From DEPARTMENT OF MEDICINE, HUDDINGE
Karolinska Institutet, Stockholm, Sweden

Translational Genomics of HIV-1 Subtype C in India: Molecular Phylogeny and Drug Resistance

Ujjwal Neogi

Karolinska Institutet

Stockholm 2013

Cover page photo: The cover picture depicts the basic viral structure of HIV-1, filled with the three colors of the Indian national flag. Orange or saffron color symbolizes *courage* and *sacrifice*, shown by individuals with HIV infection who have fought so bravely against the odds, participated in medical research, hoping for a better life for future patients. White represents *purity* and *truth*, the lofty goals of medical workers and scientists in their search for a cure for HIV. Green signifies *faith* and *prosperity*, the aspirations of the army of social and political workers who work towards reducing stigma and restoring dignity to all those infected and affected by the global HIV epidemic. Inside the virus is shown the blue-colored radial phylogenetic tree, redrawn from the paper III, representing the confluence of basic, translational, clinical and epidemiological research that is central to our ongoing battle against the global HIV epidemic.

Published by Karolinska Institutet. Printed by Universitetsservice US-AB

ISBN 978-91-7549-172-1

ওঠো ! জাগো ! এবং লক্ষ্যে না
পৌঁছানো পর্যন্ত থামিও না !

স্বামী বিবেকানন্দ

Arise! Awake! And Stop Not
Till the Goal is Reached!

Swami Vivekananda

SUMMARY

This thesis describes the translational genomics of HIV-1subtype C in India from its origin to therapeutic response with the aim to improve our knowledge for better therapeutic and preventive strategies to combat HIV/AIDS. In a systemic approach, we identified the molecular phylogeny of HIV-1 subtypes circulating in India and the time to most recent common ancestors (tMRCA) of predominant HIV-1 subtype C strains. Additionally, this thesis also studied drug resistance mutations in children, adolescents and adults, the role of host factors in evolution of drug resistance, and population dynamics of viremia and viral co-receptor tropism in perinatal transmission. Finally, the long term therapeutic responses on Indian national first-line antiretroviral therapy were also studied.

In **Paper I**, we reported an increase in the HIV-1 recombinant forms in the HIV-1 epidemiology using a robust subtyping methodology. While the study confirmed HIV-1 subtype C as a dominant subtype, its origin was dated back to the early 1970s from a single or few genetically related strains from South Africa, whereafter, it has evolved independently. In **Paper II**, the lethal hypermutations due to the activity of human apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G (hA3G) was significantly associated with antiretroviral therapy (ART) failure in Indian HIV-1 subtype C patients. The presence of M184I and M230I mutations were observed due to the editing of hA3G in the proviral compartment but stop codons were also found in the open reading frames and the same drug resistance mutations were absent in plasma virus. Therefore, it is unlikely that the viral variants which exhibit hypermutated sequences and M184I and/or M230I will mature and expand *in vivo* and hence are unlikely to have any clinical significance. The high concordance of drug resistance genotyping in the plasma and proviral compartments in therapy-naïve patients, gives weight to the idea of using whole blood for surveillance of drug resistance mutations which precludes logistic challenges of cold chain transport.

In **Papers III** and **IV**, we identified a substantial proportion of HIV-1 subtype C perinatally-infected older children who had a high burden of plasma viremia but also had high $CD4^+$ T-cell counts. In addition, older children with HIV-1 subtype C infection presented a high prevalence of predicted X4 and R5/X4 tropic strains which indicates that HIV-1 subtype C strains required longer duration of infection and greater disease progression to co-receptor transition from R5- to X4-tropic strains (**IV**). Our studies also indicate that transmitted drug resistance is low among Indian HIV-1 infected children, adolescents (**III**) and adults (**II**).

In **Paper V,** in a longitudinal cohort study, a good long-term response to the Indian national first-line therapy for a median of nearly four years with 2.8% viral failure, indicating the overall success of the Indian ART program. Our study also showed that three immunologically well patients with virological rebound and major viral drug resistance mutations (M184V, K103N and Y181C) during one study visit had undetectable viral load at their next visit. These findings suggest that use of multiple parameters like patients' immunological ($CD4^+$ T-cell count), virological (viral load) and drug resistance data should all be used to optimize the treatment switch to second line therapy.

In conclusion, this translational genomics study enhances our knowledge about the HIV-1 subtype C strains circulating in India which are genetically distinct from prototype African subtype C strains. Considerably more research using appropriate models need to be performed to understand the phenotypic and biological characteristics of these strains to guide efficient disease intervention and management strategies.

LIST OF PUBLICATIONS INCLUDED IN THE THESIS

This thesis is based on the following papers referred to in the text by their Roman numerals:

I. **Neogi U,** Bontell I, Shet A, De Costa A, Gupta S, Diwan V, Laishram RS, Wanchu A, Ranga U, Banerjea AC, Sönnerborg A. (2012) **Molecular epidemiology of HIV-1 subtypes in India: Origin and evolutionary history of the predominant subtype C.** *PLoS ONE* 7(6): e39819. doi:10.1371/journal.pone.0039819.

II. **Neogi U,** Shet A, Sahoo PN, Bontell I, Ekstrand ML, Banerjea AC, Sönnerborg A. (2013) **Human APOBEC3G-mediated hypermutation is associated with antiretroviral therapy failure in HIV-1 subtype C-infected individuals.** *J Int AIDS Soc*. 16:18472. doi: 10.7448/IAS.16.1.18472.

III. **Neogi U,** Sahoo PN, De Costa A, Shet A. (2012) **High viremia and low level of transmitted drug resistance in anti-retroviral therapy-naïve perinatally-infected children and adolescents with HIV-1 subtype C infection.** *BMC Infec Dis* Nov 22;12(1):317. doi: 10.1186/1471-2334-12-317.

IV. **Neogi U,** Sahoo PN, Arumugam K, Sönnerborg A, De Costa A, Shet A. (2012) **Higher prevalence of predicted X4-tropic strains in perinatally-infected older children with HIV-1 subtype C in India.** *J Acquir Immune Defic Syndr*. 59(4):347-353.

V. **Neogi U,** Heylen E, Shet A, Chandy S, Shamsunder R, Sönnerborg A, Ekstrand ML. (2013) **Long-term efficacy of first line antiretroviral therapy in Indian HIV-1 infected patients: A longitudinal cohort study.** *PLoS ONE* 8(1):e55421. doi: 10.1371/journal.pone.0055421.

RELATED PUBLICATIONS

1. **Neogi U,** Shet A, Shamsundar R, Ekstrand ML. (2011) **Selection of non-nucleoside reverse transcriptase inhibitor-associated mutations in HIV-1 subtype C: evidence of etravirine cross-resistance.** *AIDS* 25:1123–1126.

2. Gupta S, **Neogi U**, Hiresave S, Shet A. (2013) **High concordance of genotypic co-receptor prediction in plasma-viral RNA and pro-viral DNA of HIV-1 subtype C virus: implications in use of whole-blood DNA in resource-limited settings.** *J Antimicrob Chemother*, E-pub 30 April 2013; doi: 10.1093/jac/dkt138

3. Bachu M, Yalla S, Asokan M, Verma A, **Neogi U,** Sharma S, Murali RV, Mukthey AB, Bhatt R, Chatterjee S, Rajan RE, Cheedarla N, Yadavalli VS, Mahadevan A, Shankar SK, Rajagopalan N, Shet A, Saravanan S, Balakrishnan P, Solomon S, Vajpayee M, Satish KS, Kundu TK, Jeang KT, Ranga U. (2012) **Multiple NF-κB sites in HIV-1 subtype C LTR confer superior magnitude of transcription and thereby the enhanced viral predominance.** *J Biol Chem.* 287(53):44714-447135.

4. Shet A, **Neogi U,** Sahoo PN, De Costa A. (2012) **Effectiveness of first-line antiretroviral therapy and acquired drug resistance among HIV-1-infected children in India.** *Pediatr Infect Dis J* 2013;32: e227–e229.

5. **Neogi U,** Sahoo PN, Kumar R, De Costa A, Shet A. (2011) **Characterizations of HIV-1 subtype C *protease* gene: selection of L63P mutation in protease inhibitor-naive Indian patients.** *AIDS Res Hum Retroviruses* 27(11):1249-53.

6. **Neogi U,** Prarthana BS, Gupta S, D'souza G, De Costa A, Kuttiatt VS, Arumugam K, Shet A. (2010) **Naturally occurring polymorphisms and primary drug resistance profile among antiretroviral-naïve individuals in Bangalore, India.** *AIDS Res Hum Retroviruses* 26(10): 1097-1101.

CONTENTS

LIST OF ABBREVIATIONS

3TC	Lamivudine
ABC	Abacavir
AIDS	Acquired immunodeficiency syndrome
ANC	Antenatal clinic
APOBEC	Apolipoprotein B mRNA-editing enzyme catalytic polypeptide
APV	Amprenvir
ART	Antiretroviral therapy
ATV/r	Atazanavir/r
AZT	Zidovudine
cART	Combination antiretroviral therapy
CCR5	C-C chemokine receptor type 5
CD4	Cluster of differentiation antigen 4
COE	Centre of excellence
CRFs	Circulating recombinant forms
CXCR4	C-X-C chemokine receptor type 4
d4T	Stavudine
ddI	Didanosine
DLV	Delavirdine
DNA	Deoxyribonucleic acid
DRM	Drug resistance mutation
DRV/r	Darunavir/r
EFV	Efavirenz
EMA	European medicines agency
Env	Envelope
ETR	Etravirine
FPR	False positive rate
FPV/r	Fosamprenavir/r
FSW	Female sex worker
FTC	Emtricitabine
Gag	Glycoprotein
GoI	Government of India
GTR	General time reversible
HAD	HIV associated dementia
HIV-1	Human immunodeficiency virus type 1
IAS-USA	International AIDS society-United States of America
IDU	Injecting drug users
IDV/r	Indinavir/r
IN	Integrase
INI	Integrase inhibitor
LANL	Los Alamos National Laboratory
LPV/r	Lopinavir/r
LTR	Long terminal repeat

MEGA	Molecular evolutionary genetics analysis
MHC	Major histocompatibility complex
ML	Maximum likelihood
MNP+	Manipur network of positive people
MSM	Men who have sex with men
MTCT	Mother to child transmission
NACO	National AIDS control organization
NACP	National AIDS control program
Nef	Negative regulatory factor
NFV	Nelfinavir
NNRTI	Non-nucleoside reverse transcriptase inhibitors
NRTI	Nucleoside reverse transcriptase inhibitors
NVP	Nevirapine
ORF	Open reading frame
PCR	Polymerase chain reaction
PI	Protease inhibitor
PLHA	People living with HIV/AIDS
Pol	Polymerase
PR	Protease
Rev	Regulator of virion expression
RNA	Ribonucleic acid
RPV	Rilpivirine
RT	Reverse transcriptase
RTI	Reverse transcriptase inhibitor
RTV	Ritonavir
SDRM	Surveillance drug resistance mutation
SIV	Simian immunodeficiency virus
SIVcpz	SIVs of chimpanzee
SIVgor	SIVs of Gorilla
SIVsm	SIVs of sooty mangabeys
SQV/r	Saquinavir/r
TAM	Thymidine analogue mutations
Tat	Trans-activator of transcription
TDF	Tenofovir
TDRM	Transmitted drug resistance mutation
tMRCA	Time to most recent common ancestors
TPV/r	Tipranavir/r
UNAIDS	Joint United Nations programme on HIV/AIDS
URFs	Unique recombinant forms
US-FDA	United States Food and Drug Administrations
VAS	Visual analogues scale
Vif	Virion infectivity factor
Vpr	Viral protein R
Vpu	Viral protein unique
WHO	World health organisation

1 INTRODUCTION

1.1 THE GLOBAL HIV-1 PANDEMIC

The Human Immunodeficiency Virus type 1 (HIV-1)-induced acquired immunodeficiency syndrome (AIDS) pandemic, has emerged as a major medical and public health concern for the past three decades ever since the isolation of the virus in early 1980s [1-4]. As per the 2012 Joint United Nations programme on HIV/AIDS (UNAIDS) report, it is estimated that globally there are 34 million people living with HIV/AIDS (PLHA) at the end of 2011 [5] across almost all the parts of the world (Fig. 1). Sub Saharan Africa continues to harbor the highest burden of PLHA, which accounts for nearly 67% of the global infections [6] followed by India with 2.4 million PLHA [7]. However, wide access to antiretroviral therapy (ART) has resulted in an encouraging decline in the incidence of the number of new annual HIV infections and AIDS-related deaths worldwide.

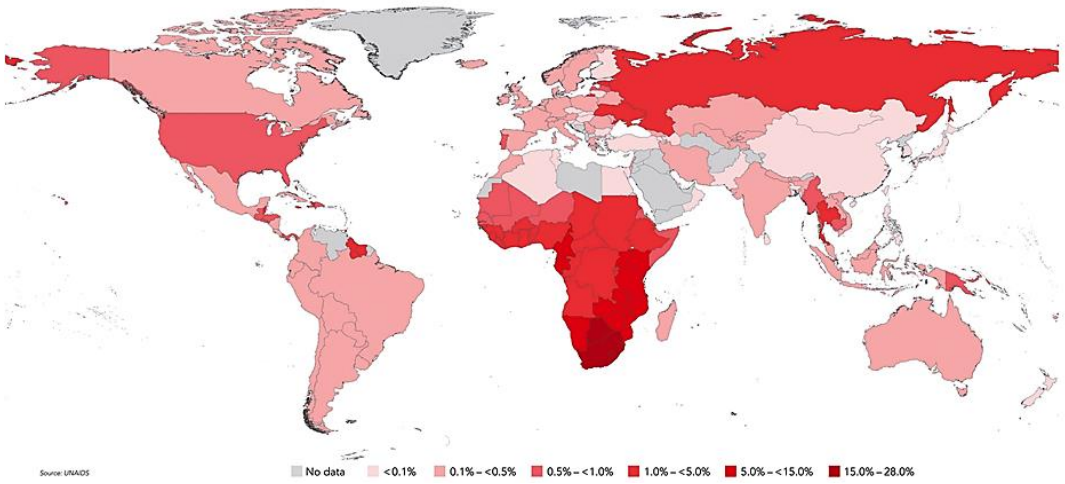

Figure 1. Global prevalence of HIV, 2009 [Source: UNAIDS]

1.1.1 HIV/AIDS EPIDEMIC IN INDIA

India is the second most populated country in the world comprising of more than one billion people that accounts for 17% of the global population (WHO Country Profile) distributed across 28 states and 7 union territories with a total land area of 1,269,219 sq miles. The first case of HIV in India was reported in 1986 among commercial sex workers in Chennai, Tamil Nadu, southern part of India [8]. Since then, it has spread to almost all parts of the country with an estimate of 2.4 million infections

as of 2009 and a prevalence of 0.31% of the general population [7]. Despite the low prevalence, demographically, India has become the third largest country with people living with HIV/AIDS after South Africa and Nigeria [7].

The epidemic is mainly restricted to high risk groups such as Female Sex Workers (FSW), Injecting Drug Users (IDU) and Men who have Sex with Men (MSM) including transgenders. The predominant mode of transmission is heterosexual (87.4%), followed by mother to child transmissions (5.4%). However, in certain regions of the country, the prevalence of HIV-infections in the IDUs and MSMs are noteworthy. Due to this heterogeneous spread of the epidemic, some geographical pockets have shown high incidence of HIV/AIDS (Fig. 2). As per 2009 estimates, the highest estimated adult HIV prevalence was in Manipur (1.40%), followed by Andhra Pradesh (0.90%), Mizoram (0.81%), Nagaland (0.78%), Karnataka (0.63%) and Maharashtra (0.55%).

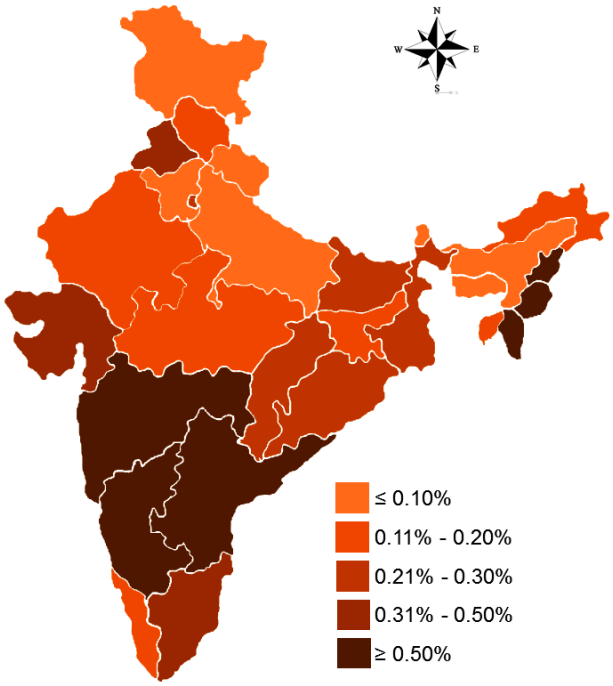

Figure 2. State wise adult HIV-1 prevalence in India (2009 estimates). Data adapted from National AIDS Control Programme Phase III, State Fact Sheets, March 2012.

Following the identification of HIV/AIDS cases, the Ministry of Health and Family Welfare, Government of India (GoI), initiated prevention and awareness programmes on HIV/AIDS through the Medium Term Plan (1990-92). This has been further extended to the National AIDS Control Programme (NACP) by three strategic phases [NACP-I (1992-99), NACP-II (1999-2006) and NACP-III (2006-2011)]. Targeted interventions among the high risk groups, executed with the assistance of various nodal non-governmental organisations and community-based organisations have shown promising results with a decline in prevalence of PHLA from 0.39% in 2004 to 0.31% in 2009. Owing to the encouraging results from the previous NACP program, the GoI has initiated the NACP-IV program which is currently in the planning phase.

1.2 THE ORIGIN AND CLASSIFICATION OF HIV

Current scientific evidence indicates that HIV entered the human population through multiple zoonotic transmissions from the Simian Immunodeficiency Virus (SIV) that infects non-human primates [9, 10]. Molecular phylogenetic studies have revealed that the likely origin of HIV-1 was an SIV isolate from the chimpanzee (SIVcpz) subspecies *Pan troglodytes troglodytes* that inhabits in the eastern equatorial forests of West Central Africa [11, 12]. HIV-2, on the other hand, has evolved from numerous cross-species transmissions of SIVs of sooty mangabeys (SIVsm) (*Cercocebus atys*) that inhabit in the western parts of Africa (Fig. 3) [13, 14].

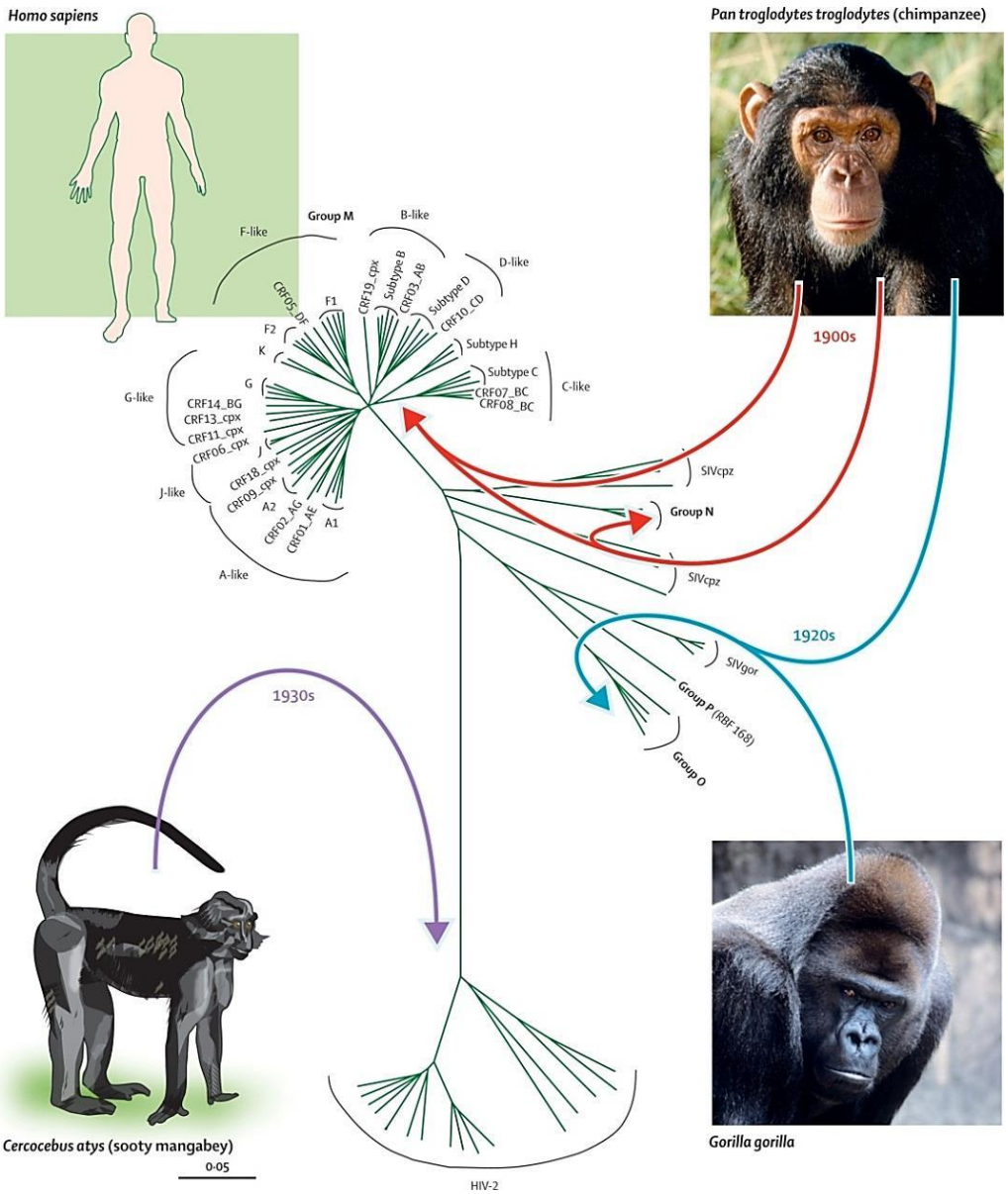

Figure 3. Cross-species transmission of HIV and distribution of subtypes and recombinant forms. Reprinted with permission from Tebit et al 2011 [15].

High genetic variability, both at the population (inter-host) as well as at the individual (intra-host) level is the hallmark of HIV that stems mainly from the inherent infidelity of the reverse transcriptase (RT) enzyme. The lack of proofreading mechanisms, coupled with an *in vivo* virus production rate exceeding 10^9 per day and the persistent nature of infection, provides tremendous scope for the generation of viral diversity [16]. In addition, the propensity for recombination also contributes to the extensive variability of HIV. Intra-subtype genetic diversity differs between 8 and 17%, while the inter-subtype genetic diversity usually differs between 17 and 35%, depending on the genomic region of HIV-1 or the subtype under study [17].

