Validation of replicative phenotyping to detect and assign HIV-1 resistance in clinical specimens

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von
Séverine Louvel
aus Frankreich

Basel 2009

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. med. Christoph Moroni
PD Dr. Thomas Klimkait
Prof. Dr. Kurt Ballmer-Hofer

Basel, den 13 November 2007

Prof. H-P Hauri
Dekan

«Ce n'est pas dans la science qu'est le bonheur, mais dans l'acquisition de la science.»

Edgar Allan Poe - *Le Pouvoir des mots*

A mes parents et mes frères,
à Aziz,
à mes amis.

Abbreviations

AIDS	Acquired ImmunoDeficiency Syndrome
ANRS	*Agence Nationale de Recherches sur le Sida*
ARV	Antiretroviral
AZT	Zidovudine
Bp	Base pairs
BSA	Bovine Serum Albumin
CiP	Calf intestinal Phosphatase
CRF	Circulating Recombinant Form
Ct	threshold cycles
ddI	Didanosine
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
EDTA	Ethylene Diamine Tetra Acidic Acid
FDA	Food and Drug Administration
HAART	Highly Active Anti-Retroviral Therapy
HIV	Human Immunodeficiency Virus
HLA	Human Leucocyte Antigen
HR1/2	Heptad Repeat
IC_{50}	Half maximal inhibitory concentration
IN	Integrase
kb	Kilo base pairs
kD	Kilo Dalton
LB	Lysogeny/Luria Broth
LTR	Long Terminal Repeat
NVP	Nevirapine
NNRTI	Non-Nucleoside Reverse Transcriptase Inhibitor
NRTI	Nucleoside/Nucleotide Reverse Transcriptase Inhibitor
ONPG	o-Nitrophenyl-β-D-Galactopyranoside
OR	Odds ratio
PBL	Peripheral Blood Lymphocytes
PBS	Phosphate Buffer Saline
Pen/Strep	Penicillin/Streptomycin
PI	Protease Inhibitor
PR	Protease
RAM	Resistance Associated Mutation
Rf	Resistance factor
RNase	Ribonuclease

RPMI-1640	medium developed at Roswell Park Memorial Institute
RT	Reverse Transcriptase, Reverse Transcription
SD	Standard Deviation
SQV	Saquinavir
Stanford DB	HIV drug resistance DataBase from Stanford University
TBE	Tris Borate EDTA
TE	Tris EDTA
3TC	Lamivudine
TAM	Thymidine Analog Mutation
VL	Viral Load
Vs	versus
w/o	without
WT	Wild-type

Table of contents

Acknowledgements

It is a healthy practice for such a work to thank at the beginning all those who, more or less directly, contributed to make it possible. Even if in my case, this list can seem longer than usual, it is with my sharpest and most sincere enthusiasm I would like to give credit to all those whom with their manner helped me to conclude this thesis. Then I wish to express my deep gratefulness:

- to Professor Thomas Klimkait, my Thesis advisor, for his trust, his patience and his generosity which contributed to this fascinating project. The full confidence that he granted me since my DESS (Diplôme d'Etudes Supérieures Spécialisées) training course, enabled me to develop a personal plan for my Thesis, specific to my aspirations. I would also like to thank him for his time as he granted me throughout these years, to have believed in my capacities and to have provided me excellent logistic conditions. Moreover, the advice throughout the writing was always clear and concise, facilitating largely the task to me leading to the production of this work. During these four years he enabled me to carry out several travels while leaving me a great freedom and high responsibilities for which I hope to have shaped up.

- to Professor Christoph Moroni, my Faculty Responsible, who gave me the opportunity to work at the Institute for Medical Microbiology in collaboration with InPheno. He allowed me to improve my scientific knowledge and to contribute to this interesting project by giving me the opportunity to join the program of the Basel University. I appreciated his advice during our personal meetings and his scientific questions helped me to develop an inquisitive mind.

- to Professor Kurt Ballmer-Hofer who accepted to read and judge my work as Referee

- to Professor Markus Affolter being my Chairman today.

I want to thank especially Doctors François Hamy and Vincent Vidal for sharing their broad scientific knowledge with me. Each of them deserves more particular thanks:

- François, you were behind my decision to come to Switzerland and more particularly to InPheno. From the beginning, you have been present for all my questions and you were a good supervisor during my DESS and my Thesis. Thank you for your support and your unnumbered ideas leading to new avenues to follow.

- Vincent, I am grateful to have shared a project with you and I appreciated your availability, your generous encouragements and the confidence that you provided me; all this with your warm accent from Toulouse.

The quality and the continuity of the data presented here are the fruit of a team work. They all contributed to a very good atmosphere in the lab and outside:

- Tatjana, more than a colleague you have now for ever a special place in my heart. You gave me the opportunity to discover your beautiful country, Croatia. I wish you all the best for your Thesis and your private life. Of course, my thanks go also to Daniel with whom I shared the hard life of a thesis student and good moments as well. Hvala!!

- Gabi, I am very happy to have met you and your “petits monstres”. Thanks for your moral support and the help in my writing. Your presence cheered me up every morning during that period and I appreciated our shared time.

- Stéphane and Vincent F., the computer team. Without your help I would never have obtained all the results shown in my Thesis. “Je vous souhaite bon vent, collègues Breton et Normand.”

- Fabian, your open-mindedness, due to your travel mood, and your kindness made you very special in the lab. You were “une bouffée d’air frais” during my working time preventing burning brain.

- Alessio and Angelika, “babies” of the lab. Both of you brought fun and were very helpful whenever needed.

- Omar, the ray of sunshine from Cuba. Now it is your turn to discover the great time of a PhD student.

Many thanks to all the IMM team and InPheno students whom I met during this time. I would like to give a special dedication to Claudia J., Katarina and Vincent B. who were a part of the team at the very beginning providing me advice and good practices.

I cordially thank the Swiss National Fond for financing my project and the DART 2006 (Drug Development for AntiRetroviral Therapies) for rewarding me with a prize for my thesis work in December 2006.

I am grateful to my friends, specially my lovely Rita for her good mood and help; and to my family: my brothers, my granny and my parents for their constant support, love and patience during all these years of studies, as well as my family in-law. “Je ne pourrai jamais oublier le soutien et l’aide très chère que vous m’avez tous apportés, ma merveilleuse famille.”

Finally I specially thank Aziz who was beside me for this long travel. “Mon amour sans toi je serais probablement allée aussi loin mais pas avec autant de bonheur, d’envie et d’énergie. Tu as su au quotidien me valoriser et me rassurer dans les moments de doute. Je suis moi aussi très fière d’avoir partagé cette aventure avec toi et de nouveaux horizons vont désormais s’offrir à nous. Je tiens à te dédicacer tout particulièrement cette oeuvre, merci sincèrement.”

I. Introduction

I.1. Acquired ImmunoDeficiency Syndrome (AIDS)

The AIDS epidemic is one of the most significant health-related events of the 20th century. Acquired immunodeficiency syndrome was first recognised as a new disease in the United States when clinicians in New York, Los Angeles, and San Francisco began to see young, homosexual men with Pneumocystis carinii pneumonia (PCP) and Kaposi's sarcoma (KS), unusual diseases for young adults not known to be immunosuppressed. The first report in the medical literature that alerted the world to this new immunodeficiency syndrome appeared in June of 1981 and described five young, homosexual men in Los Angeles with PCP[1]. Other reports followed of a similar syndrome in injecting drug users[2]. The disease was originally dubbed GRID, or Gay-Related Immune Deficiency, but health authorities soon realised that nearly half of the people identified with the syndrome were not homosexual men. In 1982, the Centers for Disease Control and Prevention (CDC) introduced the term AIDS to describe the newly recognised syndrome[3].

The prominence of homosexual men and injecting drug users in the early cases of AIDS suggested an agent that was both blood borne and sexually transmitted[4]. The majority of researchers thought that the likely agent was a sexually transmitted virus that would be found in the peripheral blood. The Human Immunodeficiency Virus (HIV) was first isolated in France in 1983 by Françoise Barré-Sinoussi in the laboratory of Luc Montagnier at the Pasteur Institute as lymphadenopathy-associated virus (LAV)[5]. Strong evidence that it was the AIDS virus did not appear until 1984, when four papers were published in one issue of Science by Robert Gallo and colleagues, who designated their isolate HTLV-III (Human T cell Leukemia/Lymphoma Virus)[6,7]. The International Committee on the Taxonomy of Viruses chose the designation Human Immunodeficiency Virus in 1986[8]. With the discovery by Montagnier's group in late 1986 of the related HIV-2 virus in West Africa, the original virus became HIV-1[9].

Retrospective studies on frozen tissue and serum samples have identified cases apparently fitting the AIDS definition as early as the 1950s and 1960s[10]: (i) a plasma sample taken in 1959 from an adult male living in the Democratic Republic of Congo, (ii) tissue samples from a 15-year-old black male from St. Louis who died in 1969 of an aggressive, disseminated KS[11], and (iii) tissue samples from a Norwegian sailor who died around 1976[10]. In each of them an HIV infection has been detected retrospectively.

Since its first recognition in 1981 and according to the latest epidemiological studies by WHO (World Health Organization) and UNAIDS (Joint United Nations Programme on HIV/AIDS) in 2006, AIDS has killed more than 25 million people making it one of the most destructive pandemics in recorded history. Actually, 40 million people are worldwide infected by HIV, and the highest prevalence is observed in Sub-Saharan Africa (Figure 1)[12].

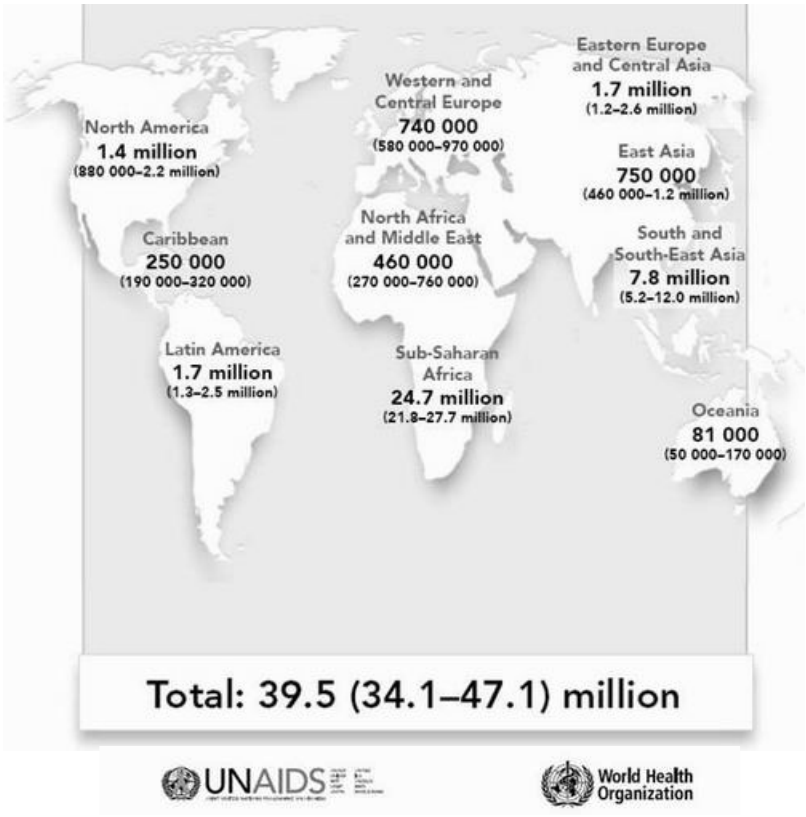


Figure 1. *Global estimates of HIV/AIDS infections as of December 2006.*

In 2006 alone, AIDS claimed an estimated 4.3 million lives, of which more than 530'000 were children. It is estimated that about 0.6% of the world's living population is infected with HIV. According to current estimates, HIV is set to infect 90 million people in Africa, resulting in a minimum estimate of 18 million orphans. Routine access to antiretroviral medication is not available in all countries, leading to still sharply rising levels of infection. Globally and in every region, more adult women (15 years or older) than ever before are now living with HIV (17.7 million or 44.8%).

An overview of HIV information about classification, virion structure, viral genome, regulation of expression and tropism is provided in the Background chapter at the end.

I.2. HIV Immunopathogenesis

Whereas HIV replication is thought to take place in activated $CD4^+$ T-lymphocytes in lymphoid tissue, other cell populations may become infected and play a role in infection persistence, like resting T-cells in the G_0 phase. Virtually every arm of the immune response may be affected by HIV infection: $CD4^+$ and $CD8^+$ T-cells, B lymphocytes, monocytes and macrophages, dendritic cells, and natural killer cells. Nevertheless the best indicator of disease course is the number of $CD4^+$ T-cells, which roughly defines the onset of AIDS when it reaches 200 cells/μL (the normal value is 1000-1200 cells per mL). The magnitude of HIV replication, reflected by plasma HIV RNA levels, is the second predictor of disease progression. The stage of infection can be determined by measuring the patient's $CD4^+$ T cell count, and the level of HIV in the blood. The initial infection with HIV generally occurs after contact with body fluids from an infected person. The first stage of infection, the primary, or acute infection, is a period of rapid viral replication that immediately follows the individual's exposure to HIV leading to an abundance of virus in the peripheral blood with levels of HIV commonly exceeding one million particles per mL (Figure 2)[13]. This response is accompanied by a marked drop in the number of circulating $CD4^+$ T cells. This acute viremia is associated in virtually all patients with the activation of $CD8^+$ T cells, which kill HIV-infected cells, and a subsequent immune response, or seroconversion.

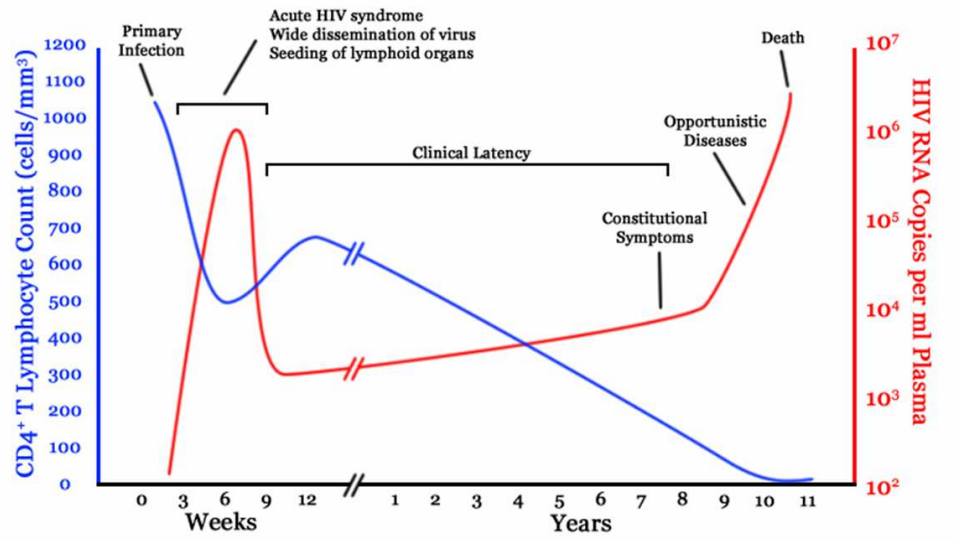


Figure 2. *Relationship between HIV copies (viral load) and CD4 counts over the average course of an untreated HIV infection; any particular individual's disease course may vary considerably.*
[http://www.edinformatics.com/biotechnology/hiv.htm]

The $CD8^+$ T cell response is thought to be important in controlling virus levels, which peak and then decline, as the $CD4^+$ T cell counts rebound to around 800 cells per mL. A good $CD8^+$ T cell response has been linked to slower disease progression and a better prognosis, though it does not eliminate the virus[14]. During this period (usually 2-4 weeks post-exposure) most individuals (80 to 90%) develop an influenza- or mononucleosis-like illness called acute HIV infection syndrome. Because of the non-specific nature of these illnesses, they are often not recognised as signs of HIV infection.

The following strong immune defence reduces the number of the viral particles in the blood stream, marking the start of the infection's clinical latency stage. Clinical latency can vary between two weeks and 20 years. During this early phase of infection, HIV is active within lymphoid organs, where large amounts of virus become trapped in the follicular dendritic cells (FDC) network[15]. The surrounding tissues that are rich in $CD4^+$ T cells may also become infected, and viral particles accumulate both in infected cells and as free virus. Individuals who are in this phase are infectious and can transmit HIV.

When $CD4^+$ T cell numbers decline below a critical level, cell-mediated immunity is lost, and infections with a variety of opportunistic microbes appear. The first symptoms often include moderate and unexplained weight loss, recurring respiratory tract infections (such as sinusitis, bronchitis, otitis media, pharyngitis), skin rashes, and oral ulcerations. Common opportunistic infections and tumors, most of which are normally controlled by robust $CD4^+$ T cell-mediated immunity, then start to affect the patient.

Like other infectious diseases many factors determine AIDS course and onset: the route and the size of the inoculum but also the virulence of the virus and the nature of the host[16]. Especially for viruses, heterogeneities in viral replicative capacity are more determinant than virulence itself. Concerning HIV, the parenteral entry is the most dangerous. The risk after transmucous infection varies according to the site of exposure, with risks of transmission through rectal exposure exceeding the risks of transmission through vaginal exposure and both of the above exceeding the risks of transmission across oral mucus. Mother-to-infant transmission is enhanced in women with high levels of plasma HIV RNA, which means that the viral load (VL) influences the clinical course[17]. Furthermore, data have been collected linking host genes with the course of AIDS (summarised in *Background, HIV tropism*).

I.3. Replication Cycle of HIV

I.3.1. Binding and Fusion

HIV is a complex retrovirus that is able to enter into macrophages and $CD4^+$ T cells by the absorption of glycoproteins on its surface to receptors on the targeted cell. gp120 is a glycoprotein embedded in the HIV envelope, and the first step in fusion involves the high-affinity attachment of its CD4 binding domains to CD4. Once gp120 is bound to the CD4 protein (Figure 3A), the envelope complex undergoes a structural change, exposing the chemokine binding domains of gp120 and allowing them to interact with the target chemokine receptor (generally either CCR5 or CXCR4 but others are known to interact) (Figure 3B). This leads to a more stable two-pronged attachment, which allows the N-terminal fusion peptide gp41 to penetrate the cell membrane (Figure 3C). Repeat sequences in gp41, HR1 and HR2 then interact, causing the collapse of the extracellular portion of gp41 into a hairpin. This loop structure brings the virus and cell membranes into close proximity (Figure 3D), allowing fusion of the membranes and subsequent entry of the viral capsid[18,19].

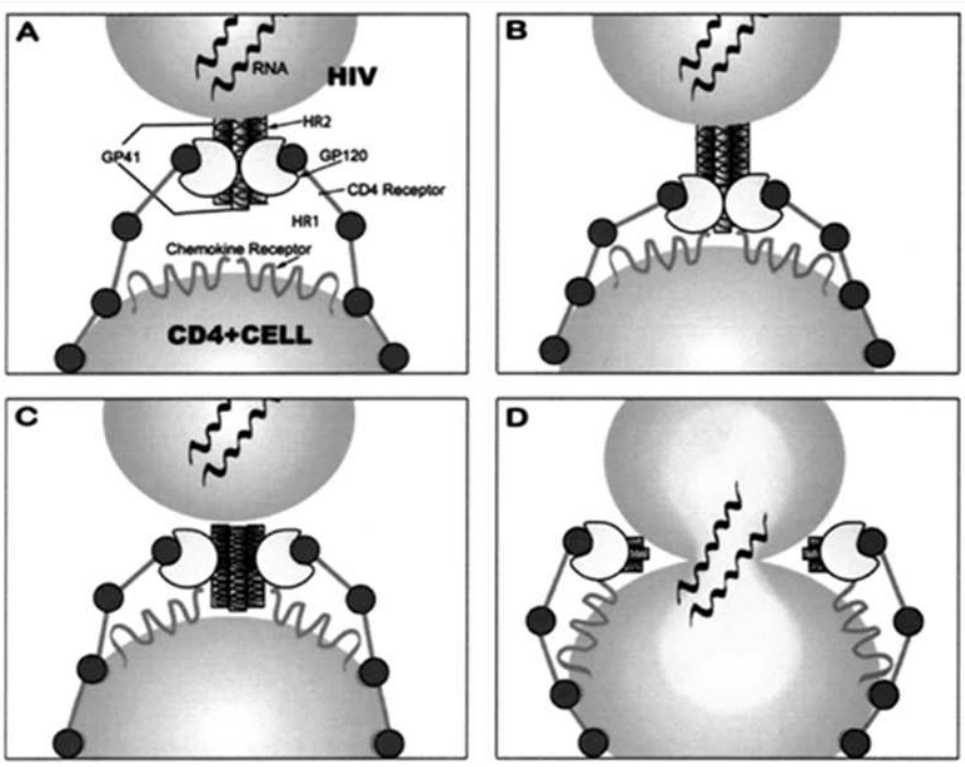


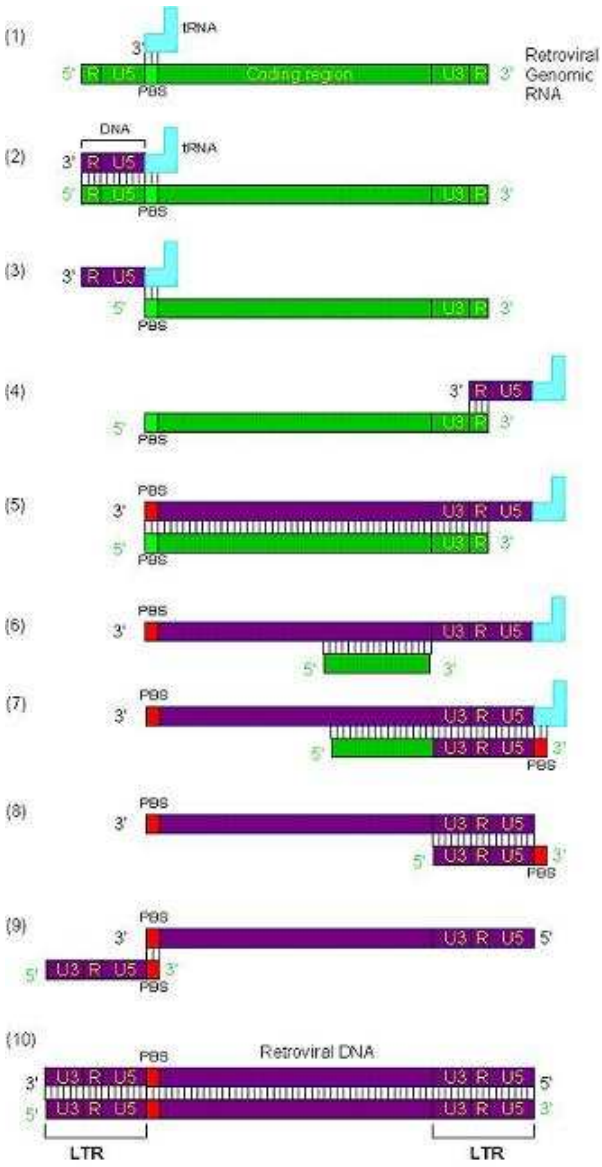
Figure 3. *Schematic representation of the binding and fusion events between the HIV particle and $CD4^+$ T cells.*
[http://www.medscape.com/viewarticle/468807_2]

Once HIV has bound to the target cell, the HIV RNA and various enzymes, including reverse transcriptase (RT), integrase (IN), ribonuclease (RNase) and protease (PR), are injected into the cell[18].

I.3.2. Reverse Transcription

Among the various characteristics of retroviruses, they are (+) RNA viruses whose genome can not serve directly as mRNA. This latter is produced by the cellular machinery without any contribution from viral polymerases. Once the viral capsid enters the cell, an enzyme called reverse transcriptase liberates the single-stranded (+) RNA from the attached viral proteins and copies it into a complementary DNA[20]. For this purpose, a particular cellular tRNA (usually W, P, or K), packaged in the capsid is essential to prime reverse transcription (Figure 4, step 1). The reverse transcriptase then makes a complementary DNA strand to form a double-stranded viral DNA intermediate (vDNA). The cDNA produced by reverse transcription is slightly longer than vRNA, due to the particular mechanism retroviruses use to transcribe RNA into DNA (Figure 4, steps 4 and 9). The vDNA is then transported into the cell nucleus.

Reverse transcription *in vivo* can be divided into ten sequential steps (Figure 4).



1. A $tRNA^{Lys}$ hybridises with the complementary PBS (Primer Binding Site) region;
2. A DNA segment is extended upstream from the tRNA;
3. Viral R and U5 sequences are removed by the RNaseH catalytic domain of RT;
4. First RT jump: DNA hybridises with the remaining R sequence at 3' end;
5. A DNA strand is extended from 3' end;
6. Most viral RNA except a polypurine tract is removed by the RNaseH;
7. A second DNA strand is extended from the viral RNA;
8. Both tRNA and the remaining viral RNA are removed by the RNaseH;
9. Second RT jump: the PBS region of the second strand hybridises with the PBS region of the first strand;
10. Extension on both DNA strands.

Figure 4. *HIV reverse transcription in vivo results in two direct long terminal repeats at the ends of the genome.*

This process of reverse transcription is extremely error-prone, and it is during this step that mutations may occur. In fact, HIV-1 is a quasispecies containing innumerable variants related to the original infecting strain. About 1% of all nucleotide positions in the RT and protease isolates from persons receiving antiretroviral therapy have detectable mixtures identified by population based-sequencing[21]. During antiretroviral therapy, mixtures occur at a higher rate (about 5%) at positions associated with drug resistance.

I.3.3. Integration

Multiple steps in this integration process are catalysed by the HIV-1 integrase[20]. The integration of HIV-1 DNA into the host chromosome is achieved by the integrase protein performing a series of DNA cutting and joining reactions (A-C).

(A) The first step in the integration process is a *3' processing*. This step requires linear double-stranded DNA with sequence specific 3' ends (CAGT), synthesised by reverse transcription from the viral RNA genome. The integrase protein removes two nucleotides from each 3' end of this viral DNA, leaving recessed CA-OH at the 3' ends.

(B) In a second step, termed *strand transfer*, the integrase protein incises chromosomal DNA at a non-defined site and produced staggered cuts, 5 bp apart. A Y-shaped, gapped, recombination intermediate results, with the 5' ends of the viral DNA strands and the 3' ends of target DNA strands remaining unjoined, flanking a gap of 5 bp.

(C) The last step involves host DNA repair synthesis in which the 5 bp gaps between the unjoined strands are filled in and then ligated. Since this process occurs at both cuts flanking the HIV genome, a 5 bp duplication of host DNA is produced at the ends of HIV-1 integration.

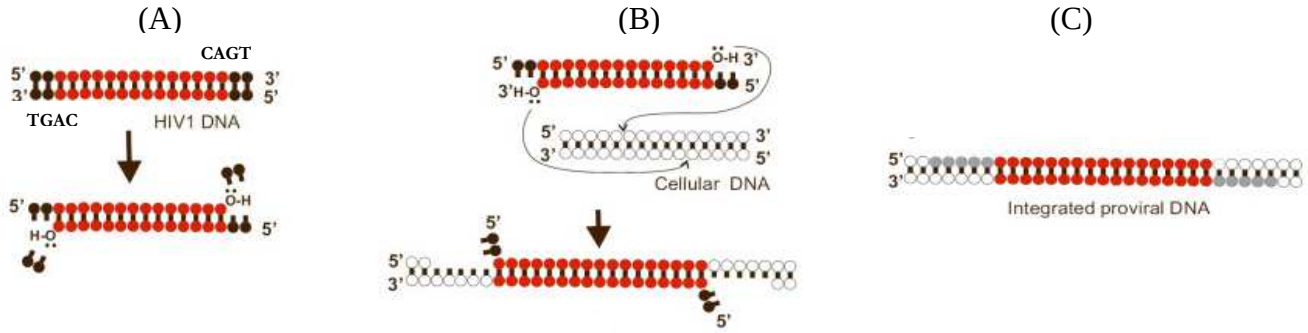


Figure 5. *Sequence of events during HIV-1 integration.*
A) integrase-catalysed 3' processing; B) integrase-catalysed strand transfer; C) DNA repair by cellular enzymes. Parts of the donor DNA that become integrated are shown in red. The acceptor DNA is shown in white. Parts of acceptor DNA repaired following the strand transfer reaction are shown in grey.

This integrated viral DNA may then lie dormant in the latent stage of HIV infection[20]. To actively produce the virus, certain cellular transcription factors need to be present, the most important of which is NF-κB, which is upregulated when T cells become activated[22] (summarised in *Background, Regulation of HIV genome expression*). This means that those cells most likely to be killed by HIV are in fact those currently fighting infection.

In this replication process, the integrated provirus is copied to mRNA which is then spliced, excising all possible introns, and translated in the regulatory proteins Tat (which encourages new virus production) and Rev. As Rev accumulates, it gradually starts to inhibit mRNA splicing[23]. At this stage, the structural proteins Gag and Env are produced from full-length mRNA (summarised in *Background, the organisation of the viral genome*). The full-length RNA is actually the viral genome; it binds to the Gag protein and is packaged into new virus particles.

I.3.4. Budding and Polyprotein Processing

The final step of the viral cycle, assembly of new HIV-1 virions, begins at the plasma membrane of the host cell. The Env polyprotein (gp160) goes through the endoplasmic reticulum (ER) and is transported to the Golgi complex where it is cleaved by a protease and processed into the two HIV envelope glycoproteins gp41 and gp120 (Figure 6). They are transported to the plasma membrane of the host cell where gp41 anchors the gp120 to the membrane of the infected cell. The Gag (p55) and Gag-Pol (p160) polyproteins also associate with the inner surface of the plasma membrane along with the HIV genomic RNA as the forming virion begins to bud from the host cell.

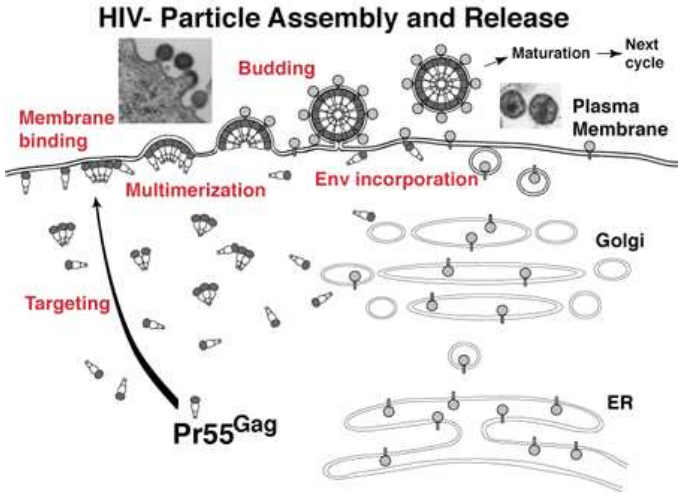


Figure 6. *The HIV-1 assembly and release pathway. [http://www.retrovirus.info/Freed_figure.html]*

Maturation either occurs in the forming bud or in the immature virion after it buds from the host cell. During maturation, HIV proteases cleave the polyproteins into individual functional HIV proteins and enzymes. The various structural components then assemble to produce a mature HIV virion which is now ready to infect another cell.

I.4. Anti-HIV Therapy and Drug resistance

I.4.1. Highly Active Anti-Retroviral Therapy (HAART)

In 1987 the U.S. Food and Drug Administration (FDA) approved the first antiviral drug zidovudine (ZDV; AZT) for use in preventing HIV replication by inhibiting the activity of the reverse transcriptase enzyme[24]. After 1991, several other nucleoside analogues became available, as well as a new class of anti-HIV drugs called the non-nucleoside analogue reverse transcriptase inhibitors (NNRTIs). Those are more quickly activated once inside the bloodstream because they don't have to be phosphorylated. Next to be developed was the class of antiviral drugs known as protease inhibitors, which were distinctly different from the reverse transcriptase inhibitors in that they prevent an already infected cell from producing a new generation of HIV.

However, the persistence of viral replication, the development of resistance and the expansion of the number of antiviral drugs necessitated a shift from monotherapy to combination therapy, in which drugs from two or more classes are used simultaneously. Theses combinations of antiretroviral drugs used in the treatment of HIV infection are called Highly Active Antiretroviral Therapy (HAART). Current HAART generally comprise three antiretroviral agents, usually two nucleoside analogues and either a protease inhibitor or a nonnucleoside reverse-transcriptase inhibitor[25].

The two classes of drugs that target the RT enzyme:

- Nucleoside and nucleotide reverse transcriptase inhibitors (NRTIs): They are prodrugs that only become effective after being converted to triphosphates. Nucleotide analogues require only two instead of three phosphorylation steps. Phosphorylated NRTIs compete with naturally occurring dNTPs (deoxynucleotide triphosphates) for incorporation into the growing DNA molecule. Because NRTIs lack a 3'OH group, their incorporation blocks further elongation of the proviral DNA and leads to interruption of the chain[26]. Today there are eight FDA-approved NRTIs: zidovudine, AZT (1987), didanosine, ddI (1991), zalcitabine, ddC (1992), stavudine, d4T (1994), lamivudine, 3TC (1995), abacavir, ABC (1998), tenofovir, TDF (2001) and emtricitabine, FTC (2003).

- Non-nucleoside RT inhibitors (NNRTIs): They inhibit the viral enzyme reverse transcriptase itself. NNRTIs are small molecules that bind to the hydrophobic pocket close to the catalytic domain of the RT which alter the conformation of the active site, causing inhibition of the enzyme[26]. Three NNRTIs have been approved to date: nevirapine, NVP (1996), delavirdine, DLV (1997) and efavirenz, EFV (1998).

There are three other classes of HIV inhibitors which target the protease enzyme or the entry/fusion step:

- Protease inhibitors (PIs): They bind to the active site of the viral PR hindering the cleavage of the viral precursor gal-pol-polyprotein by the enzyme, thereby producing immature, non-infectious viral particles. Ten PR inhibitors have been approved by the FDA: saquinavir, SQV (1995), ritonavir, RTV (1996), indinavir, IDV (1996), nelfinavir, NFV (1997), amprenavir, APV (1999), lopinavir, LPV (2000), atazanavir, ATV (2003), fosamprenavir, FPV (2003), tipranavir, TPV (2005) and darunavir, DNV (2006).

- Fusion inhibitors (FIs): They differ from NRTIs, NNRTIs and PIs, which block the replication of HIV in the infected cell. Instead, fusion inhibitors prevent HIV from entering its target cells[27]. The fusion inhibitor enfuvirtide (T-20), approved by the FDA in 2003, is currently the only registered fusion inhibitor. T20 is a synthetic peptide consisting of 36 amino acids, mimics the C-terminal HR2 domain of gp41 and competitively binds to HR1. Thus, interactions between HR1 and HR2 are blocked and the conformational change of gp41 that is necessary for fusion of virions to host cells is inhibited[28]. A single amino acid substitution in gp41 can reduce the efficacy of T-20.

- Entry Inhibitors/CCR5 coreceptor antagonist: Maraviroc, approved by FDA in August 2007, is a chemokine receptor antagonist that acts as an HIV entry inhibitor. It is designed to prevent HIV infection of CD4 T cells by blocking the CCR5 from binding to HIV. The FDA approved maraviroc for the use in combination with other antiretroviral medications for the treatment of CCR5-tropic HIV-1 (R5 virus) in adults whose viral loads remain detectable despite existing ARV treatment or who have multi-drug-resistant HIV. Fifty to 60 percent of treatment-experienced patients have a R5 virus.

I.4.2. Drug Resistance

HIV drug resistance must be distinguished from other causes of virologic failure such as non-adherence, insufficient drug levels, or inefficient combination regimens, although all these factors cause virologic failure. The expression “drug resistance” can be misleading, since it seems that resistance is an all-or-nothing phenomenon. Drug resistance is rather a continuous variable, where there is a continuum between a susceptible and a highly resistant state.

The main drivers of the development of HIV drug resistance are the high level of virus production[29] and the high error rate in reverse transcriptase activity[30,31]. These two

combined characteristics ensure that patients have a complex and diverse mixture of viral species, each differing by one or more mutations. If any of these mutations confers a selective advantage to the virus, such as a decrease in its susceptibility to antiviral drugs, the corresponding quasispecies will overtake the others, according to Darwinism[32]. However, it is not simply a case of "the survival of the fittest"; HIV infection results in survival of all major forms that have ever been generated within a patient, and replication favours the form that is fittest under the current conditions. If conditions change, previously archived variants can re-emerge[33,34].

In patients who receive HAART as the first line of anti-retroviral therapy, emergence of viral resistance is possible only if HIV continues to replicate in the presence of levels of drugs that are insufficient to block viral replication but sufficient to exert a positive selective pressure on variants with decreased drug susceptibility. Under these conditions, viruses resistant to all the components of the regimen will gradually emerge[34]. However, several reports have also described the transmission of HIV strains with resistance to single or multiple ARV[35-37]. Although most of those cases involve the transmission of strains of HIV-1 from patients in whom resistance has developed during therapy, some strains of HIV are naturally resistant to some antiretroviral drugs. For example, HIV-2 is intrinsically resistant to most NNRTIs[38] and some subtypes of HIV-1 can be less susceptible to PIs or NNRTIs than the subtype B strain[39-42].

Reduced plasma levels of one or more anti-retroviral drugs are not necessarily the result of poor compliance; they can also result from others factors. Tuberculosis (TB) is one of the main opportunistic infections among HIV-positive patients in the world, especially in sub-Saharan Africa and Asia. Unfortunately, major drug-drug interactions can occur in patients taking TB and HIV therapies, especially between rifamycins, which are the first choice drugs for the TB therapy, and HIV PR inhibitors or NNRTIs[43]. Administration of rifamycins in combination with PR inhibitors (SQV, IDV or RTV) or NNRTIs (NVP or DLV) will greatly reduce the plasma levels of the anti-retroviral agent used[43].

Nonetheless, new drugs are becoming available that appear to be substantially active against strains resistant to multiple drugs. These drugs are either new members of existing structural classes, exhibiting increased potency and improved pharmacokinetic properties, or members of new classes that are not affected by cross-resistances[34]. The new drugs, used in triple class combinations, are saved for later use as salvation therapies, to be introduced in cases of drug resistance or if treatment is first started at a very advanced stage of infection.

To better understand the HIV resistance observed for the two classes of reverse-transcriptase inhibitors which will be described in this study, I will focus on their specific

mechanisms of resistance. Two distinct mechanisms are involved in HIV resistance to NRTIs: impairment of the incorporation of the analogue into DNA and removal of the analogue from the prematurely terminated DNA chain. The resistance to NNRTIs is caused by mutations in the hydrophobic pocket, the binding site of these drugs, which reduce their binding and activity.

- Impairment of analogue incorporation: Several mutations or groups of mutations in reverse transcriptase involve steric hindrance and prevent the enzyme to incorporate a nucleotidic analogue into DNA leading to selective resistance. The most of those mutations are the M184V mutation, the Q151M complex of mutations, and the K65R mutation.

The replacement of the methionine by a valine at position 184 of the RT is the main mutation that confers high-level resistance to 3TC[44]. Methionine 184 is located at the heart of the catalytic site of RT, and its change by a valine, which has a different side chain, interferes with the proper positioning of 3TC triphosphate in the catalytic site[45].

Similarly, the group of mutations referred to as the Q151M complex can confer high-level resistance to almost all NRTIs except 3TC and TDF[46]. This pathway arises under regimens containing d4T and ddI and always starts with the Q151M substitution and is followed by the accumulation of secondary mutations enhancing resistance and increasing enzyme activity[47]. The former residue is located in the immediate vicinity of the nucleotide binding site of RT. This Q151M complex is rarely found in HIV-1 but interestingly it is markedly more frequent in HIV-2.

Another example is the K65R substitution which is occurring in patients treated with TDF or ABC. It gives resistance to most analogues, with the exception of AZT and is an antagonist mutation of the further described TAMs group[48,49].

- Removal of the analogue from the terminated DNA chain: This is associated with a group of mutations commonly termed “thymidine analogue mutations” (TAMs). Their selection is most frequent for drug combinations that include thymidine analogues such as AZT and d4T. They can promote resistance to almost all NRTIs including TDF[50-53]. These mutations occur gradually, and their order of emergence can vary[54,55]. TAMs promote resistance by fostering ATP- or pyrophosphate-mediated removal of nucleoside analogues from the 3' end of the terminated DNA strand[56,57]. ATP and pyrophosphate, which are abundant in normal lymphocytes, do not participate in the DNA-polymerisation reaction, but the structure of a reverse transcriptase expressing TAMs facilitates their entry into a site adjacent to the incorporated analogue[58,59]. In this position, ATP or pyrophosphate can attack the phosphodiester bond that links the analogue to DNA, resulting in removal of the analogue and continuation of strand polymerisation.

- Resistance to NNRTIs: The mutations that are selected for after the failure of NNRTI treatment are all located in the pocket targeted by these compounds and they reduce the affinity of the drug[60-62]. A single mutation can lead to high-level resistance to all NNRTIs, e.g. K103N and V106M, but most frequently the mutations are "drug-dependant". For example, resistance to NVP is often associated with the Y181C mutation, but other mutations such as Y188C, K103N, G190A and V106A also occur. Initial resistance to EFV is generally characterized by the K103N substitution, but the Y188L mutation is also seen. A second generation of NNRTIs, as TMC125, seems to be active against resistant viruses presenting a pattern of mutations involved in the resistance to the first-generated NNRTIs. Nevertheless the accumulation of such mutations will also decrease the effect of the TMC125.

For almost all mutations described above, a "cross-resistance" phenomenon can be observed. Its definition is a resistance to drugs to which a virus has never been exposed, resulting from mutations that have been selected for by the use of another drug. Cross-resistance is always restricted to drugs within a given class of antiretroviral agents, but all three classes of antiretroviral drugs are affected.

I.5. HIV Resistance Tests

Due to the emergence of drug resistant viruses, it became necessary and obvious to have diagnostic tools to guide the treating physician in the selection of combination regimens[25,63]. Actually there are two established assay principles for measuring resistance or sensitivity of HIV to specific antiretroviral drugs: the genotypic and the phenotypic resistance tests[64-67]. Genotype assays provide information about viral mutations that may result in changes in viral susceptibility to a particular drug or class of drugs. Phenotype assays directly determine the level of susceptibility of a patient-derived virus to specific drugs *in vitro*. Both assays are commercially available.

I.5.1. Genotyping

HIV genotypic assays detect specific mutations or nucleotide substitutions in the *gag-pol* region of the HIV-1 genome which encodes for the reverse transcriptase and protease

enzymes (the targets of current antiretroviral drugs). Specific gene sequences are compared with that of a reference (wild type) virus, and mutations associated with decreased susceptibility to specific antiretroviral drugs are identified.

HIV genotyping has historically been the more commonly used technology for drug resistance testing and generally utilises a 2-step procedure: a polymerase chain reaction (PCR) step to amplify a specific region of the HIV genome and a specific mutation detection methodology as DNA sequencing, gene chip arrays, and a line probe assay. Direct sequencing is the most commonly used. Viral genome sequencing can be performed using cycle-sequencing based kits, such as the TrueGene™ HIV-1 Genotyping kit, by Siemens Diagnostics, or the ViroSeq™ kit, by Abbott Diagnostics. Several reference laboratories also have developed their own assays, often referred to as "home-brew" assays (Virco™TYPE HIV-1, Virco; GenoSure (Plus), LabCorp; or GeneSeq, Monogram Biosciences).

The major limitation of the genotypic assays is interpretation of the results. Indeed, the value of Genotyping depends on an understanding of the relationships between identified mutations and specific drug susceptibility as well as viral fitness. Therefore, genotypic interpretation is an ideal application for a computerised expert system that accepts either a nucleotide sequence or a list of mutations and returns the predicted level of resistance. Combination of formidable amounts of data concerning the sequence of viral drug-target genes, in-vitro observations and cognate clinical outcome yielded powerful algorithms to model and therefore predict susceptibility/resistance to antiretroviral drugs. Difficulties arise in estimating the consequences of the interaction of multiple mutations on phenotype and the extent of cross-resistance among drugs in a class[64]. There are a number of on-line databases to assist in genotype interpretation either based on text (e.g. ANRS), score (e.g. Stanford DB), or fuzzy rules[68]. Other systems interpret genotypic mutation patterns with the aid of a large database of paired genotypic and phenotypic data, e.g. Geno2Pheno or virtual phenotype by Virco.

In the present study, the genotypic interpretations performed for diagnostic purposes were based on the Stanford DB algorithm (accessible from hivdb.stanford.edu), which allows inferring levels of resistance to PIs, NRTIs, and NNRTIs after submitting PR and RT sequences in text format. Stanford DB algorithm compares a submitted sequence with the consensus subtype B reference sequence, and the resulting differences are used as query parameters for interrogating the HIV Drug Resistance database. Based on the alignment, mutations as well as unusual results like frameshifts, insertions, and deletions are easily determined for each sequence. For each mutation associated with drug resistance a penalty score is assigned, and the total score for a drug is derived from the arithmetic sum. Mutations

that are known to cause hypersusceptibility to a drug have a negative score. The algorithm classifies the total score as susceptible, potential low-level resistance, low-level resistance, intermediate resistance, or high-level resistance. Ultimately, Stanford DB algorithm provides for the user an explanation to inferred drug resistance.

ANRS and Rega algorithms are two others genotypic interpretation systems that will be used in our study in comparison to the Stanford algorithm. They report their results using three susceptibility categories, therefore for comparison purposes, Stanford DB results need to be normalised (Table 1).

