Master thesis in Pure and Applied Biochemistry

Development of a novel quantitative PCR analysis method for HIV-1

Anna-Malin Hallberg

January 2021

Division of Pure and Applied Biochemistry, in cooperation with Octapharma AB

University supervisor: Lei Ye, Professor

Company supervisor: Rasmus Gustafsson, PhD

Examiner: Per-Olof Larsson, Professor emeritus

Abstract

Biological pharmaceuticals are the answer to many severe diseases of today's society. Some of these are protein-based drugs and can be produced from human blood plasma. In order to guarantee the safety of patients, absence of any pathogens needs to be ensured. Such a pathogen is HIV-1, which holds high mutation rate and a large variation in its genome.

The aim of this project was to develop novel primers and probes for detection of a wide range of subtypes of HIV-1, using the TaqMan qPCR detection system. Three targets within conserved parts of the genome were selected in the regions of *LTR5´*, *Pol* and *LTR3´*. The primers and probes were optimized regarding concentrations, salt content and compatibility.

The sensitivity of detection showed promising results with a value as low as 8.18 IU. The selectivity did not result as preferred, with the best combination of primers and probe possible to detect 7 out of 8 subtypes of the most common genotype, M. If combining the primers and probes of all targets suggested here, detection of all tested subtypes was possible.

Summarizing the results, the primers and probe targeting the *Pol* region shows promising data. Optimization regarding the sequences of the *Pol* primers and probe and additional evaluation of compatibility between all targets need to be studied, in order to see, if the method can meet the standards and further be implemented into the experimental routine for HIV-1 detection.

Popular science summary – The potential of better detection of HIV

Viruses have become a global health issue and people are constantly dying from their effects, both in rich and poor countries. This has awakened a new dimension in the urgency and interest of finding new methods of identifying infected people in order to help, treat and stop the spread of infection in the society. In order to make the world a better place for all with a better life quality.

More than 700 000 people died last year of the effects of HIV, and millions are living with the disease causing a poor life quality, worldwide. In order to be able to stop the spread of infection more accurate methods of detecting the virus are necessary. Containing the spread of the illness is paramount, and by making the contagious individuals aware of their infection, the spread will be easier to control as well as the infected individual can get treatment faster. However, many viruses of this kind cannot be cured, proper inhibitor medicines are available. Many of these treatments today have shown promising results with best effect upon usage in early stages of infection. Viruses mutate frequently, in order to avoid the immune system. This makes them a hard target, since they are changing their genetic code constantly. The numbers of infected cells can differ between individuals and in the stage of infection, which puts high standards on detection methods. Two aspects of virus detection are especially important; the possibility of detecting the virus with high specificity, able to distinguish the HIV infected cells from others, and with high sensitivity, capable of finding every infected cell even in low numbers is of great importance.

The work done during the project has shown interesting results of HIV detection. A method using genetic nucleotide sequences has been developed. The developed genetic sequences verifying the starting point of replication in DNA, called primers. Additionally, with a complementary nucleotide sequence to the DNA template, dyed with signal giving elements attached, called probes. The results suggest a detection method with high sensitivity, able to find viruses in very low concentrations within blood. In addition to, good ability of specificity, with as many as 7 out of 8 types of the most common kind of HIV, with the use of 4 replicates. In total the most promising method developed managed to find HIV in 19 out of 24 samples evaluated with prominent signals. These results indicate a highly promising method. At this stage it needs further development in order to meet the standards of European Medical Agency and Food and Drug Administration, as well as being even more precise in specificity.

Substantial theoretical knowledge of HIV and insights in possibilities of detection of the virus has been the foundation to success in the project. Different parameters such as suitable components, time, temperature, target sequences in the genetical code of the virus has been evaluated and optimized in order to develop a most positive method. Both with high specificity and low sensitivity. The aim of using the method in routine analysis is not completely reached, but the method is on the right track to fulfil the demands of it.

The research and development of new detection methods for viruses such as HIV is of great importance to the society. Infected individuals need to know about their disease and being able to get proper treatment as well as not pass the virus further on to their fellow humans. The methods can also be used to certify that donated blood is absent from pathogens. If so, the blood can be donated to those in need or used to produce biological and protein-based pharmaceuticals for severe diseases.

CONTENTS

PREFACE

This report is a result of the master thesis project that has been conducted at the Division of Pure and Applied Biochemistry, Faculty of Engineering, Lund University, in collaboration with Octapharma AB in Stockholm.