Due to this explosive genetic variability, HIV-1 genome has given rise to three distantly related groups: M (main), O (outlier) and N (non-M non-O) with unique genetic architectures [18, 19]. In 2011, a new group was identified in Cameroon, which is termed as putative or P. Interestingly, this forms a distinct cluster that includes the SIV sequences from western gorillas (SIVgor; *Gorilla gorilla gorilla*) [20, 21]. Globally, HIV-1 group M dominates the AIDS pandemic and is divided into nine pure subtypes [A to D, F to H, J, and K] and several recombinant forms [18]. The recombinants between the subtypes are designated as circulating recombinant forms (CRFs) if the strains are fully sequenced and found in three or more epidemiologically distinct individuals and unique recombinant forms (URFs) if the above criteria are not fulfilled. As of February 2013, a total of 58 CRFs have been fully sequenced and numbered sequentially from CRF01 to CRF58 (Los Alamos National Laboratory, www.hiv.lanl.gov).

Using the molecular clock dating technique, the time to the most recent common ancestor (tMRCA) of the pandemic HIV-1 group M was initially dated back to 1931 (1915-1941) [22]. After the identification of the oldest HIV-1 viruses ZR59 [23] and DRC60 [24] from the Democratic Republic of Congo and inclusion of these strains in the analysis, the tMRCA (using Bayesian skyline plot tree prior) was pushed back to 1908 (1884-1924) [24]. The tMRCA of O and N group was estimated at 1920 (1890–1940) and 1963 (1948–1977) respectively [25, 26]. The tMRCAs of HIV-2 groups A and B were dated back to 1932 (1906-1955) and 1935 (1907- 1961) respectively (Fig. 3) [26].

A recent study on the global distribution of HIV-1 subtypes over a period between 2000 and 2007 indicates that subtype C accounts for nearly half of all global infections

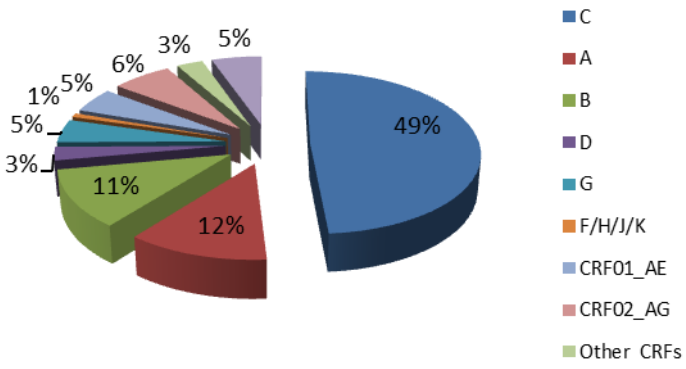

Figure 4. Distribution of major subtypes between 2000 and 2007. Data adapted from Hemelaar et al 2011 [27]

[27]. The second most common subtype is A, which infects nearly 12% of individuals followed by subtype B with 11% of total infections (Fig. 4). Among the CRFs, CRF01_AE and CRF02_AG, accounts for 5% and 6% respectively. Overall, the global distribution of HIV is broadly stable over this period [27, 28]. However, the global proportion of all CRFs combined increased by 4.5% [27].

Significant disparity has also been observed in terms of spread of HIV-1 subtype globally. The first HIV-1 subtype identified was subtype B, which is predominant in many resource-rich settings including North America, western and central Europe and Australia, along with other regions like the Caribbean and parts of Latin America, while subtype C is predominant in more resource-limited settings like Ethiopia, South Africa and India [29] (Fig. 5). Central and western Africa harbors a large diversity of subtypes and recombinant forms with CRF02_AG, subtype F, G, H, J, K along with other recombinants form. In Asia, specifically China, Myanmar, Thailand, the HIV-1 epidemic has been driven by recombinants forms (07_BC, 08_BC, 01_AE etc.).

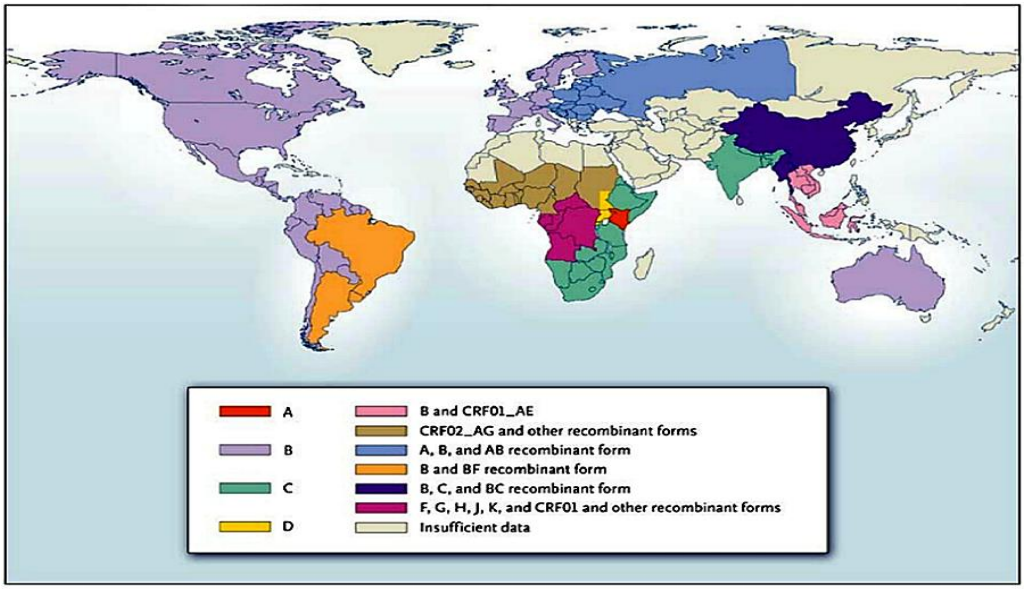

Figure 5. Distribution of HIV-1 subtypes and recombinant forms (2008). Reproduced with permission from Taylor et al 2008 [29], Copyright Massachusetts Medical Society.

1.3 HIV-1 SUBTYPE C

HIV-1 subtype C was first discovered in Ethiopia by Sönnerborg and group in 1988 [30, 31]. It shares at present 50% of the global infections. [27]. Subtype C demonstrates several interesting genotypic and phenotypic properties. Studies have shown the existence a third NF-κB site in HIV-1 subtype C long terminal repeat (LTR) whereas most non-C strains including the commonly studied subtype B viral strains have merely two NF-κB sites [32]. A more recent study from India, reported the presence of an additional 4th NF-κB in LTR, and four NF-κB strains are expanding and replacing the three NF-κB sites containing subtype C viruses [33]. Individuals who are infected with viruses harboring four NF-κB sites had high viremia compared to those individuals with virus containing only three NF-κB sites, although there was no significant difference in their CD4$^+$ T-cell counts exists [33], which gave the conceptual premise that additional NF-κB in the HIV-1 subtype C LTR might enhance the viral replication competence by enhancing the infectivity (Fig. 6).

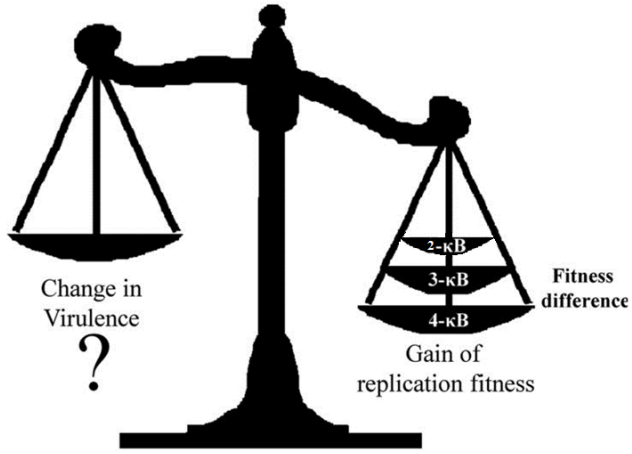

Figure 6. Schematic model portraying a novel strategy the NF-κB mediated replication competence. Reproduced with permission and modification from Bachu et al 2012 [33]

Another important difference is in neuropathogenetic ability. HIV-associated dementia (HAD) is common among untreated HIV-1 subtype B-infected individuals, but less common in subtype C infections [34]. It has been postulated that genetic differences in viral protein Tat at position 31C residue (termed as C30C31 motif) can play an important role in deduced neuropathogenesis by HIV-1. The disruption of the C30C31 motif due to the polymorphism, C31S, results in deficient monocyte chemotaxis than its subtype B counterpart [35]. A more recent study from HIV-1 subtype C from Botswana, also showed a significant linear relationship between Tat mediated LTR transactivation and patients clinical factors like plasma viral load and CD4$^+$ T-cell counts [36]. Apart from this, several subtype C specific genetic signature residues were reported clinically important reverse transcriptase (RT) and protease (PR) region of Pol [37-39]. Significant co-receptor related disparities also have been observed in HIV-1 subtype C. In the later stage of disease nearly 50% of the subtype

B strains are X4-tropic [40], while in subtype C, majority the strains use exclusively the CCR5 co-receptor [41]. Details have been described in the later section of the thesis.

Significant disparity also has been observed in the clinical course of HIV-1 infection and development of drug resistance with subtype C viruses. Studies have been shown that a substantial proportion of HIV-infected adults maintained a high viral set point after acute subtype C infection and might be responsible for rapid spread of this viruses [42]. However, a more recent study from Africa showed that subtype C and non-C subtypes do not differ significantly in terms of their viral set point post sero-conversion [43]. Thus the spread of HIV-1 subtype C strains is still inconclusive. However, there are likely that multiple factors including viral and host immuno-genetics as well as clinical and geographical disease management strategies can play a crucial role in the rapid spread and prevalence of HIV-1 subtype C strains globally [44].

Another important factor is evolution of drug resistance. HIV-1 subtype C showed a distinct drug resistant evolution pathway to accumulate different drug resistance associated mutations, which has been described in the later section of the thesis.

1.4 HIV-1 VIROLOGY: STRUCTURE AND GENOME

HIV is a lentivirus of the *Retroviridae* family and is transmitted as a single-stranded, positive-sense, enveloped RNA virus with 9.5 kb genome (Fig. 7).

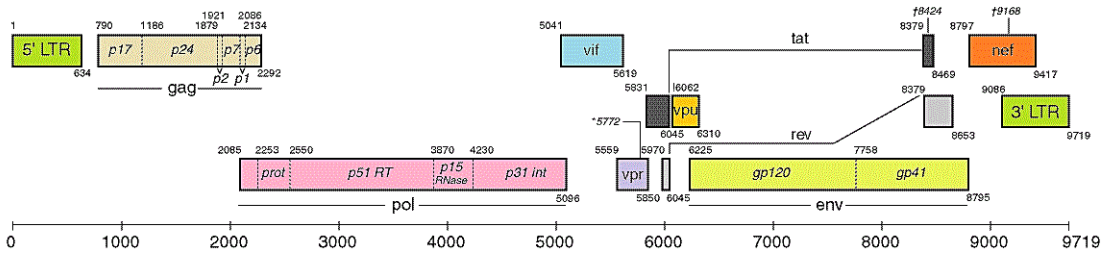

Figure 7. Model for mature HIV-1 virus. **B.** Genetic organization of HIV-1 structural and functional genes. The numbers are designated as HXB2 co-ordinate [Source: HIV Los Alamos Database, www.hiv.lanl.gov]

The RNA genome encodes three structural polyproteins (Gag, Pol and Env) and six regulatory or accessory proteins (Tat, Rev, Vif, Vpr, Vpu and Nef) flanked by non-coding long terminal repeat (LTR) at both 5' and 3' end [45-48]. The *gag* gene encodes the structural proteins matrix (MA), capsid (CA), and nucleocapsid (NC). The *pol* gene encodes the viral enzymes reverse transcriptase (RT), protease (PR), and integrase (IN). The *env* gene encodes the outer glycoprotein that mediates viral

entry of permissive cell types [49]. The regulatory proteins, trans-activator of transcription, (Tat) and regulator of expression of virion proteins, (Rev) play important roles in activation of viral transcription elongation and induction of nuclear export of intron-containing viral RNA respectively. Other accessory protein, virion infectivity factor (Vif) suppresses the human host factor, apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like 3G or F (APOBEC3G/3F) that inhibits viral infection, while viral protein R (vpr) is an enhancer of post entry infectivity and G2 cell cycle arrest. Viral protein U (Vpu) and Negative factor (Nef) play critical roles in CD4/MHC down-regulation. Nef also plays an important role in pathogenic determinations [49].

1.5 HIV-1 PATHOGENESIS

The pathogenesis of HIV-1 is complex and multifactorial involving the interplay between multiple viral and host factors [50-52]. The direct interaction between the viral envelope and its cellular receptor, CD4 along with either C-C chemokine receptor type 5 (CCR5) or C-X-C chemokine receptor type 4 (CXCR4), results in a scenario where the virus infects key cells of the adaptive immune response, and hijacks the host immune system. Following infection, a variety of intracellular mechanisms involving the host immunological factors and viral regulatory and accessory proteins are important for the clinical course of disease progression [53]. A significant disparity is observed in the disease course of HIV-infected individuals. While, those who succumb to AIDS relatively soon after infection are termed as rapid progressors and there are others, termed as long-term non-progressors who manage to evade clinical progression without therapy even after 20-25 years [54-57]. Even more interesting, a group of patients named elite controllers can control the viremia below 50 copies/ml viral load without any signs of immunodeficiency [58, 59] .

The infection and clinical progression of HIV-1 is generally divided into three phases: 1) Primary or acute infection, 2) Latency and 3) AIDS. The first phase of HIV-1 infection is termed as primary infections and occurs a few weeks after infection (Fig. 8) [60]. In this period, an explosive viral replication leading to high viremia has been observed which is termed as “acute phase viremia” and during this period the patient is highly infectious. The symptoms are generally non-specific and hence frequently not recognized as signs of HIV infection. Following this period, due to the strong host immune defense, there is a drop in the amount of viral particles in the blood to a “set

point" [61]. The risk of viral transmission between hosts largely depends on the viral set point [50]. This phase is termed as latent phase since there are new or only minor clinical symptoms and typically varies between 5 to 15 years. The last phase is AIDS which is characterized by rapid loss of $CD4^+$ T-cells and highly elevated viral load. Due to the destruction of the body's immune mechanism, AIDS-associated opportunistic infections and malignancies are manifested during this phase.

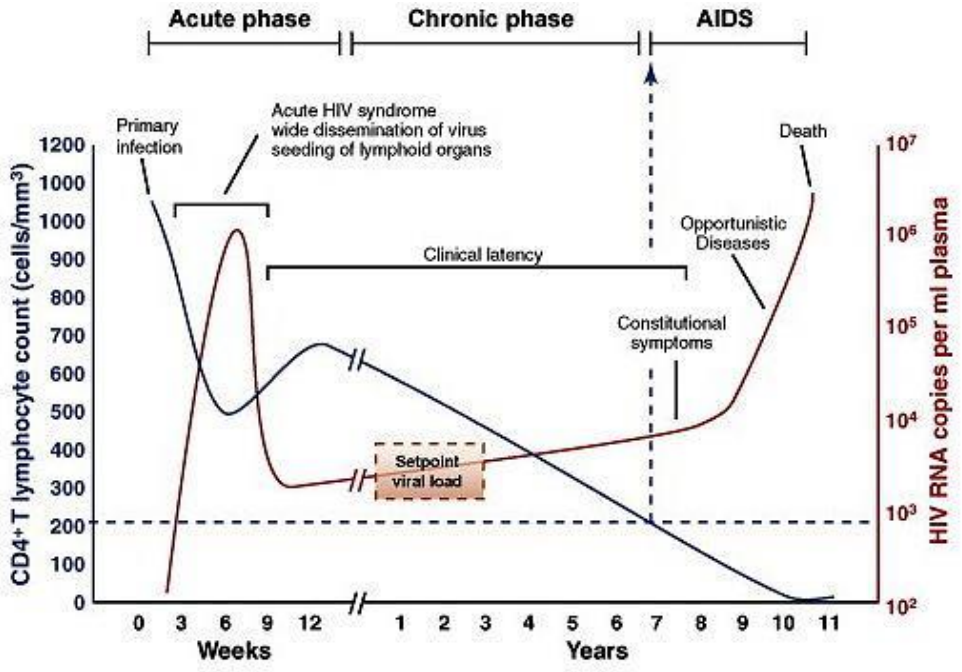

Figure 8. Natural course of HIV infection without therapy. Reprinted with permission from An and Winkler 2010 [62]

1.6 HIV-1 TRANSMISSION

The probability of HIV transmission depends on the amount of the infectious virus particles present in the body fluid, mainly blood and genital fluid in the index patient and the extent of exposure of that body fluid [50]. Also, the susceptibility of the exposed individual is also clearly important. Generally, transmission events occurs among individuals with a blood viral load >3.5 Log_{10} copies/ mL [63]. Globally, the major route of HIV transmission is via sexual intercourse (heterosexual or homosexual exposure), needle sharing among injecting drug users (IDU), unsafe needle use in medical practice, contaminated blood products through blood transfusions and mother-to-child transmission (MTCT) [53].

1.7 HIV-1 TRANSMISSION AND SUBTYPE DISTRIBUTION IN INDIA

The HIV-1 epidemic in India is heterogeneous both in terms of geography and mode of infections. Though the major route of transmission is heterosexual [7], the north-eastern states in India (mainly Manipur and Nagaland), have epidemics mostly driven by IDUs. Based on the national data 2010-11 [7] unprotected sex is the major source of spread with 87.4% heterosexual transmission and 1.3% homosexual followed by mother-to-child (5.4%) (Fig 9).

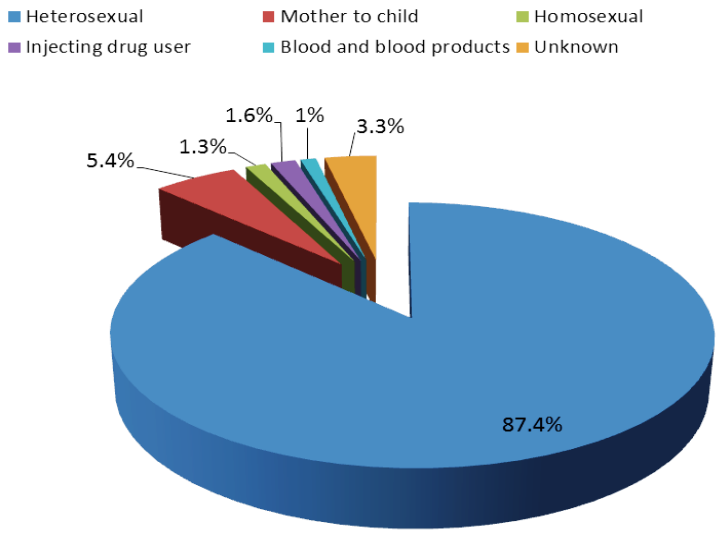

Figure 9. Mode of transmission of HIV, in India (2010-11) [Source: NACO Annual Report 2011-12]

Identical to the transmission event, data on molecular epidemiology in India indicates an epidemic dominated by HIV-1 subtype C [64-67] with occurrences of recombinant forms in some geographic pockets. In the past five years, there are increasing reports of recombinant strains in mainly northern and north-eastern India with principally CRF02_AG, URF_A1C and URF_BC [68-75]. The combined data from 53 previous studies on samples collected between 1991 and 2009 is depicted in Figure 10 [37-39, 68-70, 74-120]. The southern Indian epidemic is dominated by subtype C with sporadic reports of recombinants mainly URF_BC and URF_A1C [71, 104]. On contrary, in northern and north-eastern India, though there is dominance of HIV-1 subtype C, a high frequency of recombinants mainly URF_BC, URF_A1C, CRF01_AE and CRF02_AG are reported. However, the identification of HIV-1 subtypes in India is confounded by limitations regarding the representation of geographical regions, selection bias of the samples and small sample sizes. On the other hand, accurate virus genotyping and recombination identification techniques including the selection of genes are important. The studies in India mostly looked into a single gene, which may pose serious subtyping bias. On the contrary, selection of appropriate bioinformatics algorithms also have a founder effect in the typing of the strains [121, 122].

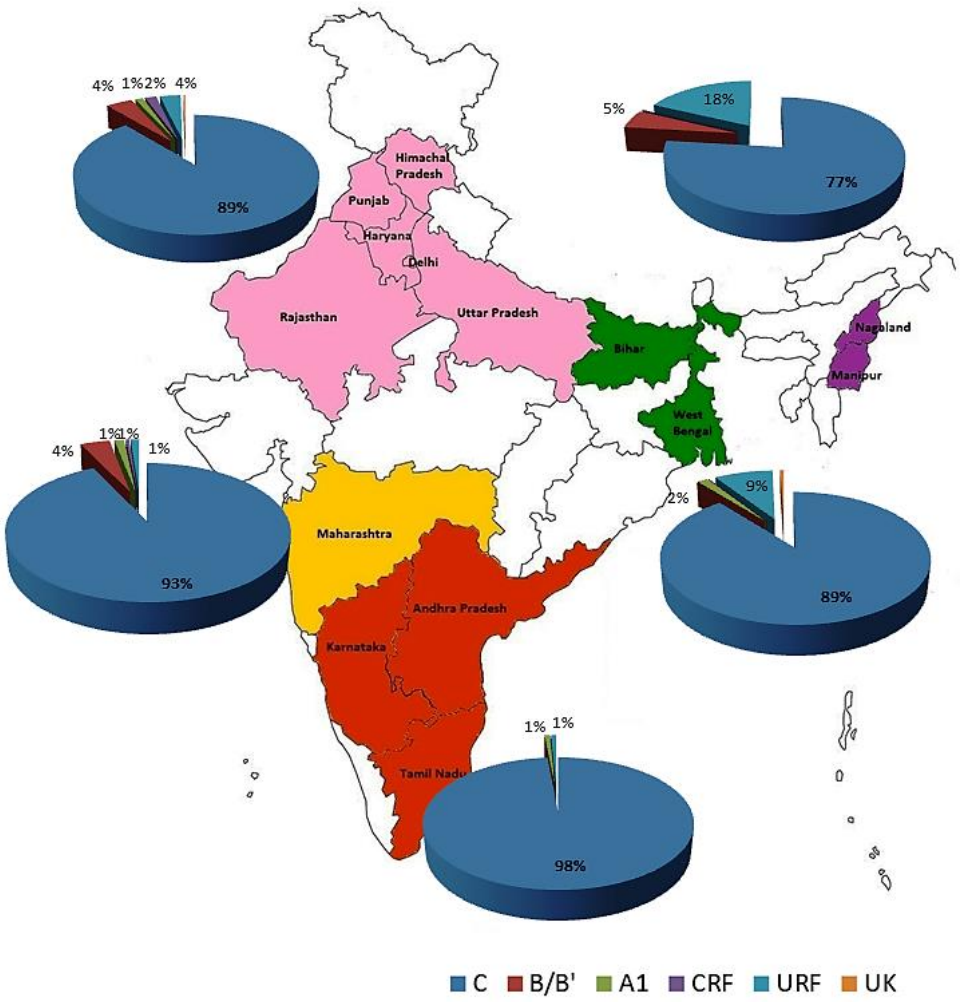

Figure 10. Molecular epidemiology of HIV-1 subtypes circulating in India (1991-2009). This includes data from all reports (n=53) in which the subtyping was determined by sequence analysis. A total of 1900 sequences of HIV-1 LTR (n=28), *gag* (n=156), *pol* (n=928), *env* (n=258), *tat* (n=102), *vpr (n=8), nef* (n=48), full length (n=13) and multiple genes (n=359) reported from 14 provinces were included in this figure. The subtype mentioned here is as reported by the authors.

