From DEPARTMENT OF LABORATORY MEDICINE

Karolinska Institutet, Stockholm, Sweden

DYNAMICS OF DRUG-RESISTANT MINOR VIRAL POPULATIONS AND PHENOTYPIC DRUG SUSCEPTIBILITY IN DIVERSE HIV-1 SUBTYPES

Shambhu Prasad Ganeshappa Aralaguppe

Karolinska Institutet

Stockholm 2019

Published by Karolinska Institutet.

Printed by E-Print AB 2019

ISBN 978-91-7831-459-1

Dynamics of drug-resistant minor viral populations and phenotypic drug susceptibility in diverse HIV-1 subtypes

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Shambhu Prasad Ganeshappa Aralaguppe

Principal Supervisor:
Prof. Anders Sönnerborg
Karolinska Institutet
Department of Laboratory Medicine
Division of Clinical Microbiology

Co-supervisors:
Docent Ujjwal Neogi
Karolinska Institutet
Department of Laboratory Medicine
Division of Clinical Microbiology

Prof. Francesca Chiodi
Karolinska Institutet
Department of Microbiology,
Tumor and Cellbiology

Prof. Vinayaka Prasad
Albert Einstein College of Medicine
Department of Microbiology and Immunology

Mentor:
Prof. Cecilia Stålsby Lundborg
Karolinska Institutet
Department of Public Health Sciences

Opponent:
Prof. Rami Kantor
Brown University
Division of Infectious Diseases

Examination Board:
Docent Christer Lidman
Karolinska Institutet
Department of Medicine Huddinge
Division of Infectious Diseases

Docent Mia Brytting
Public Health Agency of Sweden
Department of Microbiology

Prof. Sören Andersson
University of Örebro
School of Medical Sciences
Department of Microbiology

Dedicated to my family and all my teachers.

ಹುಲ್ಲಾಗು ಬೆಟ್ಟದಡಿ ಮನೆಗೆ ಮಲ್ಲಿಗೆಯಾಗು |
ಕಲ್ಲಾಗು ಕಷ್ಟಗಳ ಮಳೆಯ ವಿಧಿ ಸುರಿಯೆ ‖
ಬೆಲ್ಲ ಸಕ್ಕರೆಯಾಗು ದೀನದುರ್ಬಲರಿಂಗೆ |
ಎಲ್ಲರೊಳಗೊಂದಾಗು – ಮಂಕುತಿಮ್ಮ ‖

-ಡಿ. ವಿ. ಗುಂಡಪ್ಪ

Be a grass at the foot of mountains, and jasmine flower to thy home,
Be (strong) like a rock when fate rains torrential troubles at thou,
Be sweet (kind/savior) like jaggery to the needy and downtrodden,
Be one among all – says *Mankuthimma.*

-D.V.Gundappa

(One of the poems from collection, “Mankuthimmana kagga” by D. V. Gundappa)

ABSTRACT

Antiretroviral therapy (ART) has significantly reduced mortality in human immunodeficiency virus type 1 (HIV-1) infection, both in high- and low/middle-income countries (LMIC). However, the development of drug resistance following exposure to subtherapeutic concentrations of antiretrovirals (ARV) and transmitted drug resistance mutations (DRM) adversely affects the outcome of ART.

In my thesis, drug resistance patterns of diverse HIV-1 subtypes were described, at both genotypic and phenotypic level. A high throughput sequencing (HTS) method was utilized to amplify and sequence the *gag-pol* of different HIV-1 subtypes for identification and quantification of DRM present in <20% of the viral population. A novel bioinformatics pipeline was developed, MiDRM*pol*, which integrates genomic variations and mapping of minor populations with DRM. Phenotypic drug sensitivity assays were performed to study *in vitro* potency of newer ARV. Also, polymerase independent increase of virulence and replication competence of HIV-1, which may influence the outcome of ART, was studied.

Paper I describes the pipeline, MiDRM*pol*, which can be used without any prior knowledge and does not require on-site bioinformatics support. The raw data from HTS in the fastq format are uploaded to get an easily readable format. One of the unique features when compared to other available pipelines is the FastUniq tool, which removes the duplicate reads generated by PCR, thus reducing the pseudo-bias of few variants. Another feature is the subtype-specific adaptation during the analysis. In **Paper II** and **Paper III**, I studied the potency of three newer ARV. 4′-Ethynyl-2-fluoro-2′-deoxyadenosine (EFdA), a novel translocation defective reverse transcriptase inhibitor (RTI), was found to inhibit both wild-type and RTI resistant viruses efficiently in a subtype-independent manner.

Similarly, the second-generation integrase strand inhibitors (INSTI) cabotegravir (CAB) and bictegravir (BIC) were shown to have equal or higher potency against non-B subtypes as compared to HIV-1B. This confirms the suitability of these drugs for use in countries dominated by non-B subtypes. In **Paper IV,** I observed that a PYxE insertion in the *gag* plays a role in increased virulence and replication capacity in HIV-1C viruses and seems to be associated with suboptimal $CD4^+$ T-cell gain following ART initiation. Even though there was no effect of PYxE or PTAPP on the susceptibility to 20 ARVs, the enhanced replication capacity might increase the time to reach viral suppression during ART and thereby increased risk for the virus to develop DRM.

In conclusion, identification and quantification of DRM at frequencies <20% is a major hurdle in current ART monitoring, and our MiDRM*pol* facilitates the analysis of such HTS data. However, we found that also polymerase independent mutations which increase the replication capacity/virulence of HIV-1 may influence the outcome of ART. The new ARVs EFdA, CAB, and BIC suppress the viral load *in vitro* in a subtype independent manner. This is important since if not all subtypes are suppressed efficiently the risk of a global increase of acquired DRM would be further increased.

LIST OF SCIENTIFIC PAPERS

I. **Aralaguppe SG**, Ambikan AT, Njenda DT, Manickam A, Kumar MM, Hanna LE, Amogne W, Sönnerborg A, Neogi U. MiDRM*pol*: A high-throughput multiplexed amplicon sequencing workflow to identify protease, reverse transcriptase and integrase inhibitors minor drug resistance mutations. *Manuscript submitted*.

II. Njenda DT, **Aralaguppe SG**, Singh K, Rao R, Sonnerborg A, Sarafianos SG, Neogi U. Antiretroviral potency of 4'-ethnyl-2'-fluoro-2'-deoxyadenosine, tenofovir alafenamide and second-generation NNRTIs across diverse HIV-1 subtypes. *J Antimicrob Chemother* 2018,73:2721-2728.

III. Neogi U, Singh K, **Aralaguppe SG**, Rogers LC, Njenda DT, Sarafianos SG, Hejdeman B, Sönnerborg A. Ex-vivo antiretroviral potency of newer integrase strand transfer inhibitors cabotegravir and bictegravir in HIV type 1 non-B subtypes. *AIDS* 2018,32:469-476.

IV. **Aralaguppe SG**, Winner D, Singh K, Sarafianos SG, Quinones-Mateu ME, Sönnerborg A, Neogi U. Increased replication capacity following evolution of PYxE insertion in Gag-p6 is associated with enhanced virulence in HIV-1 subtype C from East Africa. *J Med Virol* 2017,89:106-111.

LIST OF PUBLICATIONS NOT INCLUDED IN THESIS

I. **Aralaguppe SG**, Siddik AB, Manickam A, Ambikan AT, Kumar MM, Fernandes SJ, Amogne W, Bangaruswamy DK, Hanna LE, Sonnerborg A, Neogi U. Multiplexed next-generation sequencing and de novo assembly to obtain near full-length HIV-1 genome from plasma virus. *Journal of Virological methods* 2016; 236:98-104.

II. Siddik AB, Haas A, Rahman MS, **Aralaguppe SG**, Amogne W, Bader J, Bader J, klimkait T, Neogi U. Phenotypic co-receptor tropism and Maraviroc sensitivity in HIV-1 subtype C from East Africa. *Sci Rep* 2018; 8(1):2363.

III. Kalu AW, Telele NF, **Aralaguppe SG**, Gebre-Selassie S, Fekade D, Marrone G, Sonnerborg A. Coreceptor Tropism and Maraviroc Sensitivity of Clonally Derived Ethiopian HIV-1C Strains Using an in-house Phenotypic Assay and Commonly Used Genotypic Methods. *Curr HIV Res* 2018; 16(2):113-120.

IV. Ashokkumar M, **Aralaguppe SG**, Tripathy SP, Hanna LE, Neogi U. Unique Phenotypic Characteristics of Recently Transmitted HIV-1 Subtype C Envelope Glycoprotein gp120: Use of CXCR6 Coreceptor by Transmitted Founder Viruses. *J Virol* 2018,92 Apr 13;92(9). pii: e00063-18. doi: 10.1128/JVI.00063-18. Print 2018 May 1.

CONTENTS

LIST OF ABBREVIATIONS

3TC	Lamivudine
ADR	Acquired drug resistance
AIDS	Acquired Immunodeficiency Syndrome
ART	Antiretroviral Therapy
BIC	Bictegravir
CA	Capsid
CAB	Cabotegravir
cART	Combination antiretroviral therapy
DNA	Deoxyribonucleic Acid
DRM	Drug Resistance Mutation
DTG	Dolutegravir
EFdA	4'-Ethynyl 2-fluro deoxyadenosine
EMA	European Medicines Agency
ENV	Envelope
EVG	Elvitegravir
FDA	Food and Drug Administration
FTC	Emtricitabine
GRT	Genotypic resistance testing
HIC	High income countries
HIV	Human Immunodeficiency Virus
HIV-DR	Human Immunodeficiency Virus-Drug Resistance
HTS	High throughput sequencing
IN	Integrase
INSTI	Integrase strand transfer inhibitor
LMIC	Low middle income countries
LTR	Long terminal repeats
MA	Matrix
NC	Nucleocapsid
NNRTI	Non-Nucleoside Reverse Transcriptase Inhibitor
NRTI	Nucleoside analogue Reverse Transcriptase Inhibitor
PDR	Pre-treatment drug resistance

PI	Protease inhibitors
RAL	Raltegravir
RNA	Ribose nucleic Acid
RT	Reverse transcriptase
SS	Sanger sequencing
TAF	Tenofovir alafenamide
ETR	Etravirine
VGK	Viral growth kinetics
WHO	World Health Organization

1 INTRODUCTION

1.1 General Overview

1.1.1 Global Status of HIV-AIDS

According to the United Nations Program on HIV/AIDS (UNAIDS) Global AIDS Update (2017), nearly 37 million people are living with HIV globally, of whom 21.7 million people have access to combination antiretroviral therapy (cART) (http://www.unaids.org/en). This count is significantly higher than 15.8 million in 2015, 7.5 million in 2010, and less than one million in 2000. Due to global efficient administration of ART, the mortality rates were reduced significantly to 0.9 million in 2017 from 1.9 million in 2006 (WHO HIV update, July 2018). Persons in the age group 15-49 years have a global infection prevalence of 0.8%. The significant number of people living all around the world with HIV is a significant burden to global health.

