Diploma Thesis

Ultra-deep sequencing for detection of HIV-1 subtype B drug resistance mutations

submitted by

Beata Bizon

0433324

In partial fulfillment of the requirements for the degree of

Medical Doctor (MD)

at the

Medical University of Graz

Institute of Hygiene, Microbiology and Environmental Medicine

Universitaetsplatz 4, 8010 Graz

Supervisor

Univ.-Prof. Dr. med. univ. Harald H. Kessler

Graz, May 11, 2010

Declaration

Herewith I, Beata Bizon, declare that I have written the present diploma thesis fully on my own and without any assistance from third parties. Furthermore, I confirm that no sources habe been in the preparation of the thesis other than those indicated in the thesis itself.

Graz, May 11, 2010 *Signature*

Acknowledgement

I would like to express my sincere gratitude to Univ.-Prof. Dr. med. univ. Harald H. Kessler for giving me the possibility to perform the research work and supervision of this diploma thesis.

I give my special thanks to Dr. Reinhard B. Raggam for his important mentoring during this time that allowed me to improve this work.

I would also like to thank Mag. rer. nat. Dr. sci. med. Evelyn Stelzl for helping me broaden my knowledge.

I am particularly thankful to Prim. Dr. Christian Gabriel of the Red Cross op Upper Austria for providing the facility and all the necessary tools to execute this diploma thesis.

Especially I would like to thank Dr. rer. nat. Johannes Proell for his generous support and efforts invested. Futhermore, I would like to thank Mag. Martin Danzer and Mag. Christa Hackl for their support. I am also grateful to Mag. Katja Hofer and Mag. Stephanie Stabentheiner for their time and interest.

Special thanks goes to my partner Dipl.-Ing. Peter Holler whose continuous patience and encouragement enabled me to complete this work.

CONTENTS

1 List of Figures and Tables 6

2 Abbreviations and Acronyms 9

3 Abstract 10

3.1 Abstract in German 10

3.2 Abstract in English 11

4 Background 12

5 Introduction 13

5.1 Human Immunodeficiency Virus 13

5.1.1 Structure and Morphology of HIV 13

5.1.2 Genome 14

5.1.3 Replication 15

5.1.4 Tropism 16

5.1.5 Epidemiology 16

5.1.6 Transmission 17

5.1.7 Laboratory Diagnostics 18

5.1.8 Pathophysiology and Clinical Manifestations 18

5.1.9 Treatment Options 20

5.1.10 Resistance Associated Mutations 22

6 Materials and Methods 23

6.1 Study Design 23

6.2 Clinical Samples 24

6.3 Quantitation of HIV-1 RNA 24

6.3.1 Sample Preparation 24

6.3.2 Reverse Transcription and PCR Amplification 25

6.3.3 Detection of PCR Products 25

6.3.4 Measurment of HIV-1 RNA 26

6.4 Standard Sequencing Method 26

6.4.1 Reverse Transcription and PCR Amplification 27

6.4.2 CLIP^{TM} Sequencing Reaction 27

6.4.3 Polyacrylamid Gel Electrophoresis of CLIP^{TM} Reaction Products 28

6.4.4 Data Analysis 28
6.5 Ultra-deep Sequencing 30
6.5.1 Overview 30
6.5.2 RNA Extraction from Plasma Samples 30
6.5.3 CDNA Synthesis from Extracted RNA 31
6.5.4 Amplicon Generation from cDNA 32
6.5.5 Amplicon Purification 33
6.5.6 Amplicon Quantitation by PicoGreen Flourescence 34
6.5.7 Agilent Measurement 35
6.5.8 Amplicon Dilution and Pooling 36
6.5.9 Capture Beads Wash and DNA Library Preparation 36
6.5.10 Emulsion PCR 36
6.5.11 Bead Washes and Recovery 37
6.5.12 DNA Library Bead Enrichment 37
6.5.13 Counting of the Beads 39
6.5.14 Ultra-deep Sequencing Technology 39
6.5.15 Data Analysis 41
6.5.16 Stanford University HDRM Database 42
7 Results 44
7.1 Accuracy Testing 44
7.2 Quality Parameters of the UDS Method for an Austrian Cohort 45
7.3 Distribution of HIV-1 Variants 47
7.4 Reproducibility Testing 53
8 Discussion 54
9 Conclusions 60
10 References 61

1 List of Figures and Tables

Figure 1: Structure of HIV. Available from: http://www.kon.org. Downloaded April 26, 2010

Fig. 2: Replication cycle of HIV-1. Available from: http://artforming.blogspot.com. Downloaded April 26, 2010

Fig. 3: A global view of HIV infection 2007. Available from: http://www.unaids.org/en. Downloaded December 8, 009

Fig. 4: Relationship between viral load and $CD4^+$ cell concentration during an HIV-infection. Available from: http://pathmicro.med.sc.edu/. Downloaded October 22, 2009

Fig. 5: The electropherogram of the sample V1015867. Available from Evelyn Stelzl, Institute of Hygiene, Micorbiology and Environmental Medicine of the Medical University of Graz. March 24, 2010

Fig. 6: HIV-1 gene map of the *pol* gene. Information availabe from: http://hivdb.stanford.edu. March 2, 2010

Fig. 7: HIV-1 cDNA product with eight overlapping amplicons. Available from: Alpha Site Study Protocol (HIV), Roche 454 Life Science. September 2009

Fig. 8: Schematic of the HIV-1 primer with the MID, key sequence and A and B adaptor

Fig. 9: Results of the electrophoresis on the Agilent 2100 Bioanalyzer. Obtained by the Red Cross Transfusion Service of Upper Austria. August 20, 2009

Fig. 10: DNA bound to capture and enrichment beads. Information available from: http://www.roche-applied-science.com. March 22, 2009

Fig. 11: Annealing primer bound to the DNA capture beads

Fig. 12: UDS technique which is based on sequencing-by-synthesis. Information available from: http://454.com. March 21, 2010

Fig. 13: A typical flowgram of a sequenced read. Available from: http://www.roche-applied-science.com. March 22, 2010

Fig. 14: Specimen variant pattern of the QCMD sample ENVA 09-03, reported by the AVA software. Obtained by the Red Cross Transfusion Service of Upper Austria.

Fig. 15: Amount of raw wells and ACTG that passed the filter

Fig. 16: Lanes of the sequencing run 1 and 2 and the mean value that passed the filter

Fig. 17: Mean number of mutations per drug class and patient

Fig. 18: Distribution of the drug variants according to prevalence

Fig. 19 A, B: Interassay variability- correlation between two runs: fraction (%) per hits (A) and number per hits (B)

Table 1: Results of five QCMD samples that were UDS sequenced

Tab. 2: Comparison of the PI mutations with both sequencing methods

Tab. 3: Comparison of the NRTI mutations with both sequencing methods

Tab. 4: Comparison of the NNRTI mutations with both sequencing methods

Tab. 5: Comparison of the proportion of the mutations found in the study cohort and mutations identified by Simen et al

Tab. 6: Amount and percentage of the mutations found with the UDS method

Tab 7: Results of five samples with a low viral load

Tab. 8: The manual preparation of the standard and UDS method

2 Abbreviations and Acronyms

AIDS= acquired immuno deficiency syndrome

CRF= circulating recombinant forms

db= database

ddNTP= dideoxynucleotide triphosphates

dNTP= deoxnucleotide triphosphates

emPCR= emulsion polymerase chain reaction

FIRST= Flexible Initial Retrovirus Suppressive Therapies

gp= glycoprotein

HAART= highly active antiretroviral therapy

HDRM= HIV Drug Resistance Mutations

HIV-1= human immunodeficiency virus type 1

MID= multiplex identifier

NNRTI= non nucleoside reverse transcriptase inhibitor

NRTI= nucleoside reverse transcriptase inhibitor

LTR= long terminal repeat

PCR= polymerase chain reaction

PP= polypropylene

QCMD= Quality Control of Molecular Diagnostics

PI= protease inhibitors

RT= reverse transcriptase

TBE= Tris-Borate-EDTA

TG= TruGene

UDS= ultra-deep sequencing

SPRI= Solid Phase Reversible Immobilization

3 Abstract

3.1 Abstract in German

Ziele: Die Untersuchung der Ausführbarkeit der Ultra-deep Sequenzierungsmethode (UDS) für eine verlässliche Detektion der HIV-1 Subtyp B Mutationen. Der Vergleich der Ergebnisse der UDS Methode mit jenen der routinemäßig eingesetzten Standard Sequenzierungsmethode.

Methoden: Um die Validität zu testen wurden Proben mit dem Genotyp B des Quality Control for Molecular Diagnostics (QCMD) mit der neuen UDS-Methode analysiert.

Die klinische Durchführbarkeit von 45 klinischen Routineproben wurde überprüft und die Ergebnisse mit der Standard Methode verglichen. Ergebnisse hinsichtlich der Medikamentenresistenz wurden für die UDS-Methode mittels HIV Medikamentenresistenzen-Datenbank der Universität Stanford erhalten. Für die Ergebnisse bezüglich der Standard Sequenzierung wurden die GuideLines rules (Teil der TruGene HIV-1 Untersuchungsmethode von Siemens Health Care Diagnostics in der Version 14.0) verwendet. Die Zuverlässigkeit wurde durch eine einfache Wiederholung der UDS-Sequenzierung von 21 Proben bestimmt.

Resultate: Bei der Überprüfung der Genauigkeit stimmten alle bis auf einen Mutations-Codon mit den QCMD Ergebnissen überein.

Die Analyse von 45 klinischen Routineproben zeigte mit der neuen UDS-Methode eine Gesamtanzahl von 268 Varianten mit einer Prävalenz von ≥1% von 44 verschiedenen Mutations-Codons. Beim Vergleich mit der Standard Sequenzierungsmethode wurden 170 identische Mutationen mit beiden Systemen gefunden. 98 mehr Mutationen wurden mit der UDS Methode allein und 18 Mutationen mit der Standard Methode allein gefunden. Die Präzisionstestung ergab einen Korrelationskoeffizienten von R^2=0.984.

Fazit: Die neue-UDS Methode zeigte eine gute Genauigkeit und eine bedeutende klinische Durchführung mit einer hohen Präzision der Testergebnisse. Für den Einsatz im Routinelabor ist die Steigerung des Automatisierungsgrades der Methode zur Reduktion der Durchlaufzeit notwendig.

3.2 Abstract in English

Goals: To evaluate the feasibility of the ultra-deep sequencing (UDS) method for reliable detection of HIV-1 subtype B mutations. To compare results obtained by UDS with those obtained by the routinely used standard sequencing method.

Methods: For accuracy testing, genotype B samples included in the Quality Control for Molecular Diagnostics (QCMD) were analyzed with the new UDS method. Clinical performance was tested with 45 routine clinical samples and compared to standard sequencing method. Results regarding drug resistance obtained with the UDS method were achieved by using the Stanford University HIV drug resistance database. Results regarding drug resistance obtained with the standard sequencing method were achieved by using the GuideLines rules, version 14.0, integral part of the TruGene HIV-1 assay (Siemens Health Care Diagnostics). Reproducibility was determined by repeated GS-FLX sequencing of 21 samples.

Results: When accuracy was tested, all except of one mutation codons were concordant with the QCMD results. Analysis of 45 routine clinical samples with the new UDS method revealed a total number of 268 variants with a minimum of 1% prevalence distributed to 44 different mutation codons. When results were compared with the standard sequencing method, 170 identical mutation codons with clinical relevance were found with both systems. A total of 98 mutation codons were found with the new UDS method only, and 18 mutation codons were found solely with the standard method, respectively. Reproducibility testing resulted in a correlation coefficient of R^2=0.984.

Conclusions: The new UDS method showed good accuracy and a meaningful clinical performance with high reproducibility of test results. Further automation and shorter time to results are necessary for future use in the routine diagnostic laboratory.

4 Background

For the diagnosis and the treatment of an HIV infection, an accurate and sensitive detection of drug resistant HIV strains is of major importance. Infections with drug resistant HIV strains can lead to antiretroviral treatment failure associated with increased morbidity and mortality which can be prevented by the early detection of resistant HIV-strains (Hoffmann et al, 2007; Johnson et al, 2008).

The standard method, the TruGene HIV-1 assay is based on the Sanger technique. However, this method has its limitations because it usually does not detect mutations with less than 20% of prevalence. Therefore, this presently used sequencing method may not be sufficient to detect minor drug resistance mutations. Relevant mutations can make up less than 20% of the viral population and grow rapidly under selection pressure when not detected (Simen et al 2009; Halvas et al, 2010; Peredes et al, 2009; Le et al, 2009).

This great variability is a result of a high frequency of viral replication errors that occur because of the absence of a poofreading function of the reverse transcriptase enzyme (Halvas et al, 2010; Modrow et al, 1997). The newly introduced ultra-deep sequencing (UDS) method is a sequencing method which allows for detection of resistant virus populations of less than 20% in plasma samples and is suitable for high-throughput testing. (Wang et al, 2007; Simen et al, 2009).

Since the interpretation of HIV-1 drug resistance test results is a very difficult task, in this diploma thesis two different genotyping resistance systems with its own process steps were compared. The TruGene HIV-1 assay with its own interpretation system is a proprietary drug resistance system. The UDS method uses the Stanford University HIV Drug Resistance database that is available to a much broader community of researchers and clinicians (Liu et al, 2006; Hirsch et al, 2008).

5 Introduction

5.1 Human Immunodeficiency Virus

In 1981 the Communicable Disease Center (CDC) in Atlanta, USA, observed a higher demand of Pentamidin that is prescribed for the therapy of *pneumocystis-carinii* pneumonia which is a rare opportunistic infection. Former healthy people developed infectious diseases like soor-esophagitis, toxoplasmosis of the central nervous system or malignancies like Kaposi's sarcoma with life threatening progression that are associated with a weak immune system (Mims et al, 2006).

The French scientists Françoise Barré-Sinoussi and Luc Montagnier isolated in 1983 and in 1984 the American scientist Robert Gallo a virus later called Human Immunodeficiency Virus (HIV) that is responsible for the clinical condition called the acquired immunodeficiency syndrome (AIDS). AIDS occurs when an HIV infection progresses to the last stage caused by a severe failure of the immune system (Hoffmann et al, 2008 p25-39).

5.1.1 Structure and Morphology of HIV

HIV is a member of the retrovirus family and belongs to the genus lentivirus. Infections with lentiviruses are described as a chronic infection with a long clinical latency and persisting viremia (Hoffmann et al, 2008 p59).

There are two species of HIV: HIV-1 and HIV-2. Concerning the amino acid sequence there is a homology between HIV-1 and HIV-2 of 40-60%. HIV-2 is restricted to West Africa and causes 1% of all HIV infections. HIV-1 is more virulent and infective and is accountable for the majority of HIV infections worldwide (Hoffmann et al, 2008 p59).

Three groups of HIV-1 have been identified: M (main) which is the most prevalent, N (new) and O (outlier). Group M is subdivided into 10 subtypes (A-J) which occurrence is geographically distributed. Subtype B is most widespread in North America and Europe, subtype A and C can mainly be found in Africa. The popularity of international tourism has aggravated the risk of rearrangement of distinct subtypes to circulating recombinant forms (CRFs) (Mims et al, 2006).

The diameter of the HIV is about 100nm and it is surrounded by the viral envelope. It consists of a lipoprotein layer with glycoprotein complexes which is composed of the outer membrane molecules gp120 and transmembrane molecules gp41 (Fig. 1). This complex enables the virus to attach and to fuse with humane immune cells. The envelope contains also HLA class I- and II molecules and adhesion proteins like ICAM-1. The matrix protein p17 surrounds the capsid. The viral protein p24 (core antigen) encloses two copies of the single stranded (ss) RNA. The ssRNA is bound to the nucleocapsidprotein p7 and enzymes for the development of the virus such as reverse transcriptase (RT) p66, protease p11 and integrase p32 (Hoffmann et al, 2008 p59-78).