Algorithm	**S**	**I**	**R**
ANRS	Susceptible	Potentially resistant	Resistant
Stanford DB	Scores < 20	Scores between 20 and 44	Scores ≥ 45
Rega	Sensitive	Advise against when other options available	Resistant

Table 1. *Normalised resistance predictions to enable the comparison between ANRS, Stanford DB, and Rega algorithms.*

The French ANRS HIV-1 genotypic drug resistance interpretation algorithm is gaining recognition, because the rules of the current versions are almost exclusively based on data of correlation between drug resistance mutations and virological outcome from a large database of patients failing antiretroviral therapy.

Rega takes into consideration mutations for which phenotypic drug resistance or reduced therapy response have been reported. The latest versions also took advantage from information obtained through data mining in the large databases of the group.

Both are complex algorithms, taking into account known synergistic and antagonistic effects of combinations of mutations.

The Stanford DB program allows the identification and correction of common sequence problems. Besides stop codons or frameshifts in PR or RT genes, presence of atypical mutations, that is substitutions observed in less than 0.1% of published group M HIV-1 sequences, is considered as suspect. At least, highly ambiguous nucleotides are regarded as artefacts. Indeed mixtures of two nucleotides occur commonly but the presence of mixtures with three or more nucleotides at the same position occurs rarely in high quality sequences. When a mutation at one position is present as a mixture, the algorithm assigns the same penalty score regardless of whether the mutation is present in pure form or as a mixture.

Each sequence is compared to a list of reference sequences of subtypes A, B, C, D, F, G, H, J, K, CRF01_AE, and CRF02_AG, and the subtype of the closest reference sequence is assigned to the submitted sequence. This method, however, does not accurately characterize uncommon recombinants, and subtype B protease sequences are occasionally misclassified as subtype D, because these two subtypes are very similar and PR contains less phylogenetically informative positions than RT.

I.5.2. Phenotyping

Like genotypic tests, current phenotypic assays also use PCR to amplify the *gag-pol* region of HIV-1. In these assays, however, a recombinant virus is created by introducing the protease gene and/or the reverse transcriptase gene from a clinical HIV pool into a wild type laboratory clone from which the corresponding gene(s) have been deleted. This pool of recombinant viruses is used *in vitro* to infect a cell line, and virus replication is measured in the presence and absence of a range of concentrations of different antiretroviral drugs[69]. The viral replication is reflected by the activity of a reporter gene present either in the plasmid or in the cell line, depending of the assay. Therefore, two techniques can be differentiated: a non-infectious or single cycle assay, in the former case versus a replicative assay, in the latter.

Drug susceptibility results are reported as IC_{50} values, the drug concentration required to inhibit viral growth by 50%. Changes in drug susceptibility are measured quantitatively by comparing the IC_{50} value of the patient-derived virus to that of the reference (wild type control) virus[69]. The obtained value referred as Resistance factor (Rf) is compared to the so-called cut-off value. Determination of the cut-off is crucial for the interpretation of the results and three different are currently used:

- The *technical cut-off* is a measurement of the methodological variability of the assay.

- The *biological cut-off* involves the inter-individual variability of wild type virus isolates from ART-naïve HIV patients. If the IC_{50} is below the biological cut-off, virological success is very likely. However, an IC_{50} above the biological cut-off does not allow prediction of the virological response to a drug.

- In contrast, the *clinical cut-off* indicates up to which levels of IC_{50} virological success can still be expected.

Initially, replication-competent viruses were derived directly from patient PBMCs (Peripheral Blood Mononuclear Cells) by co-cultivation and used to analyse *in vitro* their resistance profiles. The technique was labour-intensive and time-consuming for routine use[70]. Therefore, the development of recombinant virus assays generated by homologous recombination between vectors and patient sequences raised much interest[71,72]. Currently, four main laboratories worldwide perform Phenotyping for diagnostic purposes: Eurofins Scientific, Inc. (ESI) in the USA (Phenoscript™), Virco in Belgium (Antivirogram™), Monogram Biosciences in the USA (PhenoSense™), and InPheno in Switzerland (PhenoTecT). For evident reasons the majority of details regarding the employed techniques are not freely accessible. Among these four systems, different strategies are used to obtain viral PR and RT genes. For example, the two genes are taken as a unique fragment in both Monogram Biosciences and Virco assays[72,73] whereas they are separately amplified in the Phenoscript™ and the PhenoTecT systems[74]. A further difference lies in the assessment of viral activity, which is measured through expression of β-galactosidase (induced by HIV Tat) in both InPheno and Eurofins assays, while Virco analyses the cytopathic effect of HIV using tetrazolium salts, and Monogram Biosciences determines HIV survival based on luciferase activity. In the latter assay, the luciferase gene is inserted in the envelope gene of a recombinant provirus, which is therefore not viable and needs a helper virus carrying the HIV *env* gene. Evidently that system only operates for the first round of replication, when both viruses are co-transfected, therefore the assay is not replicative. On the contrary, Virco and InPheno can determine HIV subpopulations and fitness, because their assays are replicative. This is very important, because viruses with Resistance Associated Mutations (RAMs) that exhibit different *in vitro* replication kinetics compared to wild type (WT) have been reported[75,76].

I.5.3. Limitations

The current genotypic and phenotypic assays require plasma with ≥ 500-1000 copies of HIV-1 RNA per millilitre. Thus, viral load testing remains initially the method of choice to monitor for drug failure.

Furthermore, the resistance assays were developed using the more common clade B viruses of Europe and North America. The genetically diverse HIV strains from different geographic regions may be less well amplified for Genotyping and Phenotyping. This is particularly an important concern in DNA sequencing for genotypic assays which could present primer bias and therefore not providing genotypic results.

Disadvantages of phenotypic testing include the lengthy procedure (8 to 21 days) and high expense of the assay. The cost of Genotyping ranges from 350 to 500 Euro, depending on the assay and laboratory used. It is approximately twice as much for Phenotyping. Secondly, it has been described that d-drugs are difficult to assess by Phenotyping. The main limitation of this *in vitro* assay in case of d-drugs is the intracellular phosphorylation. Indeed, these antiretroviral agents are prodrugs that need to be phosphorylated inside cells in order to be effective. A suboptimal phosphorylation would render their incorporation into a DNA chain not possible and the termination of the transcription would not occur. This could be a form of cellular resistance to the drug which will prevent any detection of samples susceptible to d-drugs. However, the cells used in the rPhenotyping assay appear to allow an optimal level of phosphorylation for those drugs since in PhenoBase® a high amount of samples diagnosed susceptible to d-drugs are found.

A common denominator of genotypic tests is that they cannot detect HIV subpopulations representing less than 15% of the total viral population. Then, clinically significant minor subpopulations of virus remain not detectable by actual genotypic methodologies whereas replicative Phenotyping could assess their presence. Similarly, complex patterns like mixtures of quasispecies will be individually interpreted by phenotypic assays, while genotypic tests will associate all mutations on a single virtual genome. And last, genotypic knowledge about interactions of multiple mutations need to be regularly updated as well as the implementation of new antiretroviral agent require previous phenotypic studies.

The latest clinical recommendations do not favour Genotyping over Phenotyping or vice versa. They support both methods because they provide complementary information[77]. Many studies compared the two methods, either alone or in combination, with so-called standard of care treatment (SOC; no resistance test performed): NARVAL, CERT, Vihres (genotyping vs. phenotyping vs. SOC), VIRADAPT, GART, Havana, Argenta (genotyping vs. SOC), Kaiser, VIRA3001, CCTG575 (phenotyping vs. SOC)[78-85]. As it is often the case those studies are controversial and often produced opposite results, due to different genotypic and phenotypic systems used in different trials. In general terms, no major gain was observed by resistance testing. This may be due to the choice of the HIV populations: those studies were done in early 2000 with few drug-experienced patients and the cause of diagnosis, e.g. a resistant test done on primary infected patients, could favourably select for wild type patients. The availability of new drugs, without resistance mutation development yet, could also influence the gain provided by a resistance test.

Aims of the study

The ability of HIV-1 to develop resistance to antiretroviral drugs has limited the overall efficacy of combination therapy to suppress viral replication. Consequently, drug resistance testing has become an integral tool for many HIV specialists, particularly when managing patients experiencing failure of anti-retroviral therapy. There are two principal methodologies for assessing resistance to antiviral agents: Genotyping and Phenotyping. The relative utility of the two types of resistance tests is a topic of a heated debate.

At the very beginning, the phenotypic technique, derived directly from patient PBMCs, was labour-intensive and time-consuming for routine use. Nevertheless, new formats are currently available, as rPhenotyping, which involve the construction of recombinant virus genomes containing sequences from patient viruses encoding protease and/or reverse transcriptase genes. One aim of this study was to validate the PhenoTecT system, a replicative Phenotyping, as diagnostic tool applicable for detecting and assigning HIV resistance in clinical specimens. Furthermore, this methodology could be performed for special cases with known limitations for standard Genotyping, e.g. a low sensitivity level to detect minority species or the quasispecies misinterpretation inherent to the nature of the test.

The questions to elucidate in this study were the following:

1. Is rPhenotyping as powerful as Genotyping to reveal HIV-1 resistance in routine use? (Chapter I)

2. Can rPhenotyping overtake the limitations of standard Genotyping?
    - How are the viral mixtures influencing the resistance tests? (Chapter II)
    - Are clinically critical HIV minorities revealed by replicative Phenotyping? (Chapter III)

To find an answer to these questions will allow to provide more precise information to HIV experts and to improve our understanding of therapeutic failure. The PhenoTecT system could help treating physicians to optimise new therapy regimens for patients failing antiretroviral drugs. This study will assess the complementarity of the two methodologies, Genotyping and Phenotyping[86,87].

II. Material & Methods

II.1. Materials

II.1.1. Chemicals

Chemical	Supplier
Diagnostics	
Cobas Amplicor HIV-1 Monitor™, v1.5	Roche Molecular Diagnostics
ViroSeq™ HIV-1 Genotyping	Abbott Molecular
Standard PCR	
Platinum® PCR SuperMix	Invitrogen
iProof™ High-Fidelity DNA Polymerase	Bio-Rad
dNTPs (dATP,dCTP,dGTP,dTTP), 10mM	Sigma
RT-PCR	
SuperScript™ III Reverse Transcriptase	Invitrogen
RNaseOUT™, Recombinant Ribonuclease Inhibitor	Invitrogen
dNTPs (dATP,dCTP,dGTP,dTTP), 10mM	Sigma
Sequencing PCR	
BigDye® Terminator v3.1 Cycle Sequencing	Applied Biosystems
BigDye® v1.1/3.1 Sequencing Buffer (5X)	Applied Biosystems
Real-time PCR	
QuantiTect SYBR Green PCR kit	Qiagen
Vector Preparation	
NcoI	New England Biolabs
XmaI	New England Biolabs
Digestion buffers, 10X	New England Biolabs
Bovine Serum Albumin (BSA), 10X	New England Biolabs
Alkaline Phosphatase, Calf Intestinal (CIP)	New England Biolabs
CH_3COONa 3M, pH 5.2	Sigma
Phenol:Chloroform:Isoamyl Alcohol (25:24:1),	Sigma
Ethanol (100%)	Fluka
Cloning	
NcoI	New England Biolabs
XmaI	New England Biolabs
Digestion buffers, 10X	New England Biolabs
Bovine Serum Albumin (BSA), 10X	New England Biolabs
Clonables 2X Ligation Premix (T4 DNA ligase)	Novagen

Gel Electrophoresis

SeaKem Agarose	Cambrex
TBE buffer, 10X	Amresco
Ethidium bromide solution (10mg/mL)	Sigma
Loading Dye	Invitrogen
1kb DNA ladder (1µg/µL)	Invitrogen
100bp DNA ladder (1µg/µL)	Invitrogen

DNA Isolation and Purification

NucleoSpin® Extract II	Macherey-Nagel
Microcon® Centrifugal Filter Units YM-100	Millipore
NucleoSpin® Plasmid	Macherey-Nagel
NucleoBond® Xtra Midi Plus	Macherey-Nagel
NucleoBond® PC 500	Macherey-Nagel
NucleoBond® Finalizer Plus	Macherey-Nagel

RNA Isolation and Purification

NucleoSpin® RNA Virus	Macherey-Nagel
DNaseI, RNase free (lyophilised)	Macherey-Nagel

Site-directed Mutagenesis

QuikChange® Site-Directed Mutagenesis	Stratagene

Bacterial Culture, Competent Cells Preparation

Bacto Agar (dehydrated)	Becton-Dickinson
Bacto Tryptone (dehydrated)	Becton-Dickinson
Bacto Yeast Extract (dehydrated)	Becton-Dickinson
NaCl	Fluka
Ampicillin (sodium salt)	Sigma
Glycerol (87%)	Fluka
$CaCl_2$ dihydrate	Fluka
One Shot® TOP10 cells	Invitrogen
HB101 cells	Promega

Cell Culture

DMEM High Glucose (4.5g/L) with Stable Glutamine	BioConcept
RPMI-1640, with 25mM HEPES (w/o L-Glutamine)	BioConcept
L-Glutamine, 200mM (100X), liquid	Gibco
Opti-MEM I Reduced Serum Medium (1X), liquid	Gibco
Lipofectamine 2000	Invitrogen
Fetal Bovine Serum (Heat Inactivated)	Gibco
Trypsine/EDTA (w/o Ca^{2+}/Mg^{2+})	Gibco
β-Mercaptoethanol, 50mM (1000X)	Gibco
Pen/Strep	BioConcept
D-PBS (1X), liquid (w/o Ca^{2+}/Mg^{2+})	Gibco
Trypan Blue Stain, 0.4%	Gibco

Virus Inactivation and Cells Fixation

Formaldehyde (36.5%)	Fluka
Glutaraldehyde (25%)	Fluka
PBS (10X)	produced at IMM

Antiretrovirals

Saquinavir pills	Hoffmann – La Roche
Zidovudine pills	GlaxoSmithKline
Didanosine pills	Bristol-Myers Squibb
Zalcitabine pills	Hoffmann – La Roche
Stavudine pills	Bristol-Myers Squibb
Abacavir pills	GlaxoSmithKline
Lamivudine pills	GlaxoSmithKline
Emtricitabine pills	Gilead Sciences
Efavirenz pills	Bristol-Myers Squibb
Nevirapine pills	Boehringer Ingelheim

ONPG-Staining

Buffer Z	produced at InPheno
Buffer H	produced at InPheno
o-Nitrophenyl-β-D-Galactopyranoside (ONPG)	Amresco
β-Mercaptoethanol (100%)	Fluka

General Chemicals

Dimethyl sulfoxide (DMSO)	Riedel-de Haën
Ethanol (100%)	Fluka
Isopropanol (100%)	Fluka
Sodium hypochlorite (10%)	Fluka
Dismozon pur	Bode Chemie
Deconex 53 PLUS	Borer Chemie

II.1.2. Cell Lines

Cell Line	Characteristics	Source
SX22-1	HeLa cell line, containing an integrated plasmid with an HIV-1 LTR driving bacterial LacZ gene; naturally expressed CXCR4 coreceptor and non-naturally expressed CD4 receptor	produced by Dr. Thomas Klimkait
SX-CCR5	HeLa reporter cell line, derivative of SX22-1, carrying CCR5 coreceptor	available through Dr. Thomas Klimkait
CEM-SS	T-Lymphoblastic cell line	provided by Dr. Thomas Klimkait

II.1.3. Primers

Usually we name primers according to their 5' end position preceded by an "F" or "D" for forward or direct and an "R" for reverse primers. This notation allows calculating directly the size of a fragment from the name of primers. For mutagenesis primers, it is common to write the codon to be mutated preceded by the wild type amino acid and followed by the mutant amino acid, and to notify by sense or antisense if it is a forward or reverse primer. "D67N sense" is a forward primer used to introduce in RT gene an Asparagine at position 67 instead of an Aspartic acid. For allele-specific PCR and real-time PCR the code name is different, writing the name of the gene to be quantified (PR, RT, IN), the codon and the amino acid (for some we added internal mutations which are figured out by their number at the end).

Primer	Sequences 5' → 3'
Standard PCR/Overlap PCR/Mutagenesis	
D-2213B	agc agg atc cga tag aca agg a
R-2598L	cca tcc cgg gct tta att tta ctg g
D-2582L	att aaa gcc cgg gat gga tgg ccc aa
R-3525L	tta agt ctt tcc atg ggt cat aat a
D67NT69D sense	gcc ata aag aaa aaa aac agt gat aaa tgg aga aaa tta g
D67NT69D antisense	cta att ttc tcc att tat cac tgt ttt ttt tct tta tgg c
D67N sense	cat aaa gaa aaa aaa cag tac taa atg gag
D67N antisense	ctc cat tta gta ctg ttt ttt ttc ttt atg
T69D sense	gaa aaa aga cag tga taa atg gag aaa att
T69D antisense	aat ttt ctc cat tta tca ctg tct ttt ttc
K70R sense	ac agt act aga tgg aga aaa
K70R antisense	ttt tct cca tct agt act gt
L74V sense	aaa tgg aga aaa gtg gta gat ttc aga gaa
L74V antisense	ttc tct gaa atc tac cac ttt tct cca ttt
V75M sense	aat gga gaa aat taa tgg att tca gag aac
V75M antisense	gtt ctc tga aat cca tta att ttc tcc att
Y181C sense	cag aca tag tca tct gtc aat aca tgg atg
Y181C antisense	cat cca tgt att gac aga tga cta tgt ctg
M184V sense	tca tct atc aat acg tgg atg att tgt atg
M184V antisense	cat aca aat cat cca cgt att gat aga tga
T215F sense	tgg gga ttt ttc aca cca gac
T215F antisense	gtc tgg tgt gaa aaa tcc cca
T215Y sense	tga ggt ggg gat ttt aca cac cag aca aaa
T215Y antisense	ttt tgt ctg gtg tgt aaa atc ccc acc tca
K219Q sense	aca cca gac cag aaa cat cag
K219Q antisense	ctg atg ttt ctg gtc tgg tgt
RT-PCR	
D-1818	aga aga aat gat gac agc atg tca ggg agt
D-1826	tga tga cag cat gtc agg gag tgg ggg gac c
R-3584	tgg ctc ttg ata aat ttg ata tgt cca ttg

Sequencing PCR

D-2150	agc caa cag ccc cac cag
D-2610	gtt aaa caa tgg cca ttg aca gaa gaa a
D-3007	gga aag gat cac cag caa tat tcc a
R-2605	ggg cca tcc att cct gg
R-3010	cca tcc ctg tgg aag cac att g
R-3541	ttc tgt att tct gct att aag tct ttt gat g

Allele-specific PCR

D-2213B	agc agg atc cga tag aca agg a
D-2232	gga act gta tcc ttt agc ttc cc
R-RT181Y3	aca tac taa tca tca agg tat tga t
R-RT181C3	aca tac taa tca tca agg tat tga c
R-RT184V	cag atc cta cat aca aat cat cca c
R-RT184M	cag atc cta cat aca aat cat cca t
R-RT184V3	cag atc gta cat atg aat cat cca c
R-RT184M3	cag atc gta cat atg aat cat cca t

Real-Time PCR

F-3005	atg gaa agg atc acc agc aa
R-RT181Y3	aca tac taa tca tca agg tat tga t
R-RT181C3	aca tac taa tca tca agg tat tga c
R-3115	aca tac taa tca tca agg tat tga
R-RT184V3	cag atc gta cat atg aat cat cca c
R-RT184M3	cag atc gta cat atg aat cat cca t
R-3123	cag atc gta cat atg aat cat cca
R-3146	tct atg ctg ccc tat ttc taa
R-3151	ttt gtt cta tgc tgc cct att tc

II.1.4. Plasmids

Plasmid	Characteristics	Source
pNL4-3	Provirus derived from NY5 (5') and LAV (3') HIV-1 isolates, cloned into pUC18 at PvuII site (size: 14,877bp)[88]	produced and provided by Dr. Malcolm Martin (NIH)
pNL-NF	pNL 4-3 with reduced flanking regions of cellular DNA	produced by Dr. Thomas Klimkait
pBX_ WT	pNL-NF with elimination of BamHI site at position 8465 and introduction of BamHI/2219 and XmaI/2591	produced by Dr. Thomas Klimkait
pXN_300	Cloning cassette for the reverse transcriptase gene (between XmaI/2591 and NcoI/3510)	produced by Dr. Thomas Klimkait
pXN_WT	Wild type version of the cloning cassette for reverse transcriptase gene	produced by Dr. Thomas Klimkait

pNL-NF is a truncated version of pNL4-3, the wild type reference HIV-1 B subtype virus. To generate cloning cassette for PR and RT gene different restriction sites were engineered in pNL-NF. To allow cloning of PR fragment a BamHI and an XmaI site had been introduced upstream and downstream from PR gene, respectively; in order to have unique sites a BamHI site at position 8465 was eliminated in parallel. This new plasmid was named pBX_WT. A NcoI site downstream from RT gene was introduced in pBX_WT to permit RT fragment cloning. pXN_300, which contains in the place of the RT gene a junk DNA fragment from a reporter vector pCAT3bas, was used as cloning cassette. pXN_WT was produced cloning wild type RT in pXN_300. When the resulting XN_300 and XN_WT plasmids were transfected, the former did not produce any virus, while the latter produced virus in the PhenoTecT assay. The fitness of pXN_WT provirus is similar to the NL-NF and NL4-3 plasmids, which means that the inserted restriction sites do not have a negative impact on the virus life cycle. Since all plasmids contain pUC18 elements, they can be amplified in E. coli cells and they confer ampicillin resistance upon transformation.

All selected mutants were constructed either on a NL4-3 background by mutagenesis or on pXN_WT by overlap-PCR.

II.2. Methods

II.2.1. Molecular Biology

This study was mainly done on patient samples which were prepared as described in the diagnostic part (*§ II.2.1.1.*). In parallel in order to reduce cost, we used research kits to prepare RNA extraction or cDNA from patient samples and samples not included in the diagnostic process. The latter methodologies will be detailed in next parts. The order of the methods presented is corresponding to the “diagnostic week”: from RNA to plasmid.

II.2.1.1. Diagnostic

Sample preparation and amplification.

Plasma samples from HIV patients were received and stored at -80°C. Virus was isolated from 1mL of EDTA-plasma by centrifugation at 50,000 × g for 80 min at 4°C. Pellets were dissolved in 600 µL of guanidinium isothiocyanate lysis buffer and RNA was extracted according to the Cobas Amplicor HIV-1 Monitor™, v1.5 protocol with two washing steps and final resuspension.

Reverse transcription was performed using the ViroSeq HIV-1 Genotyping kit. Reaction volume was 20µL. Reactions were set up as it follows: 8µL HIV RT mix, 1µL RNase Inhibitor, 0.4µL DTT at 100mM, 1µL of MuLV RT and 10µL RNA template (pre-denaturated 30'' at 65°C). The standard protocol was 1h at 42°C (reverse transcription), and 5' at 99°C (inactivation of RT enzyme). Reactions were set up at room temperature.
A PCR mastermix was prepared using 29.5µL of PCR mix, 1µL UNG and 0.5µL AmpliTaq Gold (reagents from the ViroSeq HIV-1 Genotyping kit). 30µL of mastermix were added to reaction tube from the RT step. Standard cycling protocol was (in total 40 cycles): 10' at 50°C (UNG activation), 12' at 93°C (AmpliTaq Gold activation), [20'' at 93°C (denaturation), 45'' at 64°C (annealing), 3' at 66°C (extension)], and 10' at 72°C (ends polishing).
After a purification step following Microcon YM-100 Centrifugal Filter protocol, PR and RT sequences were amplified separately by nested-PCR with the Platinum® PCR SuperMix (description *§ II.2.1.3.*). Reaction volume was 50µL. We used a forward primer containing either the restriction site BamHI (D-2213b) or XmaI (D-2582L) and a reverse primer containing either XmaI site (R-2598L) or NcoI site (R-3525L), respectively for PR and RT.

These PCR products were digested by the corresponding restriction enzyme after gel extraction and inserted into cloning cassette to provide a recombinant virus (description *§ II.2.1.6.*). These steps were similar for diagnostic and research, recombinant plasmids were transfected following the PhenoTecT system (description *§ II.2.3.3.*).

Genotypic analysis by sequencing.
Sequencing reaction was set up as it follows: 8µL of the Microcon-purified PCR product (approximately 20ng) and 12µL of each primer mix (from the ViroSeq HIV-1 Genotyping kit). Standard cycling protocol was (in total 40 cycles): 10'' at 96°C (denaturation), 5'' at 50°C (annealing), and 4' at 60°C (extension). Sequencing reactions were precipitated with Ethanol prior to analysis on the Applied Biosystems 3130 Genetic Analyzer (*§ II.2.1.9.*).
Genotyping and alignment to the HXB2-reference were performed with the ViroSeq system (Abbott Molecular). Resistance-associated mutations were identified using version 4.2.6 of the Stanford algorithm (http://hivdb.stanford.edu/). The sequence was sent to this software which returns a list of mutations and their impact on resistance for each drug. The Stanford analysis is based on penalties; each mutation has a score corresponding to each drug, based on clinical data. The addition of mutation scoring per drug gives a total which is interpreted by comparison to cut-off values and provides a profile of resistance for the different drugs. There are 5 levels: susceptible, potential low-level resistance, low-level resistance, intermediate resistance and high-level resistance.

II.2.1.2. Reverse Transcription Polymerase Chain Reaction

First-Strand cDNA synthesis Step: *SuperScript™ III Reverse Transcriptase.*

Some modifications were used in contrast to provided SuperScript™ III Reverse Transcriptase protocols. Positive and negative (RNA diluent) controls were performed in parallel.

Reactions were set up for a final volume of 25µL. 12µL of isolated RNA (protocol is suited for 10pg-5µg) were put in PCR tubes with 3µL of gene-specific reverse primer at 1pmol/µL (1µM). The tubes were incubated at 70°C for 5', and for 2'30'' on ice (annealing of the primer). To each tube 8µL of master mix were added, which contains 5µL buffer (5X), 1µL dNTPs (10mM), 1µL DTT (0.1M) and 1µL RNaseOUT™, followed by a 2' incubation at 55°C. Afterwards 2µL of RT enzyme were added and reactions were incubated for 1 hour at 55°C. Reaction inactivation was accomplished at 70°C for 15'. Tubes were incubated for 15' at 4°C prior the second-strand synthesis. It is not necessary to remove the RNA strand through RNaseH treatment.

II.2.1.3. Standard Polymerase Chain Reaction

Platinum® PCR SuperMix.

Reaction volume was 50µL. Primer stock solutions were stored at -20°C at a concentration of 100µM. The working solutions were diluted 1:10. For 50µL reaction, 2µL of each primer were used, corresponding to 20pmol (final concentration of 400nM). For PCR reactions of smaller volume (25µL), 10pmol were used. The amount of DNA template depends on the molecular weight, ranging from 50ng to 10ng. The remaining volume, usually 45µL, was completed with the provided SuperMix.

This mixture contains anti-Taq DNA polymerase antibody, Mg^{2+} (1.65mM), dNTPs (220µM), and recombinant Taq DNA polymerase. It is provided at a 1.1X concentration, which means that for a 50µL reaction, 5µL can be used for primers and DNA template addition. Due to the presence of anti-Taq antibodies, the polymerase is inactive at room temperature and the reactions have to be incubated for 2' at 95°C before starting amplification. A standard amplification cycle was set up as it follows (in total 30 cycles): [30'' at 95°C, 20'' at 52°C, 1' at 72°C], and 10' at 72°C. The extension time was calculated for RT gene size (around 900bp). Reactions were set up on ice to prevent unspecific primer annealing to the template.

iProof™ High-Fidelity DNA Polymerase.

Reaction volume was 50µL. Reactions were set up as it follows: 10µL 5X iProof HF buffer, 1µL 10mM dNTPs mix (final concentration of 200µM), 2µL of each primer at 10µM, 1µL DNA template (approximately 10-50ng), 0.3µL iProof DNA polymerase, and remaining volume was filled with sterile H_2O. The standard cycling protocol was (in total 30 cycles): 30'' at 98°C, [10'' at 98°C, 20'' at 52°C, 30'' at 72°C], and 10' at 72°C. Reactions were set up on ice to prevent unspecific primer annealing to the template.

After In-House RT step, the iProof™ High-Fidelity DNA Polymerase protocol was used for the PCR. 6µL out of 25µL (SuperScript™) of first-strand cDNA synthesis reactions were used. Reactions were set up as described above in a total volume of 25µL. After this reaction the DNA concentration was usually too low to observe a band on a gel, even when the whole 25µL were loaded. Therefore, a second amplification reaction was performed with 3µL out of 25µL of 1st PCR reaction. The same PCR protocol as for the first PCR may be used or internal primers: Nested-PCR. Reaction volume was 25µL. Microcon® Centrifugal Filters Unit YM-100 were used to clean the PCR product. A 1% agarose gel was run at 250V with 5µL of purified PCR product to check for the correct size of the DNA fragment.

II.2.1.4. Gel Extraction

PCR DNA was extracted from agarose gel according to the Macherey-Nagel protocol. The DNA concentration was not quantified at this step, since the measurement was not very reproducible, due to low yield. At low concentrations this measurement has a qualitative rather than a quantitative character.

II.2.1.5. Vector Preparation

For maximal cleavage efficiency the plasmid to be cut was incubated at the optimal temperature for enzyme activity (usually 37°C) for 2 hours. 40µg of plasmid were digested in order to have a good yield. Enzymes were put in slight excess to increase digestion efficiency. However, excessive enzyme or glycerol concentrations (>5% v/v) can trigger star activity. The standard digestion was as follows: 40µL plasmid DNA (1µg/µL), 3µL Restriction enzyme 1 (30 units), 3µL Restriction enzyme 2 (30 units), 10µL BSA 10X, 10µL Buffer 10X and 34µL sterile H_2O. Afterwards 10 units of CIP phosphatase were added to remove 5´ phosphates from the cut plasmid, which prevents vector re-circularisation. The reaction was incubated for 30' at 37°C. Protein contaminations were further removed from

DNA using phenol, which denatures proteins that fall between the organic phase and the aqueous, DNA containing phase. $CHCl_3$ was used to further denature proteins and to stabilise the aqueous/organic boundary. Isoamyl alcohol was used to reduce foaming. A high-quality solution containing phenol, chloroform, and isoamyl alcohol in ratio 25:24:1 is typically used. It is saturated with TE to avoid DNA partition in the phenolic phase, and it does not contain oxidation products of phenol (damaging DNA). One volume of phenol-chloroform-isoamyl alcohol was added to 200µL DNA solution and the whole mixture was strongly vortexed and centrifuged. Supernatant, corresponding to the aqueous phase, was transferred into a new tube. In presence of monovalent cations (Na^+) and ethanol, DNA forms a pellet. Thus, 20µl 3M CH_3COONa (pH 4.6) was added together with 600µL 100% ethanol. The reaction was incubated 30' at -80°C and then centrifuged for 10' at 20,800 × g and 4°C. After supernatant removal, the DNA pellet was dried and resuspended in 100µL warm elution buffer. The entire 100µL were loaded on a 0.6% agarose gel and run at 150V, which allows a good spatial resolution. The band was then extracted according to the usual gel extraction protocol of Macherey-Nagel.

II.2.1.6. Cloning and Transformation

All inserts were digested for 1 hour with the appropriate combination of restriction enzymes at optimal temperature for enzyme activities (usually 37°C). Reaction was set up as follows: 32µL insert (5-10ng/µL), 0.3µL of each restriction enzyme, 4µL buffer 10X, 4µL BSA 10X. The enzymes were inactivated at 80°C for 20'. The theoretical molar ratio between insert and vector should be approximately 3:1. 50-70ng of vector were ligated to 3µL insert (5-10ng/µL). An equivalent volume of Clonables 2X Ligation Premix was added and the reaction was incubated at 16°C for 20'. A negative control containing the vector alone was always performed to estimate the background of vector self-ligation.

Afterwards, 100µL of fresh thawed competent cells (either HB101 or One Shot® TOP10) were added to the ligation mix and the tubes were incubated on ice for 30' (bacteria were resuspended only few times to avoid mechanical lysis). Bacteria were heat-shocked at 42°C for 1 min and put back on ice for 10'. Depending on the aim of the experiment, bacteria were either cultured in liquid LB medium or plated on LB agar. On one side, if a mixture of different inserts was cloned and had to be preserved, bacteria were incubated in 700µL LB medium w/o ampicillin for 1 hour (recovery phase) and then transferred into 2mL liquid LB medium with ampicillin at 150µg/mL and incubated overnight at 37°C on a shaking platform. On the other side, if single clones had to be isolated after heat-shock, bacteria were plated on LB agar plates containing ampicillin and incubated overnight at 37°C. The next day colonies were picked and grown in liquid culture.

II.2.1.7. Plasmid DNA Purification

DNA plasmid extraction was performed according to Macherey-Nagel protocol. For a miniprep, the starting amount of LB culture was 3 mL, for a midiprep 100mL, and for a maxiprep 250mL. The DNA content was quantified by UV spectrometry at 260nm using NanoDrop® ND-1000.

II.2.1.8. Restriction Digestion

1µL of plasmid DNA (300ng/µL) was digested with 0.3µL of each enzyme, 1µL of buffer 10X, 1µL of BSA 10X, and 6.4µL H_2O (total volume of 10µL). The reaction was incubated at 37°C for 1 hour. Subsequently 2µL loading dye 6X were added and the solution was loaded in a 0.8% agarose gel with a Size Ladder 1kb and/or 100bp in parallel.

II.2.1.9. DNA Sequencing

Sequencing for research was performed in-house using an Applied Biosystems 3130 Genetic Analyzer and the corresponding sequencing kit. Sequencing primers were ordered as HPLC purified. DNA denaturation before starting sequencing may help for long templates with complex sequences. Table 2 summarises the "rule of thumb" to prepare a dye terminator sequencing reaction. Usually plasmid DNA (about 13kb) or PCR fragments (about 1500bp) were used as template. Since the RT gene is relatively small, 4µL of Big Dye were sufficient.

Template size	**dsDNA (ng)**	**Primer (pmole)**	**Big Dye (µL)**	**Buffer 5x (µL)**	**H_2O (µL)**
100bp	10	3	4	2	to 20
1kb	80	3	4	2	to 20
10kb	800	3	6	2	to 20

Table 2. *Typical protocol for sequencing PCR.*

The PCR cycling was set up as follows (in total 40 cycles): 20'' at 96°C, 20'' at 50°C, and 4' at 60°C. For subsequent purification ethanol precipitation was the best solution for our purposes. Cold ethanol and temperature controlled lab conditions may improve results. Moreover, the use of glycogen may enhance DNA precipitation. For one sample, a fresh "stop-solution" containing 2µL 3M CH_3COONa (pH = 4.6) and 2µL 100mM EDTA was prepared. EDTA is not strictly required but can prevent enzyme mediated DNA degradation. After vortexing, 50µL of 100% ethanol were added and the solution was well mixed.

PCR products were distributed in a 96well plate (normally sequencing reactions are more than 10; hence it is more comfortable to work with a 96well format). 54µL (52µL if EDTA was not used) of "stop-solution" were added to the PCR products. The plate was covered with a plastic film and well mixed. Afterwards, it was centrifuged for 30' at 4,000 × g. Supernatant was removed by turning the plate and placing it on an absorbent paper. 150µl of ethanol 75% were added to wash the pellet; plate was centrifuged as before for 5'. After removal of the supernatant the plate was put in vacuum for 10'. Pellet was resuspended with 30µL of either H_2O or formamide (as a denaturant); the mixture was let 10' at room temperature to ensure that DNA was well solved. The polymer itself used in the sequencing machine contains ~7M urea as denaturing agent and was run at ~70°C to eliminate all hydrogen bonding associations in ssDNA.

II.2.1.10. Site-directed Mutagenesis

This method was used to introduce individual point mutations into a plasmid and was performed according to the Stratagene protocol. Briefly, the basic procedure utilises a supercoiled double-stranded DNA (dsDNA) plasmid with a target site for mutation and two synthetic oligonucleotide primers containing the desired mutation. The oligonucleotide primers, each complementary to opposite strands of the plasmid, were extended during temperature cycling by *PfuTurbo* DNA polymerase. Incorporation of the oligonucleotide primers generated a mutated plasmid containing staggered nicks. Following temperature cycling, the product was treated with *DpnI*. The *DpnI* endonuclease (target sequence: 5´-Gm6ATC-3´) is specific for methylated and hemimethylated DNA and is used to digest the parental DNA template and to select for mutation-containing synthesised DNA. The nicked plasmid DNA containing the desired mutations was then transformed into XL1-Blue supercompetent cells

II.2.1.11. Overlap-PCR

This reaction, also called fusion PCR, was used to introduce site-specific mutations in a target gene. It was normally set up in two different PCR steps. The first step was performed to produce two small fragments (Reaction 1 and 2 in Figure 7A) which were then joined by another PCR step (Reaction 3 Figure 7B). For the first reaction, a forward wild type primer annealing at the leftmost part (Primer 2) was used together with a reverse mutant primer (Primer 1a). For the second reaction a reverse wild type primer (Primer 3), this time annealing at the rightmost part, was used with a primer that is complementary to the mutant

one used in the first reaction (Primer 1b). Thus four different primers were used and the mutant primers were annealing in the middle relative to the two external wild type primers.

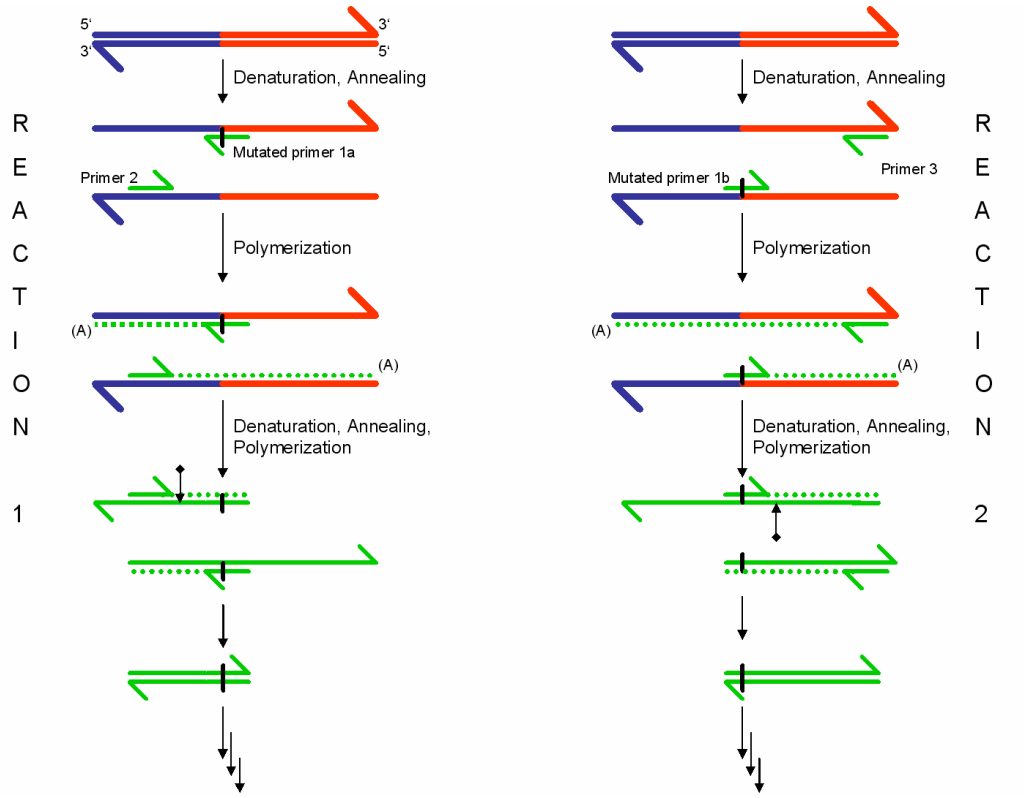


Figure 7A

Then a second PCR step was performed using the two small fragments produced during the first step and the two external primers (Figure 7B). In this way the two fragments were joined by polymerisation and further amplified.

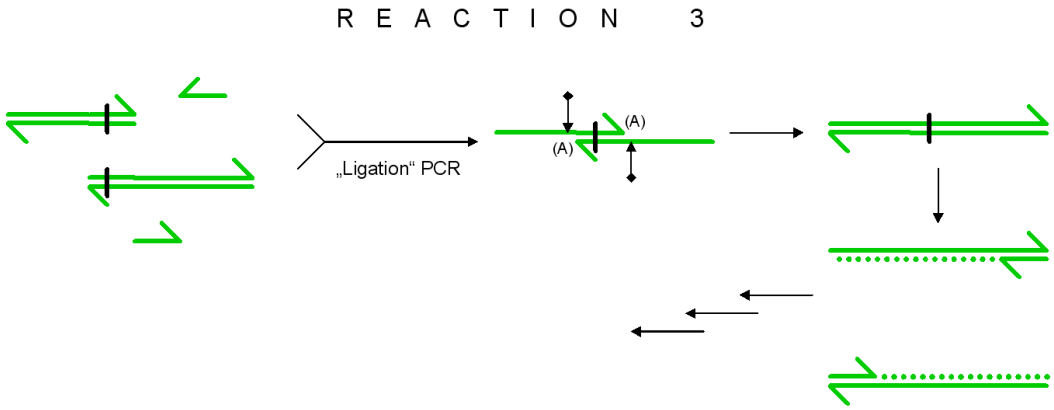


Figure 7B

Figures 7A, B. *Basic scheme of overlap-PCR (refer to text for more details).*

The first step of PCR was assembled according to the standard iProof™ High-Fidelity DNA Polymerase protocol, since the enzyme lacks terminal nucleotide transferase activity. This was absolutely necessary, otherwise the two fragments would end up with an extra nucleotide that was not complementary to the template sequence, usually an adenine ("A" in brackets in 7A, B). In this case, the second PCR amplification would fail, because there would be a mismatch exactly at the 5' end, where the enzyme should start polymerisation, unless the two fragments were produced to exactly anneal before a thymine (arrow in Figure 7A, B). To avoid this complication, it was sufficient to use for the first PCR step those enzymes which are known to have no terminal nucleotide transferase activity. For the second PCR step, this expedient was no longer necessary and enzymes with terminal nucleotide transferase activity were tolerated as well (e.g. Platinum® PCR SuperMix).

The protocol for reaction 3 was as follows: 3-4µL of fragment 1, 3-4µL of fragment 2, 2µL of each primer, 38-40µL master mix. The actual amount of the two fragments should ideally be equimolar. Yet, since their concentrations were relatively low, it was tricky to assess the mole amount. Under non-equimolar concentrations it is possible to result in more than one band after the PCR reaction (one corresponding to the coupled and the others to the uncoupled fragments) but using good resolution agarose gel electrophoresis allows to isolate the desired product.

II.2.1.12. Allele-specific PCR

This method was used to determine qualitatively the presence of resistant minorities with a mutated primer specific of the region of interest. The plasmid DNA was amplified simultaneously in two parallel reactions, by using a common upstream primer wild type and either one of two downstream primers with 3' ends matching either the wild type or the mutant sequence. In addition the mutant primer was required to also contain about three internal mutations[89]. The Platinum® PCR SuperMix was used for this purpose. Reactions were set up as follows (in total 25µL): 0.5µL plasmid DNA at 2ng/µL, 1µL forward primer (10µM), 1µL reverse primer (10µM) and 22.5µL Platinum SuperMix. The amplification was performed for 27 cycles at 95°C for 15 s, 62°C for 15 s, and 72°C for 1 min 30 s after an initial denaturation of 2 min. The crucial step was to find the optimal annealing temperature to be able to discriminate authentic from inauthentic annealing; e.g. the primer containing mutated sequence should not allow any amplification of the wild type template.
Specific amplification of the DNA fragment was verified by subsequent gel electrophoresis. For each specific PCR a negative and a positive control of specificity for amplification had to be included: wild type pNL4-3 as well as single mutated plasmid.

II.2.1.13. Mutation-specific real-time PCR assay

The amplification was performed with a QuantiTect SYBR Green PCR Kit for 50 cycles at 95°C, 10 s; 57°C, 15 s; and 72°C, 20 s after an initial 2 min step at 50°C and denaturation for 10 min at 95°C. Five microliters of the cDNA products, generated with the Viroseq HIV-1 Genotyping kit, were added to the Master mix containing: 2X QuantiTect SYBR Green PCR Master Mix , RNase-Free Water, and 7.5 pmole of each primer designed to quantify the total amount of virus as well as the percentage of specific mutant. The forward primer was designed to bind to a conserved region of *pol*, and reverse primers were designed either to quantify equally wild type and mutant template (the latter primer ends 5' to the mutant position), either to match specifically the wild type or the mutant template. Fluorescent quantification of amplicons was mediated using the Rotor-Gene™ 3000 system (Corbett Life Science). The cycle number, at which the fluorescence passed a fixed threshold line, is defined as the threshold cycle number (Ct). Copy numbers were calculated for wild type and mutant variants by interpolation of the experimentally determined threshold cycle onto standard regression curves. To ensure an accurate quantification for both, a separate standard curve was generated for each template. The distance between the Ct observed with authentic and inauthentic priming was called ΔCt and reflected the discriminatory ability.
Precise dilutions of cDNA were prepared in independent repeat experiments, and for each specimen two replicate reactions were performed for internal standard. Experiments were performed on pre-prepared mixes of wild type and mutant variants virus preparation to validate the method prior to the use on patient-derived samples. As previously described, the relative amount of mutated sequences was calculated as the ratio of the quantity of mutated sequences in the sample and the quantity of total sequences in the sample[90].