An extra thank you to supervisor Lei Ye at Division of Pure and Applied Biochemistry for his support and helping hand during the whole project, for being involved and guiding me in the right direction at difficult decisions. I would also like to thank you Rasmus Gustafsson, supervisor at Octapharma for being a great mentor and helping me with all kind of questions and situations, which has led to deeper understanding for the project. Thank you, Ernst Jonsson, for helping with the laboratory work and for always being by my side and bringing me knowledge in the practical work at R&D department at Octapharma. Finally, I would like to thank you Per-Olof Larsson for acting as my examiner.

Thank you all!

1 INTRODUCTION

1.1 BACKGROUND

Biological pharmaceuticals for treating rare diseases are often within the field of protein-based drugs, which can be produce from human blood plasma (Manning et al., 2010). For being able to use human blood as the foundation of pharmaceuticals, one needs to be certain of the quality of the blood as well as guarantee the absence of any pathogens. As a drug producer you always need to think of the safety of the patients, and by that assure that the pharmaceuticals are safe for the patient to use. Good Manufacturing Practice (GMP) is the regulation which pharmaceutical companies must comply with (Haleem et al., 2015). A critical part to consider is the fact that proteins cannot be pasteurized or sterilized since they cannot tolerate heat in the temperature range needed to certify that all pathogens are dead (Ren et al., 2018). Therefore, representative samples can be collected and tested as an alternative way to guarantee patient safeties.

The first critical step in a production line is to ensure the absence of virus in the incoming blood. This can be done by Polymerase Chain Reaction (PCR)- analysis conducted in a laboratory. The PCR analyses are screening for viruses common in blood such as Human Immunodeficiency Virus (HIV) among others (Octapharma Nordic AB, 2020).

1.2 PROJECT TASK AND PURPOSE

This master thesis project has aimed at developing a novel detection method for the Human Immunodeficiency virus type 1 (HIV-1). The method will be used in the standard routine at Octapharma when testing for different blood viruses, in the incoming blood for pharmaceutical production. The project was limited to the use of PCR methods, since equipment and tools regarding these methods has already been approved for use at Octapharma. Additionally, this includes the fact that a conducted master mix, with some already established elements is necessary for each sample investigated. This limits the possibilities of contents as well as combinations of elements in the master mix. The method developed is a Real-Time quantitative PCR (RT qPCR) analysis method which uses specific primers and TaqMan probes in order to detect the virus of interest. The RT qPCR developed has been using a standard predefined program conducted for RNA viruses (Octapharma Nordic AB, 2020). The main goal of the project has been to generate innovative primer and probe sets that will be able to detect all subtypes within the family of HIV-1 with a high sensitivity as well as high specificity.

1.3 THE SCOPE OF THE PROJECT

The development of functional primer and probe sets that are novel in their kind is the goal of the project. Priority is to achieve high degree of sensitivity, aiming at a better sensitivity than the currently used method, as well as high specificity, by that the sets can detect all subtypes of

HIV-1. In order to achieve this, optimization of previously published primer and probe sets will be conducted and evaluated with tools of bioinformatics, as well as experimental trials in the laboratory. This project can be summarized by the phases below:

- Literature study of previously published primer and probe sets and investigation of important parameters
- Plan and conduct relevant experiments for the most promising sets
- Analysis and interpretation of the experimental data
- Apply the findings to develop new method as well as optimize

Specific aims are to seek answers to these questions:

- What parameters are critical for efficient detection?
- Which parts of the HIV genome are highly conserved?
- Is it possible to detect all subtypes with only one method at the same time?

2 THEORETICAL BACKGROUND

2.1 HUMAN IMMUNODEFICIENCY VIRUS

2.1.1 A GLOBAL HEALTH PROBLEM - THE HISTORY OF THE VIRUS

The first well documented case of HIV is recorded back in 1959, when a man in the Democratic Republic of Congo was diagnosed with the disease (Worobey et al., 2008). Although some early cases were recorded, it did not become a society health problem until the late 1970s. In the early years of the 1980s cases of different rare types of cancer, pneumonia among other unspecific symptoms, which cannot be seen in people with normal immune systems were reported among mostly homosexual men, in different parts of the United States (Worobey et al., 2016). Due to this, clinicians and scientists globally used the term Acquired Immunodeficiency Syndrome (AIDS) in order to describe the conditions of the previously healthy patients (The AIDS Institute. n.d).

It is hypothesized that the sickness originated during the hunt of chimpanzees in West Africa, which existed until the first decades of the 20th century, as the most likely source of the disease (Gao et al., 1999). It is believed that HIV is a mutated form of a virus called Simian Immunodeficiency Virus (SIV) which chimpanzees and other types of monkeys can suffer from. SIV was probably transmitted to humans by infected blood from the monkeys during the slaughter (Bailes et al., 2003).