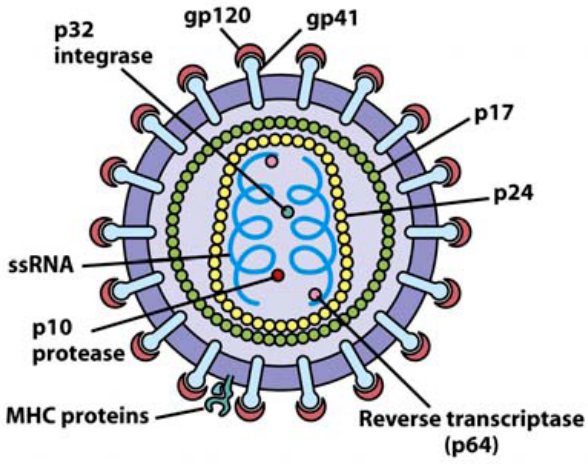

Fig. 1: Structure of HIV

5.1.2 Genome

The HIV-RNA genome is composed of three main genes: *gag* (group), *pol* (polymerase) and *env* (envelope) and six accessory genes *tat*, *rev*, *nef*, *vif*, *vpr*, and *vpu* that are regulatory genes. *Gag* and *env* genes provide information for the nucleocapsid and the envelope. The *pol* gene encodes for the RT and other enzymes. The regulatory protein *tat* and *rev* control replication and are necessary for transcription and expression of regulatory proteins. The ends of the RNA strands contain the sequence “long terminal repeat” (LTR) which is important for the integration of the HIV-RNA in the human DNA (Hoffmann et al, 2008 p59-63).

5.1.3 Replication

HIV penetrates $CD4^+$ T cells (T helper cells), monocytes, dendritic cells and macrophages by the high affine adsorption of the glycoprotein (gp) 120 to the CD4 receptor. The binding to the CD4 molecule creates a change in the conformation of gp120 causing gp41 to be exposed and allowing further consequently HIV-1 to bind to the chemokine receptors CCR5 or CXCR4. After the attachment (Fig 2), HIV inserts the fusion protein gp41 that allows the outer membrane of the virus to fuse with the cell membrane. During the endocytosis the HIV-RNA is released to the host cell and after uncoating the reverse transcription of the RNA to a complementary DNA (cDNA) molecule can be performed by the viral RT enzyme. This enzyme is extremely error prone which leads to mutations and drug resistance.

When the $CD4^+$ cell is activated i.e. during an infection the viral DNA is transported to the nucleus and integrated into the host DNA. During replication the viral DNA is transcribed to mRNA and the viral proteins systhetised. Finally, the assembly of new HIV virions is accomplished at the membrane of the host cell and the virus begins to bud from the host cell (Hahn et al, 2004; Hoffmann et al, 2008 p59-63).

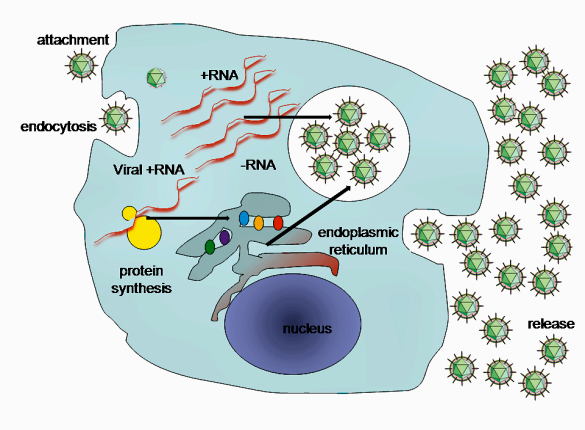

Fig. 2: Replication cycle of HIV-1

The activation of T-helper cells is not only enhanced during the reaction against HIV antigens but also during bacterial infections. The consequence of the productive replication of HIV is the destruction of T-helper cells directly through the induction of apoptosis or the increased vulnerability for cytotoxic T-lymphocytes.

Other $CD4^+$ cells like monocytes, macrophages, dendritic cells and Langerhans cells can also be infected but are not deleted and serve as a reservoir for HIV (Mims et al, 2006).

5.1.4 Tropism

Tropism describes the ability of the HIV to infect immune cells trough the binding to specific surface receptors (CD4 as well as the co-receptors CCR5 and CXCR4) to enter the cell. M-tropic isolates infect mainly peripheral blood mononuclear cells (PBMCs), monocytes and macrophages and need CCR5 for the entry to $CD4^+$ cells. Their natural ligands are the macrophage inhibitory proteins (MIP)-1α and MIP-1β which inhibit the viral entry to host cells. T-tropic isolates replicate preferably in PBMCs and utilize in addition to the CD4 receptor the CXCR4 co-receptor. The ligand for CXCR4 is the stromal cell derived factor (SDF)-1 which blocks the entry of T-tropic isolates.

People with a genetic mutation of the CCR5 co-receptor are resistant to the infection of M-tropic viruses but *in vitro* not to T-tropic isolates. This deletion is found homozygous at about 1% and heterozygous at 20% of the Caucasians (Hoffmann et al, 2008 p63-70).

The clinical relevance of the CXCR4 specific viral load to predict disease progression independent of the $CD4^+$ cell count and HIV viral load is under investigation (Weiser et al, 2008).

5.1.5 Epidemiology

From 1981 to 2007, HIV killed more than 25 million people worldwide and has caused severe demographic changes in heavily affected countries. In 2007, about 33 million people were estimated to be infected by HIV (Fig. 3). Just in this year HIV claimed about 2.7 million new HIV infections and 2 million deaths.

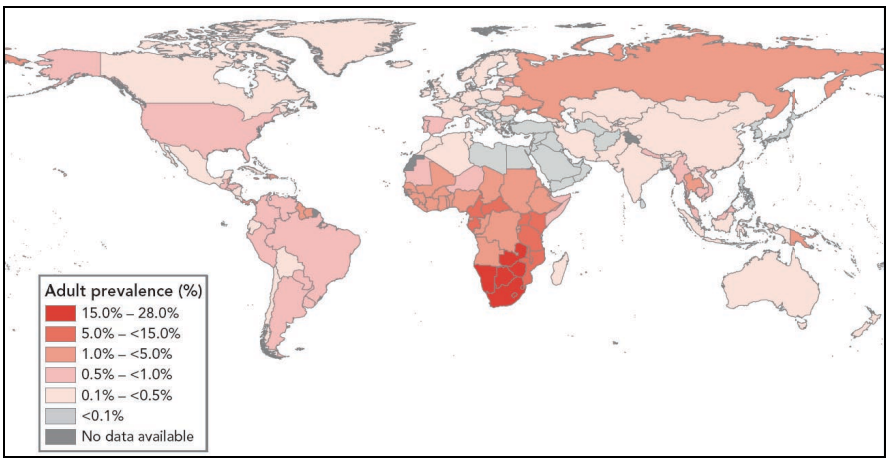

Fig. 3: A global view of HIV infection 2007, UNAIDS

From 1983 to December 1st 2009 Austria reported 2 752 patients that developed AIDS, 1 503 patients have already died on AIDS. About 12 000 to 15 000 people in Austria are infected with HIV, 48% of them live in Vienna. 79% of the HIV-patients are men. In 34% the infection is transmitted by homo/bisexual contact, in 24% by sharing injecting equipment, 21% heterosexually, and 1% vertically. Every day there are about one or two new infections in Austria (AIDS Hilfe Wien, 2009).

Fortunately, recent international epidemiological data reports that the annual number of new HIV infections is dropping in some countries in Asia, Latin America and Sub Sahara Africa. Also the number of AIDS deaths has fallen over the last ten years due to the better access and availability to antiretroviral drugs and targeted approaches, incorporating prevention, early diagnosis and counseling. Except for Sub Sahara Africa, where two third of all people affected with HIV live, HIV concerns homosexuals, sex workers and injecting drug users (Joint United Nations Programme, 2008; Quinn et al, 2008).

5.1.6 Transmission

Most of the HIV infections are acquired trough sexual transmission by close contact to infected body liquids such as semen and blood. HIV transmission depends on the susceptibility of the sexual partner as well as of the infectiousness and the concentration of the HIV in blood and semen (Pilcher et al, 2007). In higher developed countries HIV infection threatens especially homosexual men because of a more frequent practice of anal sex than heterosexual (Bouhnik et al, 2007). In contrast, the transmission from women to men is much higher in third world countries in Africa and Asia. This can be explained through a higher prevalence of sexually transmitted diseases with ulcers and vaginal discharge which harbour infected lymphocytes and monocytes (Mims et al, 2006).

Another route of transmission is the vertical mother-to-child transmission that can occur during pregnancy, at birth or during breastfeeding. The treatment with antiretroviral drugs, the caesarean section and the avoidance of breast feeding have lowered the risk of transmission (Hoffmann et al, 2008 p25-39).

Another facility of infection can appear trough direct contact to infected blood i.e. for intravenous drug users and people who receive tattoos, piercings and acupuncture procedures (Mims et al, 2006). Testing of blood and blood products for HIV has purged transmission trough blood transfusion and blood products in developed countries (Hoffmann et al, 2008 p25-39).

5.1.7 Laboratory Diagnostics

HIV tests are important for the diagnosis and for monitoring of treatment of HIV infection and for blood and organ donor screening. The replication of the virus takes place mainly during the incubation time. At the beginning of an infection the HIV genome and the p24 antigen are detectable. However, in 80% antibodies are detectable six weeks after infection. The time between the infection and the detection of antibodies is called the sero-diagnostic window. There are various diagnostic tests for the analysis of HIV-1 and HIV-2 that can detect antibodies, antigen and viral nucleic acids (Hoffmann et al, 2008 p43-56; Mims et al, 2006).

Two different standard tests for diagnosis are mandatory: a screening and a confirmatory test. The screening test detects viral specific antibodies based on ELISA (enzyme linked immunosorbant assay) technique with a very high sensitivity but low specificity. In case of a positive screening result a second ELISA test from the same serum sample is performed. If tested positive repeatedly, a confirmatory test (Western Blot) out of a new, indedpendent serum sample is necessary to confirm a HIV infection (Hoffmann et al, 2008 p43-56).

HIV infection can also be detected using real time PCR by directly measuring quantitative HIV-RNA. In special situations, i.e. suspicion of new infections, newborn of an infected mother, drug resistance or screening of blood and blood products, this assay supplements the diagnosis.

Sensitive HIV-RNA real-time polymerase chain reaction (PCR) assays are able to detect less than 20 copies/ml and serve as prognostic markers relevant for the activity of infection and monitoring of therapy (Hoffmann et al, 2008 p43-56).

5.1.8 Pathophysiology and Clinical Manifestations

The infection begins when the virus infiltrates trough body fluids. Langerhans cells present the viral antigens in regional lymph nodes where also T-lymphocytes become infected. This mechanism is responsible for the spreading in the lymph and blood vessels (Hof et al, 2009).

The acute infection appears 3-12 weeks after infiltration and is often asymptomatic. 20-30% of the patients have unspecific mononucleosis similar symptoms like fever, lymphadenopathy, diarrhea and myalgia. A rapid increase of HIV-RNA is followed by a

strong decline because of the development of an immune response that is verifiable with the detection of HIV antibodies (Hahn et al, 2004; Hof et al, 2009).

The next stage is the clinical asymptomatic HIV-infection that is characterized by a latency phase with slow decrease of $CD4^+$ lymphocytes and continuation of the viral replication. A persistent lymphadenopathy can occur. The continuous rise of the viral load and the decrease of the $CD4^+$ T cell count lead to unspecific symptoms like exhaustion, fever of unknown origin, or chronic diarrhea. Opportunistic infections like bacterial pneumonia, oropharyngeal candida infection, oral hairy leukoplakia, lung tuberculosis are more frequent (Hahn et al, 2004; Hof et al, 2009).

AIDS is the last stage of an HIV infection and appears after an incubation time of 1-14 years. This progression of the disease is characterized by opportunistic infections like *pneumocystis jiroveci* pneumonia, *toxoplasma* encephalitis, *candida* spp. infections, *herpes simplex*, *cytomegalovirus* infections and malignancies like the Kaposi sarcoma and non Hodgkin lymphomas.

The absolute count of the $CD4^+$ lymphocytes determines the grade of immunodeficiency. The appearance of AIDS associated opportunistic diseases and malignancies is more probable when the $CD4^+$ cell count has droped under 200 per µl (Fig. 4) (Hahn et al, 2004; Hof et al, 2009).

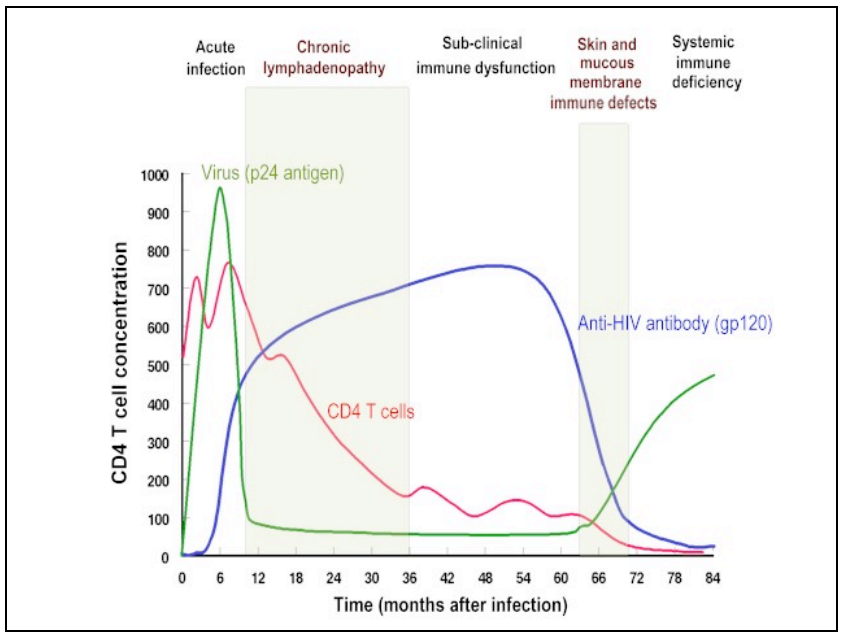

Fig. 4: Relationship between viral load and $CD4^+$ cell concentration during an HIV-infection

5.1.9 Treatment Options

There is no cure for HIV patients but a therapy that improves life quality and reduces HIV associated morbidity and mortality is available. The complete elimination of the virus is still not possible even though there is no viral load measurable. With an antiretroviral therapy the viral load increase can be suppressed in an early stage and the outbreak of the disease avoided for a long time. Current treatment for HIV infection consists of highly active antiretroviral therapy (HAART) which is a combination of several drugs that have a higher efficiency and a lower viral resistance than a monotherapy. HAART is a "cocktail" consisting of at least three drugs belonging to two types of nucleoside reverse transcriptase inhibitors (NRTIs) in combination with either a protease inhibitor (PI) or a non nucleoside reverse transcriptase inhibitor (NNRTI). There are also new classes of drugs such as entry inhibitors available, for patients with HIV infection that are resistant to other drugs (Lüllmann et al, 2006).

The decision how to treat a patient has to be resolved individually. It depends on whether it is the first therapy or the patient already has a treatment history and various resistance profiles. Measuring the amount of $CD4^+$ cells and the viral load provides information regarding the risk of disease progression and information for a possible change in therapy (Hoffmann et al, 2008 p97-270).

5.1.9.1 Nucleoside Reverse Transcriptase Inhibitors

NRTIs are nucleoside analogues that compete with the natural desoxynucleotides and stop the DNA synthesis because they are the wrong substrates for the enzyme RT (Lüllmann et al, 2006).

NRTIs belong to the first group of antiretroviral drugs that were designed. The first approved nucleoside analogue was Zidovudine (Azidothymidine) in 1987. Common side effects that are experienced at the beginning include fatigue, headache, nausea and diarrhea. However, metabolic disorders, hyperlactatemia and lactic acidosis in long term therapy can occur as the result of mitochondrial toxicity (Arenas-Pinto et al, 2008). The application of Zidovudine may induce bone marrow suppression and anemia. Other NRTI present different side effects with a higher incidence: Stavudine is responsible for dose dependent peripheral neuropathy, Zalcitabine- quick resistance, Didanosine– pancreatitis and peripheral neuropathy, Abacavir- hypersensitivity and lactic acidosis. All classical

HAART regimes contain two NRTI's as a "backbone". Most adapted and effective backbones are provided by the combination of Tenofovir plus Emtricitabine (or Lamivudine). An also favored option is the backbone Abacavir plus Lamivudine because of the lower toxicity (Lüllmann et al, 2006; Hoffmann et al, 2008 p103-111).

5.1.9.2 Protease Inhibitors

Protease inhibitors, i.e. Indinavir, Nelfinavir, Saquinavir, Nelfinavir, Amprenavir, Fosamprenavir, Lopinavir, Tipranavir and Ritonavir are abnormal peptides that inhibit the active centre of HIV protease, an enzyme that is necessary to cleave proteins into smaller fragments (Hoffmann et al, 2008 p97-270). The oral bioavailability varies from 4% Saquinavir to 80% Ritonavir. The most frequent adverse reactions are nausea and diarrhea. Indinavir leads to nephrolithiasis and hyperbilirubinemia whereas Ritonavir causes hyperlipidemia. Protease inhibitors are metabolized by cytochrome P450 and may result in serious drug interactions. Furthermore there is a high risk of the promotion to a resistant virus, especially during monotherapy and cross resistance to other PI's (Lüllmann et al, 2006; Baxter et al, 2007).