1.1.2 Structure of HIV and genome

HIV-1 is two single-stranded, positive-sense RNA virus. Its genome with a size of 9.5 kilo bases contains nine genes. The genome is flanked by a non-coding long terminal repeat (LTR), both at the 5' and 3' end [1-3] (Figure 1.1B).

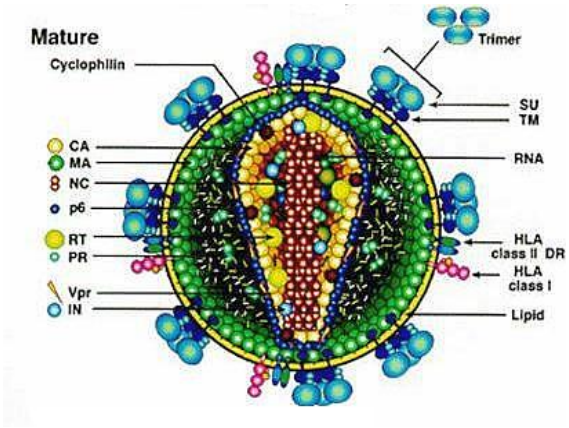
Among the nine genes, three major genes *gag*, *pol*, and *env* code for the structural proteins, viral enzymes, and envelope proteins, respectively. The structural proteins are named Capsid (CA), Nucleocapsid (NC), Matrix (MA), and p6. The *pol* gene encodes the Pol polyprotein and undergoes post-translational processing which results in three viral enzymes: Protease (PR), Reverse Transcriptase (RT), and Integrase (IN). The *env* gene encodes the polyprotein gp160, which is cleaved into gp120 and gp41. The outer glycoproteins facilitate viral entry into the host cell [1, 4-7].

In addition to the Gag, Pol and Env, several non-structural proteins are also produced. Tat and Rev are the regulatory proteins. Vpr, Vif, Vpu/Vpx, and Nef are the auxiliary proteins. Vpu and Vpx are unique for HIV-1 and HIV-2, respectively.

1.1.3 HIV-1 life cycle

The HIV-1 life cycle consists of several stages. The major steps are the entry of virus, followed by reverse transcription, integration, transcription, translation, budding and maturation. HIV-1 replicates very fast. About 10 billion nascent virions are produced in an HIV-1 infected individual and on an average, approximately 100 million cells per day get infected during acure infection. Extra-cellular mature virions first enter the host cells where viral genomic RNA is reverse transcribed into double-stranded DNA (dsDNA) by the action of RT. The IN then facilitates the next steps of transport of dsDNA into the nucleus and the integration of dsDNA into the host chromosome. In most cases, viral RNAs (structural and mRNA) are transcribed immediately from integrated viral DNA. Howerver, in a limited number of cells, the integrated DNA becomes dormant. Once transcribed, the viral RNAs are transported from the nucleus to cytoplasm and with a consequent translation of the mRNA into folded (precursor) proteins. The viral budding begins with the assembly of genomic RNA, Vpr, Vif, Nef, Gag and Env proteins and hence facilitates the formation of nascent HIV-1 particles. Finally, during the maturation stage, the viral protease cleaves the Gag and Gag-Pol polyproteins as well as the Env polyprotein [8, 9].

A.



B.

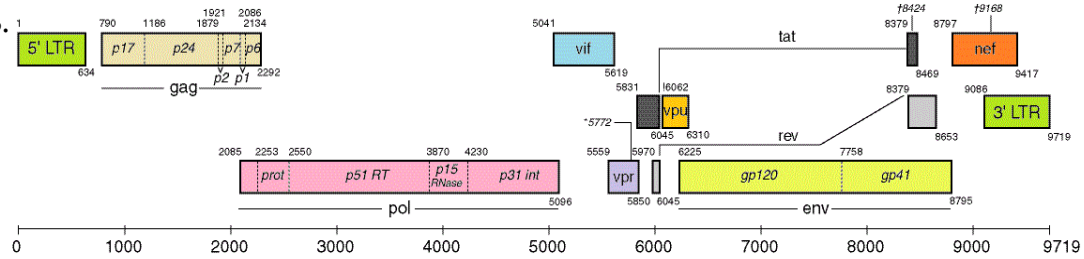


Figure 1.1 A) The model of mature HIV-1 with the coordination of structural and functional proteins. B) The structural and functional proteins orientation along with the corresponding genes of the matured HIV-1. The positions of genes are labeled with reference to HXB2 co-ordinate (Source: HIV Los Alamos Database, www.hiv.lanl.gov). Reprinted with permission.

1.1.4 HIV-1 subtypes and classification

One of the predominant features of HIV-1 that makes it so difficult to deal with is its genetic variability. Because of the lack of an effective proofreading mechanism of the RT, a high nucleotide substitution rate occurs during reverse transcription of the genomic RNA. This is one of the factors that contribute to the variability both within the same host and at a global level [10, 11]. Based on this genetic data variability, HIV-1 is classified into four groups: M (main), O (outlier), N (non-M, non-O) and another group called P (Putative) identified in Cameroon. Among these, group M is the leading one all over the world causing the majority of infections. The M-group is further classified into nine subtypes namely A, B, C, D, F, G, H, J, and K. Subtype C (HIV-1C) is the most common and is responsible for more than half of the infections worldwide. The second most predominant is subtype A [12-14]. The genetic features of the different subtypes could have implications for drug efficacy, but in general, this has not been considered to any large extent in the overall development of ART. To further characterize the potential impact of subtypes, it is worthwhile to consider their natural polymorphisms and any other resistance-mutations associated with drug usage [15, 16].

1.1.5 HIV-1 subtype C

HIV-1C was discovered in Ethiopia (HIV-$1C_{ET}$) in 1988 by Ayehunie et al. [17, 18]. It was also identified as the dominating strain in India (HIV-$1C_{IN}$). Today, HIV-1 is the most prevalent subtype globally and infects more than 50% of the global HIV population. It

seems to have originated in the 1970s from a single virus in South Africa (HIV-$1C_{ZA}$) [19]. HIV-1C demonstrates several interesting unique genotypic and phenotypic properties [20-22]. For example the existence of an additional NF-κB site in HIV-1C LTR, potentially provides a replicative advantage over other strains [23]. Also, a significantly higher number of PTAP-duplications was observed in HIV-1C infected patients than in HIV-1B infected individuals. A PYxE-insertion in Gag-p6 is found only in HIV-1C strains[24]. Distribution of PTAPd and PYxE among HIV-1B and HIV-1C is shown in Figure 1.2. A disparity has also been observed in drug resistance mutations (DRM), for example, the emergence of the important K65R mutation which is observed more frequent in therapeutic failure with HIV-1C than HIV-1B [25].

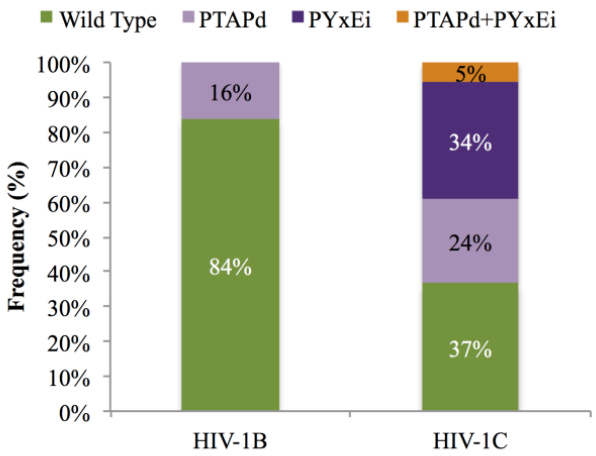


Figure: 1.2 Distribution of Gag p6 genotypes in HIV-1B and HIV-1C infected patients

1.2 Antiretroviral therapy

1.2.1 Approved antiretrovirals

Currently, there are 28 European Medicines Agency (EMA) approved drugs from six classes of drugs that target the virus at different stages of the life cycle. The drug classes and their generic names are described in Table 1.1. Genetic barrier to drug resistance, toxicity, adherence to ART, and the costs of treatment are the main parameters that influence the choice of ARV. Important aims for the development of new ARV is a high genetic barrier, diminished toxicity, and improved pharmacokinetics, to support improved long-term adherence and efficacy.

Integrase strand transfer inhibitors (INSTI) are a new class of approved ARVs, which rapidly have become a choice for first- and second line ART. It acts by inhibiting the IN thereby counteracting the integration of the viral DNA genome into the host cell's chromatin. Four drugs of this class are currently approved by EMA for use in humans: raltegravir (RAL), elvitegravir (EVG), dolutegravir (DTG), and BIC. A fifth INSTI drug CAB is under clinical trials. RAL and EVG have a low genetic barrier to drug resistance (DR), and they share a high degree of cross-resistance. The second-generation INSTIs DTG, BIC, and CAB have a substantially higher degree of the genetic barrier as compared to the first-generation INSTIs.

The development of ARVs against non- B subtype strains have not been extensively studied *in vitro*, *ex vivo* or *in vivo.* Therefore, a major and important concern is to develop effective anti retrovirals to achieve sustainable success of controling non-B subtypes .

Table 1.1 Approved antiretroviral drugs.

Entry inhibitors¤	NRTI	NNRTI	INSTI	PI
Enfuvirtide	Zidovudine	Nevirapine	Raltegravir	Saquinavir**
Maraviroc	Didanosine	Delavirdine*	Elviregravir	Indinavir**
Ibalizumab*	Zalcitabine**	Efavirenz	Dolutegravir	Ritonavir
	Stavudine**	Etravirine	Bictegravir	Nelfinavir
	Lamivudine	Rilpivirine		Amprenavir**
	Emtricitabine	Doravirine		Fosamprenavir
	Tenofovir DF			Lopinavir
	Tenofovir AF			Atazanavir
	Abacavir			Tipranavir
				Darunavir

¤Enfuvirtide: fusion inhibitor; maraviroc: CCR5 inhibitor; Ibalizumab: CD4-receptor binding inhibitor; *Approved by FDA but not EMA. **First generation ARV, not available any longer in Sweden. #DF= disoproxilfumarat; AF= alafenamide

1.2.2 Combination antiretroviral therapy

The general approach to avoid HIV-1 DR is to combine two or three ARVs. Presently it is most common to combine two NRTI and one second-generation INSTI, one PI/r (Ritonavir boosted PI) or one NNRTI. The strategy differs between high-income countries (HIC) and low-middle-income countries (LMIC). In HIC, the approach to ART regimen has a strong component of personalized medicine where treatment efficacy is monitored through measuring HIV RNA load, $CD4^+$ T-cell counts. If the viral load is high, HIV genotypic resistance testing (GRT) is performed. Important components of effective treatment are adherence support to the patients and involvement of the patient in his/her care. In many LMICs, common drug regimens are used for treating large populations instead of precision medicine, and the patients are rarely supported by physicians but rather educated by other health-care personnel or persons briefly educated about treatment. Also despite that WHO recommends HIV RNA load testing this is frequently not done due to lack of adequate resources. As a result, there is an increasing rate of both transmitted drug resistance, pretreatment drug resistance and acquired drug resistance in LMIC [26]

1.3 HIV drug resistance (DR)

A major threat towards the long-term outcome of ART is DR. DR develops when the virus is allowed to replicate at suboptimal therapeutic drug concentrations and thereby acquired drug resistance (ADR) develops. The most common reason is decreased adherence to the prescribed medications. More uncommon is drug-drug interactions or decreased absorption of one or more drugs leading to subtherapeutic drug concentrations. However, drug-resistant strains can also be transmitted, so-called transmitted drug resistance (TDR). Since it is often difficult to know, especially in LMIC, whether a patient has earlier been treated with ARV or not, the term pretreatment drug resistance (PDR) is often used. Thus, PDR may be due to

either TDR or earlier use of ARVs, which is unknown to the physicians before the start of first-line ART.

