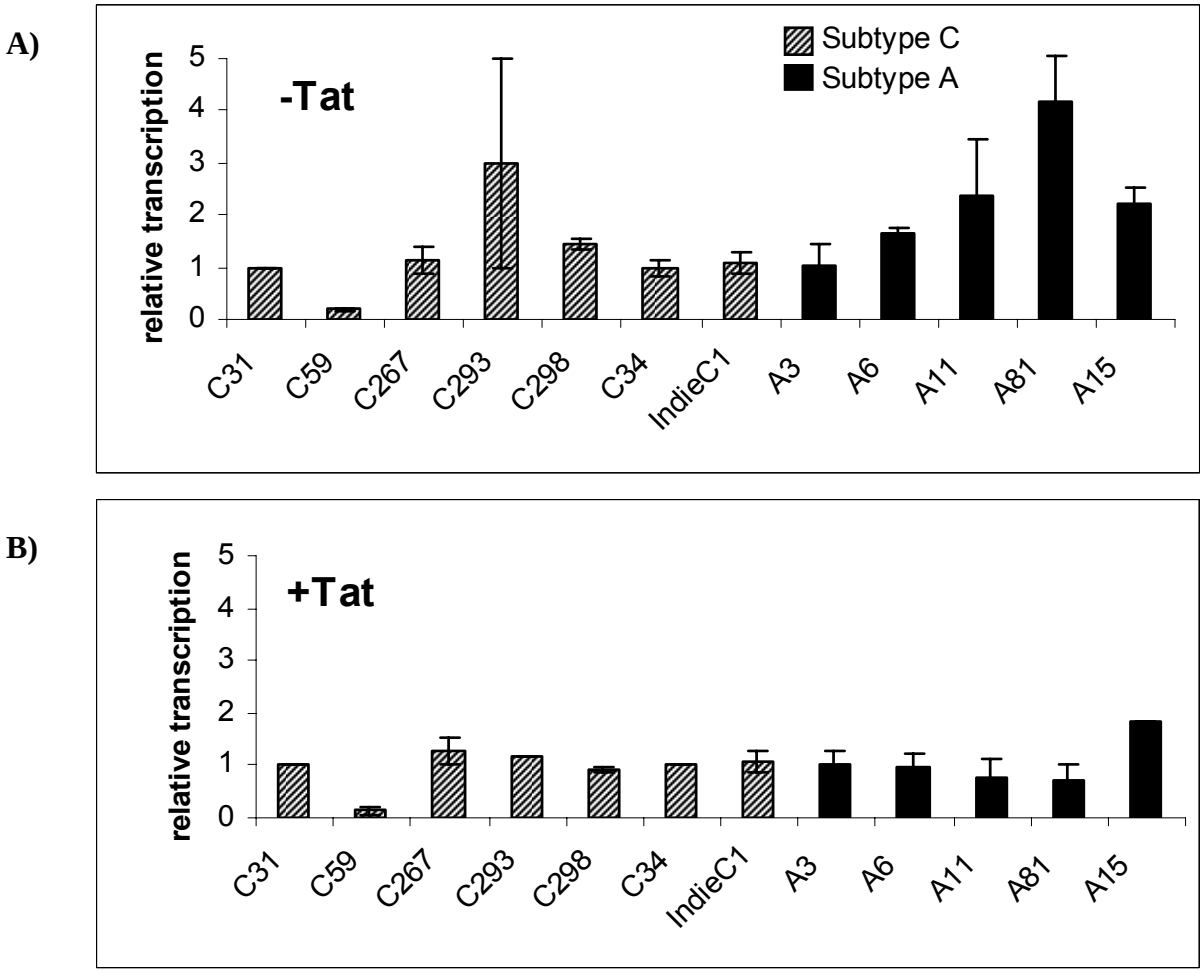

Figure 18. Comparison of transcriptional activity between subtype C and subtype A of Indian origin.

Transcriptional activity was evaluated in either the absence (A) or presence (B) of HIV-1 tat protein. A two sample t-test between the average relative value of light units of HIV-1 subtype A and subtype C LTRs was performed..(A) p=0.10, (B) p=0.59. Relative transcription refers to the average light units of each subtype-specific LTR relative to isolate C31.

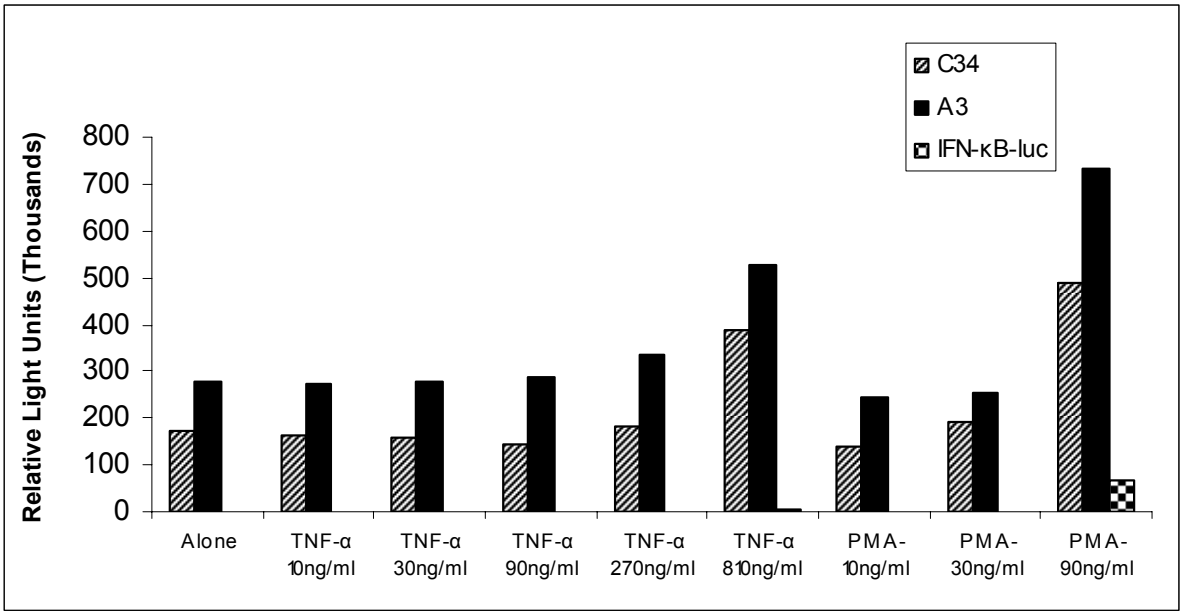

Figure 19. Transcriptional activity of HIV-1 subtype A and subtype C LTR after stimulation.

Transfected cells were cultured with increasing concentrations of TNF-α or PMA for 24 hours. Luciferase activity was measured 48 hours post-transfection. IFN-κB luciferase reporter plasmid was used as a positive control.

4.6 DISCUSSION

This study has described the HIV-1 LTR subtype-specific variability among viruses circulating in India. We did not observe any significant differences in the transcriptional activity between subtypes C and A, however we used a subtype B HIV-1 Tat protein for the transcriptional analysis and therefore speculate that the corresponding subtype specific LTR/Tat combination could potentially yield different results. Further transcriptional analysis should be performed with Tat protein from subtype C and subtype A from India. We have observed a significant difference in the *in vitro* replication fitness of subtypes C and A from India (unpublished data). In these studies subtype C primary isolates almost always outcompeted subtype A primary isolates and overall, subtype C had higher *in vitro* replication fitness in dual infection growth competition assays than subtype A. The study described here suggests that such replication fitness differences between subtype C and subtype A may not be due to differences in LTR function among these two subtypes. Furthermore, isolate C59 which displayed the lowest LTR transcriptional activity, had among the highest replication fitness values in dual infection growth competition assays. In fact a recent study which investigated HIV-1 subtype C LTRs from 4 different regions of India observed similar insertions and identified them as de facto AP-1 binding sites (102). They showed that oligonucleotides spanning the putative AP-1 site efficiently bound recombinant c-Jun (AP-1) protein although with lower affinity than wild type AP-1. AP-1 has been reported to have a positive effect on transcriptional activity (95, 126). Therefore functional studies of these subtype C LTRs are necessary to determine the effect of this insertion on transcription. These results suggest that there is a more complex interplay of virus and host cell factors which determines replication capacity and possibly the eventual spread

of a particular variant within a population. More extensive studies are necessary to determine the nature of these factors and how best to manipulate them to control the current HIV-1 epidemic.

Acknowledgements: This work was supported by the National Institutes of Health through grant U19AI51661. We thank Tianyi Wang for kindly providing the Renilla control plasmid and IFN-κB plasmid used in the reporter assays.

Chengli Shen assisted with phlylogenetic analysis and constructed phylogenetic tree.

Deena Ratner propagated, expanded and isolated HIV-1 from PBMC obtained from HIV-1 infected patients from India.

Ramadas Chatterjee provided PBMC from HIV-1 subtype C infected patients from India

Ramesh S. Paranjape and Smita S. Kulkarni provided primary isolates from HIV-1 subtype A infected patients from India.

5.0 CHAPTER THREE. HIGHER REPLICATION FITNESS AND TRANSMISSION EFFICIENCY OF HIV-1 SUBTYPE C THAN HIV-1 SUBTYPE A FROM INDIA: IMPLICATIONS FOR SUBTYPE C PREDOMINANCE

5.1 PREFACE

This chapter is adapted from a manuscript in preparation (Milka A. Rodriguez[a], Ming Ding[a], Deena Ratner[a], Yue Chen[a], Ramesh S. Paranjape[b], Smita S. Kulkarni[b], Ramdas Chatterjee[c], Patrick M. Tarwater[d] and Phalguni Gupta[a]). Work described in this chapter is in fulfillment of specific aims 2, 3 and 4.

[a]Department of Infectious Diseases and Microbiology, Graduate School of Public Health, University of Pittsburgh, PA 15261, USA

[b]National AIDS Research Institute, Pune, India

[c]Department of Virology, Chittaranjan Cancer Research Institute, Calcutta, India

[d]University of Texas, School of Public Health, El Paso, Texas

5.2 ABSTRACT

HIV-1 subtype C has been the predominant subtype throughout the course of the HIV-1 epidemic in India regardless of the geographic region of the country. In this study we have investigated the in vitro replication fitness of HIV-1 subtypes A and C from India in an effort to understand the mechanism of subtype C predominance in this country. Using a dual infection growth competition assay, we found that primary HIV-1 subtype C isolates had higher overall relative fitness in PBMC than subtype A primary isolates. Moreover using an ex vivo cervical tissue derived organ culture, subtype C isolates displayed higher transmission efficiency across cervical mucosa than subtype A isolates. To determine whether the envelope gene is responsible for increased replication fitness in PBMC and higher transmission efficiency of subtype C, we constructed a half-genome recombinant clone of subtype A in which the 3' half of the viral genome was replaced with the corresponding subtype C 3' half. This A/C recombinant virus had lower replicative fitness and transmission efficiency than the parental subtype A, but exhibited higher replicative fitness than the parental subtype C. These results suggest that the higher replication fitness and transmission efficiency of subtype C virus compared to subtype A virus from India is not due to envelope gene alone and may be due to a complex interaction between the genes located within the two halves of the viral genome. These data provide a model to explain the asymmetric distribution of subtype C over other subtypes in India.

5.3 INTRODUCTION

Human immunodeficiency virus (HIV) – 1 was introduced into the human population through 3 separate cross-species transmissions from non-human primates in Africa (1, 4). Each of these transmission events is represented by phylogenetic groups termed Group M, Group O and Group N. HIV-1 Group M, which is responsible for the global pandemic, is further subdivided into 9 genetic groups (A – D, F – H, J and K) called subtypes or clades, based on sequence diversity in the *env* gene. Subtype C has now become the most predominant HIV-1 variant and accounts for more than 50% of infections worldwide. It is mainly concentrated in South and Eastern Africa, India and China.

In India, HIV-1 has spread rapidly throughout the country where it has been reported in nearly every major city. Heterosexual transmission is the major route of transmission. There are an estimated 5 million people living with HIV-1 in India, 90-95% of which are subtype C infections. Genetic analysis of HIV-1 circulating in India indicates that the current epidemic is the result of one or very few introductions of HIV-1 into the country. Therefore the preponderance of subtype C viruses cannot be explained by new and continual introductions of HIV-1 subtype C into the country. While there may be several factors contributing to subtype C predominance, this asymmetric distribution of HIV-1 subtype C over other subtypes in India suggests that there may be a biological basis for such dissemination. Some explanations for the asymmetric distribution of HIV-1 subtypes in India include: founder effects, human genetic susceptibility and viral attributes such as transmission efficiency and replication fitness of different subtypes of HIV-1.

Fitness is an evolutionary term used to describe the ability of an organism to reproduce and adapt to its particular environment (39). For RNA viruses, fitness can be estimated by the

relative ability to produce stable infectious progeny in a given environment (39). HIV fitness can be affected by any combination of factors encountered during viral replication or host factors such as immune or drug pressure (106, 127). Several studies have investigated HIV fitness as it relates to drug resistance, disease progression and pathogenesis and also the fitness of HIV genetic variants as they relate to transmission efficiency and the global distribution of subtypes.

The seminal study examining this issue comes from the work of Quinones-Mateu *et al*. (51). Using a dual infection growth competition assay (GCA), they examined relative fitness of virus isolates obtained from slow and rapid progressors as well as those representing R5 and X4 phenotypes. They found that virus isolates from long term nonprogressors were out-competed by viruses from rapid progressors. Similarly, X4/syncytium inducing (SI) viruses generally out-competed R5/non-syncytium inducing (NSI) viruses, although some exceptions were noted. This study also showed that there is a correlation between *ex vivo* fitness and disease progression. HIV-1 isolates from progressive patients were shown to be significantly more fit than HIV-1 isolates from long-term survivors – in several cases independent of viral phenotype. This study was useful in demonstrating that fitness values may serve as a predictor of progression to AIDS. With a few HIV-1 isolates they have shown a difference in replication fitness among several non-C subtypes of HIV-1. However, given the limited number of experiments and strains used in competitive studies between different subtypes, it is not clear if a particular subtype of viruses has a clear advantage with respect to replication fitness.

In a later study, Ball *et al*. (52), compared the *ex vivo* fitness of CCR5-tropic HIV-1 isolates of subtypes B from Brazil and US and subtype C mostly from Africa in PBMC, CD4+ T cells, macrophages and Langerhans cells using a dual infection growth competition assay. All subtype C isolates were outcompeted by subtype B isolates in all cell types, except in

Langerhans cells. In the Langerhans cells, subtype C demonstrated competitive replication efficiency against subtype B. This observation is important when one considers the fact that Langerhans cells in the cervix have been implicated for HIV-1 transmission across the cervical mucosa. This study was useful in generalizing the relative fitness of HIV-1 subtype C virus, but is not sufficient to characterize the fitness of Indian type C viruses, due to the recent findings that they are different than most of the other subtype Cs in the world. We and others have shown that Indian subtype C sequences ($C_{IN,}$ C3) are distinct from subtype C sequences from 23 other countries (13, 35, 36). Specific amino acid substitutions within *env* were found to be distinguishing characteristics within this Indian type C lineage (13, 36).

In this study we have therefore compared between HIV-1 subtypes A and C from India the in vitro replication fitness in PBMC using a GCA and transmission efficiency across cervical mucosa using a cervical tissue derived organ culture. HIV-1 subtype A primary isolates were outcompeted by HIV-1 subtype C primary isolates in PBMC obtained from six US and two Indian donors. Furthermore, most of the subtype C had higher transmission efficiencies than subtype A when cervical tissues were challenged with a mixture of subtype A and subtype C viruses.

5.4 MATERIALS AND METHODS

5.4.1 Viruses.

Seven primary isolates (C31, C59, C267, C293, C298 and 1579A) were isolated from PBMC of HIV-1-infected subjects at Chittaranjan Cancer Research Institute in Calcutta, India. Virus isolation was performed as described previously (111, 112). Briefly, PBMC from HIV-infected Indian patients were cultured in the presence of phytohemagglutin (PHA)-stimulated, CD8-depleted normal donor PBMC in RPMI 1640 media supplemented with IL-2. Virus growth/expansion was monitored by HIV-1 p24 production in the culture supernatant and harvested every 5 days for 40 days. Harvested virus was expanded in phytohemagglutin (PHA)-stimulated, CD8-depleted normal donor PBMC in RPMI 1640 media supplemented with IL-2 as described previously (124). Four other primary subtype A isolates (A3, A6, A11 and A81) were obtained from the National AIDS Research Institute in Pune, India. In addition, the HIV-1 infectious molecular clone IndieC1 was obtained from the molecular clone pIndieC1 kindly provided by Dr. M. Tatsumi (117). Coreceptor usage and MT2 assay for syncytia formation were conducted as described previously (124). The infectivity titer for each isolate was determined by measuring HIV p24 production after 7 days using an endpoint dilution assay. The tissue culture dose for 50% infectivity ($TCID_{50}$) was calculated using the Spearman-Karber formula.