2.1.2 A GLOBAL HEALTH PROBLEM - THE DISEASE TODAY

HIV is still today a major global health problem, with as many as 33 million victims on its conscience. Today HIV is spreading and is infectious like no other known agent (Albert J.,

2018). It can be found on every continent of the globe as well as it is causing premature deaths. In 2019 as many as around 700 000 people died from HIV- related illnesses worldwide (UNAIDS, 2020). There is no effective cure but with great prevention treatment one can live a manageable life even with the virus and stop the process of developing the chronical sickness, AIDS (Simon et al., 2006).

At the end of 2019 as many as 38 million people were estimated to live with HIV, the main affected areas are located on the African continent (WHO, 2020). A main key factor in hindering the spread of HIV, is by distribute knowledge regarding the disease as well as making continuously testing and necessary health care highly assessable for people in the most vulnerable areas (Yahaya et al., 2010). This is an ongoing work for many HIV related help organizations worldwide, especially on the African continent. The organization UNAIDS are fighting for their ambition of the so called 90-90-90 vision. The vision includes, getting as much as 90% of infected people know about their infection and their health status. Making up to 90% of the HIV-1 positive population receive sustained antiretroviral therapy and by them, 90 % will achieve a viral suppression (UNAIDS, 2020).

Since many of the countries that hold the most active cases of HIV often have a poor infrastructure as well as health care system, it is even more challenging to get the infected people the medical care and help they need. Often roads, refrigerators, clean water and other trivial resources for medical transport and treatment are missing (Petti et al., 2006). Because of this, in combination with the fact that traveling has become more and more easily accessible, the spread and death caused by HIV and AIDS has become over the years an ongoing global pandemic. However, the African continent still holds the most spread, but it is up to all countries around the globe to help and find answers to how to solve the HIV pandemic. Because all of us can be infected (WHO, 2020).

2.1.3 A HIGHLY VIRULENT VIRUS - THE SUBGROUPS AND SUBTYPES

Based on genetic characterizations of the virus, HIV is categorized into different types called type 1 and 2. HIV-1 is more common worldwide, it accounts for 95% of people living with HIV. HIV- 2 has to a significant part showed signs of being less infective, as suggested being the reason why HIV-2 is not as frequently common as HIV-1 (O´Brien T., 1994). HIV-1 has been divided into the different subgroups M, N, O and P. Subgroups of N, O and P are very rare, and 95% of the known cases are within the Subgroup M (Osmanov et al., 2002). Phylogenetic analysis shows that the SIV found within chimpanzees belongs to group N and O (Vanden Haesevelde et al., 1996). The subgroup M has been further divided, and there are at least nine known subtypes called A, B, C, D, F, G, H, J and K. The recombinant HIV-1 are called Circulating Recombinant Form (CRFs) and originate from various subtypes. CRFs arise from the fact that the subtypes can exchange and combine their genetic material and form a hybrid (Robertson et al., 2000). Around 20% of the viruses within subgroup M are CRFs, and one common type is the AE subtype, which is derived from subtype E. It only holds a difference in its envelope protein (*env*) otherwise it is the same as subtype A. Due to this it has been re-named to AE (Seitz, 2016). Until today more than 70 different CRFs have been described and

additional CRFs are excepted within the future. It has been reported that recombination has occurred between subgroups as well, however it appears to be rare. Based on analysis, it is hypothesized that the most conserved and thereby the oldest, evolutionarily speaking, are subtypes A to D within the M subgroup (Seitz, 2016). The different subgroups as well as subtypes can be seen in Figure 2. Studies have shown that subtype B is the most common in United States as well as in Western Europe. Despite this subtype B is only responsible for 12% of all infected HIV cases worldwide. Less data regarding subtype C is available, regardless of this it is speculated that more than 50% of people living with HIV are infected by subtype C (Avert, 2019).

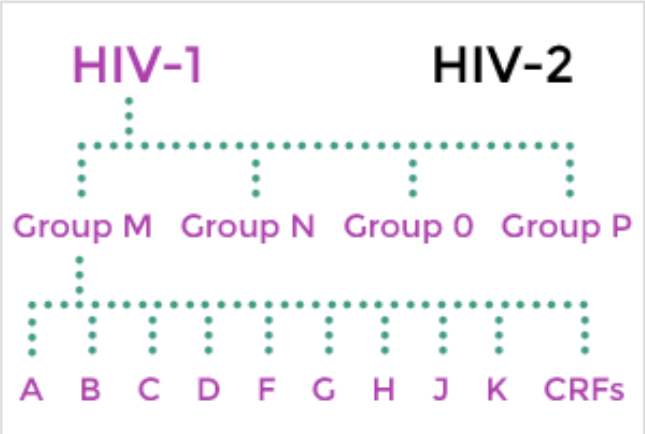

Figure 2. The figure describes the different subgroups as well as subtypes and how they are related, within the HIV-1 (Avert, 2019).