1.3.1 Error activity of RT and its role in creating the pool of quasispecies

The high replication rate of HIV-1 and the error-prone RT allows the virus to generate many new viral variants collectively called quasispecies [11, 27]. All possible mutations are generated every day in a patient, and thereby the quasispecies may have minor populations of drug-resistant variants even if the patient has never been treated [28-30]. However, if the patient is treated, but only suboptimal therapeutic drug concentrations are reached, the viral variants harboring drug resistance mutations (DRM) are selected for and may become the dominant species whereby the efficacy of the treatment are further decreased.

The impact of drug-resistant variant does not depend only on the frequency, but the mutational load (Mutational load = HIV RNA copies/mL × mutant variant frequency) better determines the absolute burden of the resistant variants [31, 32] . As a consequence, at high viral load, even <1% of minor DRM variant may overcome the threshold of drug pressure.

1.3.2 Genetic barrier of antiretroviral to HIV drug resistance

Different ARVs develops DR with different easiness and speed. It is well known that for some drugs, like 3TC, FTC, RAL, and EVG, only one mutation is sufficient to substantially decrease the efficacy of the drug. For PI/r a longer time on subtherapeutic drug concentrations are required to allow DR to develop, according to the present knowledge. For second-generation INSTI, DTG, and BIC, it was first considered that DR hardly developed at all. However, newer data show that DR can develop to these drugs also, probably even easier than to PI/r [33]. Initially DR causes a decrease in viral fitness, but with time further compensatory mutations appear that restore the decreased viral fitness and thereby allow the mutated virus to replicate at similar levels as the wild type virus.

1.4 Drug resistance testing for HIV-1

Drug resistance testing/establishing DRM libraries for HIV-1 can be performed either by phenotypic or genotypic methods. Genotypic resistance testing is widely used because of its low-cost implementation and usefulness in the clinical applications [34-38]. Phenotypic resistance testing requires high-cost laboratory setup and takes a longer time to perform. Its clinical usefulness is, therefore, limited [38, 39].

1.4.1 Genotypic resistance testing

GRT is done in the clinical routine praxis by sequencing a part of the HIV-1 genome (*pol*) whereafter the existence of previously known DRM is looked for. There are many methods to perform GRT, for example, Sanger sequencing, hybridization methods and various platforms of HTS. The most widely used method in clinical routine praxis at present is population-based Sanger sequencing, for its simplicity and reliability. But population-based Sanger sequencing cannot detect minor viral variants with DRM below 20% of the viral population. The detection rate of minor DRM variants by HTS is often set to 1-20%. However, the clinical cut off, where a resistant viral variant is clinically relevant, is presently unknown or disputed. It is likely that this is depending on the mutational load and class of drug. Hence, HTS is gaining importance for its sensitivity of detection down to 1% of the

viral population. The cost of HTS is now decreasing, but the implementation of computationally intensive bioinformatics pipelines still remains a challenge.

1.4.2 Phenotypic resistance testing

Phenotypic drug resistance testing is performed by infecting a cell line with the standardized inoculum of infectious virus particles in the presence of a range of concentrations of the antiretroviral drugs of interest. HIV DR is seldom an all-or-none phenomenon. Hence, "false" negative results are expected for certain drugs. While establishing the *in vitro* EC_{50} concentration for individual drugs, technical cut-offs/reproducibility and biological cut-offs are essential to define.

2 AIMS

The overall aim of my thesis was to study the impact of mutations in the HIV-1 genome on the sensitivity for newer antiretroviral drugs with special reference to HIV-1 non-B subtypes

Paper-I

To develop and evaluate an easy-to-use bioinformatics pipeline, MiDRM*pol*, integrating genomic variations and mapping of minor viral populations with drug resistance mutations in diverse HIV-1 subtypes after high throughput sequencing.

Paper-II

To determine the HIV-1 subtype specific antiretroviral efficacy *in vitro* of the novel drug EFdA.

Paper-III

To determine the HIV-1 subtype specific antiretroviral efficacy *in vitro* of the second-generation integrase strand transfer inhibitors DTG, CAB and BIC.

Paper-IV

To study the therapeutic response of HIV-1C harboring the PYxE insertion in Gag-p6 region.

3 MATERIALS AND METHODS

3.1 Study design and patient population

Plasma samples were collected from three different HIV-1 cohorts from Sweden (n=60), India (n=10) and Ethiopia (n=17). The Swedish cohort included ten treatment-experienced patients and the remaining patients in this thesis were treatment naïve.

3.2 Viral RNA extraction, RT PR and Nested PCR to amplify the *gag-pol*

The PCR protocol to amplify the complete HIV-1 genome was adopted from our previous publication [40] which can amplify from a wide range of HIV-1 subtypes. Briefly, viral RNA was extracted from 140 µl of plasma using QIAamp Viral RNAMini Kit (Qiagen, US) followed by cDNA synthesis using the gene-specific primer (6352R). The reaction was performed following the steps prescribed in the kit SuperScript III Reverse Transcriptase enzyme system (Invitrogen, US). RiboLock RNase inhibitor (40 U/µl; Thermo Scientific) was used in the reaction to inhibit RNase. The conditions for incubation are 25°C for 5 min followed by 55°C for 1h and finally 70°C for 10 min to terminate the reaction. This was followed by adding 1µl RNAseH (Invitrogen, US) and incubating at 37°C for 20 min.

First-round PCR was performed using 0682F and 6352R primers followed by the second round nested PCR using 0776F and 6231R primers, which yielded an amplicon of approximately 5.5 kb (Figure 3.1 A). High fidelity KAPA HiFi HotStart ReadyMix (2X) (KAPA Biosystem, MA, USA) with 15 pmol of each primer in 50 µl total reaction volume was used. PCR conditions for both the reaction was initial denaturation at 95°C for five minutes followed by 30 cycles of 98°C for 20 sec, 65°C for 15 sec and 72°C for 3 min and a final extension at 72°C for 5 min. Primer sequences are shown in Table 3.1.

Table 3.1 List of primers used for the amplification of F1 *gag-vpu*.

Primer_ID	Sequence (5′→3′)	HXB2 position
0682F	TCTCTCGACGCAGGACTCGGCTTGCTG	0682→0708
6352R	GGTACCCCATAATAGACTGTRACCCACAA	6352→6324
0776F	CTAGAAGGAGAGAGAGATGGGTGCGAG	0776→0800
6231R	CTCTCATTGCCACTGTCTTCTGCTC	6231→6207

3.3 High throughput sequencing- Illumina HiSeq2500

3.3.1 Processing of PCR products for HiSeq2500

The *gag-vpu* amplicons which were amplified using the above protocol were used to generate high throughput sequence as described in one of our previous publication [41]. Briefly, the F1 *gag-vpu* was gel-purified using QIAquick PCR purification kit (Qiagen, US) followed by fragmentation using Covaris S200 at 300bp for 75 sec with peak power- 50 and cycle/burst - 200. The library was prepared using NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (New England Biolabs, US) with multiplexed NEB next adaptors. The samples were then pooled together (either 24 or10) along with other unrelated non-viral indexed libraries. Pair-end sequencing of length 250 bp was carried out on the IlluminaHiSeq2500. Of

the 24 samples run, there was at least one culture supernatant as QC sample. To define the reproducibility, ten samples were re-sequenced.

3.3.2 MiDRM*pol* pipeline to analyze the fastq format raw file from HiSeq2500

Figure 3.1 B summarizes the bioinformatics pipeline. In brief, the raw reads were adapter trimmed using Cutadapt v1.8 program using the default setting (error rate <0.1) [42] followed by removal of the low-quality bases (phred value score <Q30) by Sickle ver 1.33 [43]. Duplicate reads were removed using FastUniq [44]. The processed reads were aligned against individual *pol* gene sequence in very sensitive local mode using Bowtie2 in order to select reads originated from *pol* gene [45]. The selected reads were then aligned against Pol protein sequence using blastx program from BlastX package [46]. The best blastx hit was chosen for each read for the amino acid counting which was performed by in-house script and the output is in easily interpretable table format.

3.4 Cloning of patient derived *gag-pol* amplicons in pNL4-3, generation of infectious virus and $TCID_{50}$ calculation

The *gag-pol* was amplified using modified primers from the first round PCR products discussed in section 3.1 as a template. Cloning of these amplicons in pNL4-3 was performed as described in Paper II and III [47, 48]. Briefly, the second-round PCR was performed using BssHII0682F (5'CTTGCTGAAGCGCGCACGGCAAGA3') and SalI5798R (5'CTGCTA TGTCGACACCCAATTCTGAAATGGATAA3') which has BssHII and SalI restriction sites, respectively. The amplified product was restriction digested followed by gel purification using the QIAquickGel Extraction Kit (Qiagen). The *gag-pol* fragment (HXB2: 0702-5798) was cloned in pNL4-3Δ*gag-pol* plasmid following digestion with BssHII and SalI (New England Biolabs, USA) and ligation using T4 DNA ligase (New England Biolabs, USA). The recombinant viruses were produced by transfecting the plasmids with HEK 293T cells using FuGENE HD Transfection Reagent (Promega, Madison, Wisconsin, USA). All the molecular clones were sequenced bi-directionally for the confirmation of proper insert. No primary DRM was observed in any of the clones' sequences. M50I was present in seven of the sequences from the HIV-1B (n=3) and the HIV-1C (n=4) strains. The tissue culture infectious dose at 50% ($TCID_{50}$) of the viruses produced was determined in TZM-bl cells using the Spearman-Karber method.