5.4.2 Growth Competition Assays.

PBMC were purified from 5 US blood bank donors (racial and ethnic backgrounds unknown) and 2 Indian donors (IRB approved) by Ficoll-Hypaque density centrifugation. CD8-depleted PBMC isolated from PBMC using immunomagnetic beads, were stimulated overnight with PHA and treated with 5ug/ml polybrene for 1 hour prior to infection. For each GCA, 3 x 10^5 PHA-stimulated cells were infected either singly or dually with subtype A and C viruses at $TCID_{50}$ ratios of 1:1 or 1:0.1 at 37°C overnight in a 48-well plate. The next day the cells were pelleted and the virus supernatant was removed. The cell pellet was then resuspended in 750ul IL-2 media and divided equally to three wells (100,000 cells/well) in a 96-well plate (Figure 20). All competition experiments were performed in duplicate. One hundred microliters of cell free supernatant was collected at 7, 14 and 21 days and analyzed for the level of subtype specific RNA by real-time RTPCR. Virus growth was monitored every 7 days by measuring HIV-1 p24 antigen in the culture supernatant.

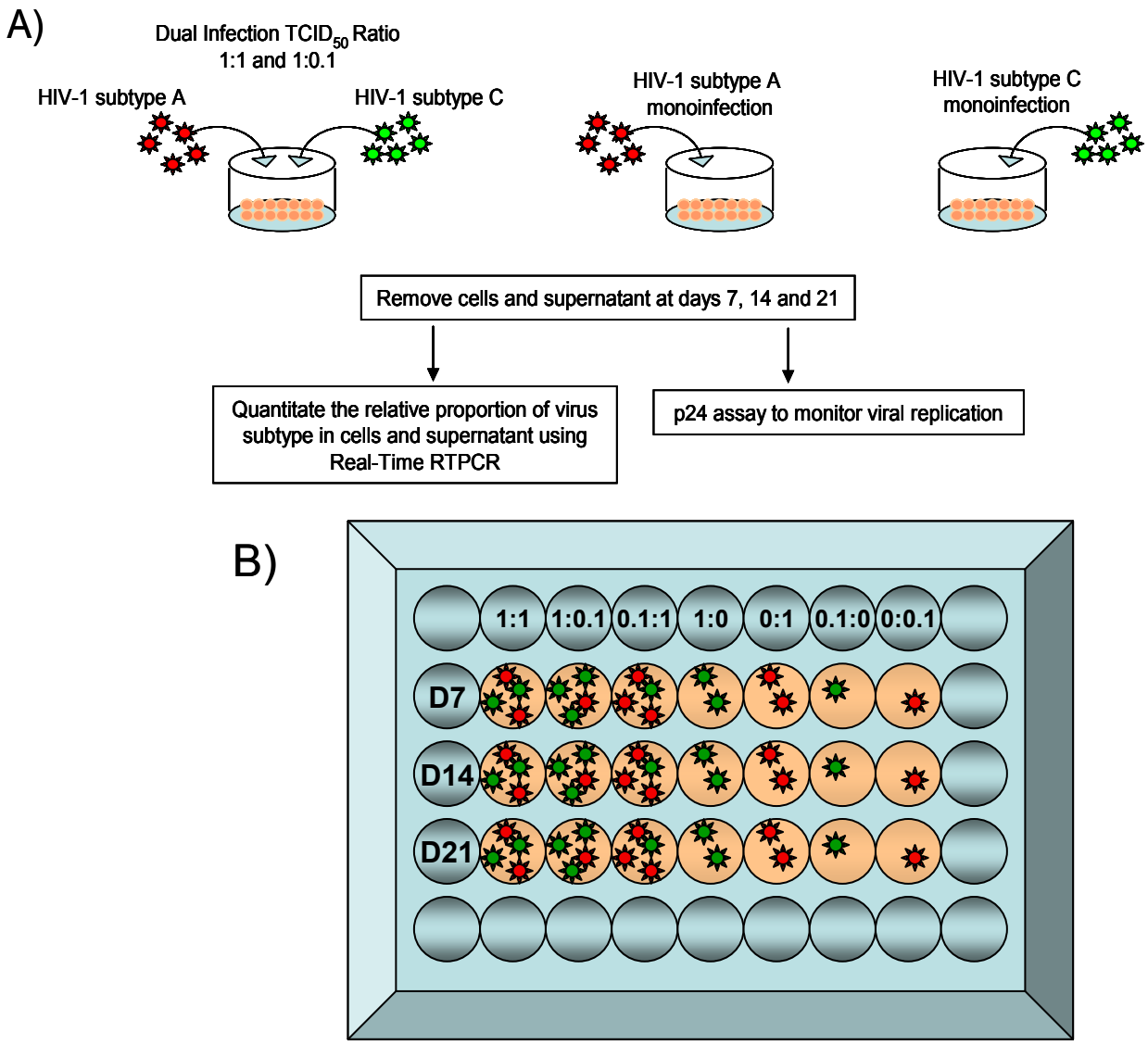

Figure 20. Schematic representation of dual infection growth competition assay procedure (A) and plate set-up (B).

5.4.3 Organ Culture.

Ectocervical tissues were obtained from HIV seronegative premenopausal women with no history of STDs or cancer, undergoing hysterectomy or anterior/posterior repair procedures. The organ culture was set up as previously described (88). Briefly, a 6mm biopsied cervical tissue was placed into the Transwell™ with the epithelial layer oriented upwards and sealed around its perimeter with 3% agarose. The tissue containing Transwell™ was then placed into a 12-well plate containing 1ml of RPMI-1640 media supplemented with 10% FBS. A Transwell with the membrane only served as a positive control while a Transwell with agarose only served as a negative control. To study the transmission of virus, 150ul of cell-free HIV-1 subtype A and C virus either alone or mixed at a ratio of 1:1, was added to the tissue containing well and the positive and negative control wells, and incubated at 37°C for 3 days. Each transmission pair was tested in at least triplicate. On the third day, the tissue was removed and the culture supernatant in the bottom chamber was centrifuged at high speed (22,00rpm) to pellet the transmitted virus. The level of subtype specific HIV-1 RNA in the pellet was quantitated using TaqMan® real-time RTPCR. To ensure that transmitted virus was not due to leaks in the system, at the end of each experiment the intactness of the tissue was evaluated by examining transmission of blue dextran through the tissue and agarose control wells.

5.4.4 RT-PCR and TaqMan real-time PCR.

To quantify the relative proportion of HIV-1 subtype A and C in culture fluid of GCA or transmission studies, subtype specific primers for the *gag*-p17 region of the genome were used in a TaqMan® real-time PCR assay. Standard curves from 10^6 to 1 copy were generated using

p1579A-1 or pIndieC1 molecular clones (117). The concentrations of the clones were determined by spectrometry at 260nm. Primers and probes for real-time PCR analysis were designed using Primer Express® software version 2.0. The subtype A specific primers used were gagAF153-178 (5'-CTGGTGAGTACGCCAATTTTTG-3') and gagAR276-258 (5'-CCCCTGGCCTTAACCGAAT-3'). The subtype C specific primers used were gagCF150-174 (5'-CGACTGGTGAGTACGCCAATTTTA-3') and gagCR270-243 (5'-GCCTTAACCTAATTTTTTCCCATTTATC-3'). The probe used was universal and recognized by both subtype A and C variants; gagP-florochrome 205-187 (5'-ATCTCTCTCCTTCTAGCCT-3'). 100-1000µl culture supernatant equivalent RNA was applied for reverse transcription using TaqMan® reverse transcription reagents (AppliedBiosystems) according to the manufacture's protocol. A 30µl TaqMan® PCR was performed by mixing 5µl cDNA with TaqMan® Universal PCR Master Mix (AppliedBiosystems), 333nM each of forward and reverse primer and 250nM FAM/ MGB labeled probe. ABI Prism 7000 Sequence Detection System was used to carry out Real-Time PCR using the following cycling condition: 50 ºC for 2 min, 95 ºC for 10 min, 45 cycles of 95 ºC for 15 sec and 60 ºC for 1 min. Serial diluted plasmid DNA ranging from 1 to 10^6 were applied to each PCR assay for a standard curve. No Template Control was included in each assay as well to guard against cross contamination. Each sample was run in triplicate. ABS Prism 7000 SDS Software (Applied Biosystems) was used for PCR data analysis and HIV copy number estimation. To quantify the relative proportion of half-genome A/C recombinant virus and parental subtype C virus using TaqMan® real-time PCR, *gag* subtype-specific primers and *gag* universal probe was used for infections with pIndieC1 parental virus and the A/C recombinant half-genome virus. To quantify parental p1579A virus and half-genome A/C recombinant virus, envelope gene subtype specific primers and probes

were used. The *env* subtype A specific primers used were 1579A-F (5'-TCTGTGTCACTTTAAATTGTAGCAATGT-3') and 1579A-R (5'-TCTGTGGTCATATTGAAAGTGCAGTT-3'). The *env* subtype A specific probe used was 1579A-P (5'-TTTACTTCCTGTGTGTTATT-3'). The subtype C specific primers used were IndC1-F (5'-CAAAGCCTAAAGCCATGTGTAAAA-3') and IndC1-R (5'-GCTCCCATTGTAGGTATTATAACTGCTA-3'). The subtype C specific probe used was IndC1-P (5'-TCTGTGTCACTTTAGAATGTAGA-3'). Standard curves, reaction conditions and cycling conditions were as described above for *gag*.

5.4.5 Construction of A/C half-genome recombinant virus.

Construction of a half genome A/C recombinant clone involved use of the available unique restrictions sites in the infectious molecular clone of subtype A HIV-1 1579A. The 3'-half (tat-3'LTR) of pIndieC1 was PCR amplified with a forward primer containing a BamHI site and a reverse primer containing a BstEII site. The PCR product was subsequently subcloned into the pCR Blunt II TOPO vector (Invitrogen). All PCR reactions were carried out using the high fidelity AccuPrime Pfx DNA polymerase (Invitrogen). p1579A and the 3'-half of pIndieC1 were digested with BamHI and BstEII and the corresponding genomic halves were ligated together. The resulting clones obtained from the ligated mixture were used in the transfection of HEK293 T cells followed by expansion of the virus containing supernatant in CD8-depleted PBMC.

5.4.6 Estimation of viral fitness.

Viral fitness was calculated by measuring relative viral RNA concentration of subtype A and C in culture supernatant as described by Quinones-Mateu *et al*. (51). Briefly, the final ratio of virus produced in dual infections (f_0) was divided by the initial ratio of virus in the inoculum (i_0) to derive a single relative fitness (w) value ($w=f_0/i_0$). An average relative fitness value (w^n) was estimated for each variant in dual competitions, where w^n = the average fitness value for each timepoint based on all TCID 50 ratios. The ratio of relative fitness values of each HIV-1 variant in the competition is a measure of the fitness difference (W_D) between both HIV-1 strains ($W_D=W_M/W_L$), where W_M and W_L correspond to the relative fitness of the more and less fit viruses, respectively. A $W_D > 1$ corresponds to the more fit variant. A $W_D <$ corresponds to the less fit variant. A $W_D = 1$ corresponds to equal fitness of the two variants (41, 51)

5.4.7 Trans suppression assay.

Culture supernatant from subtype C or A infected CD8-depleted PBMC at day 15 was filtered followed by depletion of virus by high speed centrifugation at 22,000rpm for 1 hour at 4°C. One-hundred thousand CD8-depleted PBMC were then pre-treated with the virus-free supernatant or cell culture supernatant from uninfected cells or media for 4 hours at 37°C. After pre-treatment, the cells were spun down and the supernatant removed. Pre-treated cells were infected with 200ul of $3.8x10^6$ TCID 50/ml HIV-1 subtype A (~30ng p24 protein) overnight in a 24-well plate. The next day the virus was removed and the cells were resuspended in a final volume of 10%, 20% or 50% virus-free supernatant from subtype A or C infected cells. The final culture volume

was brought up to 250ul with RPMI-1640 media supplemented with IL-2. Virus growth was estimated by HIV-1 p24 in the culture supernatant at day 7 post infection.

5.4.8 Statistical analysis.

Average RNA copy numbers were calculated from each duplicate virus in the competition. Two sample t-test of the copy number means and Mann-Whitney (nonparametric) were used for comparison of subtype A group and subtype C group at day 7 and day 14, for assessment of statistical significance.

W_D values were calculated as described above. One sample t-test was used to determine whether the log relative W_D was significantly different from $W_D = 0$. For all analyses, the level of significance was set at $p = 0.05$.

5.5 RESULTS

5.5.1 Characterization of HIV-1 isolates and subtype determination.

All of the viruses used for this study were obtained from HIV-1 infected individuals from India. HIV-1 isolates A3 - A81 are primary subtype A isolates obtained from Pune, a city in the southwest region of India. Primary HIV-1 subtype C isolates C31 - C298 and subtype A isolate A15 were obtained from Calcutta, a major city in the eastern region of India. A summary of patient information and viral characteristics is shown in Table 5. HIV-1 subtypes were assigned to each of the isolates based on DNA sequences within the *env* and *gag* genes and the LTR. To determine the coreceptor usage, U87.CD4 cells expressing either the CXCR4 or CCR5 coreceptors were infected with virus supernatant and virus growth was monitored by measuring p24 antigen in the culture supernatant. All of the primary isolates used in this study replicated more efficiently in U87.CD4 cells expressing CCR5 than CXCR4, indicating that they are all R5 tropic viruses. Consistent with CCR5 coreceptor usage, all isolates were found to be non-syncytia inducing (NSI) due to their inability to form syncytia in MT2 cells.

Table 5. Indian patient information and virus characteristics

Isolate ID#	Age	Sex	City of Isolation	Clinical Symptoms	CD4 Count	Subtype	Coreceptor Usage	Phenotype
pIndieC1	ND	ND	ND	ND	N/A	C	R5	NSI
C31	25	F	Calcutta	abdomen pain, diarrhea	343	C	R5	NSI
C34	N/A	N/A	Calcutta	N/A	N/A	C	R5	NSI
C59	16	F	Calcutta	weakness, anorexia	736	C	R5	NSI
C267	25	F	Calcutta	abdomen pain, leucorrhoea	255	C	R5	NSI
C293	24	F	Calcutta	enlarged lymph nodes, skin rashes	298	C	R5	NSI
C298	24	F	Calcutta	N/A	233	C	R5	NSI
A3	23	M	Pune	asymptomatic	N/A	A	R5	NSI
A6	30	M	Pune	asymptomatic	N/A	A	R5	NSI
A11	23	F	Pune	asymptomatic	905	A	R5	NSI
A15	16	F	Calcutta	sore throat, weight loss	477	A	R5	NSI
A81	45	M	Pune	asymptomatic	N/A	A	R5	NSI

ND or N/A – not determined, information unavailable
R5 – CCR5 coreceptor
NSI –Non-syncytia inducing

5.5.2 Standardization of HIV-1 RNA quantitation.

To quantify the viral RNA concentration in dually infected cultures we have standardized a subtype-specific real-time RTPCR that can quantify subtype A and subtype C in a mixture of these two subtypes. Using a number of TaqMan® cycling conditions, proviral DNA from subtype A and C infected PBMC and several primer-probe pairs we were able to obtain *gag* specific primer pairs and cycling conditions which could distinguish between and specifically amplify both subtypes. Figure 21A shows that using TaqMan® PCR cycling conditions, subtype C primers only amplified the HIV-1 subtype C proviral DNA. Likewise, subtype A primers only amplified HIV-1 subtype A proviral DNA. Based on these results we constructed standard curves to quantitate the relative proportion of viral variants in dual infections using full length molecular clones of HIV-1 subtype A and C. Each plasmid clone was serially diluted from 10^6 to 1 copy per reaction and analyzed in triplicate. Figure 21B shows that the assay had a linear range of detection from 10^1 to 10^6 copies/ml of subtype C and subtype A and the amplification efficiency for each subtype specific primer set was nearly identical. Furthermore, we observed that even in the presence of non-specific template, both primer sets were able to maintain subtype specificity and the same amplification efficiencies (Figure 21B).

