View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Archivio Istituzionale della Ricerca- Università del Piemonte Orientale

DIPARTIMENTO DI MEDICINA TRASLAZIONALE

Corso di Dottorato di Ricerca in Scienze e biotecnologie mediche

Ciclo XXXII

Assessment of primary mutations in treatment-naïve HIV-1 Subtype C-infected patients in Malawi

SSD - Settore Scientifico Disciplinare: MED07

Coordinatore PROF MARISA GARIGLIO

Tutor PROF GIORGIO BELLOMO

Dottorando RICHARD LUHANGA

PREFACE

The DREAM Program (*Disease Relief through Excellent and Advanced Means*) was formed and run by the Community of Sant'Egidio, a catholic international organization born in Italy in 1968. Initially was an HIV and nutrition care NGO, and now is a global health care institution for communicable and non-communicable diseases (DREAM 2.0) with programs on tuberculosis, cervical cancer, hypertension, diabetes, and others. Designed for quality and based on partnerships, is a public health program as well as applied research. All our services are free of charge.

DREAM started in Mozambique in 2002 and is now in 11 other countries with 47 health centers, 24 molecular biology laboratories and 300,000 patients assisted from 2002, 130,000 patients on HAART of whom 15,000 children. Millions of peoples reached by the program and 50,000 children were born healthy in the PMTCT program

The laboratory monitoring is part of an holistic approach that DREAM uses in routine patient care. The laboratories perform Blood count, Biochemistry, CD4 count, HIV-1 viral load, Infant virological testing and, in Malawi and Mozambique, HIV genotyping, initially with Trugene platform (Siemens Health Care Diagnostics), now with a validated in-house method.

The promotion of laboratory medicine in management of HIV positive population in Malawi is one of the major contributions that DREAM Program has impacted to the Country. When DREAM rolled out its program in 2005, Malawi did not have any molecular biology laboratory for routine patient care, but owing to the training, advocacy and setting up of model infrastructures by DREAM, the Country has registered significant developments and now 10 molecular laboratories are spread across the Country for viral load and infant virological testing.

The Malawi ART guidelines recommends an HIV genotyping test for all patients failing second line regimens before switching them to third line regimens. There is also need for genotyping services for surveillance studies. DREAM Malawi program has the only genotyping facility in the Country at its Blantyre laboratory since 2008, which is also available for all public structures. DREAM has a leading role to share expertise in HIV drug resistance services in Malawi even now, that the Country has established a genotyping lab at National HIV Laboratory in Lilongwe.

SUMMARY

The aim of the PhD study was to describe retrospectively the prevalence of transmitted HIV-1 resistance mutations in a cohort of naïve patients starting treatment at DREAM program health centers in Malawi. In Africa naïve resistance mutations rate is unknown or underestimated.

In Malawi in 2014, 1.8 million people have turned to the services of HTC (HIV testing and counseling) (Malawi National AIDS Commission). There are also important results on the treatment front: as many as 61% of adults with HIV are under anti-retroviral therapy (ART), 67.6% of whom are effective virological suppression. Moreover, the estimate frequency of Transmitted HIV drug resistance (TDR) in the Country is based on a very limited number of specimen, and more studies were needed to support the national surveillance supplying programmatic information in designing education and prevention programs as well as supporting a wise use of antiretroviral drugs by clinicians and policy makers.

Simultaneous to this study, analyzing the viral sequences, we tried to highlight some polymorphism patterns to study them and compare with other patterns related to other HIV-1 subtypes (as B, present in Western Countries, when in Malawi the major subtype described is C).

Most of the activities were developed in Blantyre DREAM laboratory. An affordable laboratory method is needed in Low Limited Setting (LLS) as Malawi, to obtain the virus sequences, in order to allow national structures to produce data locally and to dramatically improve the number of performed tests. For this reason in the PhD meantime a homemade cost effective method able to detect HIV drug resistances was developed and successfully used.

The study provides some important information in understanding the transmitted resistance rate in the Country Out of 109 samples from naïve patients, 40 (36,70%) presented at least one TDR-associated mutation, The observed frequency of nucleoside reverse transcriptase inhibitors (NRTIs) resistance mutations was: M184V (4,58%), K70R/E (3,67%),D67N (2,75%). The prevalence of resistance mutations to non-nucleoside reverse transcriptase inhibitor (NNRTI) was K103N/S (11,93%), G190A (6,42%), Y181C/V (3,67%), V106M (2,75%).

The study showed that there is an increase in NNRTI resistance and that the activity of NNRTIs is compromised because of the high level of NNRTI resistance.

However in developing African countries, the majority of first line ART regimens are nevirapine/efarinenz (NVP/EFV) based, but these drugs are responsible for selection of mutations to the entire class of NNRTI (K103N, Y181C and G190A). The study confirms the need to review the treatment protocols because the use of NNRTIs could be the cause in the increase of resistance.

Given such levels of HIV Drug Resistance (HIVDR) prevalence, HIVDR testing and surveillance capacity in Malawi should be prioritized as scale-up and the adoption of the universal ART eligibility for people living with HIV (PLHIV).

To highlight the polymorphism patterns, of the 216 codons subjected to nucleotide variations found through all the sequencing products, the most represented ones were selected to ensure greater statistical significance, reducing the number to 22 codon.

Many of these polymorphisms were found to be present in very high percentages, making it part of the consensus sequence for the subtype C.

The association of V60I with all three mutations of the NAM-1 pathway, the positive association of L228R and negative of S48T/Q with M41L and L210W are very significant. Very significant associations are found between the L228R and all 4 mutations of the NAM-2 pathway, and the associations of V60I and I135T/M with two mutations each. The negative associations with S48T/Q and K173T are also very evident. The most significant associations are those with I135T / M (positive) and with V245Q (negative). The associations highlighted by a previous study on subtype B have been compared and the only concordances found concern the association of L228R with the mutations of the NAM-2 pathway and of T39A with the mutations for NRTI.

All these findings underline how subtypes B and C are characterized by substantially different mutational profiles.

These associations are even more significant considering that three polymorphisms were found with very high frequency (81.2% for T39E/D, 55.0% for Q174T and 71.3% for R211K) enough to be considered part of the consensus sequence of the subtype C, and resulting associated with NNRTI resistance mutations, one could suspect their role in making the subtype C naturally more likely to develop resistance to NNRTIs.

SOMMARIO

Lo scopo dello studio di dottorato è stato quello di descrivere retrospettivamente la prevalenza delle mutazioni di resistenza HIV-1 trasmesse in una coorte di pazienti naïve che iniziano il trattamento presso i centri sanitari del programma DREAM in Malawi.

In Africa infatti il tasso di mutazioni di resistenza naïve è sconosciuto o sottostimato.

Nel 2014, in Malawi, 1,8 milioni di persone si sono rivolte ai servizi di HTC (test HIV e counselling) (Malawi National AIDS Commission). Ci sono stati importanti risultati sul fronte del trattamento antiretrovirale (ART): ben il 61% degli adulti con HIV è in trattamento con ART, il 67,6% dei quali mostra una efficace soppressione virologica. Inoltre, la frequenza stimata della resistenza trasmessa ai farmaci contro l'HIV (TDR) nel Paese si basa su un numero molto limitato di campioni e sono necessari ulteriori studi per supportare il programma di sorveglianza nazionale, che fornisce informazioni alla progettazione di programmi di educazione e prevenzione e per sostenere un uso saggio dei farmaci antiretrovirali da parte di medici e responsabili politici.

Contestualmente a questo studio, analizzando le sequenze virali, abbiamo provato a mettere in evidenza alcuni pattern di polimorfismi per studiarli e e confrontarli con pattern relativi ad altri sottotipi di HIV 1 (in particolare il B, maggiormente diffuso nei Paesi Occidentali, mentre in Malawi il sottotipo maggiore descritto è C).

La maggior parte delle attività sono state sviluppate nel laboratorio DREAM a Blantyre.

Per ottenere le sequenze di virus è stato necessario sviluppare un metodo di laboratorio a basso costo, per permettere a Paesi come il Malawi di ottenere le sequenze virali in modo sostenibile, al fine di consentire alle strutture nazionali di produrre dati localmente e di aumentare notevolmente il numero di test eseguiti. Per questo motivo nel dottorato di ricerca è stato sviluppato all'interno del laboratorio e utilizzato con successo un metodo economicamente efficace in grado di rilevare le resistenze ai farmaci dell'HIV.

Lo studio fornisce alcune importanti informazioni sulla comprensione del tasso di resistenza trasmessa nel Paese. Su 109 campioni di pazienti naïve, 40 (36,70%) hanno presentato almeno una mutazione associata a TDR. La frequenza osservata delle mutazioni di resistenza agli inibitori della trascrittasi inversa nucleosidica (NRTI) era per M184V (4,58%), K70R/E (3,67%),D67N (2,75%). La prevalenza di mutazioni di resistenza relativamente agli inibitori della trascrittasi inversa non nucleosidica (NNRTI) era K103N/S (11,93%), G190A (6,42%), Y181C/V (3,67%), V106M (2,75%).

Lo studio ha mostrato che c'è un aumento della resistenza NNRTI e che l'attività degli NNRTI è compromessa a causa dell'alto livello di resistenza.
In molti paesi africani in via di sviluppo, la maggior parte dei regimi ART di prima linea sono basati su neviralian/efavirenz (NVP / EFV), ma questi farmaci sono responsabili della selezione delle mutazioni dell'intera classe di NNRTI (K103N, Y181C e G190A). Lo studio conferma quindi la necessità di revisione dei protocolli di trattamento perché l'uso di NNRTI potrebbe essere la causa dell'aumento di resistenza.
Dati tali livelli di prevalenza dell'HIV Drug Resistance (HIVDR), i test di resistenza e la capacità di sorveglianza epidemiologica in Malawi dovrebbero essere prioritari, viste le nuove politiche di accesso universale al trattamento.
Per evidenziare i patterns di polimorfismo, sono stati selezionati i piu' rappresentativi tra i 216 codoni che presentavano variazioni dei nucleotidi, per garantire una maggiore significatività statistica, riducendo il numero a 22 codoni.
Molti di questi polimorfismi sono stati riscontrati in percentuali molto alte, rendendoli parte della sequenza di consenso per il sottotipo C.
L'associazione di V60I con tutte e tre le mutazioni di NAM-1, l'associazione positiva di L228R e negativa di S48T/Q con M41L e L210W sono molto significative.
Associazioni molto significative si trovano tra L228R e tutte le 4 mutazioni del pathway NAM-2 e le associazioni di V60I e I135T/M con due mutazioni ciascuna. Anche le associazioni negative con S48T/Q e K173T sono molto evidenti. Le associazioni più significative sono quelle con I135T/M (positiva) e con V245Q (negativa).
Le associazioni evidenziate da uno studio precedente sul sottotipo B sono state confrontate e le uniche concordanze riscontrate riguardano l'associazione di L228R con le mutazioni del percorso NAM-2 e del T39A con le mutazioni per NRTI.
Tutti questi dati sottolineano come i sottotipi B e C siano caratterizzati da profili mutazionali sostanzialmente diversi.
Queste associazioni sono ancora più significative considerando che questi tre polimorfismi sono stati trovati con altissima frequenza (81,2% per T39E / D, 55,0% per Q174K e 71,3% per R211K) abbastanza da essere considerati parte della sequenza di consenso del sottotipo C, e risultando associati a mutazioni di resistenza, si potrebbe sospettare il loro ruolo nel rendere il sottotipo C naturalmente più probabile che sviluppi resistenza agli NNRTI.

INTRODUCTION

1.1 Characteristics of the human immunodeficiency virus (HIV)

The human immunodeficiency virus (HIV), the etiologic agent of acquired immunodeficiency syndrome (AIDS), belongs to the Retroviridae family.
Currently two HIV serotypes are known: HIV-1, which refers to genetically related viruses found predominantly in different regions of central and southern Africa, Asia, Europe and America (1) and HIV-2 which is a distinct virus prevalent in certain West African countries (2). Although both of these viruses cause AIDS, all individuals infected with HIV-2 have a long period of latency and low mortality.

1.2 Classification of the human immunodeficiency virus

Traditionally, the Retroviruses, which until the beginning of the 80s were known and studied for neoplastic diseases caused only in some species of animals, are classified into three sub-families: Oncovirinae, Spumavirinae and Lentivirinae.
Viruses originally isolated as transforming agents are grouped into oncoviruses. They cause sarcomas, leukemia, mammary tumors and, in some cases, a variable suppression of the immune system in different species of animals.
Two human retroviruses, identified for the first time in 1980 by R.C. Gallo and his collaborators, are also part of this family: Human T Lymphoma viruses I and II (HTLV I-II) associated with T-cell lymphomas and neurological disorders.
Spumaviruses owe their name to the ability to induce vacuolar lesions that give a foamy appearance to cells grown in vitro.
Lentiviruses are associated with long-term diseases but without a direct relationship with the neoplasms. Before HIV-1 was isolated and characterized, Lentiviruses were known for diseases caused in some species of felines and ungulates. From the study of these viruses, much information has been received about the pathogenesis of AIDS.
HIV-1 is classified as a member of the Lentivirus group based on its gene organization, similarity of nucleotide sequences and the type of pathology it entails, characterized by a long phase of clinical latency. Lentiviruses are viruses that induce chronic degenerative diseases

in their hosts, preceded by a long incubation period and by a variable involvement of the immune system and the central nervous system.
Another characteristic that finally distinguishes Lentiviruses is the complexity of the genome. Human-isolated Lentiviruses are grouped into two types named HIV-1 and HIV-2 based on serological and sequence properties (16,17).
A classification based on the env gene sequences, described in the 1998 HIV Compendium (18), recognizes several HIV-1 subtypes or clades (19).
Within each subtype, there is a high degree of variability. Although mutations appear to be the most responsible factor for viral variation, it has been postulated that recombination mechanisms may also occur in individuals infected with viruses of different HIV-1 (19) and HIV-2 clades (20).
Some areas of the world mainly host a single subtype, while two or more subtypes may be prevalent in other populations (19).
Molecular epidemiological studies indicate that the pattern of global distribution and variation is due to viral migration rather than viral mutation.

1.3 Morphology and structure

1.3.1 Structure of the virions

Virions are the extracellular particles produced by cells infected with HIV-1. Figure 1 is a schematic representation of the virus and its components.

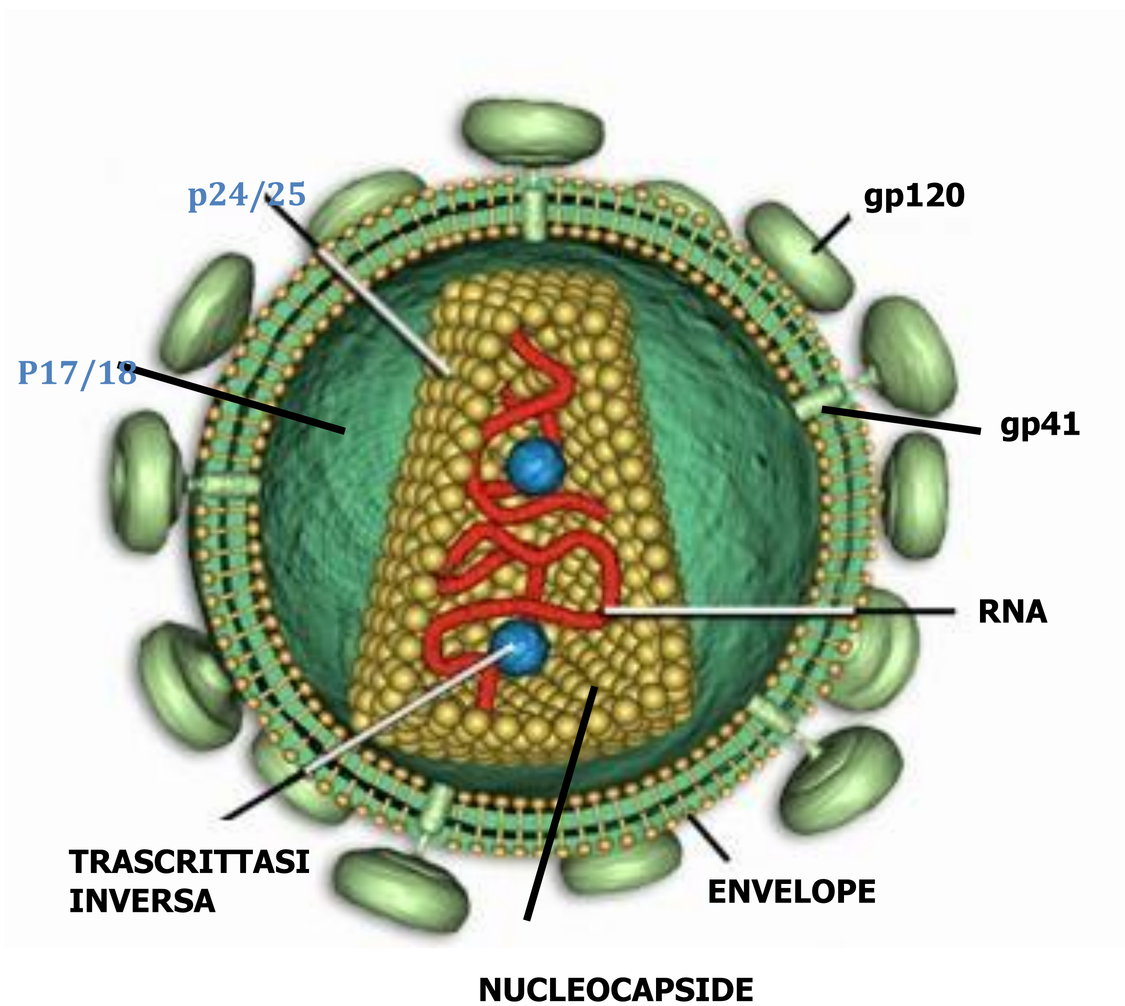


Figure 1: HIV structure

Electron microscopy shows that the virion particles have a spherical shape with a diameter of 110 nm and are formed by an external pericapside, the envelope, consisting of a double phospholipid layer on which virus-specific proteins anchor. In fact, the envelope contains about 72 triangular symmetry protuberances (21,22) each consisting of three or four heterodimers of the glycoprotein encoded by the env gene. Each heterodimer is composed of two subunits, one of which (gp120), which protrudes totally outwards, contains the domains that recognize and bind the CD4 receptor and the specific coreceptor for entry into the different target cells of the virus, while the other one (gp41), which is mostly inserted in the double phospholipidic layer, has a fusogenic activity.
The envelope also contains some cellular proteins, such as histocompatibility antigens (MHC), acquired during the budding process through the cell membrane.
Associated with the internal face of the double phospholipidic layer, a matrix protein structure is observed, consisting of the p17 protein, which forms an internal thickening of the double phospholipid layer.
Inside the viral particle is the nucleocapside or central core with a truncated cone structure. The nucleocapside crosses the entire diameter of the virion and the narrow end of the cone appears to be connected to the lipid double layer with a proteinaceous structure named CEL (core-envelope link) (23). The region between the nucleocapside and the envelope is called the paranucleoid region (24). The composition of CEL and the paranucleoid region must still be determined.
The nucleocapside consists of two single-stranded viral RNA molecules encapsulated by proteins that derive from the precursor polypeptide synthesized by the gag gene. This precursor is cut from the viral protease into four protein products: p24, the main capsid protein, whose function is to package the viral genome in virions; the p17 or matrix protein, which as mentioned is located between the nucleocapside and the virion envelope, the p9, protein of the nucleocapside that binds tightly to the genome and finally the p7 whose location is not yet clear (25). These proteins are assembled according to an icosahedral type cubic symmetry.
Inside the core, in addition to the nominal RNA, there are RNA transfer (tRNA) molecules and some viral enzymes. The tRNAs are used to trigger the replication, while the viral enzymes, which derive from the polypeptide precursor of the pol gene, are the reverse transcriptase (RT), a heterodimer consisting of two polypeptides (p51 / p66) which intervenes in the

replication of the genome, the integrase (p32) which causes the integration of viral DNA into the cellular DNA and the protease (p10) that intervenes in the maturation of the genome. An independent domain of RT carries out an additional enzyme activity of specific ribonuclease (RNAse H) that degrades the RNA in the RNA-DNA hybrid produced during the proviral genome synthesis.

1.4 Structure and organization of the viral genome

Like all retroviruses, HIV-1 presents two genomic forms: a single-stranded RNA in the extracellular phase of the life cycle of the virus and a double-stranded DNA within the cell. In the early stages of infection, virion RNA is converted into the double-stranded DNA form by the RT of the virus and is then integrated into the genome of the host cell (provirus).

The virions contain two identical copies of single-stranded RNA of approximately 9.2 Kb with positive polarity. The two strands are associated in numerous points along their length; in particular, the junction point of greater stability is located near the 5 'end of each genome. The role of diploidy is still obscure.