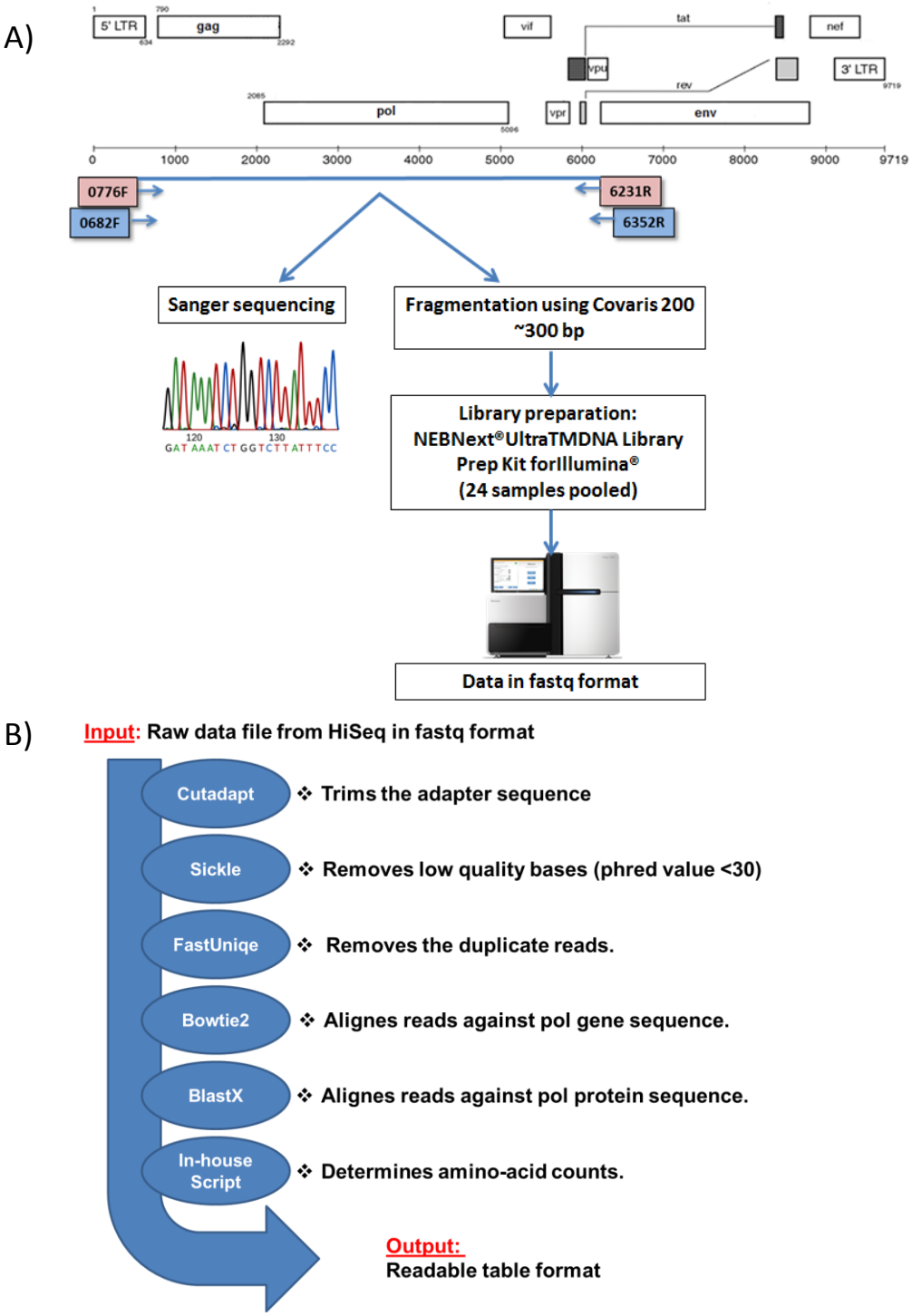


Figure 3.1. A) Schematic of *gag-pol* amplification from HIV-1 genome followed by Sanger sequencing / library preparation for HTS. B) Pipeline for the analysis of Fastq files generated from HiSeq2500.

3.5 Phenotypic Drug sensitivity assays

New generation drugs were compared with the earlier approved drugs *in vitro* against patient-derived *gag-pol* chimeric viruses from diverse subtypes (B, C, 01AE, and 02AG). The phenotypic drug sensitivity assay was performed as described in Paper-II and III [47, 48]. The drugs used for the experiments were EFdA, nevirapine (NVP), efavirenz (EFV), tenofovir alafenamide (TAF), etravirine (ETR), rilpivirine (RPV), raltegravir (RAL), elvitegravir

(EVG), DTG, CAB, and BIC. Briefly, TZM-bl cells were treated with serial dilutions of the drugs spanning from 10 to 0.0001mmol in triplicates in 96 well plate, followed by infection with a reference virus (NL4-3) or the corresponding patient-derived recombinant viruses (HIV-1B: n=6; HIV-1C: n=14; HIV-1A1: n=1; HIV-1CRF_01AE: n=1; HIV-1CRF_AG: n=2) at 0.05 MOI (multiplicity of infection) in the presence of 10 µg/ml concentration of diethylaminoethyl (DEAE) -dextran. Virus infection was quantified by measuring luciferase activity in terms of relative light units (RLU) using the Bright-Glo Luciferase Assay System (Promega) after 48 hr. Drug concentration required for virus inhibition by 50% (EC50) was calculated by a dose-response curve using nonlinear regression analysis (GraphPad Prism, version 5.01; GraphPad Software, La Jolla, California, USA). Each assay was performed using two to four biological replicates in triplicate.

3.6 *In vitro* virus replication kinetics assay

In vitro replication kinetics for the chimeric virus was performed as described in Paper-IV [24]. Briefly, three HIV-1C samples and one HIV-1B (control) were used to construct p2-INT-recombinant viruses (gag-p2/NCp7/p1/p6/pol-PR/RT/IN) in HIV-1NL4-3 backbone as described previously [49]. These viruses were used in analysis of the viral growth kinetics (VGKs) 12 days post-infection in MT4 cells (i.e., 1 X 10^6 cells at 0.01 MOI) and the replication was quantified by measuring (i) RT activity (Lenti RT activity kit, Cavidi, Sweden), (ii) p24-Gag antigen (PerkinElmer, Waltham, MA) and (iii) Virus yield based on TZM-bl cells based assay as described previously [50]. All experiments were performed in triplicate. Cell viability was determined by infecting (1 X 10^4) MT4 cells (Day 5 post-infection) and PHA-stimulated PBMCs (Day 7 post-infection) at 0.05 MOI virus, using alamarBlue Cell Viability Assay (Invitrogen, Carlsbad, CA).

3.7 Ethical Considerations

Study participants had given informed consent to participate in concordance with the Declaration of Helsinki. The studies were approved by the Regional Ethical Council in Stockholm for Swedish samples (Dnr: 2006/1367-31/4, for Paper-IV: 2005/1167-31/3), the Institutional Ethics Committee of the National Institute for Research in Tuberculosis (formerly known as Tuberculosis Research Centre ICMR; TRC-IEC No: 2009009) for Indian samples, and the Institutional Review Board Office, Faculty of Medicine, Addis Ababa University for the Ethiopian samples (Ref. No. 02/6/22/17).

4 RESULTS

Paper-I

As expected, HTS detected additional DRMs compared to population-based Sanger sequencing. The method is reliable with a cut-off value of 1% since the error rate for HTS is <1%. HTS followed by analysis using MiDRM*pol* detected and quantified DRM in 23 out of 87 samples. In contrast, population-based Sanger sequencing identified DRM in only three samples. The results from MiDRM*pol* and the comparison with PASeq output as well as DRM detected using population-based Sanger sequencing is summarized in Table 4.1. In one of the eight mother-to-child transmission (MTCT) samples, eight major RTI-DRM were detected at frequencies < 20%.

Paper-II

The *in vitro* drug sensitivity assay using 24 recombinant viruses generated from diverse HIV-1 subtypes revealed that the potency [median EC_{50} (IQR)] of EFdA [1.4 nM (0.6–2.1)] was comparable to that of TAF [1.6 nM (0.5–3.6)]. Subtype-specific differences in EC_{50} were observed for ETR (p=0.004) and RPV (p=0.017), but not for EFdA (Figure 4.1).

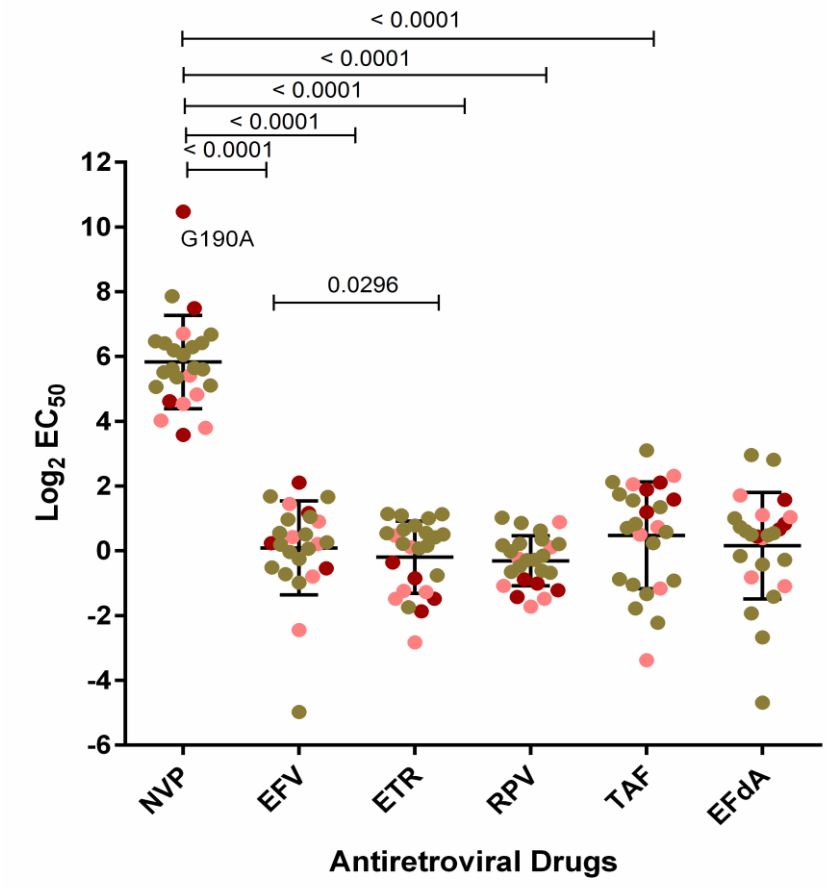


Figure 4.1 EC_{50} values of different antiretroviral drugs. Effects of EFdA and other RTIs on 24 recombinant viruses representing HIV-1B (n=6) – pink, HIV-1C (n=14) – green, and HIV-1A-like (A1, 01_AE and 02_AG; n=4) – maroon. Each dot represents the EC_{50} value for each virus.