1.8 CONSEQUENCES OF HIV-1 SUBTYPE DIVERSITY

Differential appearances of viral subtype through a dynamic genetic evolutionary process result in diversified HIV-1 pandemic. The consequential viral diversity has implications for possible differential co-receptor evolution to enter into the cells, transmission efficiency, rates of disease progression, responses to antiretroviral therapy, development of drug resistance, and vaccine development [20, 40, 41, 123-127]. The major challenges of the subtype diversity might be on therapeutic and preventive management of the disease. Is the therapeutic response subtype dependent? Does the evolution of drug resistance differ in different subtypes? The questions remain unexplained because of limited number of studies in non-subtype B dominated countries.

1.8.1 Differential co-receptor tropism

After the discovery of the CD4 molecule as the major cellular receptor for HIV entry [128, 129], the presence of secondary cellular receptors for HIV entry into the host has been identified [130, 131]. These co-receptors, particularly the chemokine receptors, C-C chemokine receptor type 5 (CCR5) and C-X-C chemokine receptor type 4 (CXCR4), have been the subject of extensive research attempting to elucidate viral entry mechanism and disease progression. Based on the co-receptor usage, HIV-1 strains are classified as R5-tropic (those that use CCR5 receptor), X4-tropic (those that use CXCR4 receptor) or dual-tropic strains (those that are able to use both the receptors) [132-134]. Also the viral population within an individual frequently consists of a mixture of such strains [40, 135]. Significant co-receptor related inconsistencies have been observed among subtypes, especially in the later stage of infection. Studies on subtype B had shown that in the later stage of disease, nearly 50% of the subtype B strains are X4-tropic [40]. It has also been shown that subtype D viruses are mostly X4- or dual-tropic throughout the course of infection compared to subtype A [134, 136-138]. On the contrary, subtype C strains frequently used exclusively the CCR5 co-receptor even in later stages of disease [41] and X4-switches have been reported to be extremely rare in Indian subtype C strains [139].

1.8.2 Efficiency in transmission

A growing body of scientific evidences suggests a differential transmission efficiency of different HIV-1 subtypes. A study from Tanzania showed a high *in utero* transmission rate of HIV-1 with subtype C compared to subtype A and subtype D or both [140]. Also, the study for finding out the reason for rapid spread of HIV-1 subtype C globally, is inconclusive. A recent study from southern Africa shows that a substantial proportion of HIV-1 subtype C-infected individuals maintain a high viral set point, which might have contributed to the high spread of HIV-1 subtype C viruses [42]. However, a study by Campbell *et al* [43] showed that subtype C and non-C subtypes do not differ significantly in terms of their viral set point after sero-conversion, which contradicts the findings of Novitsky *et al* [42]. A study conducted in Thailand among the IDUs showed an increased probability of transmission of CRF01_AE as compared to subtype B [141]. Since, multiple factors including viral and host immuno-genetics as well as clinical and geographical disease management strategies play a pivotal role in viral spread, it may be difficult to conclude whether the

differences in transmission probabilities of different subtypes were due to biologic properties associated with subtype only [44, 141].

1.8.3 Response to therapy

The scientific evidence on short term efficacy of antiretroviral therapy from resource-limited settings where non-B subtypes are prevalent have obtained comparable results to studies conducted in resource-rich settings in Europe and America [142, 143]. A recent systemic review showed that in sub-Saharan Africa, 67% of the patients achieved virological success after 2 years of therapy [143]. A study from Botswana on HIV-1 subtype C infected patients, the estimated virological failure are 22.1% and 30.1% at 3 and 5 years respectively [144]. A study from South Africa where subtype C is prevalent, reported that 61% of patients achieved virological suppression at 3 years [145]. A study from Burkina Faso reported that 81.8% of the patients remained virologically suppressed at 3 years [146]. However, a study from United Kingdom showed that viral rebound occurs more frequently in subtype C-infected patients, but viral suppression is comparable to those infected with subtype B [147]. The study also showed that patients infected with subtype C or subtype A, achieved virological suppression more rapidly than those with subtype B infection. However, the baseline viral load value is less in subtype C or A infected patients compared to subtype B. A recent study showed that a substantial proportion of HIV-1 subtype C-infected individuals maintained a high viral set point [42]. In the era of "treatment as prevention", this could make the treatment decisions more challenging. Moreover, the response to therapy may be hindered due to the subtype-related variability influencing resistance pathways.

1.8.4 Evolution of Drug Resistance

A growing body of scientific evidence suggests subtype specific differences in the evolution of drug resistance. Though the therapeutic responses are similar, there is evidently a probability of genetic metamorphoses among subtypes to yield differential patterns of resistance pathways in response to antiretroviral pressure [148-150]. The influence of drug resistance mutations on the sensitivity to different antiretroviral drugs has been well characterized in subtype B, however, this has not yet been done to the same extent for other subtypes due to its presence in resource-limited settings. Evaluation of virological and enzymatic data conceptually supports that naturally

occurring polymorphisms among different non-B subtypes can affect their susceptibility to antiretroviral therapy [148]. The resistance pathway against nucleoside reverse transcriptase inhibitors (NRTIs), even in subtype C, differs in different geographical locales. In Botswana, subtype C-infected failing patients treated with zidovudine (AZT) and didanosine (DDI) developed an atypical thymidine analogue mutation (TAM) resistance pathway (67N/70R/215Y), a mixture of TAM-1 (215Y/41L/210W) and TAM-2 (70R/67N/215F/219Q) pathways compared to subtype B [151]. However this has not been observed in subtype C in India, Malawi or South Africa when the patients were treated with same NRTIs [89, 152-154]. Another classic example is the development of K65R mutations in subtype C which confers resistance to tenofovir (TDF) and all other NRTIs except zidovudine [155-159]. The mechanistic view reveals that when subtype C reverse transcriptase (RT) synthesized DNA from subtype C RNA templates, favored pausing was seen at the nucleotide position responsible for the K65R mutation (AAG-to-AGG) but was not seen when the same enzyme used a subtype B RT template [156]. These template factors are assumed to increase the probability for the development of K65R mutation subtype C [156] which is now being observed in patients failing TDF-based first-line regimens [160] as well as in the minor viral population among the patients on cART including zidovudine or stavudine [159].

Similarly, resistance to non-nucleoside reverses transcriptase inhibitors (NNRTIs) also showed subtype dependent mutation patterns. In subtype B, the V106A mutation is selected after exposure to nevirapine (NVP) or efavirenz (EFV), while in subtype C and CRF01_AE it is V106M [89, 97, 151, 154, 161, 162].

Another important drug target is the protease region of the HIV-1 *pol* gene, which exhibits more inter-subtype variations compared to RT. Significant subtype-specific differences have been observed in this region. Analysis of 56 geographic populations identified 11 signature protease mutations at positions 12, 13, 15, 19, 35, 37, 41, 63, 64, 67 and 93. Of these, mutations in positions 63 and 93 are considered to be weak drug resistance mutations in subtype B [163]. A study from India also showed the high prevalence of L63P (63.2%) in Indian HIV-1 subtype C sequences from therapy-naïve patients indicating that L63P is most likely represent the naturally occurring polymorphism in subtype C sequences thus resistance-associated mutations in subtype B may not necessarily translate to resistance in other subtypes [39].

1.8.5 Vaccine development

The global variability of HIV-1 poses a formidable challenge for development of an effective vaccine [164]. A robust molecular epidemiology on HIV-1 subtype distribution builds the landmark of any operative large-scale immunization strategy [165]. The recent clinical trials in Thailand of ALVAC and AIDSVAX have renewed optimism in the possibility of developing an HIV-1 vaccine [166]. However, because of diverse subtypes in differential geographic locales, it is still controversial whether a single vaccine can protect against all the major subtypes and recombinant forms [165]. Due to the vast nature of the HIV-1 epidemic world-wide, it is impossible to conduct a single study to type the molecular epidemiology. Therefore, studies in different geographic pockets will be helpful for tracking the molecular epidemiology in the subtype specific epidemic. It will help in the designing of region-specific vaccines, which will be required to protect populations infected by a specific viral strains circulating in discrete geographic areas.

1.9 ANTIRETROVIRAL THERAPY

In the early 1990s, HIV-1 specific antiviral drugs were given as monotherapy. The treatment of HIV-1 infection was revolutionized in the mid-1990s, as the standard of care in HIV management advanced to include the administration of a cocktail of antiretroviral agents [167]. This introduction of combination anti-retroviral therapy (cART) has led to a dramatic reduction in the mortality and morbidity in HIV infection [167]. The six classes of anti-retrovirals that are approved by United States

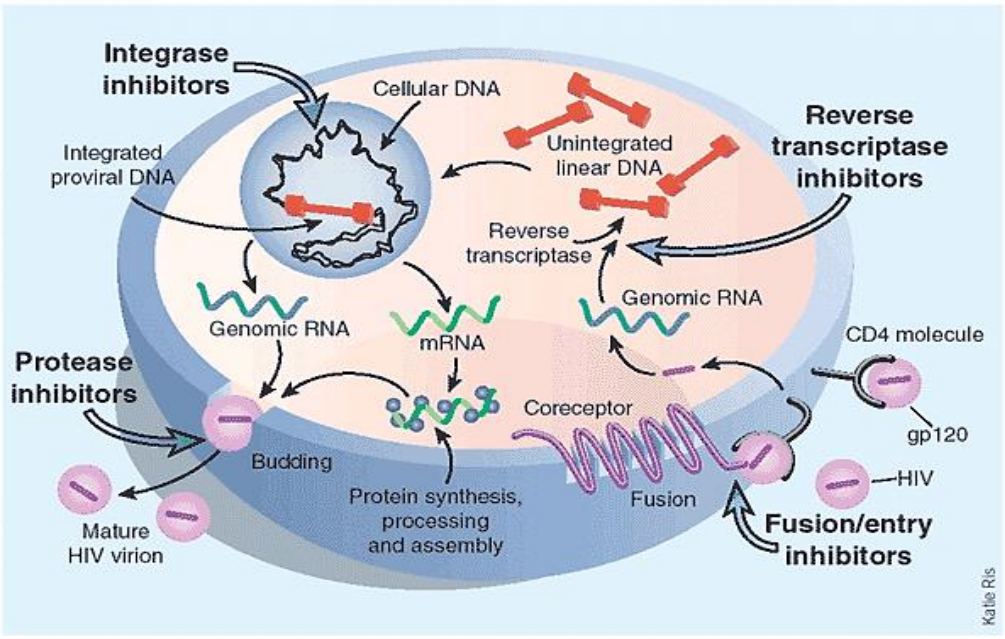

Figure 11. Targets of anti-retroviral drugs. Reprinted with permission from Fauci 2003 [1]

Food and Drug Administrations (US-FDA) and the European Medicines Agency (EMA) as of 2011 are NRTIs, NNRTIs, PIs, fusion inhibitors, integrase inhibitors, and entry inhibitor-CCR5 antagonists (Fig. 11) [156].

Current drug combinations are effective in stalling HIV replication, often leading to a 4 $\log_{10}$ copies/mL reduction in plasma viremia [168]. However these drug combinations do not eradicate the infection. The cART is usually initiated with three drugs (2 NRTIs +1 NNRTI/PI) as first-line drugs. More complex combinations with these or other classes are used for second-line or salvage therapy. The US-FDA and EMA approved drugs for HIV-1 therapy is depicted in Table 1.

Table I. Class of antiretroviral drugs use in the HIV-1treatment

Reverse Transcriptase Inhibitors (RTIs)		**Protease Inhibitors***	**Entry Inhibitors**
NRTIs	**NNRTIs**	Atazanavir (ATV)	***Fusion Inhibitor***
Zidovudine (AZT)	Rrilpivirine (RPV)	Darunavir (DRV)	Enfuvirtide, T-20
Lamivudine (3TC)	Etravirine (ETR)	Amprenvir (APV)	
Stavudine (d4T)	Delavirdine (DLV)**	Fosamprenavir (FPV)	***CCR5-Antagonist***
Emtricitabine (FTC)	Efavirenz (EFV)	Indinavir (IDV)	Maraviroc
Abacavir (ABC)	Nevirapine, (NVP)	Iopinavir (LPV)	**Integrase Inhibitor**
Tenofovir (TDF)		Nelfinavir (NFV)	Raltegravir
Didanosine (ddI)		Saquinavir (SQV)	Elvitegravir
		Tipranavir (TPV)	
		Ritonavir (RTV)	

*PIs generally prescribed with boosted RTV, ** Not approved by EMA

1.9.1 Mechanisms of anti-retroviral drugs

1.9.1.1 Reverse Transcriptase inhibitors: NRTIs and NNRTIs

There are two classes of reverse transcriptase inhibitors namely the nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs), and non-nucleoside reverse transcriptase inhibitors (NNRTIs). NRTIs were the first anti-retroviral drugs to be approved by US-FDA and EMA in 1987 [169]. To exert their anti-retroviral activity, upon entry NRTIs must first be intra-cellularly phosphorylated by cellular kinases to form their active 5'-triphosphate forms [170-173]. NRTIs interfere thereafter with the reverse transcriptase activity by competing with the natural substrates and once incorporated into the viral DNA, they act as chain terminators during RNA-dependent DNA or DNA-dependent DNA synthesis in either (–) or (+) strands of the HIV-1 proviral DNA [170, 171]. Tenofovir is the only nucleotide

reverse transcriptase inhibitor which contains a phosphate molecule in its structure and thus only requires phosphorylation to its diphosphate form for performing its antiviral activity [172]. The second class of RTI is NNRTIs which binds directly and non-competitively to the enzyme reverse transcriptase [174, 175]. The mechanisms of all NNRTIs are similar; they bind to distinct sites on the reverse transcriptase enzyme and not to the substrate (dNTP) binding site as the NRTIs. It blocks the DNA polymerase activity by altering conformational changes which disrupt the catalytic site of the enzyme [176]. The NNRTIs are inactive against HIV-2 viruses [177].

1.9.1.2 Protease Inhibitors

A protease inhibitor (PI) is a type of drug that targets the enzyme protease and a classic example of structure-based drug design [178]. HIV PR inhibitors specifically bind to the active site by imitating the tetrahedral intermediate of its substrate and disables or blocks the action of the protease [179]. Currently there are ten PIs approved by the US-FDA and EMA for clinical use [180]. The first generation PIs are saquinavir, ritonavir, indinavir, and nelfinavir. However, the clinical efficacy of the first generation of PIs was hindered because of low bioavailability, high pill burden and low genetic barrier to resistance. To overcome the issue, second generation of PIs boosted with ritonavir (amprenavir, fosamprenavir, lopinavir, atazanavir, tipranavir and darunavir), were developed. Boosting increases the plasma level of the PIs and has higher genetic barriers to the development of drug resistance mutations.

1.9.1.3 Entry Inhibitors

HIV-1 entry inhibitors can be subdivided into distinct classes based on disruption/inhibition of distinct targets/steps in the process. There are two classes of entry inhibitors approved by US-FDA and EMA; 1) Fusion Inhibitors (Enfuvirtide/T20) and 2) CCR5 antagonist (Maraviroc). Enfuvirtide (T20) is a linear 36–amino acid long peptide homologous to a segment of the HR2 region of HIV-1 gp41 which binds to the HR1 region of gp41 and blocks the formation of the 6-helix bundle necessary for fusion [172]. It is generally used as a salvage therapy on patients failed multiple therapies. On the other hand, the CCR5 antagonist, maraviroc, binds to hydrophobic pockets within the trans-membrane helices of CCR5 and acts as a negative allosteric modulator of the CCR5 receptor. Binding to CCR5 prevents HIV

gp120 from access to the co-receptor and thereby prevents the fusion process involving gp41 from proceeding further [181, 182].

1.9.1.4 Integrase Inhibitors

The first approved integrase inhibitor (II, INI or InSTI for integrase strand transfer inhibitor) is raltegravir, which contains a distinct β diketo acid moiety and has selective inhibitory activity against the strand transfer step of integration [183-185]. Recently, a second II has been approved, elvitegravir [186]. The integrase enzyme of the HIV-1 binds to the virally encoded DNA and joins it with the host chromosomal DNA followed by cellular repair activities which seals the viral DNA into the chromosome [187]. The presence of divalent cations in the catalytic core, a highly conserved region of integrase, is a key element of binding, forming covalent bonds with the phosphodiester backbone of DNA [187]. Raltegravir works by averting the formation of these covalent bonds with host DNA.

1.10 ROLL OUT OF ANTI-RETROVIRAL THERAPY IN INDIA

National AIDS Control Organisations (NACO) sponsored by Govt. of India initiated free ART services on 1st April 2004 in eight government hospitals located in six high prevalence states [188]. Since then, the program has been scaled-up rapidly to 292 (as of December 2010) Govt. and other supported health care centres providing free first line anti-retroviral therapy [7]. As of January 2012, a total of 486,173 PLHAs are receiving free anti-retroviral therapy [189]. The delivery of care and treatment services for PLHAs is provided through a three-tier structure; (i) ART centres, (ii) Link ART Centres and (iii) Centres of Excellence (COE). Free second-line ART, prioritised for women and children, is available only at 10 COEs, and as of April 2011, there are nearly 2,000 patients receiving free second line ART [190]. The ART centres are generally established in medical colleges based on the prevalence of PLHAs in the region and NACO provides support in terms of personnel (doctors, nurses, councillors etc), antiretroviral drugs, CD4 testing and drugs required for treatment of opportunistic infections.

1.10.1 Criteria for ART initiations and regimen in India

First-line ART is the initial regimen prescribed for an ART naïve patient when the patient fulfils national clinical and laboratory criteria to start ART i.e. WHO clinical stage I and II, $CD4^+$-T cell count <250 cells/mm^3, WHO Clinical stage II, $CD4^+$-T cell count <350 cells/mm^3 or WHO clinical stage IV. The first line drugs include one NNRTI, either nevirapine (NVP) or efavirenz (EFV), in combination with two NRTIs; zidovudine (AZT) or stavudine (d4T), and lamivudine (3TC). The alternate first-line is to replace AZT or d4T with tenofovir (TDF) due to intolerance to AZT/d4T [190]. The second line therapy includes AZT/d4T/TDF with 3TC/FTC along with the protease inhibitor atazanavir/ritonavir (ATV/r).

1.10.2 Therapeutic efficacy

The data on therapeutic response on first-line therapy in India is limited. A study from a southern Indian city (Chennai) with a small number of patients (n=40) using generic ART, showed a rapid suppression of HIV-1 viral load and 95% patients achieved viral load <400 copies/mL within 6 months of therapy [191]. A study from Bangalore among perinatally-infected children, revealed that 85% were virologically suppressed on first-line ART for >2years [192]. A recent study shows that at twelve months after initiation of ART, 64.6% and 75% of the patients in another two clinics (Mumbai and Chennai) achieved viral load suppression (<1000 copies/mL) [193]. However, data on long term efficacy of the national first-line regimen is lacking.

1.11 HIV-1 DRUG RESISTANCE

The development of drug resistance is a critical challenge to the success of the antiretroviral therapy. Emergence of drug resistance results in incomplete suppression of HIV replication, and is fuelled by inefficient access to viral load monitoring, drug resistance testing, adherence to ART and inappropriate drug prescriptions with monotherapy mainly, which are major problems in many resource-limited settings. The transmission of drug resistant strains may compromise the effectiveness of national ART programs, especially in resource-limited settings where only a limited number of drugs are available [194].

Several drug resistance mutations lists have been defined for surveillance of transmitted drug resistance mutations (TDRM) and acquired drug resistance mutations, such as

French National Agency for AIDS Research drug resistant algorithm [http://www.anrs.fr/]; IAS-USA Mutations Associated With Drug Resistance [195]; Stanford HIVdb drug resistance interpretation algorithm [http://hivdb.stanford.edu/]; Los Alamos National Laboratories HIV Sequence database [www.hiv.lanl.gov] and Rega Institute Drug Resistance Interpretation Algorithm [http://regaweb.med.kuleuven.be/software/regadb], which all are available to interpret drug resistance associated mutations [196]**.** These mutations lists are mainly based on the data available for HIV-1 subtype B. Although comparable patterns have been reported on non-subtype B strains [197], a selection of drug- resistance mutations in subtype B have been identified as naturally occurring polymorphisms in non-B subtypes [198, 199]. Such misinterpretation and subsequent inconsistencies in reporting rates of transmitted HIV-1 drug resistance in some of the high HIV-prevalence areas of the world, have led to reconsideration of the TDRM list to make it globally relevant.

In 2007, a standardized list of 80 RT and PR mutations was assembled by WHO to assist population based surveillance of TDRM [200]. The list was updated in 2009 and at present, there are 93 mutations in both RT (34 NRTI mutation in 15 positions and 19 NNRTI in 10 positions) and PR (40 mutations in 18 positions) regions of HIV-1 Pol [196] . The International AIDS Society-USA provides drug resistance mutations lists which are mainly useful for interpretation of drugs resistance mutations in therapy failure patients with acquired drug resistance, and are updated periodically every years (current update March 2013) [Appendix-I (RT) and II(Major-PI)] [201].

1.11.1 Transmitted drug resistance

A meta-analysis of data collected between time periods <2001, 2001-2003 and >2003 through 2009, indicated that the prevalence of transmitted drug resistance across Europe was 10.9% (95% CI: 10.6-11.3%) while in North America, it is even higher at 12.9% (95% CI: 12.2-13.7%) [202]. Combined data from studies from resource limited settings like Latin America (26 studies), Asia (23 studies) and Africa (47 studies) where the ART was introduced much later, has reported lower levels of TDRMs at 6.3%, 4.2% and 4.7% respectively [202]. However, a recent meta-regression analysis of the global trends of transmitted drug resistance estimated that in East Africa there has been an increase of TDRMs from 0.9% to 7.4% in 8 years

after ART roll-out [194]. The TDRMs found most frequently are K103N/S, Y181C/I/V and M184V/I.

1.11.2 Acquired drug resistance

Previous studies have shown that treatment interruptions, sub-optimal adherence and improper dosing can lead to acquired drug resistance [203-205]. Most of the research on acquisition of drug resistance mutations, and the impact of DRM on ART have been performed primarily for subtype B strains due to its presence in resource-rich settings. However, the major challenges remain in resource-limited settings where non-B subtypes are present. Different HIV-1 subtypes carry subtype specific genetic signature or polymorphisms in its genome that could alter the structure of the viral proteins targeted by the drugs [206]. Polymorphisms in reverse transcriptase genes generally do not occur in known drug resistance mutation sites of NRTI or NNRTI drugs. In contrast, the protease gene is more prone to subtype specific signatures in known drug resistance mutation sites of PIs. The subtype B drug resistance mutations in positions K20R, M36I and V82I have been observed as naturally occurring polymorphisms in non-B subtypes [207]. Recently, our group showed 63% prevalence of L63P mutation in PI naïve Indian subtype C sequences, a mutation that is also selected by PIs in subtype B sequences, indicating that this mutation is likely to be a naturally occurring polymorphism in Indian C strains [39].

Several studies have shown the differential accumulation of NRTIs and NNRTIs mutations in different subtypes. There are two well-known pathways to thymidine analogues conferring resistance to NRTIs, namely thymidine analogues mutation (TAM) pathway 1 and 2. TAM-1 includes mutations M41L, L210W and T215Y; while TAM-2 is characterized by mutations D67N, K70R, T215F and K219Q/E [195, 208]. In subtype B, the TAM-1 is more common than the TAM-2 pathway while TAM-2 mutations are more frequent in subtypes C and 01_AE [209]. However in real time, a mixture of these mutational pathways can also be observed. Another example is K65R, a NRTI mutation which develops rapidly in subtype C viruses when exposed to TDF, which might be due to template dependent dislocation mechanisms [158, 160, 210]. Similarly, V106M, a NNRTI associated mutation is acquired more frequently in NNRTI-exposed subtype C viruses compared to subtype B, where V106A mutation is more common [211].