Viral RNA presents the characteristics typical of most eukaryotic messenger RNAs: the CAP (m7G5'ppp5'Gmp) at the 5 'end and a tail of about 200 adenine residues at the 3' end. Also occasionally adenine residues can be methylated. A $tRNA^{lys}$ molecule is positioned near the 5 'end of each strand and serves as a primer for the synthesis of the DNA negative filament by RT.

The genomic structure of HIV-1 is shown in Figure 2.

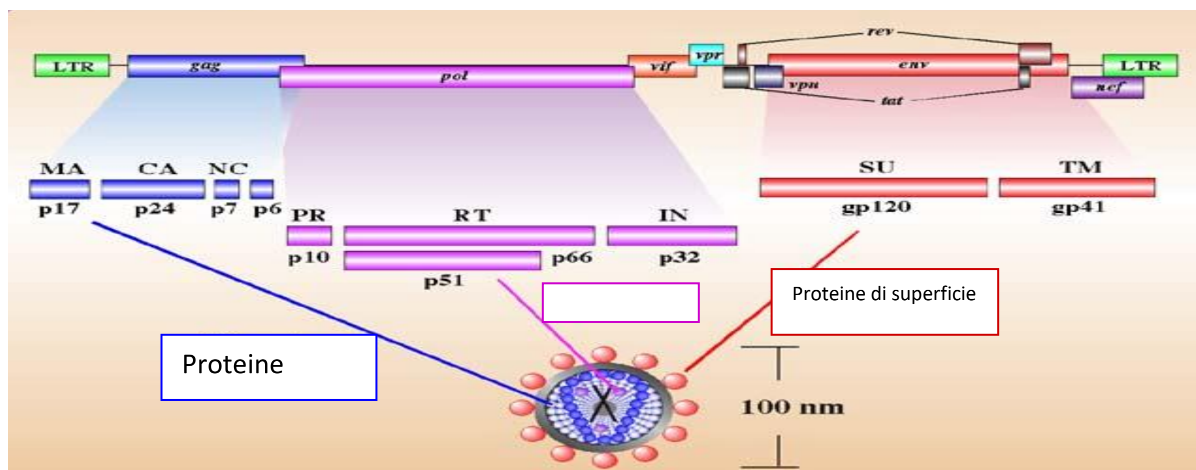


Figure 2: the genomic structure

Two repeated sequences, named long terminal repeat (LTR), are placed at the ends of the genome and flank the three major open reading frames (ORFs) of the virus: gag, pol and env and another six smaller ORFs: tat and rev, essential for viral replication, vif, vpr, vpu and nef not essential for replication, also called accessory or auxiliary genes. The presence of these six ORFs gives the virus an extraordinary level of complexity.

1.4.1 Long-terminal repeat

As in the other retroviruses, the LTR region is divided into three functionally distinct domains: U3 (-453 / + 1), R (+ 1 / + 98) and U5 (+ 99 / + 185) (from 5 'to 3').

These functional units are critical for virus integration within the host genome and contain promoters and enhancer elements recognized by viral and cellular transcription factors.

The U3 region extends from the 5 'end of the genome to the site of initiation of transcription (+1) and contains the viral promoter, consisting of three transcription domains: the modulation domain, the enhancer and the core or base domain . The first portion of U3 is the modulation region that contains *cis-acting* regulator transcription elements to which many cellular factors are bound (i.e. AP-1 and N-FAT-1). Downstream of the modulation region is the enhancer a region formed, in HIV-1, by two sites of 10 bp. These sites bind the cellular factor NF-kB which has the function of increasing the transcription activity of the virus following the activation of the same factor by some cellular regulatory proteins or other viruses. At the 3 'end of the U3 is the core containing a TATAA box region for binding to the cellular RNA polymerase II responsible for viral transcription; the viral transcription initiation site is 22 bp downstream of this region. Sites for binding the SP1 cell transcriptional factor are placed immediately after the TATAA box (26).

The region R encodes a sequence of RNA that forms a hairpin (stem-loop) structure at the 5 'end of the transcript called tat-response element (TAR). This structure binds the Tat protein of HIV-1, a powerful activator of viral transcription.

The R / U5 region is located in the leading sequence of all viral transcripts. The 3 'end of the transcripts is defined by the R / U5 edge of the 3' LTR.

Although the nucleotide sequences of the two LTRs are identical, retroviruses have mechanisms by which the 5 'LTR is used as a transcription promoter while the 3' LTR is a

signal to add the poly-A tail, in fact signals in the U3 and R region they are recognized by cellular factors that add poly-A tails to the 3 'end of viral transcripts (27).

1.4.2 Structural genes

HIV-1, like all retroviruses, has three structural genes essential for replication, called respectively: gag (group specific antigen), pol (polymerase) and env (envelope). The three genes are organized in the genome in the order 5'-gag-pol-env-3 '.
The gene gag represents the first ORF of the HIV-1 genome that encodes a peptide precursor of the internal structural proteins of the virus, of 55 kDa translated by an mRNA not subjected to splicing processes (unspliced) as long as the genome.
The precursor, the polypeptide p55, is subsequently cut from the viral protease to produce the virid capsid protein, p24, the matrix protein, p17, the nucleocapside protein p9 and the p7 previously described.
The pol gene overlaps the gag gene for about 241 bp. This gene is expressed as gag / pol fusion protein from an unspliced mRNA as long as the genome and from the segmentation of this protein the viral enzymes are derived: the protease, p10, with a autocatalytic cut, the integrase, p32, and the etherimer p66 / p51 ie the reverse transcriptase. An independent domain of the RT plays a ribonuclease function.
The env gene encodes a highly glycosylated protein precursor, gp160, translated from a polycistronic mRNA derived from a single splicing process (singly-spliced). The protein is cleaved from a host endopeptidase into two portions: the N-terminal gp120 and the C-terminal gp41 which form, through non-covalent interactions, the surface antigen of HIV-1, which mediates the entrance of the virus. Sequence comparison studies reveal a model, of the env sequence encoding gp120, of 5 variable regions (V1-V5) interspersed in conserved regions while the gp41 region is fairly conserved (28). One of these domains (V3) is highly variable and is the main target of neutralizing antibody activity.

1.4.3 Regulatory genes

The tat and rev regulatory genes control the expression of viral genes at the transcriptional and post-transcriptional levels.

The first coding exon of tat is located in the central region of the viral genome between the vpr gene and the env gene while the second exon is superimposed on the rev reading frame (open reading frame) and the env gp41 (29). The Tat protein is translated from monocistronic transcripts, produced early in the infection, subjected to multiple spliced processes (multi-spliced), and interacts with the TAR site at the 5 'end of each messenger. In this way Tat increases the levels of viral RNA by acting at the beginning of the transcription (30) or during elongation (31).

The rev gene regulates the splicing and transport of transcripts from the nucleus to the cytoplasm (32). The two coding exons of rev partially overlap with tat exons and the rev protein is also translated by multiply-spliced transcripts produced in the early stage of infection.

1.4.4 Ancillary Genes

The accessory genes of HIV-1 are vif, vpr, vpu and nef. Vpr is assembled within the virion and although most studies have not found traces of the other three proteins in the virion, it is possible that small amounts may be incorporated equally. The proteins of these genes are mostly translated from singly spliced transcripts.

Although several studies have been conducted on the function of these genes, many of these are discordant. Presumably such genes play an important role in virus-host relationships.

It is believed that the proteins encoded by the vpr and vif genes are both involved in the transmission mechanism of the virus. It seems that the vpr-encoded protein also influences the transmission rate, facilitating the nuclear localization of the viral genome, while the protein produced by the vif gene seems to be essential for virion infectivity. The protein encoded by the vpu gene, could instead be involved in the maturation process of the viral particle, facilitating its exit from the host cell.

The nef gene extends from the 3 'end of env up into the U3 domain of the 3' LTR. The transcription of the gene produces two precocious multiply spliced transcripts, independent of the rev function, one of which is monocistronic while the other is bicistronic and encodes both for nef and for rev (33).

The Nef protein has many important functions among which the activation of the cascade of enzymes that determine the activation state of the cell, in this sense the Nef protein is required for an efficient replication of HIV-1.

1.5 Life cycle of HIV

In the replicative life cycle, shown in Figure 3, of HIV-1, an early and a late phase is recognized. The early phase begins with the binding of the virion both the CD4 cell surface receptor and the coreceptor, with consequent entry of the nucleocapside into the cytoplasm. The virus RT, associated with the nucleocapside, converts the RNA into DNA that is maintained in a nucleoprotein complex and transported within the nucleus. Here the double helix of DNA is integrated into the cellular genome. The late replication phase begins with the transcription and processing of viral RNA derived from the proviral form and ending with the release of new virions from the cell.

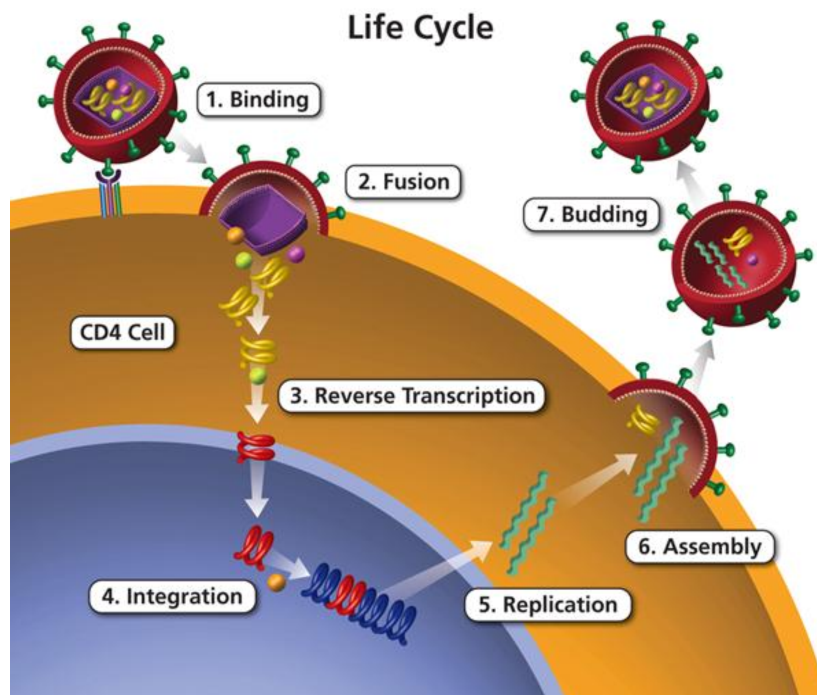


Figure 3 life cycle of HIV

1.5.1 Bonding and entry of the virion

The HIV-1 attack on cells is mediated by the initial interaction between the extracellular domain (gp120) of the virus surface antigen and the CD4 receptor located on the cell membrane of T-helper lymphocytes, macrophages, dendritic cells and microgliali (34). CD4 is a member of the immunoglobulin family and consists of four extracellular domains (D1-D4) that have a similar structure to those found in other immunoglobluline families (35). Other cell surface receptors are implicated in the attack and entry of the virion. As already mentioned, among the main molecules that can act as receptors together with CD4, it is possible to distinguish those that function as chemokine receptors; moreover, other receptors would be molecules that can facilitate the infection or that may be involved in the regulation of some steps following viral assembly, such as the fusion between cell and cell, and those that can function as primary receptors in HIV infection -1 in CD4- cells.

The existence of a CD4 co-receptor was supported by the observation that the transfection of human CD4 into murine cells (36) or a human glioma cell line (37) is not sufficient to allow entry of the virus despite the HIV-1 particles or the recombinant gp120 bind efficiently to the cells. In this regard, a protein called fusin (CXCR4) has been identified that, when coexpressed with CD4 in mouse cells, allows the fusion with cells expressing viral envelope proteins on the surface. In addition, antibodies against fusin block fusion in human cells that are normally permissive for HIV-1. This molecule belongs to the superfamily of the G protein-coupled receptors and has a close relationship with the α-chemokine CXC family receptors. It is more specifically implicated in the infection of lymphocytotropic strains and is associated with the viral phenotype SI (Syncitia Inducing).

Other chemokine receptors have been identified as coreceptors in the infection of macrofagotropic strains, in particular some cytokine receptors belonging to the α-chemokine group including CCR1, CCR20, CCR3, CCR5. Among these, CCR5 is considered to be one of the major receptors for CD4 cell infection, allowing membrane fusion and infection for the macrofagotropic HIV-1 strains, this receptor is associated with the NSI phenotype. Recent acquisitions have shown that subjects homozygous for the 32 bp deletion in the gene coding for the coronector CCR5 are frequently, but not always, resistant to HIV infection. This observation supports on one hand the hypothesis that HIV-resistant individuals do not

express CCR5 correctly, and on the other that this molecule is one of the most important HIV receptors.

Moreover, current research has developed a series of models on the possible process of attack of lymphocytotropic and macrophage-viral strains following binding of the CD4 receptor and, respectively, of CXCR4 and CCR5 receptors (34).

Other molecules may play a primary role in regulating critical steps following virus entry and cell-to-cell transmission, such as the fusion of cell membranes of CD4 T cells and HIV-1-infected monocytes and the consequent formation of syncytes . In this regard, recent studies have shown that the CD4 molecule is necessary but not sufficient to mediate the formation of syncs and that this process is regulated by the interaction between the LFA-1 adhesion molecule and its ICAM-1, ICAM-2 receptors and ICAM-3 present on lymphocytes (38). Finally, a possible receptor in the glial and neuronal cells CD4- is the glycolipid galactosyl-ceramide (GalC). In fact, it has been shown that the HIV-1 gp120 binds to this molecule with similar affinity to that of CD4, but the efficiency of the infection is lower in experimental conditions. (39,40).

As already mentioned, an important factor for the understanding of the virus-receptor binding is the genetic variability of HIV-1, a particularly important feature since variations in the nucleotide sequence can determine biologically relevant structural changes in viral proteins by modifying their functional and immunogenic properties (41, 42).

Following the recognition and binding of the target cell, the most accredited mechanism of entry of the virion (internalization) in most cells is the direct fusion of the cellular plasma membrane with the viral envelope by a process independent of pH.

The heterodimer gp120 / gp41 is kept within the envelope, through non-covalent interactions, in an oligomeric complex. In the native state, that is before the binding to the cell, each oligomeric complex presents the leucine zipper sequences of the gp41 separated from each other. Afterwards, most likely the binding to CD4 or the proteolytic cut of a membrane protease (CD26) that recognizes and cuts the sequences in the V3 loop at the level of the Arg320, induces a conformational change such that the leucine zippers in the gp41 interact and they form a fusion domain that facilitates the fusion between the two membranes.

Finally, one factor to consider is the different cellular tropism that distinguishes different HIV-1 isolates from the same infected person. Central nervous system HIV-1 isolates preferentially grow in macrophage cultures (macrophage-like strains), whereas those obtained from

peripheral blood mononuclear cells, stimulated with PHA and IL2, propagate better in T lymphocytes (lymphocytotropic strains). Blood dendritic cells and Langherans cells of the skin and mucous membranes also allow replication of many HIV-1 strains with different tropism. In contrast, follicular dendritic cells of the lymph nodes, which present the antigen to T lymphocytes in the lymphatic follicles can only bind HIV-1 to their surface without getting infected.

1.5.2 Synthesis and integration of viral DNA

The model for the reverse transcription of lentiviruses is similar to that proposed for other retroviruses and has been elucidated by analysis in cell-free systems (43). The reverse transcription of the genomic RNA is mediated by the reverse transcriptase (RT, p66 / p51) an enzyme that modulates different activities: DNA polymerase dependent RNA, DNA dependent DNA polymerase and ribonuclease (RNAase H) (44).

It is activated in the cytoplasm by a signal that is not yet well identified. The synthesis of the first DNA strand, ie the negative polarity one, begins with the annealing of the virus tRNAlys at the PBS site (primer binding site) of the RNA mold that acts as primer for the RT. The complete double helix is formed by a rather complex mechanism. After the synthesis of the filament

linear DNA is completed in the cytoplasm, the viral DNA, in the form of a nucleoprotein complex, migrates to the nucleus to be integrated into the genome of the host cell. Although both the structure and the precise composition of this preintegration complex are unknown, it is certain that viral DNA, integrase and matrix protein are part of this complex (45). The transport of the complex in the nucleus is an active process of the host cell that requires ATP but is independent of cell division (46). Lentiviruses have been proposed to replicate in differentiated non-proliferative cells in contrast to the other retroviruses.

Integration is not a random process (48). The analysis of cellular sequences flanking the provirus reveals that HIV-1 is preferentially inserted in or near two classes of repeated elements of DNA in the human genome: L1 and Alu (49). These elements are transposons and show common properties with retroviruses. Thus, the preferential integration of HIV-1 within these elements may reflect the local chromatin structure that is more susceptible to acquiring a transposon. In vitro analyzes of the proviral integration mechanism show that the

intermediate for the covalently linked provirus in the cellular genome is linear and non-circular viral DNA. However, the circular form, which can not replicate, can be found after infection and can remain stable for many days as opposed to the linear form that is rapidly degraded.

1.5.3 Expression and regulation of viral genes

The control of genome synthesis is complex and involves the combined action of viral agents acting in cis, viral transactivators and different cellular proteins. In the nucleus, the integrated proviral DNA is transcribed from the cellular RNA polymerase II to produce RNA precursors of the same length as the genome.

The expression of the provirus leads to the formation of three classes of transcripts: genomic RNA for viral progeny, messenger RNA for translation into the cytoplasm of polyproteins Gag and Gag-Pol and precursors for over 30 alternatively spliced messenger RNAs that are translated into the cytoplasm for produce env glycoproteins and accessory proteins (50). Initially in the absence of the Tat transactivator, the viral specific transcription level is low. The precursor RNAs are transported through a Rev-independent transport mechanism but before transport the spliceosomes remove the introns from the multiply spliced transcripts that are then translated to produce the regulatory Tat, Rev and Nef proteins that return to the nucleus. As the transcription proceeds, Tat levels increase and consequently the transcription activity itself increases. As Rev also accumulates, the precursor RNAs are transported to the cytoplasm by a Rev-dependent mechanism

does not require splicing before transport. In this phase, unspliced and singly spliced viral mRNAs are translated into Gag, Pol, Env, Vpr, Vpu and Vif proteins in the cytoplasm. Rev, therefore, acts as a molecular chaperone interacting with the RRE sequence present in all the transcripts in correspondence of the env gene.

1.5.4 Assembly and release of the virus

The first event in the assembly of virions is the interaction between the precursor polypeptide of the gag gene (pr55), the polypeptide precursor Gag-Pol (Pr160) and the viral genome to produce a nucleoprotein complex. This process can occur in the cytoplasm or on the membrane. The maturation of the Gag polyproteins during assembly is mediated by the

protease domain in Pr160 to produce a mature nucleocapsid. At the same time, the gp120 and gp41 are inserted into the plasma membrane and, following interaction of the matrix protein with the cytoplasmic tail of gp41, the viral nucleoprotein complex extrudes or buds through the membrane to produce a
mature virione. The precise mechanism of budding is unknown.

1.6 Role of viral enzymes

1.6.1 Reverse transcriptase

Retroviruses have a reverse transcriptase that catalyzes the synthesis of proviral DNA using viral RNA as the template. This process is complex, but
in short, it consists of three catalytic passages:

1) The production of a complementary DNA chain from the genomic RNA, which involves the formation of a RNA-DNA hybrid.
2) The removal of the RNA mold (by the ribonuclease activity of the RNase H domain of the RT), leaving the DNA with a single strand.
3) The synthesis of a second DNA chain, complementary to the first.

The double DNA chain is subsequently integrated into the chromosomal DNA by integrase. The proviral DNA becomes the template for the transcription of viral RNA for the formation of new viral particles.

The reverse transcriptase is a heterodimer composed of the subunits p55 and p66 (51)

The polypeptide contains the DNA polymerase domain (440 amino acids) and the RNase-H domain (120 amino acids). The RNase-H domain is located in the C-terminal region of the p66 subunit. Although the p51 and p66 subunits have an identical amino acid sequence, they occupy completely different relative positions. In the p66 subunit there is a split in which the RNA-DNA hybrid takes place. This structure is absent in the p51 subunit, which does not catalyze polymerization. The acid residues (aspartic acid) in position 110, 185 and 186 confer the polymeric activity.

RT activity is an essential component of HIV-1 replication.

A particular feature of the enzyme is the low degree of accuracy in the incorporation of the bases, which is fundamental for the continued survival of the virus. That is, the RT enzyme

has no 'proof reading' capability and, compared to other DNA polymerases, has a higher error frequency. This is the major factor in the generation of mutations in the viral genome. The replacement of a nucleotide can however encode the same amino acid (synonymous mutations), or it can lead to the introduction into the protein of a different amino acid (non-synonymous mutation). The amino acid change can lead to a modification of the structure and some cases of the function of the protein.