2.1.4 A HIGHLY VIRULENT VIRUS- GEOGRAPHICAL DIFFERENCES

HIV holds different subtypes which has shown diversity and distribution changes by land and region. The geographical variety of the virus may be caused due to the genetic diversity the pathogen holds because of high mutation as well as replication rate (Cuevas et al., 2015). In Europe and Northern America subtype B has shown to be the most common. For Russia, former Soviet Union and East Africa, subtype A is the most prevalent one (Taylor B. S et al, 2008). All known subtypes today have shown to be more common in some geographical places than others, but it is still unknown why it is so. There is no evidence that some subtypes should be more infectious than others, nor that genetic changes among HIV should be preferred to the geographical differences between humans and their genomes (Pebody R., 2009). The different subtypes which are most common in specific geographical regions can be seen in Figure 1.

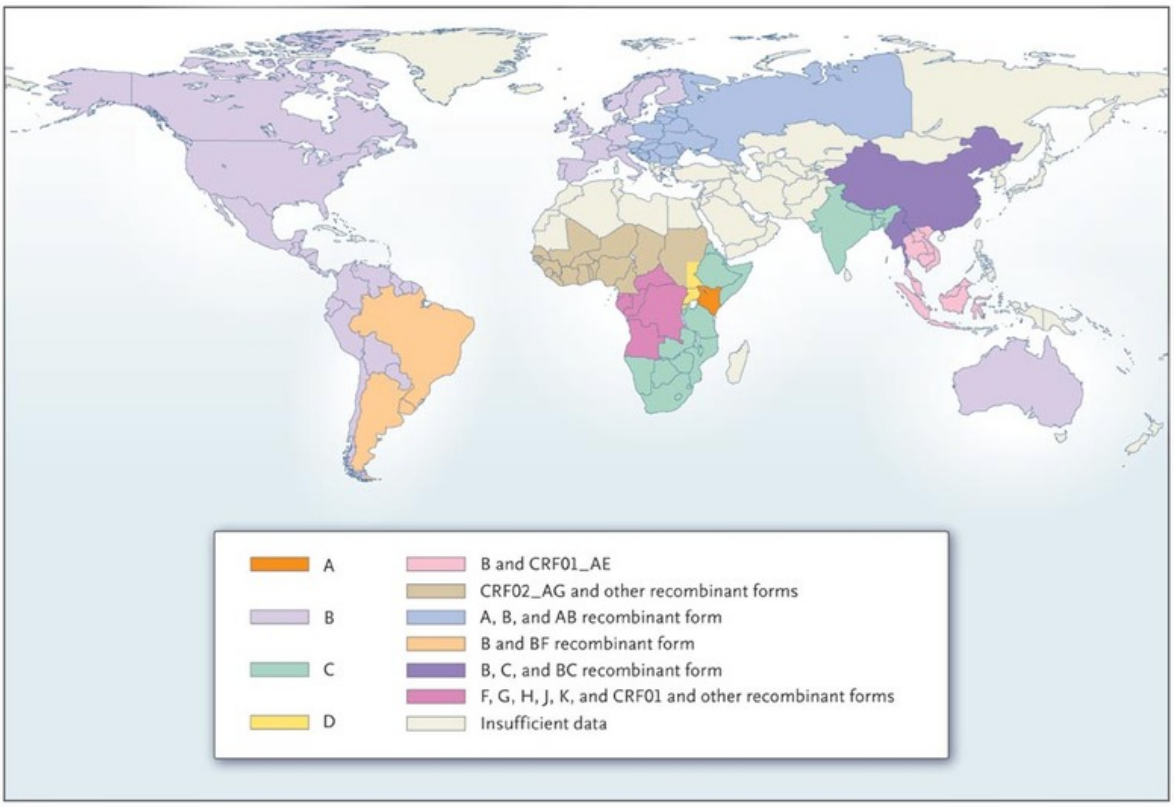

Figure 1. An overview of the most common subtypes of HIV type 1, subgroup M in a specific geographical area on the globe (Taylor B. S et al, 2008).