II.2.2. Microbiology

II.2.2.1. Bacterial Culture

LB medium with ampicillin
5g NaCl, 10g Bacto Tryptone, and 7g Bacto Yeast Extract were dissolved in 1L milliQ H_2O and autoclaved. When the solution was at room temperature, 1mL ampicillin (150mg/mL) was added: final antibiotic concentration was 150µg/mL.

LB agar plates with ampicillin
2.5g NaCl, 5g Bacto Tryptone, 3.5g Bacto Yeast Extract, and 6g Bacto Agar were dissolved in 0.5L milliQ H_2O and autoclaved. After cooling down the solution, 0.5mL ampicillin

(150mg/mL) was added: final antibiotic concentration was 150µg/mL. Approximately 20mL medium were poured in each Petri dish.

II.2.2.2. Preparation of Competent Cells

Both HB101 and One Shot® TOP10 competent cells are Amp^S and contain a recA mutation, which prevents undesirable recombination events. In addition, One Shot® TOP10 bacteria have an endA mutation that prevents carry-over of nucleases. Both cells as purchased have been made chemically competent.
Cells from this commercial source were plated on LB agar w/o ampicillin and incubated over night at 37°C. The next day one single colony was inoculated in 5mL LB medium and incubated overnight on a shaker. The over night culture was transferred into 1L LB medium; growth was followed until an OD_{260} of 0.2-0.5.
Afterwards cells were transferred in 50mL tubes and placed immediately on ice for 10'. Centrifugation was done at 4°C for 20' (1,600 × g). Supernatant was removed and cells were resuspended in 25mL ice cold 100mM $CaCl_2$ and centrifuged under the same conditions as before. After resuspending cells in 10mL ice cold 100mM $CaCl_2$, they were left on ice for 30' and then centrifuged as before. Following supernatant removal, cells were well resuspended in 50mL ice cold 100mM $CaCl_2$ with 10% glycerol and then aliquoted into PCR tubes, which were stored at -80°C. Bacteria were plated on LB agar with and w/o ampicillin to check for contaminations and concentration; moreover a transformation test with 10ng pNL-NF was done to check their transformation efficiency.

II.2.3. Cell Culture

II.2.3.1. Cell preparation

SX22-1 and SX-CCR5 are adherent cells and are grown in DMEM High Glucose (4.5g/L). CEM-SS cells are non-adherent human T-lymphocytes, which need RPMI for optimal growth. All cells were split three times a week. According to their growth kinetics SX22-1 cells were split 1:7 and SX-CCR5 cells 1:5. CEM-SS cells were diluted in an approximate way. Before splitting, cells were examined under an inverted microscope to judge their degree of confluence. For passaging adherent cells, medium was removed and cells were washed with PBS w/o Ca^{2+} and Mg^{2+}, and trypsinised. After removing the trypsin EDTA, the cells were incubated at 37°C for 5' and then resuspended in DMEM. A part of it was put in a new flask with fresh medium. Every two months a fresh vial of cells was thawed and the old

ones were discarded. Since cells were used for a longer period of time, separated media were used for each cell line to avoid cross-contaminations.
All three cell lines were stored in 1mL aliquots at -196°C (liquid nitrogen). They were thawed at 37°C and quickly transferred in culture medium, to dilute the toxic high DMSO concentration in the freezing medium to below 0.5%. Freezing solution always needs to be sterile filtrated.

II.2.3.2. Preparation of Antiretrovirals

Drug stocks were typically prepared from clinical formulation and were diluted in two formats, either as 6 or as 10 concentrations depending on the experiment. Drugs were diluted according to their solubility in polar solvents, H2O or DMSO, at 0.25% final concentration. The stock solutions were prepared at 10 mM. Moreover, each drug had its own range of concentrations depending on the IC_{50} value, in order to accurately extrapolate the inhibition curve.

II.2.3.3. Cell Transfection: PhenoTecT format

Since the PhenoTecT system has not yet been patent protected only limited information can be shared. For additional information contact Dr. Klimkait.

II.2.3.4. Virus Inactivation and Cell Fixation

Cells were fixed in 1X PBS containing 2% formaldehyde and 0.2% glutaraldehyde. Aldehydes are additive fixatives, which link membranes: the aldehyde group can combine with nitrogen atoms of proteins forming a methylene bridge. The combination of formaldehyde with glutaraldehyde takes advantage of the rapid penetration of CH_2O and the thorough cross-linking by glutaraldehyde. Furthermore, aldehydes quickly inactivate viruses. The cells, now fixed on the bottom of the wells, were washed with 1X PBS and dried.

II.2.3.5. ONPG Assay

Since the PhenoTecT system has not yet been patent protected only limited information can be shared. For additional information contact Dr. Klimkait.

The reporter cells carry an HIV-1 LTR-LacZ construct, which drives β-galactosidase expression when activated by HIV-1 Tat. In the presence of β-galactosidase, o-nitrophenyl-β-D-galactopyranoside (ONPG) is converted to galactose and o-nitrophenol (ONP). Reporter cells contain no enzymes capable of degrading ONP further. ONP is colourless like ONPG at neutral or acid pH, but in an alkaline solution it turns yellow. The amount of yellow colour is measured in a spectrophotometer at 405 nm reflecting the amount of ONP formed in a given time. Since ONP is a product of β-galactosidase activity, the spectrophotometric measurements may be used as an assay for the enzyme, which is in turn related to the amount of alive viruses (Figure 8).

Prior to measuring light absorbance, a mixture with buffer H, buffer Z, ONPG, and β-mercaptoethanol was prepared. Before adding 80µL of this solution to each well, it was gently swirled. 96well plates were then incubated for 1 hour at 37°C. The highest concentration of the positive control, an active drug against wild type gene from the plasmid cassette, was used in order to subtract the absorbance measured in this well to the one of the remaining wells. In this way background optical density can be subtracted.

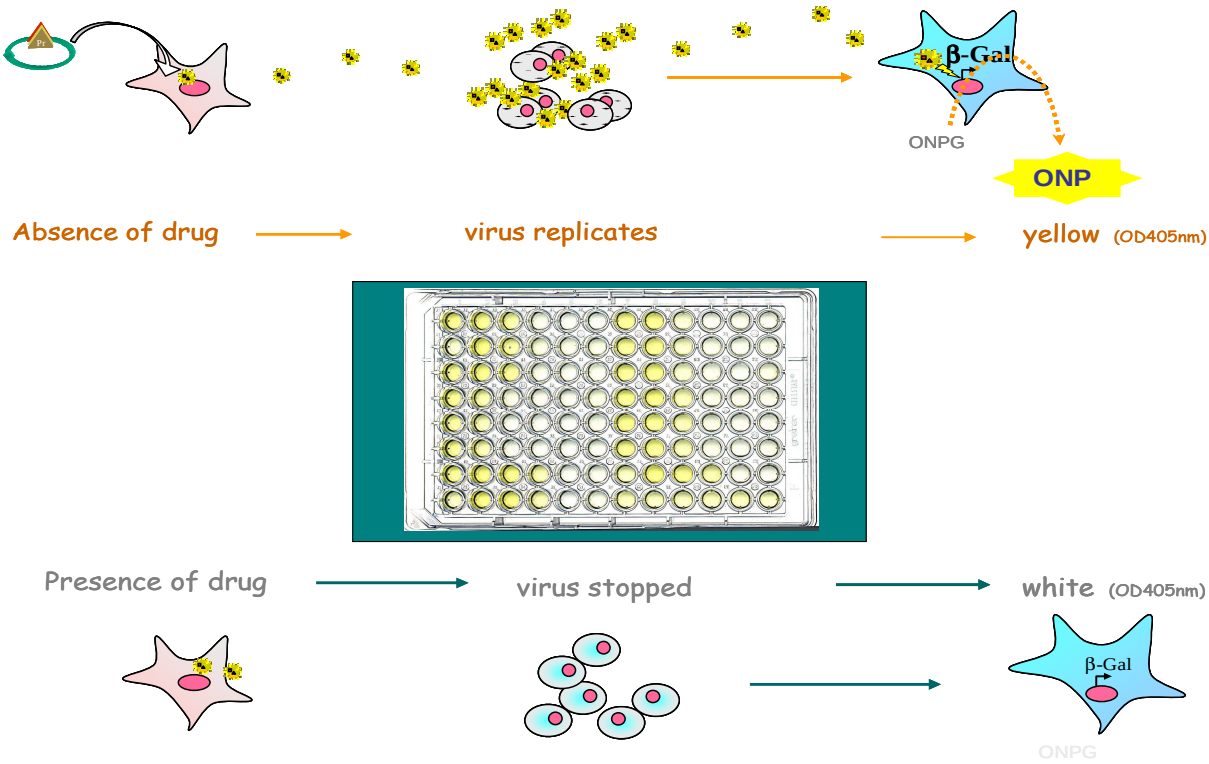


Figure 8. *Principle behind PhenoTecT assay.*

II.2.3.6. RNA Isolation and Purification

To extract HIV RNA from virions the standard Macherey-Nagel protocol was followed. Since viral supernatant came from wells containing plasmid DNA that was used to transfect SX22-1 cells, it was necessary to remove it from the RNA of the viruses. Thus, a new step was implemented in the standard protocol: after binding the viral RNA to the column, an on-column digestion was performed with DNase. DNaseI was prepared mixing 540µL RNase-free H_2O to one vial containing the lyophilised enzyme; aliquots were stored at -20°C. 10µL of the reconstituted DNaseI were added to 90µL of reaction buffer, the solution was mixed by flicking the tube and 95µL of it were put on each silica membrane. The column was then incubated at room temperature for 15'. Afterwards the standard protocol for viral RNA isolation was followed with three washing steps and final elution. The RNA content could only be qualitatively assessed by UV spectrometry at 260nm, since carrier tRNA and not glycogen was used to precipitate RNA in ethanol, leaving a background signal. On-column digestion was chosen because no re-purification of RNA with diethylpyrocarbonate (DEPC)-treated CH_3COONa would be needed after elution.

III. Chapter I – Comparative study of the results from different HIV resistance assays

III.1. Evaluation of replicative Phenotyping versus genotypic analysis based on PhenoBase®

Many laboratories in the world provide Genotyping expertise to assess resistance profiles for HIV patients. In Switzerland, there are four centres dedicated for this service: Geneva, Lausanne, Zurich and Basel. All data are collected in a common Base of the Swiss HIV cohort, to create a data-logistic network for research in the field of HIV. In Basel, new patient samples are received every week to determine their genotypic profile with the Stanford database. In parallel, a second test is performed for each patient on the 18 currently available drugs using a replicative Phenotyping (rPhenotyping system). Both resistance tests have been detailed in the introduction. The results of HIV-patients diagnosed for drug-resistance by both methodologies (>1'000 pairs of genotypic-phenotypic determinations) are computed. This helped to build the matched database 'PhenoBase®'. All information concerning each patient (personal data, physician, treatment, VL, CD4 count) are stored and are accessible in PhenoBase®, which thereby lends itself as basis to compare Genotyping and rPhenotyping.

Genotypic predictions of the Stanford DB algorithm are subdivided in 5 levels of resistance: susceptible (score 0-9), potential low-level resistance (score 10-14), low-level resistance (score 15-29), intermediate resistance (score 30-59) and high-level resistance (score ≥ 60); whereas rPhenotyping creates three categories: susceptible (S), intermediate (I) and resistant (R). To compare both diagnoses, I decided arbitrarily to also reduce Genotyping to only three classes, depending on the total scoring for all mutations:

- "Susceptible" if scoring is less than 20;
- "Intermediate" between 20 and 44;
- "Resistant" if scoring is superior or equal to 45.

In the whole study genotypic results will be presented as either Susceptible, Intermediate or Resistant, respectively footnoted Geno-S, Geno-I and Geno-R. In the same manner, phenotypic results will be introduced as Pheno-S, Pheno-I and Pheno-R.

To interrogate PhenoBase® on data from patients analysed between years 2002 and 2007 we re-interpreted the sequences with the last updated version of Stanford (version 4.2.6). In parallel, the phenotypic data were evaluated for the level of sensitivity with the cut-

offs re-calculated in December 2006. We observed an overall good agreement between Genotyping and Phenotyping especially concerning the RTIs (N= 1260 except for FTC, new drug, N= 465 and ddC, N= 690). The range of concordance for ten clinically used Reverse Transcriptase Inhibitors was from 72.3% (ABC) to 92.2% (NVP) (Figure 9). The NRTI ddC was out of this range with only 53.2% of concordance.

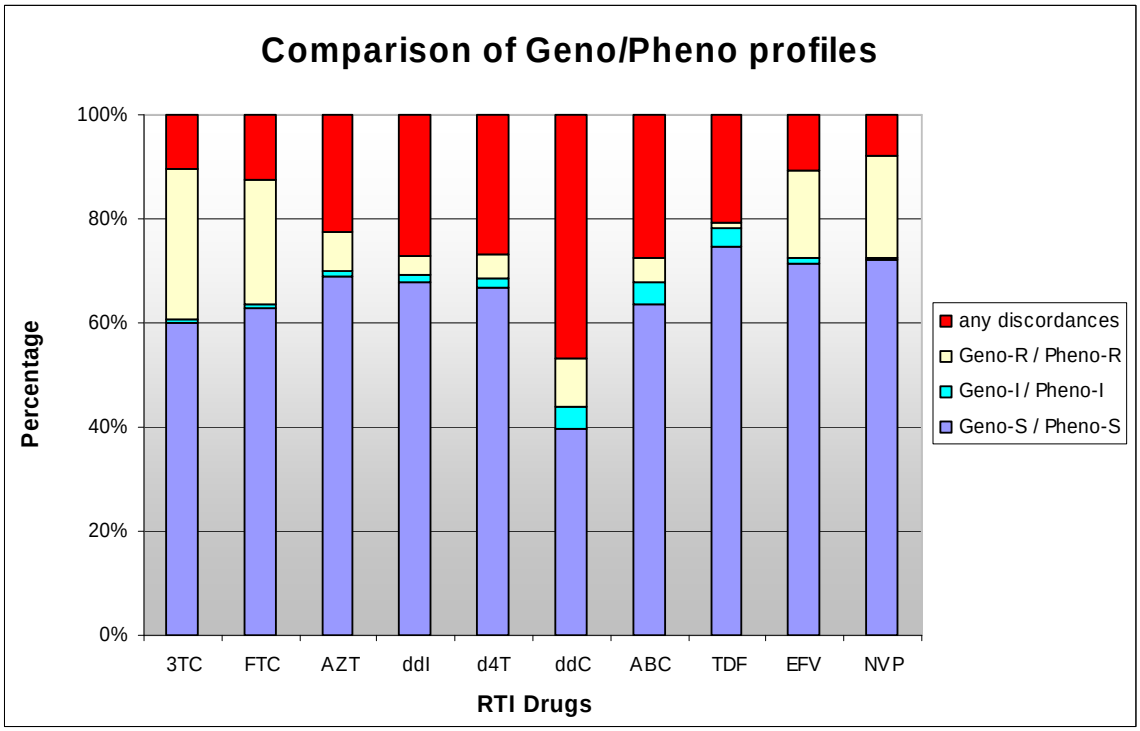


Figure 9. *Percentage of concordance and discordance between Stanford Genotyping and rPhenotyping for each NRTI and NNRTI clinically used.*

Four drugs presented a small percentage of discordance (around 10%): 3TC, FTC, NVP and EFV; while the second group showed less than 80% of concordance: d-drugs, AZT, ABC and TDF. Moreover, we observed a higher amount of Geno-R/Pheno-R in the former case. The resistance to the first group of RTIs (3TC, FTC, NVP and EFV) is mainly due to one strong mutation in the reverse transcriptase which prevents the incorporation of the drug, e.g. M184V for 3TC and FTC; K103N, V106M, Y188L and G190X for EFV and NVP. In other cases, the addition of several "weaker" mutations in the same gene leads to resistance, e.g. d-drugs, AZT, ABC and TDF. In order to better discriminate the major discordances Sensitive and Resistant, the histogram in Figure 10 was used to display the percentage of these discordances in a discrepant population.

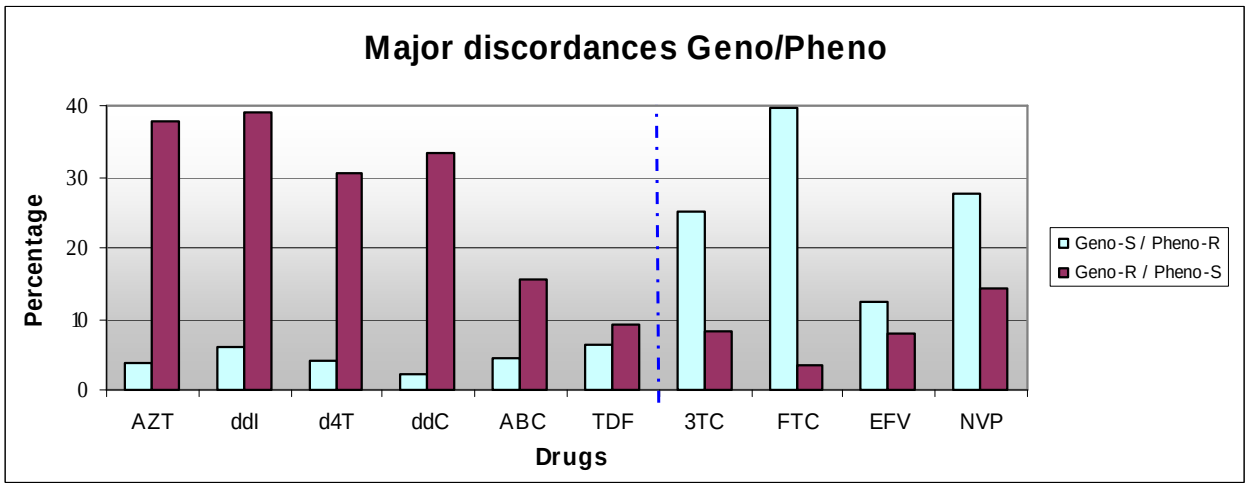


Figure 10. *Percentage of major discordance in the population presenting any discordance between Stanford Genotyping and rPhenotyping for each RTI.*

The above histogram shows the percentage of major discordance within a population of patients having different results for genotypic and phenotypic analysis on ten RTIs. Two groups of drugs are easily distinguishable separated by a blue dashed line:

- In the left part of the histogram, a higher amount of discordance is observed for the samples presenting a genotypic resistance with, at the same time, susceptibility by Phenotyping. It is particularly true for AZT and d-drugs: between 30% (d4T) and 40% (ddI) of the discordant population correspond to a prediction Geno-R/Pheno-S whereas the percentage of Geno-S/Pheno-R represents less than 6%.
- This distribution is completely inverted for the drugs plotted in the right part of the histogram which are the RTIs presenting the best correlation between Genotyping and Phenotyping. The most frequent category observed for these four ARVs is Geno-S/Pheno-R: between 12.5% (EFV) and 40% (FTC). The samples, for which Genotyping interprets resistance while in contrast Phenotyping predicts sensitivity, correspond to less than 10% of the discordant population except for NVP (14.3%).

III.2. Evolution of a Genotyping Algorithm over time

For new drugs like FTC, Stanford DB algorithm has even after many years not reached a "steady state" yet. FTC was added in Stanford DB v.3.9 on March 25, 2004 but seven months after, on October 27, 2004 many penalty scores were updated: E44A/D 8→5, T69A/I/S 5→2, T69insertion 30→20, V118I 8→5. In the version 4.1 (March 17, 2005) four further amino acid changes were included (M41L, L210W and T215Y/F), even though with rather minor scores (0→4). In addition the mutation K65N was added with a score of 15 on May 24, 2006. A similar situation is also observed for other better characterized RTIs and PIs.

The scores of EFV and NVP (NNRTIs) are more stable and in general very high (between 30 and 60). Nevertheless, several additional scores were updated: K101P (EFV and NVP) 10→20→30 on March 25, 2004 and January 20, 2006 respectively; K103T/S 30→50 for EFV and 30→60 for NVP on July 10, 2003; Y181V 10→20 for EFV and 30→60 for NVP on October 27, 2004 and G190V/T 50→60 for EFV on January 20, 2004 and for NVP on February 24, 2006. More remarkable changes were improved on amino acid V106M with a score of 15→50→60 for EFV on September 16, 2003 and April 05, 2005 respectively. It means that a given patient with just this substitution was susceptible to EFV in 2002 but intermediate resistant in 2003 and even resistant in 2005.

There is another kind of mutation which had a highly increased penalty score: the deletion at positions 67 and 69. This change had an impact on four drugs: ABC, AZT, d4T and ddI. In Stanford DB v.4.0 on October 27, 2004 the score of the deletion at position 67 was updated: 10→25 (ABC), 18→30 (AZT), 12→30 (d4T) and 5→25 (ddI). The deletion at position 69 had a stronger penalty score in January 20, 2006: 15→25 (ABC and ddI) and 15→30 (AZT and d4T). The new scoring for these deletions leads to an intermediate resistance level to each of the four mentioned RTIs while it was susceptible prior to the latest update.

Other remarkable examples were the addition of new mutations. In July 2003 two mutations were implemented: K103Q with a score of 5 for NNRTIs and K65R with a score of 15 for d4T; in March 2004 there were E44A/D and V118I with a score of 2 for ABC, AZT, d4T, ddI and TDF. In October 2004 the mutations T215S/C/E/D and K65R were added with a score of 10 for ddI and -5 for AZT respectively; in March 2005 M41L, L210W and T215F/Y gave a score of 4 for 3TC and FTC. The last addition seen until now for RTIs is the mutation K70E with a score of 8 for TDF in June 2005. Similar evolutions for PI scorings were observed (data not shown), which means that the variation of genotyping scores is a common limitation for all ARVs.

Since May 24, 2006 no other change has appeared for RTIs. Focusing on the major updates, the following changes have been made in Stanford DB algorithm from the beginning of our study (April 2003) until the end (December 2006):

Version, Date	Mutation(s)	Drugs	Old Score	New Score
Version 3.5, 2003-07-10	K65R	d4T	0	15
	K103Q	EFV-NVP	0	5
	K103T/S	EFV	30	50
		NVP	30	60
Version 3.7, 2003-09-16	V75A/I	d4T	30	20
	V106M	EFV	15	50
Version 3.9, 2004-03-25	E44A/D	ABC, AZT, TDF, ddI, d4T	0	2
	V118I	ABC, AZT, TDF, ddI, d4T	0	2
	L74I	AZT	-10	-8
	M184V	ddI	20	12
	T215F/Y	ddI	10	12
	K101P	EFV-NVP	10	20
Version 4.0, 2004-10-27	D67-	ABC	10	25
		AZT	18	30
		ddI	5	25
		d4T	12	30
	L74I	ABC	25	12
		AZT	-8	0
		ddI	50	25
	K65R	AZT	0	-5
	T69N	ddI	20	10
	T69G	ddI	5	20
	M184V	ddI	12	10
	T215F	ddI	12	15
	T215Y	ddI	12	20
	T215S/C/E/D	ddI	0	10
	T215C/E/D	d4T	5	15
	T215S	d4T	10	15
	V75A/I	d4T	20	10
	V75T	d4T	50	30
	K103Q	EFV-NVP	5	10
	Y181V	EFV	10	20
		NVP	30	60
Version 4.1, 2005-03-17	L74I	ABC	12	20
	M41L	3TC-FTC	0	4
	L210W	3TC-FTC	0	4
	T215F/Y	3TC-FTC	0	4
Version 4.1.1, 2005-04-05	V106M	EFV	50	60
Version 4.1.4, 2005-06-20	K70E	TDF	0	8
Version 4.1.7, 2006-01-20	T69-	ABC, ddI	15	25
		AZT, d4T	15	30
	L74I	ddI	25	30
	M184V	ddI	10	5
	T215F	ddI	15	20
	T215S/C/E/D	d4T	15	20
	K101P	EFV-NVP	20	30
	G190V	EFV-NVP	50	60
Version 4.1.9, 2006-02-24	Q151L	ABC, AZT, ddI, d4T	40	20
	V75A	d4T	10	15
	G190T	EFV-NVP	50	60
	V179F	NVP	5	25

Table 3. *Updates of the Stanford DB algorithm over the last years. Unchanged scores are not listed.*

A retrospective study done in collaboration with Vincent Finel[91] considered all previous Stanford DB versions since 2003 and PhenoTecT system to assess the resistance profiles of all patients in PhenoBase®. Genotypic and phenotypic predictions were compared. The finding was that, despite an overall stability of Genotyping over time, for some patients an evolution of Stanford DB predictions correlating with phenotypic results was observed. Figure 11 depicts this evolution for AZT in a pool of 17 patients.

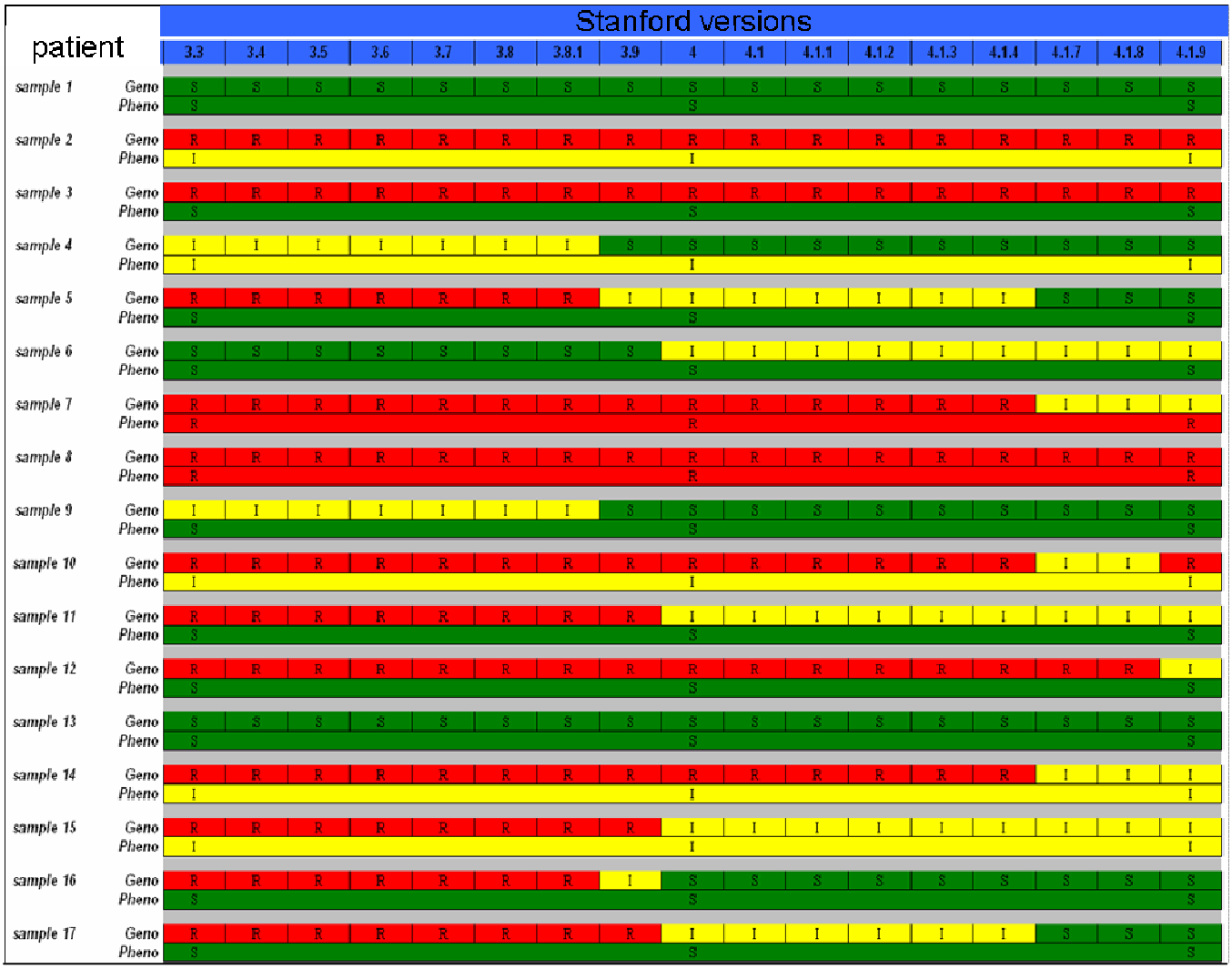


Figure 11. *Stanford DB algorithm variation over time set compared to the constant results of PhenoTecT.*

The 17 updates of Stanford DB algorithm did not have an impact on the interpretation of resistance for all samples. The profile of resistance by Genotyping is constant over time for samples 1, 2, 3, 8 and 13. Hence, this study shows that for rare samples (4, 6 and 7) the previous concordance observed between Phenotyping and Genotyping evolves into minor

discordances. The most important change pointed out in Figure 11 is the update of Genotyping leading to a "better" correlation with the PhenoTecT system. Samples 10, 14 and 15 diagnosed intermediate resistant to AZT by Phenotyping have a genotyping interpretation which evolves from resistant to intermediate. For samples 9, 11 and 12 susceptible to AZT by phenotyping, the update of penalties leads to a change from R→I (11 and 12) and I→S (9). The most remarkable evolution is observed for the samples 5, 16 and 17 susceptible to AZT by Phenotyping but resistant to the drug by Genotyping until 2004, later intermediate and finally susceptible in 2006.

III.3. Three main Algorithms to assess Genotyping

The Stanford DB algorithm is used in our laboratory to interpret the HIV sequences from patients. Many algorithms are available for the analysis of HIV sequences, and three main algorithms are employed by the different centres to assess Genotyping: ANRS, Rega and Stanford. As already described in the introduction *(I.5.1.)*, the Stanford algorithm is based on addition of mutation scorings for each drug whereas Rega and ANRS algorithms take into account the association of these mutations (antagonistic and synergistic interaction effects).

The growing number of distinct genotyping algorithms leads to different interpretations[92-94] and ultimately to different HAART prescription. This demonstrates a weakness and a limitation of those "modelling" systems. Figure 12 shows a geometric translation of a comparative discordance-assessment between three methodologies freely available through the internet (Stanford, Rega and ANRS for Genotyping) versus rPhenotyping. Below they are defined as points in space, and distances equal "percent of discordance" between two respective systems. Thereby the base of the resulting tetrahedron represents the side-by-side comparison of three genotyping algorithms, and the area within this (blue) triangle visualises and quantifies the degree of relative discordance. The addition of replicative Phenotyping as a third dimension allows the independent assessment against every algorithmic system, quantified e.g. by the volume of the tetrahedron.

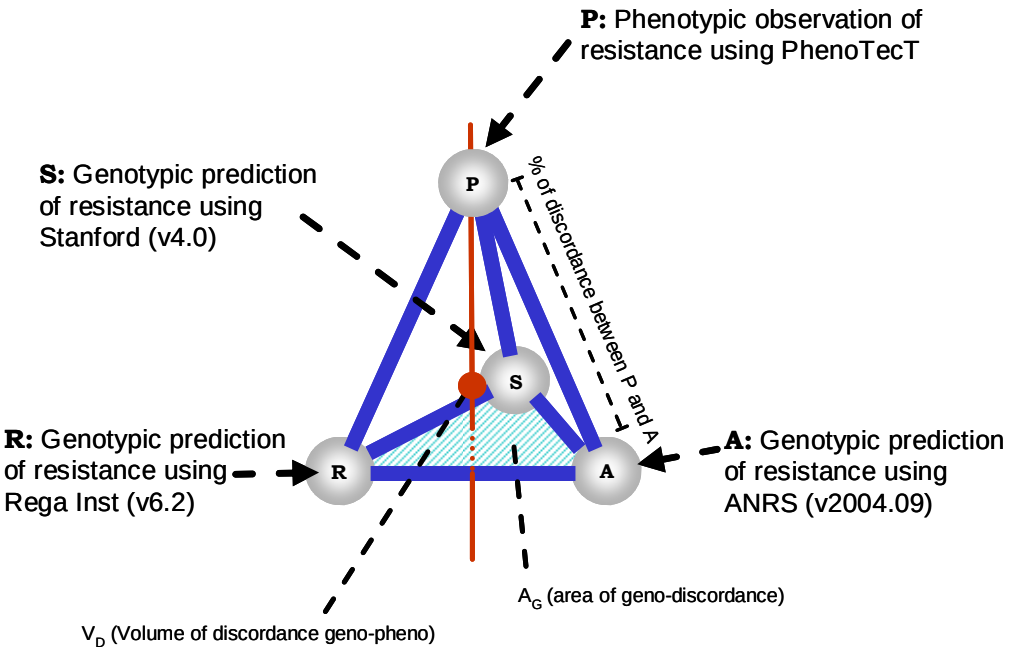


Figure 12. *Tetrahedron model displaying genotypic and phenotypic discordances.*

A retrospective study done in collaboration with Stéphane Hubert[95] considered all samples analysed between 2002 and 2004 at InPheno. The sequences were re-interpreted with the updated version of each algorithm at that time (ANRS v.2004.09; Rega v.6.2 and Stanford v.4.0) and compared with the phenotypic result sent to the physicians. For each of the RTIs a tetrahedron has been constructed (Figure 13) (in this study they supplied also the model for PIs, data not shown). FTC was not implemented in the interpretation of resistance at this time and ddC was removed from the analysis.

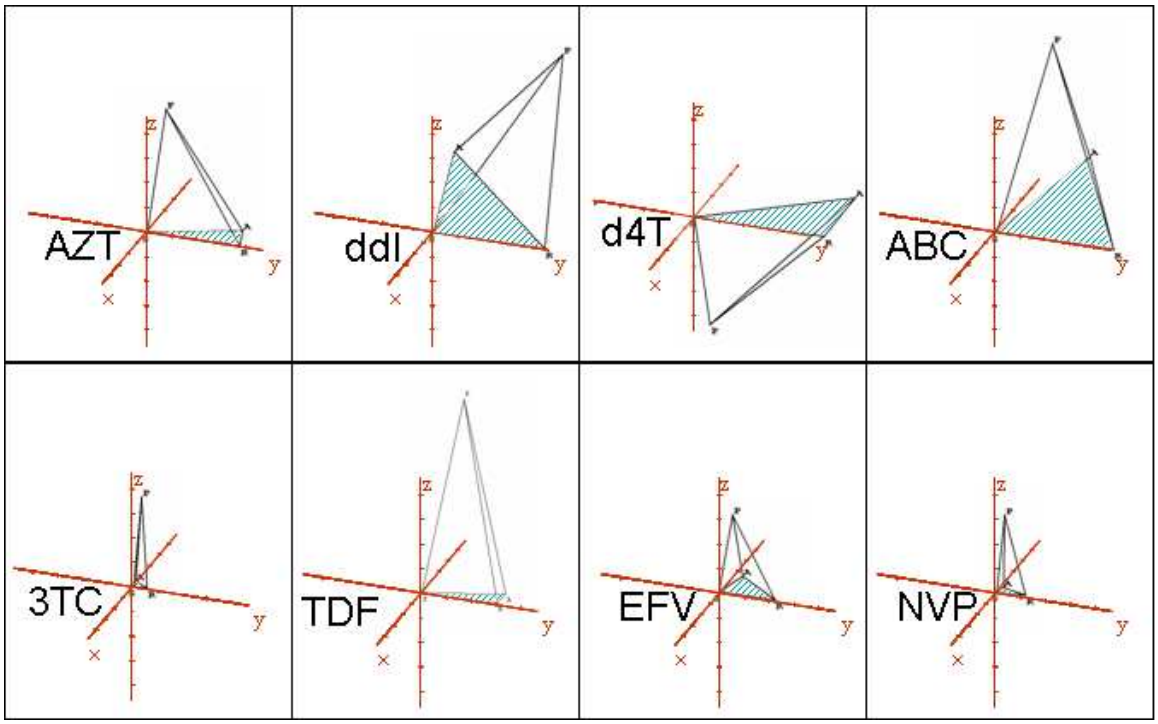


Figure 13. *Geometric translation of discordance for each of the 8 RTIs.*
The blue area represents the genotypic discordances where "x" and "y" axes delineate a plan. The "z" axis corresponds to the third dimension built by the phenotypic result compared to any Genotyping.

Tetrahedron shapes appear "stochastic" with no obvious trends and relationships for drug classes with the exception of NNRTIs and 3TC. For this group the small size of the basis reflects a good agreement among the genotypic systems and correlates with the dominant role of characterized mutations at positions 103 and 184, respectively. The wider areas are observed mainly for ddI, d4T and ABC for which the three genotypic interpretations are highly discordant (20-40%). For AZT and TDF two algorithms – ANRS and Rega – give a similar result (less than 5% of discordance) whereas Stanford shows more discrepancies with both of them (~ 15%).

The fourth method rPhenotyping is symbolised by the height of the tetrahedron. We can distinguish three categories: (i) equal distance from each of the three Genotyping (TDF, 3TC and NNRTIs); (ii) rPhenotyping is more concordant with ANRS (ABC and ddI); (iii) rPhenotyping and Stanford present less discrepancies (AZT and d4T).

These three main algorithms used to determine the profile of resistance of HIV reveal some discordances between their interpretations. One common point in their rules, which is

often blamed, is the unique reference to the B subtype of HIV to calculate the level of resistance. During several scientific conferences I attended, one of the main requests asked by scientists was the implementation of non-B subtype resistance data in the algorithms. In Switzerland the proportion of non-B subtype increases every year: 2003 & 2004 → 25%; 2005 & 2006 → 30%; 2007 → 40%. In the study done by Hubert *et al.*[95], the discordances between the three genotypic algorithms, as quantified by the base of the tetrahedron, were recalculated for each of the 8 RTIs independently for B and non-B subtypes. The obtained data reported as histogram in Figure 14 reveal a dramatic increase in disagreement between the prediction systems for non-B viruses (yellow bars). A possible explanation for this phenomenon is that mutational pattern overall differ significantly for B and non-B subtypes that are judged according to B-knowledge. However, due to the stringency of sample selection, the non-B/discordant population did not yet reach statistically exploitable frequency for 3TC and NNRTIs.

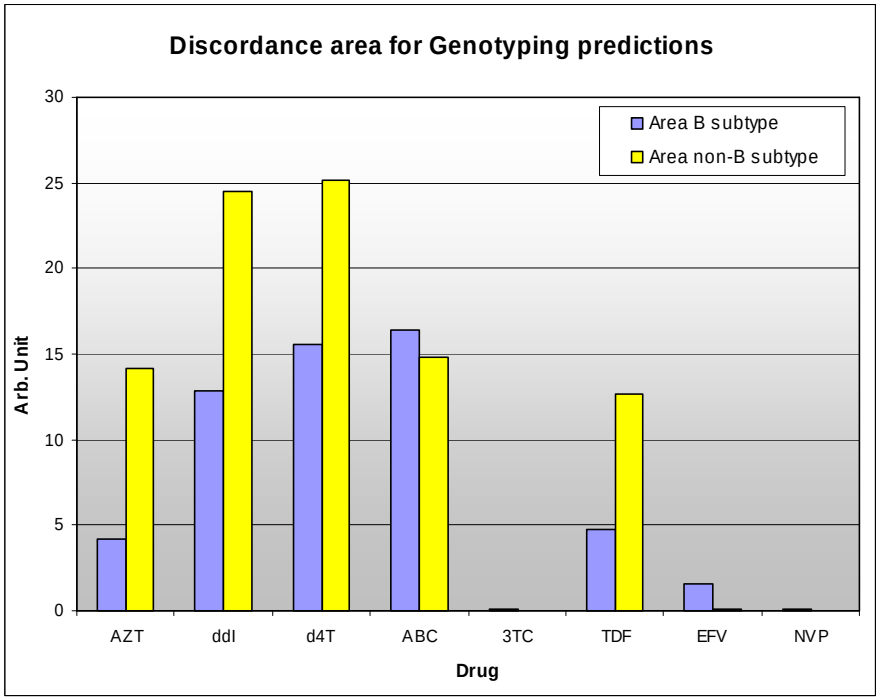


Figure 14. *Discordance between genotypic systems for B- and non-B subtypes.*

III.4. Discussion chapter I

Genotyping and rPhenotyping present overall a high degree of concordance in the resistance levels of the different RTIs. However for some samples two analyses with opposite diagnoses are provided and physicians are in the dilemma to choose the best treatment for their patients. My interest was to further analyse these discordant profiles to be able to give an adapted recommendation for treatment.

Two subgroups of RTIs have been characterized in this chapter:

- a first group presenting more than 20% of discordance between genotypic analysis and rPhenotyping and for which the major discordances were mainly Geno-R/Pheno-S. The related drugs were AZT, d-drugs, ABC and TDF. The addition of several weak mutations is necessary to prevent their action.
- a second group for which Genotyping and rPhenotyping were discordant in around 10% of samples. The statistical data shown previously indicate that these drugs (3TC, FTC and NNRTIs) had a high percentage of samples Geno-S/Pheno-R. In their case a single mutation in the RT gene can lead to HIV resistance.

The comparison of three main algorithms to assess genotypic results revealed that, in general, any of them was more concordant with phenotypic analysis. However, for the second aforementioned subgroup, all genotypic algorithms were in good agreement and presented a low level of discrepancy with rPhenotyping. The latter observation could be explained by the strong impact of several single mutations recognised by all algorithms.

Genotyping algorithms require regularly updating to: (i) better score the impact of each mutation already known; (ii) implement studied mutations newly characterized. Each genotypic interpretation is based on different rules (addition or association of mutations) but none of them implemented data from non-B subtype, which are more and more present in Europe and possibly soon also in North America. In general for RTIs neither one nor the other genotyping algorithm correlates more closely with rPhenotyping. These observations appear to indicate that multiple reasons could explain the discrepancies. This is what led me to look individually at each group of major discordance in order to identify a clinical explanation. In the next two chapters, I will dissect the major discordances observed between Stanford DB algorithm and rPhenotyping and assign them to either one of two “baskets”: Geno-R/Pheno-S and Geno-S/Pheno-R.

IV. Chapter II: Evaluation of rPhenotyping to detect complex patterns of mutations

IV.1. The presence of HIV quasispecies could influence the interpretation of resistance

IV.1.1. Background

The dynamics of HIV-1 replication are fast, and the generated viral population is also *in vivo* characterized by a high diversity[96]. This variation includes the inevitable and, in principle, predictable accumulation of mutations such as those conferring resistance to antiviral drugs. It leads to viral quasispecies which can coexist and present different patterns of mutations[21]. However, their detection is not at all trivial. As example, the sequencing methodology is based on a 'consensus' sequence of a single virtual virus meaning such Genotyping loses the ability to separate and to assign mutations to individual genomes within mixed populations. Genotypic algorithms thereby tend to overestimate resistances of viral mixtures after placing all sampled mutations into the same "virtual genome" (Figure 15).

In contrast, the robust replicative Phenotyping system PhenoTecT enables viruses to replicate up to four times, yet restricting *in vitro* evolution of new viral species. This allows to discriminate e.g. virus mixtures in plasma. Indeed, by rPhenotyping each single virus of a mixture will evolve under drug pressure and those with resistance associated mutations will be able to replicate. This is why we hypothesised that discordant profiles with Geno-R/Pheno-S could correspond to samples presenting a particular viral mixture. In fact, a homogeneous viral population should be interpreted identically by both methodologies. Nevertheless, for mixed populations there are two possibilities: either at least one single virus of the mix is resistant to the drug and the profile will be Geno-R/Pheno-R (Figure 15, Case 3); or each single virus of the mix is susceptible but carries mutations that would lead to resistance if they were all *in cis* on a single genome. This latter case will be interpreted resistant by Genotyping whereas rPhenotyping will diagnose it susceptible: Geno-R/Pheno-S (Figure 15, Case 2).

This is true for mutations with low scores, which require the presence of additional mutations before leading to drug resistance. In case of highly scored mutations Genotyping and rPhenotyping will agree in diagnosing resistance when either only "one viral species" or a mixture of virus is present.

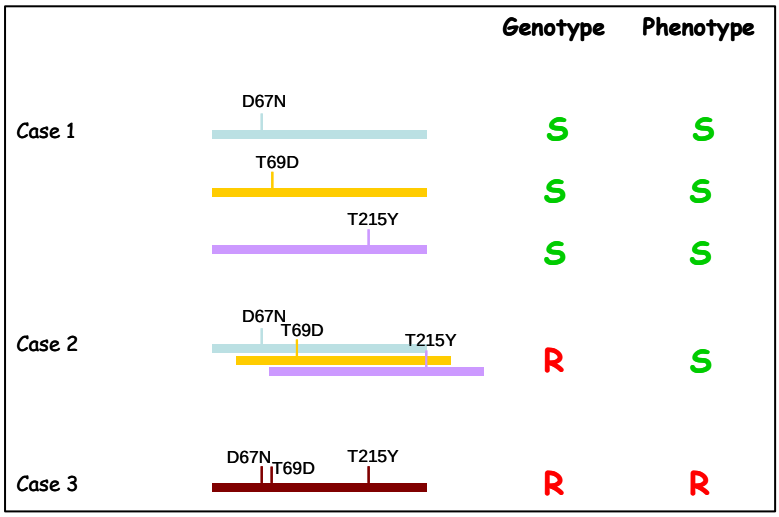


Figure 15. *Theoretical application of the addition of mutations analysed by Genotyping and Phenotyping for the assessment of AZT resistance.*
Case 1: three patients infected each by one virus carrying a single RT-mutation;
Case 2: one patient presenting viral quasispecies (three single mutated virus);
Case 3: one patient infected by one AZT-resistant virus carrying three RT-mutation in cis.