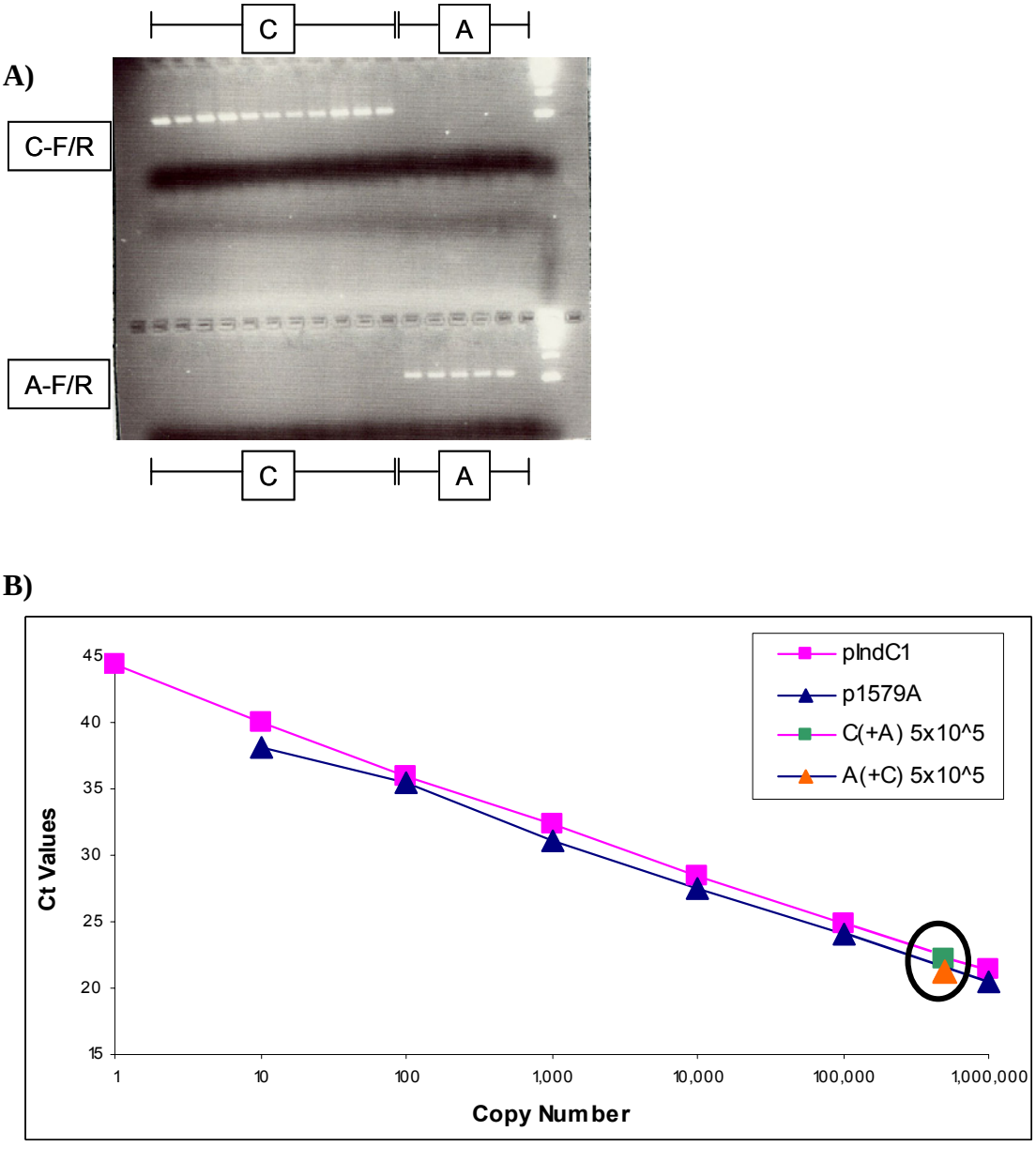

Figure 21. Optimization of primers and probes for TaqMan® real-time PCR for detection of subtypes A and C in a GCA.

A) Gel image of PCR products generated using subtype A or subtype C proviral DNA and subtype specific primers. C-F/R and A-F/R indicate subtype C or subtype A specific primers used in the PCR reaction (respectively). Lanes 1 – 11 each lane represents proviral DNA from PBMC infected with a different HIV-1 subtype C virus. Lanes 12 – 16 each lane represents proviral DNA from PBMC infected with a different HIV-1 subtype A virus. B) Standard curves of subtype A and subtype C. Oval indicates a real-time PCR assay in which primers and probes were tested by combining $5x10^5$ copies of HIV-1 subtype A and C plasmid DNA. Ct value = Cycle threshold value.

5.5.3 Estimation of viral fitness in PBMC by GCA.

It has been well established that the best way of detecting differences in replicative fitness is by dual infection growth competition assays (GCA). This is due to the fact that these assays provide an internal control for the growth environment and is better able to detect smaller differences in fitness than monoinfections (49). Several methods have been employed to estimate the fitness of viral variants in GCA. It has been recently shown that TaqMan real-time PCR is a suitable method of detection and quantification of two competing HIV-1 variants in a GCA and is less labor intensive and more sensitive than previously used techniques such as HTA (128). Therefore, in our study we have applied the TaqMan real-time PCR technique to estimate the viral fitness of subtype A and C from India in a GCA.

For our GCA experiments, we performed nineteen dual infections using seven subtype C (six primary and one molecular clone) and five subtype A isolates (four primary and one molecular clone). These competitions were performed in various combinations in six different blood donors. In these assays CD8-depleted PBMC from US blood bank donors were either dually or singly infected with each virus subtype at varying $TCID_{50}$ ratios. The proportion of each variant in the culture supernatant was measured every 7 days by TaqMan© real-time PCR. Figure 22 shows a representative GCA, showing HIV-1 RNA concentration in monoinfections and in dual infections of subtype A and C at both 1:1 and 1:0.1 TCID 50 ratios, respectively at days 7 and 14 after infection. At both of these two days tested, there was no significant difference in the relative copy numbers in monoinfections, while in dual infections subtype C clearly outcompeted subtype A by day 7 at both 1:1 and 1:0.1 (A:C) TCID 50 ratios. We observed similar results in all other competitions between subtype A and C, however in some cases at TCID 50 ratios of 1:0.1 (A:C) subtype A was not outcompeted by subtype C until day

14. Figure 23 shows a summary of all GCAs performed with subtype A and C at days 7 and 14 at a 1:1 TCID 50 ratio. The data indicate that there was a significant difference in the RNA copy number of subtype C isolates compared to subtype A isolates (at D7 $p < 0.001$, at D14 $p < 0.001$). Table 6 shows the estimated relative fitness values (w) and fitness difference (W_D) for each pair of isolates used in the GCA at day 7. With the exception of pIndieC1 in competition with the primary isolate A3, all subtype C viruses had a fitness difference value that was at least 3-fold higher than subtype A viruses. At day 7 the mean W_D for all subtype C viruses tested was 58.3 with a range of 1.0 – 255.5, while the mean W_D for all subtype A viruses tested was 0.2 with a range of 0.0 – 0.3. Similar results were obtained at day 14. The mean W_D for subtype C viruses at day 14 was 57.8 with a range of 0.4 – 226.4, while the mean W_D for subtype A viruses at day 14 was 0.4 with a range of 0.0-2.7 for day 14. Overall the mean W_D for subtype C viruses at all timepoints for both TCID50 ratios was 58.4 with a range of 0.6 – 261.3, while the overall mean W_D for subtype A viruses was 0.2 with a range of 0.0 – 1.5.

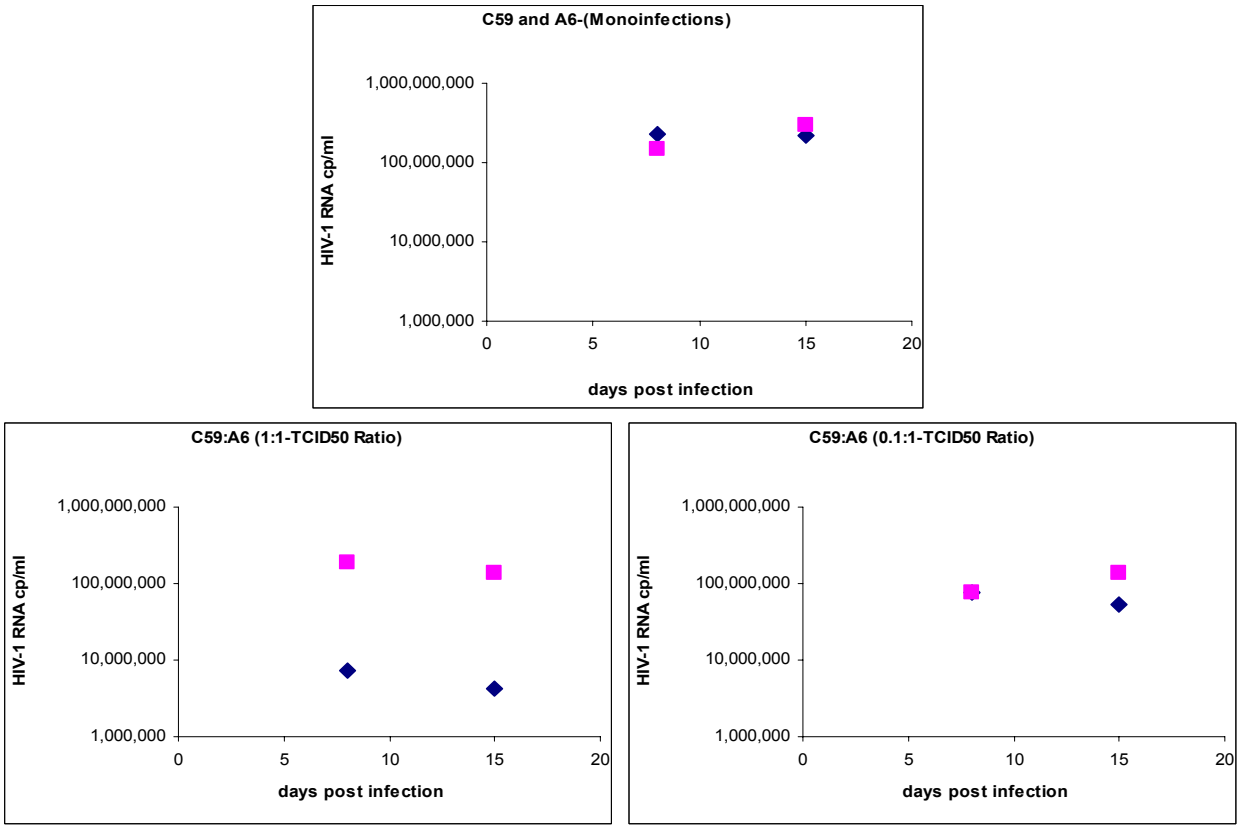

Figure 22. Representative GCA showing relative copy numbers of subtype A and C isolates in monoinfections and dual infections determined by TaqMan real-time PCR.

Pink squares = subtype C, Blue diamond = subtype A. Single pink square at day 7 of 0.1:1TCID50 ratio indicates equal number of copies for subtype A and C.

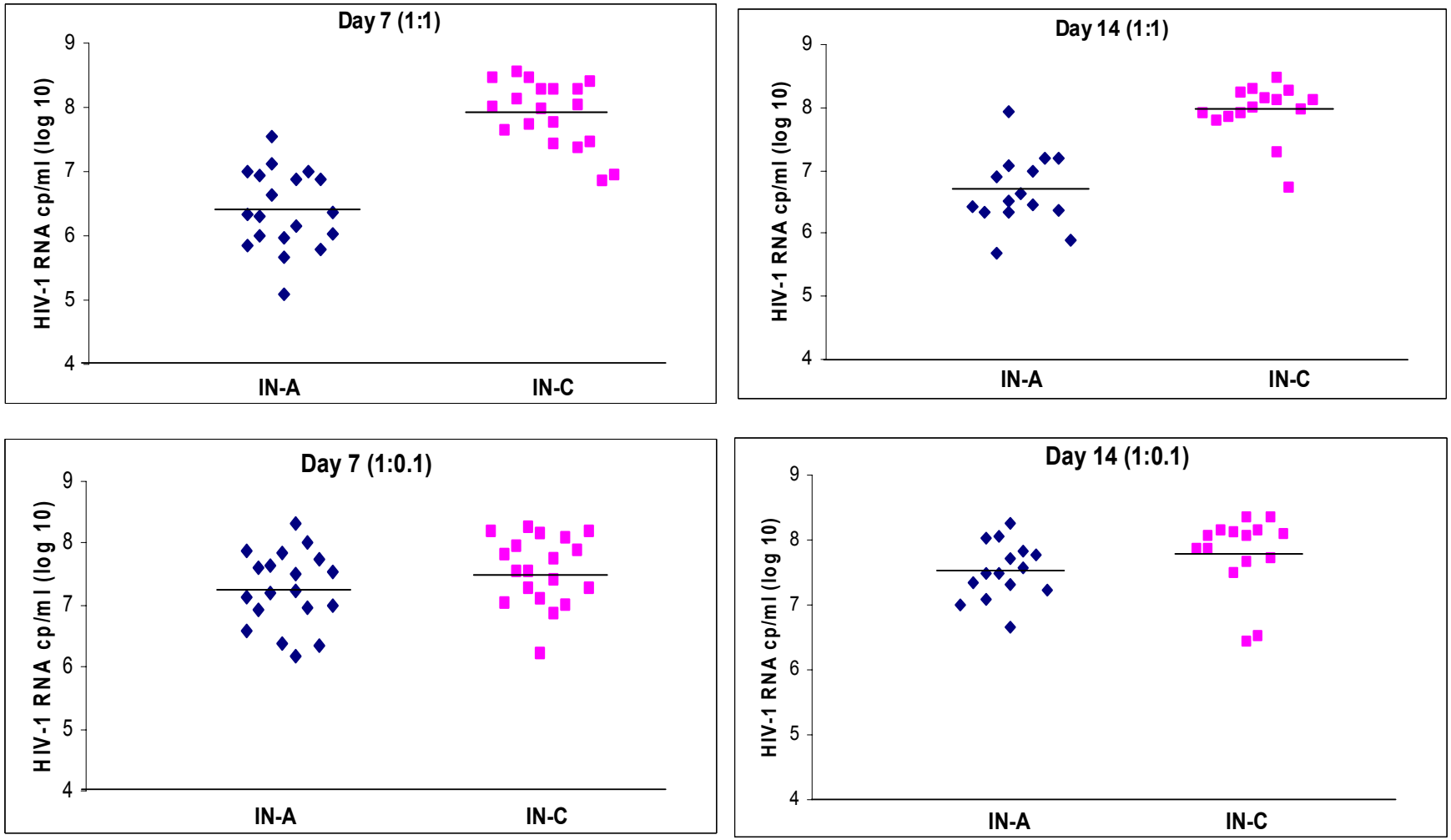

Figure 23. Summary of relative copy numbers of all GCA's performed at days 7 and 14 and 1:1 and 1:0.1 (A:C) TCID50 ratios.

Relative HIV-1 RNA copy numbers are expressed as log base 10 values. Line indicates log base 10 value of the mean of the RNA copy numbers.

Table 6. Relative fitness and relative fitness fold difference values of day 7 GCA using PBMC from US donors

GCA Pairs	Relative W_A †	Relative W_C †	Relative $W_{D(A/C)}$ ‡	Relative $W_{D(C/A)}$ ‡
C31:A6	0.7 ± 0.12	2.2 ± 0.83	0.3	3.3
C31:A11	0.2 ± 0.01	4.6 ± 0.03	0.1	19.9
C31:A15	0.1 ± 0.03	5.6 ± 0.14	0.0	175.5
C59:A6	0.3 ± 0.01	3.5 ± 0.09	0.1	10.9
C59:A11	0.2 ± 0.17	4.9 ± 1.07	0.0	28.5
C59:A15	0.0 ± 0.01	5.7 ± 0.08	0.0	255.5
C267:A3	0.1 ± 0.02	5.6 ± 0.12	0.0	110.4
C267C:A15	0.4	2.9	0.1	7.9
C298:A15	0.2	4.4	0.0	21.8
pInC1:A3	1.0 ± 0.14	1.0 ± 0.19	1.0	1.0
pInC1:A81	0.4 ± 0.07	3.0 ± 0.29	0.1	7.1

† Relative fitness $w = f_0/i_0$, where: f_0 = final ratio of virus produced in dual infections and i_0 = initial ratio of virus isolate in the inoculum. Relative fitness values are based on 1:1 and 1:0.1 TCID 50 ratios, ± standard deviation at 1:1 and 1:0.1 TCID 50 ratios.