Table 4.1 Comparison of DRMs identified by high-throughput sequencing or population sequencing as well as in PASeq and MiDRM*pol* outputs

Cohort	PID	Treatment	MiDRM*pol*				PAseq				GRT-PS	
			PI	NRTI	NNRTI	INI	PI	NRTI	NNRTI	INI	NRTI	NNRTI
Swedish	**SE03**	Experienced	None	**D67G (2.0%)** **L74I (1.9%)**	None	None	None	None	**K103R (16.9%)**	None	None	None
	SE70	Experienced	None	**D67N (99.7%)** **L210W (99.5%)** **T215Y (99.4%)**	**K103N (99.8%)** **G190A (99.6%)**	None	None	**D67N (99.9%)** **L210W (99.8%)** **T215Y (99.8%)**	**K103N (99.9%)** **G190A (99.7%)**	None	**D67N** **L210W** **T215Y**	**K103N** **G190A**
	SE73	Experienced	None	**D67N (44.2%)** **L210W (23.4%)** **T215Y (19.7%)**	**K103N (95.9%)** **G190A (35.2%)**	None	None	**D67N (49.2%)** **L210W (46.1%)** **T215Y (45.9%)**	**K103N (96.1%)** **G190A (49.6%)**	None	None	**K103N** **Y188C**
	19C	Naïve	**M46I (4.6%)**	None	None	None		Failed			None	None
	20C	Naïve	None	**T215S (2.1%)**	None	None		Failed			None	None
	26C	Naïve	None	**M184I (1.1%)**	None	None		Failed			None	None
	05C	Naïve	None	**K219R (2.3%)**	None	None		Failed			None	None
	SE07	Naïve	None	None	None	**Q148H (1.1%)**	None	None	None	None	None	None
	SE14	Naïve	None	None	None	**E138K (5.5%)**	None	None	None	**E138K (1.6%)**	None	None
	SE31	Naïve	**M46I (30.2%)**	None	None	None	**M46I (28.9%)**	None	None	None	None	None
	SE97	Naïve	**D30N (13.7%)**	None	None	None	**D30N (12.2%)**	None	None	None	None	None
	24C	Naïve	None	**T69D (99.7%)**	**M230L (99.7%)**	None	None	**D67E (77.2%)** **T69D (98.1%)** **T215A (99.4%)**	**M230L (99.8%)**	None	**T69D**	**M230L**
Indian	**IN04**	Naïve	None	None	**K101E (99.7%)**	None	None	None	**K101E (99.9%)**	None	NA	NA
	IN05	Naïve	None	**V75M (7.8%)**	None	None	None	**V75M (8.6%)**	None	None	NA	NA
	IN06	Naïve	None	None	**V106M (98.1%)** **Y181C (22.8%)**	None	None	None	**V106M (100.0%)** **Y181C (23.6%)**	None	NA	NA
	IN07	Naïve	None	**K65R (3.0%)** **K70E (3.4%)** **L74V (66.1%)** **M184V (7.5%)** **M184I (3.7%)** **T215I (3.4%)**	**V106M (3.9%)** **G190S (95.6%)**	None	None	**K65R (2.8%)** **K70E (3.2%)** **L74V (67.6%)** **M184V (8.4%)** **M184I (3.2%)** **T215I (3.3%)**	**V106M (3.6%)** **G190S (96.1%)**	None	NA	NA
	IN08	Naïve	None	None	**Y181C (99.5%)**	None	None	None	**Y181C (99.9%)**	None	NA	NA
Ethiopian	**ET160**	Naïve	**L76V (4.5%)**	None	None	None	**L76V (4.4%)**	None	None	None	None	None
	ET115	Naïve	None	**M184I (1.1%)**	None	None	None	None	None	None	None	None
	ET122	Naïve	None	None	**M230L (4.8%)**	None	None	None	None	None	None	None
	ET155	Naïve	None	None	None	**T66I (3.3%)**	None	None	None	**T66I (5.8%)**	None	None
	ET159	Naïve	None	**M184I (1.0%)**	None	None	None	None	None	None	None	None
	ET171	Naïve	None	**M184I (1.5%)**	None	**Q148H (1.2%)**	None	None	None	None	None	None

Note: None, no DRM detected; Failed, failure due to upload limit of HTS file; NA, not available.

Paper-III

The *in vitro* drug sensitivity assay using the 24 recombinant viruses revealed that the EC_{50} values for DTG, CAB, and BIC were significantly lower than those for RAL and EVG (Figure 4.2). The median EC_{50} (IQR) (nM) values of the drugs for all the 24 recombinant viruses tested were: RAL - 5.31 (1.56–6.70), EVG - 2.82 (1.69–5.39), DTG - 2.14 (1.3–2.56), CAB - 1.68 (1.34–2.55), BIC - 1.07 (0.22–2.53). No significant difference in the EC_{50} values was observed between the HIV1A-like, HIV-1B and HIV-1C for RAL, EVG, CAB, and BIC, respectively. However, DTG showed a higher potency for HIV-1A-like virus.

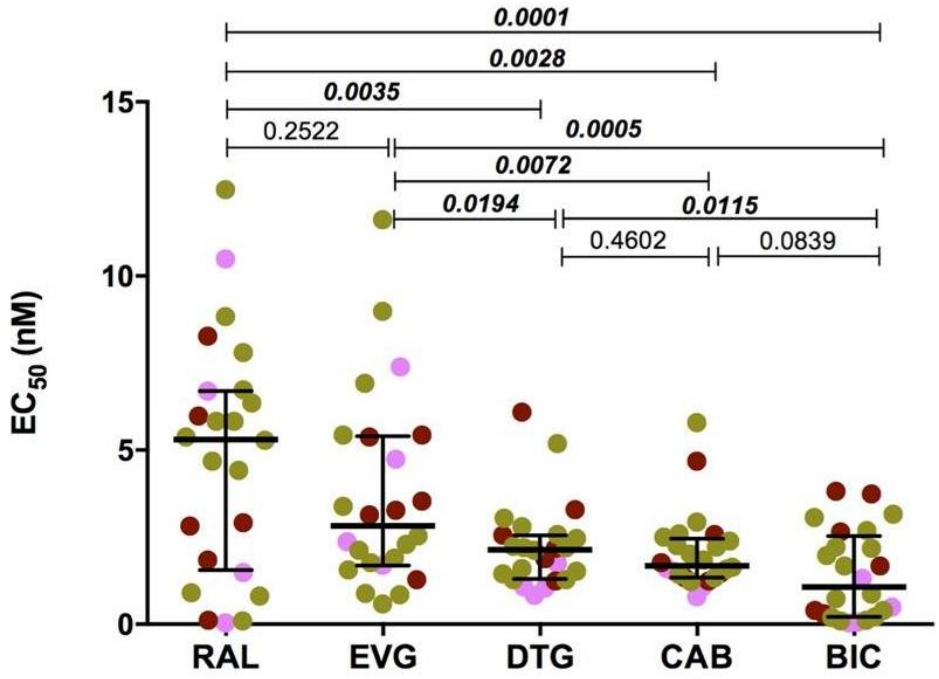


Figure 4.2 *In vitro* antiretroviral activity of INSTI. Wilcoxon signed-rank test was used to perform pair-wise analysis of the antiretroviral activity of RAL, EVG, DTG, CAB and BIC, respectively on recombinant viruses obtained from INSTI naive patients. The strains analysed were HIV-1B (n=6) – pink; HIV-1C (n=14) – green; HIV-1A-like (A1, 01_AE and 02_AG; n=4) – maroon.

Paper-IV

The higher rate of treatment failure and lower pre-therapy CD4+ T-cell counts were observed in HIV-$1C_{EA}$ infected patients compared to HIV-1B-infected patients. HIV-$1C_{EA}$–infected patients with signature C_{PYxEi}, evidenced very low pre-therapy CD4+ T-cell counts and suboptimal gain in CD4+ T-cells following therapy, as compared to the non-C_{PYxEi}-strains indicating higher virulence.

Viral growth kinetics, showed a higher replication capacity (RC) for the C_{PYxEi} viruses than the other two non-C_{PYxEi} strains as evidenced by high titer of functional RT activity on day 7 of infection (Figure 4.3A). Significantly increased *in vitro* pathogenic effect was observed in cell viability assay in HIV-1 C_{PYxEi} compared to wild type HIV-$1C_{WT}$ (Figure 4.3B).

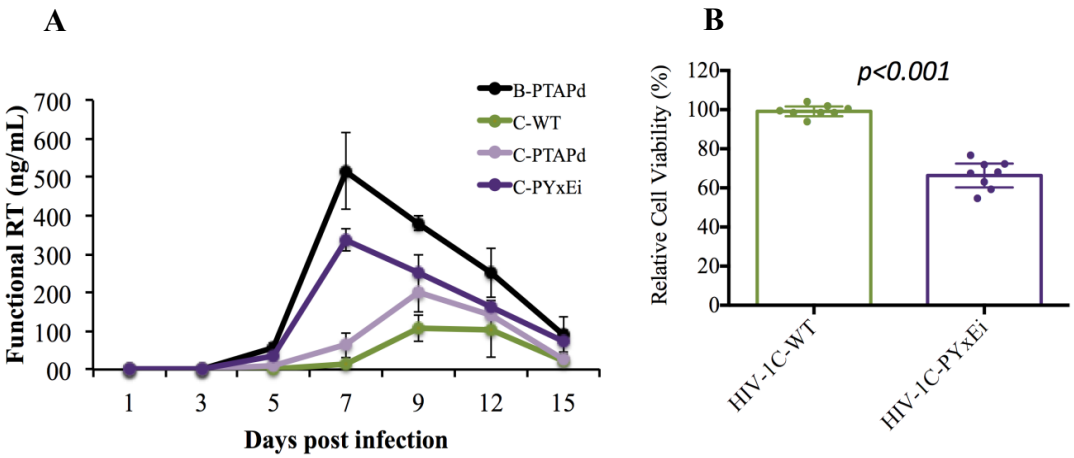


Figure 4.3 Viral growth kinetics and cell-viability assay of the p2-INT-recombinant viruses (gag-p2/NCp7/p1/p6/pol-PR/RT/IN). A) Viral growth kinetics of the p2-INT-recombinant viruses (gag-p2/NCp7/p1/p6/pol-PR/RT/IN) in MT4 cell-lines measured by functional RT assay. B) Cell-viability assay following 0.05 MOI infections in 10,000 MT4 cells. A statistically significant higher cell death was observed on five days post infection in HIV-1C_{PYxEi} viruses than wild type HIV-1C viruses.