In this section, to demonstrate the validity of our hypothesis of mixed populations, we will try to answer the following questions:

1. Do Geno-R/Pheno-S profiles correlate with the presence of viral mixtures?
2. Does rPhenotyping discriminate between triple/quadruple mutants and a viral mixture containing single mutants?
3. Do we have clinical samples in PhenoBase® illustrating this theory?

IV.1.2. Results

IV.1.2.1. Statistical data

First, PhenoBase® was interrogated to see if there is a correlation between the co-existence of at least two viruses in a sample and those patients presenting a Geno-R/Pheno-S profile for the RT gene. A viral mixture was defined as virus presenting a RT sequence with one or more characterized-mixed mutations. Samples from the years 2005 and 2006 were selected and each drug (NRTIs and NNRTIs) was separately evaluated for concordance between Genotyping and rPhenotyping (Geno/Pheno profiles). More than 5,200 distinct drug resistance profiles on RT were split into two categories according to whether they were concordant or discordant (Figure 16).

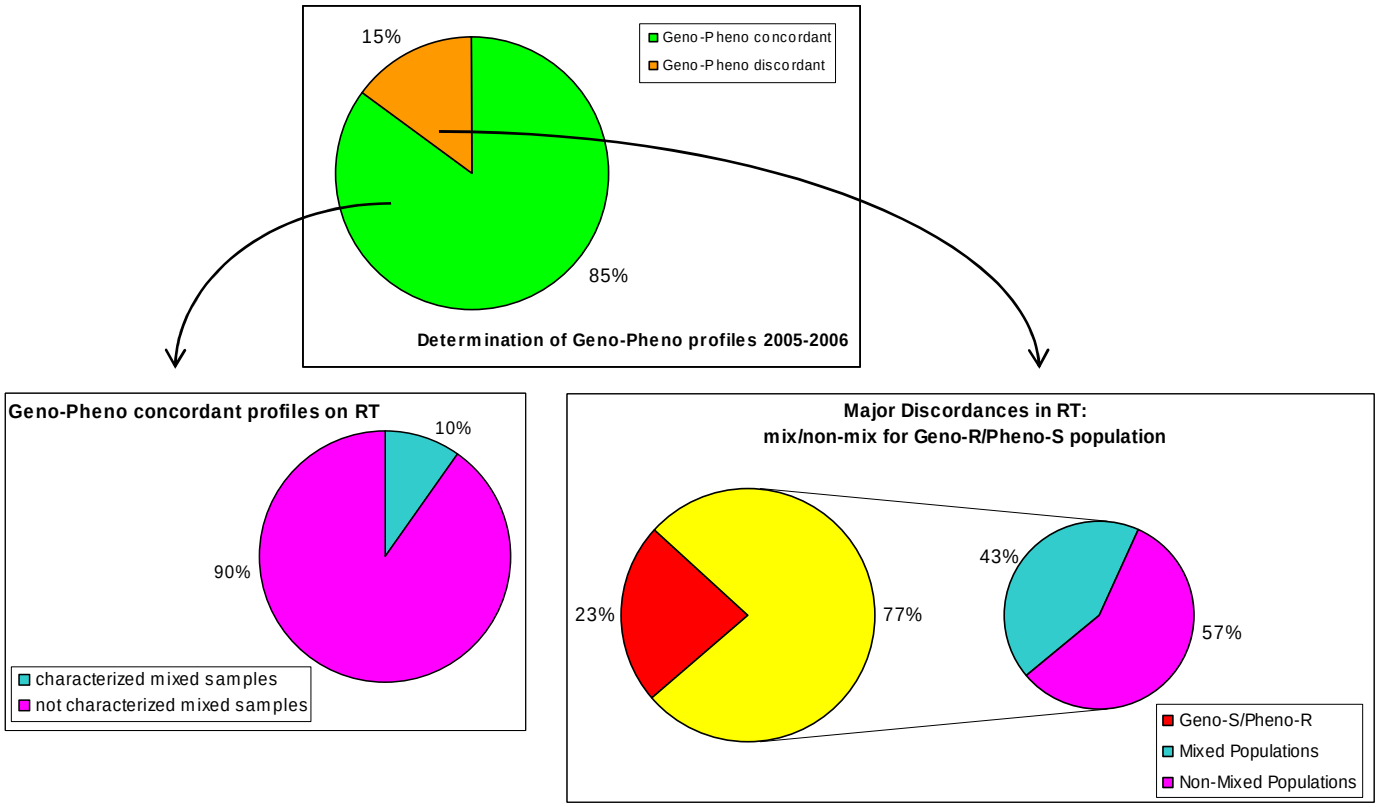


Figure 16. *Repartition of viral mixtures into concordant and discordant samples from the years 2005 and 2006.*

This first analysis (upper graph) made clear that discordance is a relatively rare observation among clinical samples in PhenoBase® (15%). To demonstrate our assumption that mixes of different viruses are one of the causes of discordance between Genotyping and Phenotyping, we further examined the concordant group dividing it in two subcategories according to whether those samples contained viral mixtures or not. We observed in

concordant RT that 10% of the samples contained clear mixes. On the other side we subdivided the Geno/Pheno discordant profiles in the first pie-chart according to their type of major discordance, i.e. Geno-R/Pheno-S (yellow) or Geno-S/Pheno-R (red). It came out that the latter occurred less frequently. We looked for mixes in the Geno-R/Pheno-S group. We found out that 43% contained virus mixes for RT. When this value was compared to the 10% found before it was possible to statistically associate samples with virus mixes and Geno-R/Pheno-S discordances.

To confirm this prospective association by statistical rules the calculation of odds was done on RT data. The odds ratio is a measure of effect size. It is defined as the ratio of the odds of an event occurring in one group to the odds of it occurring in another group. These groups might be any dichotomous classification. If the probabilities of the event in each of the groups are p (first group) and q (second group), then the odds ratio is:

$$\frac{p/(1-p)}{q/(1-q)} = \frac{p(1-q)}{q(1-p)}.$$

An odds ratio of 1 indicates that the condition or event under study is equally likely in both groups. An odds ratio greater than 1 indicates that the condition or event is more likely in the first group. And an odds ratio less than 1 indicates that the condition or event is less likely in the first group. The odds ratio must be greater than or equal to zero.

The first calculation of odds ratios was done based on discordant and concordant populations in 2005 and 2006. I looked to each amino acid of the reverse transcriptase gene to observe the presence or absence of mixes. In the table below, the number of samples is shown for each population (Geno-R/Pheno-S discordant (R/S) versus non R/S and concordant, and mixed versus non-mixed). The event is “viral mixture occurring in a first group: R/S profile and in a second group: non-R/S or concordant profiles”.

	R/S	non R/S	concordant
mix	*134*	*2767*	*2407*
non mix	*47*	*1947*	*1753*

Table 4. *Distribution of the diagnosed samples into three groups: discordant Geno-R/Pheno-S, non-discordant Geno-R/Pheno-S and concordant; and in two events: mixed population and non-mixed population.*

From the given numbers we inferred the odds ratio (OR) for mixes as follows:

- R/S versus non R/S: **OR** = (134*1947)/(47*2767) = 2
- R/S versus concordant: **OR** = (134*1753)/(47*2407) = 2.1

These odds ratios illustrated that a mixed viral population was more likely to occur in the Geno-R/Pheno-S profile than in any other category. To be interpreted resistant to a drug by Genotyping a virus should carry several characterized mutations. The hypothesis of the presence of viral mixture to explain the discrepancies Geno-R/Pheno-S would mean that these characterized mutations should be found in different viruses. This is why we focused on the specific presence of mixed characterized mutations in the studied samples. The values corresponding to this new request are summarised in Table 5.

	R/S	non R/S
characterized mix	*78*	*549*
not characterized mix	*103*	*4165*

Table 5. *Distribution of the diagnosed samples in two groups: discordant Geno-R/Pheno-S and non-discordant Geno-R/Pheno-S; and in two events: population with mixes of characterized mutations and population without mixes of characterized mutations.*

From the given numbers we inferred the odds ratio for mixes as follows:

- R/S versus non R/S: **OR** = (78*4165)/(103*549) = 5.7

The odds ratio was 5.7, showing that a viral mixture was much more likely to occur in the Geno-R/Pheno-S profile than in the non R/S profile. To conclude that it was one cause of this kind of discordance, I could show that for the samples Geno-R/Pheno-R the viral mixtures were less frequent. The values in the Table 6 represent the number of samples diagnosed, between the years 2005 and 2006, either identically by Genotyping and rPhenotyping (resistant or susceptible) or Geno-R/Pheno-S.

	R/S	R/R	S/S
characterized mix	*78*	*147*	236
not characterized mix	*103*	*318*	3397

Table 6. *Distribution of the diagnosed samples into three groups: discordant Geno-R/Pheno-S and concordant resistant or susceptible; and in two events: population with mixes of characterized mutations and population without mixes of characterized mutations.*

From the given numbers we inferred the odds ratio for mixes as follows:

- R/S versus R/R: **OR** = (78*318)/(103*147) = 1.6
- R/S versus S/S: **OR** = (78*3397)/(103*236) = 10.9

Looking especially at the characterized mutations the odds ratio between Geno-R/Pheno-S and Geno-S/Pheno-S was high: 10.9. It is due to the fact that samples sensitive to drugs are often, if not always, from patients with a primary infection or early-treated patients. They rarely have characterized mutations and multi-species, which appear under drug pressure. In case of samples genotypically and phenotypically resistant to ARV the odds ratio showed that viral mixtures were less likely to occur than in the Geno-R/Pheno-S profile.

Conclusion: The odds ratios revealed a strong correlation between the presence of viral mixtures in a sample and a discordant profile Geno-R/Pheno-S. These data confirmed the relevance of the working hypothesis. In a given sample the presence of several viruses carrying different patterns of characterized mutations could be one of the causes of Geno-R/Pheno-S profiles.

IV.1.2.2. Discrimination of rPhenotyping between mixes of single mutants versus combinations of mutations on the same virus

To determine the ability of rPhenotyping to discriminate between triple or quadruple mutants and a viral mixture of single mutants, several mutant plasmids will be engineered and tested in the phenotypic assay. Based upon PhenoBase® requests done in *chapter I*, the mutations will be selected to give genotypic resistance to AZT, which showed together with ddI, the higher number of Geno-R/Pheno-S profiles among the RTIs. Furthermore, resistance to AZT is due to the addition of at least 3 or 4 mutations, which perfectly matches the working hypothesis. A virus carrying less than three mutations will be genotypically susceptible to AZT, and the presence of several such viruses should be declared resistant by Genotyping and susceptible by rPhenotyping.

To construct relevant HIV mutants, we looked at the list of samples resistant to AZT according to Genotyping but sensitive in Phenotyping: between the years 2003 and 2006, 83 samples were diagnosed Geno-R/Pheno-S to AZT. The Table 7 summarises the distribution for each characterized RT mutation in the selected samples.

Position	41	44	62	65	67	67	69	69	70	70	74	74	75	77	100
AA	L	D	V/P	X	N	-/G/E	D	N/G	R	S/E	I	V	M/I/T	X	X
% (N=83)	54.2%	13.3%	4.8%	0.0%	69.9%	6.0%	20.5%	14.5%	36.1%	4.8%	12.0%	12.0%	3.6%	0.0%	0.0%

Position	115	116	118	151	181	184	210	210	215	215	215	219	219	219	333
AA	F	Y	I	M	C	V	W	F/G/V	F	Y	V/I/C/S	Q	E	N/R	E
% (N=83)	2.4%	1.2%	32.5%	4.8%	19.3%	68.7%	34.9%	6.0%	30.1%	49.4%	10.8%	34.9%	15.7%	8.4%	10.8%

Table 7. *Distribution of RT characterized mutations for the samples Geno-R/Pheno-S to AZT (years 2003-2006). The mutations occurring in more than 20 % are figured out in red. AA: Amino Acid. X: any mutation.*

The 10 positions highlighted in red are the amino acids more frequently mutated in samples having a discordant profile to AZT. Six of them correspond to the TAM. To see if they have all a similar impact on AZT resistance, I checked for their AZT mutation scoring in the Stanford databank: M41L (15), D67N (15), T69D (5), K70R (18), V118I (2), Y181C (-4), M184V (-8), L210W (15), T215Y/F (35) and K219Q (15). The TAMs have a penalty score from 15 to 35 which increases AZT resistance when they are associated. The two mutations T69D and V118I contribute to AZT resistance when they are present with multiple TAMs. In contrary, Y181C and M184V partially reverse AZT resistance caused by other AZT mutations (TAMs).

To demonstrate practically the observations described in Figure 15, several mutants were engineered with one, two, three or four resistance associated mutations in an NL4-3 background. Among the previous mutations with a positive scoring some TAMs were selected (67N, 215Y and 219Q) as well as one non-TAM (69D). All mutant proviruses were sequenced to be validated before being included in the next experiments. Phenotyping was performed directly on each mutant plasmid in the replicative Phenotyping system PhenoTecT in the presence of AZT. In addition, individual mutants were assessed as mixtures. As a control for the PR gene, which should not contain any mutation, viruses were grown in parallel in presence of the protease inhibitor SQV.

Every experiment was performed in triplicate, and twelve or six drug dilutions were used to extrapolate the inhibition curves in a reliable way. The total amount of transfected provirus was the same for a single virus or a mix of virus, equal to 2.5 µg. For curve fitting and statistical analysis of the experimental data XLfit version 4.1.1 was used. The entire experiment was repeated independently five times in order to reach a good degree of confidence. Figure 17 below represents the inhibition curves for triple mutant 67N-69D-215Y (green) or the mix of the three single viruses (blue) compared to a wild type reference: pBX (red).

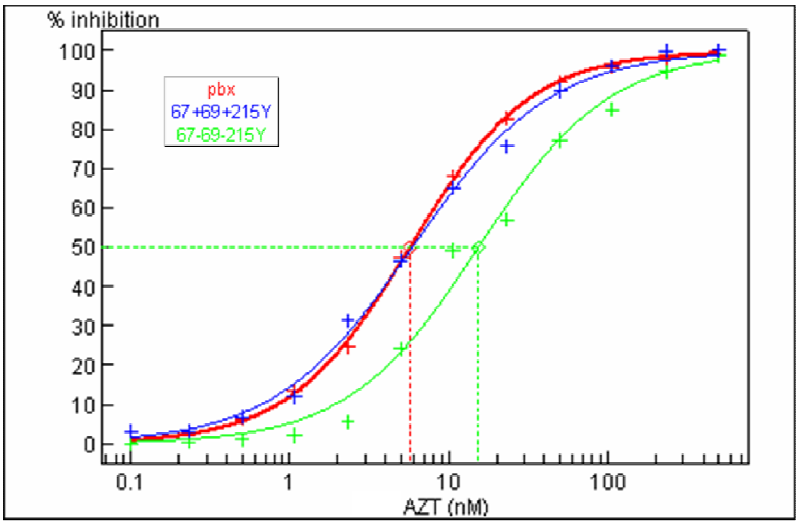


Figure 17. *Inhibition curves for one provirus carrying three mutations (green) and for a mix of the three single mutants (blue) show intermediate resistance to AZT and full susceptibility, respectively. Both viruses are SQV susceptible. Red curve = wild type virus.*

The inhibition curve corresponding to the triple mutant was shifted to the right comparing to the reference meaning a decrease in susceptibility to AZT. The IC_{50} of the reference was equal to 5.7 nM and for the triple mutant to 16.5 nM. The ratio of their IC_{50} (Rf) was: 16.5/5.7 = 2.9. To interpret this value it was compared to a cut-off value for AZT. There are two values determining the threshold between susceptibility and intermediate and between intermediate and resistance: 2.5 for the first and 4.5 for the second. It means that the triple mutant 67N-69D-215Y, with a Rf of 2.9, was intermediate resistant to AZT by rPhenotyping. To the contrary, it is clearly shown on the graph that the addition of the three single mutant 67N + 69D +215Y was susceptible to AZT with an IC_{50} of 6.5 nM (its curve superimposed the reference curve). Similar curves were obtained for the quadruple mutant 67N-69D-215Y-219Q and its respective single mutants, the shift with the reference was even wider (data not shown). rPhenotyping in these particular conditions was able to discriminate between an association of three single viruses and the derived triple mutant.

To determine the level of discrimination of the phenotypic assay, different combinations of four mutations: D67N, T69D, T215Y and K219Q were assessed in parallel. The associations of proviruses were the following:

- two double mutants 69D-215Y with 67N-219Q,
- one triple mutant 67N-69D-215Y with one single mutant 219Q,
- one quadruple mutant 67N-69D-215Y-219Q,
- four single mutants 67N + 69D + 215Y + 219Q.

The inhibition curves for AZT are shown in Figure 18.

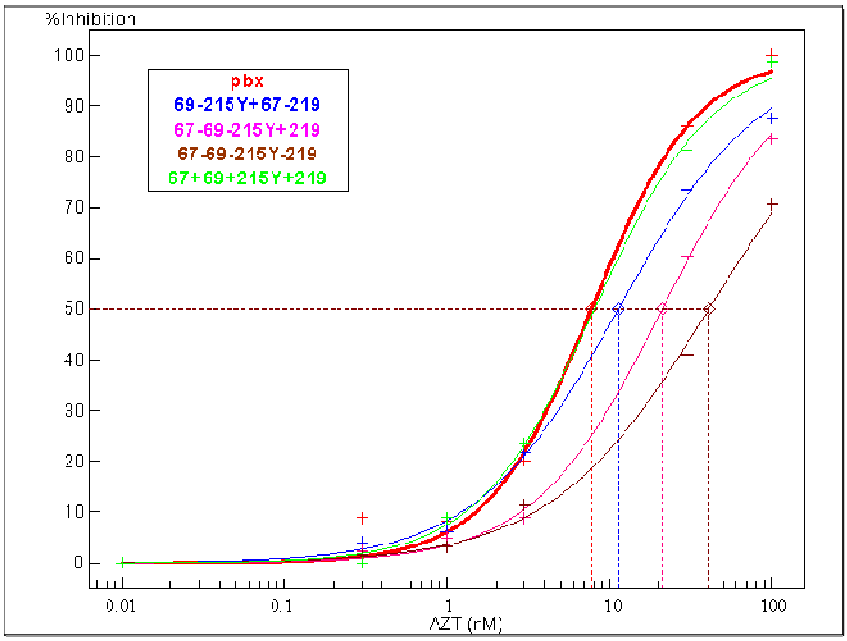


Figure 18. *Inhibition curves for four combinations of proviruses carrying mutations 67N, 69D, 215Y and 219Q show a decrease in susceptibility to AZT with the increase of mutations carried by a single virus. All viruses were susceptible to SQV. Red curve = wild type virus.*

By Genotyping (Stanford algorithm) all associations of mutated viruses are indifferently interpreted resistant to AZT; whereas by rPhenotyping we observed a clear shift between the inhibition curves. The combination of the four single mutants was fully susceptible to AZT (superimposition of the green curve with the red reference) with an IC_{50} of 8.1 nM and a Rf of 1 (IC_{50} reference = 7.7 nM). When two double mutants were co-transfected: 69D-215Y + 67N-219Q, the susceptibility to AZT was decreased compared to the reference and the four single mutants. Their IC_{50} was 11.3 nM (blue curve) giving a Rf of 1.5 which corresponded to a susceptible virus. The addition of a third mutation on the same virus, 67N-69D-215Y combined with the single mutant 219Q led to a shift to the right of the inhibition curve (pink) which provided an IC_{50} of 21.1 nM and a Rf of 2.7. This mix of triple and single mutants was intermediate resistant to AZT. The last tested virus was carrying the four mutations 67N, 69D, 215Y and 219Q *in cis*. Its inhibition curve (brown) was to the far right which produced an IC_{50} of 40.9 nM and a Rf of 5.3, meaning resistance to AZT.

The genotypic and phenotypic profiles of all the different proviruses which were mutated are summarised in Table 8 below.

mutants	D67N	T69D	T215Y	K219Q	69D-215Y	67N-219Q	67N-69D-215Y	67N-69D-215Y-219Q
Stanford Geno	S	S	I	S	I	I	R	R
rPhenotyping	S	S	S	S	S	S	I	R

mutants	69D-215Y + 67N-219Q	67N-69D-215Y + 219Q	67N-69D-219Q + 215Y	69D + 215Y	67N + 219Q	67N + 69D + 215Y	67N + 69D + 215Y + 219Q
Stanford Geno	R	R	R	I	I	R	R
rPhenotyping	S	S	S	S	S	S	S

Table 8. *Genotypic and phenotypic profiles of resistance to AZT done for single, double, triple and quadruple mutants, and for the association of the corresponding single mutated provirus. S: susceptible, I: intermediate and R: resistant.*

The main conclusion is that the PhenoTecT assay is able to discriminate between a single virus engineered with several mutations *in cis* and a mixture of virus each carrying only one of the same mutations (Figures 17 and 18). With two of the constructed mutants we could easily observe a discordance of profile done by rPhenotyping when the mutations were either engineered on a single virus or carried by several viruses. These plasmids mutated at positions 67N, 69D and 215Y or 67N, 69D, 215Y/F and 219Q were intermediate or resistant to AZT by Phenotyping whereas the addition of their respective single mutated virus was susceptible.

However, some non expected results have been observed. Indeed, according to the Stanford algorithm, mutation T215Y contributes to intermediate resistance for AZT (penalty score of 35) but the PhenoTecT test diagnosed it susceptible to AZT. This discordance could be either due to an overestimation of the impact of mutation T215Y on AZT resistance or due to an *in vitro* assay limitation. In fact, as described in *part I.4.2.*, the resistance mechanism with TAMs mediated the removal of the analogue from the terminated DNA chain. The presence of such mutations will allow the binding of ATP on the reverse transcriptase and further the removal of the RTI. A low amount of ATP in the cells used for the phenotypic assay could prevent an optimal release of the drug and lead to a false interpretation of ARV sensitivity. The same observations have been done for both double mutants.

IV.1.2.3. Dissection of clinical samples by dilution cloning

The working hypothesis: "Geno-R/Pheno-S profile could be due to the presence of viral mixture" has been demonstrated to be true for some mutated proviruses. In these plasmids only the selected mutations were present but the viral population in clinical samples may carry also additional uncharacterized mutations which could have a further influence on resistance profile. It is why I wanted to dissect several samples to see the clinical relevance of our finding. Two samples with a discordant profile Geno-R/Pheno-S for AZT were selected: MD1310-04 and 6016764. Both had a sequence showing clear double-peaks at determinant positions (Figure 19).

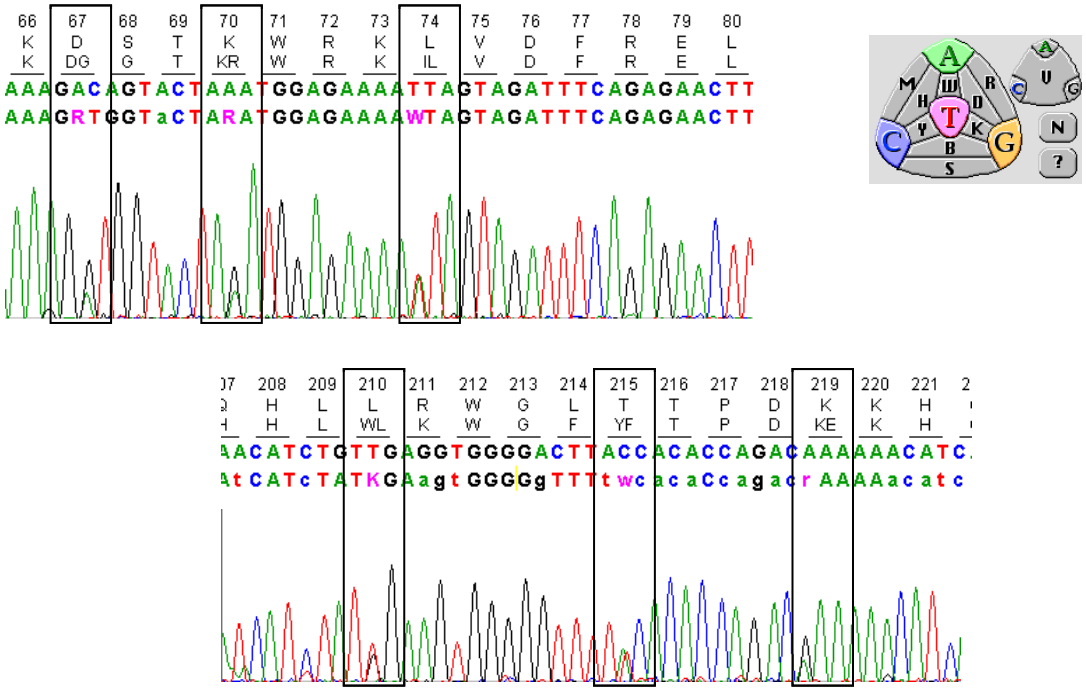


Figure 19. *Electrophoregram from the RT gene of sample 6016764 showing several mixed amino acid (20 non silent mixed mutations).*

As an internal control, in the region surrounding the patient's RT gene, which corresponds to the cloning vector, no double peak was observed and the sequence was, as expected, isogenic to pNL4-3.

Consequently, genotypic analysis was restricted to one of the numerous possible combinations of the detected mutations. The penalty scores of each mutation for a drug were all added to give a final penalty interpreted as level of resistance. The distinction of the different virus strains in the initial population could not be done by Genotyping. At present,

there is no solution for this problem. The hypothesis that more than one species is sometimes, if not always, present in the same patient was easily demonstrated through dilution cloning of bacteria, which were previously transformed with recombinant plasmids obtained from a mixture of various reverse transcriptase genes. For the two samples mentioned above, eight separate colonies were picked and sequenced.

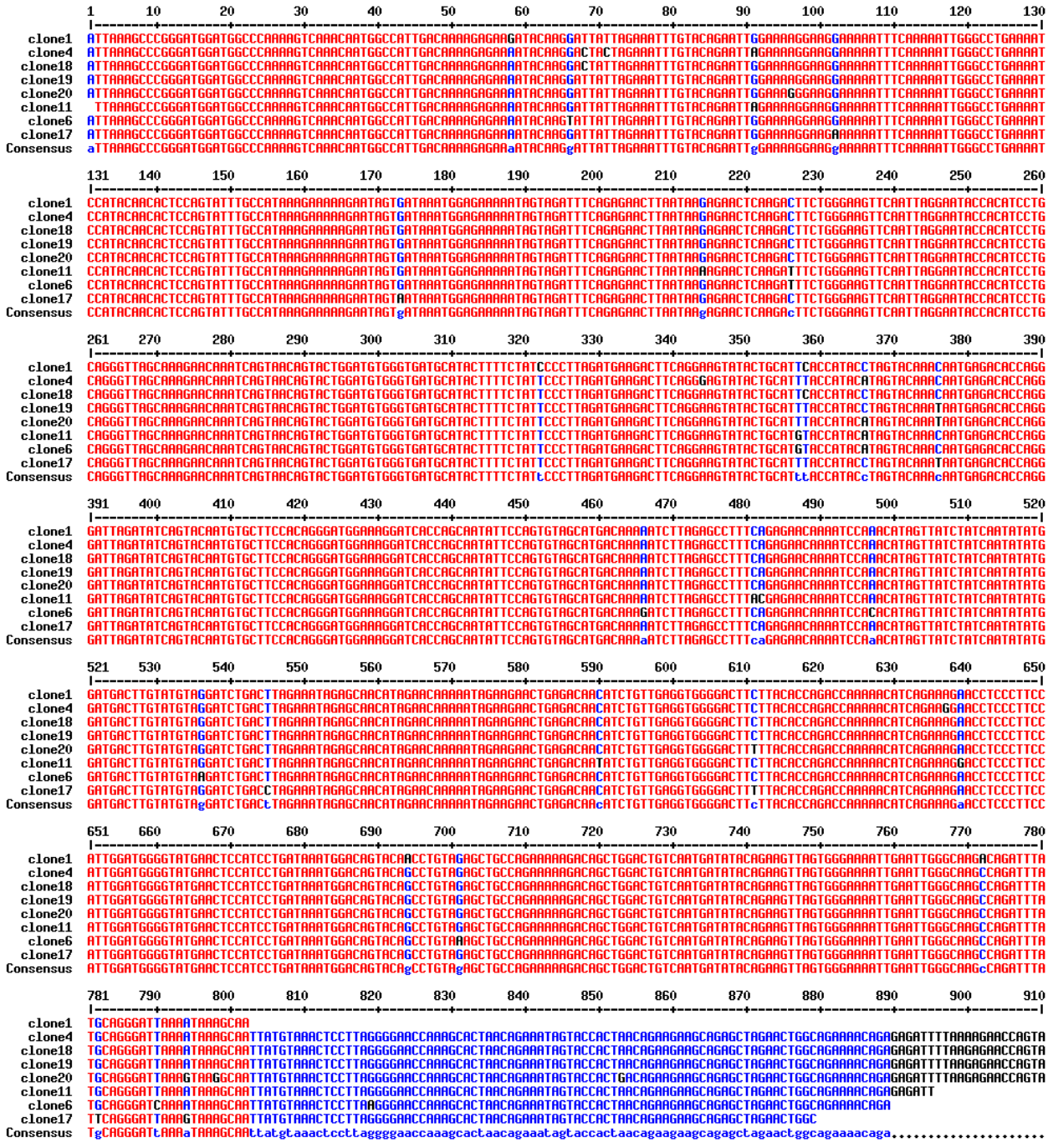


Figure 20. *Alignment, using Multalin[H], of eight clones derived from sample MD1310-04.*

Figure 21. *Alignment, using Multalin[H], of eight clones derived from sample 6016764.*

After aligning the sequences using Multalin[H], in thirty five (MD1310-04) and fifty three (6016764) distinct positions more than one nucleotide was detected, suggesting that various populations coexisted in the same patient at the time of blood draw.

A possible cause for the observed mutations could be that they were simply a PCR artefact arising during HIV cDNA amplification. For this reason, iProof High-Fidelity DNA polymerase (52-fold more accurate than Taq polymerase) was used to amplify the DNA, which was in turn produced using the Applied Biosystems kit that employs AmpliTaq Gold DNA polymerase (which has a slightly lower fidelity). This expedient should already reduce the probability to introduce errors during the amplification steps, but to rule out this hypothesis the RT gene of a clonal mutant, already sequenced, containing four defined mutations was amplified in parallel. No additional substitution, except those four, was detected by sequencing. Moreover, the clustering of different mutations observed in the clones obtained from the two patients suggests that they were not the product of a randomly occurring event, otherwise they would be evenly distributed across the whole gene.

The genotypic profiles of the different clones for the two patients selected are indicated in the Table 9 below.

sample MD1310-04

	mutations in the mix	M41L	D67N	T69D	L74I	V118I	T215FL	K219Q	101A	K103N	
CLONES	1	M41L	D67N	T69D	L74I	V118I	T215L	K219Q	101A	K103N	
	4	M41L	D67N	T69D	L74I	V118I	T215L	K219Q	101A	K103N	
	6	M41L	D67N	T69D	L74I	V118I	T215L	K219Q	101A	K103N	G190R
	11	M41L	D67N	T69D	L74I	V118I	T215L	K219Q	101A	K103N	
	17	M41L	D67N	T69N	L74I	V118I	T215F	K219Q	101A	K103N	
	18	M41L	D67N	T69D	L74I	V118I	T215L	K219Q	101A	K103N	
	19	M41L	D67N	T69D	L74I	V118I	T215L	K219Q	101A	K103N	
	20	M41L	D67N	T69D	L74I	V118I	T215F	K219Q	101A	K103N	

sample 6016764

	mutations in the mix	D67DG	K70KR	L74LI	V118VI	M184V	L210LW	T215FY		K219KE	K103KN	Y181C	G190GA
CLONES	1	D67G	K70R		V118I	M184V		T215F		K219E		Y181C	G190A
	4	D67G	K70R			M184V		T215F		K219E		Y181C	G190A
	7	D67G	K70R	L74I	V118I	M184V			T215Y		K103N		
	8	D67G	K70R		V118I	M184V	L210W		T215Y	K219E		Y181C	G190A
	11	D67G	K70R	L74I		M184V	L210W		T215Y				
	12	D67G	K70R	L74I		M184V		T215F		K219E		Y181C	G190A
	16	D67G	K70R		V118I	M184V		T215F		K219E		Y181C	G190A
	20	D67G	K70R		V118I	M184V	L210W		T215Y	K219E		Y181C	G190A

Table 9. *Profile of RT mutations for samples MD1310-04 and 6016764 and their eight clones.*

For sample MD1310-04 the eight clones carried the nine mutations detected in the mixed population. At position 215 two out of them presented the amino acid Phenylalanine instead of Leucine and one of these two had the mutation T69N but not T69D. Five clones were identical and the three others were unique meaning that four different viruses were selected in the eight clones.

In patient 6016764, Genotyping of the original sample detected 32 amino acid changes (13 characterized mutations and 19 uncharacterized). Twelve of them were unequivocal, whereas the other twenty were ambiguous because more than one residue was possible at the same position. This was a very representative example of the coexistence of more than one quasispecies in the same patient. For this patient, more than one million ($1^{12}*2^{20}$) possible different combinations could be obtained from the samples' substitutions. In the most simple (and the worst) case, all mutations would be in the same virus that coexists with the wild type. In the most optimistic case each single substitution would be carried by a different virus variant. Thus, twenty single mutant species would coexist with the wild type virus. Probably the actual situation lies between these two extremes. In case of twenty distinct single mutants, the mutations would not be detected by Genotyping, since each mutant nucleotide would only contribute to 5% of the entire peak signal. On the other side, the coexistence of the wild type virus with a single variant carrying all mutations *in cis* is questionable as well due to probabilistic calculation (probability of 0.00009%).

The sample 6016764 demonstrated that the resistance profile assessed through Genotyping represented only one of many different possible combinations. Moreover this combination might also not correspond to even one virus existing in the patient. The real solution to this endless problem is Phenotyping: all quasispecies are re-generated as separate entities as they were in the original blood sample and their susceptibility to ARVs is measured directly and irrespective of the mutant composition.

For the two aforementioned clinical samples and their eight isolated clones, phenotyping resistance profiles were assessed with AZT, ABC and ddI. SQV was used as control of non Reverse Transcriptase Inhibitor. The penalty scores and the Rfs corresponding to each clone and mixed population are summed up in the Tables 10A and 10B.

			CLONES							
		MIX	1	4	6	11	17	18	19	20
Geno	AZT	87	67	67	67	67	87	67	67	87
Pheno		2	0.8	0.9	1.3	0.9	5.9	1.1	0.4	0.9
Geno	ABC	67	57	57	57	57	67	57	57	67
Pheno		2.4	1.1	1.3	1.6	1.5	2.1	1.4	0.6	1.2
Geno	DDI	92	87	87	87	87	82	87	87	92
Pheno		1	0.9	1.3	0.9	1.6	2.1	1.9	1.5	1
	fitness	95%	132%	90%	30%	79%	102%	109%	120%	95%

Table 10A. *Penalty scores (Geno) and Rfs (Pheno) of sample MD1310-04 and its clones for three NRTIs: AZT, ABC and ddI.*
Green: ARV susceptibility, red: ARV resistance.

Seven clones out of eight had the same genotypic and phenotypic profiles as the mixed population of sample MD1310-04 (Geno-R/Pheno-S). Only clone 17 was not susceptible to AZT which is not correlating with the PhenoTecT assay performed with the pool of RT genes. The unique difference observed between the sequences of clone 17 and the mixed population was at position 69, where the mutation in clone 17 was an Asparagine instead of a Threonine, whereas for all the other clones the mutation generated an Aspartate.

			CLONES					
		MIX	1	4	7	8	12	16
Geno	AZT	88	73	56	47	88	71	73
Pheno		2	0.6	0.4	0.9	2.3	2.5	0.8
Geno	ABC	74	42	32	54	54	60	42
Pheno		2.9	2.4	1.5	2.4	3.5	3.5	2.2
Geno	DDI	77	35	25	57	47	63	35
Pheno		1.3	1.3	0.7	0.8	1.8	1.8	1.1
	fitness	77%	100%	100%	88%	100%	100%	100%

Table 10B. *Penalty scores (Geno) and Rfs (Pheno) of sample 6016764 and its clones for three NRTIs: AZT, ABC and ddI.*
Green: ARV susceptibility, yellow: ARV intermediate resistance, red: ARV resistance.

The clonal viruses 11 and 20 from sample 6016764 were too unfit to produce interpretable data. The other viruses were all susceptible to AZT and ddI, as the mixed population, with Rfs ranging from 0.4 to 2.5 for AZT and from 0.7 to 1.8 for ddI. The intermediate resistance observed in mixed populations for ABC was also shown for two clones (8 and 12). Fitness was high with a minimum of 88%.

These data on the samples MD1310-04 and 6016764 having a discordant profile Geno-R/Pheno-S for AZT were not corroborating the previous work done with constructed mutants. Indeed, mutations leading to AZT resistance were all carried by a single clonal virus indicating a resistance profile by Genotyping, and though this defined virus was still detected susceptible by Phenotyping. Three other patients were dissected in the same way but the results obtained were similar (data not shown). The hypothesis that the presence of viral quasispecies could explain a discrepancy R/S between Genotyping and rPhenotyping is not invalidated but with this study I was not able to find a clinical sample which could correlate. It suggests that for the studied clinical samples another explanation should be found to their discordant profile of resistance to NRTIs. Are the TAMs overscored by genotypic algorithms? Do interactions between TAMs and additional resistance associated mutations decrease the level of resistance to AZT?

IV.2. Interactions between Resistance Associated Mutations

IV.2.1. Introducing the subject

The impact of individual mutations on viral sensitivity to a given inhibitor is regularly re-evaluated by the different algorithms used to provide genotypic profiles. Between the years 2003 and 2007, Stanford DB updated the penalty scores of characterized mutations on PR and/or RT 23 times (as described in *chapter I, III.2.*). As a consequence, an HIV sequence analysed retrospectively with different versions of the same algorithm could be differently interpreted by Genotyping. Moreover between algorithms themselves the resistance associated mutations can be scored variously (as described in *chapter I, III.3.*). Whereas Stanford algorithm is simply adding the penalty scores of the several mutations listed for a sample, ANRS considers their interactions, e.g. the presence of at least three mutations among M41L, D67N, K70R, L210W, T215A/C/D/E/G/H/I/L/N/S/V, K219Q/E gives resistance to AZT. It is even more complicated for interpretations of ddI resistance by ANRS because the combination of some mutations could lead to an antagonistic effect, e.g. the resistance is associated to a least a score of +2 among M41L, T69D, L74V, T215Y/F, K219Q/E (score = +1) and K70R, M184 V/I (score = -1). Thereof, a patient diagnosed in two centres using different algorithms could have two genotypic interpretations showing discrepancies. Lately the compensating secondary changes were often described in studies and their role appeared to be affected by the genetic background/subtype of the virus.

Phenotyping is a cellular assay to diagnose HIV resistance. Its advantage is the direct quantification of drug sensitivity for a viral population. As shown previously, the viral heterogeneity will be kept, e.g. viral mixtures with different patterns of mutations, to assess a resistance profile corresponding to the clinical situation. Nevertheless, the intracellular nucleotide pool could matter as it could differ from the PBL (Peripheral Blood Leucocyte) cells. A higher amount would lead to an underincorporation of NRTIs providing a wrong interpretation of drug activity. In this case discordance between genotypic and phenotypic diagnoses could arise. To improve the cellular assay, different cell lines should be tested or a cellular depletion of nucleotides should be performed. Another limitation of using cellular assay is the intracellular phosphorylation level. Indeed ARVs, as AZT and d-drugs, need to be phosphorylated in the cells prior to their incorporation in the DNA neo-synthesised. A too low level of phosphorylation will inhibit the activity of this class of drugs.

Stanford algorithm, as well as ANRS and Rega, is probably not perfect but which is closer to the *in vivo* situation? Does one correlate more with rPhenotyping? The tetrahedron model, as described in *chapter I*, would pinpoint the concordance between these four assays.

The TAMs include the mutations M41L, D67N, K70R, L210W, T215Y/F and K219Q/E and the association of at least three of them lead to a resistance profile to d-drugs and to AZT by Stanford algorithm. It is why they are interesting mutations to study as part of interactions between resistance associated mutations to observe how the four HIV resistance tests will interpret them. Their interactions will be studied interrogating Stanford, Rega and ANRS algorithms in parallel and compared to the PhenoTecT results done for patient samples carrying an association of these mutations.

IV.2.2. Results

IV.2.2.1. Evolution of TAMs' penalty scores over years

Several combinations of TAMs were interpreted with four different versions of Stanford algorithm to observe evolution over years (Figure 22). I selected versions which showed highest update on penalty scores of TAMs: version 3.3 in 2003, April 25; version 3.9 in 2004, March 25; version 4.0 in 2004, October 27 and version 4.1.7 in 2006, January 20. The scoring was calculated for the resistance to ddI, presenting the strongest evolution.

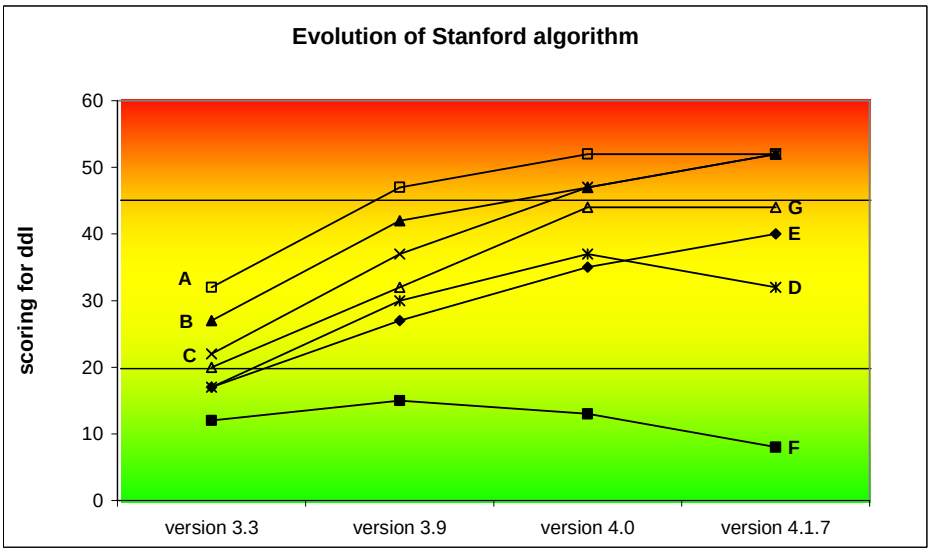


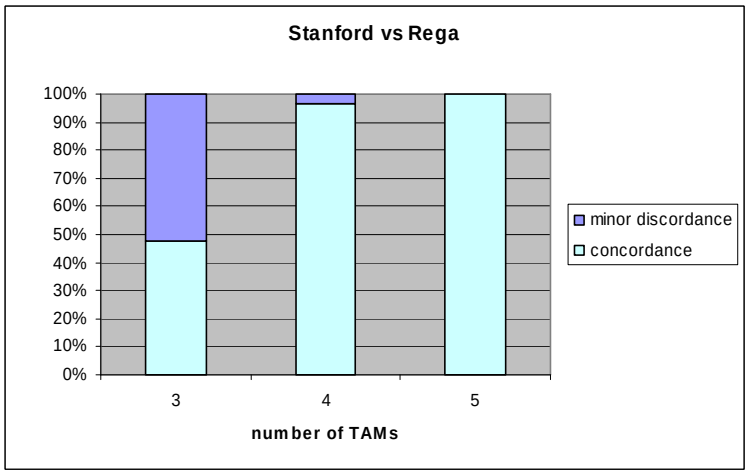
Figure 22. *Evolution of the Stanford algorithm for ddI over years for seven mutated viruses. Virus A with six mutations: 41L-67N-70R-210W-215F-219Q; Virus B with five mutations: 41L-67N-70R-210W-215F; Virus C with four mutations: 41L-67N-210W-215F; Virus D with four mutations: 41L-67N-210W-219Q; Virus E with three mutations: 41L-67N-215F; Virus F with three mutations: 67N-70R-219Q; Virus G with three mutations: 41L-210W-215Y. Scorings below threshold line at 20 are interpreted susceptible (green surface), intermediate between 20 and 45 (yellow surface) and resistant above threshold line at 45 (red surface).*

The impact on resistance level is not similar for every TAM with score ranges between 0 and 20, for 219Q and 215Y/F respectively. Virus F, a triple mutant, is interpreted susceptible to ddI with the four versions whereas virus G, also with three mutations, is susceptible to ddI with version 3.3, intermediate with versions 3.9 and 4.0, and close to resistant with the last version 4.1.7. Viruses A, B and C evolve from intermediate to resistant to ddI, and viruses D and E from susceptible to intermediate. For virus D the penalty scores had even decreased between the two last versions whereas all others were regularly increased.

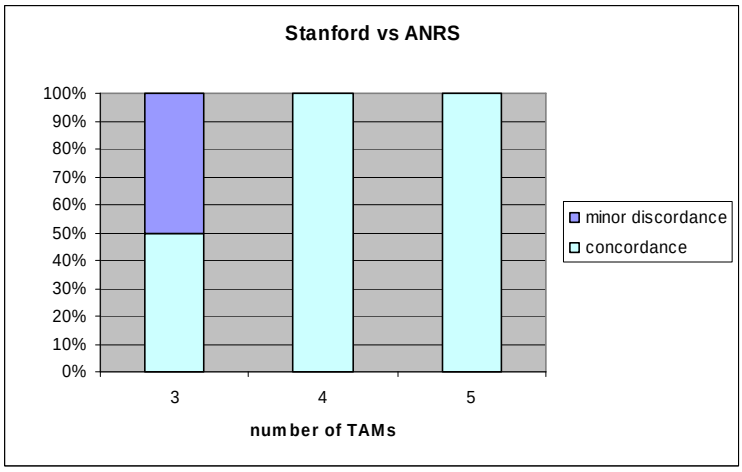
The same update of penalty scores did not have a significant effect on interpretation of AZT resistances. It could therefore not explain the Geno-R/Pheno-S profiles observed for AZT. For ddI the constant increase of penalty scores for TAMs would even raise the percentage of discrepancies Geno-R/Pheno-S since some concordant patients diagnosed susceptible to ddI in 2002 and 2003 would be discordant in 2007.