‡ Fitness difference (fold difference) $W_D = W_M/W_L$, where: W_M = relative fitness of more fit variant and W_L = relative fitness of the less fit variant

We also compared the replication fitness of subtype A and C viruses between PBMC from US and Indian donors. For this purpose we first examined the kinetics of viral replication in monoinfections of PHA-stimulated PBMC from 13 Indian blood donors and 6 US blood donors. Results showed no significant difference in the replication kinetics of subtype A and C viruses between US and Indian blood donors (data not shown). We then performed GCAs using PBMC from 2 Indian donors and 3 pairs of subtype A and C viruses with known degrees of fitness differences in PBMC of US blood donors. Our results showed that there was no significant difference in the overall W_D values between US donors and Indian donor 1 ($p = 0.39$, t test) or US donors and Indian donor 2 ($p = 0.45$, t test) (Table 6 and 7). There was also no significant difference in fitness between Indian donors 1 and 2 ($p=0.12$, t test). We did however observe a considerable difference in replication fitness among US donor 1 and both Indian donors for the A15 vs C31 competition pair at day 7. In this group, the WD for US donor 1 was 175.5 while the WD in Indian donors 1 and 2 were 9.5 and 23.1, respectively. This variability is not surprising considering that PBMC from each individual will support different levels of virus replication. But even with this variation the overall replication fitness of subtype C is higher than subtype A in PBMC from US and Indian blood donors.

Table 7. Relative fitness and relative fitness fold difference values of day 7 GCA using PBMC from Indian donors

	GCA Pairs	Relative W_A †	Relative W_C †	Relative $W_{D(A/C)}$ ‡	Relative $W_{D(C/A)}$ ‡
Indian Donor 1	**C31:A6**	0.6 ± 0.03	1.3 ± 0.24	0.5	2.2
	C31:A15	0.3 ± 0.33	3.2 ± 2.87	0.1	9.5
	C59:A11	0.2 ± 0.24	4.1 ± 2.04	0.1	18.4
Indian Donor 2	**C31:A6**	0.5 ± 0.07	2.4 ± 0.21	0.2	4.4
	C31:A15	0.2 ± 0.08	4.5 ± 0.65	0.0	23.1
	C59:A11	0.2 ± 0.01	4.6 ± 0.04	0.0	29.7

Relative fitness values (W) and fitness fold difference values (W_D) were calculated as described for Table 6.

5.5.4 Replication fitness and transmission efficiency of subtype A and C viruses using an ex vivo organ culture.

The preponderance of adult HIV-1 infections in India is due to heterosexual contact. Using a cervical tissue-derived organ culture we investigated the transmission across cervical mucosa of subtype A and C virus by dually infecting cervical tissue with either primary isolates or molecularly cloned HIV-1. After 3 days of exposure the level of transmitted virus across the cervical mucosa into the bottom well supernatant was measured by TaqMan® real-time RTPCR. Our results show that in 4 out of 6 primary isolate pairs tested, in 6 different cervical tissue specimens, subtype C virus had higher transmission efficiency (replication fitness) than subtype A virus (Figure 24). The fitness difference (WD) for subtype C viruses ranged from 0.72 – 47.03, while the WD for subtype A viruses ranged from 0.02 – 1.49. For two transmission pairs, A11:C31 and A6:C31 subtype A showed slightly higher transmission efficiency than subtype C. However, the subtype A WD values for A11 and A6 were relatively low (range 1.31 - 1.49) compared to the WD values for subtype C isolates (range 1.79 - 47.03) which showed higher transmission efficiency than subtype A. The subtype C primary isolate C31, which was outcompeted by subtype A in cervical tissue, also had the lowest WD value in PBMC GCA (Table 6). These findings imply that subtype C has higher overall transmission efficiency than subtype A although in some cases subtype A may be better transmitted to some extent.

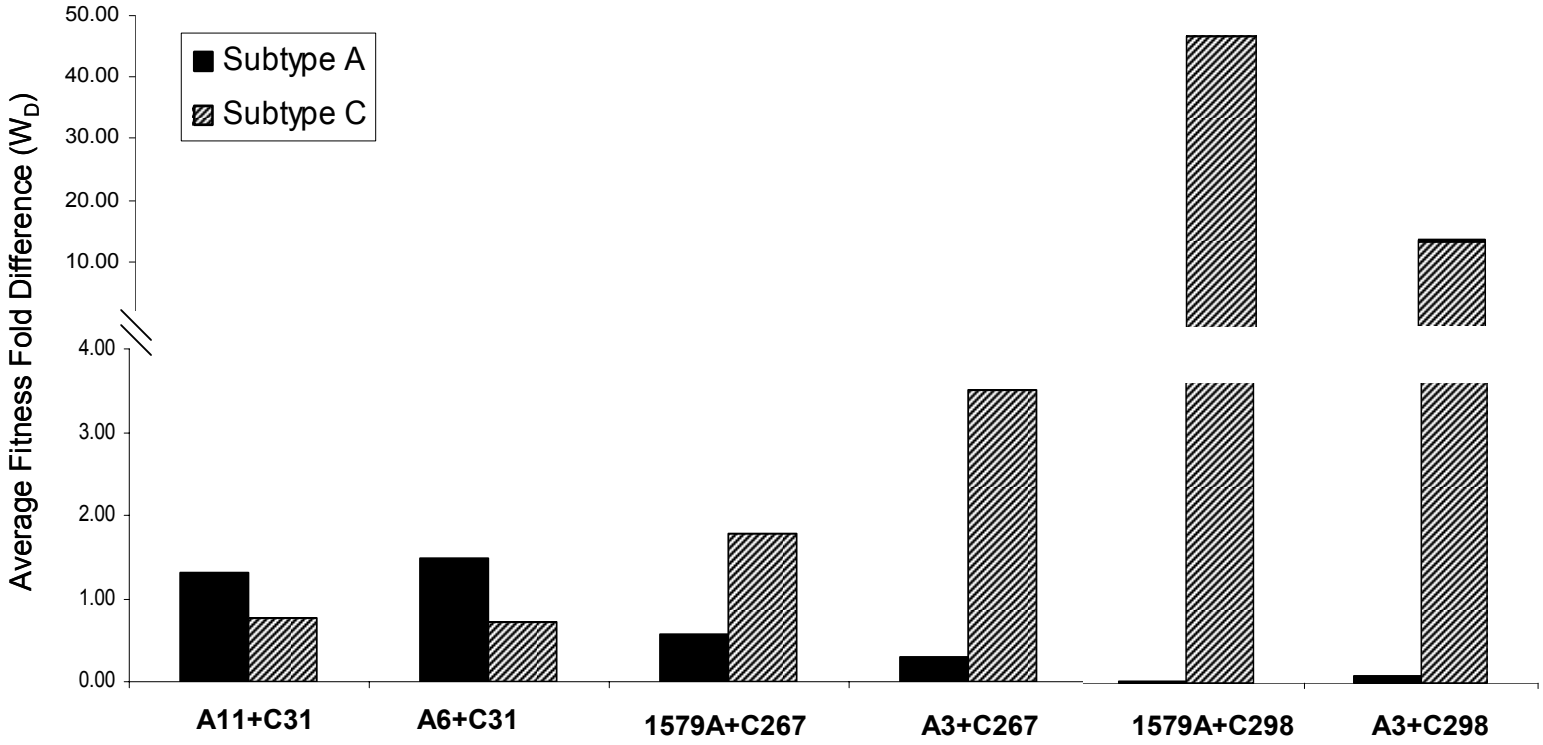

Figure 24. Transmission efficiency of subtypes A and C primary isolates across cervical mucosa

Transmission efficiencies are expressed as fitness difference (fold difference).

5.5.5 Determination of HIV-1 genomic regions responsible for enhanced replication fitness.

Recent reports have suggested that replication fitness may be mediated by the viral envelope gene (52, 106, 107). To investigate this possibility we constructed half-genome chimeric virus by replacing the 3'-half (tat-3' LTR) of an HIV-1 subtype A molecular clone (p1579A-1) with an equivalent 3' half of a subtype C clone (pIndieC1) from India to generate a full length A/C chimeric clone as described in the Materials and Methods section. The A/C chimeric virus was then tested against the parental subtype A virus in GCA using PMBC. As shown in Figure 25 at day 14, p1579A-1 out competed the A/C chimeric clone in the GCA. However, the A/C chimera outcompeted subtype C pIndieC1 in CGA. These results would suggest that genomic elements in the 3'-half of subtype C are not attributable to replication fitness differences between subtype A and C virus of Indian origin.

We also investigated the differences in transmission efficiency of parental clones and chimeric A/C virus. When tissues were dually infected with subtype A/C chimeric and parental subtype A (p1579A-1) virus we observed results similar to those obtained in GCA using PBMC in which the subtype A appeared to have higher transmission efficiency than the chimeric clone which contained the 3' half (tat – 3'LTR) of subtype C (Figure 26). However, when tissues were dually infected with parental subtype C clone and A/C chimera, the subtype C pIndieC1 had lower transmission efficiency than the A/C recombinant. These results suggest that the 3'-half of the genome works collectively with genes in the 5'-half of the genome to impart higher replication fitness.

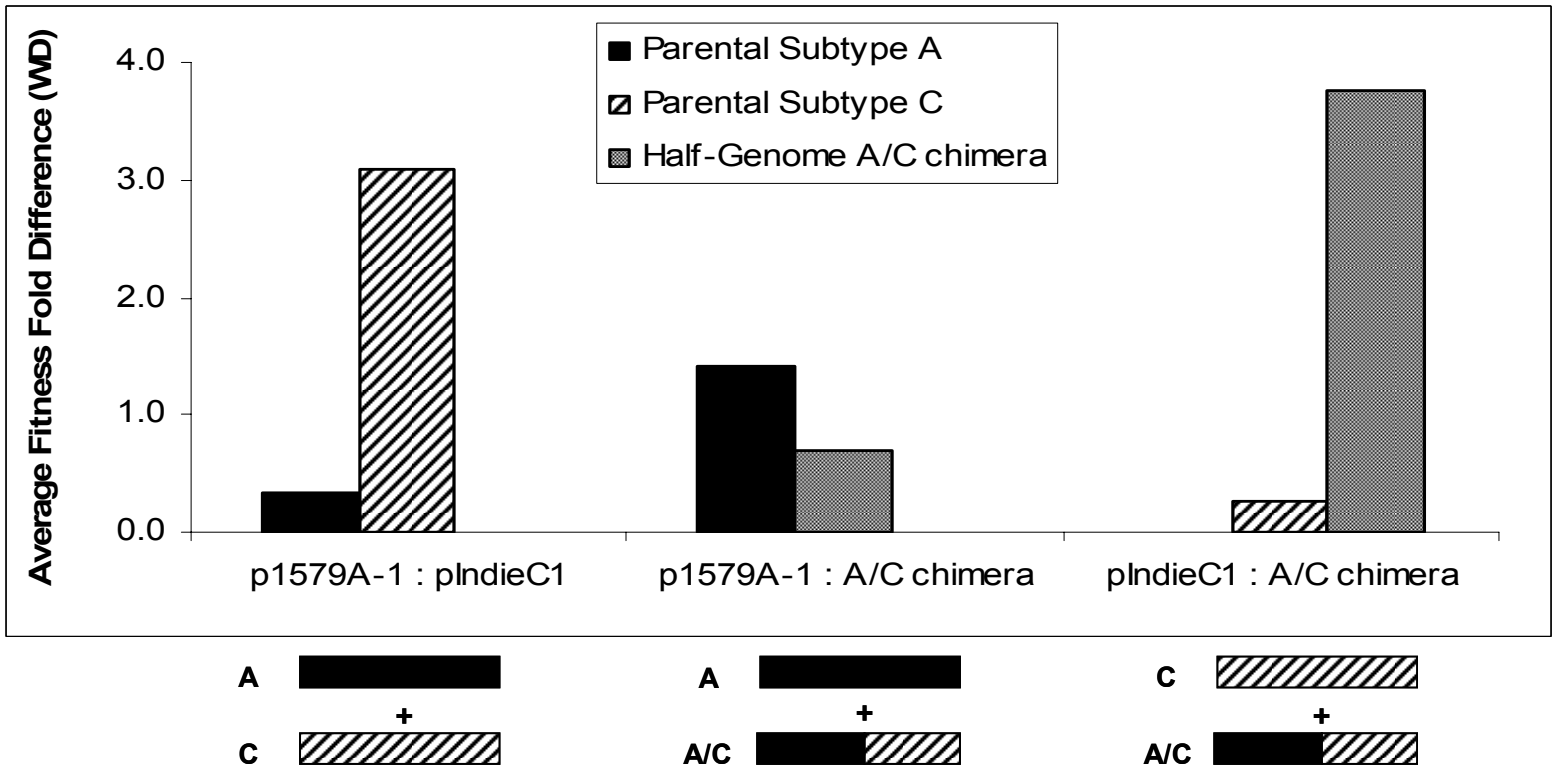

Figure 25. Replication fitness of parental and half genome chimeric cloned viruses. Replication fitness is expressed as fitness fold difference (WD) of each HIV-1 variant in a GCA performed in PBMC.

A schematic representation of HIV-1 parental and half genome clone is shown below the graph.

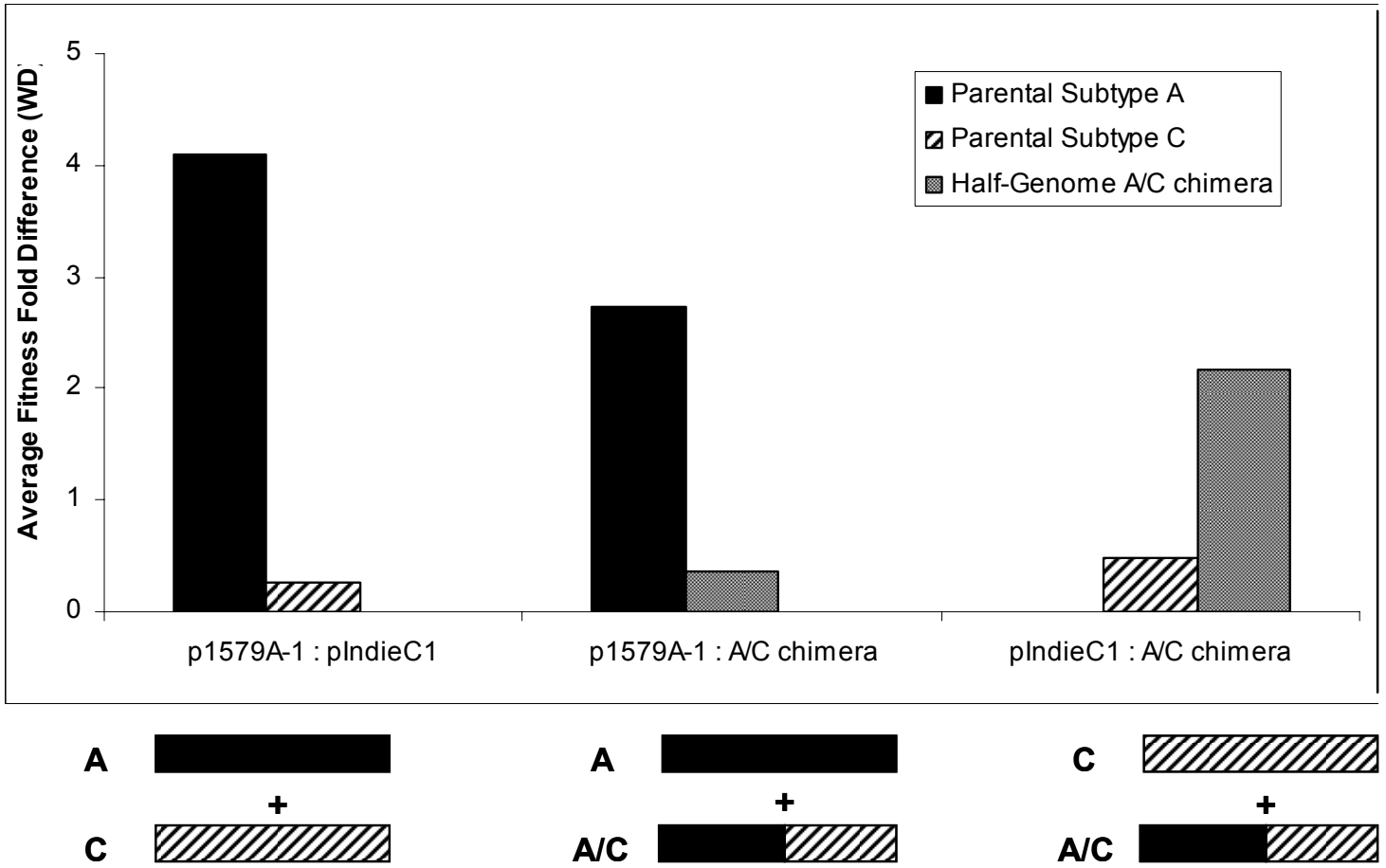

Figure 26. Transmission efficiency of subtypes A and C cloned virus and A/C chimeric virus across cervical mucosa.