1.12 DRUG RESISTANCE STUDIES IN INDIA

Despite widespread availability of antiretroviral, limited information is available on the prevalence of transmitted as well as acquired HIV-1 drug resistance in India. Surveillance and monitoring of drug resistance reports are coming only from a handful of major cities like Chennai, Pune, Mumbai, Chandigarh and Bangalore. The studies in surveillance of transmitted drug resistance among recently diagnosed ART-naïve individuals in India are limited [101, 103, 105]. However, cross-sectional studies on chronically infected HIV-1 positive study populations showed the prevalence of TDRMs between 1.6% (Mumbai, 2004) and 39% (Chandigarh, 2008) against first-line drug regimens that are widely available in India [89, 92, 94-96, 100]. The NRTI mutations are mainly T69D, D67N, M41L, M184V and T215S while the NNRTI mutations are K101E, K103N, Y181C, V106M and G190A. A recent combined study using primary clinical isolates and secondary database sequences from seven states (Tamil Nadu, Andhra Pradesh, Maharashtra, Delhi, Karnataka, Madhya Pradesh and Manipur) of India identified a low baseline prevalence of transmitted drug resistance below 5% (Neogi et al unpublished data).

Studies on the evolution of drug resistance are limited. Due to the lack of virological monitoring, most of the studies were restricted to a cross-sectional analysis on immunologically failure patients. A study from Chennai showed M184V, D67N, M41L, K70R and T215Y as major NRTI mutations while K103N, Y181C and G190A as major NNRTI mutations [212]. Similar patterns were observed from Bangalore [205] and Mumbai [89]. As a consequences of limited access of viral load monitoring, most of the time the patients are coming with either immunological or clinical failure. In those cases a wide range of cross-resistance even in the newer class of drugs were reported. A study from Bangalore reported that in a failing regimen including nevirapine or efavirenz, 40% of the subtype C infected Indian patients had at least partial resistance to etravirine, a newer class of NNRTI drug [213]. Similar observation was reported from Mumbai where the patients were exposed to NVP/EFV for 1 to 8 years, 45% were resistance to etravirine [214].

2 RATIONALE OF THE STUDY

HIV-1 translational genomics: exploring the past and present to understand the future

The genetic diversity and multiplicity of HIV is a pivotal factor in HIV vaccinology. In order to foretell the future course of HIV-1 disease management strategy in India, we need to understand the precedent history of HIV-1. Despite being the third largest country with PLHA, studies on basic and translational research on understanding HIV-1 epidemiology and biology remain inadequate in India. Attempts have been made to understand the origin of HIV-1 subtype C in India [215, 216] and the tMRCA of HIV-1 subtype C in India was dated back to the mid-to late 1970s. Although these studies have a strong scientific basis, there is an element of uncertainty in this estimate due to the inherent bias of a small sample size, use of information from a secondary database, and limited detailing of the geographic origins of samples. Furthermore, tMRCA estimates based on a single gene analysis may suffer from a limitation in that the rate of evolution of the viral genes may differ significantly due to the variable selection pressure exerted by host factors [217]. As a consequence, the actual date of viral introduction may be underestimated [218].

The continual evaluation of transmitted and acquired drug resistance mutations is the back bone of long term therapeutic success and effective disease management. The efficacy of an antiretroviral treatment regimen depends on several factors, including viral, behavioural and host genetics. The critical one being the evolution of drug resistance mutations accumulated over time, both transmitted and acquired. Unlike resource-rich countries, drug resistance genotyping at baseline is not part of standard-of-care management in India. This is mainly because of limited resources, scarcity of laboratory infrastructure and trained human resources to perform the test. Thus, drug resistance genotyping is restricted to the research arena in India and is currently not used for guiding policy on therapeutic strategies.

Similarly, data on long term virologic outcome of patients using the Indian national first-line regimen with two NRTIs and an NNRTI is limited. A previous study with a small number of patients using generic antiretroviral therapy identified a rapid suppression of HIV-1 viral load [191]. Another study showed that of the 927 patients

whose treatment outcome information was available, 71% were alive at the end of two years [219]; however, virological data was not available as viral load monitoring was not part of standard practice. The data on treatment efficacy beyond two years was not examined.

Therefore, to improve our knowledge for better therapeutic and preventive strategies to combat HIV/AIDS in India, this thesis describes the central importance of molecular epidemiology in translational genomics of HIV-1 subtype C from origin to therapeutic response. In a systemic approach, we identified the molecular epidemiology of HIV-1 subtypes circulating in India and described the tMRCA of predominant HIV-1 subtype C strains. Next, we examined drug resistance mutations in children, adolescents and adults, the role of host factors in evolution of drug resistance, population dynamics of viremia and viral co-receptor tropism in perinatal transmission. Finally, we studied long-term therapeutic responses on Indian national first-line antiretroviral therapy. This thesis thus offers a glimpse into the past and present of HIV-1 molecular epidemiology which can be pieced together as a mirror to assist in understanding the future of the epidemic for India and the world.

3 AIMS AND OBJECTIVES

The general aim of this thesis is to understand the molecular phylogeny of HIV-1 in India with its direct implications on therapeutics and vaccine development. As India is poised to provide and scale up free treatment support to the people living with HIV/AIDS and initiate sub-unit based vaccine trials using the genes from the HIV-1 subtype C viruses, my translational research has potentials to link the basic science findings with the arena of clinical investigation and public health strategies for effective therapeutic management and vaccine developments.

The specific aims are:

1. To describe the molecular epidemiology of HIV-1 subtypes in India based on samples collected from nationwide clinical cohorts between 2007 and 2011 and to reconstruct the time to the most recent common ancestor (tMRCA) of the predominant HIV-1 subtype C strains.
2. To identify the clinical significance of human APOBEC mediated drug resistance mutation and to explore use of proviral DNA genotyping for drug resistance mutation surveillance and monitoring.
3. To characterise the HIV-1 subtype C infection in Indian children and adolescents with perinatal HIV-1 transmission with regard to transmitted drug resistance and genotypic co-receptor and the association with clinical, virological and host genetic determinants.
4. To investigate the long term virological outcome among Indian HIV-1 infected patients on national first-line therapy, following initial viral suppression and to evaluate the factors associated with virological failure and emergence of drug resistance mutations.

4 MATERIALS AND METHODS

4.1 STUDY SETTINGS

The majority of the patients hailed from three states of southern India, Karnataka, Tamil Nadu and Andhra Pradesh. For **Paper I** additional samples were collected from another four states Punjab and Haryana in northern India, Manipur in north-eastern India, and Madhya Pradesh in central India (Fig. 12). The southern Indian patients attended two tertiary care hospitals, St. John's Medical College and Hospital in Bangalore and Krishna Rajendra Hospital in Mysore. The participants from northern India were from Post Graduate Institute of Medical Education and Research in Chandigarh and those from central India were attending from Ruxmiben Deepchand Gardi Medical College, Ujjain. The participants from north-eastern India were from the Manipur Network of Positive People (MNP+). A portion of the children and adolescents who participated in the study were from Sneha Care Home in Bangalore which is a faith-based residential care facility for children infected with HIV who have lost one or both parents. The participants from Sneha Care Home routinely obtain medical attention and treatment from St. John's Hospital, Bangalore.

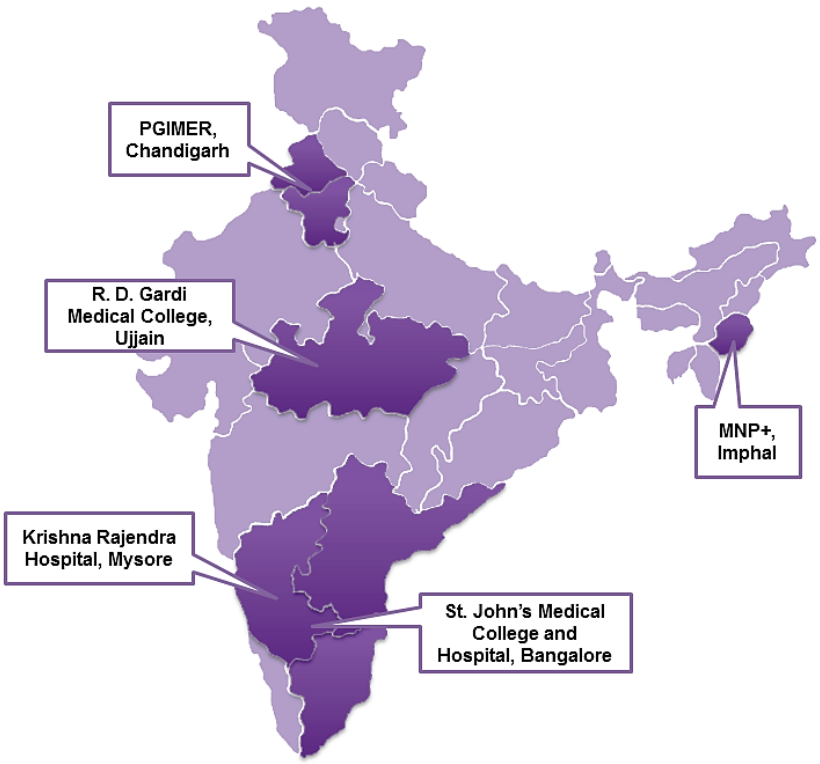

Figure 12. Study sites for Paper I from where the samples were collected. The dark colour shed represents the geographical origin of the study subjects.

4.2 STUDY DESIGN AND PATIENT POPULATION

Among the five papers presented in the thesis, four (**I-IV**) were cross-sectional studies while the final study (**V**) was a longitudinal cohort with a follow-up of two years. The total number of patients involved was 695, and includes 20% children ≤10 years of age (136/695), 7% adolescents aged 11-19 years (47/695) and 73% adults (512/695). Among them 55% (384/695) were therapy-experienced while the rest were therapy-naïve. The study design and populations for the respective papers are outlined below.

Paper I: A single peripheral blood sample was obtained from HIV-1 infected seropositive subjects (n= 168) hailing from different regions of India including northern (Punjab and Haryana), southern (Karnataka, Tamil Nadu and Andhra Pradesh), north-eastern (Manipur) and central (Madhya Pradesh) regions between October 2007 and November 2011.

Paper II: Cross-sectional blood samples were collected from 102 HIV-1 positive individuals from southern India between November 2009 and October 2011. A fraction of whole blood was kept. Plasma was separated out from the remaining part of the samples.

Paper III: A cross-sectional blood sample was obtained from 105 study subjects between March 2007 and October 2011, during a routine follow up visit in the clinic. These children were aged between 2 and 16 years and were perinatally infected, ART-naïve, and with no antenatal exposure to ART.

Paper IV: The study participants for this study were perinatally-infected children (n=139) aged 2-17 years who attended the Pediatric Infectious Disease Clinic, St. John's Medical College and Hospital, Bangalore, India and resided in southern Indian states Karnataka, Andhra Pradesh and Tamil Nadu. Cross-sectional blood samples were collected between August 2007 and May 2011 from 78 antiretroviral therapy-naive and 61 ART-experienced children.

Paper V: The patient populations were derived from a large observational clinical cohort of 533 patients from two tertiary care hospitals in southern part of India, which adhere to the national therapeutic guidelines [205, 220]. The criteria used for inclusion were: (i) Adult HIV-1 infected patients initiated on first-line ART for a minimum period of 6 months, and (ii) a viral load <100 copies/ml at the time of entry into the

study. The patients were followed up for every 3 months and evaluated for self-reported adherence, adherence barriers and other health behaviors. The blood samples were collected in every 6 months for a total follow-up period of 2 years beginning in August 2007.

Table II. Schedule of events in Paper V.

	BL	MT03	MT06	MT09	MT12	MT15	MT18	MT21	MT24
CD4 Count	X	-	X	-	X	-	X	-	X
Viral Load	X	-	X	-	X	-	X	-	X
DR Genotyping*	X	-	X	-	X	-	X	-	X
Adherence	X	X	X	X	X	X	X	X	X

***If viral load >2000 copies/mL**

4.3 PATIENTS' DEMOGRAPHY AND LABORATORY ASSESMENT

Basic demographics like age, gender, year of HIV-1 sero-diagnosis, treatment details for the therapy-experienced patients, were documented during the first study visit. Routine $CD4^+$T-cells were measured either with FACSCalibur system (BD, USA) (at Central laboratory, St. John's Hospital, Bangalore) or a single platform flow cytometric assay (PCA system; Guava Technologies Inc., Hayward, CA, USA at Molecular Diagnostics and Genetics, Reliance Life Sciences, Mumbai, India.). Viral load for the study purpose, which is not a part of the standard of care, was measured either by Abbott m2000rt system (Abbott Molecular Diagnostics, US) or an in-house real time polymerase chain reaction with TAQMAN assay (Molecular Diagnostics and Genetics, Reliance Life Sciences, Mumbai, India).

4.4 ADHERENCE MEASUREMENTS

Adherence during the past month at baseline and past three months (**V**) was measured months during the study visit using a visual analogue scale (VAS) [221]. The VAS is a horizontal line marked with numbers between 0 and 100 on which the participant use to indicate a point that best corresponds to the percentage of prescribed pills actually taken according to his/her self-assessment. Although multiple objective and subjective based self-reported adherence measures are available, the VAS has been found to be the best predictor of virological failure in this setting [222].

4.5 DEFINITIONS OF TERMS

Viral rebound: Intermittent viral rebound was defined as a single episode of detectable viral load >100 copies/mL.

Virological failure: Virological failure was defined as two successive viral load values >400 copies/mL.

Perfect Adherence: Perfect adherence was defined as a combination of two types of adherence measurements into one dichotomous variable: Percent of prescribed pills taken, assessed with the Visual Analogue Scale (VAS) and Treatment interruptions, defined as having missed medications for more than 48 hours in the past three months. To be classified as "perfectly adherent", the patients had to report 1) taking 100% of prescribed doses in the past three months (past one month at baseline) at each study visit in which they were present and 2) zero treatment interruptions during the 2-year study period.

4.6 DNA/RNA EXTRACTION

Genomic DNA was extracted from whole blood using QIAamp Blood DNA kit (Qiagen, Germany) while viral RNA was extracted using QIAamp viral RNA kit (Qiagen, Germany) and stored for polymerase chain reactions (PCR), reverse transcriptase-PCR of viral genes, drug resistance genotyping and co-receptor tropism analysis, as described in later section.

4.7 PCR AND SEQUENCING

Reverse transcriptase polymerase chain reactions were performed to prepare cDNA from plasma viral RNA using random hexamer primer (Promega, USA). HIV-1 p17 region *gag* (HXB2 position 790-1190), RT region of *pol* (HXB2 position 2598-3254) and C2V4 region of *env* (HXB2 position 7050-7550) from cDNA and whole blood genomic DNA were amplified by nested polymerase chain reaction (PCR) using iNtRON Taq Polymerase (Intron Biotech, South Korea) and the primers listed in Table III. The PCR conditions used for different genes are presented in Table IV.

Table III. Primes used for PCR amplification and sequencing.

Fragment	Primer sequence	HXB2 Pos	Paper
gag p17			
SJ017F (Outer)	5'-TCTCTCGACGCAGGACTCGGCTTGCTG-3'	682-708	I
SJ017R (Outer)	5'- TAACATTTGCATGGCTGCTTGATGTCC-3'	1366-1392	
SJ018F (Inner)	5'-CTAGAAGGAGAGAGAGATGGGTGCGAG-3'	776-800	
SJ018R (Inner)	5'-CTTGTGGGGTGGCTCCTTCTGATAATG-3'	1310-1336	
pol reverse transcriptase (RT)			
RT04F (Outer)	5'-CCTATTGAAACTGTACCAGT-3'	2559-2578	I, II, III, V
RT05R (Outer)	5'-ACTGTCCATTTATCAGGATG-3'	3252-3271	
RT07F (Inner)	5'-AAGCCAGGAATGGATGGCCCA-3'	2585-2606	
RT06R (Inner)	5'-CCATTTATCAGGATGGAGTTC-3'	3246-3266	
env C2V4 region			
IN_ED5 (Outer)	5'-ATGGGATCAAAGCCTAAAGCCATGTG-3'	6557-6582	I, IV
IN_ED12 (Outer)	5'-AGTGCTTCCTGCTGCTCCCAAGAACCCAAG-3'	7782-7811	
IN_ES7 (Inner)	5'-CTGTTAAATGGCAGTCTAGC-3'	7002-7021	
IN_ES8 (Inner)	5'-CACTTCTCCAATTGTCCCTCA-3'	7648-7668	
Note: Inner primers used for sequencing			

Table IV: PCR Conditions used for the respective gene fragments.

Proviral *gag p17, pol* reverse transcriptase (RT), *env* C2V4 region		**Paper**
First round PCR	1 x 95°C: 2 min 3 x 94°C: 1 min, 55°C: 1 min, 72°C: 3 min, 32 x 94°C: 1 min, 60°C: 1 min, 72°C: 3 min 1 x 72°C: 10 min.	I-II
Nested PCR	1 x 94°C: 2 min, 3 x 94°C: 1 min, 55°C: 30 sec, 72°C: 3 min, 32 x 94°C: 1 min, 60°C: 1 min, 72°C: 1 min 1 x 72°C: 5 min	
Plasma Viral *env* C2V4 region		
First round PCR	1 x 94°C: 5min; 39 x 94°C: 20 sec, 55°C: 30 sec, 72°C: 80 sec; 1 x 72°C: 7 min	IV
Nested PCR	1 x 94°C: 5min; 39 x 94°C: 20 sec, 50 °C: 30 sec, 72 °C: 60 sec; 1 x 72° C: 7 min	
Plasma Viral *pol* reverse transcriptase (RT)		
First round PCR	1 x 95°C: 3 sec; 39 x 95°C: 20 sec, 55°C: 30 sec, 72°C: 30 sec;	II,III, V
Nested PCR	1 x 95°C: 3min; 29 x 95°C: 20 sec, 55°C: 30 sec, 72°C: 30 sec; 1 x 72°C: 5 min	

Bidirectional population-based sequencing was performed using the forwards and reverse inner primers. The sequences were edited manually in BioEdit version 7.1.3.0 software. All the laboratory sequences were subjected to rigorous quality control using neighbor joining phylogenetic tree constructed using Molecular Evolutionary Genetics Analysis software version 5 (MEGA 5) to check for PCR cross-contamination [223]. Identical sequences were excluded from the analysis.

4.8 HIV-1 SUBTYPING (I – V)

Subtyping of specific gene fragments from clinical strains was initially performed using a maximum likelihood (ML) phylogenetic tree constructed with appropriate models, using the MEGA 5 software with reference sequences of different subtypes, downloaded from HIV-1 Los Alamos Database (LANL, www.hiv.lanl.gov). Gene segments clustering with a specific subtype were further used for recombinant screening in Recombination Identification program version 3 (RIP 3.0). To confirm subtype designation, REGA HIV-1 Subtyping Tool - version 2.0 [224] and SimPlot analysis [109] were also used. The mosaic pattern of each URF was confirmed by phylogenetic analysis of the recombination fragments.

4.9 ESTIMATING TIME TO THE MOST RECENT COMMON ANCESTORS (TMRCA) (I)

Full length subtype C available at the Los Alamos database, were selected as reference sequences since they covered the *env*, *gag* and *pol* regions investigated in this study. The phylogenetic analyses were performed in BEAST v.1.6.2 [225] using the general time reversible (GTR) substitution model with inverse gamma distribution for all three data sets (the best fitting model for all three data sets according to jModelTest [226]). Molecular clock analyses were performed to predict when lineages branch off (split) from a common ancestor in the evolutionary tree. As the different HIV-1 genes showed different rates of evolution, two relaxed molecular clock models, 'Relaxed: exponential' and 'Relaxed: log-normal' were tested in combination with four different coalescent tree priors: constant size, exponential growth, logistic growth and Bayesian skyline. The resulting log-files were analyzed in Tracer v.1.6.2 [225]. Bayes Factor analysis showed that the relaxed log-normal clock with Bayesian Skyline was the most appropriate model for all three genes. Tip dates (year of sampling for each sequence) together with a previous estimate of the age of subtype C in the year 1952 [217] were

used for calibration of the molecular clock (the prior was set to Normal distribution, 59±6 years since the last year of sampling, 2011), and rates of evolution were automatically calculated from these data. The analyses were run for 100 million generations with sampling every 10,000 generation, and the sampled trees were annotated using TreeAnnotator v1.6.2 and visualized in FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/).

4.10 DRUG RESISTANCE (II, III AND V)

The transmitted drug resistance was interpreted using World Health Organization (WHO) recommendations for surveillance of drug resistance mutations updated in 2009 (SDRM_2009) which includes 93 mutations in the 53 RT and 40 PR region of HIV-1 Pol [196]. To further increase the accuracy of this analyses, certain mutations included in the SDRM_2009 which displayed characteristics of likely polymorphisms (defined as those mutations which are normally present at a frequency of ≥0.5% in therapy-naïve patients, and at a frequency of <0.5% among those failing therapy [196]) were eliminated. The acquired drug resistance mutations were interpreted using International AIDS Society-USA update 2011 which includes 19 NRTI mutations in 16 positions, and 34 NNRTI mutations in 16 positions between amino acid residues 17 to 235 of RT region of pol [195].

4.11 CCR5Δ32 GENOTYPING (IV)

CCR5Δ32 genotyping was performed by PCR with subsequent gel electrophoresis as described previously [227]. PCR amplicons of either 262 bp (CCR5 wildtype) or 230 bp length (CCR5Δ32 deletion) were visualized in a 5% ethidium bromide-stained agarose gel.

4.12 GENOTYPIC TROPISM TESTING: SELECTION OF PREDICTION TOOL FOR SUBTYPE C (IV)

A training data set of 465 HIV-1 subtype C V3 loop sequences with known phenotypic tropism results were compiled from previously published literature [228-231] and LANL HIV sequence database. These sequences were subjected to genotypic tropism testing using the different prediction tools for the selection of tools for subtype C viruses Table V.

Table V. Performance of genotypic tools in predicting X4-tropism in HIV-1 subtype C sequences with known phenotypes.

		Phenotypic		Concordance	Sensitivity	Specificity
		X4 (n=80)	R5 (n=385)			
Geno2Pheno$_{2.5\%}$	X4	57	3	94.4	71.3	99.2
	R5	23	382			
Geno2Pheno$_{5\%}$	X4	67	7	95.7	83.8	98.2
	R5	13	378			
Geno2Pheno$_{10\%}$	X4	69	14	94.6	86.3	96.4
	R5	11	371			
Geno2Pheno$_{15\%}$	X4	69	18	93.8	86.3	95.3
	R5	11	367			
Geno2Pheno$_{20\%}$	X4	71	41	89.3	88.8	89.4
	R5	9	344			
C-PSSM	X4	68	21	92.9	85	94.6
	R5	12	364			
ds(Kernel)	X4	62	3	95.5	77.5	99.2
	R5	18	382			
11/25 rule	X4	56	3	94.2	70	99.2
	R5	24	382			
11/24/25 rule	X4	59	3	94.8	73.8	99.2
	R5	21	382			
Net Charge $\geq$ 6	X4	61	20	91.6	76.3	94.8
	R5	19	365			

Though all the genotypic tools showed a good performance in detecting X4 tropism, a better sensitivity (86.3%) and specificity (96.4%) was obtained by Geno2pheno[co-receptor] with 10% false positive rate (FPR). Thus, Geno2pheno with 10% FPR was used to determine the genotypic tropism testing which was also recommended by European guidelines for HIV-1 tropism testing [232, 233]. Minor populations with all probable combinations of V3 sequences were also included. The tropism interpretation was as follows: the sample was considered "X4-tropic", if all sequences from that patient showed FPR ≤10%; or "R5-tropic", if all the sequences from the patient showed FPR >10%. Samples which contained sequences with differing FPRs that were both ≤10% and >10% were classified as "R5/X4-tropic" samples.