5 DISCUSSION

The global use of ART has resulted in pronounced clinical advantages with decreased morbidity and mortality. However at the global level, both pretreatment and acquired HIV-1 drug resistance (HIV-DR) are increasing (https://www.who.int/hiv/topics/drugresistance/protocols /en/). Since the development of ARVs has to a large extent been based on HIV-1B, the knowledge about HIV-DR in non-B subtypes is limited. My thesis discusses some of the existing knowledge gaps in the identification and treatment of HIV-1 drug resistance.

Identification and quantification of drug-resistant viral populations before initiation of ART reduces the risk of developing virological treatment failure and acquired DRMs in patients who have PDR. In HIC, population-based Sanger sequencing has been used for a long time, both before the initiation of first-line ART and at ART failure. More recently, HTS has been developed and has so far been introduced into clinical care in a few laboratories. A major issue concerning HTS is the identification cut-offs of clinically relevant resistance for minor viral variants. Recently, a clinically relevant cut off of 5% for NNRTI resistant strains has been suggested, but there is limited knowledge for other categories of ARVs [51]. In LMIC, HIV-DR diagnostics are seldom performed in the clinical setting, but important HIV-DR surveillance programs exist, e.g., run by WHO [51].

To facilitate the analysis of viral sequences obtained through HTS in patients from both HIC and LMIC, we developed a user-friendly bioinformatics pipeline to bridge the gap between clinicians and bioinformatics expertise (**Paper I**). The output file from the high-throughput platform can be analyzed by uploading it in the MiDRM*pol* pipeline to get the final output in an easily readable table format. The overall process takes less than 10 minutes and does not require the help of onsite bioinformatics expertise. MiDRM*pol* can allow the user to choose a subtype-specific reference sequence (A, B, C, D, 01_AE, 02_AG) while variant calling followed by quantification of DRMs.

Due to the increasing global occurrence of HIV-DR and due to side effects of the available ARVs, it is of importance to develop new ARVs, preferable with a high genetic barrier and less toxicity. EFdA is one among the promising new drugs presently entering clinical phase II and III trials. The high genetic barrier of EFdA is due to its strong interactions with the active site of the RT where DRM are prone to develop. In **Paper II**, we studied its *in vitro* potency against diverse HIV-1 subtypes and showed a high potency of EFdA against several HIV-1 RTI resistant strains (NIH panel) derived from different HIV-1 subtypes. We conclude therefore that EFdA has a high inhibition potency, irrespective of the subtype. In addition, we tested EFdA against two of the viruses (SE138.4 and SE141.5) from the Swedish cohort which had the K65R mutation and found a higher efficiency (0.8 and 3.4 fold respectively) than TAF. This finding is in accordance with another study, where hypersusceptibility for EFdA was observed in virus with the K65R mutation in patients, who were failing on first-line TDF ART [52]. Thus, although EFdA is not yet approved by regulatory authorities, our data suggest that it can be a valuable addition to the presently used ARVs, both in HIC and LMIC.

The second-generation INSTIs DTG is presently rolled out on a large scale both in HIC and LMIC. However, most existing knowledge about this drug is derived from studies on HIV-

1B. Also, the former view that HIV-DR against DTG develops only with very great severity has shown not to hold through [53]. The knowledge about the two other second-generation INSTIs, CAB and BIC with regard to HIV-DR pattern is also limited. However, recently DRM was shown in some patients failing long-acting CAB and RPV within a clinical trial [54]. In **Paper III**, we studied the potency of first and second-generation INSTIs against four HIV-1 subtypes (B, C, 01_AE and 02_AG) using biochemical and cell culture assays. The integration of HIV-DNA into the host genome occurs in two major steps: 3'-end processing followed by strand transfer reaction [55]. Our results from the biochemical assays confirmed that all INSTIs are not as efficient in inhibiting 3'-end processing as compared to strand transfer activity for all subtypes, which is in agreement with previous reports for RAL, EVG [56], DTG [57] and BIC [58]. Interestingly, a natural polymorphism M50I was found in 7/24 (18%) viruses, which is in line with the reported frequency of 10-25% of INSTI-naïve patients in different subtypes according to Stanford HIV drug resistance database [59]. This polymorphism has been observed in combination with R263K in a patient who subsequently failed treatment with RAL [60]. R263K has been shown to increase resistance to DTG by 2-5 folds and decrease viral replication and strand transfer activity [61]. In viral outgrowth assay, M50I emerges after R263K [58]. Whether the presence of M50I provides a replication advantage to R263K[58] or not[62] is still under debate. The combination of R263K/M50I results in 2.8 fold reduction in BIC-susceptibility, but M50I alone does not have any effect [58]. In our *in vitro* experiments in Paper III, M50I had no effect on any of the INSTIs used.

One of the solutions for the problem of patients suboptimal adherence to ART is to develop and implement long-acting drugs to allow flexibility in dosing of at least monthly and perhaps longer. Due to a prolonged intracellular half-life of EFdA active moiety in human blood cells, it is potentially possible for this drug [63, 64]. A further two long-acting injectable antiretroviral compounds, CAB and RPV, have completed clinical trials demonstrating safety, tolerability and prolonged antiretroviral activity. In **Paper II** and **III** we showed that EFdA and CAB had potent activity against diverse HIV-1 subtypes. In contrast, we have identified that RPV may not be an optimal choice for HIV-1C [65]. Of these three anti-HIV drugs, which have the potential of being used as long-acting drugs, we feel a concern about the use of RPV where HIV-1C dominates the epidemic. However, an even higher threat against the efficacy of long-acting regimens is the high risk of developing HIV-DR to these drugs if the patient does not turn up for the regular therapeutic injections. This risk is based on the pharmacokinetic properties of the drugs where the elimination of the drugs takes a very long time, during which suboptimal drug concentrations will appear [66].

In **Paper IV**, we showed that virus with PYQE insertion had increased replication capacity compared to non-PYxE-strains *in vitro*. PYxE (where x =R/K/Q) insertion in the P6 region of the gag is one among the unique features of HIV-1C, which was reported for the first time by Neogi et al. [22]. Both the PYQE variant of PYxE in our experiments and the PYRE inserted HIV-1C reference virus (pIndieC1) in another study [67] showed higher replication capacity. The PYxE insertion is prevalent among HIV-1C from Ethiopia and Eritrea compared to South African or Indian HIV-1C strains [22]. Studies have also shown that PYxE insertion is more common in viruses from HIV-1C therapy-failure patients than in HIV-1B viruses from failing patients [22]. The mechanism has been studied by Domselaar et al., [68] where the *in silico* and *in vitro* experiments show that the PYxE insertion in gag-p6 of HIV1-C increases its binding to ALIX and enhances the viral fitness. It was observed that the increased fitness

related to PYxE could affect the sensitivity against the protease inhibitor lopinavir in the absence of any PI DRM. However, a more detailed analysis of the clinical significance of the PYxE insertion within HIV-1C gag is needed.

In conclusion, quantification of viral variants, mainly minor populations in the pool of HIV quasispecies within a patient, is a promising leap to improve the selection of optimal first-line ART, which is likely to result in fewer ART failures. There is a continous development of methods and analysis tools for the determination of HIV-DR, also with the purpose to reduce the cost of sample processing and analysis. MiDRM*pol* is one such tool based on our intentions to keep it user-friendly and free of charge, as explained in **Paper I**. The new ARV EFdA shows promising results in suppressing the viral load in *in vitro* experiments against diverse subtypes of HIV-1, also in multi-resistant strains, as well as in HIV-2 as demonstrated in other studies. **Paper II** indicates that EFdA is likely to be as efficient in non-B subtypes as in HIV-1B although this should be evaluated in more extensive clinical studies, including patients infected by diverse HIV-1 subtypes harboring different pattern of DRMs. Similarly, in **Paper III** the second-generation INSTIs CAB and BIC suppressed diverse viral variants *in vitro,* and the data give support to their future use in countries dominated by non-B subtypes. Along with the clinically well-defined DRM in the *pol* region described in **Papers I, II and III**, there are also other mutations and insertions that may or may not alter the drug susceptibility, but that also have the potential to increase the viral replication capacity increasing the risk of negative clinical consequences. One such genetic change is the PYxE insertion which is unique to HIV-1C as described in **Paper IV**. Altogether my thesis and the **Papers I-IV** show the importance of considering the efficacy of ARVs on all HIV-1 subtypes and also include analysis of viral variants in the minor viral populations. If this is done I believe that the impact of ART on the global HIV epidemic will be even larger than it is been so far.

6 CONCLUSIONS

- MiDRM*pol* is a computational as well as labor efficient bioinformatics tool that can be used without any prior knowledge in the analysis of sequences obtained by HTS and does not require onsite bioinformatics expertise.
- MiDRM*pol* facilitates the identification of DRM in diverse HIV-1 subtypes.
- Combined virological and biochemical data suggest that EFdA inhibits both wild type and RTI resistant viruses efficiently in a subtype-independent manner.
- The second-generation INSTIs, DTG, CAB, and BIC have a higher antiretroviral potency than the first generation RAL and EVG *in vitro*.
- The comparable or higher potency of DTG, CAB, and BIC against non-B subtypes supports their suitability for use in countries dominated by non-B subtypes.
- HIV-1C viruses harboring the PYxE insertion in the gag p6 region are more replication competent than non-PYxE viruses *in vitro*, which possibly could lead to an increased risk of accumulation of DRM and suboptimal CD4+ T-cell gain during treatment failure *in vivo*.

7 FUTURE PERSPECTIVES

My research has filled some knowledge gaps with regard to antiretroviral therapy and its efficacy on diverse HIV-1 subtypes but has also raised questions and need for further investigations.

Increasing migration has led to an increased circulation of recombinant HIV-1 strains globally. Therefore there is a need for improving the methods of genotypic resistance testing to consider the viral diversity better. The cost of high throughput sequencing is also presently too high to implement this method in the clinic or large surveys in LMIC.

Due to the increased global HIV-1 diversity, new potent antiretrovirals should be evaluated on a broad spectrum of HIV-1 subtypes before and after approval. Also, chemically modified forms of these drugs to increase their half-life in plasma could reduce the problem of adherence, but the high risk of resistance must be considered if patients are lost to follow up when using such new formulations.

Most of the routine HIV patient monitoring before initiation of ART and at ART failure involves population-based Sanger sequencing of *pol* region for detection but not the quantification of drug resistance mutations. With the new technical era of high throughput sequencing, the cut-offs for clinical relevant resistance mutations should be identified for each drug.

While comparing the drug resistance patterns with a reference strain, up to now the reference strain has been NL4-3, an HIV-1B strain. pMJ4 and pInide-C are available to reference strains for HIV-1C, but they produce low titer virus, limiting their usage in *in vitro* or *ex vivo* experiments even though we were successful in constructing gag-pol chimera viruses, a more efficient method for whole genome molecular cloning from other subtypes than HIV-1B is still necessary for *in vitro* studies involving non-B HIV-1 subtype virus.