Fitness fold difference (W_D) for each competition pair is indicated in the diagram below the graph.

5.5.6 Investigation of a trans suppressive mechanism of subtype A replication by subtype C in *in vitro* GCA.

As shown above we have observed that HIV-1 subtype C outcompeted HIV-1 subtype A in the GCA. We investigated the possibility of whether such competition of subtype A by subtype C virus is mediated through a trans suppressive mechanism of subtype A by a soluble factor secreted by subtype C infected cells. We tested this by growing subtype A virus in the presence of virus-free culture supernatant from subtype C-infected cells with or without pretreatment with such culture supernatant.

We observed that virus-free culture supernatant from monoinfected subtype C PBMC did not suppress the growth of subtype A, regardless of whether the cells were preincubated with the culture supernatant (Figure 27). In fact, the subtype C virus-free culture supernatant induced non–specific activation of cells and promoted increased growth of subtype A virus compared to cells infected with subtype A virus and cultured in IL-2 media only. As a control, virus-free culture supernatant from subtype A infected cells also did not suppress growth of subtype A virus, but rather promoted increased growth of subtype A virus. Such activation of viral growth did not occur in the presence of cell culture supernatant from uninfected cells. These results indicate that soluble factors such as *gag* matrix proteins or capsid protein which are produced in abundance from virus infected cells such as those infected by subtype C virus are not responsible for competitive growth of HIV-1 subtype C over subtype A in the GCA.

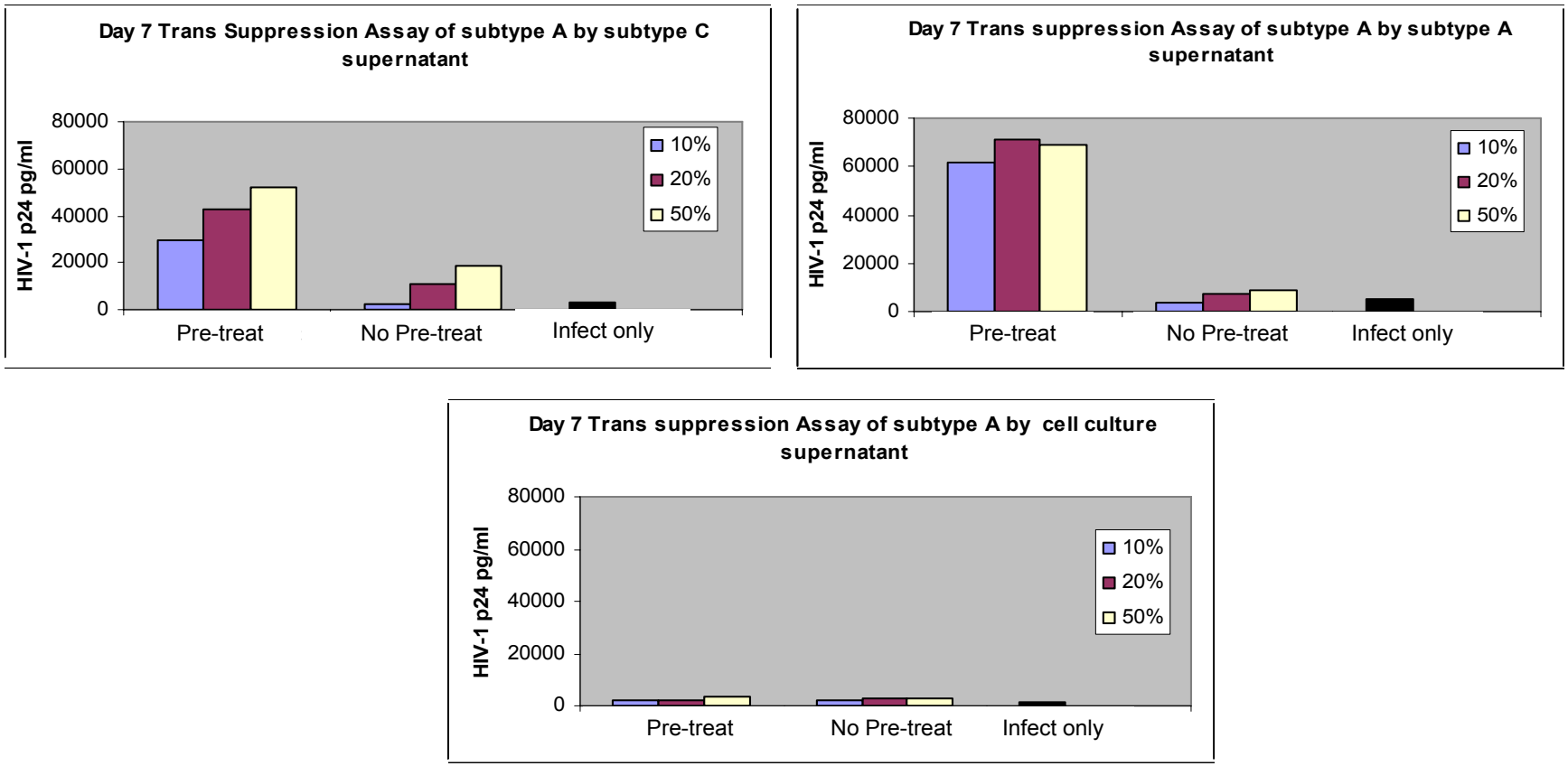

Figure 27. Trans suppression of HIV-1 subtype A by virus-free supernatant from infected cells or by culture supernatant from uninfected cells.

Cells were either pretreated with virus-free supernatant or not pre-treated, followed by infection with HIV-1 subtype A then cell culture in the presence of increasing percentages of the final culture volume of virus-free supernatant. Infect only indicates cells that were infected with HIV-1 subtype A and not pre-treated or grown in the presence of virus-free supernatant. Negative controls included 1) pre-treatment, culture in the presence of virus-free supernatant and no infection; 2) no pre-treatment, culture in the presence of virus-free supernatant and no infection and 3) cells only, no infection. All negative controls showed no virus growth as determined by HIV-1 p24 assay (p24 $\leq$ 30pg/ml at day 14, data not shown).

5.6 DISCUSSION

There is a growing body of data supporting the hypothesis that the differential spread of HIV variants, groups and subtypes in the human population may be related to differences in replication fitness and transmission efficiency (51, 52, 54, 58, 77). The seminal study of Quinones-Mateu *et al* found that there is a correlation between *ex vivo* fitness and disease progression. This study was useful in demonstrating that fitness values may serve as a predictor of progression to AIDS. These studies form the basis that in vitro replicative fitness data can be used to study the interaction between two subtypes in transmission and disease progression.

Recently using a GCA assay, Arien *et al* assigned an order of relative fitness to all classes of HIV viruses: HIV-1 group M > HIV-2 >HIV-1 group O. This study implied that the lower replicative capacity of group O and HIV-2 compared to group M viruses may have led to decreased transmission and distribution of these groups in the human population. In a recent study, *ex vivo* replicative fitness of CRF02_AG isolates was found to be higher than subtype A and G viruses from the same geographic region and was independent of the co-receptor tropism and irrespective of high or low CD4+ T cell count (55). Over the last ten years CRF02_AG has become the predominant subtype in West and West Central Africa. These results suggest that replicative fitness may have contributed to the dominant spread of CRF02_AG over A and G subtypes in West and West Central Africa (55).

Genetic analysis of HIV-1 circulating in different parts of India at different times in the last fifteen years indicates that the subtype C predominates in India with a small proportion of infection caused by subtypes A or B (13, 24, 27, 31, 129). These Indian HIV-1 are of CCR5

tropism with a NSI phenotype even when they are isolated at late stages of infection (13, 110). Additionally, we have shown that HIV-1 subtype C from India is distinct from other HIV-1 subtypes C around the world, like those from Africa and South America. HIV-1 infection in some parts of Africa started predominantly with subtype C, but later other non-C subtypes emerged to account for as much as 40% of circulating subtypes in that region. In contrast, HIV-1 subtype C in India has predominated over other subtypes throughout the entire epidemic and regardless of the geographic area (13, 24, 27, 31, 37, 129). Our results indicate that PBMC from the Indian population do not support higher replication of subtype C as compared to subtype A of Indian origin, indicating that genetics of the Indian population may not be responsible for the asymmetric distribution of subtype C over other subtypes in India. This distinct asymmetric distribution of HIV-1 subtypes in India suggests that there may be a biological basis for such dissemination. The mechanism for such an asymmetric distribution could be due to any one or combination of factors including replication fitness, transmission efficiency and founder effect. In this report we have examined the first two possibilities using subtype A and C from India.

We have used a CGA assay in CD8-depleted PBMC to study replicative fitness among these two subtypes. Our data indicate that among 11 subtype A and C pairs tested, all primary subtype C isolates had higher replicative fitness (mean fitness score of 57.7) than subtype A isolates when cells were dually infected at a ratio of 1:1. Even at 10-fold less inoculum than subtype A, subtype C isolates outcompeted subtype A in most cases. These results demonstrate that subtype C HIV-1 isolates from India have higher replicative fitness than subtype A from India in PBMC. Such replicative fitness of subtype C over subtype A also has been extended to PBMC of Indian origin, indicating that such replicative fitness of subtype C is independent of ethnic origin.

These data contradict the results obtained by Arien *et al* (53), in which subtype C viruses from Africa were found to be the least fit among all HIV-1 group M viruses in GCAs performed using PBMC and in an vitro DC-T cell transmission model.. This could be due to that fact Indian subtype C is different than subtype C from Africa and South America.. Since Langerhans cells in cervical mucosa have been implicated for HIV-1 transmission across cervical mucosa, we examined replicative fitness between subtype A and subtype C from India in a cervical tissue derived organ culture. Use of cervical tissue derived organ culture provides direct measurement of HIV transmission across the cervical mucosa. Our data indicate that in 4 out of 6 pairs tested, subtype C had higher transmission efficiencies than subtype A. Two pairs (A6:C31 and A11:C31) in which subtype C had shown lower transmission efficiency than subtype A, also showed relatively lower replicative fitness scores (3.3 and 4.7, respectively) in PMBC as compared to other competing pairs (mean 57.7). These observations along with recent a report demonstrating a correlation between subtype C infected women and increased vaginal shedding of virus (58) and another describing compartmentalization of subtype C virus in the cervix of a woman dually infected with subtype A and C viruses (130), indicate that subtype C may have transmission advantages which could have influenced its spread throughout India and regions of Africa.

To understand the molecular mechanism of higher in vitro replicative fitness of subtype C over subtype A, we have examined a trans mechanism for higher replication fitness subtype C by which a soluble factor(s) secreted by subtype C infected cells confer suppression of subtype A replication. Our data indicate that virus-free culture supernatant from monoinfected subtype C PBMC did not suppress the growth of subtype A, regardless of whether the cells were preincubated with the culture supernatant or not. In fact, the subtype C virus-free culture

supernatant induced non–specific activation of cells and promoted increased growth of subtype A compared to cells infected with subtype A and cultured in IL-2 media only.

We examined the role of LTR and envelope gene to elucidate a cis mechanism of higher replicative fitness of subtype C over subtype A. In a previous study we have shown that there was no significant difference in the structure and function of the LTR between subtypes A and C from India even though subtype C was predicted to have an additional NF-κB site (Rodriguez *et al*, ARHR in press). To understand the role of envelope gene, we have constructed recombinants between subtype A and subtype C in which the 3' half of the subtype A has been replaced with subtype C. A competition between this recombinant and parental wild type subtype A and C in a GCA assay and transmission assay across the cervical mucosa indicate that the 3'-half of the genome is not directly responsible for differences in replication fitness. In addition, the increased fitness of the A/C chimera over subtype C may indicate the possible fitness and epidemic potential of an A/C recombinant having this genomic structure, should one emerge in India. To date, only one A/C recombinant has been reported in India (ref). This is most likely due to the low frequency of subtype A viruses circulating in the country and to HIV-1 subtype determination restricted to specific regions of the genome. Further investigations of HIV-1 recombinant viruses circulating in India would help to shed light on the possible emergence of a recombinant virus variant of this nature. Finally, the higher replicative fitness and increased transmission across the mucosa of subtype C over subtype A may be due to a more complex interaction between the genes located within both halves of viral genome.

The data supporting a correlation between in vitro viral replication fitness in PBMC and the global distribution and epidemiologic spread of HIV is continuously expanding. Our investigation sheds new light on the potential role of replication fitness and transmission

efficiency to the asymmetric distribution of HIV-1 subtypes in India. This study also shows that HIV-1 subtypes with higher fitness could have a replication advantage over other subtypes which may shape the epidemic and current global subtype distribution.

Acknowledgements: This work was supported by the National Institutes of Health through grant U19AI51661.

Ming Ding performed real-time RTPCR analysis for all growth competition assays.

Deena Ratner propagated, expanded and isolated HIV-1 from PBMC obtained from HIV-1 infected patients from India. She also provided technical assistance for transmission studies.

Yue Chen provided technical assistance with construction of both p1579A and A/C half genome molecular clones.

Ramesh S. Paranjape and Smita S. Kulkarni provided primary isolates from HIV-1 subtype A infected patients from India.

Ramadas Chatterjee provided PBMC from HIV-1 subtype C infected patients from India

Patrick M. Tarwater performed statistical analyses.

6.0 OVERALL DISCUSSION AND FUTURE DIRECTIONS

6.1 SUMMARY OF FINDINGS

Genetic analysis of HIV-1 circulating in different parts of India at different times in the last fifteen years indicates that subtype C predominates with a small proportion of infections caused by subtypes A or B. Furthermore, HIV-1 subtype C from India has been shown to be distinct from other HIV-1 subtype C around the world, like those from Africa and South America. The distinct asymmetric distribution of HIV-1 subtypes in India suggests that there may be a biological basis for such dissemination of HIV-1 subtypes. Based on studies showing a correlation between in vitro replication fitness and disease progression and those suggesting a correlation between replication fitness and HIV-1 subtype geographic distribution, we hypothesized that the HIV-1 subtype C from India would display higher in vitro replication fitness than subtype A from India. To investigate this hypothesis we developed the following specific aims: 1) to construct and characterize an infectious molecular clone of HIV-1 subtype A of Indian origin; 2) to evaluate the replication fitness of HIV-1 subtypes A and C from India using an in vitro growth competition assay; 3) to compare the transmission efficiency of HIV-1 subtypes A and C from India across the mucosa of cervicovaginal tissue; and 4) to evaluate the role of the LTR and e*nv* in replication fitness and transmission efficiency across cervicovaginal tissue.

The study in chapter one describes the construction of an HIV-1 subtype A infectious molecular clone of Indian origin. To do this we cloned, sequenced and characterized virus from an HIV-1 subtype A primary isolate from India. We determined that the clone, p1579A-1, is replication competent in primary cells, that it uses CCR5 as its cellular coreceptor to gain entry into permissive cells and that it has a non-syncytium inducing phenotype, indicating low pathogenic potential. The p1579A-1 envelope gene was also found to have DNA and amino acid sequences similar to those found in HIV-1 subtypes A1 and A2 from other parts of the world. Additionally, we also found similarity in the number of potential N-linked glycoslylation sites within e*nv* between p1579A-1 and other subtype As. This study was the first of its kind to delineate and characterize the complete genomic sequence of an HIV-1 subtype A from India. Availability of this clone will be useful for the study of replication dynamics and evolution of subtypes in India and for characterizing other subtype A around the world.

The studies presented in chapter two were designed to characterize the structure and function of the LTR region to determine whether the predominance of subtype C is due to higher functional attributes of subtype C LTR compared to other subtypes in India. To do this we PCR amplified the long terminal repeat (LTR) region from a panel of subtype A and subtype C primary isolates and infectious molecular clones and cloned them into a reporter construct for functional analysis. The most notable structural difference that was observed was an additional NF-κB binding site in subtype C LTRs that was not predicted for subtype A LTRs. However, despite this additional enhancer motif, we did not observe any significant difference in the transcriptional activity between subtype A and subtype C LTRs either in the presence or absence of viral tat protein or after stimulation with TNF-alpha or PMA. These results were unexpected since several studies have reported functional differences between the LTRs of HIV-1 subtypes

from several different regions of the world under similar conditions (98, 104, 131, 132). However, in our study, no significant differences have been reported between subtypes A and C from India. These results suggest that replication capacity of Indian HIV-1 subtypes is not solely dependent on transcriptional events and that there is a more complex interplay of virus and host cell factors which determine the rate and extent of virus production.