4.13 OTHER BIOINFORMATICS ANALYSIS

4.13.1 *Intra-population divergence (II)*

The intra-population divergence (genetic distance of each of the study sequences to the Indian consensus C sequence) was calculated in MEGA 5 software.

4.13.2 *Estimation of G-to-A substitutions (II)*

To estimate G-to-A substitutions, proviral DNA sequences were aligned against the consensus Indian subtype C sequence [38]. The hA3G/F mediated GG-to-AG and GA-to-AA scores, for each sequence were calculated as described previously [234, 235]. The consolidated hA3G/F mediated G-to-A hypermutation score: [(Number of GG-to-AG or GA-to-AA substitutions/number of GG or GA in Indian consensus sequence)/(Total number of mutations/Sequence length)]. G-to-A preferences were calculated as described [234].

4.13.3 *Identification of hypermutated sequences (II)*

Hypermut 2.0 software was used to determine hA3G/F mediated hypermutated sequences [236]. To further increase the confidence in the analysis, cluster analysis of preference for G-to-A substitutions relative to consolidated hA3G/F score and sequence analysis of 21 hA3G and 20 hA3F target sites in the 17-235 aa of the RT region identified by APOBEC3G-mediated defectives (A3GD) indices [237] were also used. Hypermutations were labeled into a dichotomous variable if identified by one of the methods mentioned. Mutations were designated as lethal if there was a stop codon in the open reading frame (ORF).

4.14 STATISTICAL ANALYSIS

Differences between groups on continuous variables were examined by independent sample *t*-test or Mann–Whitney *U*-test, while categorical variables were analyzed via frequencies and cross-tabulations, with $\chi 2$ tests or Fisher's exact test to assess the significance of the associations. A Spearman rank correlation (r_s) was used for analysis of potential correlation between continuous variables. Multivariate logistic regression

was used to find the significant predictors of continuous variables. *P* value <0.05 was considered as significant.

4.15 ETHICAL CONSIDERATIONS

All studies were approved by the institutional ethical review board, St. John's Medical College Hospital, Bangalore, India (IERB/1/655/09, IERB/1/725/2010, IERB/1/621/2011(Ref 67/2002), IERB/1/74/08, IERB/1/332/06), PGIMER, Chandigarh (MED/2005/1159), National Institute of Immunology, New Delhi (IBSC/AKS/2007/22), RD Gardi Medical College, Ujjain (M/10/1533), MNP+, Imphal (MOU MNP+ dated 29/12/2010) and Committee for Human Research at University of California, San Francisco, USA (UCSF-IRB#10-03994; Ref#034080). Written informed consent was obtained from all adults and caregivers of children prior to recruitment of the children, and a verbal assent was obtained from children > 8 years of age.

5 RESULTS AND DISCUSSION

The work presented in this thesis is an investigation of the translational genomics of HIV-1 subtypes circulating in India. This is one of the most comprehensive and detailed studies performed so far in the country. The study comprises data from a large population hailing from a wide range of geographical locales which are analyzed by sophisticated bioinformatics methods with the purpose of expanding the knowledge about the HIV-1 epidemiology in India.

Paper I presents a comprehensive HIV-1 molecular epidemiology study pertaining to India. A total of 168 samples collected between 2007 and 2011 from four distant geographical regions of India were used to type the molecular epidemiology. A robust HIV-1 subtyping was done using two or three genes, *gag*, *pol*, and *env* by several methods. As expected, subtype C was found to be the dominant HIV-1 subtype in India. However, for the first time, a high prevalence (10%) of unique recombinant forms (BC and A1C) was observed when two or three genes were used instead of one gene ($p<0.01$; $p=0.02$, respectively). Of the recombinant strains, 6.0% (10/168) were identified as recombinants of subtypes B and C, whereas 4.2% (7/168) were recombinants of A1 and C. The nearest non-C sequence identified for the subtype B segments in the BC recombinant strains was most likely derived from China and Thailand while the A1 segment in A1C originated from eastern Africa (Kenya, Uganda and Tanzania).

Similarly, the rest of the papers (**II to V**) also confirmed the predominance of subtype C with more than 98% frequency. A novel unique recombinant forms CH (URF_CH) was also identified based on analysis of the RT region of the *pol* gene (**III**).

5.1 ORIGIN OF INDIAN HIV-1 SUBTYPE C

The tMRCA of Indian HIV-1 subtype C was estimated using three viral genes (**I**). The *gag* gene analysis estimated the Indian HIV-1 subtype C tMRCA to 1967 [95% CI: 1957-1975] while the *env*-analysis provided an estimated tMRCA in 1974 [95% CI: 1968-1978]. The *pol*-gene analysis was considered to provide the most reliable estimate, 1971 (95% CI: 1965-1976) (Fig. 13).

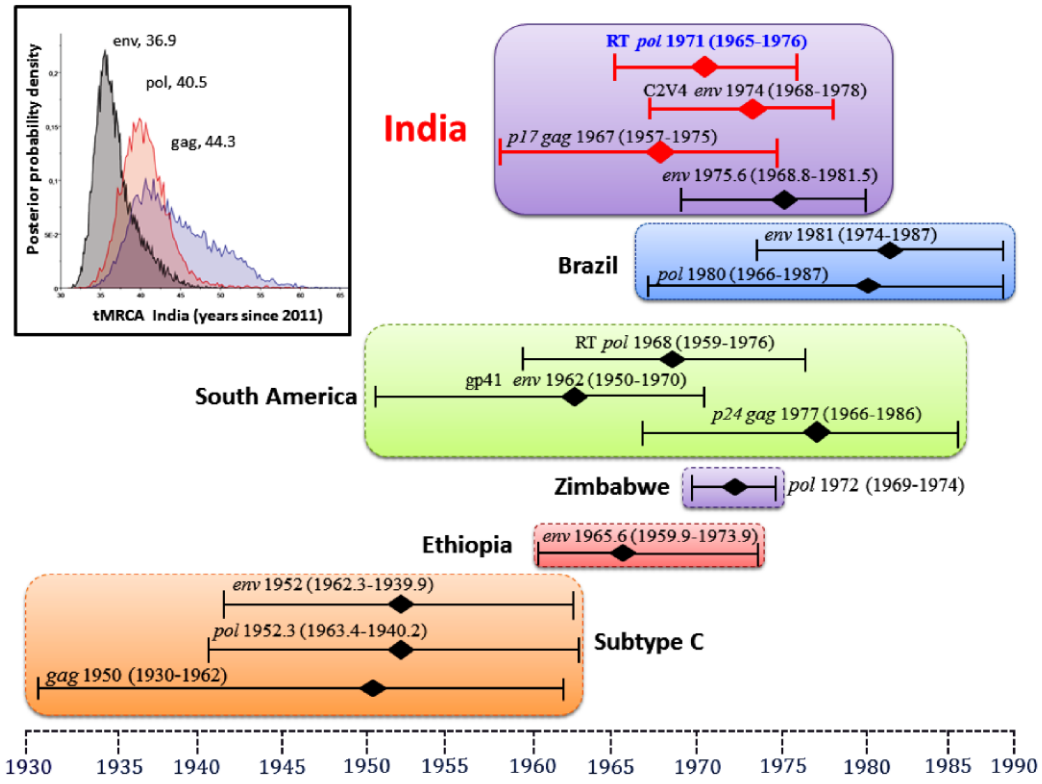

Figure 13. tMRCA of HIV-1 subtype C reported from different regions. The tMRCA of Indian subtype C identified in this study are shown in red. Posterior probability plots showing the tMRCA estimates of Indian HIV-1 subtype C (Inset): The posterior probability density plots for the three genes (*env*, *pol* and *gag*) point to a tMRCA for Indian HIV-1 subtype C from 36.9 (*env*), 40.5 (*pol*) and 44.3 (*gag*) years ago, with an overall mean for the combined graph at 40.6 years ago before 2011, that is in 1970.4 (95% CI: 1960.3–1978.1).

Our study also indicated that the Indian HIV-1 subtype C epidemic originated from a single or few genetically related African lineages, which since then has largely evolved independently as analysis of all the three genes (*gag*, *pol* and *env*) showed a large monophyletic clade of only Indian subtype C strains segregating from the African lineages. However, it was not possible to determine from which country the first introduction took place, as the main Indian clades in the *pol* tree were equidistant to all strains in the neighbouring clade. The *pol* phylogenetic tree also identified a small clade mainly consisting of the strains from western, central and north-eastern regions of the country, indicating a strong movement of strains among these regions. The effective population size in the country has been broadly stable since the 1990s as observed in the Bayesian skyline plot. The data corresponds well with the HIV-1 estimates in India, which indicates a stable or reverse HIV prevalence between 2002 and 2009 [7].

5.2 HUMAN APOBEC3G HYPERMUTATION AND DRUG RESISTANCE

In **Paper II,** we aimed to characterize the nature of APOBEC mediated hypermutations in the RT region of the *pol* gene in population of Indian HIV-1 infected patients, and the consequences in therapeutic management. The evidence from this study suggests that APOBEC3G lethal-hypermutation was associated with the use of treatment in Indian patients who failed ART and had subtype C strains. Among the therapy-experienced patients 22% (7/32) had lethal hypermutation in the proviral compartment but not in the plasma viral compartment. Though drug resistance mutations were present in the hypermutated proviral compartment, all strains had a stop codon in their ORF. Therefore, it is unlikely that such viral variants will mature and expand *in vivo* which is supported by the absence of those mutations in the plasma viral populations.

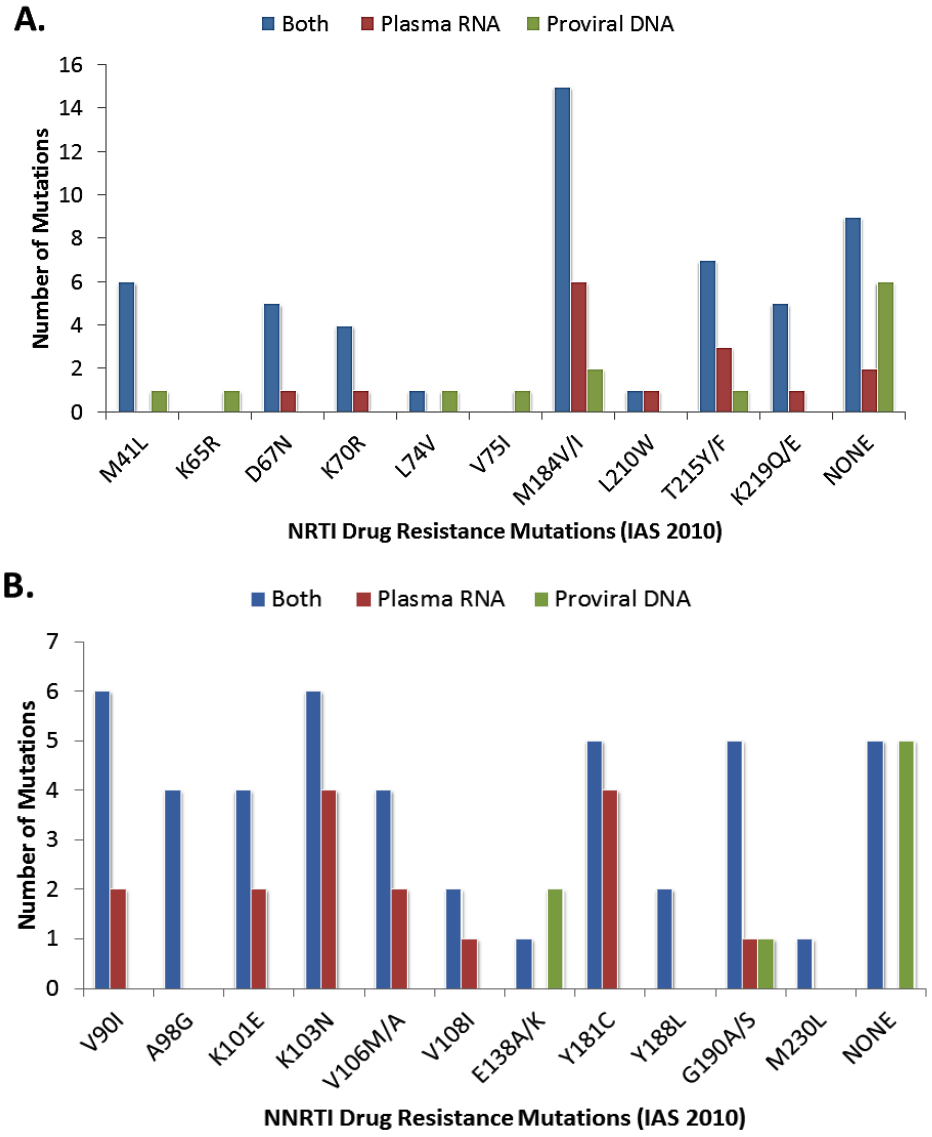

Figure 14. Distribution of drug resistance mutations **A.** NRTI mutations and **B.** NNRTI mutations in plasma viral RNA and proviral genomic DNA

A high level (94.6%; 3/56) of concordance in drug resistance was observed in therapy-naïve individuals but not in experienced individuals (43.8%; 14/32). DRMs were found in 84.4% (27/32) of the patients failing ART. Important differences were found between the plasma RNA and the proviral DNA compartments. Among the DRMs, few patients had M184V (18.75%; 6/32) and T215Y/F (9.4%; 3/32) mutations in plasma only, while M41L and K65R were observed in one patient each in provirus only (Fig. 14). Additional NRTI mutations were observed in 25% (8/32) of each in proviral and plasma viral sequences while additional NNRTI mutations were found mainly in plasma (31.2%; 10/32) and to a lesser extent in proviral DNA (12.5%; 4/32).

5.3 LIMITED TRANSMITTED DRUG RESISTANCE IN INDIA

A low level of transmitted drug resistance was observed in children and adolescents (**III**). Using the SDRM_2009 list [196] for drug resistance survillance, we found that among the 105 ART-naïve children, 3.8% (4/105) had transmitted NRTI drug resistance mutations (L74I, T69D, T215S and K219Q) and 4.8% (5/105) had NNRTI drug resistance mutations [K103N (n=2), K101E (n=2), Y181C (n=1)]. Furthermore, we analyzed 606 RT sequences obtained from treatment-experienced patients from Indian subtype C patients [downloaded from the Los Alamos database (n=102); literature survey (n=446) and unpublished from lab (n=58)], which revealed that T215S was not present in any of the sequences while L74I mutation was present in 0.3% (2/606). Therefore, we excluded the likely polymorphisms L74I and T215S, and the overall prevalence of transmitted drug resistance in this population was found to be 5.7% (6/105). We also used the IAS-USA 2011 list [195] for TDRM interpretation, and a slightly differing result was obtained. The NRTI resistance prevalence was 0.9% (1/105; K219Q) and NNRTI resistance was 5.7% [6/105; K103N (n=2), K101E (n=2), Y181C (n=1), F227C (n=1)]. The likely polymorphism F227C that confers “minor” resistance to rilpivirine, a drug that is still unavailable in India, was however excluded and the resulting resistance was the same as with the first method. Recent studies have also observed differences in prevalence of TDRM within the same cohort with individuals infected with CRF01_AE, when differing tools such as SDRM 2009 and IAS-USA 2009 are used [238]. Therefore, it is of high importance that the interpretation of prevalence of TDRM in population-based surveillance specifically non-B subtypes should be used with caution to avoid reporting falsely elevated TDR prevalence in the population.

Similarly in the adult population, transmitted drug resistance mutations were observed in both compartments in only one patient giving the prevalence of transmitted drug resistance as 1.8% (E138K, 1/56) (**II**).

5.4 HIGH VIREMIA IN SUBTYPE C INFECTED ADOLECENTS

The study presented in **Paper III**, aimed to understand the viral dynamics in perinatal transmission at the population level in therapy-naïve children and adolescents of age between 2 to 16 years. One of the most significant findings was that a substantial proportion (54%; 21/39) of adolescents with subtype C infection presenting with high viremia (median 5.14 $\log_{10}$ copies/mL; IQR 4.69-5.34), despite having a $CD4^+$ T-cell counts >350 cells/$mm^{3.}$ A low level of transmitted drug resistance in children unexposed to antenatal antiretroviral drugs was observed.

The inconsistent clinical parameters of high viral load and high $CD4^+$T-cell count co-existed in over half the adolescents studied. It is plausible that untreated adolescents with high viremia and drug resistance mutations may also play a role in fuelling the HIV epidemic. Our study also supports the idea that early initiation of ART while emphasizing good adherence may be important both from an individual angle to halt disease progression as well as from a public health perspective to attain maximum population-level reduction in HIV-1 transmission.

5.5 CO-RECEPTOR TRANSITION FROM R5- TO X4-TROPISM IN SUBTYPE C INFECTIONS WITH LONGER DURATION

In **Paper IV,** we aimed to study the predicted co-receptor tropism among the perinatally-infected children (n=72) aged between 2 and 17 years. Among children with subtype C strains (n = 71), 81.7% (58/71) were identified as R5-tropic strains with median FPR of 68.1% (range: 10.5–99.2) and 9.9% (7 of 71) were identified as X4-tropic with median FPR of 3.15% (range: 0.10–9.70). The presence of both R5-tropic and X4-tropic sequences were found in 8.5% (6 of 71) (Fig. 15).

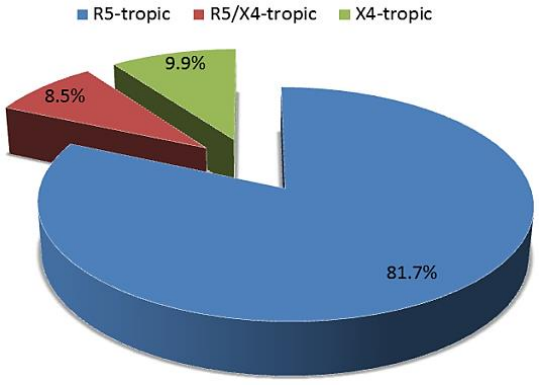

Figure 15. Distribution of R5-, X4- and R5/X4-tropic strains in the pediatric cohort.

Our data for the first time indicated that the likelihood of co-receptor transition from R5 to X4 tropism in subtype C increases with the duration of HIV infection. At univariate analysis the length of HIV-1 infection, $CD4^+$-T cell counts and inter-population genetic divergence of subtype C were the potential factors associated with co-receptor switch from R5- to X4-tropic in perinatally infected patients. However, in multivariate logistic regression, adjusting for other parameters such as gender, ART-naive status, $CD4^+$ T-cell count, log viral load, and viral divergence, we found that the length of infection was the single independent factor associated with co-receptor transition. It was also observed that the therapy-naïve children had 3.3-fold higher risk of having X4-tropic or R5/X4-tropic virus compared to therapy-experienced children although statistical significance were not obtained due to the small sample size of the therapy-experienced group. The evidence from our study also suggests that the co-receptor transitions from R5 to X4-tropic is not as uncommon in HIV-1 subtype C as previously thought [41]. It generally occurs with longer duration of infection and greater disease progression and can have therapeutic implications for the future, specifically for the use of CCR5-antagonists like maraviroc in clinical practice.

5.6 GOOD LONG TERM THERAPEUTIC RESPONSE TO NATIONAL FIRST LINE ANTIRETROVIRAL THERAPY

The **Paper V** is the first study in the country to assess the long term virological outcome among Indian HIV-1 infected patients on Indian national first-line therapy, following initial viral suppression below the detection level (<100 copies/mL). In this longitudinal cohort study, we assessed the factors associated with virological rebound, its consequences, and emergence of drug resistance mutations in a cohort of 323 patients. Our results indicate a good overall long-term response to first-line therapy reflecting an overall success of the standardized Indian national program for individuals with HIV infection. Among the patients, 75.2% (243/323) were able to maintain ≥95% adherence throughout the study and 47.4% (153/323) were able to maintain perfect adherence (100% adherence by VAS and no treatment interruptions lasting more than 48 hrs) during the same time. The median duration of viral suppression was 44 months (IQR 36-54) and 15.8% (51/323) of patients showed intermittent viral rebound without failure during this study period. Viral failure was observed in only 2.8% (9/323) of patients. Among the patients who did not achieve perfect adherence, 24.1% (41/170) experienced viral rebound or failure, which is nearly twice the proportion of perfectly

adherent patients with viral rebound or failure (12.4%; 19/153, p<0.01).Drug resistance mutations were identified in 9 of the 16 patients who had single episode of viral rebound (Fig 16). As expected [205, 213], M184V was the major NRTI drug resistance mutation seen, along with G190A/M, K103N, Y181C and V108I. Among these 16 patients, seven had successfully controlled viremia in their subsequent 6-month follow-up study visit without any change of therapy. The most interesting finding is that, among the seven patients with controlled viremia, three had multiple major NRTI/NNRTI mutations [M184V+G190A; M184V+K103N; M184V+Y181C)

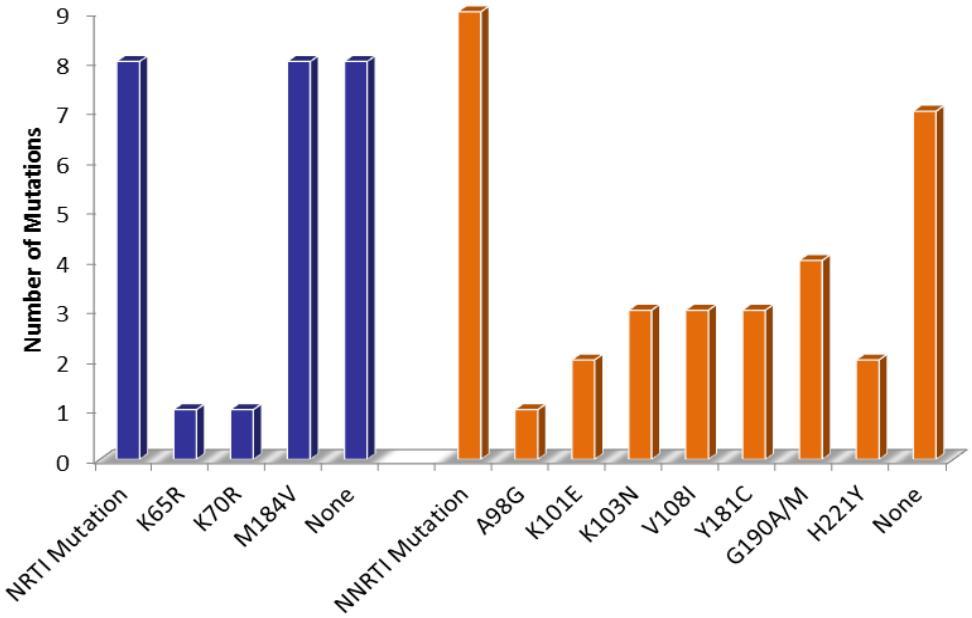

Figure 16. Distribution of drug resistance mutations among the virologically rebound/failure patients during their first viremia >2000 copies/mL.

during their viral rebound, but later during subsequent visits became suppressed without change of therapy. It is also important to note that they had maintained >95% adherence both before and after the viral rebound as well as maintained good immunological status with $CD4^+$ T-cell counts in the normal range. Studies on subtype C virus from South Africa showed that despite the presence of NNTRI mutations such as K103N, V106M and G190A, a fraction of patients were successfully re-suppressed on the same regimen [239, 240]. Studies from Sweden, Belgium and South Africa also showed that despite the presence of NRTI mutations such as M184V, viral re-suppression occurred in a fraction of patients with virological rebound [239-242].