5.1.9.3 Non Nucleoside Reverse Transcriptase Inhibitor

NNRTI's are the third class of antiretroviral drug that were developed. Nevirapin, Delaviridin and Efavirenz were first approved between 1996 and 1998. They block the reverse transcriptase by binding at a different site of this enzyme. A resistance development is expected when NNRTI's are applied as monotherapy, just like for NRTI's and PI. NNRTI's are an important part of the HIV-therapy because of the low count of daily pills and the good tolerance. Nevirapin and Efavirenz often cause a moderate rash, severe skin reactions like Steven-Johnson syndrome have also been observed. Efavirenz shows less rash but more neurological adverse effects. Furthermore, they induce the hepatic cytochrome P450 enzymes that lead to drug interactions (Lüllmann et al, 2006; Hoffmann et al, 2008 p97-270).

5.1.9.4 Entry Inhibitors

Entry inhibitors are a new class of antiretroviral drugs that affect the entry process of HIV at different steps: the attachment, binding on co-receptors and fusion of an HIV.

All three antiviral agent classes of entry inhibitors can block each of these steps separately: attachment inhibitors block the binding of the HIV surface protein gp120 to the CD4 receptor, co-receptor antagonists inhibit the binding of gp120 to either CCR5 or to CXCR4, and fusion inhibitors block the virus cell fusion. Maraviroc, a co-receptor antagonist that binds to CCR5 and Enfuvirtide a fusion inhibitor, are already approved, whereas other entry inhibitors are under investigation (Hoffmann et al, 2008 p131-139).

5.1.10 Resistance Associated Mutations

Despite of the HAART therapy a minority of patients experiences AIDS. One of the possible reasons is the virological failure due to drug resistance (Cozzi-Lepri et al, 2008).
During viral replication many mutations and furthermore resistance to treatment occur because of the lack of proofreading function of the reverse transcriptase and high viral turnover of about $1x10^{11}$ virions and $1x10^{8}$ infected cells per day (Halvas et al, 2010; Modrow et al, 1997). Many mutations accumulate during the viral replication generating "quasispecies". Under selection pressure and the use of antiretroviral drugs the outgrowth of drug resistant variants can be induced (Simen et al, 2009; Halvas et al, 2010).
HAART therapy with different antiretroviral drugs can lead to HIV-1 variants that have the same or related mutations. Such resistant viral populations are fast-growing when they are under drug selection pressure (Halvas et al, 2010; Palmer et al, 2006).
Finally, it is not only important to identify patients who are likely to progress to AIDS but also clarify whether resistance reflects to poor adherence of patients that do not take antiretroviral medication or take them intermittently (Cozzi-Lepri et al, 2008).

6 Materials and Methods

6.1 Study Design

A precise and sensitive sequencing method is crucial for diagnosis and treatment of an HIV infection. Minor drug resistant strains can lead to therapy failure when not detected. The recently introduced ultra-deep sequencing method detects minor drug resistant strains at a lower prevalence than standard sequencing methods and provides high-throughput parallel sequencing. In this retrospective study, the UDS was investigated for quantitative variant detection of protease and RT genes of the HIV-1 subtype B.

One aim of this study was to investigate the clinical feasability. Therefore 5 proficiency panels (QCMD, Glasgow, Scotland, UK) with HIV-1 subtype B of the years 2009, 2008, and 2006 were tested with the UDS method and the accuracy was determined.

For the purpose to assess the clinical performance of the assay, the remainders of 54 plasma samples that were leftover following clinical routine testing were used. Those samples were already pre-quantified and sequenced in an International Standard Organization (ISO9001, 2000)-certified laboratory, the Molecular Diagnostics Laboratory at the Institute of Hygiene, Microbiology and Environmental Medicine (IHMEM) of the Medical University of Graz with the standard sequencing method using the TruGene HIV-1 Genotyping Kit (Siemens Health Care Diagnostics) and identified as genotype B.

The samples were sent blinded to the Red Cross Transfusion Service of Upper Austria for ultra-deep sequencing with the Genome Sequencer FLX (Roche 454 Life Sciences).

The successful sequencing results of 45 samples obtained by the UDS method were compared with those obtained by the routinely used standard sequencing method.

For the standard sequencing method, drug resistance results were obtained by utilizing the GuideLines rules, version 14.0 which is an integral part of the TruGene HIV-1 assay. Drug resistance results obtained by the UDS method were detected by the means of the Stanford University HIV drug resistance mutations (HDRM) database. Furthermore, the distibution of HIV-1 variants in an Austrian cohort was estimated. Additionally, reproducibility was determined by repeated UDS sequencing of 21 study samples.

6.2 Clinical Samples

In this retrospective study plasma samples from 54 HIV-1 infected patients derived from the LKH Graz and two HIV treatment centers in Vienna between 2007 and 2009. The samples were selected according to the HIV-1 subtype B. The patients were either treated with antiretroviral drugs according to the German-Austrian recommendations (http://www.daignet.de/site-content) or were treatment naive. The HIV drug resistance testing carried out because of primary resistance gene testing, antiretroviral therapy failure or first line drug resistance. The viral load levels ranged from 1 380 to 1 210 000 HIV RNA copies/ml. The mean viral load was 190 971 copies/ ml. The plasma samples were stored in 1.5 ml Sarstedt tubes at –70°C.

6.3 Quantitation of HIV-1 RNA

For quantitation of the HIV-1 RNA in plasma the COBAS® AmpliPrep/COBAS® TaqMan® HIV-1 Test, version 2.0 (Roche Molecular Systems) was used. This established molecular assay is based on automated sample preparation followed by reverse transcription polymerase chain reaction (RT-PCR) and uses automated real time PCR technology for HIV-1 quantitation.

6.3.1 Sample Preparation

The automated sample preparation was performed on the COBAS®AmpliPrep instrument. This high-throughput sample preparation is based on an extraction method with a magnetic silica beads technology.

During HIV-1 RNA extraction, 850µl of plasma sample or control were manually transferred into input S-tubes which were bar code-labeled. Then the sample racks were placed into the COBAS®AmpliPrep instrument. The instrument was before loaded with all reagents and the sample preparation proceeded automatically.

In the first step the samples were lysed at 37°C with lysis buffer to release the nucleic acids. HIV-1 Quantitation Standard (QS) [1] targeting the *gag* sequence with non infectious amored RNA, the protease, the lysis reagents and magnetic silica beads were mixed

[1] The HIV-1 QS at a certain concentration is essential for the control of the sample preparation performance, the RT and PCR amplification and quantitation

automatically by the instrument. At an incubation temperature of 37°C the HIV-1 RNA and the HIV-1 QS RNA were bound to magnetic silica beads. The unbound plasma components such as salts and proteins were washed away by several wash steps.
After adding the elution buffer, the specimen were incubated at 80°C and eluted nucleic acids transferred into output S-tubes. Then, 50µl of the eluates with released HIV-1 RNA and HIV-1 QS RNA were transferred manually into kinetic reaction tubes (K-tubes) and mixed with 50 µl of the amplification mixture (deoxnucleotide triphosphates (dNTP's), primer of the HIV-1 *gag* and the HIV-1 LTR region, Z05 DNA polymerase, AmpErase enzyme, buffer) and transferred to the COBAS® TaqMan® 48 Analyser.

6.3.2 Reverse Transcription and PCR Amplification

The reverse transcription and PCR amplification were performed automatically on the COBAS® TaqMan® 48 Analyser with the thermostable Z05 DNA polymerase.
In the first step, the reaction mix was heated at 50°C for the primer to anneal to the HIV-1 target RNA and the HIV-1 QS RNA. With manganese Mn^{2+} and an excess of dNTP's the Z05 DNA polymerase extended the annealed primers and transcribed the target RNA reversely into cDNA.
One cycle of the PCR amplification consisted of three steps: denaturation, annealing and extension. During denaturation the double stranded (ds) DNA was separated into two single strands by the use of heat (95°C). In the second step the primer that was specific to the target sequence of the HIV/QS could anneal to the DNA at 50°C. In the last step of the PCR cycle the thermostable DNA polymerase that also has a proofreading function, extended the DNA at 72°C. 47 PCR cycles were necessary to achieve a lower quantitation limit of 15 copies/ml.

6.3.3 Detection of PCR Products

The COBAS® AmpliPrep/COBAS® TaqMan 48® HIV-1 Test version 2.0 utilized two different fluorescent dyes for the real-time detection of the amplicons. The reporter and the quencher dye were incorporated in the QS specific probes. In intact probes the reporter dye was near the quencher dye which resulted in suppression of the fluorescence light emitted by the reporter dye. The probe was degraded during PCR because of the 5'→ 3' nuclease activity of the Z05 DNA polymerase. As the reporter and quencher dye were separated, the

repoter dye showed fluorescence emission and no longer quenching. By measuring the fluorescence of target and QS DNA at different wavelengths the quantity of the amplicons that were synthesized during the PCR amplification could be obtained. The fluorescence measurement was repeated during each cycle and was proportional to the amount of the synthesized amplicons.

6.3.4 Measurment of HIV-1 RNA

The amount of HIV-1 RNA amplicons was determined by comparing the HIV-1 signal to the QS signal of each sample and control. Dilution series of the HIV-1 B QS at a kown concentration were quantified by real time PCR and the flourescence results compared to the specimen viral titer.

The baseline fluorescence level was determined with the number of cycles that crossed the threshold value (*Ct*). A sample with a high viral titer showed fluorescence emission at an earlier PCR cycle and a lower *Ct* value. Theoretically, if a maximum of efficiency of the PCR is reached, the amount of DNA doubled each cycle and after N cycles reached the 2^N-fold. A double decrease in titer indicated the half DNA amount, a 10 times increase in titer correlated with a decrease of 3.3 *Ct* values.

After about 40 cycles the exponential growth reached a plateau because of the limited amount of primers.

6.4 Standard Sequencing Method

For the standard DNA sequencing and genotyping assay, HIV-1 RNA was extracted with the MagNA Pure Compact instrument and the MagNA Pure Compact Nucleic Acid (NA) Isolation Kit I (Roche) according to the manufacturer's package insert instructions. Reverse transcription and amplification of the HIV-1 *pol* genes were performed with the TruGene™ HIV-1 Genotyping Kit (Siemens) on the GeneAmp® PCR System 9700 (Applied Biosystems). The DNA amplicons were sequenced with a modified Sanger technique which used flourescent dye labeled primers for the sequencing step (CLIP). The fragments were seperated by polyacrylamide gel electrophoresis.

The sequences were analyzed according to the GuideLines rules, version 14.0, an integral part of the TruGene HIV-1 assay and the HIV-1 subtypes determined by the means of the geno2pheno database (http://www.geno2pheno.org/).

6.4.1 Reverse Transcription and PCR Amplification

The HIV RNA was extracted from 400µl plasma and eluted in 50µl. Two master mixes I[2] and II[3] were prepared. 17µl of extracted RNA and a positive and negative control sample were added into reaction tubes with 9 µl of the master mix I and incubated at 90°C on the GeneAmp® PCR System 9700 thermocycler. After 5 minutes of incubation at 50°C, 14µl of the master mix II was added to the reaction mixture and the thermocycler program proceeded. As described in 4.3.2, the reverse transcriptase enzyme transcribed the RNA into cDNA. The reaction mixture was cooled at 50°C to allow the specific primers to anneal to the HIV-1 target RNA. A complementary DNA-RNA hybrid was generated by the RT enzyme.

The PCR amplification consisted of denaturation of the cDNA at 94°C, annealing at 57°C or 60°C and extension at 68°C according to the following RT-PCR temperature program:

One cycle of:	20 cylces of	17 cycles of:	One cycle of:
90°C for 2 min	94°C for 30 sec	94°C for 30 sec	68°C for 7 min
50°C for 60 min	57°C 30 sec	60°C for 30 sec	4°C hold
94°C for 2 min	68°C for 2 min	68°C for 2.5 min	

6.4.2 CLIP™ Sequencing Reaction

By the use of the CLIP™ reaction sequence both strands of the DNA were generated at the same time labeled with different fluorescent fragments of dyes. These generate sufficient fragments over 30 reaction cycles. The forward primer was labeled with Cy5 fluorescent dye and the reverse primer was labeled with the Cy5.5 flourescent dye.

The three different primers were designed to cover the protease and RT reading frame. The P2 primer concerned the protease codons 7 to 99, RT beginning concerned the gene codons 40 to 142 and RT middle the codons 128 to 247.

Twelve different CLIP terminator mixtures of 7µl contained one of four chain terminating dideoxynucleotide triphosphates (ddNTP's: ddATP, ddCTP, ddCTP, ddTTP), one of the primers and dNTP's were dispensed in 0.2ml PCR tubes. Then 95µl CLIP master mix[4]

---

[2] Master mix I: 43 µl PCR primer, 10.5 µl dNTP, 7 µl dithiothreitol (DTT), 3.5 µl RNAse inhibitor

[3] Master mix II: 70 µl PCR buffer, 3.5 µl RNAseIN, 7 µl RT enzyme, 17.5 µl DNA Taq polymerase

[4] CLIP™ master mix: 474µl molecular grade water, 120µl CLIP TRIS buffer, 22.5µl of CLIP AmpliTaq DNA polymerase with pyrophosphatase

containing the thermostable DNA polymerase with a high affinity to ddNTP's were mixed with 5µl of unpurified RT-PCR products. 5µl of this mixture was aliquoted to each of the 12 PCR tubes. The reaction mix was exposed to the following temperature program:

One cycle of:	30 cycles of:	One cycle of:
94°C for 5 min	94°C for 20 sec	70°C for 5 min
	56°C for 20 sec	4°C hold
	70°C for 1.5 min	

During this temperature program primers hybridized to the template DNA and were extended. They were usually terminated by the ddNTP's along the DNA sequence.

6.4.3 Polyacrylamid Gel Electrophoresis of CLIP™ Reaction Products

The CLIP™ Sequencing Reaction was terminated by addition of 14µl stop solution that contained 93% formamid and 7% dye. It was then heated up to 85°C for 3 minutes to denaturize the double stranded DNA.

The acrylamide gel in the MicroCel 500 cassettes was polymerized in Gel Toaster™ with an UV-light source. The already prepared gel contained a polyacrylamid gel (6%) with a high urea concentration to stabilize the single stranded DNA fragments.

The MicroCel 500 cassette filled with the gel was put into the Long-Read Tower Sequencer and the Tris-Borate-EDTA Buffer[5] filled into both buffer chambers.

After the 3 to 5 minutes pre-run, 2µl of the denatured DNA samples were loaded on top of the prepared gel. Then the P2, the RT beginning RT middle were loaded. A high voltage of 2000 V was applied to move the negatively charged DNA fragments towards the anode. The velocity depended on the size of the DNA fragments. In front of the anode a photosensitive sensor detected the produced fluorescent light of the four lanes, one for each of the four terminating ddNTP's.

6.4.4 Data Analysis

Raw data with an affirmative quality control result was transmitted to a computer workstation for sequence analysis with the OpenGene DNA sequencing system. This

[5] TRIS-Borate-EDTA buffer (TBE) contained 0.089 M tris-(hydroxymethyl) aminomethane, 0.089 M Borate, 0.002 M EDTA

sequencing system contained the GeneObjects DNA analysis software that obtained sequence data from the Long Read Tower Sequencers.

The fragments were assembled automatically and the aligned electropherogram patterns were displayed (Fig. 5). This was performed using a HIV-1 drug resistance and wildtype HIV-1_{LAV-1}-reference sequence.