1.6.2 Protease

The HIV-1 protease is an aspartic protease that cuts the nascent polyproteotic complex during viral replication. The enzyme is composed of two identical 99 amino acid polypeptides, with a proline at the N-terminal end and a C-terminal phenylalanine. Each monomeric subunit contains two regions with β-sheets structure and a small α-helical region.
The region comprising the active site extends from methionine to position 46
to the lysine in position 55 (52).
The study of a new class of drugs, protease inhibitors, has led to a growing interest in the mechanism of action of the protease and its role in HIV infection. In fact, the maturation events catalyzed by the enzyme are essential for the formation of infectious virions.

1.6.3 Integrase

Similarly to all retroviruses, replication processes depend on the integration of the viral genome into the chromosomes of the host cells. The HIV-1 integration processes are performed by viral integrase after a series of DNA cuts and subsequent re-joining reactions.
In recent years PCR-based assays have been developed that can measure viral integration rate in cell systems and allow the identification of compounds capable of specifically blocking integrase in vivo.

1.7 Dynamics of viral replication

The knowledge of the viral replication trend was initially based on limited virological and immunological observations (p24 antigen and CD4 lymphocyte counts) and on the clinical progression of the disease.

My colleagues and I have shown that HIV-1 replication is high and continuous at different stages of infection and that viral load can be used as a prognostic indicator of AIDS development (53). In particular, the plasma levels of the virus are constant and this implies that the virus is produced and eliminated with the same frequency, ie that the system is in equilibrium.

Two important studies published in 1995 provided the first analysis of viral dynamics through the use of antiretroviral therapy (54, 55). They reported an exponential decline in plasma viremia within two weeks of initiation of antiretroviral therapy. The decrease in plasma viraemia after treatment reflected the combination of two separate effects: the elimination of plasma-free virions and the decrease in the cells that produce the virus, resulting in the blocking of new infections.

The average life span of the virus was established by observing the progression of plasma viral decline, assuming that viral replication was completely stopped during antiretroviral therapy. The average life of the virus was about 2 days. In particular, the production of HIV-1 has been estimated at around 0.68 x 109-1.1 x 108 virions per day (54, 55).

In 1996 Perelson and colleagues, using mathematical models and non-linear regressions, calculated the amount of virions eliminated daily and the average life of virus-producing cells (56). The average life of the virus was estimated at 6 hours and therefore the production of virions was 10.3 x 109 virions per day, 15 times higher than previous estimates.

1.8 Genetic variability of HIV-1

One of the most striking features of the HIV-1 virus is its extreme genetic variability that manifests itself not only in the isolates of different individuals, but also in the different isolates of the same individual during the course of the infection.

The molecular bases of viral variation are due to the high frequency of error in the incorporation of nucleotides by the RT. These modifications are the result of mutations in viral

genes that code for these enzymes and derive from the high error frequency of viral retrotranscrypt, of about 3x10-5 substitutions per replicative cycle (57). Considering that the HIV-1 genome is about 104 bases and that the total viral load in subjects with HIV-1 infections is 109-1010 virions, the probability that in a given patient can occur every single possible nucleotide substitution is around at 104-105 times in one day (58). These data suggest the rapid onset of a genetic divergence within the viral population, generating within each HIV-infected individual a genetically different 'quasispecies', rather than a single viral population.

1.8.1 Mechanisms of viral variation

Viral variation is due to three possible mechanisms:

Mutation

The fundamental characteristic of the continuous evolution of HIV-1 is the low degree of accuracy in RT polymerization. This enzyme has a high error rate and has no 'proof-reading' activity and therefore proves to be the most responsible for generating mutations in the HIV-1 genome. These changes are usually the result of point mutations that can be divided into two categories: substitutions and insertions or deletions.

a) Replacement is the change of a nucleotide with a different one during viral DNA formation. The substitution may or may not involve coding for a new protein by providing a different 'fitness' to the virus.

b) Insertion and deletion are additions or deletions of one or more nucleotides. Insertion or deletion of nucleotides may alter the 'frame' of the mRNA reading.

In particular, the structural gene env seems to possess a greater degree of variability compared to other regions of the HIV-1 genome (46-50). Analyzing sequences of env coming from the same patient with HIV-1 infection it is possible to highlight the presence of 'quasispecie', that is of closely related but distinct viral variants, which differ from each other for about 2-5% of the sequence of env (63-65). Considering instead isolates coming from different geographic areas, the genetic variability in env can reach values of 20-30% which decrease to 6-19% for isolates coming from the same area (59-62, 66-69). For this reason, the env gene represents the most frequently used gene in lithogenic analyzes, since its high rate of variability makes it possible to more precisely evaluate the genetic distances between different isolates.

Recombination

Another mechanism of viral variation derives from recombination. This event occurs frequently during retrotranscretion, as there are two genomic RNA chains in each virion. The HIV-1 RT enzyme has the ability to transfer the synthesis of a DNA chain from one template to another.

Superinfection and recombination

Retroviral recombination requires simultaneous infection of the same cell by two different strains (ie superinfection) and the subsequent integration of two parent proviral generations in the same nucleus.
Simultaneous expression and assembly of viral RNA generates a population of first generation heterozygous viral particles. These particles in turn can infect new host cells and be retrotranscribed. In this reaction, the reverse transcript can jump from one template to another. This phenomenon is what is defined as strand-switching activity of the HIV-1 RT. In this way a second chimeric generation of provirus can be formed. The viral particles produced by the second generation will contain the recombinant forms of viral RNA (3).

1.9 Pathogenesis and natural history of infection

The main route of infection transmission is represented by homosexual and heterosexual relationships and the probability of infection depends on the number of sexual partners and the different sexual practices (70). The virus is transmitted through genital secretions, both female and male, and blood (cells, plasma and coagulation factors) (71). A strong transmission incidence is registered among drug addicts due to the use of contaminated needles. The virus can be transmitted from the infected mother to the child either by transplacental route or through childbirth following exposure to the genital tract or following breastfeeding. On the other hand, non-sexual personal contacts, exposure to saliva (72), contact with urine (73) and exposure to insects (74) have never been directly implicated in the transmission of HIV.

Little is known about the early events following HIV-1 infection. The initial target cells are most likely those of the monocyte-macrophage dendritic line such as Langherans cells and lymphocytes in the genital and rectal features. Infection of new cells in the blood and dissemination occurs through the free virus or by cell-cell interaction. The major targets of HIV-1 are the lymphoreticular system, the hematopoietic system and the nervous system. Target cells that are critical for immunopathogenesis are dendritic cells, CD4 T lymphocytes, and monocyte macrophages.

The pathogenesis of neurological diseases of HIV-1 is also unknown. The predominant cells that are infected within the nervous system are monocytes and macrophages. Within the nervous system, infected macrophages are likely to release monochins, the alpha factor of tumor necrosis or the beta factor of transformation, and viral proteins that are toxic to neural cells.

Despite the devastating effects of HIV-1 on host immunity, infected individuals develop a humoral and cellular response against HIV-1-associated antigens. Neutralizing antibodies are produced by most individuals (71) and some researchers have noted a correlation between disease progression and low titers of neutralizing antibodies (74). Envelope glycoproteins are the major targets for neutralizing antibodies and the different epitopes on gp120 have shown the ability to generate neutralizing antibodies in experimental animals. Some of these are in hypervariable regions (75) while others are in conserved areas. The reactivity to the gag gene products sought by the enzyme immunoassay (ELISA) seems to reflect the progression of the infection. In fact the decrease of circulating anti-p24 antibodies and the appearance of circulating p24 antigens often correspond to the decrease in CD4 T cells and are associated with the progression of immune dysfunction and the development of AIDS (76, 77).

Also the cellular response is directed against HIV-1 antigens, in fact, cytotoxic T lymphocytes react with the env, pol, gag and regulatory gene products.

The typical clinical course of HIV-1 infection has been clearly defined following the isolation of HIV-1 from patients with AIDS or clinical signs preceding AIDS (120). It includes a phase of primary infection, with an acute syndrome of varying severity, a prolonged period of clinical latency and a final stage of disease characterized by an increase in susceptibility to opportunistic infections and neoplastic diseases (79). A peculiar feature of HIV-1 infection is the wide variability that can be found between one patient and another in the progression of the disease. In particular, what varies considerably is the duration of the clinical latency

between the different infected persons (80) and the progression to AIDS, which occurs after an average period ranging from eight to ten years (81, 82), although the decline immunological and clinical is much more precocious in a significant proportion of patients. In this regard, it is interesting to note that in a small number (from 2 to 5%) of people infected with HIV-1 (long-term non-progressor subjects) the clinical latency phase and the immunocompetence stage extend over a period of much longer time (75, 83).

In recent years there has been much debate whether CD4 T cell depletion and AIDS pathogenesis are the result of direct cytolytic effects of HIV-1, of T-cell apoptosis by non-specific activation processes or are due to dysregulation phenomena in the production of cytokines and autoimmunity events.

According to biological and immunological molecular tests, the pathogenesis and progression of HIV-1 infection is currently described as a multi-stage process (84), in which a multitude of viral and host factors are potentially implicated. been studied in recent years in vitro and in vivo (84.85).

In this complex scenario, different biological and molecular outcomes (86) have highlighted the correlation between viral replication and infection progression by studying the association between clinical stage and viral load in progressing patients (representing the majority of people infected from HIV-1) and the dynamics of viral activity during the natural course of the disease.

1.9.1 Genetic complexity of the pandemic

The increasing availability of sequences of the entire HIV-1 genome from different continents, made possible by new assays (Heteroduplex Mobility Assay) and new methodologies (PCR-long-range and automatic capillary sequencing), has greatly enriched the sequence database of HIV (HIV Sequence Database, Los Alamos, New Mexico, USA) and has allowed the study of the changes occurred in the last decade in the distribution of distinct subtypes in different geographical areas (3, 4). The sequences from different viral isolates have in fact been classified over time into groups and subtypes based on their phylogenetic relationships.

Each subtype includes sequences that are equidistant to each other and to be assigned to each subtype the strains must be similar to each other and different from other subtypes throughout the length of the genome. To study the evolutionary distances of the sequences,

phylogenetic analysis is used. In particular these distances are represented by phylogenetic trees. They are two-dimensional graphs made up of branches connected to each other via nodes. The terminal nodes, or extremities, represent the current (or contemporary) taxa, while the internal nodes represent the ancestral ones. The phylogenetic relationships are defined through the tree topology: the distribution of the nodes and the length of the branches that connect them provide indications on the evolutionary divergence between the different taxa, allowing to identify monophyletic groups or evolutionary lines (11).

To date, three groups have been defined: M (Major), O (Outlier) and N (New). Within the M group, 11 subtypes have been classified (A, B, C, D, E, F, G, H, I, J and K). The strain E, previously identified on the basis of the env gene as a separate subtype, was found to be a recombinant based on the entire genome whose genes gag and env belong respectively to subtypes A and E. Also subtype I, initially identified in a limited number of env gene sequences has turned out to be a recombinant form in which the gag gene is a sequence composed of parts of subtype A and G, while env is a sequence distinct from the others, ie I (3, 4).

In addition to the subtypes identified to date, the so-called Recombinant Circulating Forms (CRF) are present in different geographical areas. Six have been officially classified (CRF 01-06) and 3 CRF 07-09) have been recently identified (4, 5).

Among the recombinant forms CRF 01AE and CRF_02AG play an extremely important role as they circulate at high frequency respectively in South East Asia and in sub-equatorial Africa (3-5).

One of the major characteristics of the human immunodeficiency virus is its extremely high genetic variability. This heterogeneity is the result of both the high error rate of reverse transcriptase (6), which governs the evolution of subtypes, generating the high inter-subotypic and intra-subtypical variability, and the rapid turnover of virions in individuals infected with HIV. (7, 8). Furthermore, the reverse transcriptase enzyme has recently been recognized as highly recombinogenic (3) and this characteristic appears to be the main cause of CRF generation. The recombination event requires the simultaneous infection of a cell by two different viral strains (superinfection), followed by the assembly of an RNA transcript coming from each provirus within the heterozygous virion. After the subsequent infection of a new cell the reverse transcriptase, through a strand-switching mechanism between the two RNA

templates, can generate a new retroviral DNA sequence that will be recombinant between the two parental genomes (9, 10).

These mosaic viruses are defined as recombinant and this is supported by the fact that distinct recombination points can be recognized between genomic regions belonging to different subtypes (87, 88). Today it is clearly established that recombination is a relatively common occurrence that occurs between different HIV subtypes where high prevalence cocircolane occurs. Recombination generally occurs between strains belonging to different subtypes, but it can also occur between strains belonging to the same subtype.

In 1999 the further increase in the complete genomic sequences of HIV-1 (89, 90), together with epidemiological studies on its variability and its geographical distribution (91), allowed to reconsider the contribution to the pandemic of the different variants and especially of the recombinant forms (92).

On this basis the most recent classification was introduced (92), which provides that:

1) the recombinant forms of HIV-1 with high prevalence in certain populations and geographic areas are called "Recombinant Circulating Forms" (CRF);

2) sub-sub-types can be defined on a geographical basis;

3) at least three complete genomes must be isolated in order to identify a 'new' subtype, a sub-subtype, and a CRF.

If the genome contains sequences originating from more than two subtypes, the letters are replaced by "cpx", which stands for "complex".

The designation of the subtype represented a formidable epidemiological molecular marker to follow the evolution of the HIV-1 pandemic.

On the basis of current knowledge it seems clear that the various subtypes, sub-subtypes and CRFS are the result of specific epidemiological events / phenomena.

The predominant viral forms in the global epidemic are subtypes A and C, followed by subtype B and recombinants CRF02_AG and CRF01_AE (3, 93, 94).

In Africa the subtypes A and C and CRF02_AG are the most common, however in this continent are present all the groups (M, O and N) and all the subtypes, according to the origin of the epidemic. The strains belonging to groups O and N were originally identified in Africa, were not classified into subtypes and some recent reports suggest their penetration in Europe and the United States (4). In Southern and Eastern Africa the subtype C (12, 95) predominates, in West and Central West Africa the majority of the viruses are CRF02_AG

(13). In the United States, Canada, Europe and Australia, subtype B is largely the most represented, although recent studies have shown the presence of many other subtypes belonging to group M and group O viruses (96-103). Subtype B predominates in South America, however subtypes F and C have been identified (104-106). In Asia there are many different subtypes: in India the C and A strains are prevalent today, while in South East Asia, in addition to the B that probably represented the first subtype, the subtype C and the CRF01_AE circulate. In China, further CRFs have recently been identified: CRF07_BC and CRF08_BC (14, 15).

The exact prevalence of recombinant strains is not known in detail, as the studies conducted are limited to some regions and even in these areas they may not be representative of their real distribution. In Africa, preliminary data show that the gag / env discordance, a recombination index, in the samples analyzed can vary from 10 to 40, in agreement with the studied countries or regions (13, 107-111). The recombinant strains detected by these studies are related to the subtypes that co circulate in these regions: for example in Nigeria, where the subtypes A and G circulate to the same extent, the recombinants found are CRF02_AG (107, 108). Similarly, in the Democratic Republic of the Congo, where many subtypes circulate, a high prevalence of recombinants was found among the various subtypes (112).

The global distribution of the different forms of HIV-1 is a dynamic process. The likelihood of detecting new recombinant forms will increase with increasing HIV-1 variants that will circulate together in different parts of the world. The patterns of the new mosaicisms will become increasingly complex as the recombination will involve viral strains which are in turn the result of one or more recombination events. Mosaic viruses of CRF02_AG have already been observed in some African countries (113).

The worldwide distribution of distinct HIV-1 subtypes has a major impact on epidemiology, pathogenesis, diagnosis, treatment and prevention of infection.

It has long been hypothesized that the presence of different subtypes in specific areas has important implications for pathogenesis. The subtypes differ in the number or conformation of regulatory elements such as the binding site for NF-kB and the loop of the regulatory region TAR of the genomic RNA; this suggests that this represents respectively both a replicative advantage and a greater response to the pro-inflammatory cytokine TNF-α than an up-regulation of tat-mediated transcription (114-116). On the side of the structural genes the env gene has been studied with particular reference to the use of the coreceptors by the different

subtypes. These studies demonstrated important differences between the different subtypes in coreceptorial use and consequently in cell tropism. The ability to selectively use specific coreceptors and the ability of strains to adapt in vivo may influence the ability of viral 'quasispecies' to colonize new cell types in the long term and may reflect a different course of the disease's natural history (117).

The recent estimates of the evolution of the pandemic in the world indicate that heterosexuals will play a predominant role among the categories at risk of contracting HIV infection (118). However, in both Southeast Asia and Eastern European countries flourishing local economies based on narcotics trafficking are responsible for major outbreaks (119-121).

These epidemics showed two characteristics, perhaps associated with each other, worthy of great interest: a) the emergence and rapid spread of new recombinant strains characterized by complex mosaicism in which the gene transcriptase shows numerous recombinations and b) the extraordinarily low genetic diversity of HIV-1 intrapazient characterizing these new infections when compared to other parenteral HIV-1 outbreaks (pediatric nosocomial infections, an infection in a Scottish prison, infections in Australian hemophiliacs) (3). The factors influencing these phenomena are not clarified. Although rapid diffusion of the 'founding' strain (founder), early detection of the virus after transmission and exposure of these subjects to multiple viruses have been reported, RT mosaicism has been hypothesized viral implies a reduced processivity and an alteration of affinity for the nucleotides to be incorporated, that is a decreased frequency of error in the retrotranscription.

The relationship between specific subtypes and specific disease courses remains largely to be clarified. Although studies conducted in Thailand showed that CD4 cell decline, AIDS progression time and survival time were similar in groups of subjects infected with B strain compared to those with CRF01_AE (122, 123), others Studies conducted in Africa suggest that infection sustained by C strains results in higher RNA levels and lower CD4 lymphocyte counts than infections sustained by subtypes A and D (124, 125). In this regard, new prospective studies are required in seroconverted subjects to clarify the role in disease progression of different subtypes and in particular of recombinant strains such as CRF02_AG which, based on available epidemiological data, seem to have a selective advantage to transmission in Nigeria (5).

Molecular methods (RT-PCR, branched DNA and Nucleic Acid Sequenced-Based Amplification) developed in the industrialized countries for the measurement of plasma HIV-1

viremia, which is used both for the evaluation of disease progression and for monitoring response to therapy, used subtype B and showed some limitations in the detection of non-B subtypes (126-129). Although new versions of these essays have recently been introduced, to date it is not known their ability to detect subtypes that circulate at low frequency in industrialized countries (F, H, J and K) and in particular the CRFs that are acquiring a important role in the epidemic in Africa and South East Asia and that have recently been detected in Europe (3).

1.10 Laboratory diagnosis of HIV infection

The laboratory diagnosis of HIV-1 infection is based both on the determination of virus-specific antibodies and on the detection of viral antigens and / or viral nucleic acids, as well as on virus isolation. The serology diagnosis is used for large-scale screening surveys and has continuously improving levels of sensitivity and specificity. The majority of those infected with HIV produce specific antibodies within 6-8 weeks of infection. The period of time between the onset of infection and the production of antibodies is called the "serology window". However, in some cases, the appearance of antibodies may occur within 6 months of infection. Once the anti-HIV antibodies appear, they persist throughout life, especially those directed against envelope proteins (env). The anti-p24 and anti-env antibodies (gp41 and gp160) are the most frequently detected in the early stages of seroconversion.
The isolation of the virus and the search for viral components (genomic RNA, proviral DNA and viral antigens) is essential . For example, the research of proviral DNA is particularly useful in the early stages of infection (serology window) and in the diagnosis of vertical transmission of infection, when the presence of anti-HIV antibodies transmitted from the mother to the newborn does not allow a reliable serological diagnosis (9-12 months from birth). Furthermore, the quantitative determination of viral DNA gnomic RNA has proved to be a particularly reliable virological parameter for monitoring the progression of infection and antiretroviral treatment response. Finally, the evaluation of plasma viral load is an important prognostic parameter for the evaluation of survival and the probability of vertical transmission of the virus. (130, 131)

1.10.1 Serological diagnosis

Since the discovery, in the early eighties, of human retroviruses and of the etiological relationship between human immunodeficiency virus type 1 (HIV-1) and type 2 (HIV-2) and acquired immunodeficiency syndrome (AIDS), investigations serological tests for specific antibodies have represented, and still represent, the fundamental approach to virological diagnosis of HIV infection.

Since HIV infection soon reached epidemic proportions, it soon became necessary to identify new diagnostic systems that would allow for the accurate and rapid screening of an ever-increasing number of sera. Immunoenzymatic methods (ELISA) soon became the main tool for the detection of anti-HIV antibodies, thanks to their speed and ease of execution. In addition to improving the technical characteristics of ELISA assays, linked to the introduction of automated systems, their diagnostic efficiency has been improved thanks to the use of recombinant and synthetic antigens, which have allowed to increase the specificity. The use of recombinant antigens and / or synthetic peptides has also made it possible to address the problem of the diagnosis of infection with HIV-1 strains not belonging to the main subgroup (main, M), but belonging to the O (outlier) subgroup, and from HIV-2 strains.