8 ACKNOWLEDGEMENTS

Supervisors and co-supervisors:

In the beginning, I would like to express my sincere, hearty thanks to my main supervisor **Prof. Anders** to guide me well throughout my PhD. He is the main support for all the wise decisions in the lab. With his immense experience in medicine and research, the firm and crystal clear decisions were practical. Even though the meetings were quite less often with him, he keeps the update of happenings in the lab being in tune with everyone. The beautiful messages he conveys even in the middle of the meetings and while giving the toast are priceless.

Dr.Ujjwal, my immediate boss, who is very knowledgeable, influenced a large part of my research and career life. He has a huge passion for science and research. The best thing about working with him is we get to see the suitable articles at the time we need, for example: while developing the assays for Drug sensitivity assay- we started talking about that, and he gave papers related to methods for one round assays- I made a detailed protocol (well this was not my habit, but experienced that it is really good to make one, and it stays forever). Following this we start to talk about the assaysfor protease inhibitors-which involves some important changes, instantly he suggests the articles which use the two round assays, very happy- for me, he is a walking PubMed.

I feel blessed to work with both Anders and Ujjwal. Their immense knowledge, motivation, and patience have given me more spirit to excel in the research and writing. Conducting the research couldn't be as simple as they made this for me.

I want to express my sincere thanks to co-supervisors **Prof. Francesca Chiodi** and **Prof.Vinayaka Prasad** for editing and reviewing my thesis and to be a part of my PhD thesis.

Lab mates and others

I express my thanks and regards to all the fellow lab mates and neighboring for keeping a lovely and healthy environment. **Duncan, Ashok,** and **Anoop** - for their friendly conversations and brotherly advices. I, Duncan and Ashok was working on similar kind of experiments, sometimes I spent a lot of time to do some experiments being unnecessarily extra careful, but when we started working together, it was really helpful to fix those. Altogether it was a wonderful collaboration. I thank **Sara, Maike, Xixi, Hema, Giota, Lydia**, **Katie, Noelia** for their friendly company, **Robert,** for reviewing my thesis, **Babi**- for helping with experiments and general lab routines. I also thank the internship students **Abu, Nayon, Morshed, Mira,** and **Evelyn** for the opportunity for me to teach and their help with experiments. **Amanda**-for her help in statistical analysis. **Jessica, Lena, Lise,** and **Piotr**- for their support in sample collections and processing. I thank **Bai** for her yoga sessions and **Soham** for his advice on the technical part of experiments.

Staffs: I appreciate **Arja, Marita,** and **Gudrun** for their friendly help with administrative work.

Previous mentors and lab mates

I would like to express my sincere gratitude to my first mentor and supervisor **Prof. Udaykumar Ranga**, from whom I was introduced to research and the other group members of my previous lab – **Shilpee, Anjali, Prabhu, Mahesh, Malini, Rebu, Rishi, Arun, Sutanuka** and **Deepak** for the collaborations, valuable scientific/personal suggestions and support during my tenure.

My teachers and buddies

Extra-curricular activities are something which keeps me energetic and curious about arts. I thank **Thyagarajan Sir** for his blessings of musical knowledge and my Sensei **J Bhale Rao, Ken Bokelius** and **Joachim von Rost** for their priceless training sessions in martial arts. I miss their training when I move to different places. Their training has been with me all the time and would continue to be forever. I convey my most profound respect to all of you.

I feel lucky to be in contact with all my master's batch mates, whom I can never forget. **Chetan, Jagga, Priyanka, Pooja, Gyanendra, Anudeep, Shobith, Chandini, Mitsu, Mounika, Arin and Anusha.** Their achievements and success continue to inspire me. I also feel blessed for the suggestions of **Dr.N. S. Devaki** mam who is always in touch with her students.

Not to forget friends from Vårbary hostel-**Rodrigo, Emilika, Shuvo, Shobhana, Quim, Julen** and **Xin** for the delicious international dinners. I picked various international cooking styles and cocktail recipes from all of them. I am a foodie and cocktail lover. I also thank my music friends **Praveen** and **Parvin** for their jamming sessions.

Others: I thank the anonymous reviewers of the papers for their critical analysis and questions.

My sincere thanks to **doctoral student association** and **Medicinska Föreningen** at KI who provides useful information through seminars and their support.

Grants and funding:

I acknowledge the following grants and funding: A part of the study was funded by grants from the Swedish Research Council (2016-01675) and Stockholm County Council (ALF 20160074).

Several reagents were obtained from the NIH AIDS Reagent Program. Part of the cloning and recombinant virus production was supported by the National Institute of Health RO1 grant GM118012.

I want to thank ViiV Healthcare/GSK and Gilead Sciences for providing CAB and BIC respectively.

I want to thank all the patients, nurses and clinicians who supported the InfCare system.

Family: My lovely parents **Appaji** and **Amma** for their everlasting care, best support, and understanding during tough times. Sisters, **Suneetha, Sucheta**, brother-in-laws **Manju bhava** and **Adi bhava** who are always happy and proud of my achievements, no matter how small it may seem to me. My nephews, **Shlok** and **Arav** for their love. **Arundhati** aunt, who encouraged me to read literature which has shaped a large part of my thoughts. **Basavaraj mama** and **Savitratte** for their encouragement and very practical conversations. I love you all so much !… ☺

9 REFERENCES

1. Turner BG, Summers MF. **Structural biology of HIV**. *Journal of Molecular Biology* 1999; 285(1):1-32.

2. Ganser-Pornillos BK, Yeager M, Sundquist WI. **The structural biology of HIV assembly**. *Current Opinion in Structural Biology* 2008; 18(2):203-217.

3. 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(7256):711-U787.

4. O'Doherty U, Swiggard WJ, Malim MH. **Human immunodeficiency virus type 1 spinoculation enhances infection through virus binding**. *Journal of virology* 2000; 74(21):10074-10080.

5. Kaushik R, Ratner L. **Role of human immunodeficiency virus type 1 matrix phosphorylation in an early postentry step of virus replication**. *Journal of virology* 2004; 78(5):2319-2326.

6. Samal AB, Ghanam RH, Fernandez TF, Monroe EB, Saad JS. **NMR, biophysical, and biochemical studies reveal the minimal Calmodulin binding domain of the HIV-1 matrix protein**. *The Journal of biological chemistry* 2011; 286(38):33533-33543.

7. Giagulli C, Magiera AK, Bugatti A, Caccuri F, Marsico S, Rusnati M, et al. **HIV-1 matrix protein p17 binds to the IL-8 receptor CXCR1 and shows IL-8-like chemokine activity on monocytes through Rho/ROCK activation**. *Blood* 2012; 119(10):2274-2283.

8. Coffin JM. **HIV pathogenesis. Lines drawn in epitope wars**. *Nature* 1995; 375(6532):534-535.

9. Martin N, Sattentau Q. **Cell-to-cell HIV-1 spread and its implications for immune evasion**. *Current opinion in HIV and AIDS* 2009; 4(2):143-149.

10. Goodenow M, Huet T, Saurin W, Kwok S, Sninsky J, Wain-Hobson S. **HIV-1 isolates are rapidly evolving quasispecies: evidence for viral mixtures and preferred nucleotide substitutions**. *Journal of acquired immune deficiency syndromes (1999)* 1989; 2(4):344-352.

11. Mansky LM. **Retrovirus mutation rates and their role in genetic variation**. *Journal of General Virology* 1998; 79 (Pt 6):1337-1345.

12. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, et al. **HIV-1 nomenclature proposal**. *Science* 2000; 288(5463):55-57.

13. Rambaut A, Posada D, Crandall KA, Holmes EC. **The causes and consequences of HIV evolution**. *Nature Reviews Genetics* 2004; 5(1):52-61.

14. Hemelaar J, Gouws E, Ghys PD, Osmanov S, Isolation W-UNfH, Characterisation. **Global trends in molecular epidemiology of HIV-1 during 2000-2007**. *AIDS (London, England)* 2011; 25(5):679-689.

15. Wainberg MA, Brenner BG. **The Impact of HIV Genetic Polymorphisms and Subtype Differences on the Occurrence of Resistance to Antiretroviral Drugs**. *Mol Biol Int* 2012; 2012:256982.

16. Langs-Barlow A, Paintsil E. **Impact of human immunodeficiency virus type-1 sequence diversity on antiretroviral therapy outcomes**. *Viruses* 2014; 6(10):3855-3872.

17. Ayehunie S, Johansson B, Sonnerborg A, Salminen M, Zewdie DW, Fehniger T, et al. **New subtype of HIV-1 in Ethiopia**. *Lancet* 1990; 336(8720):942.

18. Ayehunie S, Johansson B, Salminen M, Leinikki P, Sonnerborg A, Zewdie DW, et al. **HIV-1 in Ethiopia: phylogenetic divergence from other HIV-1 strains**. *Virus genes* 1991; 5(4):359-366.

19. Neogi U, Bontell I, Shet A, De Costa A, Gupta S, Diwan V, et al. **Molecular epidemiology of HIV-1 subtypes in India: origin and evolutionary history of the predominant subtype C**. *PloS one* 2012; 7(6):e39819.

20. Flys T, Marlowe N, Hackett J, Parkin N, Schumaker M, Holzmayer V, et al. **Analysis of PTAP duplications in the gag p6 region of subtype C HIV type 1**. *AIDS research and human retroviruses* 2005; 21(8):739-741.

21. Sharma S, Aralaguppe SG, Abrahams MR, Williamson C, Gray C, Balakrishnan P, et al. **The PTAP sequence duplication in HIV-1 subtype C Gag p6 in drug-naive subjects of India and South Africa**. *BMC infectious diseases* 2017; 17(1):95.

22. Neogi U, Rao SD, Bontell I, Verheyen J, Rao VR, Gore SC, et al. **Novel tetra-peptide insertion in Gag-p6 ALIX-binding motif in HIV-1 subtype C associated with protease inhibitor failure in Indian patients**. *AIDS (London, England)* 2014; 28(15):2319-2322.

23. 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**. *The Journal of biological chemistry* 2012; 287(53):44714-44735.

24. Aralaguppe SG, Winner D, Singh K, Sarafianos SG, Quinones-Mateu ME, Sonnerborg A, et al. **Increased replication capacity following evolution of PYxE insertion in Gag-p6 is associated with enhanced virulence in HIV-1 subtype C from East Africa**. *Journal of medical virology* 2017; 89(1):106-111.

25. 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 (London, England)* 2012; 26(13):1679-1684.

26. Gupta RK, Gregson J, Parkin N, Haile-Selassie H, Tanuri A, Andrade Forero L, et al. **HIV-1 drug resistance before initiation or re-initiation of first-line antiretroviral therapy in low-income and middle-income countries: a systematic review and meta-regression analysis**. *Lancet Infect Dis* 2018; 18(3):346-355.

27. Henn MR, Boutwell CL, Charlebois P, Lennon NJ, Power KA, Macalalad AR, et al. **Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection**. *PLoS pathogens* 2012; 8(3):e1002529.