IV.2.2.2. Interpretation of TAMs through three algorithms

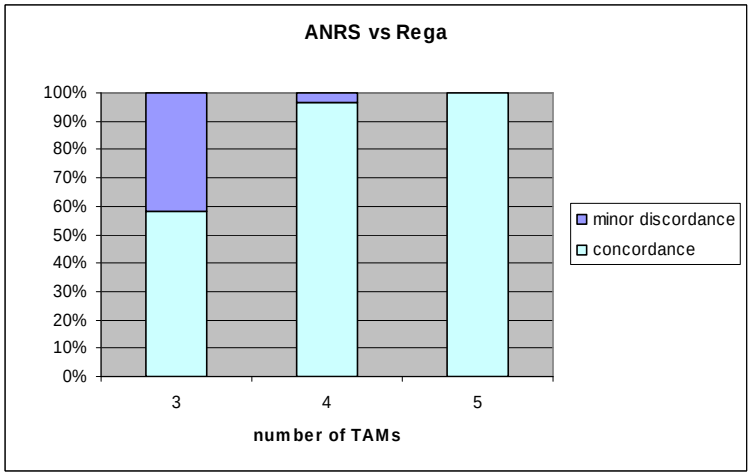
I selected samples in PhenoBase® from the years 2005 to 2007 which have at least three mutations among the main characterized TAMs (M41L, D67N, K70R, L210W, T215Y/F and K219Q/E). I re-interrogated three common algorithms used in diagnostics (ANRS, Rega and Stanford) with the sequence of each sample. The interpretation of resistance was done with the last updated version of each: Rega version 7.1 from March 2007, ANRS from July 2006 and Stanford version 4.2.6 from December 2006. I compared the profiles of resistance for AZT, ABC, ddI and d4T, and the results obtained were grouped in three classes: concordance, major discordances (Geno-R/Pheno-S or Geno-S/Pheno-R) and minor discordances (Geno-I/Pheno-S or -R and Geno-S or -R/Pheno-I). Only AZT and ddI histograms are depicted as representative examples of NRTIs.



A



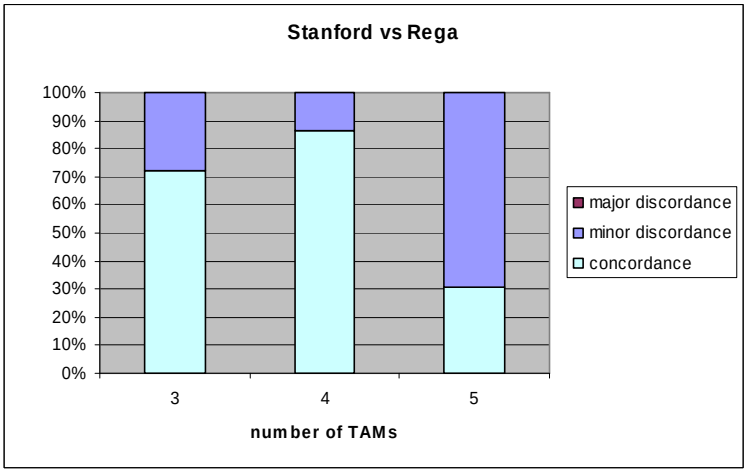
B



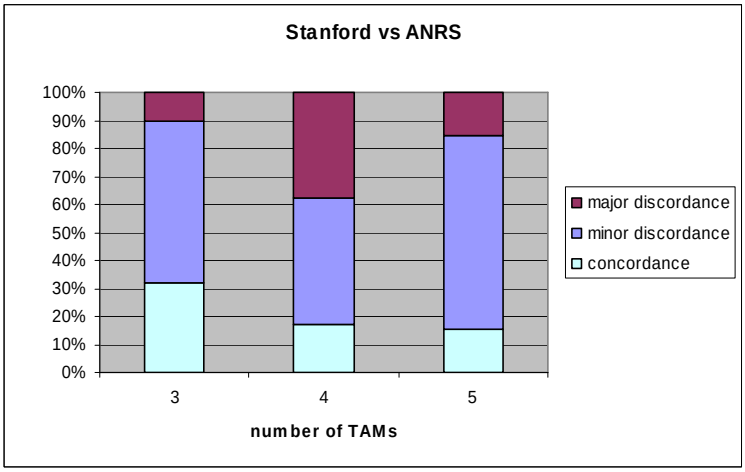
C

Figures 23 A-C. *AZT resistance predictions of different algorithms (N = 92).*

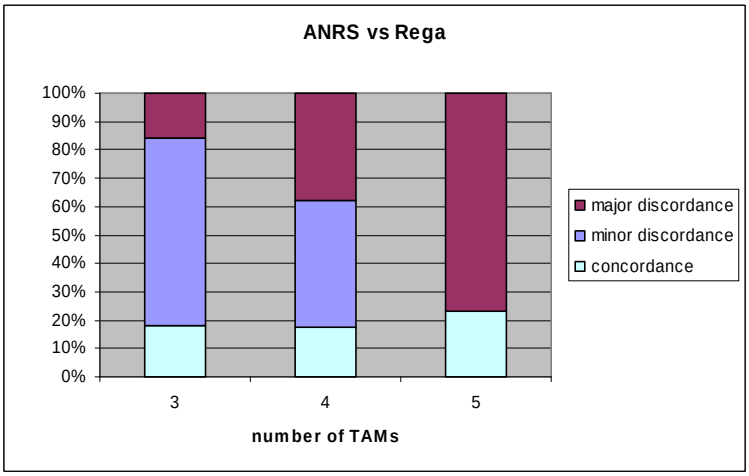
Whatever the two genotypic algorithms compared for AZT resistance, we observed perfect concordance for samples containing at least five TAMs and more than 95% for those with four TAMs. Around 40 to 50% of discordances were seen for samples with three TAMs comparing each prediction two by two. For none of the comparisons major discrepancies were identified.



A



B



C

Figures 24 A-C. *ddI resistance predictions of different algorithms (N = 92).*

Like for AZT, the comparison between Stanford and Rega indicated any major discordances and the level of concordances was still high: around 80% (except for the samples with 5 TAMs, 30%). However, when ANRS was compared to one of the others, the maximal concordance was around 30%. Between ANRS and Stanford close to 40% of major discordances were observed for samples with four TAMs and 10% to 15% for those with three and five TAMs, respectively. The comparison of ANRS versus Rega showed also a high percentage of major discordances: 15% when samples were carrying three TAMs, 38% with four TAMs and 76% with five TAMs.

For AZT, the three algorithms predicted the same impact of TAMs on resistance as the percentage of concordance was close to 100% (except for three TAMs). The minor discordances were all Intermediate/Resistant indicating anyway non-susceptible samples.

In case of ddI they were less in agreement presenting many cases of major discrepancies which were all due to a resistance profile done by Stanford or Rega but a susceptibility to ddI through ANRS. This could be due to the fact that ANRS algorithm interprets the resistance according to the interactions between the panels of mutations. It would mean either that Stanford and Rega algorithms overscored the TAMs comparing to ANRS, or that additional mutations carried by samples with TAMs were negatively scored by ANRS, and not by the two others algorithms, leading to a lower penalty score.

IV.2.2.3. Genotyping and rPhenotyping profiles for associations of TAMs

The previous data showed the importance of the choice of algorithm to interpret the TAMs' resistance to ddI (for AZT, genotypic profiles were similar for the three algorithms). Indeed, the same pattern of mutations could be interpreted susceptible to ddI by ANRS whereas Rega and Stanford predicted a resistance to this drug. For the samples previously selected *(§ IV.2.2.2*), a comparison of every genotypic interpretation with phenotypic results was performed to see if two methodologies were more concordant. All the discordances observed between the four methods were reported as a radar graph for easy visualisation (Figure 25). The scale of each radar represents the percentage of discordance observed between two tests (0, 20, 40, 60, 80 and 100%). From left to right the graphs correspond to samples carrying three, four or five TAMs.

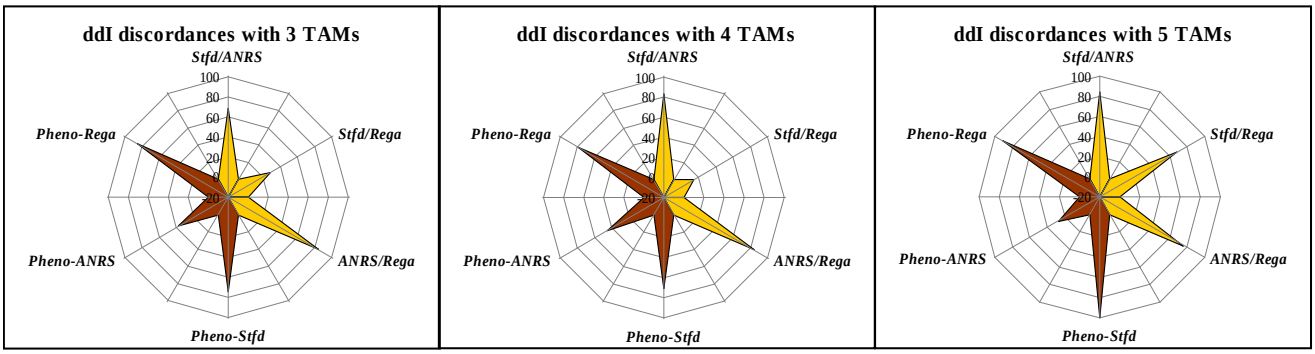


Figure 25. *Representation in "radars" of the percentage of discordance between ANRS, Rega and Stanford (Stfd) Genotyping and rPhenotyping for the NRTI ddI.*

In orange are depicted the discordance levels between genotyping algorithms and in brown the rPhenotyping/Genotyping discrepancies. Whatever the number of mutations, the highest percentage of discordance was noticed when rPhenotyping was compared to Stanford and Rega interpretations (around 80%) and similar results were observed between ANRS and Rega or ANRS and Stanford. Except for a virus with five TAMs, Stanford and Rega algorithms were in agreement for 72% to 87% of the analysis. The best correlation was shown between rPhenotyping and ANRS with 60% to 70% concordance.

For clarity, the "tetrahedron display" allows to place all genotyping algorithms as one dimension (grey triangle) and the independent Phenotyping system as a third dimension (z axis). The above results are shown in Figure 26 with a tetrahedron model.

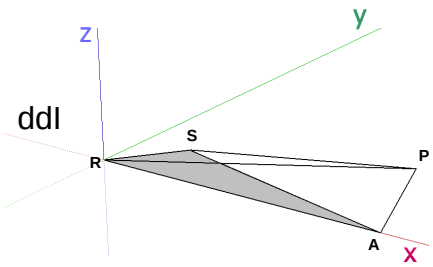


Figure 26. *Tetrahedron model of genotypic and phenotypic discordances shown with ddI. Small distance between Rega (R) and Stanford (S) shows a good concordance in opposite to far distances observed with ANRS (A). The tetrahedron summit corresponding to rPhenotyping is shift to ANRS point revealing a good agreement of these resistance tests and a high discordance with Stanford and Rega.*

The ANRS algorithm takes into account the effect of each mutation on the others (antagonistic and synergistic interaction effects) which could be one of the explanations of the high degree of discordance with the two other genotypic interpretations and the great concordance with rPhenotyping. In *chapter I*, we saw that ddI and AZT were the two NRTIs with the highest percentage of discordance corresponding for 40% of them to major discordance Geno-R/Pheno-S. Therefore, the representation in the "radar graph" was done for AZT data (Figure 27).

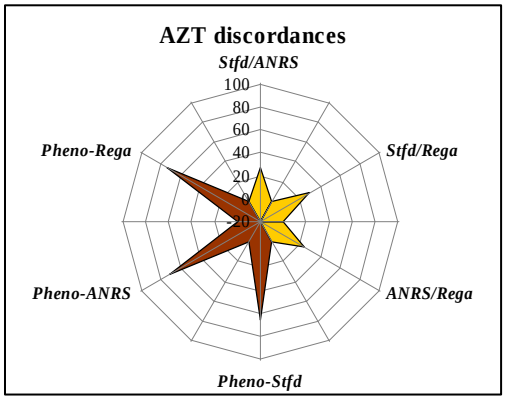


Figure 27. *Representation in "radar" of the discordance percentage between ANRS, Rega and Stanford (Stfd) Genotyping and rPhenotyping for the NRTI AZT.*

The data of the graph correspond to AZT resistance of all the samples presenting three, four or five TAMs. The level of discordance observed between rPhenotyping and Stanford Genotyping (60%) was similar comparing PhenoTecT results with either ANRS or Rega interpretations. The three algorithms used are in good agreement (more than 70%) and the discordance observed with rPhenotyping was equal whatever algorithm is compared.

IV.3. Discussion chapter II

Drug resistance testing is now a fundamental tool regularly used for treatment of HIV-infected patients and approved by regulatory bodies in many countries. Actually, Genotyping and Phenotyping are the two HIV resistance tests available in Switzerland. However, discordances between diagnostic methodologies are not infrequent, and testing HIV-resistances to specific therapeutics does not derogate from this rule. Even using the same initial methodology e.g. obtaining viral sequences in a clinical sample may lead to conflicting diagnostic recommendations depending on the algorithm used for interpretation[92,97,98]. Therefore, it was the major aim of this study to analyse comprehensively the discrepancies observed between genotypic and phenotypic analysis. In this chapter, more particularly, the interest was on samples for which Genotyping predicts resistance to the drug whereas rPhenotyping did not. The results shown in this chapter are multiple and have to be discussed in connection with new data recently published on HIV resistance.

The first working hypothesis to explain discordances Geno-R/Pheno-S – mainly observed for the NRTIs AZT, ddI, d4T and ABC – was the presence of viral mixtures. Several mutants were engineered with one to four mutations, and they were individually diagnosed by the Stanford algorithm and rPhenotyping as well as mixes of some of them. The mutations were selected to give a penalty score to AZT and d-drugs and, except for T69D, they belonged to the group of TAMs. The results showed easily the concordance between both methodologies for a virus carrying three or four of the selected mutations *in cis*. However, for a mix of single or double mutants the concordance disappears. These heterogeneous viral populations were misinterpreted resistant to AZT by Genotyping which principally adds the penalty scores of each mutation together without taking into account their presence on different HIV genomes. In the opposite, the PhenoTecT assay, which allows to measure ARV susceptibility of each individual virus species, diagnosed them susceptible. Viral mixtures could be one of the explanations to the discrepancies Geno-R/Pheno-S. However, on a small subset of "obvious samples" – carrying at least nine characterized mutations – it could not be verified. Indeed, in these discordant samples, one or more viral strains carrying enough mutations to be genotypically resistant to AZT were identified and their phenotypic analysis was still showing a susceptibility to the drug.

The selection of mutant plasmids to engineer was done prior to the availability of PhenoBase®, and I chose the most represented mutations in the samples Geno-R/Pheno-S for AZT and d-drugs. I did not consider how they were associated. In the meantime several

studies suggested the existence of two distinct clusters for the TAMs[99-102]. Mutations that occur together with T215Y, including M41L, L210W and sometimes D67N, constitute the TAM-1 cluster; mutations that occur together with K70R, including D67N, T215F and K219Q, constitute the TAM-2 cluster. The constructs done in this study turned out not to correspond with those clusters. It had been identified retrospectively by the interrogation of PhenoBase®. In fact, only few samples were found to present similar patterns but every time in association with additional TAMs. Unfortunately, our bank contained no single virus similar to those engineered. A phylogenetic tree was designed, based on the mutation patterns of the samples in PhenoBase®, to observe the pathway of the emergence of resistance associated mutations. Looking only to the emergence of TAMs, Figure 28 represents the tree starting with the mutation T215Y.

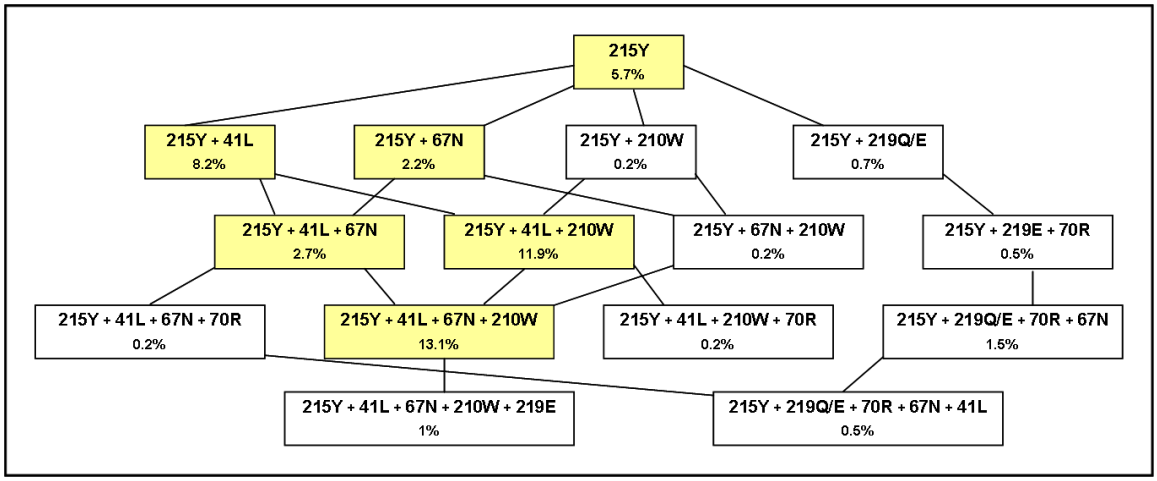


Figure 28. *Phylogenetic profile of TAMs observed in patients diagnosed between the years 2002 and 2007. The percentages are calculated comparing the number of patients carrying the indicated mutations and the total number of patients having at least one TAM.*

This model illustrates the order of appearance of mutations associated with AZT resistance in the HIV-1 RT gene during therapy of asymptomatic individuals. A pathway appears from the repartition of TAMs in diagnosed patients. In yellow are figured out the most frequent combinations of TAMs found in PhenoBase®: mutation T215Y is often followed by M41L and subsequently by L210W and D67N (corresponding to the highest values). This outcome corresponds to the TAM-1 cluster described previously. The quadruple mutant T215Y-M41L-D67N-L210W had emerged in 53 samples.

It would have been more appropriate to engineer mutant plasmids corresponding to such a pathway to validate the viral mixture hypothesis. Nevertheless the discordance of resistance profiles observed between Genotyping and rPhenotyping could not be explained by the cluster of TAMs.

The next hypothesis that could explain the different interpretations of resistance is an overscoring of TAMs by genotypic algorithms. Indeed, the Stanford algorithm updated its penalty scores for PR and/or RT mutations twenty three times between the years 2003 and 2007 based on new clinical studies. Some initially non-characterized mutations appeared to be involved in resistance mechanism and some known mutations were also updated with an increase or decrease of their penalty scores. To see the impact of these changes over years, several combinations of TAMs were interpreted with different versions of the Stanford algorithm for ddI and AZT resistance. Whereas the updates modified the interpretation of resistance to ddI, no such evolution was observed for AZT resistance. Moreover the comparison of three algorithms – Stanford, ANRS and Rega – to analyse similar sequences reflects the complexity of providing a unique rendering. It further shows that related methodologies may yield a discordant profile of resistance as shown for ddI but not for AZT. The main difference observed was ANRS susceptibility versus Rega or Stanford resistance to ddI. The fact that ANRS algorithm considers the interactions between mutations with either synergistic or antagonistic effects could be one explanation.

Recently, new studies showed that other resistance associated mutations could suppress the AZT resistance (Review articles Larder *et al.*, 1994, Sluis-Cremer *et al.*, 2000 and Goldschmidt *et al.*, 2004[103-105]). The mutation L74V alone confers ddI resistance but, when present with T215Y, it completely suppressed AZT resistance by reducing the extent of its removal by ATP-dependant phosphorolysis[106-108]. The NNRTI resistance mutations Y181C and L100I are also responsible of re-sensitisation to AZT when coexpressed with AZT-specific substitutions[109-111]. *In vitro* drug resistance studies have revealed yet two other RT mutations, K65R and M184V, which reverse AZT resistance[44,48,112-114]. One of the patients selected for the dilution cloning, patient 6016764, was carrying several TAMs: K70KR, L210LW, T215Y/F and K219KE. The sequencing of eight clones and a parallel phenotypic assay revealed that the presence of different viral genomes was not the explanation to the discordance Geno-R/Pheno-S observed (at least one clone was carrying all the mutations). However, mutations M184V and Y181C were associated with these TAMs in eight and six clones, respectively. As mentioned above, their presence could decrease the level of resistance provided by the TAMs and lead to the discordant resistance profile Geno-R/Pheno-S. To prove their possible role in the phenotypic susceptibility to AZT, I should revert them into wild type amino acids to observe an eventual increase in resistance.

Just as a re-sensitisation to AZT can occur by interactions between TAMs and other mutations, another combination has been described as a mechanism of synergy to influence AZT resistance: mutations in the C-terminal domain of the RT may significantly enhance

clinical AZT resistance (E312Q, G335C/D, N348I, A360I/V, V365I and A376S). The proviral cassette used to generate a recombinant virus for the PhenoTecT assay does not allow measuring the influence of mutations in this connection domain of the HIV-1 RT since only the N-terminal region of the RT gene (amino acid 1 to 317) is derived from the clinical sample to be tested.

Adapted versions of proviral cassettes should be considered to study by rPhenotyping the role of such interactions on AZT resistance. It could be interesting to reanalyse some discordant samples Geno-R/Pheno-S for AZT with this new version of PhenoTecT system to assess if it will influence the phenotypic result. An increase of the level of resistance would be expected. Besides, using a primer inside the RT gene, as it is in the actual format of the phenotypic assay, could lead to a preferential amplification of viral strains with perfect annealing unlike a virus carrying a mutation in the binding site of the primer. The latter viral genome will need a decreased annealing temperature to be amplified.

In fact, the sample 4007559 contained a viral population carrying seven characterized RT mutations: M41ML, A98AS, K103N, M184MV, L210LW, T215SYT and G333E. The profile of resistance to AZT was discordant: Geno-R/Pheno-S with several double peaks in the electrophoregram indicating a mixed viral population. Fourteen clones were sequenced and they all missed the TAMs and the G333E. Two different viruses had A98S, K103N and M184V; and one of them had the additional mutation M230V, previously not detected in the cDNA sequence. Retrospectively, the recombinant provirus was also sequenced and a similar result was obtained: A98S, K103N, M184V and M230MV. These results show that sequencing of cDNA and plasmid preparation may yield two different genotypic profiles. A likely explanation of this phenomenon was a bias of primer R3525L used to amplify the RT gene. Figure 29 represents the alignment of cDNA sequences from sample 4007559 with the reverse primer R3525L.

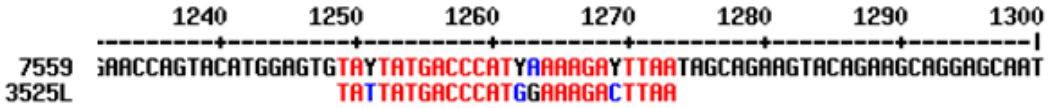


Figure 29. *Alignment, using Multalin[H], of the cDNA sequence from sample 4007559 and the reverse primer R-3525L used for the amplification of RT gene.*

For cloning a primer R3525L was used which carries the restriction site NcoI. The cDNA sequence of the sample showed a double peak C/T (Y) in the region binding the 3' end of the primer. This part of the primer is the more crucial for an efficient annealing. The

population with a cytosine was the most represented in the sample (higher peak) but, because of the mis-annealing with the primer, it will not be amplified. The alignment of the recombinant plasmid sequence with the primer R3525L showed a perfect match (data not shown) indicating that only minor species (with the thymine) could be amplified.

For this clinical sample the genotypic profile was corresponding to a mixed population while the rPhenotyping was based only on a subpopulation selected by the bias of primer. I interrogated PhenoBase® to have an overview of the samples Geno-R/Pheno-S to AZT subject to present a similar result. Out of 75 samples, seven showed a mismatch with primer R-3525L (9.3%). This could have been avoided by using a redesigned primer derived from a more conserved region. It is difficult to define such region as HIV is highly polymorphic (subtypes, …) and rapidly mutates with time in any given patient. This is a real challenge for phenotypic assays.

V. Chapter III: Detection of clinically relevant HIV minority species

V.1. Basis for existing virus mixtures

HIV infection is typically not caused by a single virus species but rather by a flurry of different populations of viruses that evolve independently and can overtime replace each other. In other words, while there may be one dominant strain of HIV, there is potentially an appearing, disappearing and even recombining of countless minority populations waiting for their selective advantage. This also means that HIV, at any time point, might have a library of several genotypes that it can select from and then use, in whatever capacity, for its own evolution. Different events could explain the presence of HIV multispecies in a patient sample:

- Since HIV has a high rate of replication and the reverse transcriptase is not able to proofread nucleotide sequences during replication[115], the genetic variability of HIV in a patient may be very high. During the course of viral replication, several mutations are accumulated leading sometimes to emerging drug-resistant mutations. The latter can replace the prevalent major virus, behaving as a reservoir from where new genotypes can be recruited depending on the pharmacological pressure;

- Moreover, homologous recombination can occur when a cell is coinfected with two different but related strains. Indeed HIV, like all retroviruses, is “diploid” (each viral particle contains two RNA strands of positive polarity, each full length and potentially able to replicate). Typically both RNA strands in a retroviral particle derive from the same parent provirus. However, if an infected cell simultaneously harbours two different proviruses, one RNA transcript from each provirus can be encapsidated into a single "heterozygous" virion. When this virion subsequently infects a new cell, the reverse transcriptase may jump back and forth between the two RNA templates so that the newly synthesised retroviral DNA sequence is recombinant between that of the two parents[116]. All subsequent progeny virions will be of this recombinant genotype. Naturally occurring recombinant HIV strains have been found in infected patients in regions of the world where multiple genotypic variants cocirculate[117];

- Superinfection is defined as the reinfection of an individual who already has an established infection with a heterologous HIV strain. Worldwide, 16 cases of HIV-1 superinfection in humans have been reported since 2002[118]. One of the viral strains could persist at a low level until viral evolution, drug treatment, or host immunity results in a fitness advantage for the minority strain.

Usually resistance development is a slow process, as it can take numerous rounds of replication and competition among the diversified strains to render one variant that has a strong survival advantage compared to other variants in the presence of the antiretroviral. Often the virus has to undergo several genetic changes in a target gene in order to evade drug-pressure[119,120]. But in some cases, full resistance can also be almost instantaneous, achieved through a single amino-acid substitution. This is best exemplified by the Methionine to Valine change at position 184 of the reverse transcriptase gene leading to high-level of resistance (>500 fold) to the two drugs 3TC and FTC[44,114,121,122]; several single mutations at others positions provide full resistance to NNRTIs as K103N/S/T, V106M, Y181C/I/V (only DLV and NVP), Y188L, G190A/C/E/Q/S/V/T, M230L[123-125].

Most *in vitro* studies find the dynamics of viral replication greatly affected by the mutation M184V[126-128], whereas in general the single mutations leading to NNRTIs resistance are not associated with reduced replication capacity[129-131]. Yet clinical and animal data are still somewhat conflicting on the impact on the course of infection *in vivo*, but it may be expected in accord with the "survival of the fittest principle" that drug pressure will drive resistant variants, even when initially present only as small proportions, to outgrow a drug-susceptible wild type. A number of reports[132-136] have mostly retrospectively shown this for clinical situations, where individual mutations could be associated with loss of drug efficacy (such as NNRTI failure following a K103N change in RT). Treatment failure to an inhibitor with the necessarily preceding presence of low percentages of the mutant virus was taken as indication for retrospectively searching for the respective mutation in plasma samples already prior to overt clinically therapy failure. At present such minority species cannot be detected early enough through conventional techniques such as single-cycle Phenotyping or population-sequencing analysis (detection limits around 15%)[137,138]. A recent report then was able to demonstrate direct detection of a resistant virus by real-time PCR of the protease mutant D30N in clinical specimens that had remained undetectable by Genotyping[139]. A highly sensitive detection e.g. by allele-specific PCR is therefore principally applicable but the essential design of the discriminating primer requires precise knowledge of the mutation to be found. Yet the large number of possible relevant mutations in the target genes of interest, PR, RT, and others to come, renders the methodology out of reach for routine diagnostics. Nevertheless it would be desirable for optimal clinical management of therapy resistance and for most effective treatment to overcome these limitations possibly reflected also in discrepancies between certain genotypic profiles (i.e. apparent absence of mutation) and accompanying clinical course (i.e. early viral escape and eventually treatment failure). As alternative, a replicative Phenotyping provides drug monitoring through several cycles of

viral replication, which mimics the dynamics of the corresponding virus inside the patient. Thereby it ought to be able to reveal the presence of minority resistant species outgrowing drug pressure. Indeed, the format of the PhenoTecT system allows each single virus to replicate separately and to infect new cells. Whereas the major population, susceptible to the antiretroviral agent, will stop replicating, the drug- resistant minor variant will continue to multiply. Even with high concentration of drug a viral population will still be detectable after four days of replication and interpreted as resistant virus species. Suspecting a resistant minority species, the replicative phenotypic determination of resistance would be discordant to the genotypic prediction of susceptibility: Geno-S/Pheno-R (Figure 30).

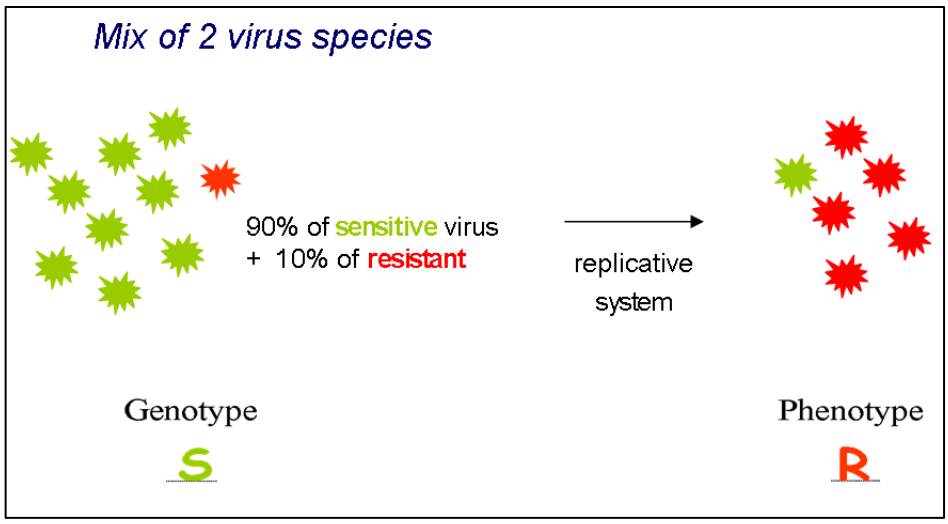


Figure 30. *Schema illustrating a viral mixture with ARV-resistant minority species.*

The primary objective of this chapter is to demonstrate the properties of rPhenotyping as “sensitive direct methodology for the detection of viral minorities in clinical samples” and, if so, to confirm the result by profiling the same original samples with most sensitive adaptations of genotyping methods such as allele-specific PCR and real-time PCR. First, distinct samples will be selected in PhenoBase® for which no mutation at position 184 of RT gene was detected following population-based sequencing despite the phenotypic assay indicated a resistance to 3TC. In a second step, allele-specific PCR methods and real-time PCR will be performed on these specific samples to identify and quantify 184V HIV variants, in order to confirm the phenotypic outcome.

V.2. Results chapter III

V.2.1. Sensitivity of the replicative format of PhenoTecT

The limit of detection of the rPhenotyping system was first determined for the fitness-affecting mutant 184V (in parallel the single mutant 103N was also tested, data not shown). A wild type reference plasmid was point-mutagenised, and defined mixes of wild type and mutated plasmids were evaluated for their sensitivity against the study drug 3TC and the unrelated RT inhibitor AZT as control as its susceptibility should remain unaffected by the mutation M184V; SQV as control-PI should also retain full drug activity on both isogenic viruses. Mixes were denoted according to the respective proportion of mutated virus they contained: 0% (pure pNL4-3 reference), 0.1%, 0.3%, 1%, 3%, …, and 100% (solely the 184V-mutant). Following four-days of viral replication, the reporter gene activity (β-Gal) was determined and expressed as percent of viral production. Figure 31 shows inhibition curves with the percent inhibition of viral replication plotted as a function of drug concentration. Results were confirmed in three independent experiments.

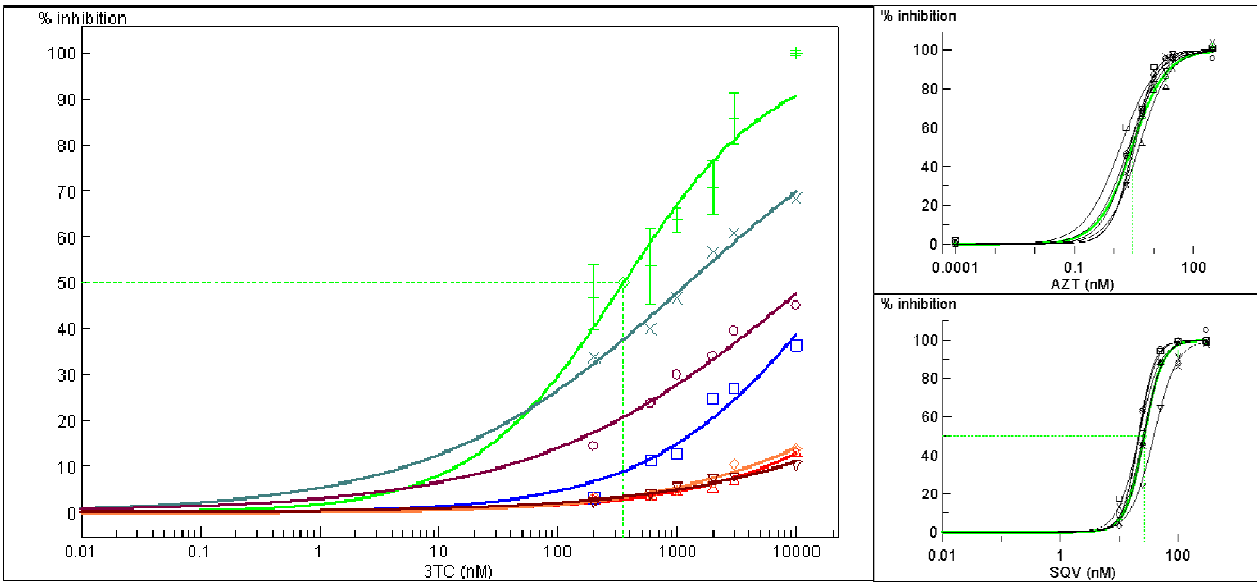


Figure 31. *Drug susceptibility of mixed populations of replicating HIV-1.*
The main graph shows the inhibition curves of viral replication for wild type alone (light green), 1% (dark green), 3% (purple), 10% (blue), 30% (orange), 50% (red) and 100% of 184V (brown). The reference IC_{50} is indicated as green dashed line. The corresponding inhibition curves in the presence of AZT (top) or SQV (bottom) are shown on the right (wild type curve in light green).

Inhibition curves demonstrated that, whereas for the non-respective control drugs AZT and SQV increasing amounts of mutant 184V had no impact on the inhibition (graphs "AZT", "SQV" on the right), the 3TC inhibition curve was significantly altered as soon as ≥ 1% of the viral population was contributed by the 3TC-resistant mutant. The IC_{50} values are summarised in Table 11 with the Rf as the ratio between the IC_{50} of the sample and the one of the wild type.

	Wild type	**1%**	**3%**	**10%**	**30%**	**50%**	**100%**
IC_{50} (nM)	354	1240	12'732	22'576	498'790	478'915	485'971
Rf	1	3.5	36.0	63.8	1'408.9	1'352.7	1'372.7

Table 11. *IC_{50} values and Rfs for the wild type, the single mutant 184V and mixed populations. "x%": percentage of mutant 184V.*

This finding is pinpointed in the representation shown in Figure 32 with the comparison between mixes re-calculated a Rf.

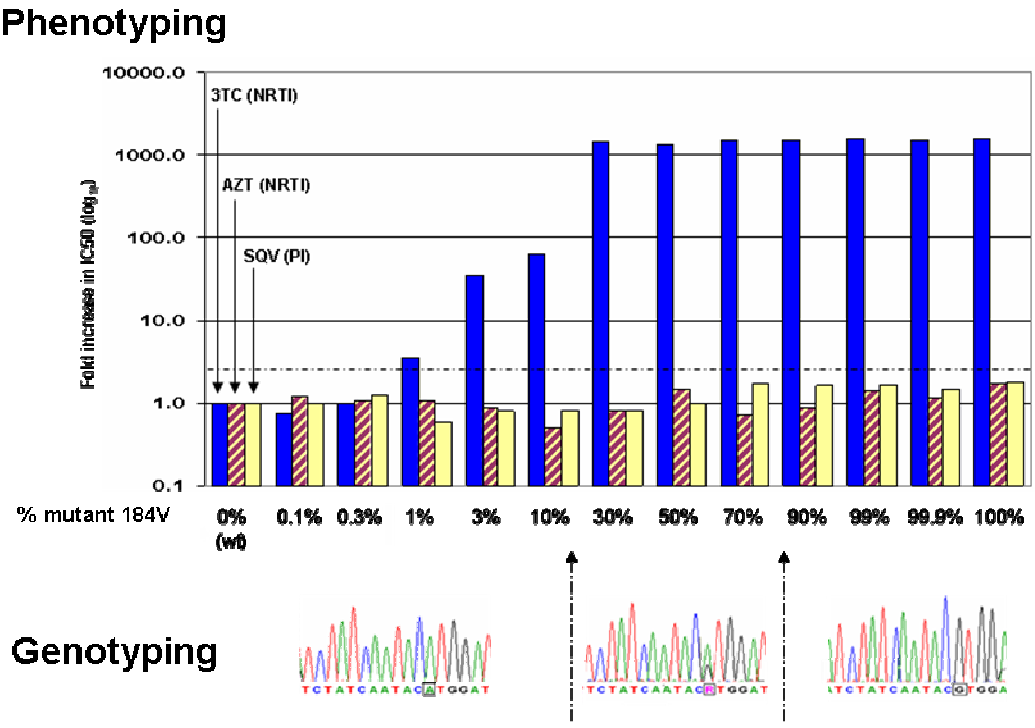


Figure 32. *Sensitivity of PhenoTecT.*
Rfs for wild type (0%) and mixes with 184V ("x%" as indicated) are plotted for 3TC (blue bars), and two non-affected controls, for RTI (AZT, hatched bars) or PI (SQV, yellow bars).

The link of the resistance factor to an established clinically diagnosed resistance allowed to define a threshold ("cut-off") value for this Rf, above which a virus is statistically considered non-susceptible. The calculated cut-off for 3TC is 3.0 (Figure 32, dashed line), indicating that in the PhenoTecT system already a proportion of 1% of mutant 184V in the mixed virus population yielded a relevant "decrease in susceptibility" to 3TC (intermediate). At low percentages of mutant the histograms of Figure 32 follow a dose-response trend as increasing proportions of mutant in the defined mixes result in increasing values for the calculated Rfs. This value reaches the plateau of full resistance once 30% of the viral population is contributed by the resistant variant. Of note, the resistance factors for the control drugs AZT and SQV remained unchanged across the entire range of mixtures. Of concern was a potential bias arising from possible effects of an altered replication capacity of the virus; however, this could be excluded as in the assay format of PhenoTecT a reduced replication (lower viral fitness) solely results in a lower absolute optical density but not in a shift in the IC_{50} as parameter of inhibition (in PhenoTecT the absolute readout for 184V was 80% of the wild type).

The identical mixes were in parallel subjected to genotypic assessment. The presence of the mutation M184V as a significant peak was confirmed unequivocally only when the mutated provirus represented $\geq$ 30% of the mix (not shown).

V.2.2. Limit of detection of minorities by allele-specific PCR

I next set out to define the technical lower limit of detection for a mutant virus within mixes. Interrogating PhenoBase® for the samples Geno-resistant to 3TC, the resulting list of mutations indicated that the sole change responsible of this resistance was mutation M184V in the RT gene (in 100% of the samples resistant to 3TC). The analysis of the sequences showed that, out of 291 patients, 89% had a Valine at position 184 in the RT gene, which was encoded by a GTG (instead of the Methionine ATG); 3.1% had a RTG meaning a mix of wild type ATG and mutant GTG. All together 92.1% of the patients presented a viral population resistant to 3TC due to a GTG codon at position 184. The others were subdivided into two categories: 6.9% had a GTA codon and 1% a GTR. Based on these data, the detection of mutant resistant to 3TC was attempted for those having a GTG codon at position 184 of the RT.

An allele-specific PCR methodology was adapted. Two oligonucleotides (R-RT184M and R-RT184V) were designed to carry at their respective 3' termini a nucleotide specific for one or the other variant to be discriminated.

- *R-RT184M:* $^{5'}$*cagatcctacatacaaatcatcc*$\underset{184M}{\underline{cat}}$ $^{3'}$ *specific to the wild type* 

- *R-RT184V:* $^{5'}$*cagatcctacatacaaatcatcc*$\underset{184V}{\underline{ca\color{red}{c}}}$ $^{3'}$ *specific to the single mutant 184V* 

The PCR was done on the wild type and mutated plasmids with a common forward primer D-2213B and either one of the two reverse primers separately. The melting temperature of these primers being approximately 59°C, it was recommended to test different annealing temperatures close to this value in order to maximise the specificity of amplification. Three different annealing temperatures were tested: 55°C, 58°C and 61°C. For each annealing temperature four PCR reactions were performed: wild type plasmid (green virus form) with either "wild type" primer R-RT184M (green dot) or "mutated" primer R-RT184V (red dot) and 184V mutant plasmid (red virus form) with both primers as well. Figure 33 shows an amplification product for each condition. It confirms like in the literature that one unique mismatch at the 3' end of the reverse primer is not sufficient to discriminate unspecific from specific amplification even with increased annealing temperature.

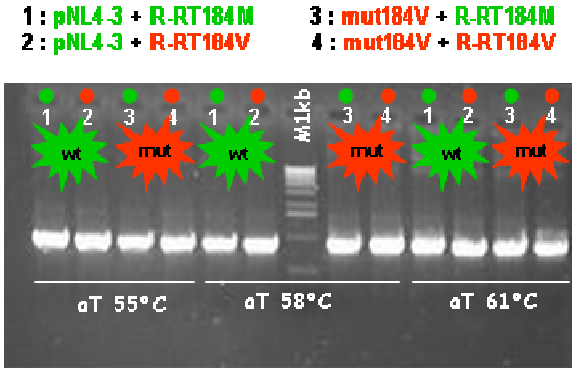


Figure 33. *Allele-specific PCR optimisation with reverse primers mutated at the 3' end. In green are represented the template and the primer wild type, in red the template and the mutated primer. aT: annealing temperature; M1kb: Size Ladder.*

In an optimised protocol, the wild type template should be amplified only with the R-RT184M primer whereas the R-RT184V primer should only amplify the 184V mutant. Korn *et al.* had described that the addition of further non-critical mismatches greatly contributes to PCR discrimination between the two specific sequences[89]. The two primer sets of this study

were further optimised by the addition of three internal mismatches to selectively amplify either one of the fragments containing 184M or 184V, respectively.

- *R-RT184M3:* $^{5'}$*cagatcgtacatatgaatcatccat* $^{3'}$ *specific to the wild type* 

184M

- *R-RT184V3:* $^{5'}$*cagatcgtacatatgaatcatccac* $^{3'}$ *specific to the single mutant 184V* 

184V

The allele-specific PCR reaction was repeated with the new sets of primers under the same conditions as the previous (as annealing temperature 65°C was implemented). Figure 34 shows the amplification products obtained at 62°C and 65°C.

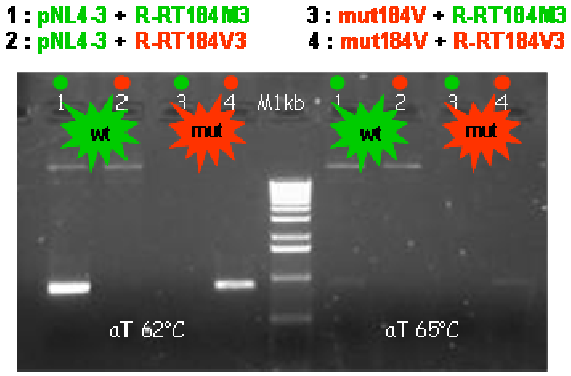


Figure 34. *Allele-specific PCR optimisation with reverse primers mutated at the 3' end and carrying three internal mismatches. In green are represented the template and primer wild type and in red the template and primer mutated.*
aT: annealing temperature; M1kb: Size Ladder.

The addition of three mutations in the middle of the primer sequences destabilises hybridisation and increases the binding specificity. At 62°C it was possible to specifically amplify the wild type template with the primer R-RT184M3 and not with the R-RT184V3, the opposite was true for the mutated plasmid. The absence of amplification of wild type with mutant primer or mutant with wild type primer served as validation.