Nonetheless, false positive results can be obtained. Therefore, the development of confirmatory assays that guarantee the specificity of the reaction between viral antigens and antibodies present in the serum remains an essential priority. The method known as Western Blot is the most widely used confirmatory test used in the diagnosis of HIV infection.

Screening essays

Rapid assays have been introduced for the screening of anti-HIV antibodies in order to have diagnostic tools even in poorly equipped environments, such as small hospitals or first aid services, or frankly inadequate, such as medical centers in countries in development, where the acquisition and maintenance of expensive and delicate equipment for automated diagnosis, such as those required for ELISA systems, are impractical. The main characteristic of these essays is the extreme simplicity and speed of execution and the visual reading of the result. Therefore, any positive result obtained with a screening method must be rechecked and then submitted to a confirmatory test.

1.10.2 Molecular diagnosis

PCR and early diagnosis

In the molecular diagnostics of HIV-1 infection the first assay introduced was represented by gene amplification by polymerase chain reaction (PCR) for the research of the proviral DNA integrated in the infected cells. This was followed by a reverse transcriptase-polymerase chain reaction (RT-PCR) assay for the detection and quantification of viral RNA in plasma by amplification of cDNA obtained in vitro following RNA sequencing retrotranscretion. The PCR reaction for the identification of HIV proviral DNA is usually carried out on DNA extracted from PBMC obtained from the peripheral blood of the test subjects. (132) Several regions of the genome can be amplified, but the most frequently used one as a target for PCR or RT-PCR reactions, it is represented by the gag gene which is one of the most conserved genomic regions.

In this context the use of rapid molecular techniques becomes fundamental: in particular the real-time PCR technique. Real-time PCR is a developed technique that, using fluorescent probes involved in the gene amplification process, allows the measurement of the polymeric chain reaction of a selected genomic segment in real time, significantly reducing investigation times. All this is possible because the progression of the polymeric reaction is associated with the increase in the fluorescence signal. Real-time PCR assays can be performed using different methods: SYBR Green, TaqMan, FRET and Beacon. In particular, the TaqMan method uses probes consisting of a short nucleotide sequence with covalently linked two fluorescent molecules: at the 5 'end a reporter (R) and at the 3' end a quencher (Q).

The short distance between the two fluorochromes causes the emission of the reporter molecule to be damped, if not completely inhibited, by the quencher. During the amplification reaction using the TaqMan method, the primers and the probe adhere to the DNA, binding to the specific target sequences. During the polymerization phase, the Taq polymerase, thanks to its 5'-exonuclease activity, removes and clasps the probe. This separation promotes the physical removal of the reporter from the quencher, resulting in the emission of a characteristic fluorescence by the R, later captured by the instrument (133).

During each cycle, the device detects the emitted fluorescence, reporting it in a graph "Fluorescence intensity / Cycles".

As the reaction progresses, the signal can be increased: it is directly proportional to the number of R molecules released in that cycle and, consequently, reflects the quantity of amplification products (133, 134).

In the first cycles the signal is very weak and can not be distinguished from the background (background signal), but with the gradual accumulation of the products it increases, initially with an exponential trend, until it stabilizes in a plateau phase (Figure 8). The saturation stage is the same for the curves of all the samples examined and is due to the exhaustion of primers, reporter or dNTPs; also the number of polymerases could prove to be limiting. The curves are separated in the growth period: this reflects differences in the initial accumulation of new DNA molecules. These differences are quantified by comparing the number of cycles required by a curve to reach a certain (reference) level of the fluorescence signal (threshold). The number of cycles used to reach the threshold is defined as the value of Ct . In this way, through the Ct value, it is possible to make considerations not only of a qualitative nature, but above all of a quantitative nature.

The introduction of this PCR method in addition to allowing an extreme rapidity of execution (results in less than 3 hours, compared to 6/10 hours of a PCR) brings with it numerous advantages including the high sensitivity, specificity (due to the use of very short but highly specific sequence probes) and reproducibility (134).

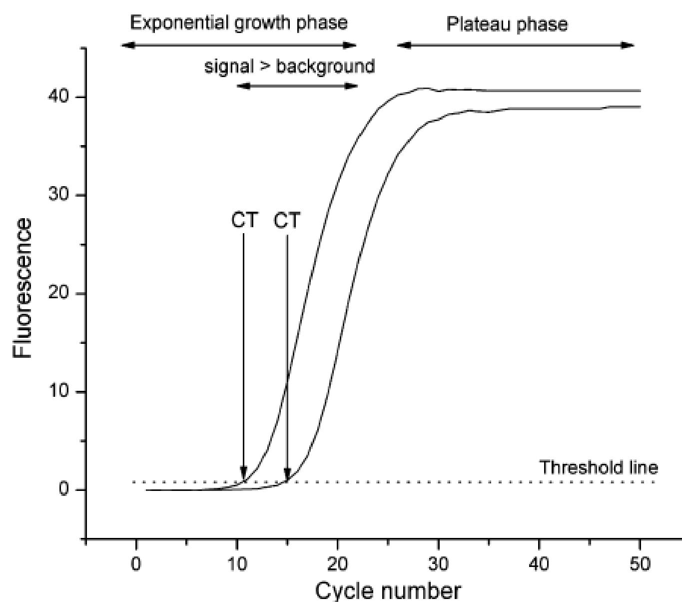


Figure 4: Real-Time PCR curve

Dosage of viral load

The availability of effective antiretroviral therapies for the treatment of HIV-1 infection necessitated monitoring of HIV infection and response to therapy. The most widely used viral replication surrogate marker is the quantitative determination of the plasma viral load (viral load) by plasma genomic HIV-RNA quantification. In fact the amount of plasma HIV-RNA correlates with the number of viral particles released in the plasma and, therefore, reflects the level of viral replication in the patient. Therefore, the reduction in viral load during therapy documents the response to treatment. If the decrease in viral load is a sign of control of the replication of the virus by the treatment, an increase in it may instead be considered in most cases an indication of poor efficacy of therapy, often due to the emergence of resistant HIV-1 strains to antiretroviral tarmacs. Therefore, the assay of plasma HIV-RNA levels is essential for individual optimization of antiviral treatment. (135)

Resistance diagnosis

Two principal commercial systems for genotyping analysis were available (ViroSeq HIV-1 genotyping system, Applied Biosystems, CA and Trugene HIV-1 genotyping kit, Visible Genetics, which was phased out in 2014). Alternatively, the assay can be performed in specialized laboratories through homemade methods. The obtained data are then analyzed by comparison of the nucleotide sequence obtained with a reference sequence of HIV-1, subtype B, using an alignment software.

These methods of identification of specific mutations present in RT and PR (both commercial and developed in specialized centers) are now widespread and allow to obtain in a short time data on the presence of drug-resistant HIV strains, which are of considerable assistance at the clinical level. However, the frequent presence of complex mutation profiles and the possibility of cross-resistance induced by single mutations between different classes of tarmacs often make interpretation of the results difficult. In this regard, it has become necessary to develop systems for interpreting sequence data.

Virtual phenotypic analysis is a system of interpretation of sequencing results developed using the genotypic sequence obtained from a given patient to predict the level of phenotypic

resistance. Data from a patient's genotypic analysis are evaluated "virtually" based on data from the phenotypic analysis of many other patients collected in a large database.
Thanks to the availability of resistance tests, therapeutic monitoring of patients with HIV infection is now easier, even if the interpretation of the results obtained with these essays is not always so immediate, and their use still requires an overall assessment of the patient's clinical situation.

NGS

Next-generation sequencing (NGS) technologies using DNA, RNA, or methylation sequencing have impacted enormously on the life sciences. NGS is the choice for large-scale genomic and transcriptomic sequencing because of the high-throughput production and outputs of sequencing data in the gigabase range per instrument run and the lower cost compared to the traditional Sanger first-generation sequencing method.(136).
Given the extreme variability of the human immunodeficiency virus (HIV) and its ability to replicate as complex viral populations, HIV variants with reduced susceptibility to antiretroviral drugs or with specific coreceptor tropism (CCR5 and/or CXCR4) may be present as minority members of the viral quasispecies. The sensitivity of current HIV genotypic or phenotypic assays is limited, and thus, these tests usually fail to detect low-abundance viral variants. Next-generation (deep) sequencing (NGS) produces an enormous amount of information that allows the detection of minority HIV variants at levels unimaginable using standard Sanger sequencing. NGS technologies continue to evolve, opening new and more affordable opportunities to implement this methodology in clinical laboratories, and HIV is not an exception. The ample use of a battery of more effective antiretroviral drugs, together with careful patient monitoring based on HIV resistance testing, has resulted in HIV-infected patients whose disease is usually well-controlled. The vast majority of adherent patients without detectable resistance become virologically suppressed; however, a subset of these patients with undetectable resistance by standard methods may fail antiretroviral therapy, perhaps due to the presence of minority HIV-resistant variants. Novel NGS-based HIV assays with increased sensitivity for identifying low-level drug resistance and/or coreceptor tropism play an important role in the success of antiretroviral treatments (137).

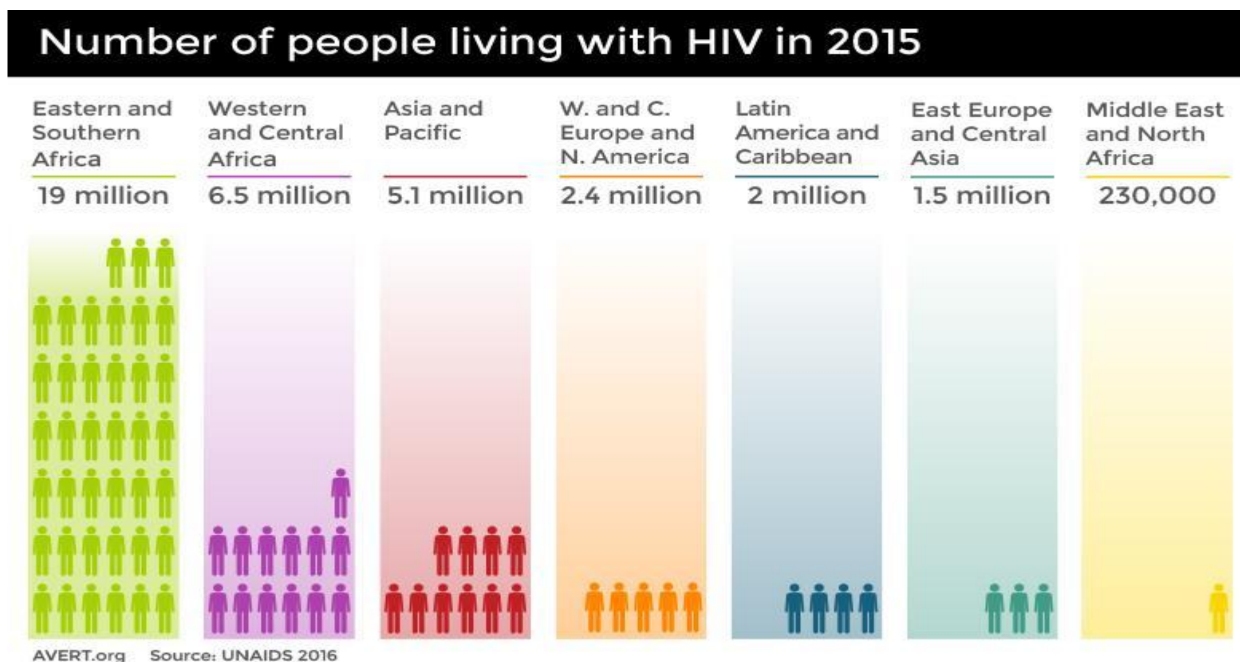
1.11 Epidemiology

1.11.1 Worldwide

The number of people living with HIV (PLHIV) in the world at the end of 2015 is estimated to amount to 36.7 million (of which about 1.8 million individuals under the age of 15), compared with 2.1 million new cases during the year (150,000 in children under 15).

Globally, HIV has a prevalence of 0.8% on the population. Italy in particular stands at 0.4%, below the world average

The latest update in June 2016 (fig 5) indicates that as many as 18.2 million people worldwide are currently being treated with antiretroviral therapy, a sensational figure considering it has doubled over the previous 5 years and leaves presage for future scenarios where almost the entire population infected can be subjected to antiretroviral therapeutic treatments in order to improve its quality and lifespan (138).



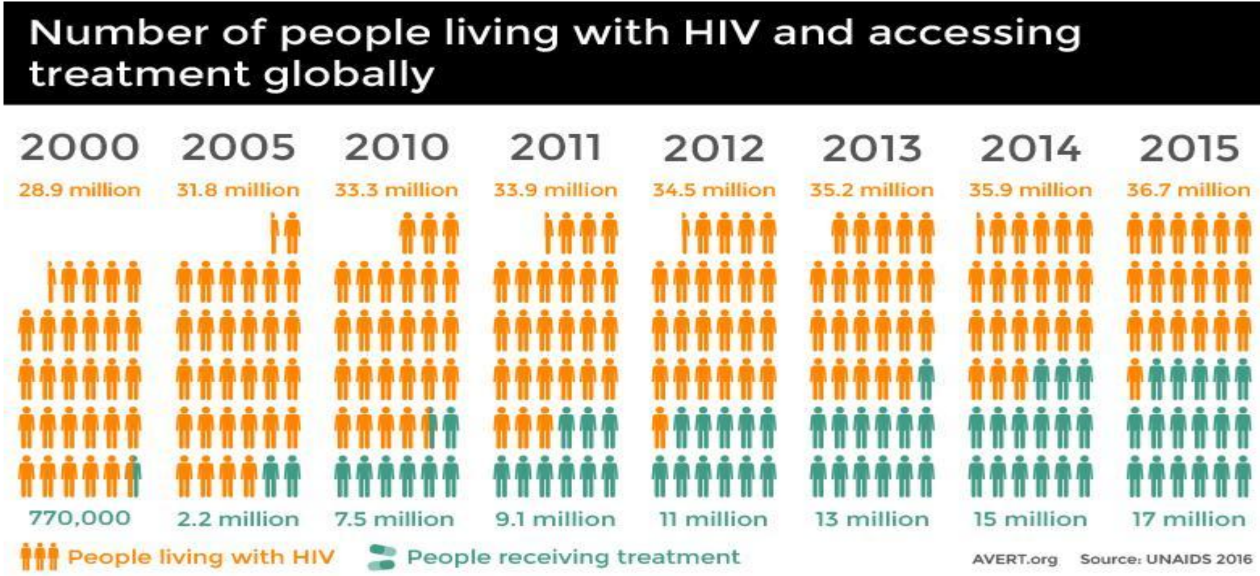


Figure 5 number of people infected and treated

The dominant HIV subtype in the Americas, Western Europe and Australasia is subtype B. As a result, the great majority of HIV clinical research has been conducted in populations where subtype B predominates, despite this subtype representing only 12% of global HIV infections. In contrast, less research is available for subtype C, although nearly 50% of all people living with HIV have subtype C. It is very common in the high prevalence countries of Southern Africa, as well as in the horn of Africa and India. The greatest diversity of subtypes is found in Cameroon and the Democratic Republic of Congo - the region where the HIV-1 epidemic originated. But migration and population mixing means geographical patterns in the distribution of subtypes are changing over time, and predicting transmission patterns in particular areas has also become more difficult.(139).

1.11.2 AFRICA and MALAWI

Malawi's HIV prevalence is one of the highest in the world, with 9.2% of the adult population (aged 15-49) living with HIV. An estimated one million Malawians were living with HIV in 2016 and 24,000 Malawians died from AIDS-related illnesses in the same year. The Malawian

HIV epidemic plays a critical role in the country's low life expectancy of just 57 years for men and 60 for women.

Over the last decade, impressive efforts to reduce the HIV epidemic have been made at both national and local levels. In 2016, 70% of people living with HIV in Malawi were aware of their status, of which 89% were on treatment, of which 89% were virally suppressed. This equates to 66% of all people living with HIV in Malawi on treatment and 59% of all people living with HIV being virally suppressed. (140,141)

New infections have dramatically declined from 98,000 new infections in 2005, to 36,000 in 2016. Malawi has also witnessed a reduction in HIV infections among children. There were 4,300 new pediatric infections in 2016, compared with 16,000 in 2010.

Malawi's HIV epidemic is generalized, which means it affects the general population as well as certain high-risk groups. Unprotected heterosexual sex between married or co-habiting partners accounts for 67% of all new HIV infections, while unprotected casual heterosexual sex accounts for 12%. Beyond this, several populations groups such as adolescent girls and young women, sex workers and men who have sex with men are particularly vulnerable to HIV.

The Malawian HIV epidemic varies greatly across the country. HIV prevalence and density is high in the urban districts of Lilongwe, Blantyre and Zomba and in the southern region of the country. (142-143)

1.12 HIV/AIDS Treatment

1.12.1 Therapeutic protocols in use in Western countries

The goal of ART therapy is to reduce morbidity and mortality related to HIV infection, and consequently improve the quality of life of the individual. The effectiveness of the regimen is assessed by the plasma virological suppression (within the first 3-6 months) and by the improvement of the immunological picture (recovery of CD4 lymphocytes).

The therapy of the naive patient, that is not yet subjected to antiviral treatments, uses the combination of multiple drugs (first-line treatment).
For example, in Italy there are 26 clinically available drugs. The choice of the appropriate combination goes from the evaluation of several factors: virological and immunological efficacy, toxicity and tolerability, cost effectiveness, possible interactions with other drugs, viral load value at the beginning of therapy, presence of possible resistance mutations specific clinical pictures and coinfections .

Although in most cases the therapeutic regimens are effective and lasting over time, a non negligible percentage of patients can experience a therapeutic failure, and the decision to change the pharmacological regimes is conditioned by the reappearance of an appreciable viral load, defined as the overcoming stable threshold of 200 copies / mL (to discriminate from so-called viremic blip: isolated and transient peaks, not indicators of stable replication recovery). Furthermore, the analysis of resistance mutations is fundamental.
The formulation of a new therapeutic formulation (called the second line) must take into account the recommendation to replace drugs that are no longer effective with drugs belonging to other classes, particularly in the presence of mutations that infuse drug cross-resistance.

1.12.2 Therapeutic schemes In Malawi

The criteria to be taken into consideration for the therapeutic approach in the developing countries are significantly different from the international guidelines, affected in particular by a smaller variety of available drugs, lower health expenditure ceilings, different nutritional conditions and environmental issues, greater difficulty in adhering to therapies and follow-up cases.
The free ART was rolled out in 2004 and since then patients starting ART are registered & followed up according to guidelines.
By end of 2017, about 1.1 million were on PLHIV and 796 100 patients are alive and on ART.
Since 2017, HIV positive patients were eligible to the antiretroviral treatment based on biological criteria as CD4 count and/or viral load, with some modification in these values threshold during the years

All pregnant women are eligible for ART .Since 2013 the Center of Disease Control of Atlanta CDC supported in Malawi the Option B+. This approach offers all HIV-positive pregnant or breastfeeding women lifesaving antiretroviral treatment (ART) for their entire lives. Option B+ simplifies the ART regimen to one pill once a day making it much easier for the patient to take and far easier to implement. This option was spread in all the Africa Countries, and was used until the adoption of the universal ART eligibility for PLHIV (test and treat approach), in 2016.
The previous schemes, were: one, colled option A, provided just a single dose of nevirapine at the delivery, and the second colled option B, provided discontinuation of NVP based ART after delivery, These schemes may have played a role in development of drug resistance mutations.
(144-146)

The ART regimens in use at the moment are:

-Regimen 1A (TDF 300/3TC 300/EFV 600) is the standard first line start regimen for all patients from 35kg

-Regimen 2A (AZT 300/3TC 150/NVP 200) is the standard first line start regimen for all patients under 35kg

- Regimen 3A (Reltagravir, Darunavir, ritonavir, abacavir 600+lamivudine 300)

-A transition to dolutegravir (DTG) based regimens for eligible patient groups as 1A is planned for January 2019

At the time of the study, the drugs available for conventional treatments consisted of a limited number of drugs, showed in the following table:

Retrotrascrittasis inhibitors	Nucleosidics **NRTI**	Stavudine (d4T)
		Zidovudine (AZT)
		Lamivudine (3TC)
		Abacavir (ABC)
		Tenofovir (TDF)
	Non-Nucleosidics **NNRTI**	Nevirapine (NVP)
		Efavirenz (EFV)
Protease inhibitors	**PI**	Atazanavir (ATV+r)
		Lopinavir (LPV+r)

The limitations in the possibilities of choice are evident, restricted to the drugs of the first generations, less expensive. In particular, Lamivudine is taken as a reference drug for all first-line treatments, because it is well tolerable and able to express a residual function even on resistant viral populations.

As a first-line treatment for adult patients above 35 kg weight the combination TDF, 3TC and EFV is recommended, while for children or adults below 35 kg the standard is AZT, 3TC and NVP (Tenofovir may have harmful effects on growing bones of children).