6 CONCLUSIONS

This thesis provides an account of HIV-1 subtypes circulating in India, including the origin, the molecular phylogeny, and characterizations of the predominant HIV-1 subtype C in relation to drug resistance mutations, role of host factor in evolution of drug resistance, viremia, co-receptor switch and the long-term therapeutic response on Indian national first-line antiretroviral therapy. This study described the central importance of molecular epidemiology in translational genomics of HIV-1 from origin to therapeutic response with an aim to improve the available therapeutic and preventive strategies for combating HIV/AIDS. The following conclusions can be drawn:

1. HIV-1 subtype C was confirmed as the dominating strain in India. However, after using comprehensive subtyping methodology, a significant increase in recombinant strains was observed. This indicates an evolving heterogeneous viral epidemic in India. This could pose serious challenges to the development of an effective vaccine that would be applicable in the whole country (**I**).
2. The Indian HIV-1 subtype C epidemic originated nearly four decades ago from a single or few genetically related African lineages, and since then largely evolved independently and formed a separate monophyletic clade segregating from the larger subtype C clade (**I**).
3. The molecular data using Bayesian Skyline plot showed a stable epidemic of HIV-1 subtype C which corresponded with the National AIDS Control Program (NACP) launched by Govt. of India in the 1990s (**I**).
4. Our study (**II**) identified that APOBEC3G mediated lethal-hypermutations were associated with treatment failure among Indian patients. Though drug resistance mutations like M184I and/or M230I, were present in the hypermutated proviral compartment, all strains had a stop codon in its open reading frame; therefore, it is unlikely that the drug resistance-containing progeny will expand *in vivo*.
5. An implication of this study (**II**) is the possibility of using proviral DNA as an alternative to plasma viral RNA for surveillance of transmitted drug resistance. As the proviral genotyping technique utilizes whole blood, it does not require cold chain transport, which facilitates studies in resource-limited settings, especially from rural and remote part of the country where logistic challenges remain.

6. The presence of key mutations like K65R and M41L in the proviral DNA compartment in therapy-failure patients may be of importance for the selection of future ART regimen (**II**).
7. The low level of transmitted drug resistance in children and adults is reassuring. Our study also suggests that interpretation of TDRM prevalence in population-based surveillance with non-B subtypes should be used with caution. Inclusion of likely polymorphisms might falsely elevate the TDR prevalence in the population (**III**).
8. More than half the adolescents living with HIV-1subtype C included in the study had inconsistent clinical parameters of high viral load and high $CD4^+$ T-cell counts. Therefore, early initiation of ART while emphasizing good adherence may be considered in this population in order to attain maximum population-level reductions in HIV-1 transmission (**III**).
9. The likelihood of the high proportion of predicted X4-tropic strains in the older children advocate that co-receptor transition can occur with longer duration of infection and greater disease progression which might have therapeutic implications for the use of CCR5-antagonist Maraviroc in future (**IV**).
10. Our study showed a good overall response with a median of four years, to a regimen containing two NRTI and one NNRTI drugs (as in the Indian national first-line antiretroviral therapy regimen), indicating the overall success of the Indian ART program. Less than perfect adherence i.e. 100% adherence and no treatment interruption for more than 48 hours increases the risk for treatment failure and subsequent drug resistance development (**V**).
11. A subset of patients with major drug resistance mutations like M184V, K103N and Y181C had suppressed viremia without change in therapy as they had optimal adherence profile throughout the study period. Our study therefore provides support for the conceptual evidence that viral load measurement can act as an indicator for adherence. Early identification of viremia followed by targeted special adherence counseling to these patients may limit the unnecessary switch to the second-line of therapy, especially in resource-limited settings, where limited drug options are available (**V**).

7 RECOMMENDATIONS FOR FUTURE STUDIES

This research has thrown up many questions which invite further investigation. A further comprehensive biological characterization of the Indian HIV-1 subtype C needs to be performed to understand this specific viral strain. The following recommendations can be made for further studies:

1. India initiated its vaccine development program as early as 1998. However, the depth of sampling (the proportion of available sequences compared to total number of estimated infections) is still very low (~0.03%). The resulting diversified epidemiology indicated by an increase in recombinant strains, may pose a serious challenge to the development of an effective vaccine that would be applicable in the country. Ongoing countrywide molecular surveillance of HIV-1 subtypes is likely to contribute towards a better understanding of the epidemiology in this region and to guide efficient disease intervention strategies.
2. The role of APOBEC3G in the evolution of drug resistance needs to be validated by further studies of the clinical consequences through analysis of a large number of patients and by appropriate *in vitro* cell culture models to gain the mechanistic view of the APOBEC-mediated drug resistance mutations.
3. High viremia in non-immunosuppressed adolescents may also play a role in fuelling the HIV epidemic. The question remains, will this group of adolescents with extended high viremia experience imminent rapid disease progression as has been seen in adult studies? What will be the therapeutic response among such patients with high viremia? Therefore, longitudinal analyses to understand the disease progression, therapeutic response and immune re-establishment need to be investigated.
4. Though the current scientific literature supports good correlation between population sequencing and interpretation of viral tropism using the Geno2Pheno system with phenotypic tests, *in vitro* studies using an appropriate cell culture model needs to be performed to validate the findings. Also, the role of minor viral populations in the evolution of co-receptor transition needs to be investigated using more sensitive techniques like next generation sequencing.
5. The low level of transmitted drug resistance in children and adults is reassuring but should not advocate complacency. As there is an increase in the therapeutic

coverage through free national ART program, regular surveillance is necessary to understand the evolution of transmitted drug resistance mutations in the population.

6. An appropriate drug resistance mutation algorithm for subtype C needs to be established for surveillance of transmitted drug resistance. Also future research should investigate the role of drug resistance and treatment outcome with HIV-1 subtype C viruses to identify the optimal time of switching failing patients to appropriate second-line regimens.
7. Finally, as Indian subtype C showed distinct genotypic characteristics compared to South African subtype C, considerably more work are to be done to determine the phenotypic and biological characteristics of these strains, in order to guide efficient disease intervention and management strategies.

8 ACKNOWLEDGEMENTS

Working on a Ph. D. thesis is a colossal task. Although the final result is officially credited to the author, it is the collective outcome of efforts of everybody who is involved in the process right from the commencement of the idea. Therefore in the following paragraphs I take the opportunity to acknowledge and appreciate those who directly or indirectly rendered their support to accomplish the task in time.

I am immensely indebted to my thesis supervisor, Prof. Anders Sonnerborg for his support and also for granting me the liberty and freedom to follow my passion throughout the thesis work. Thank you for all the scholarly guidance and expert advice. You have stimulated me to become an independent researcher and assisted me in realizing the power of critical thinking.

I would also like to extend my sincere gratitude to my co-supervisors; Dr. Irene Bontell, Dr. Ayesha De Costa, Prof. Udaykumar Ranga as well as my external mentor Prof. Akhil C Banerjea. The insightful discussions with them allowed me to understand the individual fields of basic, translational and clinical science and amalgamate them with public health. Subjecting each manuscript with their meticulous analysis and their distinct perspectives to improve upon the thesis has helped my scientific research to appeal to a wide range of audience.

This thesis would not have been possible without the help and unconditional support from Dr. Anita Shet, her consistent encouragement and positive disposition has been instrumental to the successful completion of my doctoral programme. I appreciate and acknowledge her efforts as my mentor, advisor, as well as the project leader for her concern and good wishes. Scholarly discussions and critical evaluation at each stage of my doctoral journey has helped me in improving the quality of my research. Her unwavering support and much-needed guidance during the rough phases in this journey assisted me to infuse positive strength and work with fresh enthusiasm.

The path to my doctoral degree started with a collaboration with the Division of Global Health, Karolinska Institutet. I owe a lot to Prof. Vinod Diwan and express my heartfelt gratitude for having planted the seeds of

earnestness and passion in my pursuit of knowledge, with his brief yet inspiring dialogues.

I am thankful to my fellow labmates at St. John's Medical College, Bangalore; Mr. Soham Gupta, Ms. Pravat Nalini Sahoo and Ms. Shwetha D Rao for extending their support in setting up the laboratory. The stimulating discussions with them have helped in the development of the scientific approach and attitude to reach the highest rung of my potential. A special thanks to Ms. Riya Palchaudhuri for language editing of the thesis thoroughly. I would also like to thank my colleagues at Karolinska Institute, Dr. Piotr Nowak, Dr. Samir Abdurhaman, and Dr. Babilonia Barqasho for the amicable and heartening discussions. I am also thankful to Ms. Karthika Arumugam at St. John's Research Institute and Ms. Elsa Heylen at University of California, San Francisco to help me with statistics.

It is my pleasure to acknowledge the encouraging support from my collaborators; Dr. Ajay Wanchu, Post Graduate Institute of Medical Research, Chandigarh, Dr. Vishal Diwan, R.D. Gardi Medical College, Ujjain, Dr. Ranbir Singh Laishram, Regional Institute of Medical Science, Imphal and Dr. Maria L. Ekstrand, University of California, San Francisco. My gratitude goes to them for their valuable contributions during the course of this journey.

I am also thankful to Prof. Anna Norrby-Teglund, Director of Doctoral Studies, Department of Medicine Huddinge, for all the help and support that she extended whenever I approached her with querry. I would also like to extend my sincerest thanks to Gun-Britt Eriksson of Global Health Science, not solely for her prompt administrative support but also for the kind care and concern.

It is vital to have a buoyant and pleasant social environment outside the research group. In this regard, I acknowledge my old pals and other Indian PhD students and not to forget my "Jägargatan" friends; Sandeep, Pradeep, Shahul, Uriel and Krushna. My special thanks to Mr. Prasenjit Pal to take time out and draw the cover photo which has helped in giving a very personal touch to the thesis.

I expand my thanks to Dr. Ananta Paine, Karolinska Institutet, Prof. Pintoo Bandopadhyay and Mr. Dipak Biswas of Bidhan Chandra Krishi Viswavidyalaya, West Bengal for taking time to improve upon my thesis. Their critical review and valuable suggestions has been greatly useful.

I can't imagine my current position without the love and support from my family. Above all, a special thanks to my mother, she is always excited about my achievements and that inspires me to keep going and aim higher. The sacrifices that my mother made for my education is unimaginable. Thank you 'Maa' for all your support and for believing in me.

It is my privilege to acknowledge the support of a special person in my life, whom I have met half way through my PhD journey; she is my wife, Rakhi. She has been my best friend and has supported me during very difficult times during the course of my doctoral programme. In the most positive way with her love, support, affection and caring, she has helped me to keep the spirit of science alive within me.

And now, last but not least, I would like to extend my sincere gratitude for the funding support provided by European Union Frame Work Program 7, AIDS International Training and Research Program Grant supported by National Institute of Health, US, Swedish physicians against AIDS fund, Sven Grad fund for virology research and Erasmus Mundus to make this study possible.

Special Acknowledgements

I gratefully acknowledge the generosity of the study participants who have so kindly contributed their time to help us better understand issues in HIV-1 molecular epidemiology, ART adherence and drug resistance in this setting. A special thanks to National AIDS Control Organization (NACO), Karnataka State AIDS Prevention Society (KASAPS) and the Manipur Network of Positive People (MNP+) for their big-hearted support and excellent patient care.

Ujjwal Neogi

9 REFERENCES

1. Fauci AS. HIV and AIDS: 20 years of science. *Nat Med* 2003,**9**:839-843.
2. Barresinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, *et al.* Isolation of a T-Lymphotropic retrovirus from a patient at risk for acquired immune-deficiency syndrome (AIDS). *Science* 1983,**220**:868-871.
3. Gallo RC. Historical essay - The early years of HIV/AIDS. *Science* 2002,**298**:1728-1730.
4. Montagnier L. Historical essay - A history of HIV discovery. *Science* 2002,**298**:1727-1728.
5. UNAIDS Report on the global AIDS epidemic-2012 http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2012/gr2012/20121120_UNAIDS_Global_Report_2012_en.pdf accessed on January 2013.
6. Lihana RW, Ssemwanga D, Abimiku A, Ndembi N. Update on HIV-1 Diversity in Africa: A Decade in Review. *AIDS Rev* 2012,**14**:83-100.
7. National AIDS Control Organisations: Annual Report 2010-11. Available at http://www.nacoonline.org/upload/REPORTS/NACO%20Annual%20Report%202010-11.pdf. Accessed on January 2013.
8. Simoes EAF, 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**:335-338.
9. Sharp PM, Robertson DL, Hahn BH. Cross-species transmission and recombination of AIDS viruses. *Philos Trans R Soc Lond B Biol Sci* 1995,**349**:41-47.
10. Hahn BH, Shaw GM, De Cock KM, Sharp PM. AIDS as a zoonosis: Scientific and public health implications. *Science* 2000,**287**:607-614.
11. Gao F, Bailes E, Robertson DL, Chen YL, Rodenburg CM, Michael SF, *et al.* Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* 1999,**397**:436-441.
12. Keele BF, Van Heuverswyn F, Li YY, Bailes E, Takehisa J, Santiago ML, *et al.* Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* 2006,**313**:523-526.
13. Gao F, Yue L, White AT, Pappas PG, Barchue J, Hanson AP, *et al.* Human infection by genetically diverse SIVsm-related HIV-2 in west Africa. *Nature* 1992,**358**:495-499.
14. Chen ZW, Telfer P, Gettie A, Reed P, Zhang LQ, Ho DD, *et al.* Genetic characterization of new west African simian immunodeficiency virus SIVsm: Geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey troop. *J Virol* 1996,**70**:3617-3627.
15. Tebit DM, Arts EJ. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. *Lancet Infect Dis* 2011,**11**:45-56.
16. Mansky LM. Retrovirus mutation rates and their role in genetic variation. *J Gen Virol* 1998,**79 (Pt 6)**:1337-1345.
17. Korber B, Gaschen B, Yusim K, Thakallapally R, Kesmir C, Detours V. Evolutionary and immunological implications of contemporary HIV-1 variation. *Br Med Bull* 2001,**58**:19-42.
18. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, *et al.* HIV-1 nomenclature proposal. *Science* 2000,**288**:55-57.

19. Rambaut A, Posada D, Crandall KA, Holmes EC. The causes and consequences of HIV evolution. *Nat Rev Genet* 2004,**5**:52-61.
20. Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemee V, *et al.* A new human immunodeficiency virus derived from gorillas. *Nat Med* 2009,**15**:871-872.
21. Vallari A, Holzmayer V, Harris B, Yamaguchi J, Ngansop C, Makamche F, *et al.* Confirmation of Putative HIV-1 Group P in Cameroon. *J Virol* 2011,**85**:1403-1407.
22. Korber B, Muldoon M, Theiler J, Gao F, Gupta R, Lapedes A, *et al.* Timing the ancestor of the HIV-1 pandemic strains. *Science* 2000,**288**:1789-1796.
23. Zhu TF, Korber BT, Nahmias AJ, Hooper E, Sharp PM, Ho DD. An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. *Nature* 1998,**391**:594-597.
24. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, *et al.* Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* 2008,**455**:661-664.
25. Lemey P, Pybus OG, Rambaut A, Drummond AJ, Robertson DL, Roques P, *et al.* The molecular population genetics of HIV-1 group O. *Genetics* 2004,**167**:1059-1068.
26. Wertheim JO, Worobey M. Dating the Age of the SIV Lineages That Gave Rise to HIV-1 and HIV-2. *Plos Comput Biol* 2009,**5**.
27. Hemelaar J, Gouws E, Ghys PD, Osmanov S, C W-UNHI. Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS* 2011,**25**:679-689.
28. Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* 2006,**20**:W13-23.
29. Taylor BS, Sobieszczyk ME, McCutchan FE, Hammer SM. The challenge of HIV-1 subtype diversity. *N Engl J Med* 2008,**358**:1590-1602.
30. Ayehunie S, Johansson B, Sonnerborg A, Salminen M, Zewdie DW, Fehniger T, *et al.* New subtype of HIV-1 in Ethiopia. *Lancet* 1990,**336**:942.
31. Salminen MO, Johansson B, Sonnerborg A, Ayehunie S, Gotte D, Leinikki P, *et al.* Full-length sequence of an ethiopian human immunodeficiency virus type 1 (HIV-1) isolate of genetic subtype C. *AIDS Res Hum Retroviruses* 1996,**12**:1329-1339.
32. 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**:1293-1303.
33. Bachu M, Yalla S, Asokan M, Verma A, Neogi U, Sharma S, *et al.* Multiple NF-kappaB sites in HIV-1 subtype C long terminal repeat confer superior magnitude of transcription and thereby the enhanced viral predominance. *J Biol Chem* 2012,**287**:44714-44735.
34. Rao VR, Sas AR, Eugenin EA, Siddappa NB, Bimonte-Nelson H, Berman JW, *et al.* HIV-1 clade-specific differences in the induction of neuropathogenesis. *J Neurosci* 2008,**28**:10010-10016.
35. Ranga U, Shankarappa R, Siddappa NB, Ramakrishna L, Nagendran R, Mahalingam M, *et al.* Tat protein of human immunodeficiency virus type 1 subtype C strains is a defective chemokine. *J Virol* 2004,**78**:2586-2590.
36. Rossenkhan R, Macleod IJ, Sebunya TK, Castro-Nallar E, McLane MF, Musonda R, *et al.* Tat exon 1 Exhibits Functional Diversity During HIV-1 Subtype C Primary Infection. *J Virol* 2013, **87**:5732-5745.
37. Neogi U, Gupta S, Sahoo PN, Shet A, Rao SD, Ranga U, *et al.* Genetic characterization of HIV type 1 Tat exon 1 from a southern Indian clinical

cohort: Identification of unique epidemiological signature residues. *AIDS Res Hum Retroviruses* 2012,**28**:952-956.
38. Neogi U, Prarthana BS, Gupta S, D'Souza G, De Costa A, Kuttiatt VS, *et al.* Naturally occurring polymorphisms and primary drug resistance profile among antiretroviral-naive individuals in Bangalore, India. *AIDS Res Hum Retroviruses* 2010,**26**:1097-1101.
39. Neogi U, Sahoo PN, Kumar R, De Costa A, Shet A. Characterization of HIV type 1 subtype C protease gene: selection of L63P mutation in protease inhibitor-naive Indian patients. *AIDS Res Hum Retroviruses* 2011,**27**:1249-1253.
40. Connor RI, Sheridan KE, Ceradini D, Choe S, Landau NR. Change in coreceptor use correlates with disease progression in HIV-1--infected individuals. *J Exp Med* 1997,**185**:621-628.
41. 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**:253-258.
42. Novitsky V, Ndung'u T, Wang R, Bussmann H, Chonco F, Makhema J, *et al.* Extended high viremics: a substantial fraction of individuals maintain high plasma viral RNA levels after acute HIV-1 subtype C infection. *AIDS* 2011,**25**:1515-1522.
43. Campbell MS, Kahle EM, Celum C, Lingappa JR, Kapiga S, Mujugira A, *et al.* Plasma Viral Loads During Early HIV-1 Infection Are Similar in Subtype C- and Non-Subtype C-Infected African Seroconverters. *J Infect Dis* 2013,**207**:1166-1170.
44. Neogi U, Palchaudhuri R, Shet, A. High viremia in HIV-1 subtype C and spread of the epidemic *J Infect Dis* 2013,**In Press**.
45. Turner BG, Summers MF. Structural biology of HIV. *J Mol Biol* 1999,**285**:1-32.
46. Burton DR. Structural biology - Images from the surface of HIV. *Nature* 2006,**441**:817-818.
47. Ganser-Pornillos BK, Yeager M, Sundquist WI. The structural biology of HIV assembly. *Curr Opin Struct Biol* 2008,**18**:203-217.
48. Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW, Swanstrom R, *et al.* Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* 2009,**460**:711-U787.
49. Swanson CM, Malim MH. SnapShot: HIV-1 proteins. *Cell* 2008,**133**.
50. Levy JA. Pathogenesis of human immunodeficiency virus infection. *Microbiol Rev* 1993,**57**:183-289.
51. Stevenson M. HIV-1 pathogenesis. *Nat Med* 2003,**9**:853-860.
52. Forsman A, Weiss RA. Why is HIV a pathogen? *Trends Microbiol* 2008,**16**:555-560.
53. Hansasuta P, Rowland-Jones SL. HIV-1 transmission and acute HIV-1 infection. *Br Med Bull* 2001,**58**:109-127.
54. Hogan CM, Hammer SM. Host determinants in HIV infection and disease. Part 2: genetic factors and implications for antiretroviral therapeutics. *Ann Intern Med* 2001,**134**:978-996.
55. Mikhail M, Wang B, Saksena NK. Mechanisms involved in non-progressive HIV disease. *AIDS Rev* 2003,**5**:230-244.
56. Poropatich K, Sullivan DJ, Jr. Human immunodeficiency virus type 1 long-term non-progressors: the viral, genetic and immunological basis for disease non-progression. *J Gen Virol* 2011,**92**:247-268.