As starting material for the determination of the limit of detection, the defined proviral mixes (see *part V.2.1.*) were used in the conditions described above. Figure 35 shows the amplification products, demonstrating that the primer specifically designed for solely amplifying the mutation M184V allows its detection down to 0.3% in a mixture with wild type. The corollary was similarly valid for the sensitivity of wild type detection by the specific reciprocal primer.

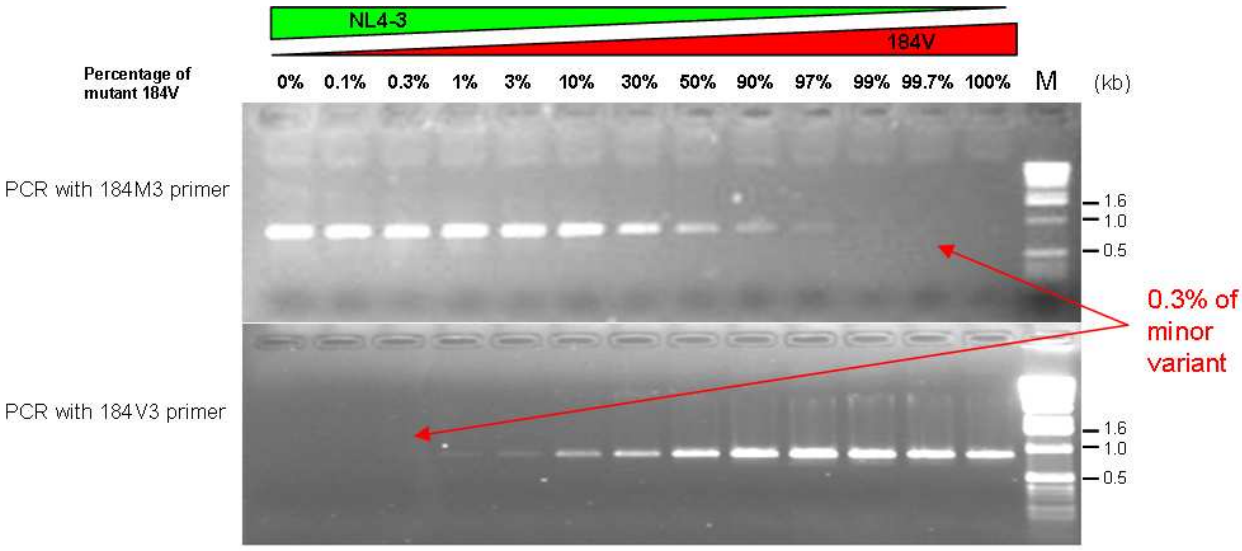


Figure 35. *Discrimination power of mutation-specific PCR.*
Aliquots of precise in vitro mixes of wild type and mutant 184V as indicated in the above graph were amplified using a primer set for wild type (184M3, top graph) or mutant (184V3, bottom graph) and analysed by gel electrophoresis. M: Size Ladder (1kb).

Subsequently a series of patient samples identified by their discordances of 3TC resistance interpretations between Genotyping and rPhenotyping were tested under the same conditions. The underlying hypothesis was that they might contain minor variants within a dominant wild type sequence. In the more recent samples from the year 2005, 18 were diagnosed Geno-S/Pheno-I or -R. I selected seven of them (A to G) as representatives of all relevant subtypes, fitness and 3TC treatment history. In the aforementioned group of 18 samples, 39% were B subtype, 22% AE subtype, 17% AG subtype and 11% F subtype. I chose three samples with B subtype virus (A, D and G), two AE subtype (C and F), one AG subtype (B) and one F subtype (E). By rPhenotyping, their viral fitness had been estimated either low (e.g. 23% for sample C) or high (e.g. 100% for samples B and G) compared to the reference. These samples were all corresponding to patients in therapy failure which were either currently under 3TC treatment, previously treated with this drug, or never treated.

The different parameters are resumed in Table 12.

patient	VL (copies/mL)	CD4 (cells/mm^3)	3TC treatment
A	100'500	42	current
B	16'900	800	no information
C	69'854	nd	current
D	690'000	61	current
E	>100'000	195	never
F	2'461	nd	current
G	30'585	nd	in past [a]

Table 12. *Clinical parameters of the seven patients in this study.*
VL, Virus Load; nd, not determined. [a] "in past" = previous treatment, not currently used.

Figure 36 clearly demonstrates for the original viral populations isolated from each one of five discordant patient samples that the expected mutant fragment could be selectively amplified with the mutant-specific primer whereas the amplification reaction using the wild type control plasmid stayed negative. In contrast, initial Genotyping had given no evidence for the alteration.

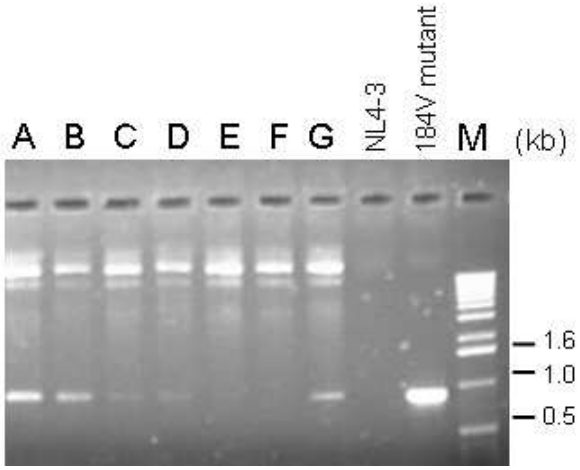


Figure 36. *PCR quantification of virus from discordant patient plasmas.*
The PCR products from seven patient-derived viruses were analysed by electrophoresis following amplification with the 184V3-mutant primer set (lanes A-G). "NL4-3" and "184V mutant" are included as respective control lanes. M: Size Ladder (1kb).

This strongly suggested that in the authentic clinical plasma samples (A, B, C, D and G) indeed the HIV-1 mutant 184V had already been established. Further support for the relevance of this amplification came from applying the same PCR amplification protocol to three randomly selected specimens from patients diagnosed as 3TC-susceptible by Genotype and PhenoTecT. In all cases the protocol failed to produce a mutant-specific band on these RNAs (data not shown).

V.2.3. Quantitative detection of minority populations with mutated sequences in clinical samples

To quantitatively estimate the proportion of 184V variant in patient blood, I developed a real-time PCR protocol distinguishing between wild type and 184V-mutant. A selective Sybrgreen assay had to be designed for overall quantification. A set of primers was utilised, for which the ΔCt between specific and non-specific annealing should be maximal (e.g. for mutant-specific primer: between 184M and 184V templates). The first step was the assay development because no assay conditions were established for the machine available in the lab (Corbett instrument: Rotor-Gene™ 3000). I had to optimise different parameters on known templates (NL4-3 and single mutant 184V) before adapting the tool for the discrepant samples Geno-S/Pheno-R to 3TC. The DNA used as a template of qPCR was a PCR product obtained from the amplification of the wild type and mutated plasmids with the primers D-1818 and R-3584. A purification step was done prior to the real-time PCR either by gel extraction with Macherey-Nagel columns or directly on Microcon YM-100 Centrifugal Filters.

- Design of the specific primer:

Based on the allele-specific PCR results, I started to test the two sets of primers carrying a mismatch at 3' end and those with the addition of three internal mutations. A forward primer F-3005 was designed in a conserved region at a distance of one hundred nucleotides upstream of the position 184. Figure 37 shows at the top the amplification of wild type (1) and 184V mutant (2) templates using the wild type primers R-RT184M (a) and R-RT184M3 (b); and at the bottom the same templates amplified with the mutant primers R-RT184V (c) and R-RT184V3 (d).

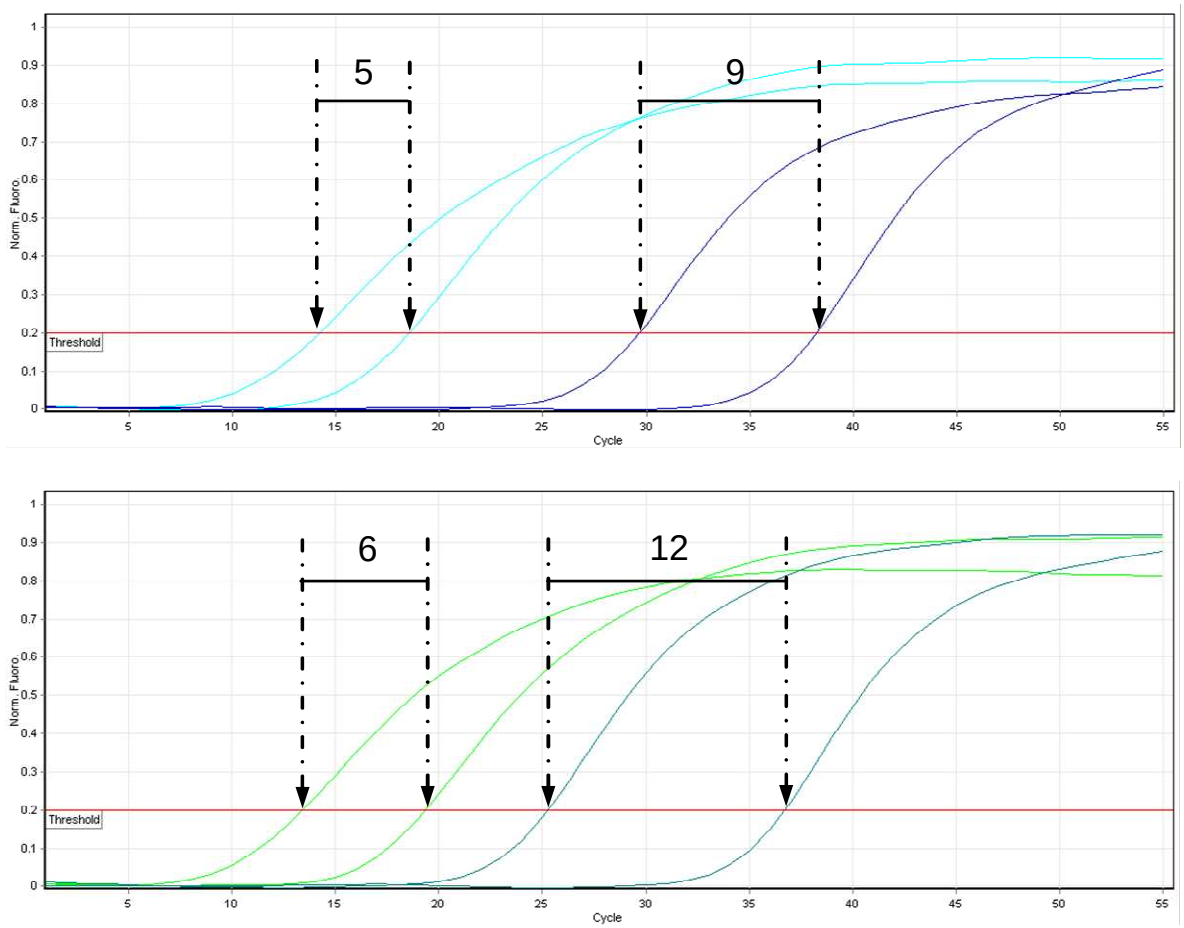


Figure 37. *qPCR optimisation with diverse couples of primers where the fluorescence is a function of the number of cycles. On top are represented the amplification curves using the specific primer 184M versus 184M3 on wild type and mutant templates; on the bottom are the curves corresponding to the specific primer 184V versus 184V3 on mutant and wild type templates.*

The top graph illustrates the discrimination observed between specific and unspecific amplification with the wild type primers. A threshold was fixed at 0.2 units of fluorescence. The first curve to appear was corresponding to the amplification of the wild type template with the primer R-RT184M and it crossed the threshold line at 14 cycles (Ct = 14) (dashed arrow). The next curve represented the amplification of mutant template with R-RT184M at Ct = 18. The difference between authentic and inauthentic priming is calculated as the distance between the Ct observed, called ΔCt, and is reflecting the discriminatory ability. Using primer R-RT184M the ΔCt value was equal to 4. The introduction of three internal mutations in this primer, R-RT184M3, destabilised binding, leading to a right shift of the curve (Ct of the wild type template = 30, instead of 14). This modification increased also the specificity of binding as revealed by the higher ΔCt equal to 8 (30→38). Similar results were

observed with the mutated primers R-RT184V and R-RT184V3. The addition of internal non-critical mutations increased the number of threshold cycles (ΔCt) from 6 to 12.

The significantly improved discrimination between the wild type sequence and the mutant 184V by the use of R-RT184V3 was required to easily distinguish low percentage of mutant in wild type background.

- Annealing temperature:

The experimental conditions were optimised and validated using different annealing temperatures with the set of primers F-3005 and R-RT184V3 on the 184V mutant template at five or three concentrations (Figure 38).

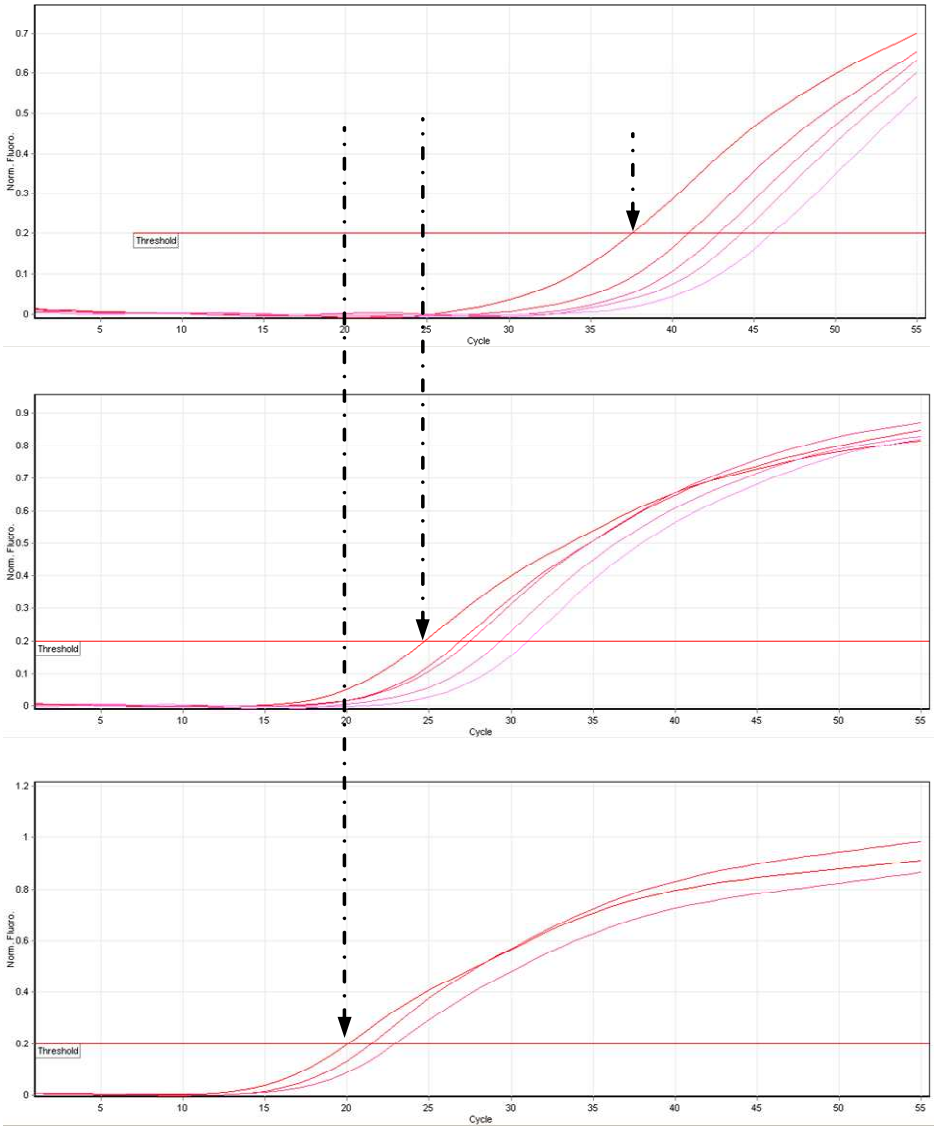


Figure 38. *qPCR optimisation with diverse annealing temperatures: 62°C on top, 60°C in the middle and 57°C on the bottom.*

The absolute amount of template in the reaction ranged from 4000 pg to 250 pg final with 2 fold dilutions. Indifferently for the three conditions, the order of apparition of the curves corresponded to the initial amount of template used (increase of Ct with a decrease of DNA template). Interestingly, the delay before the occurrence of the first curve declined with the decrease of the annealing temperature: the first dilution had a Ct of 37 at 62°C, 25 at 60°C and 20 at 57°C (dashed arrows). The shape of the curves was also better at 57°C. The choice to test 57°C was made following specific advice by the machine manufacturer. The optimal temperature to use should be between 3 and 6 degrees less than the Tm of the primers. Except at 62°C, the highest template concentration produced a curve's shape which was too flat comparing to the others indicating that the amount of starting template was also crucial for an optimal amplification.

- Template quantity:

As shown in the previous part, the amount of DNA template could influence the shape of the amplification curves. Five quantities of 184V mutant were amplified with the R-RT184V3 primer at 57°C (Figure 39).

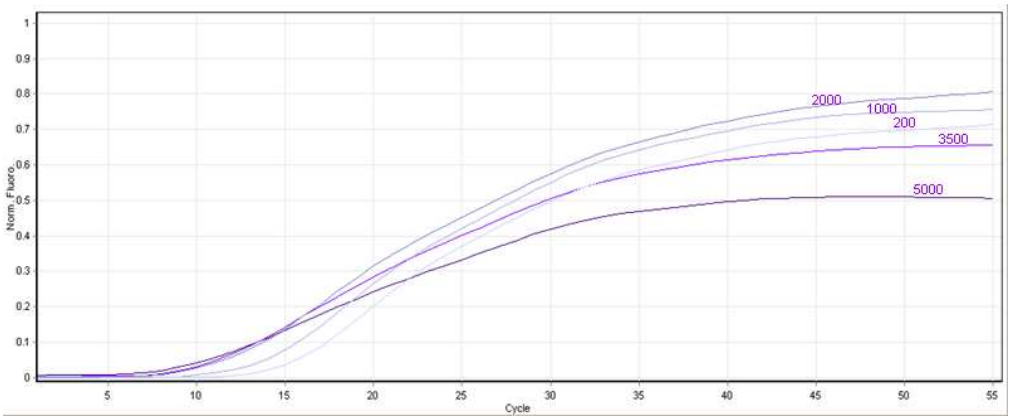


Figure 39. *qPCR optimisation with five different amounts of DNA template (184V mutant): 5000 pg, 3500 pg, 2000 pg, 1000 pg and 200 pg.*

The highest concentrations, 5000 and 3500 pg, generated flat curves whereas between 2000 pg and 200 pg the curves had a similar shape and the ΔCt were reflecting the difference of template quantity.

As fourth parameter the size of the PCR product of the real-time PCR reaction was naturally assessed (data not shown). The results indicated that an optimal result was obtained for PCR product size between 100 and 150 bp. The efficiency of the qPCR was higher when the PCR amplicon had a size of 120 nucleotides compared to one of 150 nucleotides (with the primers F-3005 and R-3151). With the couple of primers used to detect the mutation M184V (F-3005 and R-RT184V3), the size was exactly 118 bp.

The same conditions were then applied for deliberate mixtures of NL4-3 and mutant 184V at different ratios in order to discriminate minority species from non-specific annealing. The mix was performed with PCR products obtained with the primers D-1818 and R-3584.

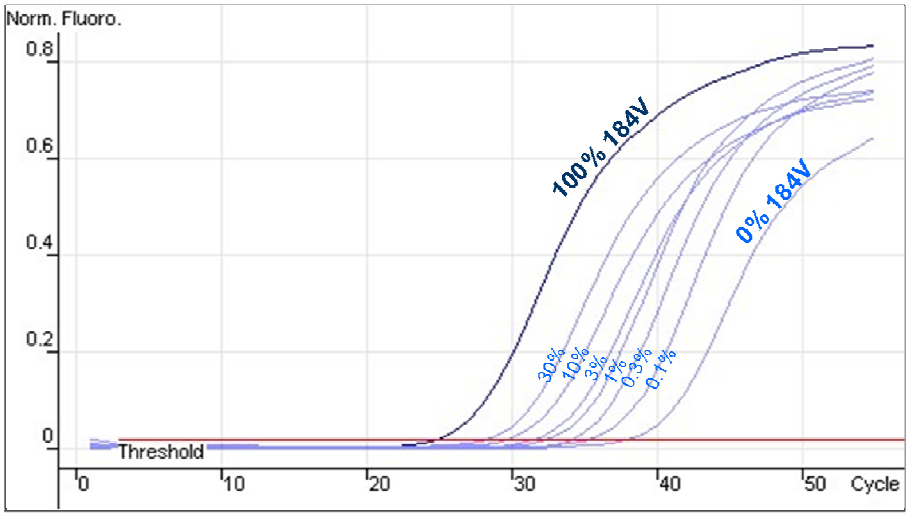


Figure 40. *Real-time PCR on mixed populations of wild type and 184V mutant. The amplification was performed at 57°C with the primers F-3005 & R-RT184V3. The fluorescence, representing the amount of DNA, is plotted as a function of cycle numbers (linear scale).*

In the graph above the first curves to appear were, as expected, for the highest percentage of mutant 184V; the last curves represented the lowest mutant percentage. The system could easily quantitatively distinguish with high discrimination minor populations from each other: all mixes containing between 0.1% and 30% mutant were detected. On the other hand, once the target virus constituted >90% of a mixed population the mutant-specific primer for 184V reached the upper limit of resolution of the system (data not shown).

The respective threshold cycle for a first detection of a PCR amplicon for the 184V mutant was plotted as a function of the relative proportion of mutant 184V in a defined mix

(Figure 41). Each calculation was adjusted for the absolute quantity of a viral population present in the reaction by means of a second, parallel PCR using primers in regions isogenic for both proviral plasmids (F-3005 and R-3123). The primer R-3123 is similar to the R-RT184V3 without the last nucleotide at the 3' end.

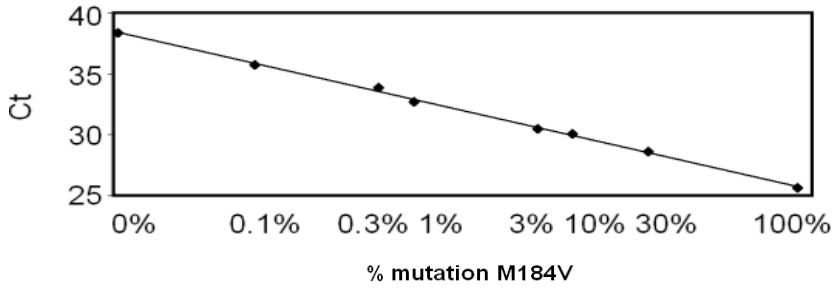


Figure 41. *Real-time PCR results plotting for precise in vitro mixes of wild type and mutant the threshold value as function of the relative amount of mutant 184V.*

The results showed that the ΔCt between 0.1% and 0% mutant was still more than two cycles and therefore discernable. This demonstrated reliable and quantitative identification even of small sub-populations in a mixed HIV sample. These data were conversely verified further by using a primer pair specific for detecting and quantifying small amount of wild type in a mixed sample with mutant virus (not shown). Using multiple repeats of duplicate determinations, as described by Hance *et al.*[90], a lower limit of sensitivity was determined using pure wild type ("0% mutant") with the mutation-specific primer pair: wild type and mutant could not be discriminated anymore once the mutant constituted in a mix less than 0.02% ± 0.03% (mean ± 2SD, n=9). These values correlated well with those previously reported for protease[90]: cut-off value of 0.05%, above which a sample was considered positive for a mutant.

The settings established by this *in vitro* validation were applied to the five discordant patient samples which had already been validated as containing 3TC-resistant minority species by allele-specific PCR (Figure 42).

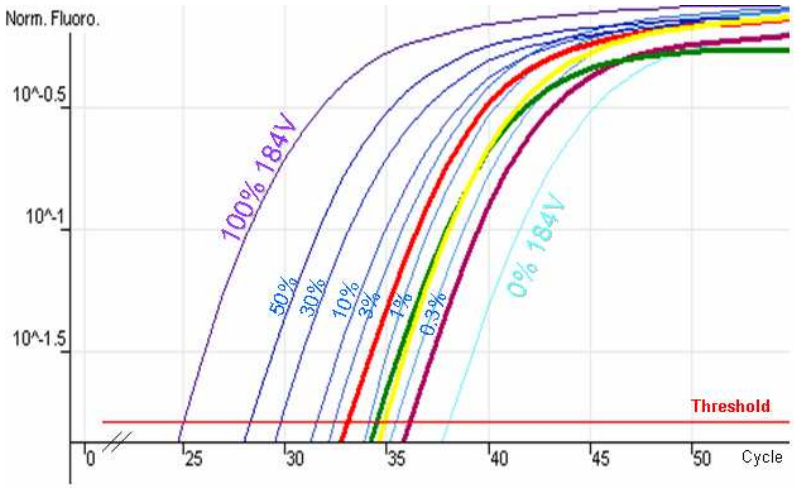


Figure 42. *Real-time PCR on mixed populations of wild type and 184V mutant and on four patient samples. The amplification was performed at 57°C with primers F-3005 & R-RT184V3. The fluorescence, representing the amount of DNA, is plotted as a function of cycle numbers (logarithmic scale). In red: patient G, in green: patient C, in yellow: patient D and in purple: patient A.*

Like for the range of precise *in vitro* mixes between wild type and mutant 184V (Figure 43, inset graph), the respective threshold cycle for the first detection of a PCR amplicon for the 184V mutant was plotted as a function of the relative proportion of mutant 184V in a defined mix (Figure 43). Each calculation was adjusted for the absolute quantity of a viral population present in the reaction by means of a second, parallel PCR using primers in regions isogenic for both proviral plasmids.

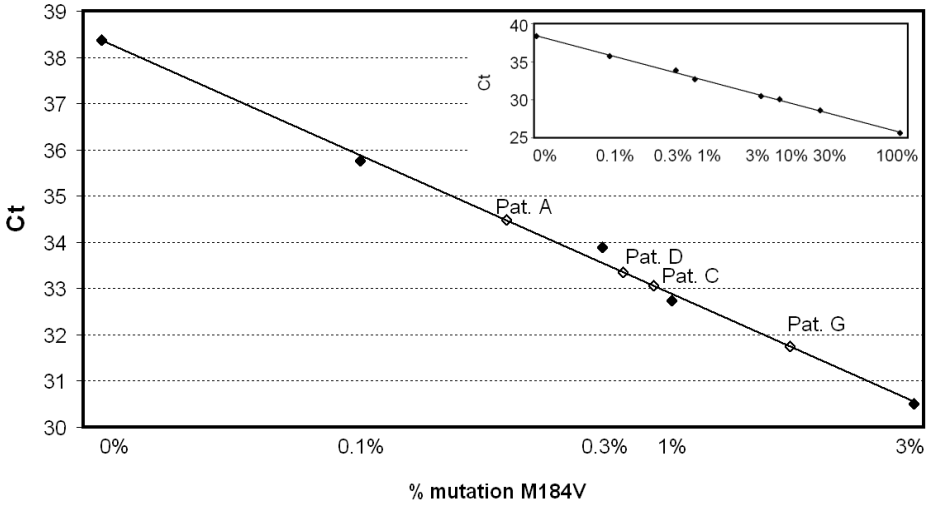


Figure 43. *Real-time PCR results plotting for precise in vitro mixes of wild type and mutant and for four patient samples the threshold value as function of the relative amount of mutant 184V. Pat: Patient.*

I was able to demonstrate the presence of 3TC-resistant viral species in four authentic clinical samples. The percent mutated sequences in the circulating pool of viruses in the studied patients were calculated each by duplicate determinations in four independent experiments. Analysis of plasma samples from patients A, C, D and G verified indeed the presence of the 184V-mutant in the original specimen (Table 13).

samples	threshold cycle (Ct)		% M184V
	template-specific primer	mutant-specific primer	
100%	15.85	25.66	100.00
30%	15.68	28.12	24.88
10%	14.96	28.95	9.97
3%	14.21	30.02	3.02
1%	15.44	32.26	0.91
0.3%	14.74	33.38	0.28
0.1%	15.79	35.28	0.10
0%	16.41	37.82	0.02
G	17.12	32.63	1.51
C	19.97	35.34	0.61
D	17.84	34.45	0.50
A	15.72	34.16	0.23

Table 13. *Threshold cycles for deliberate HIV mixes and for patient samples.*
From the titration curve depicted above, key values were extracted to show expected versus measured percentage of mutant in mixed populations (top part of table), and to quantify the relative amount of mutant virus in patient-derived samples G, C, D, A.

The relative abundance of a mutant behind a wild type background was 0.25%, 0.53%, 0.85% and 1.5% (average of 4 experiments in duplicate) for patients A, D, C and G, respectively. The low abundance of mutant either hints an early phase of emergence of mutant 184V in these clinical samples or indicates resistant viruses persisting after cessation of 3TC treatment. For the fifth sample, patient B, no 184V mutant could be called. Ct curves for quantification (non-mutant primers) were shifted to the far right and were shallower in shape, possibly indicating suboptimal annealing. The mutant-specific amplification curve could not be separated from the curve of <0.01% mutant. As a result the data could not be interpreted for resistance determination. This could happen e.g. in case that further mutations in the respective primer regions occurred.

V.2.4. Case Report

One female patient, who had been diagnosed with HIV-1 in September 1996, displayed a viral load of 70'004 RNA copies/mL and a CD4 count of 293 cells/mL (12% CD4) at baseline before starting treatment. Genotypic resistance testing was performed and indicated the presence of a subtype B strain with resistance mutations V179I within the RT, M36I and L63P within protease (interpreted according to the Stanford DB v.3.8 algorithm). It resulted in a resistance score of 4 for all the PIs but a null score for the RTIs. In parallel, rPhenotyping also diagnosed a fully susceptible virus and therefore the combination TDF + 3TC + NVP was started in November 2003. Two years later, no response was observed despite good adherence to therapy. Resistance testing showed the acquisition of the mutations K65R, Y115F, Y181C and M184V within the RT, conferring resistance to TDF, ABC, NVP and 3TC, respectively. The same resistances were observed by rPhenotyping. The follow-up of the patient is described in Figure 44. Therapy was changed to 3TC monotherapy at month 27 (arrow head in Figure 44) but no information on the clinical reasons was available. This step resulted in a continuous virological increase and a decline of the immunological response. In December 2006 (month 39), the viral load was detected at 6'472 RNA copies/mL. After a new HIV resistance test showing 3TC and NVP resistance (only by rPhenotyping for the second), a third-line regimen was started with the combination 3TC + AZT + LPV/r. Clear virologic response was indicated by an immediate drop in VL to 40 RNA copies/mL after the first month of third-line therapy.

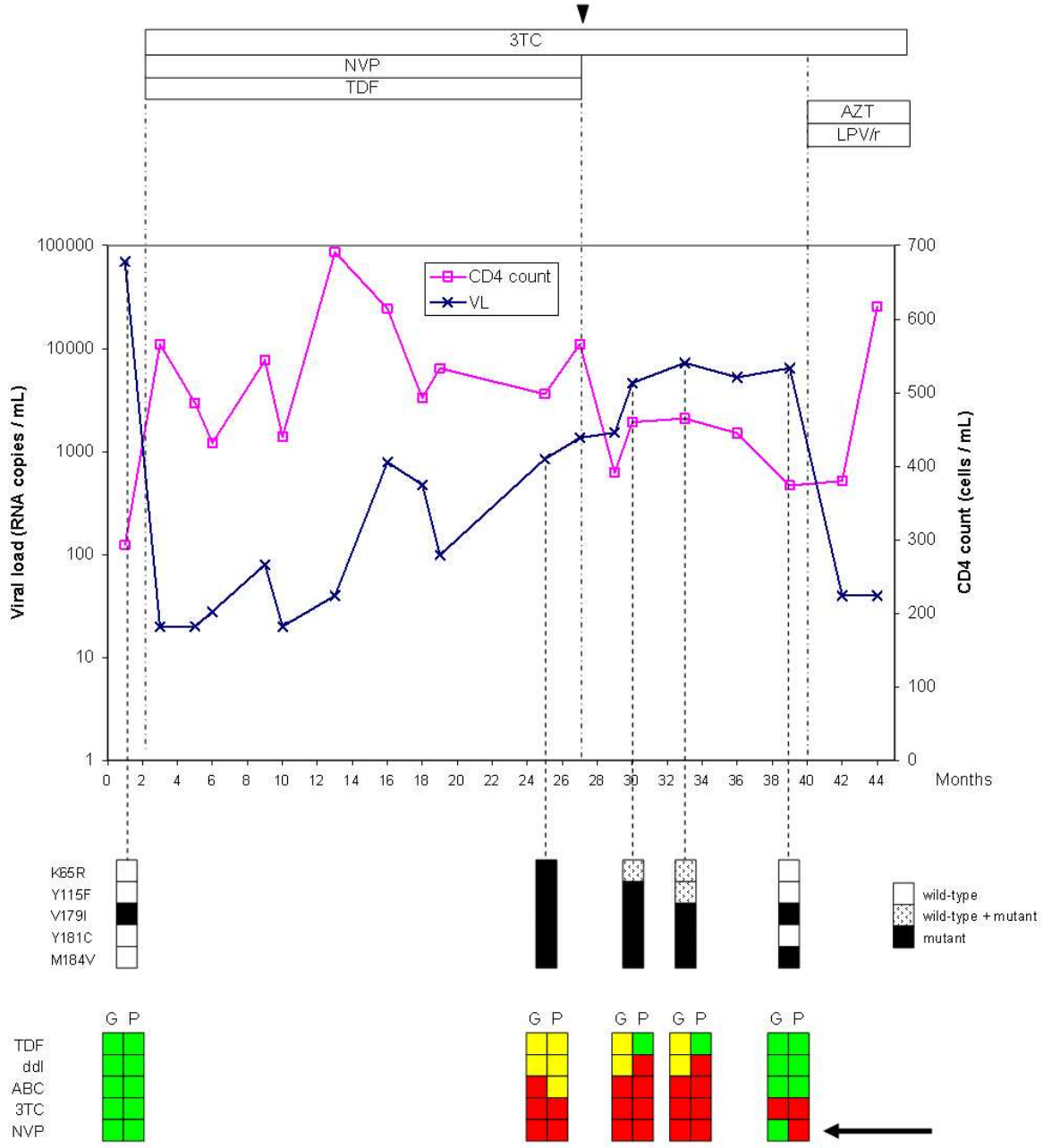


Figure 44. *Follow-up of the patient. The history of CD4 count, viral load and therapy are depicted (the first month corresponds to October 2003). Below the graph, the presence of the wild type and/or mutant amino-acid at resistance-associated positions within RT as observed by population-based sequencing of the corresponding samples is displayed. The mutations M36I and L63P in the protease gene did not lead to resistance to PIs. The genotypic (G) and phenotypic (P) interpretations are represented for each drug having a decrease in susceptibility during treatment. Three levels of resistance are depicted: susceptible (green), intermediate resistance (yellow) and resistance (red).*

After the cessation of TDF treatment at month 27 (arrow head in Figure 44), a mixture of the wild type (K) and mutant (R) amino acid at the RT position 65 was already observed three months later (small peak corresponding to the wild type codon and a high peak for the mutant, month 30 in Figure 45). For each genotypic resistance determination, an electrophoretic representation corresponding to the region of amino acid 65 in the RT gene is shown in Figure 45. At month 33, only a small peak indicating mutant amino acid (R) was depicted on electrophoregram. One year after the therapy change, a pure wild type was detected at position 65 of the RT indicating apparent TDF and ddI susceptibilities.

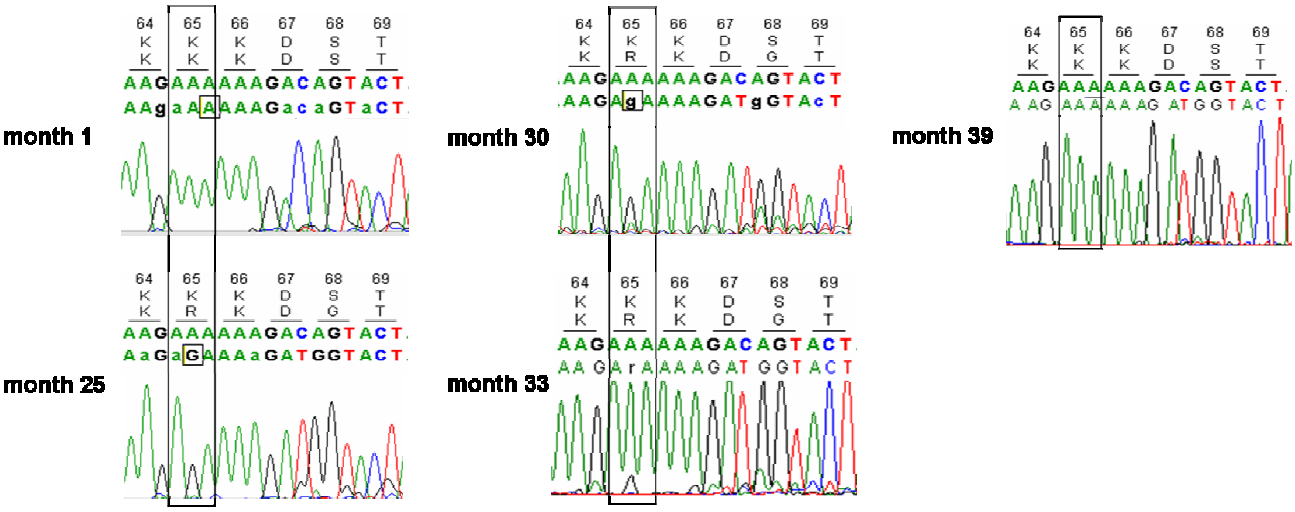


Figure 45. *Electrophoregram representing a part of the RT gene from the patient's viral population showing mixed amino acids at position 65.*
Genotypic determinations of HIV resistance was performed: just before to start a first-line treatment (month 1), after two years of tritherapy (TDF + 3TC + NVP) (month 25), after starting a 3TC monotherapy (months 30 and 33) and one year after beginning the second-line treatment (month 39).

Similarly, at position 115 of the RT the mutant amino acid (F) slowly disappeared (month 33) to be replaced by pure wild type (Y) at month 39 leading to ABC susceptibility (electrophoregrams not shown).

Genotyping and rPhenotyping of the patient samples were in good agreement for the detection of resistance to 3TC, ABC and ddI at any resistance test. This was not the case for the two antiretroviral agents TDF and NVP.

In case of TDF resistance, genotypic and phenotypic interpretations concluded as intermediate resistance at month 25 but since the mixture of wild type and mutant 65R was pinpointed by genotypic analysis (months 30 and 33), the PhenoTecT system did not reveal resistance to TDF. A study had been done on pre-defined mixture of wild type and 65R mutant plasmids to detect the resistance to TDF by rPhenotyping (data not shown). Only a pure mutant population was detected resistant whereas the different mixes were all susceptible to the drug. A possible explanation to the non-detection of the 65R variant present in a viral population with wild type virus could be its low fitness (20%) compared to the reference[140-142]. Due to the format of the phenotypic assay, the transfection of the first cells, before further incorporation of ARV, will generate a high background of fit wild type virus which will prevent the good detection of the new TDF resistant minority species.

Concerning the mutation Y181C, one year after the stop of NVP therapy it was not detected anymore by Genotyping but the PhenoTecT system was still able to detect a resistance to NVP (black arrow). Looking to the electrophoretic peaks at months 30 and 33, a wild type variant was appearing whereas the 181C variant declined progressively, as shown with the mutant 65R. Therefore, the discrepancy observed between Genotyping and rPhenotyping at month 39 could be explained by the persistence of NVP-resistant HIV-1 species in the viral population at a level non-detectable by population-based sequencing (less than 15%).

To demonstrate the relevance of the 181C minority species in the discordant patient sample at month 39, an allele-specific PCR was adapted to the Y181C mutation in the same way as for the detection of minor 184V variants. A reverse primer R-RT181C3 was designed to carry at its 3' terminus one nucleotide specific of the variant to be discriminated and three internal additional mutations. In PhenoBase®, the mutation Y181C is in 92.1% due to the substitution of the codon TAT (Tyrosine) to TGT (Cysteine). It is also the case of this patient; this is why the primer designed was carrying a Cytosine in 3' end. The forward primer used was D-2232. The optimisation was done on known mixes of wild type (NL4-3) and 181C mutant plasmids (Figure 46, right part). An unspecific amplification was observed for a product of approximately 1600 bp. By alignment of the primer sequence with the reference NL4-3, I discovered that the last nine nucleotides in the 3' end could also anneal at the C terminal region of RT (around the amino acid 415). This annealing will provide a PCR product of 1590 bp which corresponds to the band observed on the agarose gel. The expected amplification product had a size of approximately 900bp (black arrow on the left).

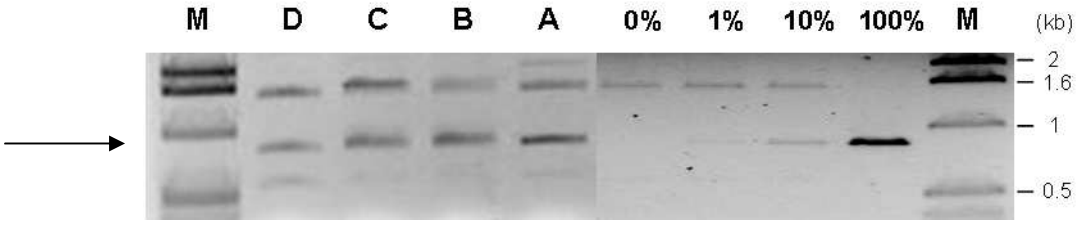


Figure 46. *Mutation-specific PCR of precise in vitro mixes of wild type and mutant 181C (right side) and of plasma from the patient (left side, lanes A-D).*
Sample A was diagnosed at month 25, sample B at month 30, sample C at month 33 and sample D at month 39. The PCR products were analysed by electrophoresis following amplification with the 181C3-mutant primer (annealing at 62°C). M: Size Ladder (1kb).

The percentages indicated on the top of the gel picture correspond to the proportion of mutant plasmid in the mix. The 0% (pure wild type) was not amplified validating the specificity of the amplification. The detection of mutant 181C was possible down to 1% in a mixture with wild type. In parallel four patient samples tested previously by genotypic and phenotypic assays were also amplified under the same conditions. The protocol revealed the presence of the 181C variant in all samples A, B, C and D with a poorer amplification for the last. The three former had been already interpreted resistant to NVP by the Stanford algorithm based on the presence of the mutation Y181C. But for the latter, sample D, no Y181C mutation had been detected at the time of population-based Genotyping while rPhenotyping had diagnosed a resistance to NVP. The amplification of sample D by a sensitive mutation-specific PCR confirmed the phenotypic result and the presence of a persistent minor 181C variant in the viral population. The use of NVP in a future treatment scheme would lead instantaneously to the re-emergence of the 181C variant as a major species.

V.3. Discussion chapter III

The earliest possible detection of HIV resistance is fundamental for the HIV infected patient in order to adapt an appropriate therapy preventing viral escape. Even patients never treated are more and more diagnosed for HIV resistance. Recent evidence suggests that 1 in 10 new infections involve resistant strains of virus[143]. The standard genotypic resistance test is an easy and rapid methodology to detect the outbreak of resistance associated mutations. However, its low sensitivity is a limit to the detection of virus present at less than 15% of the total population as in particular clinical situations e.g. minor resistant species. These minority populations remain therefore undetected and can continue to evolve and ultimately emerge as dominant viral populations leading to a rapid failure of therapy.

In this study we intended to demonstrate that a replicative format of rPhenotyping was able to detect the otherwise missed presence of relevant, drug-resistant minority species in clinical specimens. Such variants in authentic clinical samples could explain a part of the discrepancies Geno-S/Pheno-R observed between Genotyping and rPhenotyping. The 3TC-resistant and NNRTIs-resistant variants are good examples to study as for these antiretroviral agents, and not for the others, the presence of specific single mutations results in a high-level of resistance. An independent technical proof of the presence or absence of these unique mutations being available by applying a highly sensitive allele-specific real time PCR protocol, it is possible to confirm the high sensitivity of rPhenotyping to detect minority species. In this study the 3TC-resistant variant 184V was specifically used since it is the unique mutation responsible for this resistance. In case of the NNRTIs, several individual mutations exist, that can lead to resistance and which could underline the presence of minority species. The highly sensitive rPhenotyping answer may be of general utility for diagnostic use as the methodology does not require an algorithmic interpretation and is thereby principally applicable with no difference between the drug classes. For such special cases the substantial gain in sensitivity might merit the principally more complex format of Phenotyping.

Seven patients discordant Geno-S/Pheno-I or -R for 3TC were selected in PhenoBase®. By a first methodology, mutation-specific PCR, five of these patient samples were identified as presenting 3TC-resistant variant 184V. An allele-specific real-time PCR confirmed in a second time the presence of such variants in four of them. When linking the findings of this study to treatment history as listed in Table 12 (see page 94) this further confirmed that a resistance to 3TC had either arisen only recently, currently in ‘statu

nascendi' or as non-prominent virus representatives e.g. (still) restricted to a putative small compartment (patients A, C, D). In contrast, previous therapy episodes with 3TC e.g. for patient G suggest that resistant viruses had been selected during this time and were still detectable months after discontinuation.