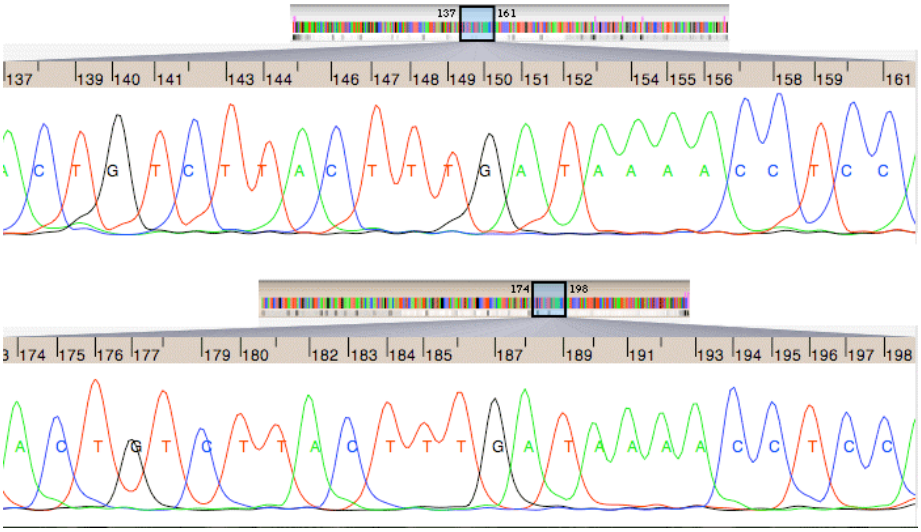

Fig. 5: The electropherogram of the sample V1015867 of the P2 region in the forward (upper) and reverse (lower) sequencing direction with a good alignment and basecalling.

After the alignment was complete, the clinically relevant drug resistance mutations were identified with the TRUGENE HIV-1 Module for GuideLines Rules 14.0. These drug resistance algorithms were first established by an international panel of independent experts in 1999 (Reid, 2002), are annually updated and are US Food and Drug Administration (FDA) cleared.

Finally the drug resistance was interpreted with the GuideLines Rules 14.0 and an interpretative resistance report file of each sample was generated.

Important sequence criteria for the quality control of the CLIP sequencing and detection of contamination of the samples were a negative and a positive control sample. The negative control sample showed only one primer peak and no further base calling sequences. The positive control sample of a non infectious HIV-1 subtype B had no more than two stops per assay.

6.5 Ultra-deep Sequencing

6.5.1 Overview

After HIV-1 RNA was extracted with the manual High Pure Viral Nucleic Acid Kit (Roche) and transcribed to cDNA, three gene specific cDNAs per sample were generated with the FastStart HiFi polymerase (Roche). PCR was performed to receive eight overlapping amplicons of the PR and RT genes at the size between 195 and 418 bp. The amplicons were coded with the multiplex identifier (MID's), a specific oligonucleotide sequence, for the accurate assignment during the sequencing process. The amplicons were purified with magnetic SPRI beads (Agencourt AMPure) and quantified with the Quant-iT^{TM} PicoGreen$^{®}$ dsDNA Assay Kit (Invitrogen). For the DNA library the eight amplicons were diluted at the same concentration and pooled afterwards. Specific forward and reverse adapters were added and the DNA library fragments fixed on capture beads. The PCR reaction occurred in a microreactor developed by emulsification. After breaking of the emulsion and after the enrichment the beads were counted on the Z1 counter (Beckman Coulter). Then the beads were loaded on a picotiter plate device (Roche) and sequencing was performed on the Genome Sequencer (GS) FLX (Roche 454 Life Sciences).
The aligned sequences were compared to reference sequences with the Amplicon Variant Analyzer software (Roche) and drug resistance variants were provided by the Standford HIV Drug Resistance Mutations Database (last updated 2009).

6.5.2 RNA Extraction from Plasma Samples

For the extraction and purification of HIV-1, the High Pure Viral Nucleic Acid (NA) Large Volume Kit[6] was used. The HIV was isolated from 1ml plasma. 1ml plasma was centrifuged for one hour with high speed at 4°C in a 1.5ml tube. The samples were incubated for 10 min at 72°C. The supernatant was aspirated and RNA mixed with 200µl of the phosphate buffered Saline PBS binding buffer (pH 7.2, Gibco) supplemented with polyA-RNA and 50µl proteinase K for cell lysis. Then the RNA was mixed with isopropanol (Sigma) and bound to a High Pure filter. The washing step consisted of

[6] High Pure Viral Nucleic Acid Kit contained binding buffer, polyA-RNA carrier, proteinase K, inhibitor removal buffer, wash buffer, elution buffer

centrifugation for 1 minute at 8 000g and washing with 500µl inhibitor removal buffer and twice with 450µl wash buffer. After the proteins, salts and contamination were washed away the High Pure filter was put into a new 1.5ml microcentrifuge tube. The purified viral nucleic acids were then eluted with 75µl elution buffer to a total extracted RNA volume of 75µl.

6.5.3 CDNA Synthesis from Extracted RNA

Drug resistance mutations of the three important drug classes (PI, NRTI and NNRTI) concern the investigated HIV *pol* region (Fig. 6):

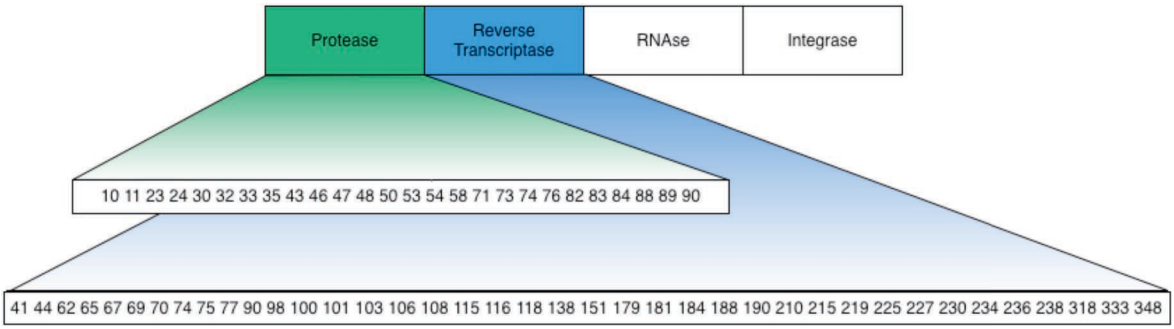

Fig. 6: HIV- gene map of the *pol* gene with the PR and RT- coding regions. Codons for drug resistance mutations (Stanford) are shown in the expanded segment

To analyse potential mutations postitions of whole PR gene (codon 10 to 93) and RT gene (codon 41 to 348) the RNA was transcribed reversely to cDNA using three cDNA primers (4R-HIVB, 5R-HIVB, BR-HIVB). Consequently eight partly overlapping amplicon primers (A-C, 1-5) were designed of the three cDNA specific amplicons that covered the demanded *pol* gene region (Fig. 7).

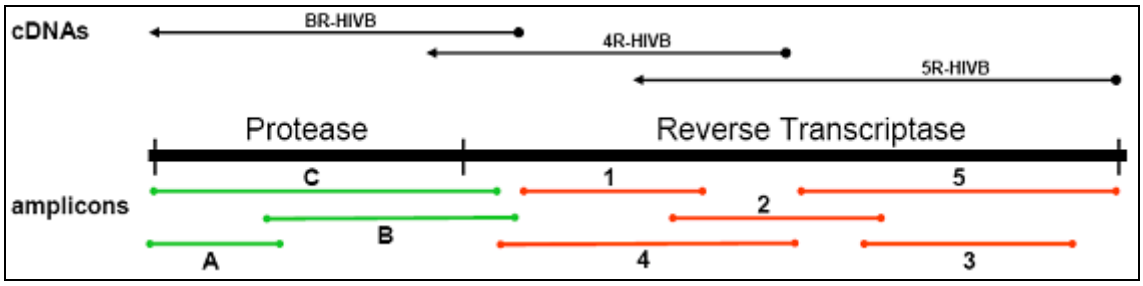

Fig. 7: HIV-1B cDNA products BR-HIVB, 4R-HIVB and 5R-HIVB and eight overlapping amplicons of the PR and RT genes

The maximal read length that can be sequenced by the GS FLX is about 600 bp but reliable and accurate results can only be achieved in exceptional cases. Therefore, it is recommended to use amplicons with a smaller range, i.e. 250 bp.

The amplicons sequenced in this study had an average length of 308 bp and ranged from 195 to 418 bp. The three protease amplicons showed the following read length: A= 195 bp, B= 282 bp, C= 354 bp and the five reverse transcriptase amplicons showed 1= 258 bp, 2= 286 bp, 3= 270 bp, 4= 418 bp, 5= 402 bp (Fig. 7).

For the reverse transcription to cDNA, the FastStart High Fidelity PCR System was used. The FastStart HiFi polymerase is a proofreading enzyme that offers high fidelity and reduces the error rate during reverse transcription and PCR amplification.

11.5µl eluted RNA and 0.5µl of each specific cDNA primers were added into each well of a 96-well plate. The 96-well plate was put in to the 9800 Fast Thermal Cycler (Applied Biosystem) for 10 minutes at 65°C to allow the cDNA primer to anneal to the RNA. Subsequently, 8µl of master mix[7] were added into each well and the following reverse transcription thermocycler run was performed:

One cycle of:

50°C for 60 min

85°C for 5 min

4°C hold

Finally 1µl RNAse H, an endoribonuclease that degrades the RNA into smaller components, was added into each well and put on the thermalcycler at 37°C for 20 minutes.

6.5.4 Amplicon Generation from cDNA

To assure the right assignment of the different samples in a mixed pool during the ultra-deep sequencing process, a nucletide barcode called the multiplex identifier (MID) (Fig. 8) was inserted. Bar coding allows many samples to be analyzed on one picotiter plate which is time and cost-reducing per each sample (Hoffmann et al, 2007).

During the amplicon generation microtiter plates with eight different dried down primers (Roche) for the PR and RT region that included each twelve different MID's for twelve samples were used. The primers that were used also included adaptors that were 5'extensions with binding sites for the forward (A) and reverse (B) primers (Fig. 10).

[7] The master mix contained 420µl trancriptor buffer, 105µl DTT (dithiothreitol), 210µl dNTP, 52,5µl protector RNAase inhibitor, 52,5µl transcriptor RT enzyme

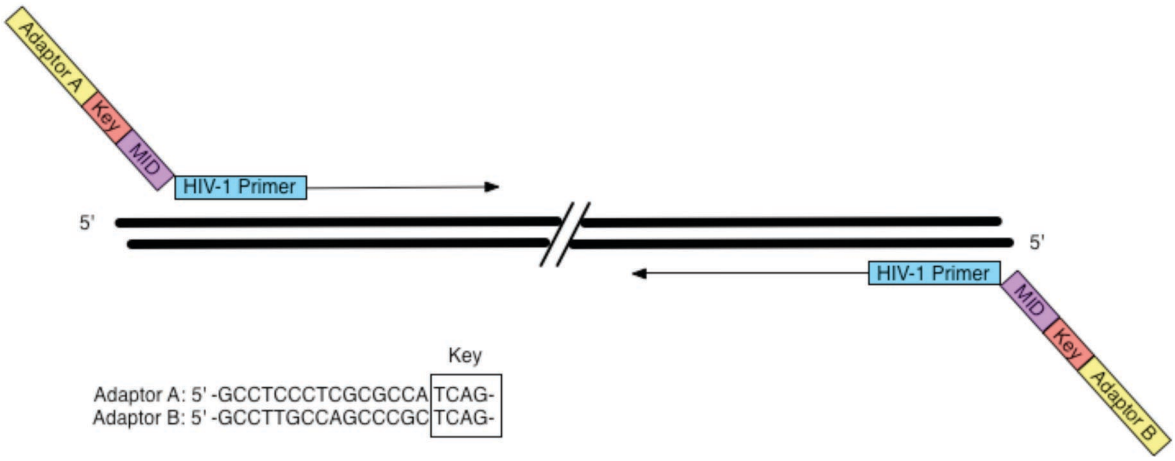

Fig. 8: HIV-1 primers for the amplicon generation include also the MID, the key sequence and an A and B adaptor

For the amplicon generation, 22µl PCR master mix[8] with the DNA polymerase and 3µl cDNA were mixed and the following PCR thermocycler program run performed:

One cycle of:	40 cycles of:	One cycle of:
94°C for 3 min	94°C for 15 sec	72°C for 8 min
	55°C for 30 sec	4°C hold
	72°C for 30 min	

6.5.5 Amplicon Purification

The PCR amplicons were purified with the AmPure kit (Agencourt AMPure) that based on the Solid Phase Reversible Immobilization SPRI para-magnetic beads purification technology. The SPRI beads have a polystyrene core that is first coated with a layer of iron for the magnetic response and finally encapsulated with a polymer layer. The polymer coating and small bead size of 1µm lead to a higher specific binding to nucleic acids larger than 100bp. The purification process consisted of three steps: immobilizing of the PCR amplicons, removal of the contaminations and amplicon elution.

First, 22.5µl of molecular grade water and 22.5µl of the PCR-product were added into each well on a polypropylene (PP)-plate. 72µl SPRI beads were added into each well and mixed 12 times using a pipette. Then the plate was placed on a 96-well magnetic ring stand (Ambion) and the magnetic field pulled the beads with the PCR amplicons out of the

[8] Master mix contained 105µl dNTP mix, 525µl FastStart 10x buffer, 52.5µl FastStart Hifi polymerase, 3937.5µl molecular grade water

solution. After incubation for 10 minutes the supernatant was removed with a good portion of contamination such as unincorporated dNTP's, primers and salts. Subsequently the plate was removed from the magnetic ring stand and the microparticels were washed twice with 100µl of 70% ethanol (Merck). The plate was then placed on the magnetic ring stand and the supernatant removed again. The pellets were dried at 40°C on a heat block. 10µl Tris-ethylenediaminetetraacetic acid (TE) (Fluka analytical) with pH 8 was added to each well to elute the PCR amplicons from the magnetic beads. Finally the purified amplicons were transferred to a new PCR plate.

6.5.6 Amplicon Quantitation by PicoGreen Flourescence

The Quant-iT™ PicoGreen® dsDNA Assay Kit (Invitrogen) contains an ultrasensitive fluorescent stain used for measuring the nucleic acid concentration. PicoGreen is a fluorochrome that has a high affinity to dsDNA and similar characteristics to SYBR-Green I. The advantage of PicoGreen over other fluorescent dyes is the low limit of detection down to 25pg/ml of dsDNA with a standard spectrofluorometer.

The first step of the protocol was the dilution of the PicoGreen dye 1:200 with TE buffer (200mM Tris-HCl, 20 mM EDTA, ph 7.5, Molecular Probes). 99µl of TE buffer and 1µl of DNA amplicons were transferred into each well on the 96-well black flourometer plate. Then 100µl of dye solution was added. Standard curves were constructed by serial dilution of the Quant-iT™ PicoGreen® dsDNA λ DNA Standard.

The black flourometer plate was analyzed on the Flourometer infinite M200 (Tecan) with the appropriate software application. The excitation wavelength of the PicoGreen stained specimen was 480nm and the emission wavelength of 520nm was recorded. The data was saved as an Excel file. The dilution was calculated as follows: With the intensity of the standard DNA samples (y) the correlation between the intercept d and the slope k for y= k*x+d was calculated. For the DNA concentration the formula x= (y-d)/k was used.

The following formula was used for molecules/µl:

$$\text{Molecules/µl} = \frac{sample conc[ng/\mu l] \text{ x } 6{,}022 \cdot 10^{23}}{656 \cdot 109 \text{ x amplicon length [bp]}} \qquad \frac{N}{V} = \frac{cN_A}{656 \cdot 10^9}$$

6.5.7 Agilent Measurement

PCR amplicons that were quantified below 5ng/µl were further evaluated on the Agilent 2100 Bioanalyzer (Agilent Technologies). The Agilent 2100 Bioanalyzer utilizes microchannels to separate nucleic acids fragments electrophoretically. This instrument yields information of the quantity of the dsDNA samples as well as quality by the determination of the amount and type of contamination. The size of the fragments that can be analyzed range from 25 to 1000 bp and the concentration of the DNA vary between 0.1-50ng/µl.If the band of expected size was present in the amplicon samples, the samples were used for amplicon pooling.

First the Agilent DNA 1000 chip[9] was placed on the Chip Priming Station (Agilent Technologies). 9µl of a gel-dye mix was pipetted in a marked well and and the priming station closed. After the distribution of the gel-dye mix on the DNA chip, the priming station was opened and 9µl of a gel-dye mix was put in two other wells. Then 5µl of the marker was pipetted in the ladder and the 12 sample wells. Finally 1µl of the sample ladder and 1µl of 12 purified amplicon samples were added into the marker and the chip run in the Agilent 2100 Bioanalyzer within 5 minutes. With the DNA-1000 series II assay the amplicon of an expected size was quantified.