57. Brostrom C, Visco-Comandini U, Yun Z, Sonnerborg A. Longitudinal quantification of human immunodeficiency virus type 1 DNA and RNA in long-term nonprogressors. *J Infect Dis* 1999,**179**:1542-1548.
58. Rachinger A, Navis M, van Assen S, Groeneveld PH, Schuitemaker H. Recovery of viremic control after superinfection with pathogenic HIV type 1 in a long-term elite controller of HIV type 1 infection. *Clin Infect Dis* 2008,**47**:e86-89.
59. Lambotte O, Boufassa F, Madec Y, Nguyen A, Goujard C, Meyer L, *et al.* HIV controllers: a homogeneous group of HIV-1-infected patients with spontaneous control of viral replication. *Clin Infect Dis* 2005,**41**:1053-1056.
60. Lindback S, Thorstensson R, Karlsson AC, von Sydow M, Flamholc L, Blaxhult A, *et al.* Diagnosis of primary HIV-1 infection and duration of follow-up after HIV exposure. Karolinska Institute Primary HIV Infection Study Group. *AIDS* 2000,**14**:2333-2339.
61. Mellors JW, Kingsley LA, Rinaldo CR, Jr., Todd JA, Hoo BS, Kokka RP, *et al.* Quantitation of HIV-1 RNA in plasma predicts outcome after seroconversion. *Ann Intern Med* 1995,**122**:573-579.
62. An P, Winkler CA. Host genes associated with HIV/AIDS: advances in gene discovery. *Trends Genet* 2010,**26**:119-131.
63. Cohen MS, Gay CL. Treatment to prevent transmission of HIV-1. *Clin Infect Dis* 2010,**50 Suppl 3**:S85-95.
64. Kandathil AJ, Ramalingam S, Kannangai R, David S, Sridharan G. Molecular epidemiology of HIV. *Indian J Med Res* 2005,**121**:333-344.
65. Lakhashe S, Thakar M, Godbole S, Tripathy S, Paranjape R. HIV infection in India: epidemiology, molecular epidemiology and pathogenesis. *J Biosci* 2008,**33**:515-525.
66. Neogi U, Sood V, Banerjee S, Ghosh N, Verma S, Samrat S, *et al.* Global HIV-1 molecular epidemiology with special reference to genetic analysis of HIV-1 subtypes circulating in North India: functional and pathogenic implications of genetic variation. *Indian J Exp Biol* 2009,**47**:424-431.
67. Neogi U, Sood V, Ronsard L, Singh J, Lata S, Ramachandran VG, *et al.* Genetic architecture of HIV-1 genes circulating in north India & their functional implications. *Indian J Med Res* 2011,**134**:769-778.
68. Tripathy SP, Kulkarni SS, Jadhav SD, Agnihotri KD, Jere AJ, Kurle SN, *et al.* Subtype B and subtype C HIV type 1 recombinants in the northeastern state of Manipur, India. *AIDS Res Hum Retroviruses* 2005,**21**:152-157.
69. Neogi U, Sood V, Chowdhury A, Das S, Ramachandran VG, Sreedhar VK, *et al.* Genetic analysis of HIV-1 Circulating Recombinant Form 02_AG, B and C subtype-specific envelope sequences from Northern India and their predicted co-receptor usage. *AIDS Res Ther* 2009,**6**:28.
70. Sarkar R, Pal R, Bal B, Mullick R, Sengupta S, Sarkar K, *et al.* Genetic Characterization of HIV-1 Strains Among the Injecting Drug Users in Nagaland, India. *Open Virol J* 2011,**5**:96-102.
71. Ragupathy V, Casado C, Jacob SM, Samuel NM, Lopez-Galindez C. Circulation of HIV-1 subtype A within the subtype C HIV-1 epidemic in Tamil Nadu, India. *J Med Virol* 2012,**84**:1507-1513.
72. Sarkar R, Sarkar K, Brajachand Singh N, Manihar Singh Y, Mitra D, Chakrabarti S. Emergence of a unique recombinant form of HIV-1 from Manipur (India). *J Clin Virol* 2012, **55**:274-277.
73. Sarkar R, Sarkar K, Singh NB, Singh YM, Chakrabarti S. Near full-length genomic characterization of a HIV type 1 BC recombinant strain from Manipur, India. *Virus Genes* 2012, **45**:201-206.

74. Mullick R, Sengupta S, Sarkar K, Saha MK, Chakrabarti S. Phylogenetic analysis of env, gag, and tat genes of HIV type 1 detected among the injecting drug users in West Bengal, India. *AIDS Res Hum Retroviruses* 2006,**22**:1293-1299.
75. Neogi U, Sood V, Goel N, Wanchu A, Banerjea AC. Novel HIV-1 long terminal repeat (LTR) sequences of subtype B and mosaic intersubtype B/C recombinants in North India. *Arch Virol* 2008,**153**:1961-1966.
76. Pfutzner A, Dietrich U, von Eichel U, von Briesen H, Brede HD, Maniar JK, *et al.* HIV-1 and HIV-2 infections in a high-risk population in Bombay, India: evidence for the spread of HIV-2 and presence of a divergent HIV-1 subtype. *J Acquir Immune Defic Syndr* 1992,**5**:972-977.
77. 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**:2161-2168.
78. 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**:299-307.
79. 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**:1199-1202.
80. 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**:941-944.
81. Halani N, Wang B, Ge YC, Gharpure H, Hira S, Saksena NK. Changing epidemiology of HIV type 1 infections in India: evidence of subtype B introduction in Bombay from a common source. *AIDS Res Hum Retroviruses* 2001,**17**:637-642.
82. Kalpana A, Srikanth T, Abhay J, Sushama J, Swarali K, Ramesh P. gp120 sequences from HIV type 1 subtype C early seroconverters in India. *AIDS Res Hum Retroviruses* 2004,**20**:889-894.
83. Kurle S, Tripathy S, Jadhav S, Agnihotri K, Paranjape R. Full-length gag sequences of HIV type 1 subtype C recent seroconverters from Pune, India. *AIDS Res Hum Retroviruses* 2004,**20**:1113-1118.
84. Gupta RM, Prasad VV, Rai A, Seth P. Analysis of HIV type 1 subtype C full-length gag gene sequences from India: novel observations and plausible implications. *AIDS Res Hum Retroviruses* 2005,**21**:1066-1072.
85. Sengupta S, Jana S, Sarkar K, Bhattacharya SK, Chakrabarti S. Determination of gag subtypes of HIV type 1 detected among female sex workers in Calcutta, India. *AIDS Res Hum Retroviruses* 2005,**21**:806-809.
86. Agnihotri KD, Tripathy SP, Jere AP, Kale SM, Paranjape RS. Molecular analysis of gp41 sequences of HIV type 1 subtype C from India. *J Acquir Immune Defic Syndr* 2006,**41**:345-351.
87. Kumar M, Jain SK, Pasha ST, Chattopadhaya D, Lal S, Rai A. Genomic diversity in the regulatory nef gene sequences in Indian isolates of HIV type 1: emergence of a distinct subclade and predicted implications. *AIDS Res Hum Retroviruses* 2006,**22**:1206-1219.
88. Khan IF, Vajpayee M, Prasad VV, Seth P. Genetic diversity of HIV type 1 subtype C env gene sequences from India. *AIDS Res Hum Retroviruses* 2007,**23**:934-940.
89. Deshpande A, Jauvin V, Magnin N, Pinson P, Faure M, Masquelier B, *et al.* Resistance mutations in subtype C HIV type 1 isolates from Indian patients of

Mumbai receiving NRTIs plus NNRTIs and experiencing a treatment failure: resistance to AR. *AIDS Res Hum Retroviruses* 2007,**23**:335-340.

90. Sen S, Tripathy SP, Chimanpure VM, Patil AA, Bagul RD, Paranjape RS. Human immunodeficiency virus type 1 drug resistance mutations in peripheral blood mononuclear cell proviral DNA among antiretroviral treatment-naive and treatment-experienced patients from Pune, India. *AIDS Res Hum Retroviruses* 2007,**23**:489-497.
91. Soundararajan L, Karunaianandham R, Jauvin V, Schrive MH, Ramachandran R, Narayanan PR, *et al.* Characterization of HIV-1 isolates from antiretroviral drug-naive children in southern India. *AIDS Res Hum Retroviruses* 2007,**23**:1119-1126.
92. Arora SK, Gupta S, Toor JS, Singla A. Drug resistance-associated genotypic alterations in the pol gene of HIV type 1 isolates in ART-naive individuals in North India. *AIDS Res Hum Retroviruses* 2008,**24**:125-130.
93. Kandathil AJ, Kannangai R, Abraham OC, Sudarsanam TD, Pulimood SA, Sridharan G. Genotypic resistance profile of HIV-1 protease gene: a preliminary report from Vellore, south India. *Indian J Med Microbiol* 2008,**26**:151-154.
94. Lall M, Gupta RM, Sen S, Kapila K, Tripathy SP, Paranjape RS. Profile of primary resistance in HIV-1-infected treatment-naive individuals from Western India. *AIDS Res Hum Retroviruses* 2008,**24**:987-990.
95. Iqbal HS, Solomon SS, Madhavan V, Solomon S, Balakrishnan P. Primary HIV-1 drug resistance and polymorphic patterns among injecting drug users (IDUs) in Chennai, Southern India. *J Int Assoc Physicians AIDS Care (Chic)* 2009,**8**:323-327.
96. Kandathil AJ, Kannangai R, Abraham OC, Rupali P, Pulimood SA, Verghese VP, *et al.* The frequency of HIV-I drug resistance mutations among treatment-naive individuals at a tertiary care centre in south India. *Int J STD AIDS* 2009,**20**:522-526.
97. Deshpande A, Jeannot AC, Schrive MH, Wittkop L, Pinson P, Fleury HJ. Analysis of RT sequences of subtype C HIV-type 1 isolates from indian patients at failure of a first-line treatment according to clinical and/or immunological WHO guidelines. *AIDS Res Hum Retroviruses* 2010,**26**:343-350.
98. Dutta Choudhury S, Chaudhury AK, Kalra R, Andrabi R, Wig N, Biswas A, *et al.* Antiretroviral drug resistance mutations in the reverse transcriptase gene of HIV-1 isolates from Northern Indian patients: a follow-up study. *Arch Virol* 2010,**155**:563-569.
99. Mullick R, Sengupta S, Sarkar K, Chakrabarti S. Molecular characterization of tat gene and long terminal repeat region of human immunodeficiency virus type-1 detected among the injecting drug users (IDUs) of Manipur, India: identification of BC recombinants. *Virus Res* 2010,**147**:195-201.
100. Deshpande A, Karki S, Recordon-Pinson P, Fleury HJ. Drug resistance mutations in HIV type 1 isolates from naive patients eligible for first line antiretroviral therapy in JJ Hospital, Mumbai, India. *AIDS Res Hum Retroviruses* 2011,**27**:1345-1347.
101. Thorat SR, Chaturbhuj DN, Hingankar NK, Chandrasekhar V, Koppada R, Datkar SR, *et al.* Surveillance of transmitted HIV type 1 drug resistance among HIV type 1-positive women attending an antenatal clinic in Kakinada, India. *AIDS Res Hum Retroviruses* 2011,**27**:1291-1297.

102. Saini S, Bhalla P, Gautam H, Baveja UK, Pasha ST, Dewan R. Resistance-Associated Mutations in HIV-1 among Patients Failing First-Line Antiretroviral Therapy. *J Int Assoc Physicians AIDS Care (Chic)* 2012,**11**:203-209.
103. Sinha S, Ahmad H, Shekhar RC, Kumar N, Dar L, Samantaray JC, *et al.* Prevalence of HIV Drug Resistance Mutations in HIV Type 1 Isolates in Antiretroviral Therapy Naive Population from Northern India. *AIDS Res Treat* 2012,**2012**:905823.
104. Siddappa NB, Dash PK, Mahadevan A, Desai A, Jayasuryan N, Ravi V, *et al.* Identification of unique B/C recombinant strains of HIV-1 in the southern state of Karnataka, India. *AIDS* 2005,**19**:1426-1429.
105. Chaturbhuj DN, Hingankar NK, Srikantiah P, Garg R, Kabra S, Deshmukh PS, *et al.* Transmitted HIV drug resistance among HIV-infected voluntary counseling and testing centers (VCTC) clients in Mumbai, India. *AIDS Res Hum Retroviruses* 2010,**26**:927-932.
106. 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**:685-690.
107. 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**:1435-1441.
108. Ahmad KM, Mujtaba S, Das R, Zafrullah M, Sehgal S, Jameel S. nef sequences of primary HIV type 1 isolates from northern India. *AIDS Res Hum Retroviruses* 1998,**14**:1491-1493.
109. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, *et al.* Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 1999,**73**:152-160.
110. 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**:1219-1227.
111. Deshpande A, Recordon-Pinson P, Deshmukh R, Faure M, Jauvin V, Garrigue I, *et al.* Molecular characterization of HIV type 1 isolates from untreated patients of Mumbai (Bombay), India, and detection of rare resistance mutations. *AIDS Res Hum Retroviruses* 2004,**20**:1032-1035.
112. Bhanja P, Sengupta S, Singh NY, Sarkar K, Bhattacharya SK, Chakrabarti S. Determination of gag and env subtypes of HIV-1 detected among injecting drug users (IDUs) in Manipur, India: evidence for intersubtype recombination. *Virus Res* 2005,**114**:149-153.
113. Sengupta S, Jana S, Roy P, Sarkar K, Bhattacharya SK, Chakrabarti S. Phylogenetic analysis of the p24-p7 region of the human immunodeficiency virus type 1 gag gene to determine subtype distribution among female sex workers in Calcutta, India. *J Clin Microbiol* 2005,**43**:5787-5791.
114. Sengupta S, Khetawat D, Jana S, Sarkar K, Bhattacharya SK, Chakrabarti S. Polymorphism of HIV-1 gag (p17) gene from female sex workers in Calcutta, India. *Arch Virol* 2005,**150**:2117-2124.
115. Bhanja P, Sengupta S, Banerjee D, Sarkar K, Jana S, Chakrabarti S. Detection of intersubtype recombinants with respect to env and nef genes of HIV-1 among female sex workers in Calcutta, India. *Virus Res* 2007,**130**:310-314.
116. Lakhashe S, Tripathy S, Paranjape R, Bhattacharya J. Evidence of a novel B/C recombinant exhibiting unique breakpoints of near full-length HIV type 1

genome from Northeastern India. *AIDS Res Hum Retroviruses* 2008,**24**:229-234.

117. Lakhashe S, Tripathy S, Paranjape R, Bhattacharya J. Characterization of B/C recombinants of near full-length HIV type 1 from northeastern India with mosaics identical to ARE195FL but with a different ancestral origin. *AIDS Res Hum Retroviruses* 2008,**24**:92-99.
118. Bano AS, Sood V, Neogi U, Goel N, Kuttiat VS, Wanchu A, *et al.* Genetic and functional characterization of human immunodeficiency virus type 1 VprC variants from north India: presence of unique recombinants with mosaic genomes from B, C and D subtypes within the open reading frame of Vpr. *J Gen Virol* 2009,**90**:2768-2776.
119. Sen S, Tripathy SP, Sahni AK, Gupta RM, Kapila K, Chopra GS, *et al.* Human immunodeficiency virus type 1 gp 41 mutations in proviral DNA among antiretroviral treatment-naive individuals from India. *AIDS Res Hum Retroviruses* 2009,**25**:521-523.
120. Neogi U, Sharma Y, Sood V, Wanchu A, Banerjea AC. Diversity of HIV type 1 long terminal repeat (LTR) sequences following mother-to-child transmission in North India. *AIDS Res Hum Retroviruses* 2010,**26**:1299-1305.
121. Rainwater S, DeVange S, Sagar M, Ndinya-Achola J, Mandaliya K, Kreiss JK, *et al.* No evidence for rapid subtype C spread within an epidemic in which multiple subtypes and intersubtype recombinants circulate. *AIDS Res Hum Retroviruses* 2005,**21**:1060-1065.
122. Zhang M, Foley B, Schultz AK, Macke JP, Bulla I, Stanke M, *et al.* The role of recombination in the emergence of a complex and dynamic HIV epidemic. *Retrovirology* 2010,**7**:25.
123. Vasan A, Renjifo B, Hertzmark E, Chaplin B, Msamanga G, Essex M, *et al.* Different rates of disease progression of HIV type 1 infection in Tanzania based on infecting subtype. *Clin Infect Dis* 2006,**42**:843-852.
124. Baeten JM, Chohan B, Lavreys L, Chohan V, McClelland RS, Certain L, *et al.* HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. *J Infect Dis* 2007,**195**:1177-1180.
125. Kiwanuka N, Laeyendecker O, Robb M, Kigozi G, Arroyo M, McCutchan F, *et al.* Effect of human immunodeficiency virus Type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. *J Infect Dis* 2008,**197**:707-713.
126. Kiwanuka N, Robb M, Laeyendecker O, Kigozi G, Wabwire-Mangen F, Makumbi FE, *et al.* HIV-1 viral subtype differences in the rate of CD4+ T-cell decline among HIV seroincident antiretroviral naive persons in Rakai district, Uganda. *J Acquir Immune Defic Syndr* 2010,**54**:180-184.
127. Kim JE, Beckthold B, Chen Z, Mihowich J, Malloch L, Gill MJ. Short communication: identification of a novel HIV type 1 subtype H/J recombinant in Canada with discordant HIV viral load (RNA) values in three different commercial assays. *AIDS Res Hum Retroviruses* 2007,**23**:1309-1313.
128. Dalgleish AG, Beverley PC, Clapham PR, Crawford DH, Greaves MF, Weiss RA. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* 1984,**312**:763-767.
129. Klatzmann D, Champagne E, Chamaret S, Gruest J, Guetard D, Hercend T, *et al.* T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* 1984,**312**:767-768.

130. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 1996,**272**:872-877.
131. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, *et al.* Identification of a major co-receptor for primary isolates of HIV-1. *Nature* 1996,**381**:661-666.
132. Sander O, Sing T, Sommer I, Low AJ, Cheung PK, Harrigan PR, *et al.* Structural descriptors of gp120 V3 loop for the prediction of HIV-1 coreceptor usage. *PLoS Comput Biol* 2007,**3**:e58.
133. Dybowski JN, Heider D, Hoffmann D. Prediction of co-receptor usage of HIV-1 from genotype. *PLoS Comput Biol* 2010,**6**:e1000743.
134. Raymond S, Delobel P, Mavigner M, Cazabat M, Encinas S, Souyris C, *et al.* CXCR4-using viruses in plasma and peripheral blood mononuclear cells during primary HIV-1 infection and impact on disease progression. *AIDS* 2010,**24**:2305-2312.
135. Doranz BJ, Rucker J, Yi Y, Smyth RJ, Samson M, Peiper SC, *et al.* A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell* 1996,**85**:1149-1158.
136. Huang W, Eshleman SH, Toma J, Fransen S, Stawiski E, Paxinos EE, *et al.* Coreceptor tropism in human immunodeficiency virus type 1 subtype D: high prevalence of CXCR4 tropism and heterogeneous composition of viral populations. *J Virol* 2007,**81**:7885-7893.
137. Kaleebu P, Nankya IL, Yirrell DL, Shafer LA, Kyosiimire-Lugemwa J, Lule DB, *et al.* Relation between chemokine receptor use, disease stage, and HIV-1 subtypes A and D: results from a rural Ugandan cohort. *J Acquir Immune Defic Syndr* 2007,**45**:28-33.
138. Wambui V, Kiptoo MK, Kinyua JG, Odera I, Muiruri P, Muge E, *et al.* Predicted HIV-1 coreceptor usage among Kenya patients shows a high tendency for subtype d to be cxcr4 tropic. *AIDS Res Ther* 2012,**9**:22.
139. Neogi U, Prarthana SB, D'Souza G, Decosta A, Kuttiatt VS, Ranga U, *et al.* Co-receptor tropism prediction among 1045 Indian HIV-1 subtype C sequences: Therapeutic implications for India. *AIDS Res Ther* 2010,**7**:24.
140. Renjifo B, Gilbert P, Chaplin B, Msamanga G, Mwakagile D, Fawzi W, *et al.* Preferential in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D. *AIDS* 2004,**18**:1629-1636.
141. Hudgens MG, Longini IM, Jr., Vanichseni S, Hu DJ, Kitayaporn D, Mock PA, *et al.* Subtype-specific transmission probabilities for human immunodeficiency virus type 1 among injecting drug users in Bangkok, Thailand. *Am J Epidemiol* 2002,**155**:159-168.
142. Ivers LC, Kendrick D, Doucette K. Efficacy of antiretroviral therapy programs in resource-poor settings: a meta-analysis of the published literature. *Clin Infect Dis* 2005,**41**:217-224.
143. Barth RE, van der Loeff MF, Schuurman R, Hoepelman AI, Wensing AM. Virological follow-up of adult patients in antiretroviral treatment programmes in sub-Saharan Africa: a systematic review. *Lancet Infect Dis* 2010,**10**:155-166.
144. Bussmann H, Wester CW, Ndwapi N, Grundmann N, Gaolathe T, Puvimanasinghe J, *et al.* Five-year outcomes of initial patients treated in Botswana's National Antiretroviral Treatment Program. *AIDS* 2008,**22**:2303-2311.
145. Innes C CS, Maraisane M, Senoge S, Seatlanyane P, *et al.* 24 and 36 month outcomes among HIV-infected individuals starting antiretroviral therapy in a

workplace program in South Africa. AIDS 2008 - XVII International AIDS Conference Abstract no. MOPE0046. 2008.
146. Konate I, Traore L, Ouedraogo A, Sanon A, Diallo R, Ouedraogo JL, *et al.* Linking HIV prevention and care for community interventions among high-risk women in Burkina Faso--the ARNS 1222 "Yerelon" cohort. *J Acquir Immune Defic Syndr* 2011,**57 Suppl 1**:S50-54.
147. Geretti AM, Harrison L, Green H, Sabin C, Hill T, Fearnhill E, *et al.* Effect of HIV-1 subtype on virologic and immunologic response to starting highly active antiretroviral therapy. *Clin Infect Dis* 2009,**48**:1296-1305.
148. Martinez-Cajas JL, Pai NP, Klein MB, Wainberg MA. Differences in resistance mutations among HIV-1 non-subtype B infections: a systematic review of evidence (1996-2008). *J Int AIDS Soc* 2009,**12**:11.
149. Lai MT, Lu M, Felock PJ, Hrin RC, Wang YJ, Yan Y, *et al.* Distinct mutation pathways of non-subtype B HIV-1 during in vitro resistance selection with nonnucleoside reverse transcriptase inhibitors. *Antimicrob Agents Chemother* 2010,**54**:4812-4824.
150. Wainberg MA, Brenner BG. Role of HIV Subtype Diversity in the Development of Resistance to Antiviral Drugs. *Viruses* 2010,**2**:2493-2508.
151. Novitsky V, Wester CW, DeGruttola V, Bussmann H, Gaseitsiwe S, Thomas A, *et al.* The reverse transcriptase 67N 70R 215Y genotype is the predominant TAM pathway associated with virologic failure among HIV type 1C-infected adults treated with ZDV/ddI-containing HAART in southern Africa. *AIDS Res Hum Retroviruses* 2007,**23**:868-878.
152. Hosseinipour MC, van Oosterhout JJ, Weigel R, Phiri S, Kamwendo D, Parkin N, *et al.* The public health approach to identify antiretroviral therapy failure: high-level nucleoside reverse transcriptase inhibitor resistance among Malawians failing first-line antiretroviral therapy. *AIDS* 2009,**23**:1127-1134.
153. Kumarasamy N, Madhavan V, Venkatesh KK, Saravanan S, Kantor R, Balakrishnan P, *et al.* High frequency of clinically significant mutations after first-line generic highly active antiretroviral therapy failure: implications for second-line options in resource-limited settings. *Clin Infect Dis* 2009,**49**:306-309.
154. Marconi VC, Sunpath H, Lu Z, Gordon M, Koranteng-Apeagyei K, Hampton J, *et al.* Prevalence of HIV-1 drug resistance after failure of a first highly active antiretroviral therapy regimen in KwaZulu Natal, South Africa. *Clin Infect Dis* 2008,**46**:1589-1597.
155. Miller MD, Margot N, McColl D, Cheng AK. K65R development among subtype C HIV-1-infected patients in tenofovir DF clinical trials. *AIDS* 2007,**21**:265-266.
156. Coutsinos D, Invernizzi CF, Xu H, Moisi D, Oliveira M, Brenner BG, *et al.* Template usage is responsible for the preferential acquisition of the K65R reverse transcriptase mutation in subtype C variants of human immunodeficiency virus type 1. *J Virol* 2009,**83**:2029-2033.
157. Coutsinos D, Invernizzi CF, Xu H, Brenner BG, Wainberg MA. Factors affecting template usage in the development of K65R resistance in subtype C variants of HIV type-1. *Antivir Chem Chemother* 2010,**20**:117-131.
158. Coutsinos D, Invernizzi CF, Moisi D, Oliveira M, Martinez-Cajas JL, Brenner BG, *et al.* A template-dependent dislocation mechanism potentiates K65R reverse transcriptase mutation development in subtype C variants of HIV-1. *PLoS One* 2011,**6**:e20208.
159. Recordon-Pinson P, Papuchon J, Reigadas S, Deshpande A, Fleury H. K65R in subtype C HIV-1 isolates from patients failing on a first-line regimen including