The studies in chapter three sought to investigate the in vitro replication fitness of subtypes A and C of Indian origin in an effort to understand the mechanism of subtype C predominance in India. Using a dual infection growth competition assay (GCA) and a panel of subtype A and subtype C primary isolates and infectious molecular clones of both subtypes, we found that subtype C primary isolates outcompeted subtype A primary isolates in PBMC by as much a 255-fold when competed at equal infectivity. Likewise, using an ex vivo cervical tissue derived organ culture transmission model which closely mimics the in vivo cervicovaginal tissue architecture, the majority of subtype C displayed higher transmission efficiency than subtype A. We also observed that there was no significant difference in overall fitness values when competitions were performed using US or Indian PBMC. This result served to rule out the possibility of genetic factors that could be predisposing the Indian population to support higher subtype C replication.

This investigation reports genetic and functional characterization of HIV-1 subtype A and subtype C of Indian origin which has contributed new HIV-1 genomic sequence information that is important for characterizing HIV-1 subtypes and which may be necessary for design of subtype-specific vaccines and antiviral compounds. Additionally, this investigation sheds new light on the potential role of replication fitness and transmission efficiency as contributing factors to HIV-1 subtype distribution in India. Here, unlike previous findings, in which subtype

C has been shown to be the least fit among all HIV-1 group M subtypes, we report a higher replication fitness of subtype C over subtype A. This has implications to the unique properties of subtype C from India compared to other HIV-1 group M subtypes. Based on the findings from this study we propose an epidemiological disease model for India in which HIV-1 subtype C has acquired higher replication capacity and increased transmissibility compared to subtype A. These properties have been advantageous and beneficial for subtype C spread throughout human hosts in India and has resulted in the current predominance of subtype C in India.

6.2 PUBLIC HEALTH SIGNIFICANCE

HIV-1 is a rapidly mutating and highly adaptable pathogen which results in an increase in viral genetic diversity and potentially an evolutionary edge (11). This high genetic diversity has resulted in multiple subtypes, circulating recombinant forms (CRF), unique recombinant forms (URF) and intrapatient quasispecies. This variability highlights the difficulty and importance of obtaining broad protection against HIV-1 in the form of vaccines, antiviral drugs and microbicides. The findings presented in this study have increased our understanding of the mechanism of asymmetric distribution of HIV-1 subtypes in India by highlighting the importance and uniqueness of HIV-1 subtypes of Indian origin. As we have described in this work, HIV-1 subtype C of Indian origin has higher replication fitness and increased transmissibility compared to subtype A of Indian origin. Therefore, control strategies, such as the design and use of antiviral compounds, in India should be predominantly geared towards subtype C since it may continue to predominate over subtype A and other subtypes circulating in this country. As well, vaccine and antiviral compound design efforts should encompass all regions of the genome since we have shown that the envelope gene may not be the only determining factor of replication and transmission differences. Although a direct correlation has not been made between in vitro replication fitness and global subtype distribution, investigations of fitness differences among HIV-1 subtypes in India should be ongoing. Our findings that A/C recombinant virus had higher replication fitness than subtype C virus may be especially important for future projections considering that within the last 10 years, CRFs have become predominant in regions where two or more subtypes co-circulate (133). These types of studies

may be useful in forecasting the emergence of genetic variants with altered biologic properties that could eventually lead to their rapid spread and predominance.

6.3 THE IMPORTANCE OF STUDYING HIV-1 SUBTYPES

6.3.1 HIV-1 subtypes and neutralizing antibody responses

Neutralizing antibodies (Nabs) against viral envelope proteins (*env*) provide the first line of adaptive immune defense against HIV-1 infection by blocking the infection of susceptible cells (134). Opposing studies have shown neutralization activities by serum samples and monoclonal antibodies (MAbs) against HIV-1 virus subtypes to be both cross-reactive (134-136) and subtype specific (91, 134). Recently Binley *et al* (134), conducted a large-scale and comprehensive investigation to evaluate MAbs from subtype B-infected donors and a subtype B HIV^{+} plasma against a panel of viruses from diverse backgrounds. Overall, MAbs directed against gp120 or gp41 or the CD4 binding domain exhibited greater activity against subtype B than non-B viruses. Out of eight MAbs tested, only two, b12 and 4E10, both directed at the CD4 binding site, were effective against subtype C viruses. Based on this study it seems important to continue to investigate neutralizing antibody responses to non-B subtype viruses to determine those that are the most broadly reactive. This is especially true in the case of subtype C which is the most predominant subtype worldwide.

6.3.2 HIV-1 subtypes and cell-mediated immune responses

It is generally accepted that cross-reactive cytotoxic T lymphocytes are present in individuals infected by non-B subtypes (137, 138). However, intrasubtype CTL responses are usually stronger and more frequently detected than cross-clade (subtype) reactivities (5). Furthermore, differences in the frequency and relative magnitude of subtype-specific responses has been reported (139). These differences were generally observed based on the level of diversity/conservation of the epitope. Subtype-specific CTL epitopes have also been reported (140). Despite the observed subtype cross-reactive responses, a successful vaccine has not been developed.

6.3.3 HIV-1 subtypes and drug resistance

Virtually all available evidence suggests that all HIV-1 subtypes display similar sensitivities to antiviral drugs, but viruses from some subtypes may have a greater propensity to develop resistance to certain drugs or may contain inherent variability rendering them resistant (141). Most available antiretroviral drugs are directed towards the RT (reverse transcriptase) and PR (protease) proteins of Pol and are based on HIV-1 subtype B sequences (141). Variations in these regions may therefore affect drug susceptibility and the development of drug resistance. In the case of subtype C viruses, studies have reported more rapid selection of resistance variants to NNRTI and significantly lower concentration of drug required for development of resistance to NNRTIs as compared to subtype B (142-144). There have also been reports of development of novel resistance mutations of subtype C isolates to the NNRTIs nevirapine and efavirenz (142). In addition diminished drug sensitivity due to rapid growth kinetics has been reported for other

HIV-1 subtypes (145). These findings are especially significant due to the substantial number of people living with HIV/AIDS in developing countries that likely do not have access to highly active antiretroviral therapy (HAART).

6.3.4 HIV-1 and microbicides

Microbicides are products that can be applied to vaginal or rectal mucosal surfaces with the goal of preventing or reducing the transmission of sexually transmitted diseases including HIV-1 (62, 67). The development of a microbicide against HIV-1 is especially important since the majority of HIV-1 infections are transmitted sexually. If effective, a microbicide may offer an alternative method of protection to a vaccine which to date, has not been developed and may take several more years. There is very little data comparing HIV-1 subtype and microbicide activity. One study which reported the antiviral activity of commercial microbicides against HIV-1 subtypes A, B and C reported similar activity of each microbicide against each HIV-1 variant (167). In some cases however, subtype C virus infection was inhibited the least compared to the other subtypes.

Microbicides can be classified by their primary mechanism of action. Vaginal defense enhancers help maintain the vaginal pH in an acidic range or facilitate colonization of vaginal flora with lactobacilli (67). Agents that act in this respect are Acidform™(146), BufferGel™(147), and Lactobacilli (*Lactobacillus crispatus*)(148-150).

Surfactants or detergents disrupt microbial cell membranes. While they can be effective against the HIV-1 membrane, surfactants can also disrupt cell membranes which can lead to toxicity thus increasing the risk of infection. Unfortunately this was the situation observed in clinical trials of the surfactant Nonoxynol-9 (N-9) leading to its termination as a potential

microbicide (151, 152). Other surfactants that have been better tolerated than N-9 are C31G (Savvy™) (153, 154) and the Invisible condom™ (sodium lauryl sulfate, SLS) (155). Of these two, SLS, whose mechanism involves a combination of solubilization of viral envelopes and denaturation of envelope or capsid proteins; is currently in a phase 1/2 study (156). Each of the above mentioned agents have the potential to perform equally well against various HIV-1 subtypes.

Entry or fusion inhibitors target viral epitopes or cell receptors to prevent infection of cells. The agents that are particularly directed against viral epitopes probably have the most potential to display varying degrees of efficacy due to the highly variable nature of HIV – especially in the envelope gene. Entry inhibitors directed against viral epitopes in gp120 or gp41 include polyanions such as PRO-2000 (157), Dextrin sulfate (158-160), Cellulose sulfate (161), carrageenans (Carraguard™) (162) and cellulose acetate phthalate (CAP) (163). Cyanovirin-N (CV-N) is an agent that binds high mannose residues in the HIV envelope preventing fusion with the host cell membrane. CV-N would probably be more effective against a broad range of HIV-1 subtypes however the cost of this agent for optimal protection is probably unrealistic for the target population in greatest need. PSC-RANTES is a second generation synthetic CCR5 antagonist. In vitro studies using PBMC targets have shown that it is active against all HIV subtypes (164) and studies in macaques displayed protection against vaginal challenge with a SHIV isolate (157).

Finally, replication inhibitors act by preventing reverse transcription of the viral RNA genome. Reverse transcriptase (RT) inhibitors can be delivered topically or orally. These types of agents include Tenofovir (67) and UC-781 (165). There are several advantages to the use of these drugs as microbicides, however the major disadvantage is the possibility of the emergence

of drug resistant mutants. As discussed above in section 6.2.3, subtype C viruses may respond differently to these drugs and may have a propensity to develop resistance mutations sooner than other HIV-1 subtypes.

Taken together our findings highlight the importance of continued investigation and characterization of genetic differences of HIV-1 subtypes and how these differences could affect replication capacity which in itself could prompt the emergence of immune escape and drug resistant variants.

6.4 FUTURE DIRECTIONS

6.4.1 Construction of *env* recombinant viruses

There are only a few recent studies which have shown that replication fitness is mediated by e*nv*. Using an HIV-1 neutral backbone such as pNL4-3 or HXB2, a comprehensive panel of Indian subtype A and subtype C e*nv* recombinant clones can be generated to investigate differences in replication fitness solely based on e*nv*. Additionally, e*nv* chimeras can be constructed to determine the region of the envelope gene responsible for enhanced replication fitness.
Subsequently point mutations can be made to pinpoint DNA and amino acid sequences involved in this mechanism/fitness advantage.

6.4.2 Investigation of early timepoints in fitness

We have reported that the "winner" in a dual infection is generally determined by day 7 when two variants are competed at equal TCID 50 ratios. We do not know however, the nature of viral dynamics at early timepoints. Has the winner already been determined after the first round of replication or does it take several hours or days? Investigations into the early events during the competition would address these concerns. As well, looking into the amount of virus that has infected cells at early timepoints by quantitation of viral RNA inside cells would explain if it is viral burst size from infected cells that allows the winner an advantage over the outcompeted variant.

6.4.3 Continuing investigations of the long terminal repeat

Several studies have reported differences in the transcriptional activity of HIV-1 subtype-specific LTR based on cell type, host cell signaling pathways and *tat* protein transactivation. Macrophages are also an important subset of cells found in the subepithelial mucosa that are susceptible to HIV-1 infection by CCR5 using viruses. LTR reporter constructs can be investigated in macrophages, using a subtype-matched *tat* protein and in the presence of biologically relevant concentrations of stimuli for transcriptional activity.

6.4.4 Replication fitness in different cell types

As discussed for the LTR, it is also important to investigate the replication fitness dynamics of primary isolates in other cell types such as macrophages. These cells are present in the cervical mucosa and blood and would also elucidate mechanisms of in vivo replication fitness differences.

6.4.5 Subtype C response to antiretroviral compounds

The foundation of all of these investigations is to make significant contributions to the field of HIV-1 research which would aid in therapeutic drug and vaccine design. Therefore based on the data presented here the response of HIV-1 subtype C viruses of Indian origin to antiretroviral drugs and microbicides is a next logical step. It may be revealed that not only is Indian subtype C genetically and functionally unique from other subtype C but that it also displays a unique pattern of drug recognition and/or resistance. This type of investigation would put into perspective the additional importance of necessary compounds designed specifically for subtype C of Indian origin. This should be of the highest priority given the vast number of people in India infected with this subtype of HIV-1.

BIBLIOGRAPHY

1. McCutchan FE. Global epidemiology of HIV. J Med Virol 2006;78 Suppl 1:S7-S12.

2. Levy JA. HIV and the pathogenesis of AIDS. 3rd ed. Washington, D.C.: ASM Press; 2007.

3. UNAIDS/WHO (2006) AIDS Epidemic Update; December 2006. www.unaids.gov.

4. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, *et al*. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature 1999;397(6718):436-41.

5. Thomson MM, Perez-Alvarez L, Najera R. Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. Lancet Infect Dis 2002;2(8):461-71.

6. Biebricher CK, Eigen M. What is a quasispecies? Curr Top Microbiol Immunol 2006;299:1-31.

7. Geretti AM. HIV-1 subtypes: epidemiology and significance for HIV management. Curr Opin Infect Dis 2006;19(1):1-7.

8. Gottlinger HG. The HIV-1 assembly machine. Aids 2001;15 Suppl 5:S13-20.

9. Karageorgos L, Li P, Burrell C. Characterization of HIV replication complexes early after cell-to-cell infection. AIDS Res Hum Retroviruses 1993;9(9):817-23.

10. Greene WC, Peterlin BM. Charting HIV's remarkable voyage through the cell: Basic science as a passport to future therapy. Nat Med 2002;8(7):673-80.

11. Henry KR, Weber J, Quinones-Mateu ME, Arts EJ. The impact of viral and host elements on HIV fitness and disease progression. Curr HIV/AIDS Rep 2007;4(1):36-41.

12. Pantaleo G, Graziosi C, Fauci AS. New concepts in the immunopathogenesis of human immunodeficiency virus infection. N Engl J Med 1993;328(5):327-35.

13. Shankarappa R, Chatterjee R, Learn GH, Neogi D, Ding M, Roy P, *et al*. Human immunodeficiency virus type 1 env sequences from Calcutta in eastern India:

identification of features that distinguish subtype C sequences in India from other subtype C sequences. J Virol 2001;75(21):10479-87.

14. Grossman Z, Meier-Schellersheim M, Paul WE, Picker LJ. Pathogenesis of HIV infection: what the virus spares is as important as what it destroys. Nat Med 2006;12(3):289-95.

15. Picker LJ, Watkins DI. HIV pathogenesis: the first cut is the deepest. Nat Immunol 2005;6(5):430-2.

16. Brenchley JM, Schacker TW, Ruff LE, Price DA, Taylor JH, Beilman GJ, *et al*. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. J Exp Med 2004;200(6):749-59.

17. Daar ES, Moudgil T, Meyer RD, Ho DD. Transient high levels of viremia in patients with primary human immunodeficiency virus type 1 infection. N Engl J Med 1991;324(14):961-4.

18. Cavert W, Notermans DW, Staskus K, Wietgrefe SW, Zupancic M, Gebhard K, *et al*. Kinetics of response in lymphoid tissues to antiretroviral therapy of HIV-1 infection. Science 1997;276(5314):960-4.

19. Steinbrook R. HIV in India--a complex epidemic. N Engl J Med 2007;356(11):1089-93.

20. Arora P, Cyriac A, Jha P. India's HIV-1 epidemic. Cmaj 2004;171(11):1337-8.

21. Mills E. More on India's HIV-1 epidemic. Cmaj 2005;173(1):17-8; author reply 18.

22. Cohen J. HIV/AIDS in India. HIV/AIDS: India's many epidemics. Science 2004;304(5670):504-9.

23. Delwart EL, Mullins JI, Gupta P, Learn GH, Jr., Holodniy M, Katzenstein D, *et al*. Human immunodeficiency virus type 1 populations in blood and semen. J Virol 1998;72(1):617-23.

24. Maitra A, Singh B, Banu S, Deshpande A, Robbins K, Kalish ML, *et al*. Subtypes of HIV type 1 circulating in India: partial envelope sequences. AIDS Res Hum Retroviruses 1999;15(10):941-4.

25. Mandal D, Jana S, Bhattacharya SK, Chakrabarti S. HIV type 1 subtypes circulating in eastern and northeastern regions of India. AIDS Res Hum Retroviruses 2002;18(16):1219-27.

26. Mandal D, Jana S, Panda S, Bhattacharya S, Ghosh TC, Bhattacharya SK, *et al*. Distribution of HIV-1 subtypes in female sex workers of Calcutta, India. Indian J Med Res 2000;112:165-72.

27. Tripathy S, Renjifo B, Wang WK, McLane MF, Bollinger R, Rodrigues J, *et al*. Envelope glycoprotein 120 sequences of primary HIV type 1 isolates from Pune and New Delhi, India. AIDS Res Hum Retroviruses 1996;12(12):1199-202.

28. Cassol S, Weniger BG, Babu PG, Salminen MO, Zheng X, Htoon MT, *et al*. Detection of HIV type 1 env subtypes A, B, C, and E in Asia using dried blood spots: a new surveillance tool for molecular epidemiology. AIDS Res Hum Retroviruses 1996;12(15):1435-41.