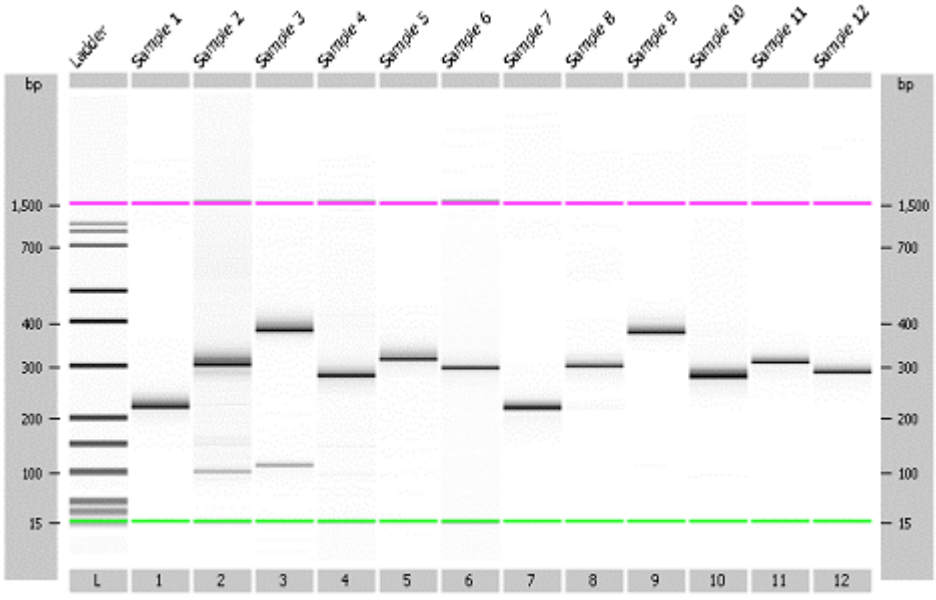

Fig. 9: Results of the electrophoresis on the Agilent 2100 Bioanalyzer: Sample 2 and 3 show primer dimers at a size of about 100 bp

[9] Agilent DNA 1000 Kit contained DNA chips, syringe kit, reagents: DNA ladder, DNA marker, DNA dye concentrate, DNA gel matrix

6.5.8 Amplicon Dilution and Pooling

After calculation the DNA molecules were diluted with TE buffer at a concentration of $1 \cdot 10^9$ molecules/µl. An equal volume of each of eight amplicons per sample was mixed. If some amplicons were not detected, the volumes of overlapping amplicons were augmented so that the whole PR and RT genes regions were covered. Nine samples showed a very low DNA concentration for which the repetition of the extraction was not possible and therefore were not suitable for ultra-deep sequencing. A maximum of twelve samples with different MID per plate were pooled together. Each of the six finally achieved pools was loaded on a lane of the picotiter plate device.

6.5.9 Capture Beads Wash and DNA Library Preparation

To reduce errors during the sequencing run, forward (A) and reverse (B) amplicons were used. The amplicons were sequenced in both, forward and reverse directions. The capture beads have bound on their surface area oligonucleotides that were complementary to adaptor A and B of the amplicons.

For each amplicon pool, 45µl capture beads dilution[10] that contained 900 000 A and B capture beads were used. The capture beads were washed twice with 500µl of capture bead wash buffer. The volume was determined and the capture beads aliquotted in 0.2ml tubes to an end volume of approximately 20µl. Then the DNA library that was repeatedly diluted at the concentration of $2x10^5$molecules/µl with bead wash buffer. Finally 4.5µl of $2x10^5$molecules/µl library template were mixed to A and B DNA capture beads for emulsion PCR to achieve one copy per bead.

6.5.10 Emulsion PCR

In order to increase the sensitivity of the sequencing step and to detect viral quasispecies with a low prevalence, a clonal amplification on beads, the emulsion (em) PCR, was performed.

During the emPCR, 250µl Mock Amplification mix was added into each emulsion oil tube and placed on Tissue Lyser for 25 oscillations/sec and the reaction homogenized for 5

[10] GS emPCR Kit II (Amplicon A, Paired End)and GS emPCR Kit III (Amplicon B, Paired End) contained bead recovery reagents II and III, clonal amplification reagents II and III, enrichment beads, emulsion oil

minutes. The live amplification mix [11]contained primer A and B that were bound to the 5' extensions of the DNA library. 160µl of the live amplification mix was added to the annealed library beads. The bead/amplification mix was put into the emulsion tube and placed on the tissue lyser for 15 oscillations/sec for 5 minutes. Thereby droplets were produced that served as microreactors when the PCR amplification occurred.
The emulsion reactions were transfered into a transparent 96 well plate (Qiagen) and the amplification program started:

One cycle of:	40 cycles of:	13 cycles of:
94°C for 4 min	94°C for 30 sec	58°C for 6 min
	58°C for 60 sec	10°C hold
	68°C for 90 sec	

During the first cycle at 94°C the in average one single DNA strand was bound to the complementary primer on the capture beads in an aqueous droplet. A huge amount of copies of each single DNA strand were produced clonally during the emPCR.

6.5.11 Bead Washes and Recovery

The emPCR was broken with a Swin-Lok filter syringe and washed three times with 9 ml isopropanol, once with 6 ml bead wash buffer and once with 9ml enhancing fluid. For the bead recovery the beads were aspirated with a syringe with Swin-Lok filter and expelled into a falcon tube. After centrifugation at 700rpm for 15 minutes the supernatant was removed.

6.5.12 DNA Library Bead Enrichment

First of all, the enrichment beads (=streptavidin beads) were washed. Per each emulsion reaction, 1 ml of enhancing fluid was mixed with 20µl of enrichment beads. After vortexing these paramagnetic beads were pelleted using a Magnetic Particle Collector (MPC) (Dynal). The supernatant was then removed and the beads resuspended in 100µl of enhancing fluid.

---

[11] Live amplification mix**:** 2905.9µl, $MgSO_4$: 160µl, amplification primer mix 33.28µl, Platinum HiFi Taq Polymerase: 96µl, PPiase: 4.8µl

Enrichment beads were coated with streptavidin that that fixed the DNA library. The DNA library was ligated to short adapters A and B adaptor, containing a 5' biotin tag (Fig. 10).

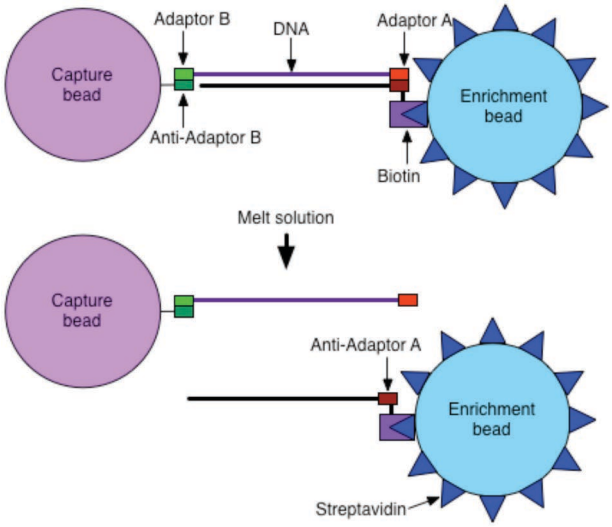

Fig. 10: Simplified, the dsDNA is bound to capture beads and to enrichment beads by the streptavidin-biotin binding

During the enrichment, 100µl of washed enrichment beads were mixed with amplified DNA beads and placed on a rotator. Then the beads were washed three times with 1ml of enhancing fluid. The tubes were placed on the MPC and the supernatant that contained capture beads without amplified DNA was removed.

Then the collection of the enriched DNA capture beads was performed: the pellets were resuspended in 700µl of melt solution[12], vortexed und pelleted with the MPC. The supernatant with the single stranded DNA was put to a separate 1.5ml tube and the DNA beads melt again. Then the DNA beads were washed three times with 100µl annealing buffer.

For primer annealing, 15µl of annealing buffer and 3µl of sequencing primer (Fig. 11) were added to the DNA capture beads and one cycle of primer annealing program (65°C for 5 min, 50°C for 1 min and 40°C for 1 min) performed. The beads were finally washed once with 100µl and once with 200µl annealing buffer and resuspended in 100µl annealing buffer.

[12] Melt solution was prepared with 0.125ml of NaOH (10N) and 9.875ml of PCR-H_20

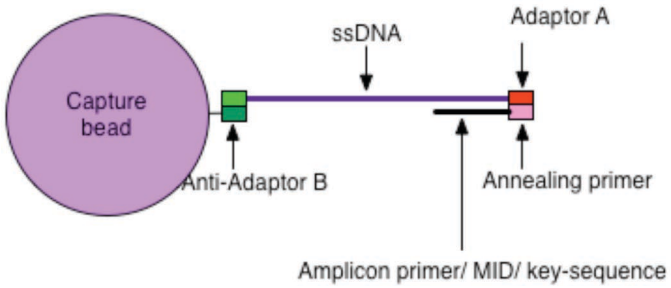

Fig. 11: The annealing (=sequencing) primer is complementary to adaptor A and B

6.5.13 Counting of the Beads

The achieve the right amount of beads that contain clonally amplified DNA for sequencing and to examine the beads loss during the bead recovery and enrichment, the DNA beads were counted on the Z1 counter (Beckman Coulter) before and after the annealing PCR.
According to manufacturer's package insert, 5-10% of this amout were left after enrichment and recommended for ultra-deep sequencing. Subsequently, the remaining forward and reverse beads of each amplicon pool were pooled together to achieve about 230 000 beads.

6.5.14 Ultra-deep Sequencing Technology

The first steps of the sequencing run involved the pre-wash of the GS FLX instrument and a wash step of the picotiter plate (PTP) with the BB2[13], using the GS-FLX LR70 Sequencing Kit (70x75) [14].
For the first layer of each of the four lanes on the PTP, 230 000 DNA beads with 10 000 control DNA beads with a known sequence were prepared. Then for the second layer packing beads that prevented the DNA beads of any movement when distributed on the picotiter plate were washed three times with BB2. 360µl packing beads were mixed with 80µl of the DNA polymerase containg bead incubation mix (BIM)[15]. As the last layer of the PTP, the enyme beads were prepared. The enzyme beads contained ATP sulfurylase and luciferase, enzymes necessary for the sequencing-by-synthesis reaction. They were

[13] BB2, the bead buffer 2 contained 200ml bead buffer and 34µl apyrase

[14] GS-FLX LR70 Sequencing Kit (70x75) contained reagents and enzymes, bead and wash buffer, pre-wash buffer tubes, buffer CB, Sipper tubes, packing beads

[15] BIM contained 785µl BB3, 75µl polymerase cofactor, 150µl DNA polymerase

washed three times with BB2 and then resuspended in 1ml BB2. 920µl enzyme beads and 980µl BB2 were then combined.

The PTP was loaded first with 900µl DNA bead suspension and 960µl BB2 per each lane. The PTP, an optical fiber chip, consisted of stamped in wells with a diameter of 44µm. The DNA beads with a diameter of 25-30µm were ideally placed into these wells. After 10 minutes of gravity depostion, the supernatant was removed and the second layer of packing beads was distributed on the PTP. The PTP was then centrifuged at 2800rpm for 10 minutes to fill out the whole wells with beads. In a typical sequencing run, 20-25% of the 1.6 million wells on a 70x75 mm picotiter plate device were filled out with DNA beads.

Subsequently the last layer of the enzyme beads was distributed on the PTP that was important for the sequencing-by-synthesis reaction (Fig 12).

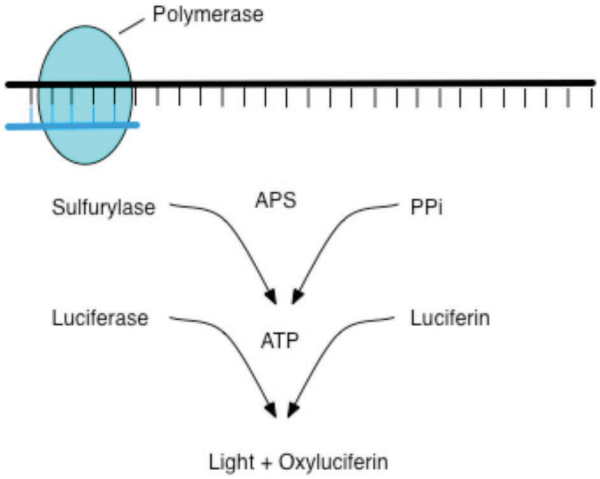

Fig. 12: The ultra-deep sequencing technique which is based on sequencing-by-synthesis

The fluidic sub-system that contained the flow reagents such as the four nucleotide triphosphates (dCTP, dGTP, dATP and dTTP), DTT, Adenosine 5′-phosphosulfate *(*APS), apyrase and buffers were put into the Genome Sequencer FLX instrument. Then the PTP device was loaded into the GS FLX instrument and the sequencing program started.

During the sequencing program the sequencing reagents were delivered in a repetitive order into the PTP device.

The sequencing-by-synthesis reaction occurred when the DNA polymerase elongated the DNA sequencing template from the annealed primer by the addition of nucleotides.

When a nucleotide was incorporated into the growing DNA strand, pyrophosphate was released (Fig. 12). The ATP sulfurylase converted its substrates pyrophosphate and APS to Adenosine-5'-triphosphate (ATP). ATP on the other hand activated the enzyme luciferase

to produce oxyluciferin and a detectable light signal. The sequencing byproducts such as unused nucleotides were washed away with the enzyme apyrase.

A CCD camera recorded each signal light from each well on the PTP device. The light intensity was directly proportional to the numbers of nucleotides that were inserted. In a region where i.e. three dCTP were incorporated, the light signal was strong. The light intensity also gave information about the position of the bead on the PTP.

6.5.15 Data Analysis

For the processing the Genome Sequencer software and for data analysis of amplicon variants the GS Amplicon Variant Analyzer (AVA) Software was used.

The processing of the data consisted of the run-time phase and the post-run phase. During the run-time phase the image acquisition of raw images, the image processing and signal processing were performed. This resulted in Standard Flowgram Format (SFF) files that contained flowgrams (Fig. 13).

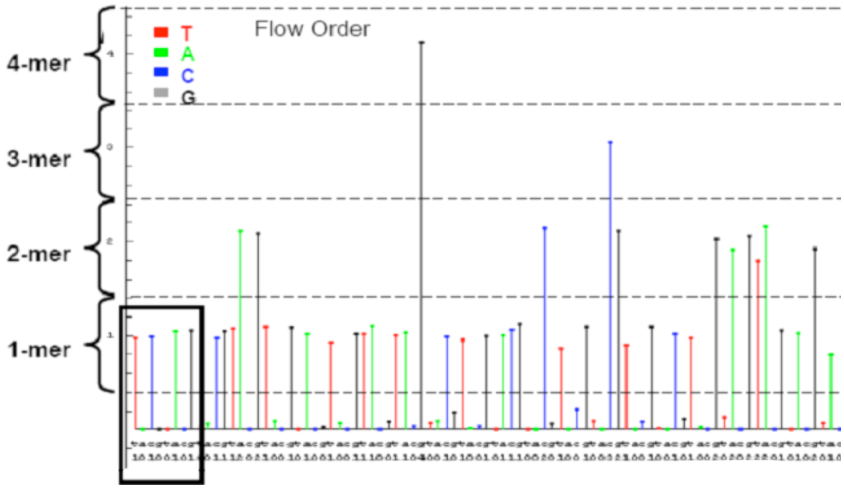

Fig. 13: A typical flowgram that shows the intensity of the light signal when a nucleotide is incorporated. The signal level was determined and controlled by the use of parallel sequenced control bead signals and depends on the number of nucleotides incorporated.

To qualify for being a high quality read each sequence has to pass diverse filters. Filters for the match to the MID's and the key sequence, filters for the amplification primers and signal intensity filter were checked.

Each flowgram begun with the four base key sequences that was for the DNA beads the TCAG and for the control beads ACTG sequence.

The Genome Sequencer software provided for the post-run phase the GS De Novo Assembler, the Reference Mapper and the Amplicon Variant Analyzer. The De Novo Assembler aligned the individual sequences into "contigs" which are overlapping mulitple reads and created a consensus sequence. The Reference Mapper compared the consensus sequence to a known reference sequence of HIV-1 HXB2. With the Amplicon Variant Analyzer the mutations were searched and quantified relative to the reference sequence. The variant results were displayed as forward and reverse denominations. The denominations were aligned reads of the overlapping amplicon regions. The so called forward and reverse "hits" per denomination were the variants being relevant for drug resistance.

With the use of avaProjectManager.pl a data file of the results was created and exported into Excel. The column "variant pattern" is coded as follows: m means that there is a match between the bases of the reference sequence and s means that there is a substitution of the base. The brackets following these characters show the codon position where the match or substitution is localized. For example the variant pattern m(994)s(995,G)m(996) means that at the codon position 994 is a match, at the position 995 the base is substituted with a guanine and at the position 996 is a match. The variant pattern of three bases resulted in some cases in the same amino acid. Each mutation was included in the results.