d4T or AZT: comparison of Sanger and UDP sequencing data. *PLoS One* 2012,**7**:e36549.
160. Sunpath H, Wu B, Gordon M, Hampton J, Johnson B, Moosa MY, *et al.* High rate of K65R for antiretroviral therapy-naive patients with subtype C HIV infection failing a tenofovir-containing first-line regimen. *AIDS* 2012,**26**:1679-1684.
161. Hsu LY, Subramaniam R, Bacheler L, Paton NI. Characterization of mutations in CRF01_AE virus isolates from antiretroviral treatment-naive and -experienced patients in Singapore. *J Acquir Immune Defic Syndr* 2005,**38**:5-13.
162. Rajesh L, Karunaianantham R, Narayanan PR, Swaminathan S. Antiretroviral drug-resistant mutations at baseline and at time of failure of antiretroviral therapy in HIV type 1-coinfected TB patients. *AIDS Res Hum Retroviruses* 2009,**25**:1179-1185.
163. Huang A, Hogan JW, Istrail S, Delong A, Katzenstein DA, Kantor R. Global analysis of sequence diversity within HIV-1 subtypes across geographic regions. *Future Virol* 2012,**7**:505-517.
164. Carr JK. Viral diversity as a challenge to HIV-1 vaccine development. *Curr Opin HIV AIDS* 2006,**1**:294-300.
165. Salemi M. Toward a robust monitoring of HIV subtypes distribution worldwide. *AIDS* 2011,**25**:713-714.
166. Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, Paris R, *et al.* Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N Engl J Med* 2009,**361**:2209-2220.
167. Collier AC, Coombs RW, Schoenfeld DA, Bassett RL, Timpone J, Baruch A, *et al.* Treatment of human immunodeficiency virus infection with saquinavir, zidovudine, and zalcitabine. AIDS Clinical Trials Group. *N Engl J Med* 1996,**334**:1011-1017.
168. Wong JK, Gunthard HF, Havlir DV, Zhang ZQ, Haase AT, Ignacio CC, *et al.* Reduction of HIV-1 in blood and lymph nodes following potent antiretroviral therapy and the virologic correlates of treatment failure. *Proc Natl Acad Sci U S A* 1997,**94**:12574-12579.
169. Young FE. The role of the FDA in the effort against AIDS. *Public Health Rep* 1988,**103**:242-245.
170. Mitsuya H, Broder S. Inhibition of the in vitro infectivity and cytopathic effect of human T-lymphotrophic virus type III/lymphadenopathy-associated virus (HTLV-III/LAV) by 2',3'-dideoxynucleosides. *Proc Natl Acad Sci U S A* 1986,**83**:1911-1915.
171. Gao WY, Agbaria R, Driscoll JS, Mitsuya H. Divergent anti-human immunodeficiency virus activity and anabolic phosphorylation of 2',3'-dideoxynucleoside analogs in resting and activated human cells. *J Biol Chem* 1994,**269**:12633-12638.
172. Warnke D, Barreto J, Temesgen Z. Antiretroviral drugs. *J Clin Pharmacol* 2007,**47**:1570-1579.
173. Arts EJ, Hazuda DJ. HIV-1 Antiretroviral Drug Therapy. *Cold Spring Harb Perspect Med* 2012,**2**:a007161.
174. Grob PM, Wu JC, Cohen KA, Ingraham RH, Shih CK, Hargrave KD, *et al.* Nonnucleoside inhibitors of HIV-1 reverse transcriptase: nevirapine as a prototype drug. *AIDS Res Hum Retroviruses* 1992,**8**:145-152.
175. Merluzzi VJ, Hargrave KD, Labadia M, Grozinger K, Skoog M, Wu JC, *et al.* Inhibition of HIV-1 replication by a nonnucleoside reverse transcriptase inhibitor. *Science* 1990,**250**:1411-1413.

176. Spence RA, Kati WM, Anderson KS, Johnson KA. Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors. *Science* 1995,**267**:988-993.
177. Witvrouw M, Pannecouque C, Switzer WM, Folks TM, De Clercq E, Heneine W. Susceptibility of HIV-2, SIV and SHIV to various anti-HIV-1 compounds: implications for treatment and postexposure prophylaxis. *Antivir Ther* 2004,**9**:57-65.
178. Wlodawer A, Vondrasek J. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annu Rev Biophys Biomol Struct* 1998,**27**:249-284.
179. Drag M, Salvesen GS. Emerging principles in protease-based drug discovery. *Nat Rev Drug Discov* 2010,**9**:690-701.
180. Wensing AM, van Maarseveen NM, Nijhuis M. Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. *Antiviral Res* 2010,**85**:59-74.
181. Dragic T, Trkola A, Thompson DA, Cormier EG, Kajumo FA, Maxwell E, *et al.* A binding pocket for a small molecule inhibitor of HIV-1 entry within the transmembrane helices of CCR5. *Proc Natl Acad Sci U S A* 2000,**97**:5639-5644.
182. Tsamis F, Gavrilov S, Kajumo F, Seibert C, Kuhmann S, Ketas T, *et al.* Analysis of the mechanism by which the small-molecule CCR5 antagonists SCH-351125 and SCH-350581 inhibit human immunodeficiency virus type 1 entry. *J Virol* 2003,**77**:5201-5208.
183. Hazuda D, Blau CU, Felock P, Hastings J, Pramanik B, Wolfe A, *et al.* Isolation and characterization of novel human immunodeficiency virus integrase inhibitors from fungal metabolites. *Antivir Chem Chemother* 1999,**10**:63-70.
184. Espeseth AS, Felock P, Wolfe A, Witmer M, Grobler J, Anthony N, *et al.* HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. *Proc Natl Acad Sci U S A* 2000,**97**:11244-11249.
185. Hazuda DJ, Young SD, Guare JP, Anthony NJ, Gomez RP, Wai JS, *et al.* Integrase inhibitors and cellular immunity suppress retroviral replication in rhesus macaques. *Science* 2004,**305**:528-532.
186. Shimura K, Kodama E, Sakagami Y, Matsuzaki Y, Watanabe W, Yamataka K, *et al.* Broad antiretroviral activity and resistance profile of the novel human immunodeficiency virus integrase inhibitor elvitegravir (JTK-303/GS-9137). *J Virol* 2008,**82**:764-774.
187. Hicks C, Gulick RM. Raltegravir: the first HIV type 1 integrase inhibitor. *Clin Infect Dis* 2009,**48**:931-939.
188. Care, Support and Treatment, National AIDS Control Programme, Phase-III, India. Available at http://nacoonline.org/upload/Care%20&%20Treatment/Monograph%20on%20HIV_CST%20division%20_21st%20June%202010.pdf, Accessedon 17 October 2012.
189. Department of AIDS Control Ministry of Health & Family Welfare Annual Report 2011-12. Available at http://www.nacoonline.org/upload/Publication/Annual%20Report/NACO_AR_Eng%202011-12.pdf accessed on December 2012.
190. National AIDS Control Organisation. National Guidelines on Second-line ART for adults and adolescents (April 2011). Available at http://delhisacs.org/commonincludes/pg_line/2nd_line_art.pdf. accessed on December 2012.

191. Kumarasamy N, Vallabhaneni S, Flanigan TP, Balakrishnan P, Cecelia A, Carpenter CC, *et al.* Rapid viral load suppression following generic highly active antiretroviral therapy in Southern Indian HIV-infected patients. *AIDS* 2005,**19**:625-627.
192. Shet A, Neogi U, Sahoo PN, De Costa A. Effectiveness of First-Line Antiretroviral Therapy and Acquired Drug Resistance among HIV-1-Infected Children in India. *Pediatr Infect Dis J* 2012.
193. Hingankar NK, Thorat SR, Deshpande A, Rajasekaran S, Chandrasekar C, Kumar S, *et al.* Initial virologic response and HIV drug resistance among HIV-infected individuals initiating first-line antiretroviral therapy at 2 clinics in Chennai and Mumbai, India. *Clin Infect Dis* 2012,**54 Suppl 4**:S348-354.
194. Gupta RK, Jordan MR, Sultan BJ, Hill A, Davis DH, Gregson J, *et al.* Global trends in antiretroviral resistance in treatment-naive individuals with HIV after rollout of antiretroviral treatment in resource-limited settings: a global collaborative study and meta-regression analysis. *Lancet* 2012,**380**:1250-1258.
195. Johnson VA, Calvez V, Gunthard HF, Paredes R, Pillay D, Shafer R, *et al.* 2011 update of the drug resistance mutations in HIV-1. *Top Antivir Med* 2011,**19**:156-164.
196. Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, *et al.* Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. *PLoS One* 2009,**4**:e4724.
197. Montes B, Vergne L, Peeters M, Reynes J, Delaporte E, Segondy M. Comparison of drug resistance mutations and their interpretation in patients infected with non-B HIV-1 variants and matched patients infected with HIV-1 subtype B. *J Acquir Immune Defic Syndr* 2004,**35**:329-336.
198. Ross L, Lim ML, Liao Q, Wine B, Rodriguez AE, Weinberg W, *et al.* Prevalence of antiretroviral drug resistance and resistance-associated mutations in antiretroviral therapy-naive HIV-infected individuals from 40 United States cities. *HIV Clin Trials* 2007,**8**:1-8.
199. Snoeck J, Kantor R, Shafer RW, Van Laethem K, Deforche K, Carvalho AP, *et al.* Discordances between interpretation algorithms for genotypic resistance to protease and reverse transcriptase inhibitors of human immunodeficiency virus are subtype dependent. *Antimicrob Agents Chemother* 2006,**50**:694-701.
200. Shafer RW, Rhee SY, Pillay D, Miller V, Sandstrom P, Schapiro JM, *et al.* HIV-1 protease and reverse transcriptase mutations for drug resistance surveillance. *AIDS* 2007,**21**:215-223.
201. Johnson VA, Calvez, V., Günthard, H.F., Paredes, R., Pillay, D., Shafer, R.W., Wensing, A.M., Richman, D.D. . Update of the Drug Resistance Mutations in HIV-1: March 2013. *Top Antivir Med* 2013,**21**:4-12.
202. Frentz D, Boucher CA, van de Vijver DA. Temporal changes in the epidemiology of transmission of drug-resistant HIV-1 across the world. *AIDS Rev* 2012,**14**:17-27.
203. Sethi AK, Celentano DD, Gange SJ, Moore RD, Gallant JE. Association between adherence to antiretroviral therapy and human immunodeficiency virus drug resistance. *Clin Infect Dis* 2003,**37**:1112-1118.
204. Roge BT, Barfod TS, Kirk O, Katzenstein TL, Obel N, Nielsen H, *et al.* Resistance profiles and adherence at primary virological failure in three different highly active antiretroviral therapy regimens: analysis of failure rates in a randomized study. *HIV Med* 2004,**5**:344-351.
205. Ekstrand ML, Shet A, Chandy S, Singh G, Shamsundar R, Madhavan V, *et al.* Suboptimal adherence associated with virological failure and resistance

mutations to first-line highly active antiretroviral therapy (HAART) in Bangalore, India. *Int Health* 2011,**3**:27-34.
206. Santos AF, Soares MA. HIV Genetic Diversity and Drug Resistance. *Viruses* 2010,**2**:503-531.
207. Holguin A, Paxinos E, Hertogs K, Womac C, Soriano V. Impact of frequent natural polymorphisms at the protease gene on the in vitro susceptibility to protease inhibitors in HIV-1 non-B subtypes. *J Clin Virol* 2004,**31**:215-220.
208. Johnson VA, Brun-Vezinet F, Clotet B, Gunthard HF, Kuritzkes DR, Pillay D, *et al.* Update of the Drug Resistance Mutations in HIV-1. *Top HIV Med* 2008,**16**:138-145.
209. Kantor R, Katzenstein DA, Efron B, Carvalho AP, Wynhoven B, Cane P, *et al.* Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration. *PLoS Med* 2005,**2**:e112.
210. Brenner BG, Oliveira M, Doualla-Bell F, Moisi DD, Ntemgwa M, Frankel F, *et al.* HIV-1 subtype C viruses rapidly develop K65R resistance to tenofovir in cell culture. *AIDS* 2006,**20**:F9-13.
211. Grossman Z, Istomin V, Averbuch D, Lorber M, Risenberg K, Levi I, *et al.* Genetic variation at NNRTI resistance-associated positions in patients infected with HIV-1 subtype C. *AIDS* 2004,**18**:909-915.
212. Vidya M, Saravanan S, Uma S, Kumarasamy N, Sunil SS, Kantor R, *et al.* Genotypic HIV type-1 drug resistance among patients with immunological failure to first-line antiretroviral therapy in south India. *Antivir Ther* 2009,**14**:1005-1009.
213. Neogi U, Shet A, Shamsundar R, Ekstrand ML. Selection of nonnucleoside reverse transcriptase inhibitor-associated mutations in HIV-1 subtype C: evidence of etravirine cross-resistance. *AIDS* 2011,**25**:1123-1126.
214. Anquetil D, Deshpande A, Zongo D, Le Bihan L, Pinson PR, Fleury HJ. Susceptibility to etravirine of HIV type 1 subtype C isolates from nevirapine/efavirenz-experienced patients: comparative interpretation of ANRS and STANFORD algorithms. *AIDS Res Hum Retroviruses* 2012,**28**:1793-1797.
215. Tee KK, Pybus OG, Li XJ, Han X, Shang H, Kamarulzaman A, *et al.* Temporal and spatial dynamics of human immunodeficiency virus type 1 circulating recombinant forms 08_BC and 07_BC in Asia. *J Virol* 2008,**82**:9206-9215.
216. Shen C, Craigo J, Ding M, Chen Y, Gupta P. Origin and dynamics of HIV-1 subtype C infection in India. *PLoS One* 2011,**6**:e25956.
217. Abecasis AB, Vandamme AM, Lemey P. Quantifying differences in the tempo of human immunodeficiency virus type 1 subtype evolution. *J Virol* 2009,**83**:12917-12924.
218. Veras NM, Gray RR, Brigido LF, Rodrigues R, Salemi M. High-resolution phylogenetics and phylogeography of human immunodeficiency virus type 1 subtype C epidemic in South America. *J Gen Virol* 2011,**92**:1698-1709.
219. Bachani D, Garg R, Rewari BB, Hegg L, Rajasekaran S, Deshpande A, *et al.* Two-year treatment outcomes of patients enrolled in India's national first-line antiretroviral therapy programme. *Natl Med J India* 2010,**23**:7-12.
220. Shet A, DeCosta A, Heylen E, Shastri S, Chandy S, Ekstrand M. High rates of adherence and treatment success in a public and public-private HIV clinic in India: potential benefits of standardized national care delivery systems. *BMC Health Serv Res* 2011,**11**:277.
221. Giordano TP, Guzman D, Clark R, Charlebois ED, Bangsberg DR. Measuring adherence to antiretroviral therapy in a diverse population using a visual analogue scale. *HIV Clin Trials* 2004,**5**:74-79.

222. Ekstrand ML, Chandy S, Heylen E, Steward W, Singh G. Developing useful highly active antiretroviral therapy adherence measures for India: the Prerana study. *J Acquir Immune Defic Syndr* 2010,**53**:415-416.
223. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* 2011,**28**:2731-2739.
224. de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, *et al.* An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics* 2005,**21**:3797-3800.
225. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 2007,**7**:214.
226. Posada D. jModelTest: phylogenetic model averaging. *Mol Biol Evol* 2008,**25**:1253-1256.
227. Oh DY, Jessen H, Kucherer C, Neumann K, Oh N, Poggensee G, *et al.* CCR5Delta32 genotypes in a German HIV-1 seroconverter cohort and report of HIV-1 infection in a CCR5Delta32 homozygous individual. *PLoS One* 2008,**3**:e2747.
228. Coetzer M, Cilliers T, Ping LH, Swanstrom R, Morris L. Genetic characteristics of the V3 region associated with CXCR4 usage in HIV-1 subtype C isolates. *Virology* 2006,**356**:95-105.
229. Michler K, Connell BJ, Venter WD, Stevens WS, Capovilla A, Papathanasopoulos MA. Genotypic characterization and comparison of full-length envelope glycoproteins from South African HIV type 1 subtype C primary isolates that utilize CCR5 and/or CXCR4. *AIDS Res Hum Retroviruses* 2008,**24**:743-751.
230. Choge I, Cilliers T, Walker P, Taylor N, Phoswa M, Meyers T, *et al.* Genotypic and phenotypic characterization of viral isolates from HIV-1 subtype C-infected children with slow and rapid disease progression. *AIDS Res Hum Retroviruses* 2006,**22**:458-465.
231. Papathanasopoulos MA, Cilliers T, Morris L, Mokili JL, Dowling W, Birx DL, *et al.* Full-length genome analysis of HIV-1 subtype C utilizing CXCR4 and intersubtype recombinants isolated in South Africa. *AIDS Res Hum Retroviruses* 2002,**18**:879-886.
232. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of HIV coreceptor usage. *Nat Biotechnol* 2007,**25**:1407-1410.
233. Vandekerckhove LP, Wensing AM, Kaiser R, Brun-Vezinet F, Clotet B, De Luca A, *et al.* European guidelines on the clinical management of HIV-1 tropism testing. *Lancet Infect Dis* 2011,**11**:394-407.
234. Pace C, Keller J, Nolan D, James I, Gaudieri S, Moore C, *et al.* Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with APOBEC3G and vif genetic variation. *J Virol* 2006,**80**:9259-9269.
235. Ulenga NK, Sarr AD, Hamel D, Sankale JL, Mboup S, Kanki PJ. The level of APOBEC3G (hA3G)-related G-to-A mutations does not correlate with viral load in HIV type 1-infected individuals. *AIDS Res Hum Retroviruses* 2008,**24**:1285-1290.
236. Rose PP, Korber BT. Detecting hypermutations in viral sequences with an emphasis on G -> A hypermutation. *Bioinformatics* 2000,**16**:400-401.
237. Gifford RJ, Rhee SY, Eriksson N, Liu TF, Kiuchi M, Das AK, *et al.* Sequence editing by Apolipoprotein B RNA-editing catalytic component-B and

epidemiological surveillance of transmitted HIV-1 drug resistance. *AIDS* 2008,**22**:717-725.

238. Sungkanuparph S, Kiertiburanakul S, Sukasem C, Chantratita W. Discrepancies between WHO 2009 and IAS-USA 2009 lists for determining the rate of transmitted HIV-1 drug resistance: a prospective study. *J Acquir Immune Defic Syndr* 2012,**59**:e3-5.
239. Hoffmann CJ, Charalambous S, Sim J, Ledwaba J, Schwikkard G, Chaisson RE, *et al.* Viremia, resuppression, and time to resistance in human immunodeficiency virus (HIV) subtype C during first-line antiretroviral therapy in South Africa. *Clin Infect Dis* 2009,**49**:1928-1935.
240. El-Khatib Z, Delong AK, Katzenstein D, Ekstrom AM, Ledwaba J, Mohapi L, *et al.* Drug resistance patterns and virus re-suppression among HIV-1 subtype C infected patients receiving non-nucleoside reverse transcriptase inhibitors in South Africa. *J AIDS Clin Res* 2011,**2**.
241. Aleman S, Soderbarg K, Visco-Comandini U, Sitbon G, Sonnerborg A. Drug resistance at low viraemia in HIV-1-infected patients with antiretroviral combination therapy. *AIDS* 2002,**16**:1039-1044.
242. Verhofstede C, Van Wanzeele F, Van Der Gucht B, Pelgrom J, Vandekerckhove L, Plum J, *et al.* Detection of drug resistance mutations as a predictor of subsequent virological failure in patients with HIV-1 viral rebounds of less than 1,000 RNA copies/ml. *J Med Virol* 2007,**79**:1254-1260.

10 APPENDICES

Appendix I: Mutations in the *reverse transcriptase* gene concomitant with resistance to reverse transcriptase inhibitors, NRTI and NNRTI (adapted from International AIDS society-USA list of DRM: March 2013 Update)

Drugs	Resistance Mutations
Nucleoside and Nucleotide Analogue Reverse Transcriptase Inhibitors (NRTIs)	
Abacavir (ABC)	K65R, L74V, Y115F, M184V
Didanosine (ddI)	K65R, L74V
Emtricitabine (FTC)	K65R, M184V/I
Lamivudine (3TC)	K65R, M184V/I
Stavudine (d4T)	M41L, K65R, D67N, K70R, L210W, T215Y/F, K219Q/E
Tenofovir (TDF)	K65R, K70E
Zidovudine (AZT)	M41L, D67N, K70R, L210W, T215Y/F, K219Q/E
69 insertion complex	M41L, A62V, 69ins, K70R, L210W, T215Y/F, K219Q/E
151 complex	A62V, V75I, F77L, F116Y, Q151M
TAM-1	T215Y, M41L, L210W
TAM-2	K70R, D67N, T215F, K219Q
Non-nucleoside Analogue Reverse Transcriptase Inhibitors (NNRTIs)	
Efavirenz (EFV)	L100I, K101P, K103N/S, V106M, V108I, Y181C/I, Y188L, G190S/A, P225H, M230L
Nevirapine (NVP)	L100I, K101P, K103N/S, V106M/A, V108I, Y181C/I, Y188C/L/H, G190A, M230L
Etravirine (ETR)	V90I, A98G, L100I, K101E/H/P, V106I, E138A/G/K/Q, V179D/F/T, Y181C/I/V, G190S/A, M230L
Rilpivirine (RPV)	K101E/P, E138A/G/K/Q/R, V179L, Y181C/I/V, Y188L, H221Y, F227C, M230I/L

Abbreviations: TAM- Thymidine Analogue Mutation
Amino acid abbreviations: A-Alanine; C-Cysteine; D-Aspartate; E-Glutamate; F-Phenylalanine; G-lycine; H-Histidine; I-Isoleucine; K-Lysine; L-Leucine; M-Methionine; N-Asparagine; P-Proline; Q-Glutamine; R-Arginine; S-Serine; T-Threonine; V-Valine; W-Tryptophan; Y-Tyrosine.

Appendix II: Major mutations in the *protease* gene concomitant with resistance to protease inhibitors (adapted from International AIDS society-USA list 2013 and Stanford Drug Resistance Database, http://hivdb.stanford.edu).

Drugs	Resistance Mutations
Atazanavir/r (ATV/r)	V32I, M46I/L, I47V, G48V/M, I50L, I54V/T/A/M, V82A/T/F/S, I84V, N88S, L90M
Darunavir/r (DRV/r)	V32I, I47V/A, I50V, I54L/M, L76V, V82F, I84V,
Fosamprenavir/r (FPV/r)	V32I, M46I/L, I47V/A, I50V, I54V/T/A/L/M, L76V, V82A/T/F/S, I84V, L90M
Indinavir/r (IDV/r)	V32I, M46I/L, I47V/A, I54V/T/A, L76V, V82A/T/F/S, I84V, N88S, L90M
Lopinavir/r (LPV/r)	V32I, M46I/L, I47V/A, G48V/M, I50V, I54V/T/A/L/M, L76V, V82A/T/F/S , I84V, L90M
Nelfinavir/r (NFV/r)	D30N, M46I/L, I47V/A, G48V/M, I54V/T/A/L/M, V82A/T/F/S , I84V, N88D/S, L90M
Saquinavir/r (SQV/r)	G48V/M, I54V/T/A/M, V82A/T, I84V, N88S, L90M
Tipranavir/r (TPV/r)	V32I, M46I/L, I47V/A, I54V/T/A/M, V82L/T, I84V,