29. Dietrich U, Grez M, von Briesen H, Panhans B, Geissendorfer M, Kuhnel H, *et al*. HIV-1 strains from India are highly divergent from prototypic African and US/European strains, but are linked to a South African isolate. Aids 1993;7(1):23-7.

30. Gadkari DA, Moore D, Sheppard HW, Kulkarni SS, Mehendale SM, Bollinger RC. Transmission of genetically diverse strains of HIV-1 in Pune, India. Indian J Med Res 1998;107:1-9.

31. Delwart EL, Shpaer EG, Louwagie J, McCutchan FE, Grez M, Rubsamen-Waigmann H, *et al*. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 env genes. Science 1993;262(5137):1257-61.

32. Grez M, Dietrich U, Balfe P, von Briesen H, Maniar JK, Mahambre G, *et al*. Genetic analysis of human immunodeficiency virus type 1 and 2 (HIV-1 and HIV-2) mixed infections in India reveals a recent spread of HIV-1 and HIV-2 from a single ancestor for each of these viruses. J Virol 1994;68(4):2161-8.

33. Ljungberg K, Hassan MS, Islam MN, Siddiqui MA, Aziz MM, Wahren B, *et al*. Subtypes A, C, G, and recombinant HIV type 1 are circulating in Bangladesh. AIDS Res Hum Retroviruses 2002;18(9):667-70.

34. Sahni AK, Prasad VV, Seth P. Genomic diversity of human immunodeficiency virus type-1 in India. Int J STD AIDS 2002;13(2):115-8.

35. Harris ME, Maayan S, Kim B, Zeira M, Ferrari G, Birx DL, *et al*. A cluster of HIV type 1 subtype C sequences from Ethiopia, observed in full genome analysis, is not sustained in subgenomic regions. AIDS Res Hum Retroviruses 2003;19(12):1125-33.

36. Lal RB, Chakrabarti S, Yang C. Impact of genetic diversity of HIV-1 on diagnosis, antiretroviral therapy & vaccine development. Indian J Med Res 2005;121(4):287-314.

37. Jameel S, Zafrullah M, Ahmad M, Kapoor GS, Sehgal S. A genetic analysis of HIV-1 from Punjab, India reveals the presence of multiple variants. Aids 1995;9(7):685-90.

38. Tsuchie H, Saraswathy TS, Sinniah M, Vijayamalar B, Maniar JK, Monzon OT, *et al*. HIV-1 variants in South and South-East Asia. Int J STD AIDS 1995;6(2):117-20.

39. Domingo E, Holland JJ. RNA virus mutations and fitness for survival. Annu Rev Microbiol 1997;51:151-78.

40. Yuste E, Sanchez-Palomino S, Casado C, Domingo E, Lopez-Galindez C. Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. J Virol 1999;73(4):2745-51.

41. Chao L. Fitness of RNA virus decreased by Muller's ratchet. Nature 1990;348(6300):454-5.

42. Clarke DK, Duarte EA, Elena SF, Moya A, Domingo E, Holland J. The red queen reigns in the kingdom of RNA viruses. Proc Natl Acad Sci U S A 1994;91(11):4821-4.

43. Domingo E, Menendez-Arias L, Quinones-Mateu ME, Holguin A, Gutierrez-Rivas M, Martinez MA, *et al*. Viral quasispecies and the problem of vaccine-escape and drug-resistant mutants. Prog Drug Res 1997;48:99-128.

44. Holland JJ, de la Torre JC, Clarke DK, Duarte E. Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses. J Virol 1991;65(6):2960-7.

45. Novella IS, Duarte EA, Elena SF, Moya A, Domingo E, Holland JJ. Exponential increases of RNA virus fitness during large population transmissions. Proc Natl Acad Sci U S A 1995;92(13):5841-4.

46. Domingo E, Escarmis C, Sevilla N, Moya A, Elena SF, Quer J, *et al*. Basic concepts in RNA virus evolution. Faseb J 1996;10(8):859-64.

47. Quinones-Mateu ME, Arts EJ. HIV-1 fitness: implications for drug resistance, disease progression and global epidemic evolution. In: Kuiken C, Foley, B., Hahn, B., Korber, B., Marx, PA., McCutchan, F., Mellors, JW., Mullins, JI., Sodoroski, J., and Wolinksy, S., editor. HIV Sequence Compendium 2001. Los Alamos, NM: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory; 2001. p. 134-170.

48. Zhu T, Wang N, Carr A, Nam DS, Moor-Jankowski R, Cooper DA, *et al*. Genetic characterization of human immunodeficiency virus type 1 in blood and genital secretions: evidence for viral compartmentalization and selection during sexual transmission. J Virol 1996;70(5):3098-107.

49. Quinones-Mateu ME, Arts EJ. Virus fitness: concept, quantification, and application to HIV population dynamics. Curr Top Microbiol Immunol 2006;299:83-140.

50. Dyer WB, Geczy AF, Kent SJ, McIntyre LB, Blasdall SA, Learmont JC, *et al*. Lymphoproliferative immune function in the Sydney Blood Bank Cohort, infected with natural nef/long terminal repeat mutants, and in other long-term survivors of transfusion-acquired HIV-1 infection. Aids 1997;11(13):1565-74.

51. Quinones-Mateu ME, Ball SC, Marozsan AJ, Torre VS, Albright JL, Vanham G, *et al*. A dual infection/competition assay shows a correlation between ex vivo human immunodeficiency virus type 1 fitness and disease progression. J Virol 2000;74(19):9222-33.

52. Ball SC, Abraha A, Collins KR, Marozsan AJ, Baird H, Quinones-Mateu ME, *et al*. Comparing the ex vivo fitness of CCR5-tropic human immunodeficiency virus type 1 isolates of subtypes B and C. J Virol 2003;77(2):1021-38.

53. Arien KK, Gali Y, El-Abdellati A, Heyndrickx L, Janssens W, Vanham G. Replicative fitness of CCR5-using and CXCR4-using human immunodeficiency virus type 1 biological clones. Virology 2006;347(1):65-74.

54. Konings FA, Burda ST, Urbanski MM, Zhong P, Nadas A, Nyambi PN. Human immunodeficiency virus type 1 (HIV-1) circulating recombinant form 02_AG (CRF02_AG) has a higher in vitro replicative capacity than its parental subtypes A and G. J Med Virol 2006;78(5):523-34.

55. Njai HF, Gali Y, Vanham G, Clybergh C, Jennes W, Vidal N, *et al*. The predominance of Human Immunodeficiency Virus type 1 (HIV-1) circulating recombinant form 02 (CRF02_AG) in West Central Africa may be related to its replicative fitness. Retrovirology 2006;3:40.

56. Kanki PJ, Hamel DJ, Sankale JL, Hsieh C, Thior I, Barin F, *et al*. Human immunodeficiency virus type 1 subtypes differ in disease progression. J Infect Dis 1999;179(1):68-73.

57. Neilson JR, John GC, Carr JK, Lewis P, Kreiss JK, Jackson S, *et al*. Subtypes of human immunodeficiency virus type 1 and disease stage among women in Nairobi, Kenya. J Virol 1999;73(5):4393-403.

58. John-Stewart GC, Nduati RW, Rousseau CM, Mbori-Ngacha DA, Richardson BA, Rainwater S, *et al*. Subtype C Is associated with increased vaginal shedding of HIV-1. J Infect Dis 2005;192(3):492-6.

59. Poonia B, Wang X, Veazey RS. Distribution of simian immunodeficiency virus target cells in vaginal tissues of normal rhesus macaques: implications for virus transmission. J Reprod Immunol 2006;72(1-2):74-84.

60. Gray RH, Wawer MJ, Brookmeyer R, Sewankambo NK, Serwadda D, Wabwire-Mangen F, *et al*. Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1-discordant couples in Rakai, Uganda. Lancet 2001;357(9263):1149-53.

61. Pope M, Haase AT. Transmission, acute HIV-1 infection and the quest for strategies to prevent infection. Nat Med 2003;9(7):847-52.

62. Shattock RJ, Moore JP. Inhibiting sexual transmission of HIV-1 infection. Nat Rev Microbiol 2003;1(1):25-34.

63. Meng G, Wei X, Wu X, Sellers MT, Decker JM, Moldoveanu Z, *et al*. Primary intestinal epithelial cells selectively transfer R5 HIV-1 to CCR5+ cells. Nat Med 2002;8(2):150-6.

64. 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):367-77.

65. Paxton WA, Martin SR, Tse D, O'Brien TR, Skurnick J, VanDevanter NL, *et al*. Relative resistance to HIV-1 infection of CD4 lymphocytes from persons who remain uninfected despite multiple high-risk sexual exposure. Nat Med 1996;2(4):412-7.

66. Van Herrewege Y, Michiels J, Waeytens A, De Boeck G, Salden E, Heyndrickx L, *et al*. A dual chamber model of female cervical mucosa for the study of HIV transmission and for the evaluation of candidate HIV microbicides. Antiviral Res 2007;74(2):111-24.

67. McGowan I. Microbicides: a new frontier in HIV prevention. Biologicals 2006;34(4):241-55.

68. Haase AT. Perils at mucosal front lines for HIV and SIV and their hosts. Nat Rev Immunol 2005;5(10):783-92.

69. Morrow G, Vachot L, Vagenas P, Robbiani M. Current concepts of HIV transmission. Curr HIV/AIDS Rep 2007;4(1):29-35.

70. Hladik F, Sakchalathorn P, Ballweber L, Lentz G, Fialkow M, Eschenbach D, *et al*. Initial events in establishing vaginal entry and infection by human immunodeficiency virus type-1. Immunity 2007;26(2):257-70.

71. Boggiano C, Littman DR. HIV's vagina travelogue. Immunity 2007;26(2):145-7.

72. Gupta P, Collins KB, Ratner D, Watkins S, Naus GJ, Landers DV, *et al*. Memory CD4(+) T cells are the earliest detectable human immunodeficiency virus type 1 (HIV-1)-infected cells in the female genital mucosal tissue during HIV-1 transmission in an organ culture system. J Virol 2002;76(19):9868-76.

73. Miller CJ, Li Q, Abel K, Kim EY, Ma ZM, Wietgrefe S, *et al*. Propagation and dissemination of infection after vaginal transmission of simian immunodeficiency virus. J Virol 2005;79(14):9217-27.

74. Soto-Ramirez LE, Renjifo B, McLane MF, Marlink R, O'Hara C, Sutthent R, *et al*. HIV-1 Langerhans' cell tropism associated with heterosexual transmission of HIV. Science 1996;271(5253):1291-3.

75. van Harmelen J, Wood R, Lambrick M, Rybicki EP, Williamson AL, Williamson C. An association between HIV-1 subtypes and mode of transmission in Cape Town, South Africa. Aids 1997;11(1):81-7.

76. Mastro TD, Kunanusont C, Dondero TJ, Wasi C. Why do HIV-1 subtypes segregate among persons with different risk behaviors in South Africa and Thailand? Aids 1997;11(1):113-6.

77. Arien KK, Abraha A, Quinones-Mateu ME, Kestens L, Vanham G, Arts EJ. The replicative fitness of primary human immunodeficiency virus type 1 (HIV-1) group M, HIV-1 group O, and HIV-2 isolates. J Virol 2005;79(14):8979-90.

78. Braathen LR, Bjercke S, Thorsby E. The antigen-presenting function of human Langerhans cells. Immunobiology 1984;168(3-5):301-12.

79. Kawamura T, Cohen SS, Borris DL, Aquilino EA, Glushakova S, Margolis LB, *et al*. Candidate microbicides block HIV-1 infection of human immature Langerhans cells within epithelial tissue explants. J Exp Med 2000;192(10):1491-500.

80. Soto-Ramirez LE, Tripathy S, Renjifo B, Essex M. HIV-1 pol sequences from India fit distinct subtype pattern. J Acquir Immune Defic Syndr Hum Retrovirol 1996;13(4):299-307.

81. Hu J, Gardner MB, Miller CJ. Simian immunodeficiency virus rapidly penetrates the cervicovaginal mucosa after intravaginal inoculation and infects intraepithelial dendritic cells. J Virol 2000;74(13):6087-95.

82. Spira AI, Ho DD. Effect of different donor cells on human immunodeficiency virus type 1 replication and selection in vitro. J Virol 1995;69(1):422-9.

83. Veazey RS, Marx PA, Lackner AA. Vaginal CD4+ T cells express high levels of CCR5 and are rapidly depleted in simian immunodeficiency virus infection. J Infect Dis 2003;187(5):769-76.

84. Piguet V, Blauvelt A. Essential roles for dendritic cells in the pathogenesis and potential treatment of HIV disease. J Invest Dermatol 2002;119(2):365-9.

85. Bobardt MD, Chatterji U, Selvarajah S, Van der Schueren B, David G, Kahn B, *et al*. Cell-free human immunodeficiency virus type 1 transcytosis through primary genital epithelial cells. J Virol 2007;81(1):395-405.

86. Howell AL, Edkins RD, Rier SE, Yeaman GR, Stern JE, Fanger MW, *et al*. Human immunodeficiency virus type 1 infection of cells and tissues from the upper and lower human female reproductive tract. J Virol 1997;71(5):3498-506.

87. Blauvelt A, Glushakova S, Margolis LB. HIV-infected human Langerhans cells transmit infection to human lymphoid tissue ex vivo. Aids 2000;14(6):647-51.

88. Collins KB, Patterson BK, Naus GJ, Landers DV, Gupta P. Development of an in vitro organ culture model to study transmission of HIV-1 in the female genital tract. Nat Med 2000;6(4):475-9.

89. Maher D, Wu X, Schacker T, Horbul J, Southern P. HIV binding, penetration, and primary infection in human cervicovaginal tissue. Proc Natl Acad Sci U S A 2005;102(32):11504-9.

90. Narimatsu R, Wolday D, Patterson BK. IL-8 increases transmission of HIV type 1 in cervical explant tissue. AIDS Res Hum Retroviruses 2005;21(3):228-33.

91. Moore JP, Trkola A, Korber B, Boots LJ, Kessler JA, 2nd, McCutchan FE, *et al*. A human monoclonal antibody to a complex epitope in the V3 region of gp120 of human immunodeficiency virus type 1 has broad reactivity within and outside clade B. J Virol 1995;69(1):122-30.

92. Collins KB. A novel in vitro organ culture model for the investigation of HIV-1 transmission in the female genital tract [Thesis M S --University of Pittsburgh 1999]; 1999.

93. Delwart EL, Herring B, Rodrigo AG, Mullins JI. Genetic subtyping of human immunodeficiency virus using a heteroduplex mobility assay. PCR Methods Appl 1995;4(5):S202-16.

94. Ramirez de Arellano E, Soriano V, Alcamil J, Holguin A. New findings on transcription regulation across different HIV-1 subtypes. AIDS Rev 2006;8(1):9-16.

95. De Arellano ER, Soriano V, Holguin A. Genetic analysis of regulatory, promoter, and TAR regions of LTR sequences belonging to HIV type 1 Non-B subtypes. AIDS Res Hum Retroviruses 2005;21(11):949-54.

96. Naghavi MH, Schwartz S, Sonnerborg A, Vahlne A. Long terminal repeat promoter/enhancer activity of different subtypes of HIV type 1. AIDS Res Hum Retroviruses 1999;15(14):1293-303.

97. De Baar MP, De Ronde A, Berkhout B, Cornelissen M, Van Der Horn KH, Van Der Schoot AM, *et al*. Subtype-specific sequence variation of the HIV type 1 long terminal repeat and primer-binding site. AIDS Res Hum Retroviruses 2000;16(5):499-504.

98. van Opijnen T, Jeeninga RE, Boerlijst MC, Pollakis GP, Zetterberg V, Salminen M, *et al*. Human immunodeficiency virus type 1 subtypes have a distinct long terminal repeat that determines the replication rate in a host-cell-specific manner. J Virol 2004;78(7):3675-83.

99. Montano MA, Nixon CP, Essex M. Dysregulation through the NF-kappaB enhancer and TATA box of the human immunodeficiency virus type 1 subtype E promoter. J Virol 1998;72(10):8446-52.

100. Centlivre M, Sommer P, Michel M, Ho Tsong Fang R, Gofflo S, Valladeau J, *et al*. HIV-1 clade promoters strongly influence spatial and temporal dynamics of viral replication in vivo. J Clin Invest 2005;115(2):348-58.

101. Hunt GM, Johnson D, Tiemesse CT. Characterisation of the long terminal repeat regions of South African human immunodeficiency virus type 1 isolates. Virus Genes 2001;23(1):27-34.