Alternative combinations are considered in particular conditions and it is also indicated to provide for severe drug intolerance responses with replacement with other regimens, possibly without interruption.

The possibility of incurring a therapeutic failure is monitored by the viremia dose, performed after 1 year from the start of therapy, and once a year in consecutive years. It is recommended to set up a second-line treatment if the viral load is higher than 1,000 copies / mL, or between 1000 and 4999 copies / mL for two successive repetitions, provided that in any case a satisfactory adherence to the treatments by the subjects in the months prior to the test.

1.13 DRUG RESISTANCE HIV STUDY

1.13.1.Resistance mutations

We can classify mutations according to the family of drugs that cause the onset and to which they confer resistance.

Mutations vs NRTI:

The **NAM** complex (nucleoside analogue mutations (NAMs) called either thymidine analogue mutations (**TAM**s). It is a group of 6 mutations that tend to appear sequentially and cumulatively, giving total resistance to Zidovudine AZT and Stavudine d4T and partial resistance to other NRTIs excluding Lamivudine 3TC. They give the RT the ability to remove the nucleoside terminator by a phosphololysis reaction. It has been proved that these mutations appear to follow two distinct pathways with which mutations are associated, thus distinguishing between **NAM-1** (M41L, L210W and T215Y / D) and **NAM-2** (D67N, K70R, T215F / I and K219Q / E). Once four or more of these mutations are present, the resistance is total.

The M184V mutation, which appears to be the most frequent, makes HIV totally resistant to Lamivudine and Emtricitabine, but at the same time more sensitive to the action of AZT, d4T and TDF. It is expressed with an increase in viral transcription fidelity and therefore is associated with a decrease in the onset of new mutations and the synthesis rate of RT in general.

The mutations of the 65-75 region (K65R, T69D / N, L74I / V, V75A / M / T) which confer resistance to ddC, ddI, TDF and d4T.

The Q151M mutation, rarer but associated with a total Multi Drug Resistance (MDR) framework for all NRTIs. It appears to appear in association with other mutations such as NAM, insertions in 69, M184, V75, E44 or V118.

Mutations vs NNRTI:

Among the most frequent and able to confer cross-resistance to many of the drugs in this class, mutations in the 98-108 region (A98G, L100I, K101E, K103N, V106A, V108I). In particular, K103N confers absolute multiple resistance for NNRTIs.

Mutations in the 181-190 region also tend to give cross-resistance. These are mutations Y181C / I, Y188L / C / H and G190A / S.
The mutations in the 225-236 region (P225H, M230L, P236L) are more rare and give resistance only to one or two NNRTI drugs.

Mutations vs PI:
Primary mutations (D30N, G48V, I50V / L, V82A / T / F / S, I84V / A / C, L90M). Among these some confer complete resistance to only one or two drugs and partial to others (D30N for Nelfinavir, G48V for Saquinavir, I50V for Amprenavir, V82A for Ritonavir and Indinavir, L90M for Saquinavir and Nelfinavir), while I84V gives cross resistance towards all PI inhibitors except Lopinavir.
The secondary (L24I, V32I, L33F, M46I / L, I47V, F53L, I54V / T / A / S / M / L, G73C / S / T / A, N88S / D) instead consolidate the resistances of the primary mutations or give new partial resistances.
Various natural polymorphisms are also known to show a slight increase in the effects of previous mutations if they are present in combination (L10I, K20R, M36I, L63P, A71V, V77I).

Mutations vs FI (Fusion Inhibitors):
These drugs target the gp41 protein, so resistance mutations will be identified in the env region of the HIV genome.
Several mutations in the 36-45 region are known that make the virus resistant to the effects of Enfurtivide (G36D / E / S, V38A / M / E, Q40H, N42T / D, N43D / K, L45M).

Interaction with resistance pathways

There appear to be associations between the new mutations and the NAM-1 and NAM-2 pathways.
In detail, mutations T39A, K43E / Q, K122E, E203K, and H208Y are for example associated with NAM-1 mutations. They are also located very close to the interaction site with the nucleotides to be incorporated, suggesting that they can contribute to mediate the effect of resistance to NRTIs.
Instead the K20R and D218E mutations are significantly associated with the NAM-2 pattern. In particular, D218E seems to correlate with the use of Zidovudine, an observation in line with

the results of another recent study demonstrating that prolonged exposure to Zidovudine would seem to correlate with an increase in the probability of observing the NAM-2 pathway.
The co-presence of K43E, K122E, or H208Y alone or as a cluster with NAM-1 has been associated with higher viremia and fewer CD4 cells for therapeutic failure, and this fact may suggest a compensatory role for these mutations, which leads to improved viral replication.
It is also conceivable that the class 1 and 2 mutations contribute to a further increase in the level of resistance. The phenotypic resistance data in the Stanford HIV Drug Resistance Database support this hypothesis. The presence, either individual or combined, of K43E, K122E, and H208Y with and without NAM-1 has been associated with a high increase in zidovudine resistance.
It has recently been shown that the H208Y mutation, frequently selected in combination with the NAM pathway under combined therapy with zidovudine and lamivudine, is associated with an increase in the level of resistance to zidovudine (147).

Polymorphisms

single nucleotide polymorphism (SNP) is a genetic polymorphism between two genomes that is based on deletion, insertion, or exchange of a single nucleotide.
genetic polymorphism is the occurrence, together in the same population, of two or more genetically determined phenotypes.

Special features of subtype C

For many years now it has begun to understand how the different viral subtypes have differences not only at the epidemiological level, but also at the clinical level, in particular in the acquisition of resistance mutations and differences in the pathways.
For example, some strains of subtype C spontaneously present the G190A polymorphism, thus becoming naturally resistant to some NNRTIs (Efavirenz and Nevirapine).
Even synonymous mutations may be significant, in the C strains, the V106M mutation is common in subjects treated with Efavirenz. Clinical relapse is important because V106M confers cross-resistance to the entire NNRTI class.

In recent years we are witnessing the growth of a substantial body of studies concerning the characteristics of the subtype C, especially by the countries most involved in the issue, such as South Africa and India.

Another case concerns the high frequency of occurrence of nevirapine resistance mutations following the administration of a single dose of nevirapine in women to prevent mother-child transmission of HIV (148,149). These studies have found the high resistance rate found in African women at the sixth to eighth week after the administration of nevirapine, over 40% (in some cases, as for Malawi, almost 70%), if the rate of mutation is significantly lower in the other genotypes, around 20%.

Repeated studies in recent years have led to some interesting conclusions: the natural polymorphisms present in subtype C in codons 64, 65 and 66 of RT could play an important role in the acquisition of resistance against tenofovir disoproxil-fumarate (TDF) (150, 151) Indeed, the characteristic adverse mutation to this nucleoside analogue is K65R; therefore, the polymorphisms close to that codon could potentially be attributed to the high rate of resistance development observed in sub-Saharan regions, where TDF is one of the standards of therapeutic strategies.

The work of White E. et al (152) solved the dilemma by demonstrating that by analyzing the data, an apparent risk of developing TDF resistance more than twice on the part of the C subtype compared to subtype B was observed. However, stratifying the data for the clinical condition -demographic of the subjects to the study, the differences between the two groups vanish. By adjusting the data in this way, the confounding factors such as adherence, viral load baseline and tightness of resistance monitoring, apparently constituted a bias against the studies comparing two different cohorts: subjects affected by subtype C in countries in development and subjects with subtype B in countries with better health standards.

There are three main categories of HIV drug resistance (HIVDR):

1. Acquired HIV drug resistance (ADR) develops when HIV mutations emerge due to viral replication in individuals receiving ARV drugs.
2. Transmitted HIV drug resistance (TDR) is detected in ARV drug-naive people with no history of ARV drug exposure. TDR occurs when previously uninfected individuals are infected with virus that has drug resistance mutations.

3. Pretreatment HIV drug resistance (PDR) is detected in ARV drug-naive people initiating ART or people with prior ARV drug exposure initiating or reinitiating first-line ART. PDR is either transmitted or acquired drug resistance, or both. PDR may have been transmitted at the time of infection (i.e. TDR), or it may be acquired by virtue of prior ARV drug exposure (e.g. in women exposed to ARV drugs for the prevention of mother-to-child transmission of HIV, in people who have received pre-exposure prophylaxis, or in individuals reinitiating first-line ART after a period of treatment interruption without documented virological failure). ARV drug-naive applies to people with no history of ARV drug exposure.

1.13.2. Relevance of Transmitted Drug Resistance mutations study (TDR) ..

Although all the 3 categories are a public health concern as for example transmitted resistance has the potential to more rapidly reverse the effectiveness of first-line antiretroviral therapy at the population level. Persons with transmitted drug resistance begin antiretroviral therapy with a lower genetic barrier to resistance, a higher risk of virologic failure, and a higher risk of developing resistance even to those drugs in their regimen that were originally fully active. Surveillance of transmitted resistance can supply programmatic information in designing education and prevention programs as well as supporting a wise use of antiretroviral drugs by clinicians and policy makers.

Of the 26 countries with national HIVDR surveys that are completed or ongoing, 14 have reported data to WHO (fig.6) (153)

Fig. 1: Implementation of WHO pretreatment HIV drug resistance surveys, 2014–2017[6]

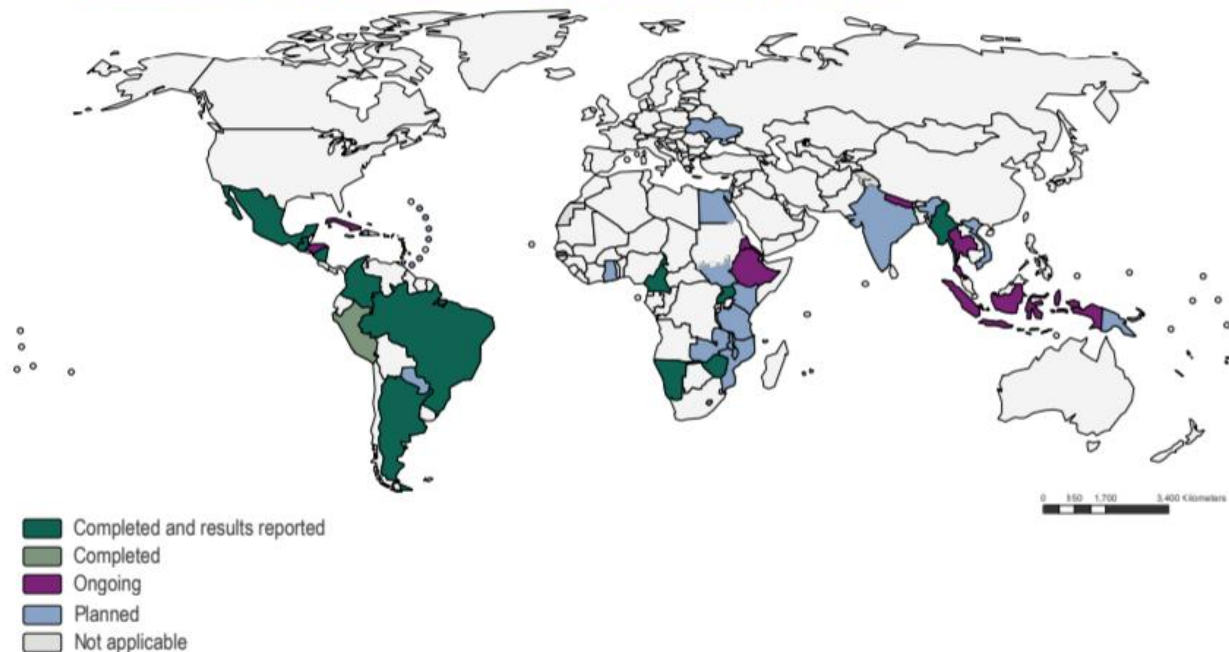


figure 6 .. implementation of WHO pretreatment resistance surveys 2014-2017

Levels of pretreatment resistance to efavirenz or nevirapine, the most affordable and widely used drugs in first-line ART, reached 10% or above in six out of 11 countries that reported PDR survey data.

Population-based HIV Impact Assessment surveys in Malawi.

The frequency of drug resistance mutations among incident HIV infections Population-based HIV Impact Assessment (PHIA) surveys are used to estimate the national incidence of HIV among adults (15–59 years), and the national prevalence of viral load suppression (HIV RNA<1.000 copies/ml). Unfortunately the size of the sample is very reduced (n=26)- (report who 2107)

In Malawi there are no laboratories equipped for the study of viral resistance, and the samples for epidemiological surveillance are sent abroad with the support of the CDC (Atlanta) From a clinical point of view, the test is not used for pre-ART evaluation (antiretroviral therapy), nor for the transition from the first to the second therapeutic line (for which the viral load results are required). The clinical resistance test begins to be required for the transition

from the second to a possible third line with integrase inhibitors that are starting to be available, but with much difficulty

Anyway, the annual HIV incidence among 15-49 year-olds in Malawi's PHIA survey was 0.32% (95% CI 0.16–0.48). Among those specimens (n=26) classified as recently infected, and with genotype data available, DRMs were detected in four out of 26 specimens. One sequence had high-level EFV/NVP resistance with K103R/S; two had low-level resistance to EFV and intermediate to NVP with A98G; and a fourth had multiple mutations impacting both NRTIs and NNRTIs (A62V, K65N, D67I, A98G, K103NE, Y181C, G190A and H221Y
The prevalence of HIVDR observed among recently infected people in Malawi is broadly consistent with the prevalence of HIVDR observed in the PDR surveys in the African Region. As mentioned above, the overall prevalence of viral load suppression among adults reporting current ART was nearly 90% (153)

Aim of the thesis

Main objective

The study general objective was the assessment of primary mutations in treatment-naïve HIV-1 Subtype C-infected patients in Blantyre, Balaka, Machinga and Mangochi districts in Malawi identifying major transmitted drug resistant mutation frequency (TDR).

Related objectives

Determine primary mutations, accessory mutations and polymorphisms from resistance data of HIV-1 Subtype C infected patients in Malawi

Explore any association between polymorphisms and accessory mutations with primary resistant mutations in HIV-1 Subtype C that could lead to mutational pathways and compare the observed mutational pathways in Subtype C with published mutational pathways in Subtype B

Evaluate a cost effective In-House Plasma Method for HIV-1 Drug Resistance Genotyping

Materials and methods

First step

In this laboratory-based retrospective study, we analyzed HIV sequences from 109 naive patients enrolled in the DREAM Program prior to treatment initiation between 2013 and 2015. This cohort did not have prior exposure to ARVs.

Eligible patients responded to the following characteristics:

- documented positivity for HIV
- age over 15 years
- ability to understand and accept informed consent
- patients belonging to the DREAM program from Blantyre, Balaka, Machinga and Mangochi districts
- availability of a plasma aliquot of at minimum 1 ml

The DREAM program offers free of charge state-of-the-art treatment, following National guidelines, and laboratory monitoring (including viral load and CD4 counts, basic biochemistry, and complete blood counts) was done on a routine basis for all patients.

For Viral Load determination, EDTA blood samples were collected and received at the laboratory within 4 hours of collection, immediately separated by centrifugation. Two aliquots of plasma from each sample were prepared and stored at −80◦ for a minimum of one day and a maximum of 3 months. One aliquot was used for routine determination with the Abbott m2000 system and the second processed for genotyping. The HIV-1 RNA quantification was performed with the Abbott Real-Time HIV-1 assay (m2000rt; Abbott Molecular Diagnostics, Mississauga, Ontario, Canada) according to the manufacturer's instructions. Purified RNA was obtained from a 0.6 mL plasma sample by using the m2000sp automated extractor, and qRT-PCR amplification and detection were done using the fully automated m2000rt instrument.

Viral sequences of the pol gene were obtained from plasma (HIV-RNA) using the Trugene assay (Siemens Diagnostics) after extraction from whole blood (QIAamp DNA Blood Mini Kit,

Qiagen, Hilden, Germany). This assay is a Sanger sequencing method that uses labeled fluorescent primers instead of labeled terminators.

The assay consists of several processes:

- Reverse Transcription and PCR Amplification: Reverse transcription of target RNA to generate cDNA using Reverse Transcriptase (RT) and PCR amplification of target cDNA using HIV-1 subtype B specific primers;
- CLIP Reaction: CLIP sequencing of the PCR amplicons using HIV-I subtype B specific primers;
- Gel Preparation, Polymerization, and Electrophoresis of CLIP Reaction Products: Separation of the CLIP sequencing reactions by electrophoresis on a polyacrylamide gel, and detection by laser-induced fluorescence; and
- Data Analysis - Genotyping: Analysis of the forward and reverse CLIP sequences using the OpenGene system software. The end result is a TRUGENE HIV-1 Resistance Report.

All the tests were performed in the Blantyre DREAM laboratory.
Among the 109 specimens, 61 were from females patients and 48 from males.

Second step

In a second step, we considered a second group of 425 sequences from a cohort of both naive and experienced patients (patients awaiting ART treatment or already undergoing therapy), to study different mutational pathways related to polymorphisms.
Eligible patients responded to the following characteristics:
- documented positivity for HIV
- age over 15 years
- ability to understand and accept informed consent
- patients belonging to the DREAM project with a detectable viremia of> 1000 copies / mL
- availability of a plasma aliquot corresponding to the sample tested for viremia

The samples were all subjected to sequence analysis in the pro-pol region of the viral genome.
For each sample, the whole obtained sequence was analyzed, and the information

corresponding to the amino acid sequence of the over 1000 codons was analyzed and elaborated to highlight the positive and negative associations between the different groups of mutations.

The polymorphisms were included in the study on the basis of their absolute frequency in the sample cohort: all the polymorphisms with a prevalence greater than 5% were considered, a level sufficient to guarantee a consistent statistical significance.

Statistic analysis

The data were analyzed using SPP software package version 20.0 (Chicago, IL)

Optimization of the homemade method

The optimization constitutes a series of processes and activities that ensured that the method was working in our laboratory. It included evaluating technical steps of the method and logistics, because in this method we have only one step that had to be done externally (sequence reaction reading), because there is no sequencer instrument in the country.

- Technically, we found all the steps feasible from purification of RNA from plasma to amplification and sequencing.
- Logistically, we identified a laboratory in South Africa, (Ingaba Biotechnical Industries) to read our sequences. Since we are exporting amplification products, this did not pose any documentation and declaration challenges. We got our read sequences from the reference lab within a week, and did not affect our overall turnaround time (TAT) of having resistance results within two weeks.

DREAM - Home Made Method of HIV Drug Resistance Testing

DREAM has been performing drug resistance testing in Malawi since 2008 using TRUGENE genotyping kit (Siemens Healthcare Diagnostics). The test was used for routine patient management and studies. This method was phased out in 2015. Through partnership with the University of Siena, DREAM implemented a homemade method of genotyping in 2016. The motivation behind the development of the new method arose from lack of in-country capacity for genotyping services and the need of genotyping services in the country for patient care and studies having an ART program for over 14 years. In addition, the Country needs for a

cost effective method of genotyping. The current basic commercial method like Abbott ViroSeq costs over $200 per patient. The DREAM homemade method costs under $95 and is suitable for limited resource countries. Developing this method was possible in our laboratory taking advantage of a lots of technical knowledges after having worked on the commercial method (TRUGENE sequencing procedure) for over 8 years (2008-2015).

After providing informed consent, EDTA blood samples are collected and 2 aliquots of plasma from each sample are prepared and stored at -80ºC. One aliquot is used for routine RNA quantification and the second one is genotyped for participants eligible for the study.

The RNA quantification is done with the Abbott Real-Time HIV-1 assay (m2000rt: Abbott Molecular Diagnostics, Mississauga, Ontario, Canada) according to the manufacturer's instructions. For HIVDR genotyping, viral nucleic acids are extracted from plasma specimens with QIAamp®Viral RNA kit (Hidden, Germany) according to manufacturer's instructions.

HIVDR genotyping of the protease (PR) and reverse transcriptase (RT) regions of HIV-1 pol gene are obtained using homemade method.

Reverse transcription-polymerase chain reaction (RT-PCR)

Briefly, the entire PR and amino acids 1-250 of RT is amplified using reverse transcription (random primer Promega Corporation 2800 Woods Hollow Road Madison, WI 53711 USA)

Following RNA extraction, the RT-PCR was performed immediately: Add 20 µL of the viral RNA to 0.2 mL reaction tubes which were then placed into a thermal cycler(Geneamp PCR System 9700, Applied Biosystems) to relax the RNA secondary structure programmed at 70 °C x 5 minutes.