28. Coffin JM. **HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy**. *Science* 1995; 267(5197):483-489.

29. Ribeiro RM, Bonhoeffer S, Nowak MA. **The frequency of resistant mutant virus before antiviral therapy**. *AIDS (London, England)* 1998; 12(5):461-465.

30. Kearney M, Palmer S, Maldarelli F, Shao W, Polis MA, Mican J, et al. **Frequent polymorphism at drug resistance sites in HIV-1 protease and reverse transcriptase**. *AIDS (London, England)* 2008; 22(4):497-501.

31. Kozal MJ, Chiarella J, St John EP, Moreno EA, Simen BB, Arnold TE, et al. **Prevalence of low-level HIV-1 variants with reverse transcriptase mutation K65R and the effect of antiretroviral drug exposure on variant levels**. *Antiviral therapy* 2011; 16(6):925-929.

32. Lataillade M, Chiarella J, Yang R, DeGrosky M, Uy J, Seekins D, et al. **Virologic failures on initial boosted-PI regimen infrequently possess low-level variants with major PI resistance mutations by ultra-deep sequencing**. *PloS one* 2012; 7(2):e30118.

33. Mesplede T, Leng J, Pham HT, Liang J, Quan Y, Han Y, et al. **The R263K Dolutegravir Resistance-Associated Substitution Progressively Decreases HIV-1 Integration**. *MBio* 2017; 8(2).

34. Durant J, Clevenbergh P, Halfon P, Delgiudice P, Porsin S, Simonet P, et al. **Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial**. *Lancet* 1999; 353(9171):2195-2199.

35. Baxter JD, Mayers DL, Wentworth DN, Neaton JD, Hoover ML, Winters MA, et al. **A randomized study of antiretroviral management based on plasma genotypic antiretroviral resistance testing in patients failing therapy. CPCRA 046 Study Team for the Terry Beirn Community Programs for Clinical Research on AIDS**. *AIDS (London, England)* 2000; 14(9):F83-93.

36. Tural C, Ruiz L, Holtzer C, Schapiro J, Viciana P, Gonzalez J, et al. **Clinical utility of HIV-1 genotyping and expert advice: the Havana trial**. *AIDS (London, England)* 2002; 16(2):209-218.

37. Cingolani A, Antinori A, Rizzo MG, Murri R, Ammassari A, Baldini F, et al. **Usefulness of monitoring HIV drug resistance and adherence in individuals failing highly active antiretroviral therapy: a randomized study (ARGENTA)**. *AIDS (London, England)* 2002; 16(3):369-379.

38. Meynard JL, Vray M, Morand-Joubert L, Race E, Descamps D, Peytavin G, et al. **Phenotypic or genotypic resistance testing for choosing antiretroviral therapy after treatment failure: a randomized trial**. *AIDS (London, England)* 2002; 16(5):727-736.

39. Cohen CJ, Hunt S, Sension M, Farthing C, Conant M, Jacobson S, et al. **A randomized trial assessing the impact of phenotypic resistance testing on antiretroviral therapy**. *AIDS (London, England)* 2002; 16(4):579-588.

40. Grossmann S, Nowak P, Neogi U. **Subtype-independent near full-length HIV-1 genome sequencing and assembly to be used in large molecular epidemiological studies and clinical management**. *Journal of the International AIDS Society* 2015; 18:20035.

41. Aralaguppe SG, Siddik AB, Manickam A, Ambikan AT, Kumar MM, Fernandes SJ, et al. **Multiplexed next-generation sequencing and de novo assembly to obtain near full-length HIV-1 genome from plasma virus**. *Journal of virological methods* 2016; 236:98-104.

42. Martin M. **Cutadapt removes adapter sequences from high-throughput sequencing reads**. *EMBnet journal* 2011; 17(1):pp. 10-12.

43. Joshi N, Fass J. **Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)[Software]**. In; 2011.

44. Xu H, Luo X, Qian J, Pang X, Song J, Qian G, et al. **FastUniq: A Fast <italic>De Novo</italic> Duplicates Removal Tool for Paired Short Reads**. *PloS one* 2012; 7(12):e52249.

45. Langmead B, Salzberg SL. **Fast gapped-read alignment with Bowtie 2**. *Nature methods* 2012; 9(4):357-359.

46. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. **Basic local alignment search tool**. *J Mol Biol* 1990; 215(3):403-410.

47. Neogi U, Singh K, Aralaguppe SG, Rogers LC, Njenda DT, Sarafianos SG, et al. **Ex-vivo antiretroviral potency of newer integrase strand transfer inhibitors cabotegravir and bictegravir in HIV type 1 non-B subtypes**. *AIDS (London, England)* 2018; 32(4):469-476.

48. Njenda DT, Aralaguppe SG, Singh K, Rao R, Sonnerborg A, Sarafianos SG, et al. **Antiretroviral potency of 4'-ethnyl-2'-fluoro-2'-deoxyadenosine, tenofovir alafenamide and second-generation NNRTIs across diverse HIV-1 subtypes**. *The Journal of antimicrobial chemotherapy* 2018; 73(10):2721-2728.

49. Weber J, Vazquez AC, Winner D, Rose JD, Wylie D, Rhea AM, et al. **Novel method for simultaneous quantification of phenotypic resistance to maturation, protease, reverse transcriptase, and integrase HIV inhibitors based on 3'Gag(p2/p7/p1/p6)/PR/RT/INT-recombinant viruses: a useful tool in the multitarget era of antiretroviral therapy**. *Antimicrobial agents and chemotherapy* 2011; 55(8):3729-3742.

50. Selhorst P, Vazquez AC, Terrazas-Aranda K, Michiels J, Vereecken K, Heyndrickx L, et al. **Human immunodeficiency virus type 1 resistance or cross-resistance to nonnucleoside reverse transcriptase inhibitors currently under development as microbicides**. *Antimicrobial agents and chemotherapy* 2011; 55(4):1403-1413.

51. Inzaule SC, Hamers RL, Noguera-Julian M, Casadella M, Parera M, Kityo C, et al. **Clinically relevant thresholds for ultrasensitive HIV drug resistance testing: a multi-country nested case-control study**. *Lancet HIV* 2018; 5(11):e638-e646.

52. Michailidis E, Ryan EM, Hachiya A, Kirby KA, Marchand B, Leslie MD, et al. **Hypersusceptibility mechanism of Tenofovir-resistant HIV to EFdA**. *Retrovirology* 2013; 10:65.

53. Wainberg M, Anstett K, Mesplede T, Quashie P, Han Y, Oliveira M. **The R263K mutation in HIV integrase that is selected by dolutegravir may actually prevent clinically relevant resistance to this compound**. *Journal of the International AIDS Society* 2014; 17(4 Suppl 3):19518.

54. O. Huang; Krischan J. Hudson; David A. Margolis; Kimberly Smith; Peter E. Williams11, William Spreen; SSJ-FA-VGJRGRABMDMMGLVPJMMJ. **Long-acting cabotegravir + rilpivirine as maintenance therapy: ATLAS week 48 results.** In: *Conference on Retroviruses and Opportunistic Infections (CROI)*. Seattle 2019.

55. Craigie R. **HIV integrase, a brief overview from chemistry to therapeutics**. *The Journal of biological chemistry* 2001; 276(26):23213-23216.

56. Marinello J, Marchand C, Mott BT, Bain A, Thomas CJ, Pommier Y. **Comparison of raltegravir and elvitegravir on HIV-1 integrase catalytic reactions and on a series of drug-resistant integrase mutants**. *Biochemistry* 2008; 47(36):9345-9354.

57. Metifiot M, Johnson BC, Kiselev E, Marler L, Zhao XZ, Burke TR, Jr., et al. **Selectivity for strand-transfer over 3'-processing and susceptibility to clinical resistance of HIV-1 integrase inhibitors are driven by key enzyme-DNA interactions in the active site**. *Nucleic acids research* 2016; 44(14):6896-6906.

58. Tsiang M, Jones GS, Goldsmith J, Mulato A, Hansen D, Kan E, et al. **Antiviral Activity of Bictegravir (GS-9883), a Novel Potent HIV-1 Integrase Strand Transfer Inhibitor**

with an Improved Resistance Profile. *Antimicrobial agents and chemotherapy* 2016; 60(12):7086-7097.

59. Ceccherini-Silberstein F, Malet I, D'Arrigo R, Antinori A, Marcelin AG, Perno CF. **Characterization and structural analysis of HIV-1 integrase conservation**. *AIDS Rev* 2009; 11(1):17-29.

60. Garrido C, Villacian J, Zahonero N, Pattery T, Garcia F, Gutierrez F, et al. **Broad phenotypic cross-resistance to elvitegravir in HIV-infected patients failing on raltegravir-containing regimens**. *Antimicrobial agents and chemotherapy* 2012; 56(6):2873-2878.

61. Mesplede T, Quashie PK, Osman N, Han Y, Singhroy DN, Lie Y, et al. **Viral fitness cost prevents HIV-1 from evading dolutegravir drug pressure**. *Retrovirology* 2013; 10:22.

62. Wares M, Mesplede T, Quashie PK, Osman N, Han Y, Wainberg MA. **The M50I polymorphic substitution in association with the R263K mutation in HIV-1 subtype B integrase increases drug resistance but does not restore viral replicative fitness**. *Retrovirology* 2014; 11:7.

63. Markowitz M, Sarafianos SG. **4'-Ethynyl-2-fluoro-2'-deoxyadenosine, MK-8591: a novel HIV-1 reverse transcriptase translocation inhibitor**. *Current opinion in HIV and AIDS* 2018; 13(4):294-299.

64. Jacobson JM, Flexner CW. **Universal antiretroviral regimens: thinking beyond one-pill-once-a-day**. *Current opinion in HIV and AIDS* 2017; 12(4):343-350.

65. Neogi U, Haggblom A, Singh K, Rogers LC, Rao SD, Amogne W, et al. **Factors influencing the efficacy of rilpivirine in HIV-1 subtype C in low- and middle-income countries**. *The Journal of antimicrobial chemotherapy* 2016; 71(2):367-371.

66. Spreen WR, Margolis DA, Pottage JC, Jr. **Long-acting injectable antiretrovirals for HIV treatment and prevention**. *Current opinion in HIV and AIDS* 2013; 8(6):565-571.

67. Chaturbhuj DN, Nirmalkar AP, Paranjape RS, Tripathy SP. **Evaluation of a cost effective in-house method for HIV-1 drug resistance genotyping using plasma samples**. *PloS one* 2014; 9(2):e87441.

68. van Domselaar R, Njenda DT, Rao R, Sonnerborg A, Singh K, Neogi U. **HIV-1 subtype C with PYxE insertion has enhanced binding of Gag-p6 to host cell protein ALIX and increased replication fitness**. *Journal of virology* 2019.