6.5.16 Stanford University HDRM Database

For the detection of relevant major and minor PI and NNRTI and NRTI resistance mutations the HIV Drug Resistance Mutations (HDRM) Database of the Stanford University (http://hivdb.stanford.edu/) Version 6.0.7 updated on 2009 Dec 28 was used.

The Stanford HDRM database is a publicly available database for research and educational purposes in contrast to the proprietary TruGene HIV-1 assay interpretation system.

Drug resistance mutations are defined as the association between the antiretroviral treatment and the variants that emerge (genotype-treatment correlation), the in vitro drug susceptibility results to the genotype (genotype-phenotype correlation) and the virological response to a new treatment (genotype-clinical outcome correlations) (Shafer, 2006).

The database website consists of three programs. The HIVseq program compares user-submitted protease and RT sequences to a reference sequence and interrogates mutations. The second program, the HIVdb provides a drug interpretation of the user-submitted data with a penalty score. This score reports the susceptibility of a drug (Rhee et al, 2003). The

last program, the HIValg allows the user to compare the results to any resistance algorithm that is publicly available, also developing new algorithms with the use of the Algorithm Specification Interface (ASI) (Betts et al, 2003).

The TG GuideLines Rules 14.0 (Siemens) and Stanford HDRM were evaluated for the FDA-approved antiretroviral drugs. They characterize almost the same antiretroviral drugs. Also the amino acid positions that are associated with drug resistance of both systems were investigated and were partly discordant.

Both systems include eight equal Ritonavir boosted PI's (Fosamprenavir, Atazanavir, Darunavir, Indinavir, Lopinavir, Nelfinavir, Saquinavir, Tipranavir). The six NRTI's (Abacavir, Didanosine, Lamivudine/Emtricitabine, Stavudine and Tenofovir) are also included in both drug resistance interpretations systems.

Three equal NNRTI's (Efavirenz, Nevirapine and Etravirine) are mentioned with both interpretations systems except for Delaviridine that was referred additionally by the Stanford database.

The discrepancies among the interpretation systems result from the complexity of the genotyping results and the use of different algorithms for susceptibility levels. An agreement between the HIV databases to a consensus algorithm for the drug resistance interpretation is nevertheless necessary and also evaluated, i.e. by the IAS (Kijak et al, 2003).

7 Results

7.1 Accuracy Testing

For accuracy testing, five samples with genotype B obtained by the QCMD proficiency program were analyzed on the Genome Sequencer FLX. All relevant drug resistance mutations obtained by using the Stanford HDRM database were found. All other mutations except of one were found with the UDS method (Tab. 1).

Tab. 1: Mutations that were ultra-deep sequenced: non relevant drug resistance mutations are marked red. The green marked mutations were found additionally with the UDS method.

PI Mutation position	Amino acid	number
10	I	3
11	**I**	**1**
13	V	2
20	**K**	**1**
35	**D**	**4**
36	I	3
43	T	1
46	I	3
54	V	3
62	V	3
63	P	3
63	**A**	**1**
71	T	3
77	I	1
82	A	3
90	M	3
93	L	4
93	**L**	**1**
Total		43
NRTI Mutation codons		
98	**S**	**1**
179	**V**	**2**
181	**Y**	**4**
Total		6
NNRTI Mutation codons		
41	L	1
67	G	1
116	**F**	**1**
184	V	1
210	W	1
215	Y	1
Total		6

With the UDS method, non relevant PI mutations such as 20K, 35D and 63A, NRTI mutations such as 179V and 181YI and the NNRTI mutation 116F were non identified

originally because the were not on the Stanford HDRM list. They were all but one identified on a subsequent anaysis. One silent mutation, 20K, was not detected correctly according to QCMD. The consensus sequence of the nucleotides was defined as ARR (R=A or G) by the QCMD. However, the sequencing of this mutation with TruGene at the Institute of Hygiene of the Medical University of Graz resulted in AAG, which is equal to the results of the ultra-deep sequencing. Because the UDS results were identical to the TruGene results, the same performance score of 315 was achieved. The maximum score was 316.

Three mutations (marked green in Tab. 1) were found additionally with the UDS method. The PI mutation 93L of the ENVA06-02 sample with a viral load of 66 900 copies/ml was found at a prevalence of 2.7%. The PI mutation 11I of the ENVA06-01 sample with the viral load of 8 090 copies/ml was also found at a low prevalence of 1.4%. The NRTI mutation 98S of the ENVA06-01 sample was detected at a prevalence of 99.3%.

The viral load varied from 4 680 to 66 900 copies/ ml and accounted in average 19 648 copies/ml. With the UDS method, 8 636 reads/sample and 43 183 reads for all samples in total were analyzed. 70 959 forward and reverse reads per variant region (also called denominations) and 39 591 forward and reverse drug resistance variants (also called hits) were sequenced in total.

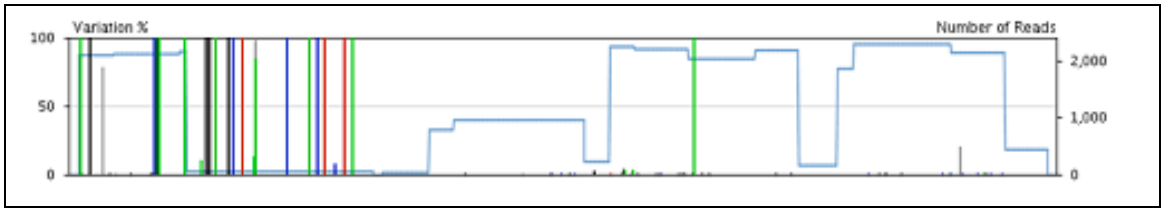

Fig. 14: The variation histogram shows multiple alignments of the variants of the individual ENVA 09-03 sample

7.2 Quality Parameters of the UDS Method for an Austrian Cohort

The sequence reaction carried out in the first run yielded 868 322 beads run and 775 502 beads in the second run (Fig. 15, 16). During the first run, 616 585 passed the key pass filter and 247 541 passed the quality filter which consisted of a perfect match to the key sequence and the primer region. During the second run, 560 096 passed the key, and 253 565 passed the quality filter. This quality control resulted in a 28.5% elimination of the beads during the first run and 32.7% of the beads in the second run.

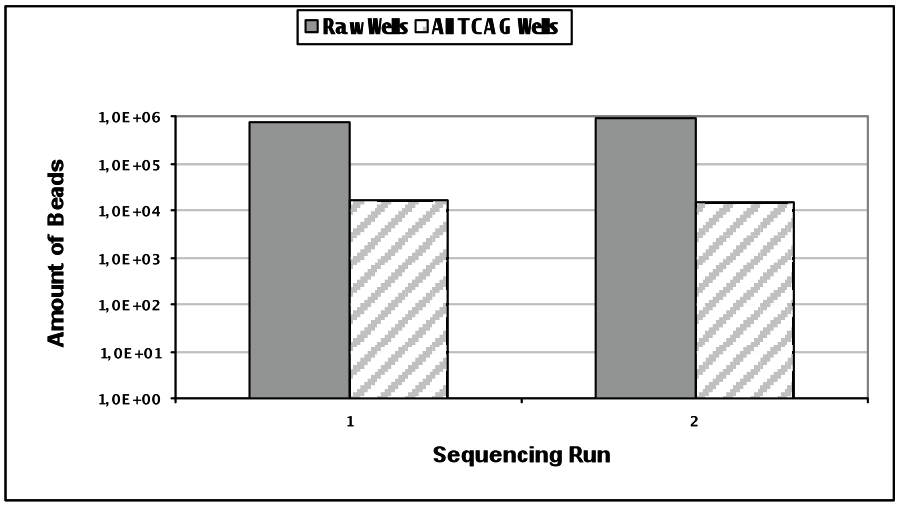

Fig. 15: Amount of raw wells that started the sequencing reaction and TCAG wells that passed the filter

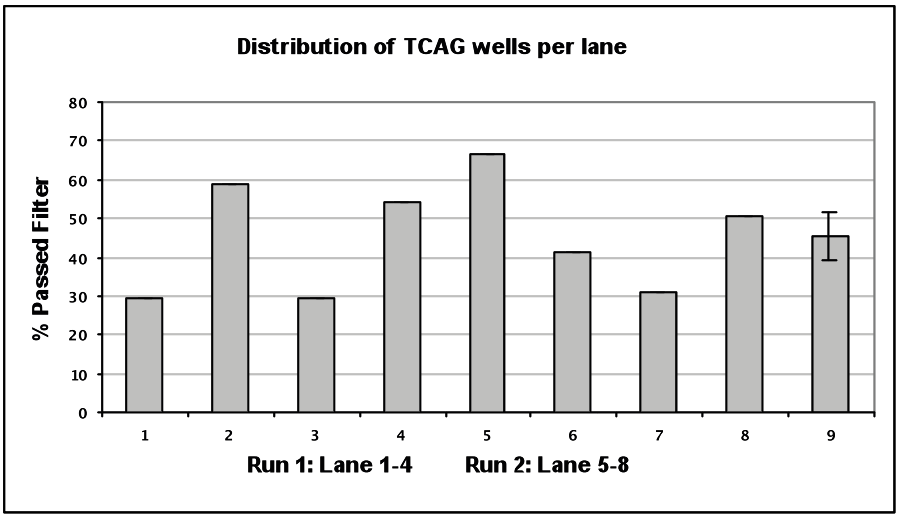

Fig. 16: Lanes of the sequencing run 1(pool 1-4) and run 2 (pool 5-8) that passed the filter (%) more detailed and the mean value (lane 9)

After multiple reads were aligned, the data was assembled into "contigs". The reads were mapped to the HIV-1 HXB2 reference sequence, high quality reads were filtered generating a high confidence consensus sequence against the reference sequence.

The total number of forward and reverse reads for 45 samples was 3 518 636 (1 268 636 forward and 225 228 reverse reads). An average GS-FLX run yielded for 45 samples 282.9 forward and 502.7 reverse reads per amplicon region, which is 785.6 reads for both directions. Precisely, 180 084 variants (hits) for each relevant drug resistance mutation were found.

7.3 Distribution of HIV-1 Variants

To estimate the distribution of HIV-1 variants in the investigated Austrian patient's cohort (n=45), the results of the standard sequencing based on the GuideLines Rules 14.0 (Siemens) algorithm included in the TruGene (TG) Assay and the results of the UDS based on the Stanford database mutations list (last updated Dec 28, 2009) were compared.
In total, with both sequencing methods, 170 identical mutations were detected (Tab. 2-4).
98 mutations were identified only with ultra-deep sequencing and 18 only with the standard sequencing method according to the genotyping systems as complete packages. 44 diverse mutation codons were observed. The criteria for the UDS method were defined as ≥1% hits and 10 reads for a mutation.

To analyze the mean number of drug resistance mutations detected by the standard sequencing method and UDS, the paired Student T test was performed. The mean number of mutations of each sample per drug class (PI, NRTI and NNRTI) was calculated in a 95% confidence interval (CI) and the P values were presented two tailed (Fig. 17).

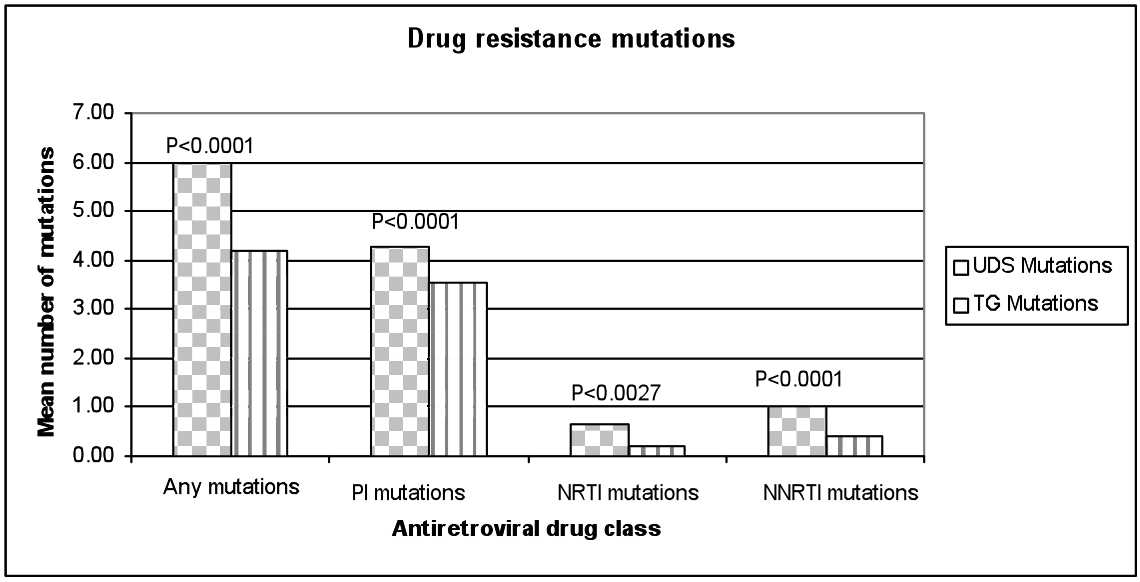

Fig. 17: Mean number of mutations per drug class and patient

The total number of mutations detected with the standard sequencing method shows a mean number of 4.18 mutations per sample (95% CI, 3.67-4.69) and 5.96 for the UDS method (95% CI, 5.20-6.71), respectively (P<0.0001).

The mean for the 160 PI mutations found with the standard method is 3.56 mutations (95% CI, 2.52-4.59) per sample and 4.29 mutations (95% CI, 3.04-5.54; P<0.0001) for the UDS method.

Eleven NRTI mutations only were found with the standard sequencing method with a mean value of 0.24 (95% CI, 0.17-0.32) and 29 were detected with the UDS method, with a mean of 0.64 (0.46-0.83; P<0.0027).

Seventeen NNRTI drug resistance mutations were found with the standard sequencing method with a mean of 0.38 (95% CI 0.27-0.49) and 46 were detected with the UDS method with a mean of 1.02 (95% CI, 0.72-1.32; P<0.0001).

Tab. 2 shows the comparison of PI resistance mutations according to standard sequencing method with TruGene analysis and the UDS method with Stanford HDRM at the prevalence ≥1% and >20%. The last column shows drug resistance mutations that were found with both methods.

Tab. 2: Comparison of the PI mutations with both sequencing methods

Sequencing Results PI					
Mutation codons	AA	Standard	UDS ≥1%	UDS >20%	both
10	I	11	14	11	11
10	V	2	3	2	2
11	I	2	1	1	2
13	V	9	13	8	9
15	**V**	7	0	0	0
16	**A**	0	1	0	0
16	E	4	10	4	4
19	**I**	5	0	0	0
20	R	1	3	1	1
33	V	0	1	0	0
35	G	0	2	0	0
36	I	9	20	9	9
36	**T**	0	1	0	0
36	V	0	1	0	0
43	T	2	3	3	2
60	E	4	4	4	4
62	V	23	23	18	23
63	**T**	4	0	0	0
63	P	30	35	31	30
69	**K**	1	0	0	0
71	T	5	6	5	5
71	V	1	2	1	1
77	I	17	22	17	17
82	**I**	0	3	2	0
85	V	0	1	0	0
89	M	1	1	1	1
93	L	22	23	22	22
Total		160	193	140	143

As can be seen on Tab. 2, 160 mutations were found with the standard method, 143 of them were equal with the UDS method. 193 mutations were found with the UDS method at a prevalence ≥1% and 140 at a prevalence >20% with partially good concordance to the standard method.

The blue marked mutations 15V, L19I, L63T and H69K were only relevant for the TruGene guidelines. All other TruGene mutations were found with UDS, except for those seventeen that were only included in the TG assay. With the Stanford HDRM list, five red marked mutations (16A, 36T, 82I) were found additionally with the UDS method (Tab. 2)

The TruGene result described the mutation variant 63P/T. This was counted only once according to the first amino acid i. e., 63P. Ten UDS PI mutations resulted in the same amino acid but had a different base. Each mutation was accounted individually.

Tab. 3 indicates the sequencing results of the NRTI for both sequencing methods. Eleven NRTI mutations that were found with the standard method are in high concordance with the twelve UDS mutations found at a prevalence ≥20%.