The master mix for 1° PCR was prepared with: 5 µl ImProm-II 5X Reaction Buffer (ref#: M289C, Promega, at -30°C), 1.8 µl MgCl2 (25mM, ref#: A351H, Promega, at -30°C) 4 µl of dNTP mix (1.25mM, Promega, at -30°C), 1.0 µl random primer (500µg/mL, ref#: C118A, Promega, at -30°C), 0.5 µl RNasin® ribonuclease Inhibitor (40U/µL, ref#: N211B, Promega, at -30°C) and 1 µl ImProm-II™ Reverse Transcriptase, (ref#: M314C, Promega, at -30°C) making a final volume of 30 µl. Following the denaturation of extracted RNA in the thermocycler, 14 µl of the master mix was added to each tube and run at the following PCR conditions:

N° cycles	
1	25°C x 5 min

1	37°C x 45 min
1	80°C x 5 min

The RT-PCR product (cDNA) was stored at 2-8°C in the refrigerator pending nested PCR

Nested PCR

The nested PCR is performed with primers 533 and 534 (nPCR1) and primers 535 and 536 (nPCR2) as showed **in tab 1** using Go Taq® G2 Hot start polymerase kit (Promega Corporation, 2800 Woods Hollow Road Madison, WI 53711-5399 USA).

Primer	Direction	Sequence (5'.........>3')	Location on HXB2	Purpose
P533	antisense	GCTAYTAARTCTTTTGWTGGGTCATA	3504-3519	1° PCR
P534	Sense	AAAARGGYTGTTGGAAATGTGG	2018-2039	1° PCR
P535	Sense	GARAGRCAGGCTAATTTTTTGGGA	2071-2095	2° PCR
P536	antisense	TCTTTTARAATYTCCCTGTTYTCTGC	3459-3484	2° PCR
P213	Sense	GCTTTTATTTTTTCTTCTGTCAATGGCCA	2619-2647	Sequencing
P219	antisense	TTTCCAATTAGTCCTATTGARACTGTRCC	2547-2575	Sequencing
P535	Sense	GARAGRCAGGCTAATTTTTTGGGA	2071-2095	Sequencing
P536	Antisense	TCTTTTARAATYTCCCTGTTYTCTGC	3459-3484	Sequencing

tab. 1 sequence of primers used in nested PCR

The master mix for1° PCR was prepared with: 10 µl 5X Colourless Go Taq Flexi Buffer (ref#: M890A, Promega, at -30°C), 5 µl MgCl2 (25mM, ref#: A351H, Promega, at -30°C) 3.2 µl of dNTP mix (1.25mM, Promega, at -30°C), 0.3 µl primer 533 , 0.3 µl primer 534, 0.2 µl Go Taq® G2 Hot Start Polymerase (5U/µL ref#: M7408, Promega at -30°C), 26 µl Nuclease free water (ref#: 1039498,QIAGEN at 15-25°C, making a final volume of 50 µl. Then 45 µL of the 1° PCR master mix was distributed into each 0.2 mL reaction tubes to which 5 µL of template cDNA was added. The samples were run in a thermocycler programmed at:

N° cycles	Denaturation	Annealing	Extension
1	94°C x 4 min		

5	95°C x 40 sec	56°C x 30 sec	68°C x 1 min 45 sec
25	95°C x 40 sec	54°C x 30 sec	68°C x 1 min 45 sec
1			54°C x 1 min
			68°C x 8 min

The DNA was stored at 4 °C pending 2° PCR

The master mix for2° PCR was prepared with: 6 µl 5X Green Go Taq Flexi Buffer (ref#: M891A, Promega, at -30°C), 3 µl $MgCl_2$ (25mM, ref#: A351H, Promega, at -30°C) 0.8 µl of dNTP mix (1.25mM, Promega, at -30°C), 0.3 µl primer 535 , 0.3 µl primer 536, 0.2 µl Go Taq® G2 Hot Start Polymerase (5U/µL ref#: M7408, Promega at -30°C), 17.4 µl Nuclease free water (ref#: 1039498,QIAGEN at 15-25°C, making a final volume of 28.4 µl. Then 28 µL of the 2° PCR master mix was distributed into each 0.2 mL reaction tubes to which 2 µL of DNA template was added. The samples were run in a thermocycler programmed at:

N° cycles	Denaturation	Annealing	Extension
1	94°C x 4 min		
5	95°C x 40 sec	56°C x 30 sec	68°C x 1 min 45 sec
29	95°C x 40 sec	54°C x 30 sec	68°C x 1 min 45 sec
1			54°C x 1 min
			68°C x 8 min

The DNA was stored at -20 °C pending confirmation of amplification PCR products by agarose gel electrophoresis.

Detection of PCR products on agarose gel

The confirmation of PCR amplification results was obtained by electrophoretic run on 1.5% agarose gel. During the preparation 10 µl of Atlas Clear DNA Stain (cat# BH40501, BIOATLAS, at 4°C) was incorporated, and was intercalated between the bases of the nucleic acid and allowed the visualization of the DNA under UV exposure. Moreover, it was possible to recognize the presence and size of the amplification product by a molecular ladder loaded simultaneously with the samples (marker 100, BenchTop 100bp DNA ladder, Ref: G210A, Promega corporation USA). The expected PCR product that was required for sequencing measured 1.08kb and fell just below the 1.2kb marker

PCR purification

The PCR products were purified using EXOSAP-IT Purification Kit (Affymetrix, Santa Clara, California, USA) 2uL of ExoSAP-IT was aliquoted for 5uL of PCR product, vortexed, spin and run in the thermocycler on the following program:

N° cycles	
1	37°C x 15 min
1	80°C x 15 min

Cycle sequencing

The purified PCR product was directly sequenced in both directions using BigDye™ Terminator Cycle Sequencing Ready Reaction Kit version 3.1 (Applied Biosystems, Foster City, CA, USA). The master mix for sequencing reaction for each primer was prepared with: 3.2 µl primer (1pmol/ uL), 2.3 µl Nuclease free water (ref#: 1039498, QIAGEN), 1 µl of BigDye buffer (5X), 0.5 µl of BigDye terminator mix, making a final volume of 10 µl. Then 7 µL of the cycle sequencing master mix was distributed into appropriate wells of a 96-well reaction plate to which 3 µL of purified template DNA was added. The samples were run in a thermocycler programmed at:

N° cycles	Denaturation	Annealing	Extension
1	96°C x 1 min		
25	96°C x 10 sec	50°C x 5 sec	60°C x 4 min

The sequencing PCR products were stored at -20 °C pending sequencing purification.

Sequencing purification

Sequencing purification was done with BigDye® XTerminator™ Purification Kit (20 mL product code 437487, Life Technologies) to remove dye terminators. The procedure involved: Aliquot 10uL to each well in the PCR reaction plate with sequencing PCR products, followed

by dispensation of 45uL of SAM solution to each well. The sealed reaction plate was vortexed for 30 min at 2000rpm and spun for 3 min at 3000rpm. A new seal was then applied and the plate was run in a sequencer to generate sequence reads.

Sequence Detection

Sequence reads were done externally at Ingaba Biotechnical Industries, In Pretoria, South Africa. The shipment was by FedEx courier.

Sequence Analysis and Interpretation

DNAStar Lasergene was used to edit sequences and generate consensus sequences before analysis.
Sequence traces are aligned and consensus sequence constructed using Seqman Pro software (Madison, Wisconsin, USA).

Interpretation of results was done with International Antiviral Society-USA list-2016 and Stanford HIV Drug Resistance Database. (www.stanford.org)

Quality Assurance
In-house positive and negative controls were prepared and included in each RT-PCR reaction run exactly as testing specimens to check the validity of the run. Negative control was (HIV negative) human plasma stored at -80°C. The positive control was normal human plasma spiked with a plasmid PLN4P12, stored at -80°C. The run was valid when there was no visible band in the negative control and the positive control showed amplification and visible correct sized band (1.06kb) during gel electrophoresis. The test was registered in the external quality assessment scheme with a United Kingdom based Quality for Molecular Diagnostics (QCMD).

General characteristics of the method

The homemade method was optimized and validated on plasma specimens. The Abbott m2000sp extractor was found to be equally effective in producing quality nucleic acids for sequencing. It can detect group M & O It has sensitivity for VL>1000 copies

The Ingaba Biotechnical Industries (Pty) Ltd in Pretoria, South Africa was identified to read the sequences at a cost of $3.41 per unit. The overall total cost per test was $95. Turnaround time of receiving sequence data was 7 days. Seqman Pro 14 Software, was selected owing to its cost effectiveness and I got trained on how to use it at Milano Sacco

Validation of the method

The objective was to validate DREAM homemade method as one of cost effective methods for HIV genotyping using plasma samples.

The validation includes establishment of analytical performance characteristics such as sensitivity, specificity, precision, reproducibility and accuracy:

- Amplification sensitivity was evaluated on HIV-1 positive plasma samples (n50) with known viral loads ranging from about 1000 to 5.8 million RNA copies/mL
- Amplification specificity was evaluated on molecularly confirmed HIV-1 negative samples (n10)
- The precision and reproducibility were carried out on 3 samples with 5 replicates (n15) representing varying viral loads (low VL about 1000, medium VL about 10 000, high VL about 150 000 copies/mL)
- Accuracy is planned to be assessed by comparing paired sequences (n10) generated by DREAM homemade method and the method used by CDC as a gold standard.
- Another way of checking accuracy was to get registered to an external quality assurance scheme, which we did with Quality Control for Molecular Diagnostics (QCMD), a London based provider.

The method showed:

- 100% sensitivity on samples with HIV-1 viral load >1000 copies/mL
- 100% specificity on HIV-1 negative samples
- precision and reproducibility data showed high nucleotide sequence identities, >99.79+/-0.10, and 99.75+/-0.10 respectively

RESULTS

1.0 Frequency of Transmitted Drug Resistance in HIV Naïve patients

The study objective, as we said, was to describe the frequencies of major HIV-1 drug resistance associated mutations (RAM) and those listed for surveillance of transmitted drug resistance (TDR) in a cohort of drug-naïve patients.

Specifically, we calculated the proportion of subjects with transmitted drug resistance mutations TDR, and the proportion of subjects with NRTI-TDR, NNRTI-TDR, PI-TDR.

In this laboratory-based retrospective study, we analyzed HIV sequences from 109 patients enrolled in the DREAM Program prior to treatment initiation between 2013 and 2015. This cohort did not have prior exposure to ARVs. The HIV drug resistance genotyping was performed at Blantyre DREAM Laboratory.

Among the 109 specimens, 61 were from females patients and 48 from males.
The baseline characteristics are shown in the table 1 below and the results of baseline characteristics for the population are reflecting the common Malawian HIV infected patients.

Variables at baseline	
Observation time	2013-2015
Gender	Female (61), male (48), total 109
Age (years)	33 (IQ:26-41)
CD4 count (cell/mm^3)	554 (IQ:385-728)
HIV Viral Load (Log_{10} copies/mL)	3.97 (IQ:3.39-4.43)
Haemoglobin (g/dL)	11.30 (IQ:10.00-12.65)
Body Mass Index(Kg/m^2)	21.95 (IQ:18.86-26.17)

Table 1: patient characteristics

Out of 109 patients, 40 (36,70%) presented at least one TDR-associated mutation as shown in table 2

Resistance mutation	Abs frequency (n:109)	% frequency
TDR	40	36,70
NRTI-TDR	13	11,93
NNRTI-TDR	32	29,36
NRTI & NNRTI	5	4,58

Table 2: Transmitted Drug Resistance mutation frequency

1.1 Frequency of nucleotide reverse transcriptase inhibitors (NRTIs) resistance mutations in naïve patients

The observed frequency of nucleoside reverse transcriptase inhibitors (NRTIs) resistance mutations was M184V (4.58%), K70R/E (3,67%), D67N (2,75%) and T69D (1,83%)

Table 3: frequency of NRTIs mutations in naïve patients

NRTI resistance mutation	Frequency (N)
M184V	4,58% (5/109)
K70R/E/T	3,67% (4/109)
D67N/G	2,75% (3/109)
T69D/N	1,83% (2/109)
K219Q/E	0,92% (1/109)

M184V causes high-level in vitro resistance to 3TC and FTC and low-level resistance to ddI and ABC. However, it increases susceptibility to AZT and TDF.
K219Q/E are accessory TAMS associated with reduced susceptibility to AZT and possibly d4T (stavudine, nucelosidic analogue).

D67N is a non-polymorphic TAM associated with low-level resistance to AZT and d4T.
When present with other TAMs, it contributes to reduced susceptibility to ABC, ddI, and TDF.
T69D is a non-polymorphic mutation that reduces susceptibility to ddI and possibly d4T.
K70R causes intermediate resistance to AZT and possibly low-level resistance to D4T, DDI, ABC and TDF.

1.2 Frequency of non-nucleotide reverse transcriptase inhibitors (NNRTIs) resistance mutations in naïve patients.

The prevalence of TDR to non-nucleoside reverse transcriptase inhibitor (NNRTI) was K103N/S (11,93%), G190A (6.42%), Y181C/V (3,67%), V106M (2,75%), K101E (3,67%), V179D/T (4,59%) and E138K/Q/R/G/A/S (6.42%).
Other observed NNRTI mutations were observed at a frequency of <5%.

NNRTI resistance mutation	Frequency (N)
K103H/N/S/T	11,93% (13/109)
E138K/Q/R/G/A/S	6,42% (7/109)
V179D/T	4,59% (5/109)
K101H/E/P/Q	3,67% (4/109)
Y181C/V	3,67% (4/109)
G190A/S	6,42% (7/109)
V106M/A	2,75% (3/109)
M230I/L	1,83% (2/109)

K103N indicates resistance to efavirenz (EFV) and nevirapine (NVP). The presence of 11.93% of this mutation in naïve patients is worrying because in the programmatic antiretroviral program being implemented in Malawi, 93% of patients are on first line regimen that contains TDF/3TC/EFV (5A).and the rest are on 2A regimen (AZT/3TC/NVP).

K103N is a non-polymorphic mutation that causes high-level resistance to NVP and EFV. K103S is a non-polymorphic mutation that causes high-level resistance to NVP and intermediate resistance to EFV. Because K103S is a 2-bp change from the wildtype K and a 1-bp change from K103N, patients with K103S may be likely to have once had K103N.

Y181C indicates possible resistance to EFV and the presence of this mutation may indicate undetected presence of other NNRTI resistance associated mutations.

G190 A/S is a nonpolymorphic mutation selected by NVP and EFV. Alone G190A reduces NVP susceptibility >50-fold and EFV susceptibility 5 to 10-fold.

V106M is a nonpolymorphic mutation selected primarily by EFV and NVP. It is particularly common in subtype C viruses. It causes >30-fold reduced susceptibility to NVP and EFV . K101E usually occurs in combination with other NNRTI-resistance mutations. Alone it reduces susceptibility to NVP by 3 to 10-fold, to EFV by 1 to 5-fold. K101H is a nonpolymorphic mutation selected in patients receiving NVP, EFV, and ETR. When it occurs in combination with other NNRTI-resistance mutations, K101H is associated with reduced susceptibility to NVP and EFV. K101P is a nonpolymorphic two base-pair mutation selected in persons receiving each of the NNRTIs except possibly DOR for which few data are available. It reduces NVP, EFV, and RPV susceptibility by >50-fold. K101Q is a minimally polymorphic accessory mutation weakly selected in patients receiving NVP and EFV. It appears to have minimal, if any, detectable effect on NNRTI susceptibility

There were 5 specimens (4,58%) that had both NRTI and NNRTI resistance.

Frequency of resistances to some specific drugs as zidovudice, lamivudine, tenofovir , nevirapine are shown in table 5.

Resistance frequency to AZT is 5,5 % (6/109) while the frequency of resistance against Lamivudine 3TC is 4,6% (5/109). The frequency for Tenofovir TDF is low and is value is around 1,8% (2/109) while we detected a very high percentage for frequency of resistance against NVP/EFV ; 27,5% of naïve subjects shown a resistance against these drugs (30/109).

Res_AZT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Absent	103	94.5	94.5	94.5
	Present	6	5.5	5.5	100.0
	Total	109	100.0	100.0	

Res_3TC

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Absent	104	95.4	95.4	95.4
	Present	5	4.6	4.6	100.0
	Total	109	100.0	100.0	

Res_TDF

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Absent	107	98.2	98.2	98.2
	Present	2	1.8	1.8	100.0
	Total	109	100.0	100.0	

Res_NVP_EFV

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Absent	79	72.5	72.5	72.5
	Present	30	27.5	27.5	100.0
	Total	109	100.0	100.0	

Table 5: resistance frequencies to some specific drugs

With the prevalence of this mutation in about a quarter of the HIV positive naïve population , the initiated ART will not maximally suppress the viral replication.

2.0 Frequencies of resistance mutations and polymorphisms

The 425 specimens were studied by stratifying them into 4 subgroups for each association:

- who presented a given resistance mutation and a given polymorphism,
- who only had the mutation,
- who only the polymorphism,
- who none of the two.

This is to test every possible mutation / polymorphism association between the genetic variants included in the study.

For each pair of data, the statistical significance of the difference in the relative frequencies in the 4 subgroups was calculated using a statistical method (two-frequency test), in order to highlight any positive or negative correlations between polymorphisms and known resistance mutations. Only the correlations that showed a high degree of significance were considered.

Resistance mutations were subdivided into some groups: mutations of the NAM-1 and NAM-2 pathways, the M184V mutation, the other resistances against NRTIs and those against NNRTIs. Mutations for Protease Inhibitors (PIs) that emerged from sequencing are also reported, although they have not been considered for association studies in this work.

2.1 Frequency of polymorphisms

Of the 216 codons subjected to nucleotide variations found through all the sequencing products in the RT region, the most represented ones were selected to ensure greater statistical significance, reducing the number to 22 codons.

Polymorphisms at frequencies greater than 10% were included in the study, with the exception of the K43E / D mutation, for which a previous study had shown interesting results and, despite being present in less than 10% of the sequences, was found with an acceptable frequency .

	ABSOLUTE FREQUENCIES (n=425)	PERCENTAGE FREQUENCIES %
T39E/D	345	81.18
E40A/D/H	62	14.59
K43/E/D	20	4.71
S48/T/Q	346	81.41
V60I/V	106	24.94
E122K/G/P	85	20.00
D123S/G	272	64.00
I135T/M	133	31.29
S162Y/D/H	97	22.82
K166R/E	49	11.53
K173T/A/D	389	91.53
Q174K/R	234	55.06
D177E/K/G	314	73.88

I178L/M	130	30.59
G196R/M/E	64	15.06
T200A	406	95.53
I202I/V/M	44	10.35
Q207E/D	393	92.47
R211K	303	71.29
L214F	331	77.88
L228R	50	11.76
V245Q/T/E	359	84.47

Many of these polymorphisms were found to be present in very high percentages, making it part of the consensus sequence for the subtype C. They were however included in the study to investigate the existence of any natural predispositions related to those codons. Many of the mutations taken into consideration from the previous work and found to be involved in significant associations with resistance mutations, were found in very low percentages in the HIV subtype C samples, so that they could not be considered statistically significant and excluded from the elaborations. However, these differences are equally important, because they suggest that the different subtypes could use different mutational pathways during the acquisition of the resistance.

2.2 Associations between polymorphisms and resistance mutations

To test the statistical significance of all possible associations, the sample cohort was divided between those who tested positive for a new mutation or polymorphism, and those who did not present the new mutation or polymorphism. The incidence frequencies of the various resistance mutations in the two separate groups were then calculated to determine the presence and strength of the associations.

We have thus identified positive associations, that is, the presence of the new mutation or polymorphism is accompanied by a greater frequency of resistance mutations; and negative associations, in which the presence of polymorphism is linked to a lower frequency of resistance mutations, suggesting the possibility that the two types of mutation may be to some extent mutually excludable.
The following tables show all the associations founded and subdivided by the various groups of resistance mutations defined previously.

NAM-1				
	Positive associations		**negative associations**	
M41L	V60I I178L/M L214F L228R V245Q	P<0,001 P<0,05 P<0,05 P<0,05 P<0,05	S48T/Q	P<0,05
L210W	V60I G196R/M/E L228R	P<0,05 P<0,05 P<0,01	S48T/Q	P<0,001
T215Y	V60I	P<0,05		

The association of V60I with all three mutations of the NAM-1 pathway, the (positive) association of L228R and (negative) of S48T/Q with M41L and L210W are very significant.

NAM-2		
	Positive associations	**negative associations**

D67N/G	V60I L228R	P<0,001 P<0,01	S48T/Q K173T	P<0,001 P<0,005
K70R/E/T	V60I I135T/M I178L/M L228R	P<0,05 P<0,01 P<0,05 P<0,001	S48T/Q K173T V245Q	P<0,001 P<0,005 P<0,05
T215F/I	L228R	P<0,01	S48T/Q	P<0,01
K219Q/E	I135T/M L228R	P<0,05 P<0,001	S48T/Q K173T L214K	P<0,001 P<0,01 P<0,001

Very significant associations are found between the L228R and all 4 mutations of the NAM-2 pathway, and the associations of V60I and I135T/M with two mutations each. The negative associations with S48T/Q and K173T are also very evident.