102. Choudhury S, Montano MA, Womack C, Blackard JT, Maniar JK, Saple DG, *et al*. Increased promoter diversity reveals a complex phylogeny of human immunodeficiency virus type 1 subtype C in India. J Hum Virol 2000;3(1):35-43.

103. Munkanta M, Handema R, Kasai H, Gondwe C, Deng X, Yamashita A, *et al*. Predominance of three NF-kappaB binding sites in the long terminal repeat region of HIV Type 1 subtype C isolates from Zambia. AIDS Res Hum Retroviruses 2005;21(10):901-6.

104. Jeeninga RE, Hoogenkamp M, Armand-Ugon M, de Baar M, Verhoef K, Berkhout B. Functional differences between the long terminal repeat transcriptional promoters of human immunodeficiency virus type 1 subtypes A through G. J Virol 2000;74(8):3740-51.

105. Montano MA, Novitsky VA, Blackard JT, Cho NL, Katzenstein DA, Essex M. Divergent transcriptional regulation among expanding human immunodeficiency virus type 1 subtypes. J Virol 1997;71(11):8657-65.

106. Rangel HR, Weber J, Chakraborty B, Gutierrez A, Marotta ML, Mirza M, *et al*. Role of the human immunodeficiency virus type 1 envelope gene in viral fitness. J Virol 2003;77(16):9069-73.

107. Marozsan AJ, Moore DM, Lobritz MA, Fraundorf E, Abraha A, Reeves JD, *et al*. Differences in the fitness of two diverse wild-type human immunodeficiency virus type 1 isolates are related to the efficiency of cell binding and entry. J Virol 2005;79(11):7121-34.

108. John TJ, Babu PG, Jayakumari H, Simoes EA. Prevalence of HIV infection in risk groups in Tamilnadu, India. Lancet 1987;1(8525):160-1.

109. Simoes EA, Babu PG, John TJ, Nirmala S, Solomon S, Lakshminarayana CS, *et al*. Evidence for HTLV-III infection in prostitutes in Tamil Nadu (India). Indian J Med Res 1987;85:335-8.

110. Cecilia D, Kulkarni SS, Tripathy SP, Gangakhedkar RR, Paranjape RS, Gadkari DA. Absence of coreceptor switch with disease progression in human immunodeficiency virus infections in India. Virology 2000;271(2):253-8.

111. Balachandran R, Thampatty P, Enrico A, Rinaldo C, Gupta P. Human immunodeficiency virus isolates from asymptomatic homosexual men and from AIDS patients have distinct biologic and genetic properties. Virology 1991;180(1):229-38.

112. Balachandran R, Thampatty P, Rinaldo C, Gupta P. Use of cryopreserved normal peripheral blood lymphocytes for isolation of human immunodeficiency virus from seropositive men. J Clin Microbiol 1988;26(3):595-7.

113. Bjorndal A, Deng H, Jansson M, Fiore JR, Colognesi C, Karlsson A, *et al*. Coreceptor usage of primary human immunodeficiency virus type 1 isolates varies according to biological phenotype. J Virol 1997;71(10):7478-87.

114. O'Marro SD, Armstrong JA, Asuncion C, Gueverra L, Ho M. The effect of combinations of ampligen and zidovudine or dideoxyinosine against human immunodeficiency viruses in vitro. Antiviral Res 1992;17(2):169-77.

115. Bachmann MH, Delwart EL, Shpaer EG, Lingenfelter P, Singal R, Mullins JI. Rapid genetic characterization of HIV type 1 strains from four World Health Organization-sponsored vaccine evaluation sites using a heteroduplex mobility assay. WHO Network for HIV Isolation and Characterization. AIDS Res Hum Retroviruses 1994;10(11):1345-53.

116. Kalia V, Sarkar S, Gupta P, Montelaro RC. Rational site-directed mutations of the LLP-1 and LLP-2 lentivirus lytic peptide domains in the intracytoplasmic tail of human immunodeficiency virus type 1 gp41 indicate common functions in cell-cell fusion but distinct roles in virion envelope incorporation. J Virol 2003;77(6):3634-46.

117. Mochizuki N, Otsuka N, Matsuo K, Shiino T, Kojima A, Kurata T, *et al*. An infectious DNA clone of HIV type 1 subtype C. AIDS Res Hum Retroviruses 1999;15(14):1321-4.

118. Chen M, Shi C, Kalia V, Tencza SB, Montelaro RC, Gupta P. HIV gp120 V(1)/V(2) and C(2)-V(3) domains glycoprotein compatibility is required for viral replication. Virus Res 2001;79(1-2):91-101.

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

120. Kuiken CL, Foley B, Hahn B, Korber B, Marx PA, McCutchan F, Mellors JW, and Wolinsky S, Eds. HIV Sequence Compendium 2001. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, LA-UR 02-2877.

121. Yang Z, Engel JD. Human T cell transcription factor GATA-3 stimulates HIV-1 expression. Nucleic Acids Res 1993;21(12):2831-6.

122. Liu YZ, Latchman DS. The octamer-binding proteins Oct-1 and Oct-2 repress the HIV long terminal repeat promoter and its transactivation by Tat. Biochem J 1997;322 (Pt 1):155-8.

123. Zhang M, Genin A, Cron RQ. Overexpression of octamer transcription factors 1 or 2 alone has no effect on HIV-1 transcription in primary human CD4 T cells. Virology 2004;321(2):323-31.

124. Rodriguez MA, Chen Y, Craigo JK, Chatterjee R, Ratner D, Tatsumi M, *et al*. Construction and characterization of an infectious molecular clone of HIV-1 subtype A of Indian origin. Virology 2006;345(2):328-36.

125. Ramirez de Arellano E, Martin C, Soriano V, Alcami J, Holguin A. Genetic analysis of the long terminal repeat (LTR) promoter region in HIV-1-infected individuals with different rates of disease progression. Virus Genes 2007;34(2):111-6.

126. Pereira LA, Bentley K, Peeters A, Churchill MJ, Deacon NJ. A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. Nucleic Acids Res 2000;28(3):663-8.

127. Bates M, Wrin T, Huang W, Petropoulos C, Hellmann N. Practical applications of viral fitness in clinical practice. Curr Opin Infect Dis 2003;16(1):11-8.

128. Weber J, Rangel HR, Chakraborty B, Tadele M, Martinez MA, Martinez-Picado J, *et al*. A novel TaqMan real-time PCR assay to estimate ex vivo human immunodeficiency virus type 1 fitness in the era of multi-target (pol and env) antiretroviral therapy. J Gen Virol 2003;84(Pt 8):2217-28.

129. Baskar PV, Ray SC, Rao R, Quinn TC, Hildreth JE, Bollinger RC. Presence in India of HIV type 1 similar to North American strains. AIDS Res Hum Retroviruses 1994;10(8):1039-41.

130. Iversen AK, Learn GH, Skinhoj P, Mullins JI, McMichael AJ, Rambaut A. Preferential detection of HIV subtype C' over subtype A in cervical cells from a dually infected woman. Aids 2005;19(9):990-3.

131. Montano MA, Nixon CP, Ndung'u T, Bussmann H, Novitsky VA, Dickman D, *et al*. Elevated tumor necrosis factor-alpha activation of human immunodeficiency virus type 1 subtype C in Southern Africa is associated with an NF-kappaB enhancer gain-of-function. J Infect Dis 2000;181(1):76-81.

132. Roof P, Ricci M, Genin P, Montano MA, Essex M, Wainberg MA, *et al*. Differential regulation of HIV-1 clade-specific B, C, and E long terminal repeats by NF-kappaB and the Tat transactivator. Virology 2002;296(1):77-83.

133. Najera R, Delgado E, Perez-Alvarez L, Thomson MM. Genetic recombination and its role in the development of the HIV-1 pandemic. Aids 2002;16 Suppl 4:S3-16.

134. Binley JM, Wrin T, Korber B, Zwick MB, Wang M, Chappey C, *et al*. Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. J Virol 2004;78(23):13232-52.

135. Bures R, Morris L, Williamson C, Ramjee G, Deers M, Fiscus SA, *et al*. Regional clustering of shared neutralization determinants on primary isolates of clade C human immunodeficiency virus type 1 from South Africa. J Virol 2002;76(5):2233-44.

136. Trkola A, Pomales AB, Yuan H, Korber B, Maddon PJ, Allaway GP, *et al*. Cross-clade neutralization of primary isolates of human immunodeficiency virus type 1 by human monoclonal antibodies and tetrameric CD4-IgG. J Virol 1995;69(11):6609-17.

137. Coplan PM, Gupta SB, Dubey SA, Pitisuttithum P, Nikas A, Mbewe B, *et al*. Cross-reactivity of anti-HIV-1 T cell immune responses among the major HIV-1 clades in HIV-1-positive individuals from 4 continents. J Infect Dis 2005;191(9):1427-34.

138. Zhao S, Zhai S, Zhuang Y, Wang S, Huang D, Kang W, *et al*. Inter-clade cross-reactivity of HIV-1-specific T cell responses in human immunodeficiency virus type 1 infection in China. Curr HIV Res 2007;5(2):251-9.

139. McKinnon LR, Ball TB, Kimani J, Wachihi C, Matu L, Luo M, *et al*. Cross-clade CD8(+) T-cell responses with a preference for the predominant circulating clade. J Acquir Immune Defic Syndr 2005;40(3):245-9.

140. Dorrell L, Dong T, Ogg GS, Lister S, McAdam S, Rostron T, *et al*. Distinct recognition of non-clade B human immunodeficiency virus type 1 epitopes by cytotoxic T lymphocytes generated from donors infected in Africa. J Virol 1999;73(2):1708-14.

141. Wainberg MA. HIV-1 subtype distribution and the problem of drug resistance. Aids 2004;18 Suppl 3:S63-8.

142. Brenner B, Turner D, Oliveira M, Moisi D, Detorio M, Carobene M, *et al*. A V106M mutation in HIV-1 clade C viruses exposed to efavirenz confers cross-resistance to non-nucleoside reverse transcriptase inhibitors. Aids 2003;17(1):F1-5.

143. Loemba H, Brenner B, Parniak MA, Ma'ayan S, Spira B, Moisi D, *et al*. Polymorphisms of cytotoxic T-lymphocyte (CTL) and T-helper epitopes within reverse transcriptase (RT) of HIV-1 subtype C from Ethiopia and Botswana following selection of antiretroviral drug resistance. Antiviral Res 2002;56(2):129-42.

144. Loemba H, Brenner B, Parniak MA, Ma'ayan S, Spira B, Moisi D, *et al*. Genetic divergence of human immunodeficiency virus type 1 Ethiopian clade C reverse transcriptase (RT) and rapid development of resistance against nonnucleoside inhibitors of RT. Antimicrob Agents Chemother 2002;46(7):2087-94.

145. Palmer S, Alaeus A, Albert J, Cox S. Drug susceptibility of subtypes A,B,C,D, and E human immunodeficiency virus type 1 primary isolates. AIDS Res Hum Retroviruses 1998;14(2):157-62.

146. Garg S, Anderson RA, Chany CJ, 2nd, Waller DP, Diao XH, Vermani K, *et al*. Properties of a new acid-buffering bioadhesive vaginal formulation (ACIDFORM). Contraception 2001;64(1):67-75.

147. Zeitlin L, Hoen TE, Achilles SL, Hegarty TA, Jerse AE, Kreider JW, *et al*. Tests of Buffergel for contraception and prevention of sexually transmitted diseases in animal models. Sex Transm Dis 2001;28(7):417-23.

148. Antonio MA, Hillier SL. DNA fingerprinting of Lactobacillus crispatus strain CTV-05 by repetitive element sequence-based PCR analysis in a pilot study of vaginal colonization. J Clin Microbiol 2003;41(5):1881-7.

149. Lagenaur LA, Berger EA. An anti-HIV microbicide comes alive. Proc Natl Acad Sci U S A 2005;102(35):12294-5.

150. Martin HL, Richardson BA, Nyange PM, Lavreys L, Hillier SL, Chohan B, *et al*. Vaginal lactobacilli, microbial flora, and risk of human immunodeficiency virus type 1 and sexually transmitted disease acquisition. J Infect Dis 1999;180(6):1863-8.

151. Hillier SL, Moench T, Shattock R, Black R, Reichelderfer P, Veronese F. In vitro and in vivo: the story of nonoxynol 9. J Acquir Immune Defic Syndr 2005;39(1):1-8.

152. Van Damme L, Ramjee G, Alary M, Vuylsteke B, Chandeying V, Rees H, *et al*. Effectiveness of COL-1492, a nonoxynol-9 vaginal gel, on HIV-1 transmission in female sex workers: a randomised controlled trial. Lancet 2002;360(9338):971-7.

153. Krebs FC, Miller SR, Catalone BJ, Welsh PA, Malamud D, Howett MK, *et al*. Sodium dodecyl sulfate and C31G as microbicidal alternatives to nonoxynol 9: comparative sensitivity of primary human vaginal keratinocytes. Antimicrob Agents Chemother 2000;44(7):1954-60.

154. Patton DL, Sweeney YT, Balkus JE, Hillier SL. Vaginal and rectal topical microbicide development: safety and efficacy of 1.0% Savvy (C31G) in the pigtailed macaque. Sex Transm Dis 2006;33(11):691-5.

155. Haineault C, Gourde P, Perron S, Desormeaux A, Piret J, Omar RF, *et al*. Thermoreversible gel formulation containing sodium lauryl sulfate as a potential contraceptive device. Biol Reprod 2003;69(2):687-94.

156. Piret J, Desormeaux A, Bergeron MG. Sodium lauryl sulfate, a microbicide effective against enveloped and nonenveloped viruses. Curr Drug Targets 2002;3(1):17-30.

157. Lederman MM, Veazey RS, Offord R, Mosier DE, Dufour J, Mefford M, *et al*. Prevention of vaginal SHIV transmission in rhesus macaques through inhibition of CCR5. Science 2004;306(5695):485-7.

158. Bakobaki JM, Lacey CJ, Bukenya MI, Nunn AJ, McCormack S, Byaruhanga RN, *et al*. A randomized controlled safety and acceptability trial of dextrin sulphate vaginal microbicide gel in sexually active women in Uganda. Aids 2005;19(18):2149-56.

159. McCormack S, Jespers V, Low-Beer N, Gabe R, Kaganson N, Chapman A, *et al*. A dose-ranging phase I study of dextrin sulphate, a vaginal microbicide, in HIV-negative and HIV-positive female volunteers. Sex Transm Dis 2005;32(12):765-70.

160. Stafford MK, Cain D, Rosenstein I, Fontaine EA, McClure M, Flanagan AM, *et al*. A placebo-controlled, double-blind prospective study in healthy female volunteers of dextrin sulphate gel: a novel potential intravaginal virucide. J Acquir Immune Defic Syndr Hum Retrovirol 1997;14(3):213-8.

161. Scordi-Bello IA, Mosoian A, He C, Chen Y, Cheng Y, Jarvis GA, *et al*. Candidate sulfonated and sulfated topical microbicides: comparison of anti-human immunodeficiency virus activities and mechanisms of action. Antimicrob Agents Chemother 2005;49(9):3607-15.

162. Coggins C, Blanchard K, Alvarez F, Brache V, Weisberg E, Kilmarx PH, *et al*. Preliminary safety and acceptability of a carrageenan gel for possible use as a vaginal microbicide. Sex Transm Infect 2000;76(6):480-3.

163. Neurath AR, Li YY, Mandeville R, Richard L. In vitro activity of a cellulose acetate phthalate topical cream against organisms associated with bacterial vaginosis. J Antimicrob Chemother 2000;45(5):713-4.

164. Torre VS, Marozsan AJ, Albright JL, Collins KR, Hartley O, Offord RE, *et al*. Variable sensitivity of CCR5-tropic human immunodeficiency virus type 1 isolates to inhibition by RANTES analogs. J Virol 2000;74(10):4868-76.

165. Borkow G, Barnard J, Nguyen TM, Belmonte A, Wainberg MA, Parniak MA. Chemical barriers to human immunodeficiency virus type 1 (HIV-1) infection: retrovirucidal activity of UC781, a thiocarboxanilide nonnucleoside inhibitor of HIV-1 reverse transcriptase. J Virol 1997;71(4):3023-30.

166. de Witte L, Nabatov A, Pion M, Fluitsma D, de Jong MA, de Gruijl T, et al. Langerin is a natural barrier to HIV-1 transmission by Langerhans cells. Nat Med 2007;13(3):367-71.

167. Rohan LC, Ratner D, McCullough K, Hiller SL, Gupta P. Measurement of anti-HIV activity of marketed vaginal products and excipients using a PBMC-based in vitro assay. Sex Transm Dis 2004;31(3):143-8.