Tab. 3: Comparison of NRTI mutations with both sequencing methods

Sequencing Results NRTI					
Mutation codons	AA	Standard	UDS ≥1%	UDS >20%	both
41	L	1	1	1	1
62	V	0	1	0	0
65	R	0	2	0	0
67	**E**	0	1	0	0
67	**G**	0	1	1	0
69	N	2	2	1	2
69	D	0	2	0	0
69	**S**	0	1	1	0
70	E	0	1	0	0
75	**L**	0	1	0	0
118	I	1	2	1	1
181	**C**	1	0	0	0
184	I/V	1	2	1	1
184	V	0	4	1	0
210	**W**	1	1	1	1
210	S	0	1	0	0
215	S	3	3	3	3
215	D	0	1	0	0
215	F	0	1	0	0
219	E	1	1	1	1
Total		11	29	12	10

The relevant drug resistance mutation 181C which is marked green in Tab. 3 was found only with the standard method. Five red marked mutations (67E, 67G, 69S, 75L, 210W) were listed only in the Stanford HDRM list. The mutation mixture 184I/V was counted as 181I. In total, 29 NRTI mutations were found with the UDS method.

As shown in Tab. 4, 17 NNRTI drug resistance mutations detected with the standard sequencing method and the 46 with the UDS method.

Tab. 4: Comparison of the NNRTI mutations detected with both sequencing methods

Sequencing Results NNRTI					
Mutation codons	AA	Standard	UDS ≥1%	UDS >20%	both
90	I	3	4	3	3
98	**S**	0	1	1	0
101	Q	2	2	2	2
101	**R**	0	4	2	0
103	**E**	0	4	0	0
103	**R**	0	3	2	0
103	T	0	2	0	0
106	A	1	1	1	1
106	I	3	3	2	3
108	I	0	1	0	0
138	A	4	7	5	4
138	K	0	1	0	0
179	D	1	2	2	1
179	I	0	4	4	0
179	T	0	1	0	0
190	A	2	3	2	2
190	E	0	1	0	0
227	L	1	1	1	1
238	**R**	0	1	1	0
Total		17	46	28	17

Thirteen mutations of the codons 98S, 101R, 103E, 103R, 238R were described additionally with the UDS method. Unlike the PI's where at the prevalence >20% less mutations were found with the UDS method (149 vs 160), the NNRTI were found more often with the UDS method (29 vs 18) than with the standard sequencing method. All in all, NNRTI mutations were detected more frequently with the UDS method (47.3%) than with the standard method.

The prevalence of the mutations of the Austrian cohort were compared to an American study, Simen et al, 2009 based on the Flexible Initial Retrovirus Suppressive Therapies (FIRST) Study. The different distribution of each drug class is shown in Tab. 5.

Tab. 5: The left column shows the number of the study mutations, the right column shows the number of mutations identified by Simen et al, 2009

	Study cohort		Simen et al, 2009	
Drug class	Standard	UDS	Standard	UDS
PI	160 (85.1%)	193 (72.0%)	6 (15.4%)	12 (13.8%)
NRTI	11 (5.9%)	29 (10.8%)	16 (41.0%)	36 (41.4%)
NNRTI	17 (9.0%)	46 (17.2%)	17 (43.6%)	39 (44.8%)
Total	188	268	39	87

The study cohort shows the highest prevalence for PI's (85.1% and 72.0%) that in contrast are represented in Simen et al, with 15.4% and 13.8%. The NRTI appear with the lowest prevalence (5.9% and 10.8%) in the study cohort and represent about 41% in in Simen et al. The percentage of the NNRTI shows 9.0% with the standard method and like for the NRTI's almost twice as much (17.3%) for the UDS method in the study. Simen et al, shows similar percentages for both drug classes (NRTI and NNRTI) and methods.

Using the UDS method based on the Stanford HDRM, 32.8% of low-abundance mutations (≤20% prevalence) were detected additionally. Regarding PI- NRTI- and NNRTI, 27.5%, 58.6% and 39.1% more resistance mutations were recognized at low prevalence levels with the UDS method, respectively. A more detailed subdivision is pointed out in Fig. 18:

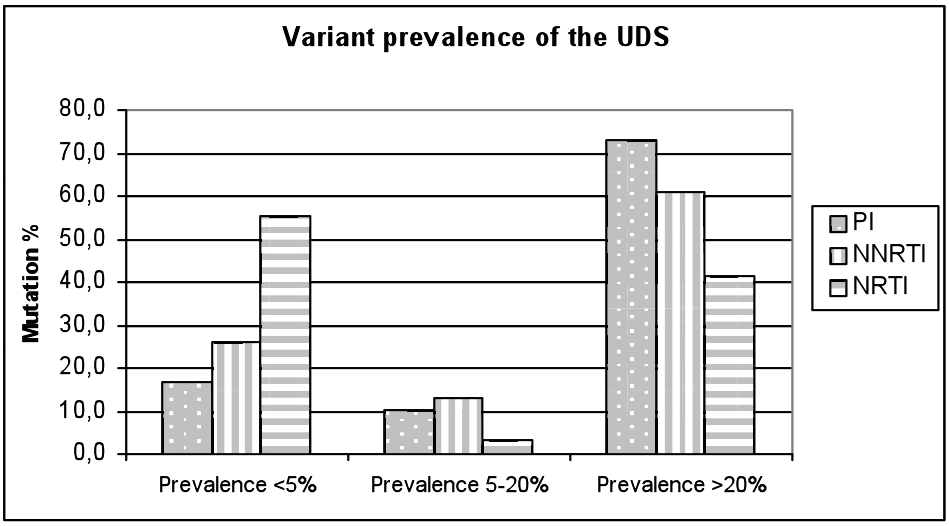

Fig. 18: Distribution of the drug variants at a prevalence level of <5%, 5- 20% and >20% of the viral population

Tab. 6 shows the number and percentage of detected PI, NRTI and NNRTI mutations divided into groups of <5%, 5-20% and >20% of prevalence.

Tab. 6: Amount and percentage of mutations found with the UDS method

	Prevalence <5%		**Prevalence 5-20%**		**Prevalence >20%**	
	Amount	%	Amount	%	Amount	%
PI resistance mutations	32	16,6	20	10,4	141	73,1
NRTI resistance mutations	16	55,2	1	3,4	12	41,4
NNRTI resistance mutations	12	26,1	6	13,0	28	60,9
All muations	60	22,4	27	10,1	181	67,5

In this study cohort the major fraction of drug resistance mutations for PI's and NNRTI's was found with the UDS at a prevalence >20%. The lowest numbers of resistance mutations were detected within the range of 5 to 20%. From the low abundance variants, the majority was found below 5%. 55.2% of total NRTI resistance mutations were found <5% prevalence.

Five samples that were ultra-deep sequenced had a low viral load (below 10 000 copies/ml). As shown in Tab. 7, all mutations that were found with the standard sequencing method could also be detected with the UDS method except one.

Tab 7. Five samples sequenced with a low viral load

		UDS method			**Standard sequencing**		
Sample	Viral load	PI	NRTI	NNRTI	PI	NRTI	NNRTI
4	1 380	3	0	0	2	0	0
7	1 570	2	1	2	2	0	0
9	4 220	5	0	0	2	0	0
49	3 440	5	7	1	4	3	0
50	4 110	4	2	6	3	0	2
		19	10	9	13	3	2

The percentage of additionally detected mutations with the UDS method does not correlate well with the above described proportions: 19 vs 13 PI mutations, 10 vs 3 NRTI and 9 vs 2 NNRTI's. The number is however too small for a statistically significant result.

The NRTI mutation Y181C of sample number 49 is according to both sequencing systems a clinically relevant drug resistance mutation. This was the only mutation that was found with the standard sequencing method solely.

Two samples that had a high number of mutations detected with standard sequencing method (sample nr. 9 and 49) also showed a high number of mutations in ultra-deep sequencing.

7.4 Reproducibility Testing

Reproducibility was evaluated by repeated ultra-deep sequencing of 21 samples. The DNA library preparation was equal for two runs, the runs varied in the repeated emPCR preparation of the beads. The linear regression of the variants prevalence showed a correlation value of R^2=0.969 (Fig. 19 A). This indicated a good concordance of the prevalence of variants for both runs. However, the correlation between the number of reads per variant (hits) showed a poorer correlation of R^2=0.749. This is compatible with differences in DNA capture bead preparation (Fig. 19 B) which is stated normal according to the manufacturer's protocol.

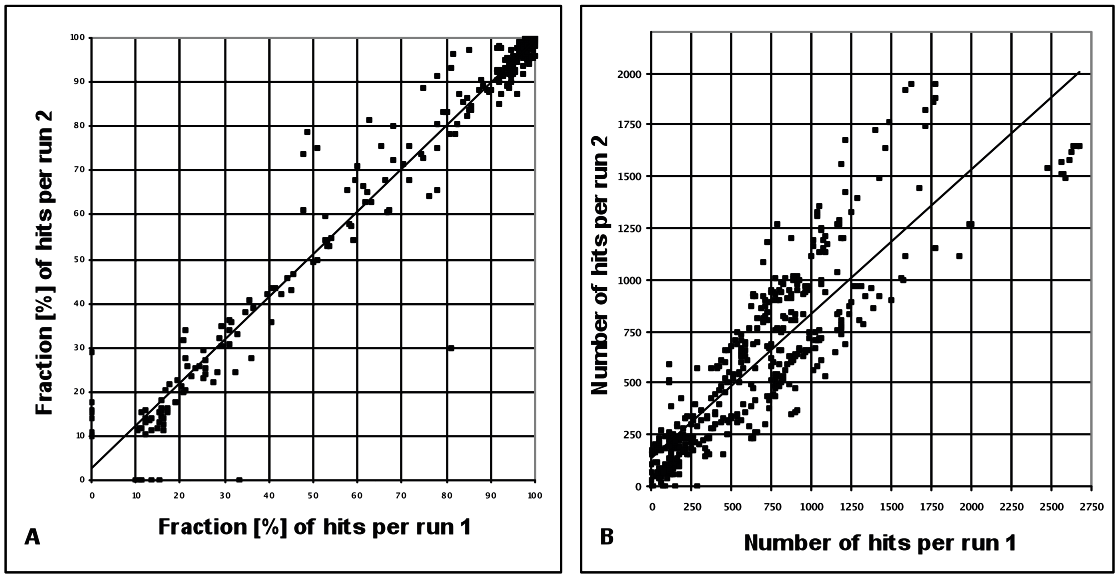

Fig. 19: Correlation between two runs: Fig. A represents the fraction (%) between two runs with a R^2=0.969 and Fig. B the hits with R^2=0.749

8 Discussion

The aim of this diploma thesis is the comparison of two different genotyping resistance systems as complete packages (extraction, reverse transcription, amplification, sequencing and corresponding data analysis together) against each other. So the given discussion on the differences of these systems is held on the complete sequencing methods with all their procedure. The different process steps are listed in particular in Tab. 8.
Other studies that also evaluated the UDS method, i.e. Le et al, 2009 and Simen et al, 2009 used the same extraction volume and the same extraction kit (QIAmp RNA Mini kit, Qiagen) for a direct comparison of the standard and UDS method. In this study, the extraction volume and the extraction kit were different, due to different protocols.

The accuracy was tested with five QCDM samples. All but one silent mutation were found according to the QCMD results. The performance score of the detected mutations was found in 99% agreement with the QCMD. The mutations were identified according to the International AIDS Society (IAS) USA drug resistance table 2005.
Two mutations were found additionally show a prevalence of 2.7% and 1.4% which is acceptable because this is the low abundance area. In contrast, one mutation was found at a high prevalence of 99.3% in advance. This mutation was according to the Stanford HDRM list relevant, but not according to the IAS (Johnson et al, 2005).

In this retrospective study the sequencing data of the 45 specimen of an Austrian cohort with the HIV-1 subtype B was anaylzed. The specimens were all sequenced previously with the standard sequencing and specimen with one or more drug resistance mutations were collected and ultra-deep sequenced. In total, 170 clinically relevant drug resistance mutations were detected identically with both methods.
With the UDS method a significantly larger number of 98 mutations were detected additionally of which 23 were not listed in the TG guidelines. With the standard sequencing method, 18 mutations were detected in advance. 17 of them were not listed on the Standford HDRM list, and hence not identified.

To verify the hypothesis that with the UDS method more mutations were identified, the paired Student T test was chosen. In a normal distribution of drug resistance mutations per patient, a 95% confidence interval and the p values were calculated two tailed (Fig. 17). As expected the UDS method shows a higher mean number of drug resistance mutations per sample (5.96 vs 4.18 mutations) but has a higher standard deviation (2.63 vs 1.75). The higher sensitivity of the UDS method leads to a higher number of detected drug resistance mutations, especially in low prevalence field. Mutations that were identified with the UDS method at a prevalence >20% were listed in Tab 2-4 and allowed a direct comparison to the standard method, that in general doesn't detect mutations in the low prevalence field. More PI's were found with the standard method (160 vs 140) at a prevalence >20% but the concordance at the mutation codon level was nevertheless noticeable. On the other hand, the NRTI's and NNRI showed better concordance of the standard method with the UDS >20%.

Simen et al, evaluated if the new UDS method is a more sensitive method for the detection of low abundance mutations. Therefore, UDS was performed on a subset of the randomized FIRST Study on 264 samples. The previously untreated US participants received one of three different initial antiretroviral therapy strategies with a NRTI as a background. The participants were followed up for five years, with the determination of CD4 cell count, HIV-RNA level and history of an AIDS-defining event. 35 participants (13.6% of 258) showed resistance mutations with the standard method, 73 participants (28.3% of 264) showed resistance mutations with the UDS method, respectively. More than one mutation according to Standford HDRM was found in 11 samples (4.3%) by the standard sequencing, and in 22 samples (8.5%) with the UDS method. 33 mutations were detected with both methods, two had a mutation detected only by the standard sequencing method and 40 had a mutation detected with the UDS method. The results of this study were compared to the Austrian cohort in Tab. 5. The high number of PI is noticeable with the Austrian cohort (84% and 72%), whereas the American shows a high percentage of NRTI (41%) and NNRTI (44%) mutations.

Recent studies, e.g. Le et al, suggest that minor drug resistance mutations are associated with a higher risk of virological failure of the antiretroviral therapy.

This study results show that about a third (32.8%) of clinically relevant drug resistance mutations that were identified additionally with the UDS method are in the low prevalence area comprising less than 20% prevalence. More precisely, 27.5% PI's, 58.6% NRTI's and 39.1% NNRTI's were found with the UDS method at a prevalence ≤20%.

Le et al, detected in 22 subjects 36% (90 of 247 mutations) minor drug resistance mutations that were unrecognized with standard sequencing and correlated with virolical failure and historical antiretroviral use.

Simen et al, showed that 39.8% (45 of 113) mutations were found in the low abundance region. Simen et al, demonstrates that the UDS method identified a significantly larger proportion of low abundance drug resistant HIV variants in chronically HIV-infected and ART-naive persons that may not be detected with the standard method. Additionally, this study showed that all individuals with high and low prevalence NNRTI-resistance mutations detected with UDS method experienced virological failure.

The results of this study show a high amount of 39.1% of minor NNRTI variants that according to Simen et al, lead to a subsequent virological failure. Other recent studies (Johnson et al, 2008 and Paredes et al, 2010) also provide proof that especially NNRTI mutations are associated with poor treatment success. The study samples with low abundance NNRTI mutations should be therefore further evaluated because of the importance referring to the clinical outcome.

Because of the low number of participants with PI mutations a significant association to the virological failure was not possible in the study by Simen et al. In the Austrian cohort however, a very high number of PI mutations was found with both methods, which is in good concordance to Le et al. that also shows that the highest number of mutations is contributed to PI.

A combination of two NRTI's form the backbone of each antiretroviral therapy and are more effective regarding drug resistance than a single NRTI agent (Dietrich, 2006). The reduced risk of resistance occurance in the HAART therpy leads to a more intensive investigation of other drug classes such as PI and NNRTI. However, the influence of low abundance NRTI should be more intensive evaluated in future studies.

Of all detected low abundance drug resistance mutations (32.8%), the majority was found at a prevalence <5% as shown in Fig. 18 and Tab. 6. 10.1% were found at a prevalence form 5 to 20% and a much bigger part of 22.4% under 5% prevalence of all drug resistance

mutations. The results were in agreement to Simen et al, that showed that the major part of minority drug resistance mutations is at a prevalence level of <5% and accounted 39.8% of all mutations in that study. Le et al, detected 62% of the low abundance mutations at levels 1 to 5% and 38% at levels 5 to 20%.