M184V (Lamivudine)				
	Positive associations		**negative associations**	
M184V	I135T/M L228R	P<0,001 P<0,05	S48T/Q K166R/E L214F V245Q	P<0,05 P<0,05 P<0,05 P<0,001

The most significant associations are those with I135T/M (positive) and with V245Q (negative).

Other NRTI

	Positive associations		negative associations	
E44D	V60I K166R/E	P<0,05 P<0,005	S48T/Q	P<0,01
A62V	I135T/M R211T/S	P<0,05 P<0,001	R211K	P<0,05
K65R	Q174K I202V/M	P<0,05 P<0,05	V60I/V	P<0,05
T69D/N	V60I L228R	P<0,05 P<0,001	K173T L214F	P<0,001 P<0,05
L74I			V245Q	P<0,05
V75T/I	G196R/M/E	P<0,01	S48T/Q T200A	P<0,001 P<0,001
F77L	E122K R211T/S	P<0,05 P<0,05	T39E/D T200A	P<0,05 P<0,01
V90I/W	T39E/D S162Y/D/H I202V/M R211K	P<0,05 P<0,01 P<0,05 P<0,05	S48T/Q	P<0,01
Y115F				
F116Y	G196R/M/E I202V/M	P<0,05 P<0,005	D123S/G T200A	P<0,05 P<0,05
V118I	V60I D123S/G	P<0,01 P<0,05	S48T/Q I177E/K/G T200A L214F	P<0,05 P<0,001 P<0,05 P<0,001
Q151M	I135T/M G196R/M/E R211T/S I202V/M	P<0,05 P<0,001 P<0,001 P<0,001		

Among the most relevant associations there is a good level of significance between L228R (positive) and K173T (negative) with the T69D/N mutation. A good number of negative associations for S48T/Q and T200A are also observed.

NNRTI				
	Positive associations		**negative associations**	
A98G	V60I I178L/M R211K L228R	P<0,001 P<0,05 P<0,01 P<0,001	S48T/Q	P<0,05
K101H/E/P	V60I E122K/G/P	P<0,05 P<0,05	S48T/Q	P<0,01
K103H/N	T39E/D I135T/M	P<0,01 P<0,05	G196R/M/E	P<0,05
V106M			L214F	P<0,005
V106A	I202V/M R211T/S L228R	P<0,01 P<0,001 P<0,05	S48T/Q L214F	P<0,01 P<0,001
V108I	V60I D123S/G I135T/M I178L/M L228R	P<0,01 P<0,05 P<0,005 P<0,05 P<0,001	S48T/Q K173T	P<0,001 P<0,001
Y181C/V	I135T/M K166R/E I178L/M R211K L228R	P<0,05 P<0,05 P<0,005 P<0,001 P<0,001	S48T/Q	P<0,001
Y188L/H	I178L/M	P<0,05		
G190A/S	E122K/G/P Q174K	P<0,05 P<0,001	K166R/E	P<0,01
P225H	L228R	P<0,01	V245Q	P<0,05

Many relevant associations emerged from the comparisons between resistance mutations to NNRTIs and new polymorphisms. In particular the associations of: T39E/D with K103H/N, R211K and L228R with Y181C/V, E122K and Q174K with G190A/S. Moreover, S48T/Q is also associated in this case in a negative way with different resistance mutations.

The associations highlighted by a previous study on subtype B (**154**) have been collected in this table and combined with the most solid evidence emerging from this work

	Previous study	**Experimental study**
Positive association with NAM-1	T39A K43E/Q/N H208Y R211K L214F K223E/Q	V60I (L228R)
Positive association with NAM-2	E203K/D L214L D218E L228H	L228R (V60I) (I135T/M)
Positive association with M184V	K101E/Q/P E203K/D L214L D218E	I135T/M L228R
Positive association for other mutation NRTI	T39A K43E/Q/N K101E/Q/P E122E/P E203K/D H208Y R211K L214L D218E H221Y K223E/Q	T39A V60I I135T/M G196R/M/E I202V/M R211T/S
Negative association with NAM1		S48Q/T
Negative association with NAM2	R83K	S48Q/T K173T
Negative association with M184V		V245Q S48T/Q K166R/E L214F
Negative association for other mutation with NRTI		S48T/Q K173T I177E/K/G T200A L214F

The only concordances found in the comparison with the previous study concern the association of L228R with the mutations of the NAM-2 pathway and of T39A with the mutations for NRTI.

It is to be considered that a good part of the new mutations considered from the previous study have not been analyzed because they were found to be too low, and therefore cannot be reconfirmed by our study. Vice versa, many of the high frequency polymorphisms included in our study did not appear in the work on subtype B.

NNRTI	
Positive association	**Negative association**
V60I/V E122K/G/P I135T/M Q174K I178L/M R211K L228R	S48T/Q K166R/E L214F

The associations with the NNRTI mutations have not been studied in the previous work on the subtype, so there is no direct comparison.

CONCLUSIONS

The study showed that in Malawi there is an increase in NNRTI resistance and that the activity of NNRTIs is compromised because of the high level of threshold NNRTI resistance.
However in developing African countries, the majority of first line ART regimens are NVP/EFV based, even if these drugs are responsible for selection of mutations to the entire class of NNRTI (K103N, Y181C and G190A).
It should be considered that in Malawi, like in developing countries , the resistance test in a patient that has to start the ART therapy is not provided. This could cause a choice of a sub-optimal therapy for the patient since the viral sequencing is done only in case of treatment failure and the epidemiological surveillance is just starting now.
The study confirms a review of treatment protocols because the use of NNRTIs could be the cause in the increase of resistance, and confirms the need of the planned shift to a first line with integrase inhibitors drugs.
The study suggests also that a sequencing test should be done before starting the therapy in order to choose an optimal treatment protocol and avoid the insurgence of problems during the patient care program. An affordable sequencing method could support this approach.
Given such levels of HIVDR prevalence, HIVDR testing and surveillance capacity in Malawi should be prioritized as the universal access will be adopted for all HIV people.

Following WHO guidance, NNRTI TDR was classified in Malawi as moderate (5%– 15%) but we found a high percentage of 29% for NNRTI resistance mutations in our cohort.
NRTI and PI TDR was classified as low (<5%) but in our study we found a moderate percentage of 11,9% for NRTI. In the case of PI, the percentage detected in our study was 1,8% corresponding to the WHO low classification. The differences could also depend on the sample size since our population was composed by 109 patients while WHO analyzed in Malawi only 26 sequences.

TDR	WHO Malawi classification	DREAM RESULTS
NRTI	low	moderate
NNRTI	moderate	high

WHO classification:<5% low, between 5 and 15 % moderate, >20% high.

The sequencing of 425 samples of HIV-infected subjects of subtype C residing in Malawi allowed us to determine the frequencies of classical resistance mutations and other non-characterized mutations, or polymorphisms. We investigated whether some of these polymorphisms are associated with the presence of resistance mutations.

We also performed a comparison with a previous study performed in Novara on subjects infected with HIV strains of subtype B, to highlight any differences with the already known associations. From the information obtained, the polymorphisms studied have been grouped into some groups.

Many of the polymorphisms considered in our study have never been shown in association with resistance mutations in subtype B. One of the most exemplary cases is S48T/Q, which is negatively associated with mutations of all the considered resistance classes: NAM-pathways (NAM-1 and NAM-2), the M184V and the other mutations for NRTI. S48T/Q would therefore play a role of "good mutation", i.e. its presence seems to be protective against various resistance mutations.

In some cases a discordant behavior has been evidenced in the associations between subtype C and B. The polymorphisms T39A, K43E/D, R211K and L214F are associated in the subtype B to the mutation pathway NAM-1, while in the samples of the subtype C there are no association with those mutations. The M184V resistance mutation due to its ability to induce complete resistance alone to Lamivudine and Emtricitabine was considered separately from other NRTI mutations. None of the associations related to it in subtype B were confirmed, while new ones emerged, positive for I135T/M and L228R, and negative for V245Q, S48T/Q, K166R/E and L214F. All these considerations underline how subtypes B and C are characterized by substantially different mutational profiles. It should be noted that many of the polymorphisms taken into consideration by the previous study were not found with sufficient frequencies to be included in this work, and vice versa, many of the polymorphisms evaluated in the present work had not been previously. However this proves once more the hypothesis that mutations and polymorphisms arise differently between the two subtypes. The associations regarding resistance mutations for NNRTI with subtype B could not be compared because they have not yet been investigated, but equally interesting data emerged.

Three polymorphisms positively associated to mutations that could determine a reduction in susceptibility to Efavirenz (EFV) and Nevirapine (NVP) were found with very high frequency (81.2% for T39E/D associated to K103H/N, 55.0% for Q174K associated to G190A and 71.3% for R211K strongly associated to Y181C) enough to be considered part of the consensus sequence of the subtype C.For their association with resistance mutations, it could be suspected that they have a role in making the subtype C naturally more likely to develop resistance to NNRTIs. An interesting consideration can also be made regarding the synonymous V106V mutation. The GTA variant is common in subtype B and "prepares" this codon for a possible evolution to V106A, responsible for resistance to Nevirapine. Instead, in subtype C we find the GTG triplet, which offers a lower genetic barrier for the V106M mutation, responsible for NNRTI multi-resistance.

Combining the results we obtained through the naive patients cohort and the polymorphisms analysis. we can conclude that guidelines for HIV treatment should be globally revised.

Indeed, studying the cohort of naive patients, we identified an important frequency on some drug resistance mutations and we discovered also different polymorphisms pathways between subtype B and C: our hypothesis is that different subtypes behave differently in the induction of the resistances.

The fact that polymorphisms and resistance mutations are so different between B and C subtypes, indicates that mutational pathways are different and it is possible that while treatment with NVP and EFV in the subtype B rarely induces resistances, maybe in subtype C instead it could favor them more easily.

Since all the scientific studies about treatment guidelines for HIV have always been based preferentially on subtype B virus, diffused in western world, similar studies should be conducted on other subtypes and guidelines should be modified and target subtype-based guidelines should be created.

Future perspectives

Especially for polymorphisms study, were registered some difficulties in collecting samples after initiation of the treatment from patients we enrolled for naive analysis, because following the therapeutic protocol, we have to wait one year to have back the patient for molecular test. For this reason this is still an ongoing activity and it will be object of future researches in our laboratory.

In the same time, we developed an homemade sequencing method that is a very important goal in a setting like Malawi, where commercial methods, developed and successfully used in Western Countries, are not affordable to be used in routine activity. This effort was an ethic commitment for us, in order to allow a country as Malawi to have the possibility to implement a tests also for HIV1 drug resistance diagnostic and studies. For this reason a lot of time, efforts, and grants (external to the PhD course) were employed and we are proud of this result. The next goal is to apply the same in house method to a different type of samples than plasma as Dry Blood Spot (DBS). We will increase dramatically the accessibility to the test (as it’s happening for the viral load test), especially in Malawi, where the population lives in rural areas.

ACKNOWELGEMENTS

I thank all the people and institutions that have contributed to the accomplishment of this academic exercise, which is dedicated to all the clients in the DREAM Program in Africa whose value of life is fundamental for the Community of Sant'Egidio.
In a special way, I express my deepest gratitude to Dr Susanna Ceffa for creating, defining and giving me a unique space and status in the development of laboratory medicine in resource limited countries like Malawi. I have proudly worked as a researcher and manager my success in the execution these roles has relied on the advice and mentorship I got from her. I also owe my greatest appreciation to my colleague and DREAM International Supervisor, Dr Davide Brambilla. Since 2007, Dr Brambilla has been involved in the practical realization of first rate diagnostics as well as reliable and affordable technologies in Africa.
I recognize the technical support and expertise I got from Professor Maurizio Zazzi from University of Siena and Dr Paolo Ravanini from the University of Piemonte Orientale. Your unwavering support was a source of motivation and drive to actualization of my research findings for the people of Africa.
I wish to thank all staff of Blantyre DREAM Laboratory in Malawi and Laboratory of Molecular Biology at University Hospital of Piemonte Orientale in Italy for providing me with an enabling friendly environment to work.
To my wife Lizzie, and my children Pachalo and Pemphero, I appreciate your patience and encouragement when I was immersed in the rigour of this academic pursuit.
I would like to thank the Community of Sant'Egidio for giving me this opportunity to further my studies as a part of a continuous program to develop and advance the role of the laboratories in patient management. DREAM is a program designed for quality, partnerships, participation, public health and applied research. Training and capacity building of local staff is a core element of DREAM model and is a testament to the progressive and visionary agenda in the global fight against HIV infection, communicable and non-communicable diseases in countries with few resources.

Thank you mother Community of Sant'Egidio.

Bibliography

1. **Jones WK and Curran JW.** Epidemiology of AIDS and HIV infection in industrialized countries. In: Broder S, Merigan T.C., Bolognesi D., eds.Textbook of AIDS medicine. Baltimore: Williams and Wilkins 1994, 91-108.
2. **Quinn TC.** Population migration and the spread of types 1 and 2 human immunodeficiency viruses. Natl Acad Sci USA 1994; 91:2407-2414.
3. **McCutchan F.** Understanding the genetic diversity of HIV. AIDS 2000, 14:S31-S44.
4. **Robertson DL, Anderson JP, Bradac JA, et al.** HIV-1 nomenclature proposal. Science 2000; 288:55-6.
5. **Peeters M.** Recombinant HIV sequences: their role in the global epidemic Los Alamos National Laboratory Los Alamos, New Mexico.
6. **Preston BD, Poiesz BJ, Loeb LA.** Fidelity of HIV-1 reverse transcriptase. Science 1988; 242:1168-1172.
7. **Ho DD, Neumann AU, Perelson AS, et al.** Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature 1995; 373:123-6.
8. **Wei X, Ghosh SK, Taylor ME, et al.** Viral dynamics in human immunodeficiency virus type 1 infection. Nature 1995; 373:117-122.
9. **Hu WS, Temin HM.** Retroviral recombination and reverse transcription. Science 1990; 250:1227-1233.
10. **Stuhimann H, Berg P.** Homologous recombination of copackaged retrovirus RNAs during reverse transcription. J Virol 1992; 66:2378-2388.
11. **Cavalli-Sforza L, Menozzi P, Piazza A.** Reconstruction of human evolution: bringing together genetic, archeological, and linguistic data. Proc Natl Acad Sci USA. 1988; 85:6002-6006.
12. **Heyndrickx L, Janssens W, Zekeng L, et al.** Simplified strategy for detection of recombinant human immunodeficiency virus type 1 group M isolates by gag/env heteroduplex mobility assay. Study Group on Heterogeneity of HIV Epidemics in African Cities. J Virol 2000; 74:363-370.
13. **Montavon C, Toure-Kane C, Liegeois F, et al.** Most env and gag subtype A HIV-1 viruses circulating in West and West Centrai Africa are similar to the prototype AG recombinant virus IbNg. J Acquir Immune Defic Syndr 2000; 23:363-374.

14. **Su L, Marcus Graf, Yiianzhi Zhang, et al.** Characterization of a virtually full-length human immunodefìciency virus type 1 genome of a prevalent intersubtype (c/b) recombinant strain in china. J Virol 2000; 74:11367-70.
15. **Piyasirisilp S, McCutchan FE, Carr JK, et al.** A recent outbreak of human immunodefìciency virus type 1 infection in Southern China was initiated by two homogeneous, geographically separated strains, circulating recombinant form AE and a novel BC. J Virol 2000; 74:11286-11295.
16. **Myers G.** HIV: between past and future. AIDS Res Hum Retroviruses 1994; 10:1317-1324.
17. **Myers G, and Pavlakis GN**. Evolutionary potential of complex retroviruses. In: Levy J.A, Ed. The retroviridae. New York: Plenum. Press 1992:51-105.
18. **Korber B, Kuiken C, Foley B, et al.** Human Retroviruses and AIDS 1998- A compilation and analysis of nucleic acic and amino acid sequences. Los Alamos National Laboratory Los Alamos, New Mexico.
19. **Louwagie J, Janssen W, Mascola J, et al.** Genetic diversity of the envelope glycoprotein from human immunodefìciency virus type 1 replication. J Virol 1995; 69:263-271.
20. **Gao F, Yue L, Robertson DL, et al**. Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. J Virol 1994; 68:7433-7447.
21. **Gelderblom HR.** Assembly and morphology of HIV: potential effect of structure on viral function. AIDS 1991; 5:617-638
22. **Ozel M, Pauli G, Gelderblom HR.** The organization of the envelope projections on the surface of HIV. Arch Virol 1988; 100:255-266.
23. **Hoglund S, Ofverstedt LG, Nilsson A, et al**. Spatial visualization of the maturing HIV-1 core and its linkage to the envelope. AIDS Res Hum Retroviruses 1992; 8:1-7.
24. **Yu OC, Matsuda Z, Yu X, et al.** An electrol-lucent region within the virion distinguishes HIV-1 from HIV-2 and simian immunodeficiency viruses. AIDS Res Hum Retroviruses 1994; 10:757-761.
25. **Marquet R, Baudin F, Gabus C, et al**. Dimerization of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanism. Nucleic Acids Res 1991; 19:2349-2357.

26. **Jones KA and Peterlin BM**. Control of RNA iniziation and elongation at the HIV-1 promoter. Annal Rev Biochem 1994; 63:717-743.
27. **Cherrington J and Gamen D**. Regulation of polyadenylation in human immunodeficiency virus (HIV): contributions of promoter proximity and upstream sequences. EMBO J 1992; 11:1513-1524.
28. **Myers G. Korber B, Wain-Hohson S, et al.** Human retroviruses and AIDS. A compilation and analysis of nucleic acid and amino acid sequences. Los Alamos, NM: Alamos National Laboratory, 1994.
29. **Arya SK, Guo C, Jsephs SF, et al.** Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). Science 1985; 229:69-73.
30. **Hauber J, Perkins A, Heimer EP, et al.** Trans-activation of human immunodeficiency virus gene expression is mediated by nuclear events. Proc Natl Acad Sci USA 1987; 84:6364-6368.
31. **Kao SY, Calman AF, Luciw PA, et al.** Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature 1987; 330:489-493.
32. **Parslow TG.** Post-transcriptional regulation of human retroviral gene expression. In: Cullen B.R., ed. Human retroviruses 1993, Oxford: IRL Press:101-136.
33. **Colombini S, Arya S, Reitz MS, et al.** Structure of simian immunodeficiency viruses regulatory genes. Proc Natl Acad Sci USA 1989; 86:4813-4817.
34. **Weiss S, Konig B, Muller HJ, et al.** Synthetic human tRNAlys3 and natural bovine tRNAlys3 interact with HIV-1 reverse transcriptase and serve as specific primers for retro viral cDNA synthesis. Gene 1992; 111:183-197.
35. **Capon DJ and Ward RHR**. The CD4-gpl20 interaction and AIDS pathogenesis. Annual Rev Immunol 1991, 9:649-678.
36. **Maddon PJ, Dalgleih AG, McDougal JS, et al.** The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. Cell 1986; 47:333-348.
37. **Chesebro B, Buller R, Portis J, et al.** Failure of human immunodeficiency virus entry and infection in CD4-positive human brain and skin cells. J Virol 1990; 64:215-221.
38. **Butini L, De Fougerolles AR, Vaccarezza M, et al**. Intercellular adhesion molecules ICAM-1 ICAM-2 and ICAM-3 function as counter-receptors for lymphocyte function-associated molecule 1 in human immunodeficiency virus-mediated syncytia formation. Eur J Immunol 1994; 24:2191-2195.

39. **Harouse JM, Bhat S, Spitalnik SL, et al.** Inhibition of entry of HIV-1 in neural cell lines by antibodies against galactosyl ceramide. Science 1991; 253:320-323.
40. **Bhat S, Spitalnik SL, Gonzalez-Scarano F, et al.** Galactosyl ceramide or a derivative is an essential component of the neural receptor for human immunodeficiency virus type 1 envelope glycoprotein gp120. Proc Natl Acad Sci USA 1991; 88:7131-7134.
41. **Hahn BH, Shaw GM, Taylor ME, et al.** Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk for AIDS. Science 1996; 232:548-53.
42. **Kusumi K, Conway B, Cunningham S, et al.** Human immunodeficiency virus type 1 envelope gene structure and diversity *in vivo* and after cocultivation in vivo. J Virol 1992; 66:875-885.
43. **Peliska JA and Benkovic SJ.** Mechanism of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. Science 1992, 258:1112-1118.
44. **Baltimore D.** RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature 1970; 226:1209-1211.
45. **Bukrinsky MI, Sharova N, McDonald TL, et al.** Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. Proc Natl Acad Sci USA 1993; 90:6125-6129.
46. **Gally P. Swingler S, Aiken C, et al.** HIV-1 infection of non-dividing cells: C- terminal tyrosine phosphorylation of the viral matrix protein is a key regulator. Cell 1995; 80:379-388.
47. **Lewis PF, and Emerman M**. Passage through mitosis is required for oncoretroviruses by not for the human immunodeficiency virus. J Virol 1994; 8:39-48.
48. **Varmus H and Brown P.** Retroviruses. In: Berg D.E., Howe M.M., eds. Mobile DNA. Washington, DC: American Society for Microbiology 1989; 53-108.
49. **Stevens SW and Griffìth JD.** Human immunodeficiency virus type 1 may preferentially integrate into chromatin occupied by L1Hs ripetitive elements. Proc Natl Acad Sci USA 1994; 91:5557-5561.
50. **Contag CH, Dewhurst S, Vigilanti GA, et al.** Simian immunodeficiency virus (SIV) from Oid Work monkeys. In: Gallo RC., Jay G, eds. The human retroviruses. San Diego: Academic. Press. 245-276. 1991.