As indication for therapeutic failure, the selected patients presented detectable plasma virus loads between 2'400 (F) and >100'000 c/mL (A). The persisting mutant viruses were therefore likely to in part contribute to the unsuccessful viral suppression. In contrast, for the sample from patients E no 184V virus was detected, and, according to the treatment record, this patient had never received 3TC. This very case might illustrate the limits of a PCR-based verification as such a protocol is only suitable for perfectly matching sequences. M184V-based resistance to 3TC is mainly caused by an ATG to GTG transition (for >90% of all matching Geno-R/Pheno-R patients in the database). Nevertheless, a small proportion of patient-derived resistant viruses rather carry a GTA codon (Val). This double mutated codon (ATG→GTA) will be missed by the chosen strategy.

Similarly mismatches in forward or reverse amplification primers will prevent PCR amplification. For the allele-specific PCR the forward primer was annealing in the P6 region upstream of the protease gene. The template used to perform this amplification was the recombinant plasmid carrying only a part of the RT gene from the patient. Therefore the P6 gene was "wild type" for the seven isolates. In case of the reverse primers R-RT184V3 and R-3123 (similar to the R-RT184V3 without the last nucleotide at the 3' end) they were designed based on the sequence of the wild type NL4-3. It is likely that these two oligonucleotides did not hybridise to the HIV-1 targets of the patients E and F, because of a too high number of mismatches in the primer binding site. Indeed the sequence of subject E presented a mutation in codon 190 (GGG instead of GGA) and subject F had one in codon 186 (GAC instead of GAT) (Figure 47). Although these mutations were identified in the major variant we can extrapolate that minor populations might carry them as well. This is probably the reason why no amplification was observed after allele-specific PCR for patients E and F.

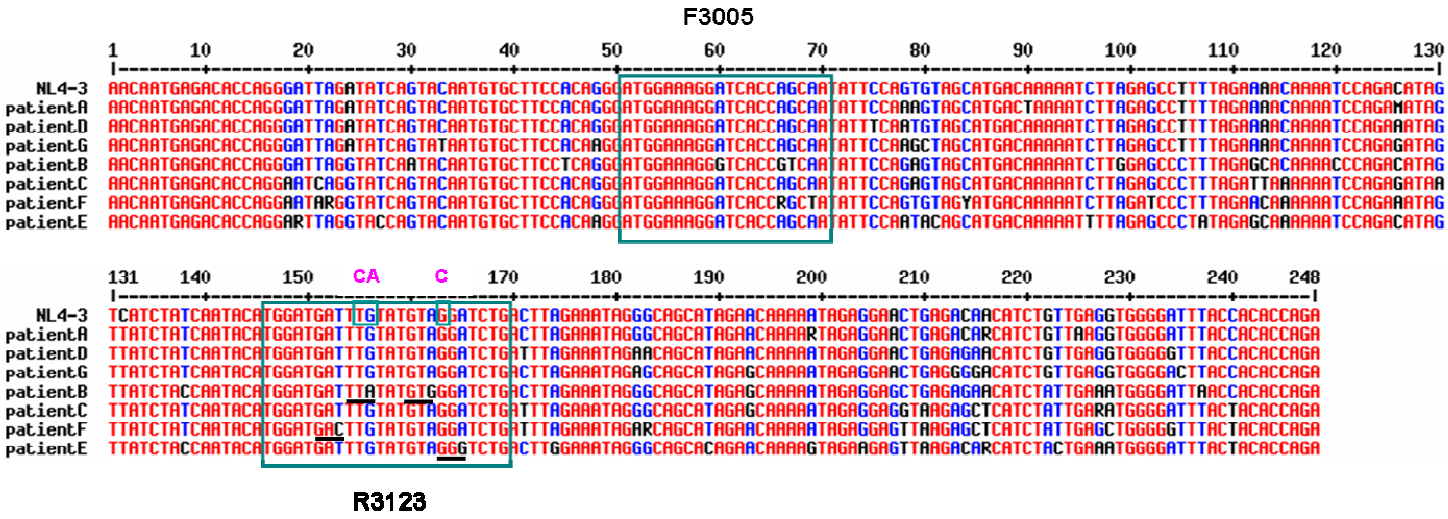


Figure 47. *Alignment with Multalin[H] of cDNA sequences corresponding to a part of the RT gene of patients A-G and the reference NL4-3 with the oligonucleotides F-3005 and R-3123. The nucleotides in pink are the three internal mutations present in the sequence of the primer R-RT184V3 versus the target.*

For the sample from patient B amplification by allele-specific PCR was successful, but real-time PCR repeatedly failed. Subtype analysis revealed that the virus was of AG subtype. As the consensus of this HIV recombinant presents with three mismatches in the forward quantification primer (F3005 in Figure 47) the subtype would explain the observed failure to quantify; this was not genotypically verified for the virus minority in sample B. The alignment of the population-based sequence from patient B with the reverse primer (Figure 47) revealed as well two mis-annealing with the wild type reference. In the allele-specific real time PCR reaction three mismatches are needed to specifically amplify the targeted variant. Due to a mutated nucleotide in codon 187 (TTA instead of TTT) only two internal mutations are observed between reverse primer and sample B. Nevertheless a third is contributed by a new mutation in codon 189 (GTG instead of GTA). Finally to amplify by allele-specific PCR the sample from patient B, this reverse primer should be matching but it will require another forward primer.

The last point discussed in the *chapter II* was the mis-annealing of the primers used to amplify the RT gene in the rPhenotyping protocol (D-2582L and R-3525L). This could lead to an arbitrary selection of a subpopulation corresponding perfectly to the primer sequence unlike another presenting a mismatch in this particular region; primer bias issue. The samples diagnosed Geno-S/Pheno-R to a drug could illustrate this issue by a preferential amplification of a minor variant resistant to the ARV. To ensure that detection of a minor variant is not due to a preselection, the sequence of the seven patients involved in the study on minority species

were aligned with the sequence of the two primers aforementioned. Any mismatch was observed meaning that the major variant is perfectly matching the primer (data not shown).

The detection of viral minority species carrying mutation M184V was possible for samples A, C, D and G with 0.25%, 0.85%, 0.53% and 1.5%, respectively (average of 4 experiments in duplicate). However, a small risk of underestimating one or the other HIV-1 variant cannot completely be excluded. For patients C and D, the ordering physician had suspected the transmission of a resistant virus; it is why a resistance test was prescribed. In these cases a transmitted mutant virus will be hidden by the wild type virus which is likely to be more fit. Patient G had previously received 3TC treatment which might have generated a mutant 184V variant. This variant might even after therapy cessation still be detectable as minority species in the absence of 3TC pressure. Patient A is currently under 3TC treatment along with AZT and EFV and therapy is beginning to fail: Genotyping detected a full sensitive virus whereas rPhenotyping revealed a resistance to 3TC. The detection of this 184V minor resistant virus at 0.25%, however, can not explain therapy failure but viral resistance due to other non-canonical, resistance-associated mutations is not unlikely. This patient is infected by a subtype AG strain and, as already mentioned, the algorithms on non-B virus are not predictive enough to ensure that non-characterized mutations which may be less relevant for subtype B did not have an impact on non-B variants of HIV.

The allele-specific PCR and real-time PCR methodologies were able to demonstrate the presence of 184V minor variant in a wild type background with a limit of sensitivity of 0.3% and 0.05%, respectively. These methods are of course not suitable for a diagnostic routine on patient samples. The high sequence variability of HIV-1 makes it necessary to adjust the primer sequences for each individual isolate and to establish standard curves for correct quantification with wild type and mutant template from the sample under investigation. In this study, they were used as a tool to prove the particular acuity of the PhenoTecT system in detecting therapy-threatening minorities in clinically relevant therapy situations.

The detection and characterization of minor drug-resistant variants is important for the clinical management of HIV infection. Using rPhenotyping it is possible to detect viral populations resistant either to 3TC/FTC or to NNRTIs as minority species not observed by standard Genotyping. Their early detection could help avoiding viral rebound as a consequence of treatment with an ineffective regimen. Minor virus variants arising in a patient’s circulation could provide an explanation for the onset of progression towards therapeutic failure since their early presence may provide for viral optimisation, e.g. of a

viruses replicative capacity[144] that could occur unnoticed in the background. It appears logical that in clinical settings such viral populations might be particularly prevalent in non-compliant patients, during drug holidays, under structured treatment interruption strategies[90,145,146], or as early events in the context of transmitted mutant HIV variants with suboptimal replicative fitness[147,148]. A significant fitness-cost would lead to a more or less stable balance with wild type viruses and in the absence of drug pressure prevents them from easily becoming the predominant virus form in a patient. It will be interesting to apply the technology (and hypothesis) to studies on transient therapies, e.g. the short-term administration of NVP to pregnant women for prevention of mother-to-child transmission for further investigation of the presence and persistence of viral minorities with replication properties inferior to wild type[149-152]. Additional studies are needed to determine the frequency of specimen variants in clinical situations.

The detection of such minor resistant variants is more and more explored showing their importance. In fact, other methodologies have been described in several studies for detecting low-frequency drug resistance mutations. In 2005, Palmer *et al.* detected by SGS (Single Genome Sequencing) minor variants carrying M41L, T69N, L74V, K103N, Y181C, or T215Y which were not detected by standard genotyping analysis[153]. The AS-PCR (Allele-Specific PCR) was also able to show the reemergence of virus present (i) before STI that encoded 3-drug class resistance on the same genome[154] or (ii) after single intrapartum dose of NVP for the prevention of mother-to-child transmission of HIV[151]. All these numerous methodologies could determine the prevalence of such HIV minor variants carrying mutations associated to ARV resistance in authentic clinical specimens; nevertheless this requires knowing in advance the mutations to look for while rPhenotyping can be used as diagnostic tool to detect them.

Another explanation to the samples with Geno-S/Pheno-R discordance could be the presence of mutations which previously had not been associated with drug resistance; uncharacterized mutations. Novel NNRTI resistance mutations have been characterized in patients with primary HIV infection. In fact, some studies revealed the presence of viral strains with reduced susceptibility to NNRTIs, which were not carrying detectable amounts of any of the mutations associated to NNRTIs resistance[155-157]. These viral strains had instead amino acid substitutions at positions 135 and 283 in subtype B virus[158] and the mutations 135L, 139V and 245T in an isolate of subtype D virus[155]. By *in vitro* mutagenesis studies, the effects of these mutations indicated a susceptibility to the drug lower than the wild type control NL4-3. Reverse genetics on viral population presenting those mutations led to a re-sensitisation to the ARV. Several specific combinations of mutated amino acids at these sites

were tested to assess their possible synergy on drug resistance: 135L+283I, 135M+283I and 135T+2183I[40,156]; 135L+245T[155]. All samples diagnosed Geno-S/Pheno-R to EFV and/or NVP in PhenoBase® were compared to those concordant Geno-S/Pheno-S for the four amino acids at positions 135, 139, 245 and 283. Each subtype has been separately analysed. Unfortunately, even if at least one of these four mutations was present in 50% to 100% of the samples, depending on the subtype, no correlation could be identified between their presence and the discordant profile Geno-S/Pheno-R. Effectively, similar mutations were also found in concordant Geno-S/Pheno-S samples. The presence of others mutations which could interact negatively on resistance level would have to be considered. The unique mutation observed in case of discordance Geno-S/Pheno-R, and never in susceptible concordant samples, was the I135R. To confirm a possible impact of this substitution at position 135 of RT gene, an in-depth study should be performed; if so, the genotypic algorithms will have to be updated in order to take into account the weight of this mutation in NNRTIs resistance.

Concluding remarks

The present study underscores the considerable ability of HIV-1 to generate resistance mutations, eventually resulting in the coexistence of numerous viral quasispecies with distinct resistance genotypes throughout the evolution of resistance. Thus, they appear to have played an active role in the development of resistance in the patients studied here and may accelerate the process. The development of strategies for the evaluation and treatment of patients with drug resistance that adequately take into consideration this additional level of complexity might produce clinical benefit.

Genotypic algorithms are self-optimising systems that improve their performance over time. A significant discordance was observed between interpretations inferred by different genotyping algorithms. This raises the urgent requirement to define common rules and possibly even new criteria to be considered. This will increase coherence of results and therapy advices within national and international studies and cooperation between centres. Nevertheless some intrinsic shortcoming remain that necessitate the use of alternative dissecting methods such as rPhenotyping: Genotyping principally fails to assign new mutations or novel resistance patterns provoked by new drugs, and systems are limited in the ability to follow viral speciation during extended treatment periods as this leads to the formation of viral minorities. The evolution of a growing complexity of viral genomes in any given patient requires dissection and assignment of mutations to individual viral genomes. Data presented in this study demonstrate that the principle of a functional assessment by rPhenotyping can strongly detect: (i) minor HIV populations with a sensitivity level of less than 1%, (ii) viral mixtures by analysing each single virus separately, (iii) viruses carrying any known mutations for actual drug and (iv) viruses presenting mutations with synergistic or antagonistic interactions. This renders it, for these applications, superior to Genotyping. Nevertheless further studies will have to show how this gain in sensitivity will translate into clinical utility and improved disease management (ongoing project with the Swiss HIV cohort and the group of M. Battegay, USB).

In the future the treatment patients should benefit from those improvements, which will assist in identifying new drug-options. In addition, these observations will form the basis for an application in other viral diseases, notably HBV and HCV, where similar issues are likely to arise with the introduction of potent antiviral drugs.

Background

The Human Immunodeficiency Virus (HIV)

❖ HIV classification

According to the International Committee on Taxonomy of Viruses (ICTV), HIV was classified as a member of the genus Lentivirus, part of the family of Retroviridae. Infections with Lentiviruses typically show a chronic course of disease, a long period of clinical latency, persistent viral replication and involvement of the central nervous system. Lentiviruses are transmitted as single-stranded, positive-sense, enveloped RNA viruses. The HIV genome is a dimer of a two (+) ssRNA molecules, which are not 100% homologous, held together by hydrogen bonds. The viral RNAs have a 5' cap and a 3' poly-(A) tail like mRNAs. Since HIV mutates very quickly, there are many different strains, even in a single infected person. Based on genetic similarities, the numerous virus strains are classified into types, groups, and subtypes.

There are two principal types of HIV: HIV-1 and HIV-2. HIV-1 is thought to have originated in southern Cameroon after jumping from wild chimpanzees (pan troglodytes troglodytes) to humans during the twentieth century[159,160]. HIV-2 may have originated from the Sooty Mangabey (Cercocebus atys), an Old World monkey of Guinea-Bissau, Gabon, and Cameroon[161]. Both types are transmitted from human to human by sexual contact, through blood, or from mother to child, and they appear to cause clinically indistinguishable AIDS. However, it seems that HIV-2 is less easily transmitted, and the period between initial infection and illness may be longer for HIV-2. Worldwide, the predominant virus is HIV-1, and generally when people refer to HIV without specifying the type of virus they will be referring to HIV-1 (as it will be in this report). The less common HIV-2 type is concentrated in West Africa and is rare elsewhere.

Using electron microscopy, HIV-1 and HIV-2 show a close morphological similarity. However, they differ with regard to the molecular weight of their proteins and have differences in their accessory genes. Both HIV-1 and HIV-2 replicate in $CD4^+$ T-cells and are regarded as pathogenic in infected persons, although the actual immune deficiency may be less severe in HIV-2-infected individuals.

HIV-1 can be classified into three groups: M(ajor), O(utlier) and N(ew), which was only discovered in 1998[162]. These three groups may represent three separate introductions of SIV Simian Immunodeficiency Virus) into humans. Group O seems to be restricted to Africa and group N is even rarer. The next topics will relate to HIV-1 group M only. Within group

M nine genetically distinct subtypes (or clades) are known: A, B, C, D, F, G, H, J, and K[163] (Figure 48).

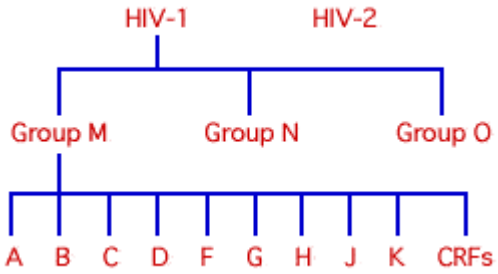


Figure 48. *Phylogenetic relationship of HIV.*

Occasionally, two viruses of different subtypes meet in the same cell in an infected person and mix their genetic material to create a new hybrid virus (a process similar to sexual reproduction, and sometimes called "viral sex")[164]. Many of these new strains do not survive for long, but those that infect more than one person are known as "circulating recombinant forms" or CRFs. For example, the CRF A/G is a genetic hybrid of subtypes A and G. One of the CRFs is called A/E because it is thought to have resulted from recombination between subtypes A and E. However, no pure form of subtype E has been found yet. In fact, CRF A/E is more correctly called CRF01_AE (http://www.hiv.lanl.gov/content/hiv-db/CRFs/CRFs.html). The HIV-1 subtypes and CRFs are very unevenly distributed throughout the world (Figure 49), and the most prevalent are subtypes B (found mainly in North America and Europe), A and D (found mainly in Africa), and C (found mainly in Africa and Asia). In Europe subtype B remains the most common but other subtypes are becoming more frequent and now account for at least 25% of new infections.

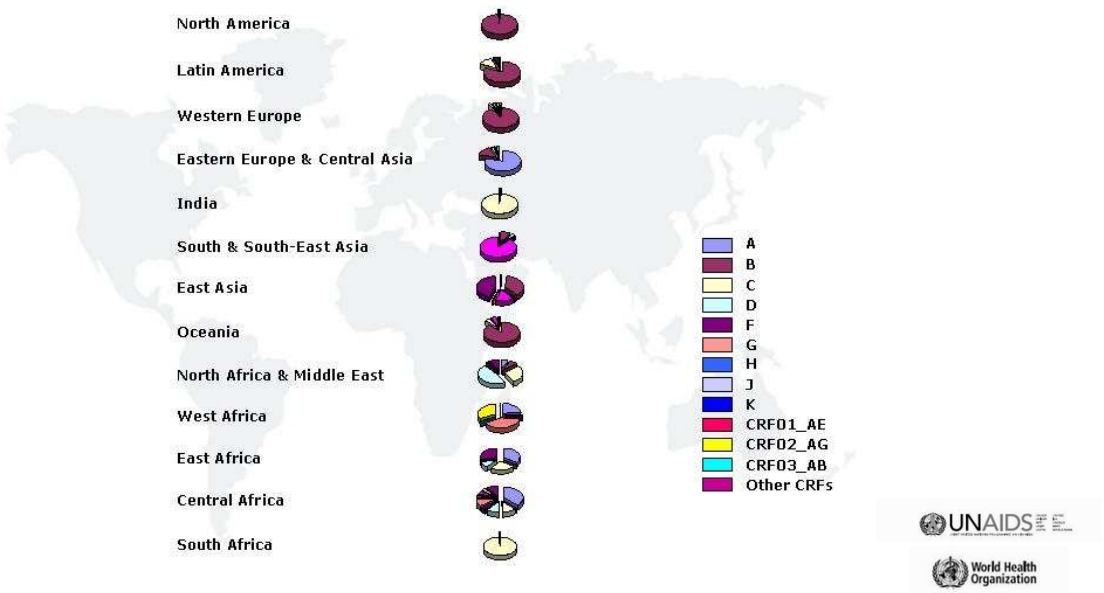


Figure 49. *Global geographical distribution of HIV-1 genetic forms.*

❖ Its morphological structure

HIV-1 viral particles have a diameter of 100 nm and are surrounded by a lipoprotein membrane. Each viral particle contains 72 glycoprotein complexes, which are integrated into this lipid membrane, and are each composed of trimers of an external glycoprotein gp120 and a transmembrane spanning protein gp41. The bonding between gp120 and gp41 is only non-covalent and therefore gp120 may be shed spontaneously within the local environment. Soluble glycoprotein gp120 may also be detected in the serum[165] as well as within the lymphatic tissue of HIV-infected patients[166]. During the process of budding, the virus incorporates different host proteins from the membrane of the host cell into its lipoprotein layer, such as HLA class I and II proteins, or adhesion proteins such as ICAM-1 (InterCellular Adhesion Molecule) that may facilitate adhesion to other target cells. The matrix protein p17 is fixed to the inside of the viral lipoprotein membrane via a myristoilation anchor. The p24 core antigen contains two copies of HIV-1 RNA. The HIV-1 RNA is part of a protein-nucleic acid complex, which is composed of the nucleoprotein p7 and the reverse transcriptase p66. The viral particle contains all the enzymatic equipment that is necessary for replication: a reverse transcriptase, an integrase p32 and a protease p11 (overview in Gelderblom *et al.*,1989[167]) (Figure 50).

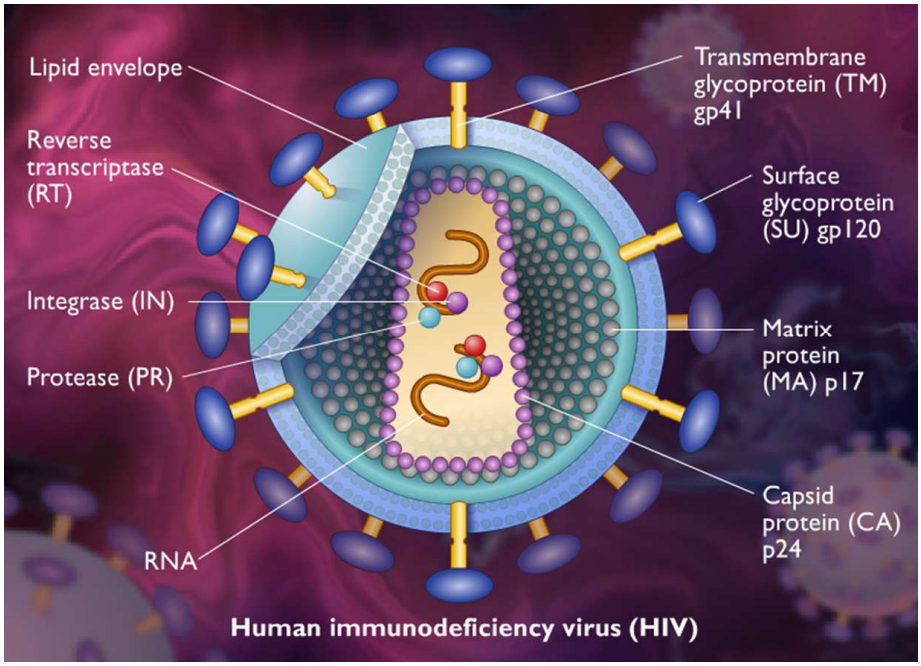


Figure 50. *Structure of an HIV virion particle.*
[http://highered.mcgraw-hill.com/olc/dl/120088/micro41.swf]

❖ The organisation of the viral genome

Most replication competent retroviruses depend on three genes: gag, pol and env. **gag** means "**g**roup-specific-**a**nti**g**en", **pol** represents "**pol**ymerase" and **env** is for "**env**elope" (overview in Wong-Staal *et al.*, 1992[168]) (Figure 51). The "classical" organisational scheme of a retroviral genome is: $^{5'}$LTR-gag-pol-env-LTR$^{3'}$. The LTR ("long terminal repeat") regions represent the two end parts of the viral genome, that are connected to the cellular DNA of the host cell after integration and do not encode for any viral proteins. They play a crucial role in the regulation of the transcription. The gag and env genes code for the nucleocapsid and the glycoproteins of the viral membrane; the pol gene codes for the reverse transcriptase and other enzymes.

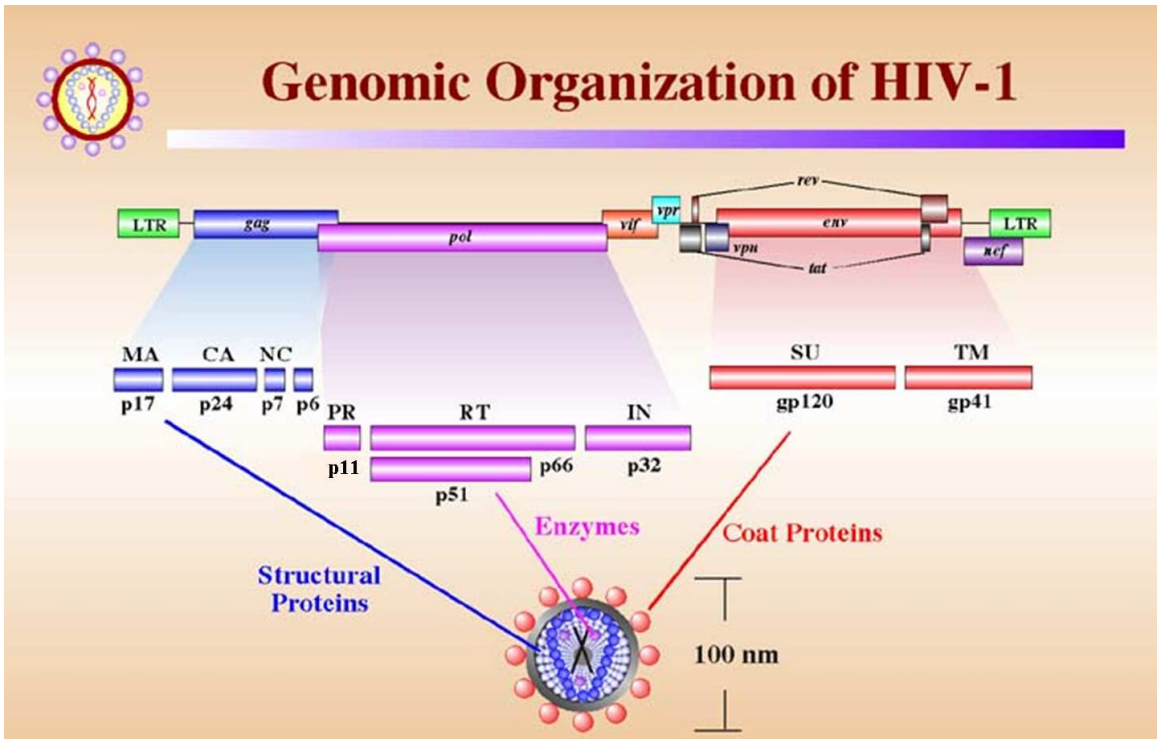


Figure 51. *HIV and its genes.*
[http://www.stanford.edu/group/virus/retro/2005gongishmail/HIV.html]

In addition, HIV-1 contains six genes: tat, rev, nef, vif, vpr and vpu (or vpx in the case of HIV-2) in its approximately 9.8 kb RNA that contribute to its genetic complexity. Nef, vif, vpr and vpu were classified as accessory genes in the past, as they are not absolutely required for replication *in vitro*, while tat and rev are classified as regulatory genes.
The HIV genome has nine open reading frames (leading to nine primary translation products) but 15 proteins are made as a result of cleavage of three of the primary products.

The production of several polypeptides from the same mRNA, through differential splicing and initiation of translation (e.g. Rev and Tat) or ribosomal frameshift, enables to compress maximal information into a relatively small genome. Between *gag* and *pol* genes of HIV, a so-called slippery sequence has been characterized: most ribosomes encountering this sequence, translate it without difficulty reaching finally a stop codon. Some of them will, however, slip back one nucleotide allowing the translation of Pol in addition to Gag but with a different reading frame. The slippery sequence alone is not sufficient to produce this effect; indeed an RNA pseudoknot downstream from it, allows ribosomes pausing longer at that sequence increasing frameshift frequency[169,170]. Ribosomal frameshift enables the production of an excess of Gag in comparison with Gag/Pol, which contains the protease protein needed to cleave Gag and Gag/Pol polyproteins at nine positions to produce structural proteins and additional enzymes.

The structural genes:

- The gag gene gives rise to a precursor protein of 55 kD also called p55. After budding of the viral particle, p55 is cleaved by the virally encoded aspartic protease (a product of the pol gene) during the process of viral maturation into four smaller proteins: MA (matrix, p17), CA (capsid, p24), NC (nucleocapsid, p7) and p6, as depicted in Figure 51[171].

- The Gag-Pol precursor (p160) is generated by a ribosomal frame shifting event[169], which is triggered by a specific cis-acting RNA motif[172]. When ribosomes encounter this motif, they shift approximately 5% of the time to the pol reading frame without interrupting translation. During viral maturation, the virally encoded protease cleaves the Pol polypeptide away from Gag and further digests it to separate the protease (p11), RT (p51), RNaseH (p15), and integrase (p32) activities. These cleavages do not all occur efficiently, e.g. 50% of the RT protein remains linked to RNaseH as a single polypeptide (p65).

- The 160 kD Env (gp160) is expressed from singly spliced mRNA. First synthesised in the endoplasmic reticulum, Env migrates through the Golgi complex for glycosylation. A cellular protease cleaves gp160 to generate gp41 and gp120. Functional, mature Env exists as a trimer on the surface of infected cells and virions[173].

The regulatory genes:

- Tat is a transcriptional transactivator that is essential for HIV-1 replication[174]. Tat is a RNA binding protein, unlike conventional transcription factors which interact with DNA[175,176]. Tat binds to a short-stem loop structure, known as the transactivation response element (TAR), that is located at the 5' terminus of the LTR. The binding of Tat to TAR activates transcription from the HIV LTR at least 1000-fold.

- Rev is a 13 kD sequence-specific RNA binding protein[177]. Produced from fully spliced mRNAs, Rev acts to induce the transition from the early to the late phase of HIV gene expression[178]. Rev binds to a 240-base region of complex RNA secondary structure, called the Rev response element (RRE), which is located within the second intron of HIV within the coding region of gp41[179]. The binding of Rev to the RRE facilitates the export of unspliced and incompletely spliced viral RNAs from the nucleus to the cytoplasm.

The accessory genes:

- Nef (an acronym for negative factor) is a 27 kD protein that is encoded by a single exon that extends into the 3' LTR. Nef, an early gene of HIV, is the first viral protein to accumulate to detectable levels in a cell following HIV-1 infection[178]. Nef has been shown to have multiple activities, including the downregulation of the cell surface expression of CD4[180], the perturbation of T cell activation[181-183], and the stimulation of HIV infectivity[184].

- The Vpr protein is incorporated into viral particles. Vpr plays a role in the ability of HIV to infect non-dividing cells such as macrophages[185] and can also block cell division[186].

- The 16 kD Vpu polyprotein is an integral membrane phosphoprotein that is primarily localised in the internal membranes of the cell[187]. The two functions of Vpu, the down-modulation of CD4 and the enhancement of virion release, can be genetically separated[188].

- Vif is a 23 kD polypeptide that is essential for the replication of HIV in peripheral blood lymphocytes, macrophages, and certain cell lines[189].

❖ Regulation of HIV Gene Expression

The regulation of HIV gene expression is accomplished by a combination of both cellular and viral factors. HIV gene expression is regulated at both the transcriptional and post-transcriptional levels. The HIV genes can be divided into the early genes and the late genes[190,191]. The early genes, *tat*, *rev*, and *nef*, are expressed in a Rev-independent manner after a full splicing. The mRNAs encoding the late genes, *gag*, *pol*, *env*, *vpr*, *vpu*, and *vif* are singly spliced or unspliced, and they require Rev in order to be exported to the cytoplasm for a further translation.

Transcription of the proviral genome:

HIV transcription is mediated by a single promoter in the 5'LTR. Expression from the 5'LTR generates a 9-kb primary transcript that has the potential to encode all nine HIV genes.

The primary transcript is roughly 600 bases shorter than the provirus. It can be spliced into more than 30 mRNA species[192] or packaged as full length without further modification into virion particles (to serve as the viral RNA genome).

The LTRs are composed of three subregions designated U3, R, and U5[193]. The U3 region (for unique 3' sequence) is approximately 450-basepairs (bp) in length and is located at the 5' terminus of each LTR. The central region of each LTR contains the 100-bp R (for repeated sequence) region and the U5 region (for unique 5' sequence) is 180-bp in length. The 5' LTR directs transcription, while the 3' LTR directs cleavage of the primary transcript and addition of a poly-(A) tail. In the 5'LTR, the U3 region contains promoter elements recognised by several transcription factors involved in the formation of the transcriptional machinery whereas the U5 region contains the Tat binding site (TAR) and packaging sequences of HIV. The 3' end of U5 is defined by the location of a lysyl tRNA binding site. The lysyl tRNA acts as primer for reverse transcription.

Regulation of Transcription:

As illustrated in Figure 52, the 5' LTR also functions as enhancer, containing more than one element recognised by cellular transcriptional activators (e.g. AP-1, NFAT-1, USF-1, Ets-1, NF-κB, Sp1, TBP, LBP-1).

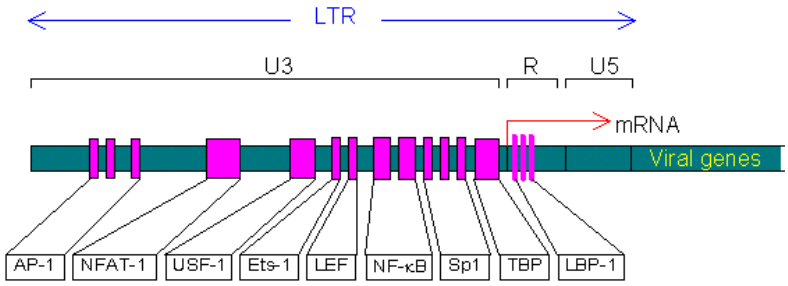


Figure 52. *Broad dissection of the 5' HIV LTR.*

The key factors among the DNA binding sites required for the activation of the transcription of the HIV provirus are those for the NF-kappa B family of transcription factors[194]. Two adjacent NF-kappa B sites are present in the U3 region of the HIV-1 LTR. The NF-kappa B protein allows the virus to be responsive to the activation state of the infected T cell. Stimulation of the T cell receptor (TCR) causes the inactive form of NF-kappa B, localised in the cytoplasm, to be translocated into the nucleus where it induces the expression of a series of T cell activation-specific genes. NF-kappa B and subsequent activation of HIV transcription can also be induced by the cytokines tumor necrosis factor alpha (TNF alpha)[195] and interleukin-1 (IL-1)[196].

The initial activation of the HIV LTR is a consequence of inducible and constitutive cellular transcription factors. Activation of the LTR by cellular transcription factors leads primarily to the generation of short transcripts up to a "strongstop" signal[197]. However, a low level of read through allows the production of the Tat protein. The trans-acting regulatory protein Tat then interacts with the TAR element in the U5 region of the 5' LTR to greatly increase the levels of long transcription of viral RNAs, and the elongation capacity of RNA polymerase II is improved[197,198]. The Tat protein thus plays a key role in the activation and maintenance of high levels of transcription from the proviral DNA. Since synthesis of Tat is itself under the control of LTR promoters, HIV transcription results in a positive feedback.

mRNA splicing and cellular localisation:

The primary HIV-1 transcript contains multiple splice donors (5' splice sites) and splice acceptors (3' splice sites), which can be processed to yield more than 30 alternative mRNAs[192]. Many of the mRNAs are polycistronic; i.e., they contain the open reading frame of more than one protein. The polycistronic mRNAs typically express a single gene product. Open reading frame choice is governed by the efficiency of the initiation codon and the proximity of the initiation codon to the 5' end of the mRNA[199].

Tat and Rev are the two transcripts of HIV (together with Nef) to be fully spliced and so they are exported to the cytoplasm soon after host cell infection. As an initial event after cellular infection, intron-containing RNAs must be completely spliced before they can exit the nucleus. This regulation is essential because it prevents the translation of intronic sequences contained in partially spliced mRNAs. Rev binds to a secondary structure of unspliced or partially spliced mRNAs, the RRE, and allows their export from the nucleus (Rev has both nuclear import and export signals)[179]. This export allows those viral RNAs to bypass the normal "check point" of RNA splicing. The fully spliced viral mRNAs exit the nucleus by using the export pathway followed by the majority of cellular mRNAs.

HIV-1 mRNAs fall into three size classes:

1. Fully spliced RNA. These mRNAs have spliced out both introns of HIV and have the potential to express Nef, and the two-exon forms of Tat and Rev. These heterogeneous mRNAs do not require the expression of the Rev protein for export and translation.
2. Incompletely spliced RNA. These RNAs can potentially express Env, Vif, Vpu, Vpr, and the single-exon form of Tat. These heterogeneous mRNAs are 4- to 5-kb long and retain the second intron of HIV.
3. Unspliced RNA. The unspliced 9-kb primary transcript can be expressed to generate the Gag and Gag-Pol precursor proteins or be packaged into virions to serve as the genomic RNA.

Whereas the proteins encoded by the fully spliced mRNAs, Nef, Tat, and Rev, can be produced immediately, and are thus early viral gene products, threshold levels of Rev are necessary for exporting intron-containing HIV mRNAs, explaining why those encode the viral late gene products[177,200].

❖ HIV tropism

The term *viral tropism* refers to the cell types which get infected by HIV. HIV can infect a variety of immune cells such as CD4+ T cells, macrophages, and microglial cells. HIV-1 entry to macrophages and CD4+ T cells is mediated through interaction of the virion envelope glycoproteins (gp120) with the CD4 molecule on the target cells and, in addition, with a chemokine coreceptor: generally CCR5 or CXCR4 (fusin)[201].

The former nomenclature of Macrophage strains (M-tropic) of HIV-1, or non-syncitia-inducing strains (NSI), indicated viruses that use the β-chemokine receptor CCR5 for entry (and sometimes the CCR3 coreceptor). They are able to replicate in the macrophages and in the primary $CD4^+$ T cells[202]. This CCR5 coreceptor is used by almost all primary HIV-1 isolates regardless of the viral genetic subtype. The macrophages appear to be the first cells infected by HIV and perhaps the source of HIV production when $CD4^+$ cells become depleted in the patient. Macrophages and microglial cells are the cells infected by HIV in the central nervous system.

T-tropic isolates, or syncitia-inducing (SI) strains, replicate in primary CD4+ T cells and use the α-chemokine receptor, CXCR4, for entry[202-204]. They can also enter macrophages by using both CCR5 and CXCR4 coreceptors.

Dual-tropic HIV-1 strains are thought to be transitional strains of the HIV-1 virus and thus are able to use either CCR5 or CXCR4 as coreceptors for viral entry.

HIV that use only the CCR5 receptor are today termed R5-tropic, those that only use CXCR4 are termed X4-tropic, and those that use both, X4R5.

Sexual intercourse is the major mode of HIV transmission. Both X4- and R5-tropic HIV are represented in the genital fluid which is passed from partner to partner. The virions can then infect numerous cellular targets and disseminate into the organism. However, a selection process leads to a predominant transmission of the R5 virus through this pathway[205-207]. How this selective process works is still under investigation. In patients infected with subtype B HIV-1, there is often a coreceptor switch in late-stage disease and T-tropic variants appear that can infect a variety of T cells through CXCR4[208]. These variants then replicate

more aggressively with greater virulence that causes rapid T cell depletion, immune system collapse, and opportunistic infections that mark the advent of AIDS[209]. Thus, during the course of infection, viral adaptation to the use of CXCR4 instead of CCR5 may be a key step in the progression to AIDS. A number of studies with subtype B-infected individuals have determined that between 40 and 50% of AIDS patients can harbour viruses of the SI, and presumably the X4, phenotype[210,211].

The host genes, as the chemokine receptors, can interfere in the course of AIDS. People homozygous for a 32bp deletion (the so-called Δ32 mutation) in the chemokine receptor gene for CCR5, which results in failure of surface expression of this key viral coreceptor, are protected nearly at 100% from an HIV infection, considering that transmitted viruses are mainly CCR5-tropic[212]. Among HIV seronegative individuals, who are at high risk of infection, only a minority have a homozygous Δ32 mutation, indicating that other mechanisms determine protection from HIV infection. Single-nucleotide polymorphisms (SNPs) in the *CCR5* promoter have shown that alleles containing an adenine at position -2459 display higher *CCR5* promoter activities than alleles with a guanine. HIV infected individuals homozygous for the A allele (*-2459A*/*A*) progressed more rapidly to AIDS than those who were homozygous for the G allele (*-2459G*/*G*)[213]. A C-to-G substitution at position -28 in the promoter of RANTES, a ligand of the CCR5 coreceptor, supports increased expression of RANTES and is associated with a decreased rate of $CD4^+$ T-cell decline in HIV infected people[214]. Moreover, people homozygous for a non-coding sequence in the gene for stromal cell-derived factor 1 (SDF-1), the natural ligand for the CXCR4 coreceptor, were found to have a delayed risk for progression to AIDS[215]. Certain human leucocyte antigen (HLA) alleles indicate greater or lesser risks of disease progression; in addition viral mutations that impair the binding of specific viral peptides to HLA of a given patient are correlated with a higher VL[216].

In general, variation in the host immune system could explain the difference in disease progression observed among different patients, but also the enormous adaptability of the viral envelope, especially the variation of its glycosylation patterns, is responsible for immune escape[217,218].

❖ AntiRetroVirals

Nucleoside/Nucleotide Reverse Transcriptase Inhibitors (NRTIs)

Brand Name	Generic Name(s)	Structure	Manufacturer Name	Approval Date
Retrovir	zidovudine, azidothymidine, AZT, ZDV	Thymidine Analog	GlaxoSmithKline	19-Mar-87
Videx	didanosine, dideoxyinosine, ddI	Inosine Analog	Bristol Myers-Squibb	9-Oct-91
Hivid	zalcitabine, dideoxycytidine, ddC	Cytidine Analog	Hoffmann-La Roche	19-Jun-92
Zerit	stavudine, d4T	Thymidine Analog	Bristol Myers-Squibb	24-Jun-94
Epivir	lamivudine, 3TC	Cytidine Analog	GlaxoSmithKline	17-Nov-95
Ziagen	abacavir sulfate, ABC	Guanosine Analog	GlaxoSmithKline	17-Dec-98
Viread	tenofovir disoproxil fumarate, TDF	Adenine Analog	Gilead	26-Oct-01
Emtriva	emtricitabine, FTC	Cytidine Analog	Gilead Sciences	02-Jul-03

Nonnucleoside Reverse Transcriptase Inhibitors (NNRTIs)

Brand Name	Generic Name	Manufacturer Name	Approval Date
Viramune	nevirapine, NVP	Boehringer Ingelheim	21-Jun-96
Rescriptor	delavirdine, DLV	Pfizer	4-Apr-97
Sustiva	efavirenz, EFV	Bristol Myers-Squibb	17-Sep-98

Protease Inhibitors (PIs)

Brand Name	Generic Name(s)	Manufacturer Name	Approval Date
Invirase	saquinavir mesylate, SQV	Hoffmann-La Roche	6-Dec-95
Norvir	ritonavir, RTV	Abbott Laboratories	1-Mar-96
Crixivan	indinavir, IDV,	Merck	13-Mar-96
Viracept	nelfinavir mesylate, NFV	Agouron Pharmaceuticals	14-Mar-97
Agenerase	amprenavir, APV	GlaxoSmithKline	15-Apr-99
Kaletra	lopinavir and ritonavir, LPV/RTV	Abbott Laboratories	15-Sep-00
Reyataz	atazanavir sulfate, ATV	Bristol-Myers Squibb	20-Jun-03
Lexiva	Fosamprenavir Calcium, FOS-APV	GlaxoSmithKline	20-Oct-03
Aptivus	tipranavir, TPV	Boehringer Ingelheim	22-Jun-05
Prezista	Darunavir, DNV	Tibotec, Inc.	23-Jun-06

Fusion Inhibitors

Brand Name	Generic Name	Manufacturer Name	Approval Date
Fuzeon	enfuvirtide, T-20	Hoffmann-La Roche & Trimeris	13-Mar-03

Entry Inhibitors - CCR5 co-receptor antagonist

Brand Name	Generic Names	Manufacturer Name	Approval Date
Selzentry	maraviroc	Pfizer	06-August-07

References

1. Pneumocystis pneumonia--Los Angeles. *MMWR Morb Mortal Wkly Rep.* 1981; 30: 250-2.
2. Masur H, Michelis MA, Greene JB, et al. An outbreak of community-acquired Pneumocystis carinii pneumonia: initial manifestation of cellular immune dysfunction. *N Engl J Med.* 1981; 305: 1431-8.
3. Update on acquired immune deficiency syndrome (AIDS)--United States. *MMWR Morb Mortal Wkly Rep.* 1982; 31: 507-8, 13-4.
4. Sonnabend J, Witkin SS and Purtilo DT. Acquired immunodeficiency syndrome, opportunistic infections, and malignancies in male homosexuals. A hypothesis of etiologic factors in pathogenesis. *Jama.* 1983; 249: 2370-4.
5. Barre-Sinoussi F, Chermann JC, Rey F, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science.* 1983; 220: 868-71.
6. Popovic M, Sarngadharan MG, Read E, et al. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science.* 1984; 224: 497-500.
7. Gallo RC, Salahuddin SZ, Popovic M, et al. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science.* 1984; 224: 500-3.
8. Coffin J, Haase A, Levy JA, et al. What to call the AIDS virus? *Nature.* 1986; 321: 10.
9. Guyader M, Emerman M, Sonigo P, et al. Genome organization and transactivation of the human immunodeficiency virus type 2. *Nature.* 1987; 326: 662-9.
10. Huminer D, Rosenfeld JB and Pitlik SD. AIDS in the pre-AIDS era. *Rev Infect Dis.* 1987; 9: 1102-8.
11. Garry RF, Witte MH, Gottlieb AA, et al. Documentation of an AIDS virus infection in the United States in 1968. *Jama.* 1988; 260: 2085-7.
12. Greener R. Why companies should intervene: a case study of the costs of HIV / AIDS to employers. Economic impact: Southern Africa. *AIDS Anal Afr.* 1998; 8: 3-4.
13. Piatak M, Jr., Saag MS, Yang LC, et al. High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. *Science.* 1993; 259: 1749-54.
14. Pantaleo G, Demarest JF, Schacker T, et al. The qualitative nature of the primary immune response to HIV infection is a prognosticator of disease progression independent of the initial level of plasma viremia. *Proc Natl Acad Sci U S A.* 1997; 94: 254-8.
15. Burton GF, Keele BF, Estes JD, et al. Follicular dendritic cell contributions to HIV pathogenesis. *Semin Immunol.* 2002; 14: 275-84.
16. Pedraza MA, del Romero J, Roldan F, et al. Heterosexual transmission of HIV-1 is associated with high plasma viral load levels and a positive viral isolation in the infected partner. *J Acquir Immune Defic Syndr.* 1999; 21: 120-5.
17. Garcia PM, Kalish LA, Pitt J, et al. Maternal levels of plasma human immunodeficiency virus type 1 RNA and the risk of perinatal transmission. Women and Infants Transmission Study Group. *N Engl J Med.* 1999; 341: 394-402.
18. Chan DC and Kim PS. HIV entry and its inhibition. *Cell.* 1998; 93: 681-4.
19. Wyatt R and Sodroski J. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science.* 1998; 280: 1884-8.
20. Zheng YH, Lovsin N and Peterlin BM. Newly identified host factors modulate HIV replication. *Immunol Lett.* 2005; 97: 225-34.
21. Shafer RW, Hertogs K, Zolopa AR, et al. High degree of interlaboratory reproducibility of human immunodeficiency virus type 1 protease and reverse transcriptase sequencing of plasma samples from heavily treated patients. *J Clin Microbiol.* 2001; 39: 1522-9.
22. Hiscott J, Kwon H and Genin P. Hostile takeovers: viral appropriation of the NF-kappaB pathway. *J Clin Invest.* 2001; 107: 143-51.
23. Pollard VW and Malim MH. The HIV-1 Rev protein. *Annu Rev Microbiol.* 1998; 52: 491-532.
24. Mitsuya H, Weinhold KJ, Furman PA, et al. 3'-Azido-3'-deoxythymidine (BW A509U): an antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus in vitro. *Proc Natl Acad Sci U S A.* 1985; 82: 7096-100.
25. Yeni PG, Hammer SM, Carpenter CC, et al. Antiretroviral treatment for adult HIV infection in 2002: updated recommendations of the International AIDS Society-USA Panel. *Jama.* 2002; 288: 222-35.
26. Pomerantz RJ and Horn DL. Twenty years of therapy for HIV-1 infection. *Nat Med.* 2003; 9: 867-73.
27. Wild CT, Shugars DC, Greenwell TK, et al. Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. *Proc Natl Acad Sci U S A.* 1994; 91: 9770-4.
28. Kilby JM, Hopkins S, Venetta TM, et al. Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. *Nat Med.* 1998; 4: 1302-7.