One major finding of this study is that the new ultrasensitive UDS method has shown to be satisfying to detect minor drug resistance variants with a low level of prevalence of 1%-20% which is the limitation of the standard sequencing method. However, the sensitivity of the UDS method is limited by two reasons: first, the viral load and second, number of reads per nucleotide position.

For the standard sequencing method, a minimum of viral load in plasma samples according to Kuritzkes et al, 2003 and the package insert instructions of the TruGene^{TM} HIV-1 Genotyping Assay was 1 000 copies/ml from extracted RNA which indicates a high sensitivity.

As reported by Wang et al, 2007 and Margulies et al, 2005 and recommended by the manufacturer's package insert, samples containing at least 10 000 copies/ml should be used to gain a more accurate proportion of variants down to 1% prevalence. Wang et al, 2007 characterizes the error rate that can occur during the PCR amplification and ultra-deep sequencing. The low abundance variants at a prevalence of ≥1% can be detected reliably with an error rate of 0.1% when a viral load of at least 10 000 copies/ml is used.

Despite the fact that the samples are bound to a high viral load, the original amount of plasma needed for sequencing is nearly the same for both methods as showed below.

400µl plasma of each specimen was used in the standard method and eluted in 50µl elution buffer. One third of this amount (17µl) of this was used for further sequencing. Effectively 133µl are sufficient for a single sequencing run.

In contrast, a large volume of 1ml plasma was used for the UDS method and eluted in 75µl elution buffer. Only 15% of this amount (11.5µl) were used for further sequencing which means that 150µl of the origin plasma is needed for a sequencing run. All in all, the UDS method needs about 10% more plasma for the analysis. The high original plasma volume for the UDS method might be beneficial for the reliable detection of minor variants. Le et al, and Simen et al, used a plasma volume of 140µl because the same volume was used for the standard sequencing. However, Simen et al, advised to extract the HIV RNA from larger volumes.

Because of the high extraction volume of 1ml for the UDS method, in this study also samples with a viral load beneath 10 000 copies/ml were sequenced. Five samples with a viral load beneath 10 000 copies/ml had no sufficient DNA content (<5ng/ml) when the amount was measured with the PicoGreen fluorescence. They yielded no sequencing results and were excluded from the further analysis.
Five other samples with RNA copy numbers ranging from 1 380 to 10 000 were sequenced successfully to investigate the influence of the viral load on the detection of variants.
Le et al, ultra-deep sequenced six samples with a viral load below 10 000 copies/ml. Like Le et al, these study samples may not represent the actual proportion of drug resistance mutations.
One of these samples had the clinically relevant mutation Y181C that was identified with the standard sequencing but not with the UDS method. According to the Standford HDRM list and the TG guidelines, this mutation was relevant. The viral load of this sample was 3 440 copies/ml. This confirms the assumption that a high viral load is crucial for the correct detection of HIV-variants.
Concerning the total number of mutations detected by both sequencing methods of the five samples, with the standard method 18 mutations and with the UDS method 38 mutations were found which is almost the double amount. In comparison to the total proportions, this samples show deviations and may also not describe the actual proportion. However, the sample number is too small to draw significant conclusions and should be evalutated for a higher number of samples.

The sensitivity of the UDS method depends also on the number of reads per nucleotide position. In this study, the mean amount of reads per nucleotide position was 785.6 in forward and reverse direction while i.e. Simen et al. in contrast has an average coverage of 1700 reads. To achieve a higher sensitivity for the UDS method, especially low abundance mutations, the number of reads should be increased and new introductions, like the GS FLX Titanium series reagents applied.

The main disadvantage of the new UDS method for the application in clinical routine testing is the long and error prone manual DNA library preparation. Tab. 8 shows the different processing steps for the standard and for the UDS method. In total the standard method requires 9.25 hours for one sample. Up to eight samples can be processed parallel

in a single run with minimal variation of the total preparation time. With the UDS method, one sample can be analyzed in 37 hours. The additional preparation time per sample is also longer because of the timewithin consuming manual steps.

Tab. 8: The manual preparation of the UDS method

Steps	TruGene Assay	Time (h)	UDS method	Time (h)
Sample	MagNA Pure Kit I	0.75	High Pure Viral Kit	2
Amplification	RT-PCR	4	RT-PCR	4
Purification	-	-	Agencourt AMPure	12
Quantification	-	-	PicoGreen	2
Sample prep.	-	-	Dilution	4
Sequencing	CLIP sequencing	4	emPCR UDS	11
Data analysis	OpenGene software GuideLines 14.0	0.5	AVA software Stanford HDRM	2

Therefore, improvements to reduce the time table of the work steps and to increase the practicability of the sample preparation are necessary.
In the following, I want to discuss and mention briefly some of the future techniques, which are already partly used and partly under development. They are thought to help streamline the cumbersome library preparation process for the ultra-deep sequencing.

One automated high throughput sample preparation platform acceptable for the UDS method is the Biomek® FX^P (Beckman Coulter). After the automated nucleic acid preparation, the DNA can be purified with the SPRI technology (Agencourt) (www.beckmancoulter.com).
A further suitable alternative for the automated amplicon library preparation is the Freedom EVO automated liquid handling workstation by Tecan, which is a global provider for Liqid handling instruments. This workstation is combined to the Magnetic Seperation Module I (Chemagen) to yield a high-throughput purification of the beads assimilable to the SPRI purification technology (Agencourt). With an adequate software application and integrated robotic arms the working process on the pipetting platform from the nucleic acid extraction to the quantification of the PCR products can be performed automatically. A bar-coding system for the identification of samples is also used (Hol. 2005).

The Rem e System (Roche 454), a liquid handling accessory was developed to automate several steps of the emPCR (enrichment and primer annealing). The module is compatible with other liquid handling workstations such as the Biomek® FX^P (Beckman Coulter) (www.454.com).
The Fluidigm Access Array™ System offers possiblity for automated high throughput DNA amplicon library preparation with the integrated Fluidigm Access Array™ Integrated Fluidic Circuit (IFC). The samples can be bar-coded on a single-use micro fluidic chip, allowing up to 48 samples to be combined with 48 primers for each and hence prepared parallel using reaction volumes of 5-30nl approximately with enormous effects on costs and time savings. Various combination possibilities are still under development (Wang et al. 2009).

The Genome Sequencer FLX that replaced the first generation GS20 Sequencer is being further improved. The recent technical innovations on the GS FLX instrument are the GS FLX Titanium series reagents that can sequence a higher read length. This sequencing system that is continously under development contains also simplified preparation steps as well as an automated emPCR preparation.

9 Conclusions

In conclusion, minor drug resistance variants can increase the risk of virologic failure to the HAART therapy as shown in a growing number of recent studies. More sensitive and rapid sequencing methods may detect commonly unrecognized mutations, increase tournaround time and can therefore improve the clinical outcome. In addition, as resistance testing becomes generally accepted, a better agreement of the impact of drug resitance mutations could be achieved and universally valid algorithms established.
However, regarding UDS further development including automation of the method is necessary for future clinical application.

10 References

Arenas-Pinto A, Lactic Acidosis International Study Group. Risk factors for lactic acidosis and severe hyperlactataemia in HIV-1-infected adults exposed to antiretroviral therapy. AIDS 2007;21:2455-2464

AIDS Hilfe Wien, AIDS Statistik, Zahlen und Fakten, 2009 [updated ?;downloaded December 29, 2009]; Available from: http://www.aids.at/index.php?id=15

Baxter JD, Schapiro JM, Boucher CA, Kohlbrenner VM, Hall DB, Scherer JR, Mayers DL. Genotypic changes in human immunodeficiency virus type 1 protease associated with reduced susceptibility and virologic response to the protease inhibitor tipranavir. J Virol. 2006;80(21):10794-10801

Bennett DE, Camacho RH, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, Heneine W, Kantor R, Jordan MR, Schapiro JM, Vandamme AM, Sandstrom P, Boucher CAB, van de Vijver D, Rhee SY, Liu TF, Pillay D, Shafer RW. Drug resistance mutations for surveillance of transmitted HIV-1 drug resistance: 2009 update. PloS One. 2009;4(3):e4724

Betts BJ, Shafer RW. Algorithm Specification Interface for Human Immunodefiency Virus Type 1 Genotypic Interpretation. J Clin Microbiol. 2003;41(6):2792–2794

Bouhnik AD, Preau M, Schiltz MA, Lert F, Obadia Y, Spire B. Unprotected sex in regular partnerships among homosexual men living with HIV: a comparison between sero-noncorcondant and seroconcordant couples. AIDS 2007;21:S43-S48

Cozzi-Lepri A, Phillips AN, Clotet B, Mocroft A, Ruiz L, Kirk O, Lazzarin A, Wiercinska-Drapalo A, Karlsson A, Lundgren JD. Detection of HIV drug resistance during antiretroviral treatment and clinical progression in a large European cohort study. AIDS 2008;22:2187-2198

Dietrich DT. Disease Management Constructing Optimal NRTI-Based Combinations: Past, Present, and Future. Journal of International AIDS society 2006;8:16

Grant RM, **Kuritzkes DR**, **Johnson VA**, **Mellors JW**, **Sullivan JL**, **Swanstrom R**, **D'Aquila RT**, **Van Gorder M**, **Holodniy M**, **Lloyd Jr RM Jr**, **Reid C**, **Morgan GF**, **Winslow DL**. Accuracy of the TRUGENE HIV-1 Genotyping Kit. J Clin Microbiol. 2003;41:1586-1593

GuideLines Rules 14.0 for the use of the TRUGENE HIV-1 Module 14.0. OpenGene-Software 4.1. www.siemens.com/diagnostics. 2009 © 2009 Siemens Healthcare Diagnostics Inc. PN: 300-0365-00 Rev. 01 2009-02

Hahn H, Falke D, Kaufmann SHE, Ullmann D. Medizinische Mikrobiologie und Infektiologie. 5th ed. Heidelberg. Springer Verlag. 2004;577-590

Halvas EK, Wiegand A, Boltz VF, Kearney M, Nissley D, Wantman M, Hammer SM, Palmer S, Vaida F, Coffin JM, Mellors JW. Low Frequency Nonnucleoside Reverse-Transcriptase Inhibitor–Resistant Variants Contribute to Failure of Efavirenz-Containing Regimens in Treatment-Experienced Patients. J Infect Dis. 2010;201(5):672-680

Hof H, Dörries R. Medizinische Mikrobiologie. 4th ed. Stuttgart. Georg Thieme Verlag. 2009;233-240

Hoffmann C, Rockstroh JK, Kamps BS. HIV.Net 2008. Wuppertal-Beyenburg. Steinhäuser Verlag. 2008

Hoffmann C, Minkah N, Leipzig J, Wang G, Arens MQ, Tebas P, Bushman FD. DNA bar coding and pyrosequencing to identify rare HIV drug resistance mutations. Nucleic Acids Res. 2007;35(13);e91

Hol F. From Primary Blood to DNA-Fully Automated DNA Extraction. Application note Stragene. BIOforum Europe 6/2005

Kijak GH, Rubio AE, Pampuro SE, Zala C, Cahn P, Galli R, Montaner JS, Salomón H. Discrepant results in the interpretation of HIV-1 drug-resistance genotypic data among widely used algorithms. HIV Medicine 2003;4(1):72-78

Kuritzkes DR, Grant RM, Feorino P, Griswold M, Hoover M, Young R, Day S, Lloyd Jr RM Jr, Reid C, Morgan GF, Winslow DL. Performance characteristics of the TRUGENE HIV-1 Genotyping Kit and the Opengene DNA Sequencing System. J Clin Microbiol. 2003;1594-1599

Liu TF, Shafer RW. Web Resources for HIV Type 1 Genotypic-Resistance Test Interpretation. Clin Infect Dis. 2006;42(11):1608-1618

Mims C, Dockrell HM, Goering RV, Roitt I, Wakelin D, Zuckerman M. Medizinische Mikrobiologie Infektiologie. 2nd ed. München. Urban & Fischer Verlag. 2006;283-293

Modrow S, Falke D. Molekulare Virologie. Ulm. Spektrum Akademischer Verlag GmbH. 1997;302-303

Joint United Nations Programme on HIV/AIDS (UNAIDS) [homepage], Geneva: World Health Organisation; Status of the global HIV epidemic. *2008 Report on the global AIDS epidemic*. [daily updated; downloaded November 28, 2009]. Available from: http://www.unaids.org/en/

Johnson JA, Li JF, Wei X, Lipscomb J, Irlbeck D, Craig C, Smith A, Bennett DE, Monsour M, Sandstrom P, Lanier ER, Heneine W. Minority HIV-1 Drug Resistance Mutations Are Present in Antiretroviral Treatment–Naïve Populations and Associate with Reduced Treatment Efficacy. PLoS Med 2008;5(7):e158

Johnson VA, Brun-Vézinet F, Clotet B, Conway B, Kuritzkes DR, Pillay D, Schapiro J, Telenti A, Richman D. Update of the Drug Resistance Mutations in HIV-1: 2005. International AIDS Society–USA: Topics in HIV Medicine. 2005;13(1)

Le T, Chiarella J, Simen BB, Hanczaruk B, Egholm M, Landry ML, Dieckhaus K, Rosen MI, Kozal MJ. Low-Abundance HIV Drug-Resistant Viral Variants in Treatment-Experienced Persons Correlate with Historical Antiretroviral Use. PLoS One 2009;4(6):e6079

Lüllmann H, Mohr K, Lutz H. Pharmakologie und Toxikologie. 16thed. Stuttgart. Georg Thieme Verlag. 2006;481-491

Marguilies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature 2005;437(7057):376-380

Palmer S, Boltz V, Maldarelli F, Kearney M, Halvas EK, Rock D, Falloon J, Davey RT Jr, Dewar RL, Metcalf JA, Mellors JW, Coffin JM. Selection and persistance of non-nucleoside reverse transcriptase inhibitor-resistant HIV-1 in patients starting and stopping non-nucleoside therapy. AIDS 2006;20(5):701-710

Paredes R, Lalama CM, Ribaudo HJ, Schackman BR, Shikuma C, Giguel F, Meyer WA3rd, Johnson VA, Fiscus SA, D'Aquila RT, Gulick RM, Kuritzekes DR. AIDS Clinical Trials Group (ACTG) A5095 Team. Preexisiting minority drug-resistant HIV-1 variants, adherence, and risk of an antiretroviral treatment failure. J Infect Dis. 2010;201:662-671

Pilcher CD, Joaki C, Hoffman IF, Francis Martinson EA, Mapanje C, Stewart PW, Powers AK, Galvin S, Fiscus SA, Cohen MS. Amplified transmission of HIV-1: Comparison of HIV-1 concentrations in semen and blood during acute and chronic infection. AIDS 2007;21:1723-1730

Quinn TC. HIV epidemiology and the effects of antiviral therapy on long term consequences. AIDS 2008;22:S7-S12

Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW. Human immunodeficiency virus reverse transcriptase and protease sequence database. Nucleic Acids Res. 2003. 31(1):298-303

Reid C, Bassett R, Day S, Larder B, De Gruttola V, Winslow D. A dynamic rules-based interpretation system derived by an expert panel is predictive of virological failure. Antivir Ther. 2002;7:S91

Shafer RW. Rationale and Uses of a Public HIV Drug-Resistance Database. J Infect Dis. 2006;194:S51-58

Simen B, Simons JF, Huppler Hullsiek K, Novak RM, MacArthur RD, Baxter JD, Huang C, Lubeski C, Turenchalk GS, Braverman MS, Desany B, Rothberg JM, Egholm M, Kozal MJ. Low-Abundance Drug-Resistant Viral Variants in Chronically HIV-Infected, Antiretroviral Treatment–Naive Patients Significantly Impact Treatment Outcomes. J Infect Dis. 2009;199:693-701

Wang C, Mitsuya Y, Gharizadeh B, Ronaghi M, Shafer RW. Characterization of mutation spectra with ultra-deep pyrosequencing: Application to HIV-1 drug resistance. Genome Research. 2007;17:1195-1201

Wang J, Lin M, Crenshaw A, Hutchinson A, Hicks B, Yeager M, Berndt S, Huang WY, Hayes BB, Chanock SJ, Jones RC, Ramakrishnan R. High-troughput single nucleotide polymorphism genotyping using nanofluidic Dynamic Arrays. BMC Genomics. 2009;10:561

Weiser B, Philpott S, Klimkait T, Burger H, Kitchen C, Bürgisser P, Gorgievski M, Perrin L, Piffaretti J-C, Ledergerber B, Swiss HIV Cohort Study. HIV-1 coreceptor usage and CXCR4-specific viral load predict clinical disease progression during combination antiretroviral therapy. AIDS 2008;22:469-479