51. **Kohistaedt LA, Wang J, Friedman JM, et al.** Crystal structure of a 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science 1992; 256:1783-1790.
52. **Wlodawer A and Erickson JW.** Structure-based inhibitors of HIV-1 protease. Ann Rev Biochem 1993; 62:543-585.
53. **Ho DD, Moudgil T, Aia M, et al.** Quantitation of human immunodeficiency virus type 1 in the blood of infected persons. N Engl J Med 1989; 321:1621-25.
54. **Ho DD, Neumann AU, Perelson AS, et al.** Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature 1995; 373:123-126.
55. **Wei X, Ghosh SK, Taylor ME, et al.** Viral dynamics in human immunodeficiency virus type 1 infection. Nature 1995; 373:117-122.
56. **Perelson AS, Neumann AU, Markowitz M, et al.** HIV-1 dynamics in vivo: virion clearance rate, infected cell lifespan, and viral generation time. Science 1996; 271:1582-1586.
57. **Mansky LM, Tennn HM.** Lower *in vivo* mutation rate of Human Immunodeficiency Virus type 1 than that predicted from the fidelity of purified reverse transcriptase. J Virol 1995; 69:5087-5094.
58. **Coffin JM.** HIV population dynamics *in vivo*: implications for genetic variation, pathogenesis, and therapy. Science 1995; 267:483-9.
59. **Levy JA.** HIV and the pathogenesis of AIDS. ASM Press. Washington, D.C. 1994.
60. **Sharp PM, Roberto DL, Geo F, et al.** Origins and diversity of human immunodefìciency viruses. AIDS 1994; 4:S27-S42.
61. **Myers G, Korber BTM, Smith RF, et al.** Human Retroviruses and AIDS Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, NM, 1994.
62. **Korber BTM, Ailen EE, Farmer AD, et al.** Heterogeneity of HIV-1 and HIV-2. AIDS 1995; 9:S5-S18.
63. **Louwangie J, McCutchan FE, Peteers M, et al.** Comparison of *gag* genes from seventy international HIV-1 isolates provides evidence of multiple genetic subtypes. AIDS 1993; 7:769-780.
64. **Wain-Hobson S.** The fastest genome evolution ever described: HIV variation in situ. Curr Opin Genetics and Development 1993; 3:878-883.

65. **Deiwart EL, Shaper EG, Louwagie J, et al.** Genetic relationships determined by a DNA heteroduplex mobility analysis of HIV-1 env genes. Science 1993; 262:1257-1261.
66. **Deiwart EL, Sheppard HW, Walker BD, et al.** Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assay. J Virol 1994; 68:6672-6683.
67. **Weniger BG, Takebe Y, Ou CY, Yamazaki S.** The molecular epidemiology of HIV in Asia. AIDS 1994; 8:S13-S28.
68. **Louwagie J, Janssens W, Mascola J, et al.** Genetic diversity of the envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. J Virol 1995; 69:263-271.
69. **Kalish ML, Baldwin A, Raktham S, et al.** The evolving molecular epidemiology of HIV-1 envelope subtypes in injecting drug users in Bangkok, Thailand: Implications for HIV vaccine trials. AIDS 1995; 9: 1350-1360.
70. **Anderson RM and May RM.** Epidemiological parameters of HIV transmission. Nature 1998; 313: 514-519.
71. **Ho DD, Sarngadharan MG, Hirsch MS, et al.** HIV neutralizing antibodies recognize several conserved domains on the envelope glycoproteins. J Virol 1987; 61:2024-2028.
72. **Ho DD, Byington RE, Schooley RT, et al.** Infrequency of isolation of HTLV-III virus from saliva in AIDS. N Engl J Med 1985; 313: 1606.
73. **Skolnik P, Kosloff B, Bechtel LJ, et al.** Absence of infectious HIV-1 in the urine of seropositive viremic subjects. J Infect Dis 1989 ; 160:1056-60.
74. **Groopman JE, Benz PM, Ferriani R, et al.** Characterization of serum neutralization response to the human immunodefìciency virus (HIV). AIDS Res Hum Retroviruses 1987; 3:71-85.
75. **Rusche JR, Javahetian K, McDanal C, et al.** Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind a 24-amino acid sequence of the viral envelope, gp120. Proc Natl Acad Sci USA 1988; 85:3198-3202.
76. **Allain JP, Laurian Y, Paul DA, et al.** Long term evaluation of HIV antigen and antibodies to p24 and gp41 in patients with hemophilia-potential clinical importance. N Engl Med 1987; 317:1114-21.

77. **De Wolf F, Lange JMA, Houweling JTM, et al.** Numbers of CD4+ cells and the levels of core antigens of and antibodies to the human immunodeficiency virus as predictors of AIDS among seropositive homosexual men. J Infect Dis 1998; 158:615-622.
78. **Albert J, Gaines H, Sonnerborg A, et al.** Isolation of human immunodeficiency virus (HIV) from plasma during primary HIV infection. J Med Virol 1987; 23:67-73.
79. **Fauci AS.** The human immunodeficiency virus: infectivity and mechanism of pathogenesis. Science 1988; 239:617-622.
80. **Fauci AS, Schnittman SM, Poli G, et al.** Immunopathogenic mechanism in human immunodeficiency virus (HIV) infection. Ann Intern Med 1991; 114:678-693.
81. **Buchbinder SP, Katz MH, Hessol NA, et al.** Long term HIV-1 infection without immunologic progression. AIDS 1994; 8:1123-1128.
82. **Lemp GF, Paye F, Rutheford GW, et al.** Progression of AIDS morbidity and mortality in San Francisco. JAMA 1990; 263:1497-1401.
83. **Lifson AR, Buchbinder SP, Sheppard HW, et al.** Long term human immunodeficiency virus infection in asymptomatic homosexual men with normal CD4+ lymphoyte counts: immunologic and virologic characteristics. J Infect Dis 1991 ; 163:959-965.
84. **Fauci AS.** Multifactorial nature of human immunodeficiency virus disease: implications for therapy. Science 1993; 262:1011-1018.
85. **Greene WC.** The molecular biology of human immunodefìciency virus type 1 infection. N Engl J Med 1991; 324:308-317.
86. **Bagnarelli P, Valenza A, Menzo S, et al.** Dynamics of molecular parameters of human immunodefìciency virus type 1 activity in vivo. J Virol 1994; 68:2495-2501.
87. **Carr JK, Salminen MO, Koch C, et al.** Full genome sequences of human immunodeficiency virus type 1 subtype G and A/G intersubtype recombinants. Virology 1998; 247:22-31.
88. **Gao F, Robertson DL, Carruthers CD, et al.** A comprehensive panel of near-full-lenght clones and reference sequences for non-subtype B isolates of human immunodeficiency virus type. J Virol 1998; 72:5680-5698.
89. **McCutchan FE, Salminen MO, Carr JK, et al.** HIV-1 genetic diversity. AIDS 1996; 10:S13-S20.

90. **Korber BKC, Foley B, Hahn B, et al.** Human Retroviruses and AIDS. Los Alamos, NM:Theoretical Biology and Biophysics Group; 1997.
91. **Korber BKC, Hoelscher M.** HIV-1 subtypes: implications for epidemiology, pathogenicity, vaccines, and diagnostic. Workshop of the European Commission (DG XII, INCO-DC) and the Joint United Nations Program on HIV/AIDS. Dar es Salaam, Tanzania.
92. **Robertson DL, Anderson JP, Bradac JA, et al.** HIV-1 nomenclature proposal. In Human Retroviruses and AIDS. Edited by Kuiken CL, Foley B, Hahn B et al. Los Alamos, NM:Theoretical Biology and Biophysics Group;1999.
93. **Peeters M, Sharp P.** The genetic diversity of HIV-1: the moving target. AIDS 2000; 12:334-41.
94. **Workshop report from the European Commission and the Joint United Nations Programme on HIV/AIDS.** HIV-1 subtypes: implications for epidemiology, pathogenicity, vaccines and diagnostics. AIDS 1997, 11:UNAIDS17-UNAIDS36.
95. **Van Harmelen JH, Van der Ryst E, Loubser AS, et al.** A predominantly HIV type 1 subtype C-restricted epidemic in South African urban populations. AIDS Res Hum Retroviruses 1999; 15:395-398.
96. **Artenstein AW, Coppola J, Brown AE, et al.** Multiple introductions of HIV-1 subtype E into the western hemisphere. Lancet 1995; 346:1197-1198.
97. **Brodine SK, Mascola JR, Weiss PJ, et al.** Detection of diverse HIV-1 genetic subtypes in the United States. Lancet 1995, 346:1198-1199.
98. **Brodine SK, Shaffer RA, Starkey MJ, et al.** Drug resistance patterns, genetic subtypes, clinical features, and risk factors in military personnel with HIV-1 seroconversion. Ann Intern Med 1999; 131:502-506.
99. **Womack C, Roth W, Newmann C, et al.** Identification of non-B human immunodeficiency virus type 1 subtypes in rural Geòrgia. J Infect Dis 2001; 183:405-410.
100. **Aiaeus A, Leitner T, Lidman K, et al.** Most HIV-1 genetic subtypes have entered Sweden. AIDS 1997; 11:199-202.
101. **Heyndrickx L, Janssens W, Coppens S, et al.** HIV type 1 C2V3 env diversity among Belgian individuals. AIDS Res Hum Retroviruses 1998; 14:1291-1296.

102.**Lasky M, Perret JL, Peeters M, et al.** Presence of non-B subtypes and divergent subtypes B strains of HIV-1 in individuals infected after overseas deployment. AIDS 1997; 11:43-51.
103.**Simon F, Loussert-Ajaka I, Damond F, et al.** HIV type 1 diversity in northern Paris, France. AIDS Res Hum Retroviruses 1996; 12:1427-1233.
104.**Janini LM, Tanuri A, Schechter M, et al.** Horizontal and vertical transmission of human immunodeficiency virus type 1 dual infections caused by viruses of subtypes B and C. J Infect Dis 1998; 177:227-231.
105.**Marquina S, Leitner T, Rabinovich RD, et al.** Coexistence of subtypes B, F, and as B/F env recombinant of HIV type 1 in Buenos Aires Argentina. AIDS Res Hum Retroviruses 1996; 12:1651-1654.
106.**Russell KL, Carcame C, Watts DM, et al.** Emerging genetic diversity of HIV-1 in South America. AIDS 2000; 14:1785-1791.
107.**McCutchan FE, Carr JK, Bajani M, et al.** Subtype G and multiple forms of A/G intersubtype recombinant human immunodeficiency virus type 1 in Nigeria. Virology 1999; 254:226-234.
108.**Peeters M, Esu-Williams E, Vergne L, et al.** Predominance of subtype A and G HIV type 1 in Nigeria, with geographical differences in their distribution. AIDS Res Hum Retroviruses 2000; 16:315-25.
109.**Renjifo B, Chaplin B, Mwakagile D, et al.** Epidemic expansion of HIV type 1 subtype C and recombinant genotypes in Tanzania. AIDS Res Hum Retroviruses 1998; 14:635-638.
110.**Toure-Kane C, Montavon C, Faye MA, et al.** Identification of all HIV type 1 group M subtypes in Senegal, a country with low and stable seroprevalence. AIDS Res Hum Retroviruses 2000; 16:603-609.
111.**Vidal N, Martino P, Mulanga-Kabeya C, et al.** Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. J Virol 2000; 74:10498-10507.
112.**Vidal N, Mulanga-Kabeya C, Nzila N, et al.** The identification of a complex env subtype E HIV-1 virus from the Democratic Republic of Congo, recombinant with A, G, H, J, K and unknown subtypes. AIDS Res Human Retroviruses 2000; 16:1190-95.

113.**Janssens W, Salminen MO, Laukkanen T, et al.** Near full-length genome analysis of HIV type 1 CRF02_AG subtype C and CRF02_AG subtype G recombinants. AIDS Res Hum Retroviruses 2000, 16:1183-1189.
114.**Jeeninga RE, Hoogenkamp M, Armand-Ugon M, et al.** Functional differences between the long terminal repeat transcriptional promoters of human immunodefìciency virus type 1 subtypes A through G. J Virol 2000; 33:195-97.
115.**Montano MA, Nixon CP, Essex M.** Dysregulation through the NF-kB enhancer and TATA box of the human immunodefìciency virus type 1 subtype E promoter. J Virol 1998; 72:8446-8452.
116.**Montano MA, Nixon CP, Ndung'u T, et al.** Elevated tumor necrosis factor-α activation of human immunodeficiency virus type 1 subtype C in Southern Africa is associated with an NF-kB enhancer gain-of-function. J Infect Dis 2000; 181:76-81.
117.**Broder CC and Jones-Trower A.** Coreceptor use by primate lentiviruses. Los Alamos, the Human Retroviruses and AIDS 1999 compendium, 517-541.
118.**Rezza G and Padian N.** Epidemiology. AIDS 2000; 14: S45-46.
119.**Beyrer C, Razak MH, Lisam K, et al.** Overland heroin traffiking routes and HIV-1 spread in south and south-east Asia. AIDS 2000; 14:75-83.
120.**Kato K, Shiino T, Kusagawa S, et al.** Genetic similarity of HIV type 1 subtype in a recent outbreak among injecting drug users in northern Vietnam to strains in Guangxi Province of Southern China. AIDS Res Hum Retroviruses 1999; 15:1157-1168.
121.**Yu XF, Chen J, Shao Y, et al.** Emerging HIV infections with distinct subtypes of HIV-1 infection among injection drug users from geographically separate locations in Guangxi Province, China. J Acquir Immune Defic Syndr 1999; 22:180-188
122.**Piyasirisilp S, McCutchan F, Carr JK, et al.** Injecting drug users in Guangxi Province, China harbour two highly homogeneous, geographically separated HIV-1 strains: CRFAE and a novel BC recombinant. XIII Intemational Conference on AIDS. Durban, July 2000 [abstract A2051].
123.**Amornkul PN, Tansuphasawadikul S, Limpakarnjanarat K, et al.** Clinical disease associated with HIV-1 subtype B and E infection among 2104 patients in Thailand. AIDS 1999; 13:1963-1969.

124.**Kilmarx PH, Limpakarnjanarat K, Kaewkungwal J, et al.** Disease progression and survival with human immunodeficiency virus type 1 subtype E infection among female sex workers in Thailand. J Infect Dis 2000; 34:221-28.
125.**Neilson JR, John GC, Carr JK, et al.** Subtypes of human immunodeficiency virus type 1 and disease stage among women in Nairobi, Kenya. J Virol 1999; 73:4393-4403.
126.**Kaleebu PFN, Mahe C, Yirrell D, et al.** The role of HIV-1 envelope subtypes A and D on disease progression in a large cohort of HIV-1 positive individuals in Uganda. XIII International Conference on AIDS. Durban, July 2000 [abstract A3 079].
127.**Chew CB, Herring BL, Zheng F, et al.** Comparison of three commercial assay for the quantification of HIV-1 RNA in plasma from individuals infected with different HIV-1 subtypes. J Din Virol 1999; 14:87-94.
128.**Parekh B, Phillips S, Granarie TC, et al.** Impact of HIV type 1 subtype variation on viral RNA quantitation. AIDS Res Hum Retroviruses 1999; 21:340-45.
129.**Pasquier C, Sandres K, Salama G, Puel J.** Using RT-PCR and bDNA assay to measure non-clade B HIV-1 subtype RNA. J Virol Methods 1999; 15:133-142.
130.**Arens M.** Human Immunodeficiency Virus (HIV) and other human retroviruses. In Essential of Diagnostic Virology, STORCH GA (Ed). Churchill Livingstone, New York 2000, pp. 249-270.
131.**Cohen OJ, Fauci AS.** Pathogenesis and medicai aspects HIV-1 infection. In Fieids, Virology, KNIPE DM, HOWLEY PM (Eds), vol. 2, 4th ed. Lippincott Williams & Wilkins, Philadelphia 2001, pp. 2043-2094.
132.**Owens DK, Holodniy M, Garber AM, et al.** Polymerase chain reaction for the diagnosis of HIV infection in adults. A meta-analysis with recommendations for clinical practice and study design. Ann Intern Med 1996; 124:803-815.
133.**Bustin SA, Mueller R.** Real-time reverse transcription PCR (qRT-PCR) and its potential use in clinical diagnosis. Clinical Science 2005; 109:365-379.
134.**Andrade MJ, Kubista M, Bengtsson M, et al.** The real-time polymerase chain reaction. Mol Aspects Med 2006; 27:95-125.
135.**Swanson B.** HIV plasma viral load in the clinical setting: measurement and interpretation. J Assoc Nurses AIDS Care 1997; 8:21-23.
136.**Jerzy K Kulski** Next Generation Sequencing - Advances, Applications and Challenges, Chapter: 1, Publisher: InTech, Editors:, pp.3-60

137.**Gibson**, **Richard & Schmotzer, Christine & Quinones-Mateu, Miguel**. (2014). Next-Generation Sequencing to Help Monitor Patients Infected with HIV: Ready for Clinical Use?. Current infectious disease reports. 16. 401. 10.1007/s11908-014-0401-5.

138.http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf

139.**Fox, Juliea,c; Castro, Hannahb,c; Kaye, Stevea; McClure, Myraa; Weber, Jonathan Na; Fidler, Saraha** on behalf of the UK Collaborative Group on HIV Drug Resistance AIDS Epidemiology of non-B clade forms of HIV-1 in men who have sex with men in the UK: September 24th, 2010 - Volume 24 - Issue 15 - p 2397–2401

140.http://www.who.int/countries/mwi/en/

141. 2014 National HIV and AIDS Estimates, National AIDS Commission, Ministry of Health, Malawi

142.**Eveline Geubbels and Cameron Bowie** Epidemiology of HIV/AIDS in adults in Malawi Malawi Med J. 2006 Sep; 18(3): 111–133.

143.http://www.unaids.org/sites/default/files/country/documents/MWI_narrative_report_2015.pdf

144.**Palombi L, Galluzzo CM, Andreotti M, Liotta G, Jere H, Sagno JB, Luhanga R, Mancinelli S, Amici R, Marazzi MC, Vella S, Giuliano M.** Drug resistance mutations 18 months after discontinuation of nevirapine-based ART for prevention of mother-to-child transmission of HIV in Malawi. J Antimicrob Chemother. 2015 Oct;70(10):2881-4

145.**Mancinelli S et al.** Virological Response and Drug Resistance 1 and 2 Years Post-Partum in HIV-Infected Women Initiated on Life-Long Antiretroviral Therapy in Malawi. AIDS Res Hum Retroviruses. (2016)

146.**Hosseinipour M et al**. J Viral Suppression and HIV Drug Resistance at 6 Months Among Women in Malawi's Option B+ Program: Results From the PURE Malawi Study. Acquir Immune Defic Syndr. (2017)

147.https://hivdb.stanford.edu/

148.**Johnson JA, Li JF, Morris L, et al.** Emergence of drug-resistant HIV-1 after intrapartum administration of single-dose nevirapine is substantially underestimated. J Infect Dis 2005; 192: 16-23.

149.**Eshleman SH, Hoover DR, Chen S, et al.** Nevirapine (NVP) resistance in women with HIV-1 subtype C, compared with subtypes A and D after administration of single-dose NVP. J Infect Dis 2005; 192: 30-6.].
150.**Kantor R, Smeaton L, Vardhanabhuti S, et al.** Pretreatment HIV drug resistance and HIV-1 subtype CARE OF CYCLING (ACTG A5175) clinical trial. Clin Infect Dis 2015; 60: 1541-9.
151.**Theys K, Vercauteren J, Snoeck J, et al.** HIV-1 subtype is an independent predictor of reverse transcriptase mutation K65R in HIV-1 patients treated with combination antiretroviral therapy including tenofovir. Antimicrob Agents Chemother 2013; 57: 1053-6..
152.**White E, Smit E, Churchill D, et al.** No evidence that HIV-1 subtype C infection compromises the efficacy of tenofovir-containing regimens: cohort study in the United Kingdom. J Infect Dis 2016; 214: 1302-8.

153.**World Health Organization, United States Centers for Disease Control and Prevention, The Global Fund to Fight AIDS, Tuberculosis and Malaria:** HIV drug resistance report 2017.
154.**Ravanini P, Godino A, Crabu MG et al**. POSTER P829 4th European Congress of Virology, 2011