29. Haase AT. Population biology of HIV-1 infection: viral and CD4+ T cell demographics and dynamics in lymphatic tissues. *Annu Rev Immunol.* 1999; 17: 625-56.
30. Preston BD, Poiesz BJ and Loeb LA. Fidelity of HIV-1 reverse transcriptase. *Science.* 1988; 242: 1168-71.
31. Roberts JD, Bebenek K and Kunkel TA. The accuracy of reverse transcriptase from HIV-1. *Science.* 1988; 242: 1171-3.
32. Darwin C. The origin of species by means of natural selection: Or, the preservation of favored races in the struggle for life. *John Murray publishers: London.* 1859:
33. Siliciano JD and Siliciano RF. A long-term latent reservoir for HIV-1: discovery and clinical implications. *J Antimicrob Chemother.* 2004; 54: 6-9.
34. Clavel F and Hance AJ. HIV drug resistance. *N Engl J Med.* 2004; 350: 1023-35.
35. Yerly S, Kaiser L, Race E, et al. Transmission of antiretroviral-drug-resistant HIV-1 variants. *Lancet.* 1999; 354: 729-33.
36. Little SJ, Holte S, Routy JP, et al. Antiretroviral-drug resistance among patients recently infected with HIV. *N Engl J Med.* 2002; 347: 385-94.
37. Grant RM, Hecht FM, Warmerdam M, et al. Time trends in primary HIV-1 drug resistance among recently infected persons. *Jama.* 2002; 288: 181-8.
38. Witvrouw M, Pannecouque C, Van Laethem K, et al. Activity of non-nucleoside reverse transcriptase inhibitors against HIV-2 and SIV. *Aids.* 1999; 13: 1477-83.
39. Shafer RW, Eisen JA, Merigan TC, et al. Sequence and drug susceptibility of subtype C reverse transcriptase from human immunodeficiency virus type 1 seroconverters in Zimbabwe. *J Virol.* 1997; 71: 5441-8.
40. Descamps D, Collin G, Letourneur F, et al. Susceptibility of human immunodeficiency virus type 1 group O isolates to antiretroviral agents: in vitro phenotypic and genotypic analyses. *J Virol.* 1997; 71: 8893-8.
41. Palmer S, Alaeus A, Albert J, et al. Drug susceptibility of subtypes A,B,C,D, and E human immunodeficiency virus type 1 primary isolates. *AIDS Res Hum Retroviruses.* 1998; 14: 157-62.
42. Descamps D, Apetrei C, Collin G, et al. Naturally occurring decreased susceptibility of HIV-1 subtype G to protease inhibitors. *Aids.* 1998; 12: 1109-11.
43. Pozniak AL, Miller R and Ormerod LP. The treatment of tuberculosis in HIV-infected persons. *Aids.* 1999; 13: 435-45.
44. Boucher CA, Cammack N, Schipper P, et al. High-level resistance to (-) enantiomeric 2'-deoxy-3'-thiacytidine in vitro is due to one amino acid substitution in the catalytic site of human immunodeficiency virus type 1 reverse transcriptase. *Antimicrob Agents Chemother.* 1993; 37: 2231-4.
45. Sarafianos SG, Das K, Clark AD, Jr., et al. Lamivudine (3TC) resistance in HIV-1 reverse transcriptase involves steric hindrance with beta-branched amino acids. *Proc Natl Acad Sci U S A.* 1999; 96: 10027-32.
46. Iversen AK, Shafer RW, Wehrly K, et al. Multidrug-resistant human immunodeficiency virus type 1 strains resulting from combination antiretroviral therapy. *J Virol.* 1996; 70: 1086-90.
47. Kosalaraksa P, Kavlick MF, Maroun V, et al. Comparative fitness of multi-dideoxynucleoside-resistant human immunodeficiency virus type 1 (HIV-1) in an In vitro competitive HIV-1 replication assay. *J Virol.* 1999; 73: 5356-63.
48. Parikh UM, Bacheler L, Koontz D, et al. The K65R mutation in human immunodeficiency virus type 1 reverse transcriptase exhibits bidirectional phenotypic antagonism with thymidine analog mutations. *J Virol.* 2006; 80: 4971-7.
49. Antinori A, Trotta MP, Nasta P, et al. Antiviral efficacy and genotypic resistance patterns of combination therapy with stavudine/tenofovir in highly active antiretroviral therapy experienced patients. *Antivir Ther.* 2006; 11: 233-43.
50. Coakley EP, Gillis JM and Hammer SM. Phenotypic and genotypic resistance patterns of HIV-1 isolates derived from individuals treated with didanosine and stavudine. *Aids.* 2000; 14: F9-15.
51. Picard V, Angelini E, Maillard A, et al. Comparison of genotypic and phenotypic resistance patterns of human immunodeficiency virus type 1 isolates from patients treated with stavudine and didanosine or zidovudine and lamivudine. *J Infect Dis.* 2001; 184: 781-4.
52. Shafer RW, Winters MA, Jellinger RM, et al. Zidovudine resistance reverse transcriptase mutations during didanosine monotherapy. *J Infect Dis.* 1996; 174: 448-9.
53. Larder BA and Kemp SD. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science.* 1989; 246: 1155-8.
54. Boucher CA, O'Sullivan E, Mulder JW, et al. Ordered appearance of zidovudine resistance mutations during treatment of 18 human immunodeficiency virus-positive subjects. *J Infect Dis.* 1992; 165: 105-10.
55. Richman DD. Susceptibility to nucleoside analogues of zidovudine-resistant isolates of human immunodeficiency virus. *Am J Med.* 1990; 88: 8S-10S.
56. Meyer PR, Matsuura SE, Mian AM, et al. A mechanism of AZT resistance: an increase in nucleotide-dependent primer unblocking by mutant HIV-1 reverse transcriptase. *Mol Cell.* 1999; 4: 35-43.

57. Arion D, Kaushik N, McCormick S, et al. Phenotypic mechanism of HIV-1 resistance to 3'-azido-3'-deoxythymidine (AZT): increased polymerization processivity and enhanced sensitivity to pyrophosphate of the mutant viral reverse transcriptase. *Biochemistry.* 1998; 37: 15908-17.
58. Boyer PL, Sarafianos SG, Arnold E, et al. Selective excision of AZTMP by drug-resistant human immunodeficiency virus reverse transcriptase. *J Virol.* 2001; 75: 4832-42.
59. Chamberlain PP, Ren J, Nichols CE, et al. Crystal structures of Zidovudine- or Lamivudine-resistant human immunodeficiency virus type 1 reverse transcriptases containing mutations at codons 41, 184, and 215. *J Virol.* 2002; 76: 10015-9.
60. Ren J, Nichols C, Bird L, et al. Structural mechanisms of drug resistance for mutations at codons 181 and 188 in HIV-1 reverse transcriptase and the improved resilience of second generation non-nucleoside inhibitors. *J Mol Biol.* 2001; 312: 795-805.
61. Joly V and Yeni P. [Non-nucleoside reverse transcriptase inhibitors]. *Ann Med Interne (Paris).* 2000; 151: 260-7.
62. Zhou Z, Madrid M, Evanseck JD, et al. Effect of a bound non-nucleoside RT inhibitor on the dynamics of wild-type and mutant HIV-1 reverse transcriptase. *J Am Chem Soc.* 2005; 127: 17253-60.
63. Weinstein MC, Goldie SJ, Losina E, et al. Use of genotypic resistance testing to guide hiv therapy: clinical impact and cost-effectiveness. *Ann Intern Med.* 2001; 134: 440-50.
64. Wilson JW and Bean P. A physician's primer to antiretroviral drug resistance testing. *AIDS Read.* 2000; 10: 469-73, 76-8.
65. Wilson JW. Update on antiretroviral drug resistance testing: combining laboratory technology with patient care. *AIDS Read.* 2003; 13: 25-30, 35-8.
66. Youree BE and D'Aquila RT. Antiretroviral resistance testing for clinical management. *AIDS Rev.* 2002; 4: 3-12.
67. Hanna GJ and D'Aquila RT. Clinical use of genotypic and phenotypic drug resistance testing to monitor antiretroviral chemotherapy. *Clin Infect Dis.* 2001; 32: 774-82.
68. Rhee SY, Gonzales MJ, Kantor R, et al. Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucleic Acids Res.* 2003; 31: 298-303.
69. Hirsch MS, Brun-Vezinet F, D'Aquila RT, et al. Antiretroviral drug resistance testing in adult HIV-1 infection: recommendations of an International AIDS Society-USA Panel. *Jama.* 2000; 283: 2417-26.
70. Japour AJ, Mayers DL, Johnson VA, et al. Standardized peripheral blood mononuclear cell culture assay for determination of drug susceptibilities of clinical human immunodeficiency virus type 1 isolates. The RV-43 Study Group, the AIDS Clinical Trials Group Virology Committee Resistance Working Group. *Antimicrob Agents Chemother.* 1993; 37: 1095-101.
71. Kellam P and Larder BA. Recombinant virus assay: a rapid, phenotypic assay for assessment of drug susceptibility of human immunodeficiency virus type 1 isolates. *Antimicrob Agents Chemother.* 1994; 38: 23-30.
72. Hertogs K, de Bethune MP, Miller V, et al. A rapid method for simultaneous detection of phenotypic resistance to inhibitors of protease and reverse transcriptase in recombinant human immunodeficiency virus type 1 isolates from patients treated with antiretroviral drugs. *Antimicrob Agents Chemother.* 1998; 42: 269-76.
73. Petropoulos CJ, Parkin NT, Limoli KL, et al. A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. *Antimicrob Agents Chemother.* 2000; 44: 920-8.
74. Zennou V, Mammano F, Paulous S, et al. Loss of viral fitness associated with multiple Gag and Gag-Pol processing defects in human immunodeficiency virus type 1 variants selected for resistance to protease inhibitors in vivo. *J Virol.* 1998; 72: 3300-6.
75. Keulen W, Back NK, van Wijk A, et al. Initial appearance of the 184Ile variant in lamivudine-treated patients is caused by the mutational bias of human immunodeficiency virus type 1 reverse transcriptase. *J Virol.* 1997; 71: 3346-50.
76. Croteau G, Doyon L, Thibeault D, et al. Impaired fitness of human immunodeficiency virus type 1 variants with high-level resistance to protease inhibitors. *J Virol.* 1997; 71: 1089-96.
77. Vandamme AM, Sonnerborg A, Ait-Khaled M, et al. Updated European recommendations for the clinical use of HIV drug resistance testing. *Antivir Ther.* 2004; 9: 829-48.
78. Meynard JL, Vray M, Morand-Joubert L, et al. Phenotypic or genotypic resistance testing for choosing antiretroviral therapy after treatment failure: a randomized trial. *Aids.* 2002; 16: 727-36.
79. Wegner SA, Wallace MR, Aronson NE, et al. Long-term efficacy of routine access to antiretroviral-resistance testing in HIV type 1-infected patients: results of the clinical efficacy of resistance testing trial. *Clin Infect Dis.* 2004; 38: 723-30.
80. Durant J, Clevenbergh P, Halfon P, et al. Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial. *Lancet.* 1999; 353: 2195-9.

81. Baxter JD, Mayers DL, Wentworth DN, et al. A randomized study of antiretroviral management based on plasma genotypic antiretroviral resistance testing in patients failing therapy. CPCRA 046 Study Team for the Terry Beirn Community Programs for Clinical Research on AIDS. *Aids.* 2000; 14: F83-93.
82. Tural C, Ruiz L, Holtzer C, et al. Clinical utility of HIV-1 genotyping and expert advice: the Havana trial. *Aids.* 2002; 16: 209-18.
83. Cingolani A, Antinori A, Rizzo MG, et al. Usefulness of monitoring HIV drug resistance and adherence in individuals failing highly active antiretroviral therapy: a randomized study (ARGENTA). *Aids.* 2002; 16: 369-79.
84. Cohen CJ, Hunt S, Sension M, et al. A randomized trial assessing the impact of phenotypic resistance testing on antiretroviral therapy. *Aids.* 2002; 16: 579-88.
85. Haubrich RH, Kemper CA, Hellmann NS, et al. A randomized, prospective study of phenotype susceptibility testing versus standard of care to manage antiretroviral therapy: CCTG 575. *Aids.* 2005; 19: 295-302.
86. Parkin N, Chappey C, Maroldo L, et al. Phenotypic and genotypic HIV-1 drug resistance assays provide complementary information. *J Acquir Immune Defic Syndr.* 2002; 31: 128-36.
87. Gallego O, Martin-Carbonero L, Aguero J, et al. Correlation between rules-based interpretation and virtual phenotype interpretation of HIV-1 genotypes for predicting drug resistance in HIV-infected individuals. *J Virol Methods.* 2004; 121: 115-8.
88. Adachi A, Gendelman HE, Koenig S, et al. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J Virol.* 1986; 59: 284-91.
89. Korn K, Reil H, Walter H, et al. Quality control trial for human immunodeficiency virus type 1 drug resistance testing using clinical samples reveals problems with detecting minority species and interpretation of test results. *J Clin Microbiol.* 2003; 41: 3559-65.
90. Hance AJ, Lemiale V, Izopet J, et al. Changes in human immunodeficiency virus type 1 populations after treatment interruption in patients failing antiretroviral therapy. *J Virol.* 2001; 75: 6410-7.
91. Finel V, Hubert S, Louvel S, et al. Evolution of genotyping algorithm and comparison to phenotyping HIV resistance over the last three years. Paper presented at: 2nd Swiss Workshop on Basic HIV Research, 2006; Murten.
92. Ravela J, Betts BJ, Brun-Vezinet F, et al. HIV-1 protease and reverse transcriptase mutation patterns responsible for discordances between genotypic drug resistance interpretation algorithms. *J Acquir Immune Defic Syndr.* 2003; 33: 8-14.
93. Vergne L, Snoeck J, Aghokeng A, et al. Genotypic drug resistance interpretation algorithms display high levels of discordance when applied to non-B strains from HIV-1 naive and treated patients. *FEMS Immunol Med Microbiol.* 2006; 46: 53-62.
94. Snoeck J, Kantor R, Shafer RW, et al. Discordances between interpretation algorithms for genotypic resistance to protease and reverse transcriptase inhibitors of human immunodeficiency virus are subtype dependent. *Antimicrob Agents Chemother.* 2006; 50: 694-701.
95. Hubert S, Louvel S, Brondani V, et al. Resistance prediction for non-B subtype HIV1. Paper presented at: 3rd European HIV drug resistance workshop, 2005; Athens.
96. Coffin JM. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science.* 1995; 267: 483-9.
97. Kijak GH, Rubio AE, Pampuro SE, et al. Discrepant results in the interpretation of HIV-1 drug-resistance genotypic data among widely used algorithms. *HIV Med.* 2003; 4: 72-8.
98. Sturmer M, Doerr HW, Staszewski S, et al. Comparison of nine resistance interpretation systems for HIV-1 genotyping. *Antivir Ther.* 2003; 8: 239-44.
99. Hu Z, Giguel F, Hatano H, et al. Fitness comparison of thymidine analog resistance pathways in human immunodeficiency virus type 1. *J Virol.* 2006; 80: 7020-7.
100. Marcelin AG, Delaugerre C, Wirden M, et al. Thymidine analogue reverse transcriptase inhibitors resistance mutations profiles and association to other nucleoside reverse transcriptase inhibitors resistance mutations observed in the context of virological failure. *J Med Virol.* 2004; 72: 162-5.
101. Hanna GJ, Johnson VA, Kuritzkes DR, et al. Patterns of resistance mutations selected by treatment of human immunodeficiency virus type 1 infection with zidovudine, didanosine, and nevirapine. *J Infect Dis.* 2000; 181: 904-11.
102. Yahi N, Tamalet C, Tourres C, et al. Mutation patterns of the reverse transcriptase and protease genes in human immunodeficiency virus type 1-infected patients undergoing combination therapy: survey of 787 sequences. *J Clin Microbiol.* 1999; 37: 4099-106.
103. Larder BA. Interactions between drug resistance mutations in human immunodeficiency virus type 1 reverse transcriptase. *J Gen Virol.* 1994; 75 (Pt 5): 951-7.

104. Sluis-Cremer N, Arion D and Parniak MA. Molecular mechanisms of HIV-1 resistance to nucleoside reverse transcriptase inhibitors (NRTIs). *Cell Mol Life Sci.* 2000; 57: 1408-22.
105. Goldschmidt V and Marquet R. Primer unblocking by HIV-1 reverse transcriptase and resistance to nucleoside RT inhibitors (NRTIs). *Int J Biochem Cell Biol.* 2004; 36: 1687-705.
106. St Clair MH, Martin JL, Tudor-Williams G, et al. Resistance to ddI and sensitivity to AZT induced by a mutation in HIV-1 reverse transcriptase. *Science.* 1991; 253: 1557-9.
107. Sharma PL and Crumpacker CS. Attenuated replication of human immunodeficiency virus type 1 with a didanosine-selected reverse transcriptase mutation. *J Virol.* 1997; 71: 8846-51.
108. Miranda LR, Gotte M, Liang F, et al. The L74V mutation in human immunodeficiency virus type 1 reverse transcriptase counteracts enhanced excision of zidovudine monophosphate associated with thymidine analog resistance mutations. *Antimicrob Agents Chemother.* 2005; 49: 2648-56.
109. Larder BA. 3'-Azido-3'-deoxythymidine resistance suppressed by a mutation conferring human immunodeficiency virus type 1 resistance to nonnucleoside reverse transcriptase inhibitors. *Antimicrob Agents Chemother.* 1992; 36: 2664-9.
110. Byrnes VW, Emini EA, Schleif WA, et al. Susceptibilities of human immunodeficiency virus type 1 enzyme and viral variants expressing multiple resistance-engendering amino acid substitutions to reserve transcriptase inhibitors. *Antimicrob Agents Chemother.* 1994; 38: 1404-7.
111. Selmi B, Deval J, Alvarez K, et al. The Y181C substitution in 3'-azido-3'-deoxythymidine-resistant human immunodeficiency virus, type 1, reverse transcriptase suppresses the ATP-mediated repair of the 3'-azido-3'-deoxythymidine 5'-monophosphate-terminated primer. *J Biol Chem.* 2003; 278: 40464-72.
112. White KL, Chen JM, Feng JY, et al. The K65R reverse transcriptase mutation in HIV-1 reverses the excision phenotype of zidovudine resistance mutations. *Antivir Ther.* 2006; 11: 155-63.
113. Boyer PL, Sarafianos SG, Arnold E, et al. The M184V mutation reduces the selective excision of zidovudine 5'-monophosphate (AZTMP) by the reverse transcriptase of human immunodeficiency virus type 1. *J Virol.* 2002; 76: 3248-56.
114. Tisdale M, Kemp SD, Parry NR, et al. Rapid in vitro selection of human immunodeficiency virus type 1 resistant to 3'-thiacytidine inhibitors due to a mutation in the YMDD region of reverse transcriptase. *Proc Natl Acad Sci U S A.* 1993; 90: 5653-6.
115. Mansky LM. Retrovirus mutation rates and their role in genetic variation. *J Gen Virol.* 1998; 79 (Pt 6): 1337-45.
116. Hu WS and Temin HM. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc Natl Acad Sci U S A.* 1990; 87: 1556-60.
117. Robertson DL, Hahn BH and Sharp PM. Recombination in AIDS viruses. *J Mol Evol.* 1995; 40: 249-59.
118. Smith DM, Richman DD and Little SJ. HIV superinfection. *J Infect Dis.* 2005; 192: 438-44.
119. Molla A, Korneyeva M, Gao Q, et al. Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. *Nat Med.* 1996; 2: 760-6.
120. Condra JH, Holder DJ, Schleif WA, et al. Genetic correlates of in vivo viral resistance to indinavir, a human immunodeficiency virus type 1 protease inhibitor. *J Virol.* 1996; 70: 8270-6.
121. Schuurman R, Nijhuis M, van Leeuwen R, et al. Rapid changes in human immunodeficiency virus type 1 RNA load and appearance of drug-resistant virus populations in persons treated with lamivudine (3TC). *J Infect Dis.* 1995; 171: 1411-9.
122. Quan Y, Gu Z, Li X, et al. Endogenous reverse transcription assays reveal high-level resistance to the triphosphate of (-)2'-dideoxy-3'-thiacytidine by mutated M184V human immunodeficiency virus type 1. *J Virol.* 1996; 70: 5642-5.
123. Clotet B. Efavirenz: resistance and cross-resistance. *Int J Clin Pract Suppl.* 1999; 103: 21-5.
124. Casado JL, Hertogs K, Ruiz L, et al. Non-nucleoside reverse transcriptase inhibitor resistance among patients failing a nevirapine plus protease inhibitor-containing regimen. *Aids.* 2000; 14: F1-7.
125. Bacheler L, Jeffrey S, Hanna G, et al. Genotypic correlates of phenotypic resistance to efavirenz in virus isolates from patients failing nonnucleoside reverse transcriptase inhibitor therapy. *J Virol.* 2001; 75: 4999-5008.
126. Frost SD, Nijhuis M, Schuurman R, et al. Evolution of lamivudine resistance in human immunodeficiency virus type 1-infected individuals: the relative roles of drift and selection. *J Virol.* 2000; 74: 6262-8.
127. Sharma PL and Crumpacker CS. Decreased processivity of human immunodeficiency virus type 1 reverse transcriptase (RT) containing didanosine-selected mutation Leu74Val: a comparative analysis of RT variants Leu74Val and lamivudine-selected Met184Val. *J Virol.* 1999; 73: 8448-56.
128. Back NK, Nijhuis M, Keulen W, et al. Reduced replication of 3TC-resistant HIV-1 variants in primary cells due to a processivity defect of the reverse transcriptase enzyme. *Embo J.* 1996; 15: 4040-9.
129. Uhlmann EJ, Tebas P, Storch GA, et al. Effects of the G190A substitution of HIV reverse transcriptase on phenotypic susceptibility of patient isolates to delavirdine. *J Clin Virol.* 2004; 31: 198-203.

130. Stup TL, Pallansch LA, Buckheit RW. The Consequence of Resistance-Engendering Mutations on In Vitro Replication and Fitness of HIV-1. Paper presented at: 39th Interscience Conference on Antimicrobial Agents and Chemotherapy, 1999; Toronto.
131. Huang W, Parkin NT, Lie YS, et al. A Novel HIV-1 RT Mutation (M230L) Confers NNRTI Resistance and Dose-Dependent Stimulation of Replication. Paper presented at: 40th Interscience Conference on Antimicrobial Agents and Chemotherapy, 2000; Toronto.
132. Charpentier C, Dwyer DE, Mammano F, et al. Role of minority populations of human immunodeficiency virus type 1 in the evolution of viral resistance to protease inhibitors. *J Virol.* 2004; 78: 4234-47.
133. Roquebert B, Malet I, Wirden M, et al. Role of HIV-1 minority populations on resistance mutational pattern evolution and susceptibility to protease inhibitors. *Aids.* 2006; 20: 287-9.
134. Dykes C, Najjar J, Bosch RJ, et al. Detection of drug-resistant minority variants of HIV-1 during virologic failure of indinavir, lamivudine, and zidovudine. *J Infect Dis.* 2004; 189: 1091-6.
135. Lecossier D, Shulman NS, Morand-Joubert L, et al. Detection of minority populations of HIV-1 expressing the K103N resistance mutation in patients failing nevirapine. *J Acquir Immune Defic Syndr.* 2005; 38: 37-42.
136. Mellors J, Palmer S, Nissley D, et al. Low frequency NNRTI-resistant variants contribute to failure of efavirenz-containing regimens. Paper presented at: HIV DRP Symposium on Antiviral Drug Resistance, 2003; Chantilly, Virginia.
137. Gunthard HF, Wong JK, Ignacio CC, et al. Comparative performance of high-density oligonucleotide sequencing and dideoxynucleotide sequencing of HIV type 1 pol from clinical samples. *AIDS Res Hum Retroviruses.* 1998; 14: 869-76.
138. Schuurman R, Brambilla D, de Groot T, et al. Underestimation of HIV type 1 drug resistance mutations: results from the ENVA-2 genotyping proficiency program. *AIDS Res Hum Retroviruses.* 2002; 18: 243-8.
139. Detsika MG, Chandler B, Khoo SH, et al. Detection and quantification of minority HIV isolates harbouring the D30N mutation by real-time PCR amplification. *J Antimicrob Chemother.* 2007; 60: 881-4.
140. Perez-Bercoff D, Wurtzer S, Compain S, et al. Human immunodeficiency virus type 1: resistance to nucleoside analogues and replicative capacity in primary human macrophages. *J Virol.* 2007; 81: 4540-50.
141. Weber J, Chakraborty B, Weberova J, et al. Diminished replicative fitness of primary human immunodeficiency virus type 1 isolates harboring the K65R mutation. *J Clin Microbiol.* 2005; 43: 1395-400.
142. White KL, Margot NA, Wrin T, et al. Molecular mechanisms of resistance to human immunodeficiency virus type 1 with reverse transcriptase mutations K65R and K65R+M184V and their effects on enzyme function and viral replication capacity. *Antimicrob Agents Chemother.* 2002; 46: 3437-46.
143. Cane PA. Stability of transmitted drug-resistant HIV-1 species. *Curr Opin Infect Dis.* 2005; 18: 537-42.
144. Balejko A. [Importance of hospital schools in the process of educating chronically ill children]. *Pol Tyg Lek.* 1973; 28: 1435-7.
145. Martinez-Picado J, Morales-Lopetegi K, Wrin T, et al. Selection of drug-resistant HIV-1 mutants in response to repeated structured treatment interruptions. *Aids.* 2002; 16: 895-9.
146. Metzner KJ, Bonhoeffer S, Fischer M, et al. Emergence of minor populations of human immunodeficiency virus type 1 carrying the M184V and L90M mutations in subjects undergoing structured treatment interruptions. *J Infect Dis.* 2003; 188: 1433-43.
147. Metzner KJ, Rauch P, Walter H, et al. Detection of minor populations of drug-resistant HIV-1 in acute seroconverters. *Aids.* 2005; 19: 1819-25.
148. Wensing AM and Boucher CA. Worldwide transmission of drug-resistant HIV. *AIDS Rev.* 2003; 5: 140-55.
149. Flys T, Nissley DV, Claasen CW, et al. Sensitive drug-resistance assays reveal long-term persistence of HIV-1 variants with the K103N nevirapine (NVP) resistance mutation in some women and infants after the administration of single-dose NVP: HIVNET 012. *J Infect Dis.* 2005; 192: 24-9.
150. 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.
151. Jourdain G, Ngo-Giang-Huong N, Le Coeur S, et al. Intrapartum exposure to nevirapine and subsequent maternal responses to nevirapine-based antiretroviral therapy. *N Engl J Med.* 2004; 351: 229-40.
152. Palmer S, Boltz V, Martinson N, et al. Persistence of nevirapine-resistant HIV-1 in women after single-dose nevirapine therapy for prevention of maternal-to-fetal HIV-1 transmission. *Proc Natl Acad Sci U S A.* 2006; 103: 7094-9.
153. Palmer S, Kearney M, Maldarelli F, et al. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *J Clin Microbiol.* 2005; 43: 406-13.
154. Benson CA, Vaida F, Havlir DV, et al. A randomized trial of treatment interruption before optimized antiretroviral therapy for persons with drug-resistant HIV: 48-week virologic results of ACTG A5086. *J Infect Dis.* 2006; 194: 1309-18.

155. Gao Y, Paxinos E, Galovich J, et al. Characterization of a subtype D human immunodeficiency virus type 1 isolate that was obtained from an untreated individual and that is highly resistant to nonnucleoside reverse transcriptase inhibitors. *J Virol.* 2004; 78: 5390-401.
156. Brown AJ, Precious HM, Whitcomb JM, et al. Reduced susceptibility of human immunodeficiency virus type 1 (HIV-1) from patients with primary HIV infection to nonnucleoside reverse transcriptase inhibitors is associated with variation at novel amino acid sites. *J Virol.* 2000; 74: 10269-73.
157. Little SJ, Daar ES, D'Aquila RT, et al. Reduced antiretroviral drug susceptibility among patients with primary HIV infection. *Jama.* 1999; 282: 1142-9.
158. Furusaki A, Takayanagi H and Tsukada M. Theory of quantum conduction of supercurrent through a constriction. *Phys Rev Lett.* 1991; 67: 132-35.
159. Gao F, Bailes E, Robertson DL, et al. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature.* 1999; 397: 436-41.
160. Keele BF, Van Heuverswyn F, Li Y, et al. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science.* 2006; 313: 523-6.
161. Reeves JD and Doms RW. Human immunodeficiency virus type 2. *J Gen Virol.* 2002; 83: 1253-65.
162. Simon F, Mauclere P, Roques P, et al. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat Med.* 1998; 4: 1032-7.
163. Janssens W, Buve A and Nkengasong JN. The puzzle of HIV-1 subtypes in Africa. *Aids.* 1997; 11: 705-12.
164. Burke DS. Recombination in HIV: an important viral evolutionary strategy. *Emerg Infect Dis.* 1997; 3: 253-9.
165. Oh SK, Cruikshank WW, Raina J, et al. Identification of HIV-1 envelope glycoprotein in the serum of AIDS and ARC patients. *J Acquir Immune Defic Syndr.* 1992; 5: 251-6.
166. Sunila I, Vaccarezza M, Pantaleo G, et al. gp120 is present on the plasma membrane of apoptotic CD4 cells prepared from lymph nodes of HIV-1-infected individuals: an immunoelectron microscopic study. *Aids.* 1997; 11: 27-32.
167. Gelderblom HR, Ozel M and Pauli G. Morphogenesis and morphology of HIV. Structure-function relations. *Arch Virol.* 1989; 106: 1-13.
168. Wong-Staal F and Haseltine WA. Regulatory genes of human immunodeficiency viruses. *Mol Genet Med.* 1992; 2: 189-219.
169. Jacks T, Power MD, Masiarz FR, et al. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature.* 1988; 331: 280-3.
170. Chamorro M, Parkin N and Varmus HE. An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA. *Proc Natl Acad Sci U S A.* 1992; 89: 713-7.
171. Gottlinger HG, Sodroski JG and Haseltine WA. Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proc Natl Acad Sci U S A.* 1989; 86: 5781-5.
172. Parkin NT, Chamorro M and Varmus HE. Human immunodeficiency virus type 1 gag-pol frameshifting is dependent on downstream mRNA secondary structure: demonstration by expression in vivo. *J Virol.* 1992; 66: 5147-51.
173. Bernstein HB, Tucker SP, Kar SR, et al. Oligomerization of the hydrophobic heptad repeat of gp41. *J Virol.* 1995; 69: 2745-50.
174. Ruben S, Perkins A, Purcell R, et al. Structural and functional characterization of human immunodeficiency virus tat protein. *J Virol.* 1989; 63: 1-8.
175. Feng S and Holland EC. HIV-1 tat trans-activation requires the loop sequence within tar. *Nature.* 1988; 334: 165-7.
176. Roy S, Delling U, Chen CH, et al. A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation. *Genes Dev.* 1990; 4: 1365-73.
177. Zapp ML and Green MR. Sequence-specific RNA binding by the HIV-1 Rev protein. *Nature.* 1989; 342: 714-6.
178. Kim SY, Byrn R, Groopman J, et al. Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. *J Virol.* 1989; 63: 3708-13.
179. Malim MH, Hauber J, Le SY, et al. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature.* 1989; 338: 254-7.
180. Aiken C, Konner J, Landau NR, et al. Nef induces CD4 endocytosis: requirement for a critical dileucine motif in the membrane-proximal CD4 cytoplasmic domain. *Cell.* 1994; 76: 853-64.
181. Luria S, Chambers I and Berg P. Expression of the type 1 human immunodeficiency virus Nef protein in T cells prevents antigen receptor-mediated induction of interleukin 2 mRNA. *Proc Natl Acad Sci U S A.* 1991; 88: 5326-30.

182. Skowronski J, Parks D and Mariani R. Altered T cell activation and development in transgenic mice expressing the HIV-1 nef gene. *Embo J.* 1993; 12: 703-13.
183. Baur AS, Sawai ET, Dazin P, et al. HIV-1 Nef leads to inhibition or activation of T cells depending on its intracellular localization. *Immunity.* 1994; 1: 373-84.
184. Miller MD, Warmerdam MT, Gaston I, et al. The human immunodeficiency virus-1 nef gene product: a positive factor for viral infection and replication in primary lymphocytes and macrophages. *J Exp Med.* 1994; 179: 101-13.
185. Heinzinger NK, Bukinsky MI, Haggerty SA, et al. The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. *Proc Natl Acad Sci U S A.* 1994; 91: 7311-5.
186. Rogel ME, Wu LI and Emerman M. The human immunodeficiency virus type 1 vpr gene prevents cell proliferation during chronic infection. *J Virol.* 1995; 69: 882-8.
187. Sato A, Igarashi H, Adachi A, et al. Identification and localization of vpr gene product of human immunodeficiency virus type 1. *Virus Genes.* 1990; 4: 303-12.
188. Schubert U, Bour S, Ferrer-Montiel AV, et al. The two biological activities of human immunodeficiency virus type 1 Vpu protein involve two separable structural domains. *J Virol.* 1996; 70: 809-19.
189. Strebel K, Daugherty D, Clouse K, et al. The HIV 'A' (sor) gene product is essential for virus infectivity. *Nature.* 1987; 328: 728-30.
190. Pandori MW, Fitch NJ, Craig HM, et al. Producer-cell modification of human immunodeficiency virus type 1: Nef is a virion protein. *J Virol.* 1996; 70: 4283-90.
191. Schwartz O, Marechal V, Danos O, et al. Human immunodeficiency virus type 1 Nef increases the efficiency of reverse transcription in the infected cell. *J Virol.* 1995; 69: 4053-9.
192. Schwartz S, Felber BK, Benko DM, et al. Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. *J Virol.* 1990; 64: 2519-29.
193. Starcich B, Ratner L, Josephs SF, et al. Characterization of long terminal repeat sequences of HTLV-III. *Science.* 1985; 227: 538-40.
194. Nabel G and Baltimore D. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature.* 1987; 326: 711-3.
195. Okamoto T, Matsuyama T, Mori S, et al. Augmentation of human immunodeficiency virus type 1 gene expression by tumor necrosis factor alpha. *AIDS Res Hum Retroviruses.* 1989; 5: 131-8.
196. Kobayashi N, Hamamoto Y, Koyanagi Y, et al. Effect of interleukin-1 on the augmentation of human immunodeficiency virus gene expression. *Biochem Biophys Res Commun.* 1989; 165: 715-21.
197. 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: 489-93.
198. Southgate CD and Green MR. The HIV-1 Tat protein activates transcription from an upstream DNA-binding site: implications for Tat function. *Genes Dev.* 1991; 5: 2496-507.
199. Schwartz S, Felber BK and Pavlakis GN. Mechanism of translation of monocistronic and multicistronic human immunodeficiency virus type 1 mRNAs. *Mol Cell Biol.* 1992; 12: 207-19.
200. Wu Y and Marsh JW. Gene transcription in HIV infection. *Microbes Infect.* 2003; 5: 1023-7.
201. Chan DC, Fass D, Berger JM, et al. Core structure of gp41 from the HIV envelope glycoprotein. *Cell.* 1997; 89: 263-73.
202. Coakley E, Petropoulos CJ and Whitcomb JM. Assessing chemokine co-receptor usage in HIV. *Curr Opin Infect Dis.* 2005; 18: 9-15.
203. Deng H, Liu R, Ellmeier W, et al. Identification of a major co-receptor for primary isolates of HIV-1. *Nature.* 1996; 381: 661-6.
204. Feng Y, Broder CC, Kennedy PE, et al. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science.* 1996; 272: 872-7.
205. Zhu T, Mo H, Wang N, et al. Genotypic and phenotypic characterization of HIV-1 patients with primary infection. *Science.* 1993; 261: 1179-81.
206. Zhu T, Wang N, Carr A, et al. Genetic characterization of human immunodeficiency virus type 1 in blood and genital secretions: evidence for viral compartmentalization and selection during sexual transmission. *J Virol.* 1996; 70: 3098-107.
207. van't Wout AB, Kootstra NA, Mulder-Kampinga GA, et al. Macrophage-tropic variants initiate human immunodeficiency virus type 1 infection after sexual, parenteral, and vertical transmission. *J Clin Invest.* 1994; 94: 2060-7.
208. Clevestig P, Maljkovic I, Casper C, et al. The X4 phenotype of HIV type 1 evolves from R5 in two children of mothers, carrying X4, and is not linked to transmission. *AIDS Res Hum Retroviruses.* 2005; 21: 371-8.
209. Moore JP. Coreceptors: implications for HIV pathogenesis and therapy. *Science.* 1997; 276: 51-2.

210. Karlsson A, Parsmyr K, Aperia K, et al. MT-2 cell tropism of human immunodeficiency virus type 1 isolates as a marker for response to treatment and development of drug resistance. *J Infect Dis.* 1994; 170: 1367-75.
211. Koot M, van 't Wout AB, Kootstra NA, et al. Relation between changes in cellular load, evolution of viral phenotype, and the clonal composition of virus populations in the course of human immunodeficiency virus type 1 infection. *J Infect Dis.* 1996; 173: 349-54.
212. Samson M, Libert F, Doranz BJ, et al. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature.* 1996; 382: 722-5.
213. Kawamura T, Gulden FO, Sugaya M, et al. R5 HIV productively infects Langerhans cells, and infection levels are regulated by compound CCR5 polymorphisms. *Proc Natl Acad Sci U S A.* 2003; 100: 8401-6.
214. Liu H, Chao D, Nakayama EE, et al. Polymorphism in RANTES chemokine promoter affects HIV-1 disease progression. *Proc Natl Acad Sci U S A.* 1999; 96: 4581-5.
215. Winkler C, Modi W, Smith MW, et al. Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. ALIVE Study, Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC). *Science.* 1998; 279: 389-93.
216. Carrington M and O'Brien SJ. The influence of HLA genotype on AIDS. *Annu Rev Med.* 2003; 54: 535-51.
217. Koch M, Pancera M, Kwong PD, et al. Structure-based, targeted deglycosylation of HIV-1 gp120 and effects on neutralization sensitivity and antibody recognition. *Virology.* 2003; 313: 387-400.
218. Losman B, Bolmstedt A, Schonning K, et al. Protection of neutralization epitopes in the V3 loop of oligomeric human immunodeficiency virus type 1 glycoprotein 120 by N-linked oligosaccharides in the V1 region. *AIDS Res Hum Retroviruses.* 2001; 17: 1067-76.

Curriculum Vitae

Séverine LOUVEL

CURRENT STATUS

Name: Séverine LOUVEL
Date of Birth: 1977, 25th March
Place of Birth: St Malo (France)
PhD student at the Institute for Medical Microbiology
University of Basel
Phone: +41 (0)61 267 09 80 (lab)
+33 (0)3 89 69 03 07 (home)
+33 (0)6 83 87 09 27 (cell)
E-mail: severine.louvel@unibas.ch

EDUCATION

Since January 2004	**Universität Basel** PhD Student at the Faculty of Science (Phil II)	Basel SWITZERLAND
2002-2003	**University of Science and Technology** (Lille II) Post graduate in cellular and molecular engineering, with honors (mention très bien)	Lille FRANCE
2000-2002	**University of Science and Technology** (Lille II) Graduated in cellular and molecular Biology, with honors (mention bien)	Lille FRANCE
1999-2000	**Institute of pharmaceutical Chemistry** Diploma of Drug Sciences and Biological Sciences, with honors (mention très bien)	Lille FRANCE
1997-1999	**University Institute of Technologies** Technician Degree in biological and biochemical analysis	Quimper FRANCE
1995-1997	**University of Medicine**	Rennes FRANCE

EXPERIENCE

Since January 2004 — **Universität Basel** — Basel, SWITZERLAND
Institut für Medizinische Mikrobiologie

Title of the Thesis Project:
Validation of replicative phenotyping to detect and assign HIV-1 resistance in clinical specimens.

2003 — **University of Science and Technology** (Lille II) — Lille, FRANCE
Department of Molecular and Cellular Biology
Training course at InPheno AG, Basel, Switzerland

Title of the Master Project:
Study of the phenotypic HIV resistance testing:
- discordance with genotyping in heterogeneous viral populations
- optimisation of a new methodology to assess HIV fitness

This project was financed by a personal grant I obtained from the CROUS (Centre Regional des Oeuvres Universitaires et Scolaires), known as: “bourse de voyage”

Sept 00-Jun 01 — **University of Lille** — Lille, FRANCE
Institut de Chimie pharmaceutique Albert Lespagnol
Training course at INSERM, unit 377, Lille, France

Title of the DU Project:
Isolation and study of the murine mucin genes. Screening of genomic banks and study of a cosmid carrying a part of muc4 murine gene.

2001 — **University Institute of Technologies** — Quimper, FRANCE
Department of biological and biochemical analysis
Training course at Laval University, Quebec, Quebec

Title of the DUT Project:
Cloning of the major capsid protein of the bacteriophage KSY1.

This project was financed by a personal grant I obtained from the CROUS (Centre Regional des Oeuvres Universitaires et Scolaires), known as: “bourse de voyage”

PUBLICATION

- Detection of drug-resistant HIV minorities in clinical specimens and therapy failure. Séverine Louvel, Manuel Battegay, Pietro Vernazza, Thomas Bregenzer, Thomas Klimkait, François Hamy, and the Swiss HIV cohort study. (submitted)

MEETINGS

- 2d European HIV drug resistance workshop, Rome, Italy, March 2004.
  Poster presentation: In vitro assessment of NRTI-discordances between Genotypic and Phenotypic resistance testing using defined complex mutants or viral mixtures as found in clinical samples.

- XV International AIDS Conference, Bangkok, Thailand, July 2004.
  Poster presentation: Dissection of HIV-1 resistance profiles comparing genotypic analysis with replicative Phenotyping.

- Saaten, Switzerland, January 2004.
  Poster presentation: Comparison of HIV-1 resistance profiles for PI and RTI using genotypic algorithms (ANRS, Stanford, Rega) and replicative phenotyping (PhenoTecT).

- 3d European HIV drug resistance workshop, Athens, Greece, March 2005.
  Poster presentation: Detection of viral minority species using resistance testing.

- HIV Highlights meeting, Amsterdam, Netherlands, June 2005.

- 2nd Swiss Workshop on Basic HIV Research, Murten, Switzerland, May 2006.
  Poster presentation: Revealing drug resistant minority variants using replicative phenotyping.

- HIV DART 2006 Symposium, Cancun, Mexico, December 2006.
  Poster presentation: Replicative Phenotyping (PhenoTecT) detects critical HIV minorities in plasma. Awarded with the chairmen's poster award offered by Robert Murphy and Raymond F. Schinazi on behalf of the organising committee of HIV DART 2006.