2.1.5 A HIGHLY VIRULENT VIRUS – THE CHARACTERIZATIONS

HIV is a lentivirus and belongs to the family of *Retroviridae*, which often is characterized by their long incubation time. It is known that lentiviruses infect a broad range of mammalian species and has been identified worldwide (Haase A., 1986). The lentivirus holds the properties of converting its RNA into cDNA by the reverse transcriptase it contains. By that it can integrate a substantial part of its viral cDNA into the infected cell´s DNA. It has been verified that the virus upon infection can attacks both dividing and non-dividing cells, making it a very efficient gene deliverer as well as letting it infect cardiomyocytes and stem cells (Li. R and Wiesel. R, 2014). For HIV-1 to enter a cell, possible host cell needs to have a CD4 receptor. The cells that holds the properties of CD4 receptors are often T helper cells as well as monocytes, both plays an important role in the immune system (Toes et al., 1999).

2.1.6 A HIGHLY VIRULENT VIRUS - THE GENOME

The HIV genome is about 9 200 nucleotides in size and consists of two identical single strands of RNA which carries the genetic information (Li et al., 1992). The RNA is safely kept within the envelope which builds up the round shaped virus. The envelope is the most crucial element of the virus, since it has glycoproteins which takes part in the attaching as well as the invading process to a host cell. Within the envelope is the matrix, its main task is to hold the glycoproteins to the virus. Further, inside the matrix itself lays the viral core which is the home for the RNA and some enzymes, e.g. reverse transcriptase, that are essential for the virus to replicate when entered the host cell (Payseur, n.d). Inside the cell itself lays different vital enzymes for the replication as well as structure for the virus. The virus structure can be seen in Figure 3.

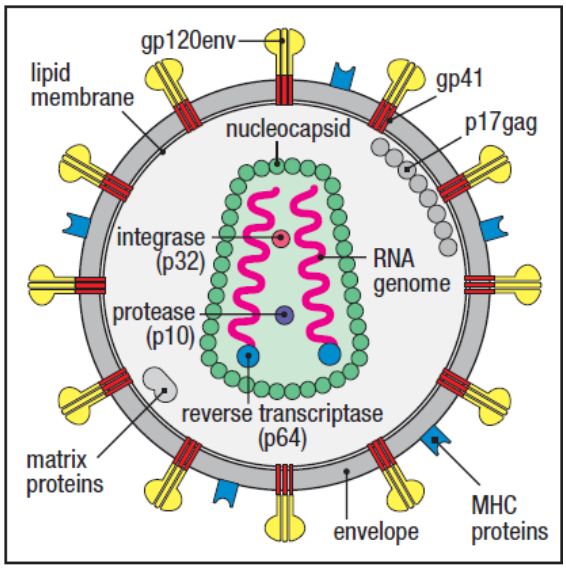

Figure 3. Illustration of the virus HIV-1. The virus particles are round in shape and measures a diameter of 120nm. The virus particle is approximately 60 times smaller than a T helper cell. The three essential viral proteins for the virus: reverse transcriptase, integrase and protease are all shown. On the envelop surface of the particle are the glycoproteins gp41 and gp120 present, both play an important role upon attachment of a CD4-receptor on the host cell (Murphy K. and Weaver C., 2017, p. 575).

The cDNA of HIV is edged on both sides of Long Terminal Repeat (*LTR*) sequences. The role of *LTR* is still unknown (Richetta et al., 2019). At the 5´region of the genome starts the transcription of the viral genetic code, upon binding of a primer. The reading frames direction is 5´to 3´and start with the gene gag after the *LTR* sequence. This part mainly codes for the proteins of the outer core membrane as well as the capsid protein (Briggs et al., 2004). The *gag* gene is followed by the *Pol* gene which is coding for the reverse transcriptase, RNase and the enzymes protease (Weber et al., 2003). Next to the *Pol* gene lays the smaller gene *env*, which encodes for the highly important envelope proteins (Lopez-Vergès et al., 2006). Apart from these essential structural proteins, the genome of HIV encodes for various regulatory proteins, such as: *vif* (viral infectivity factor), *vpr* (virus protein r), *vpu* (virus protein unique), *tat* (trans activator protein), *rev* (RNA splicing-regulator) and *nef* (negative regulating factor). The whole genome and the different genes as well as how the genes are disposed, can be seen in Figure 4.

Rev and *tat* are believed to play an important role for the initiation of HIV replication whereas the other genes have a necessary effect on pathogenesis, viral replication as well as budding. The genome structure of SIV, which is the believed ancestor of HIV, is undistinguished from HIV type 1 (Seitz, 2016).

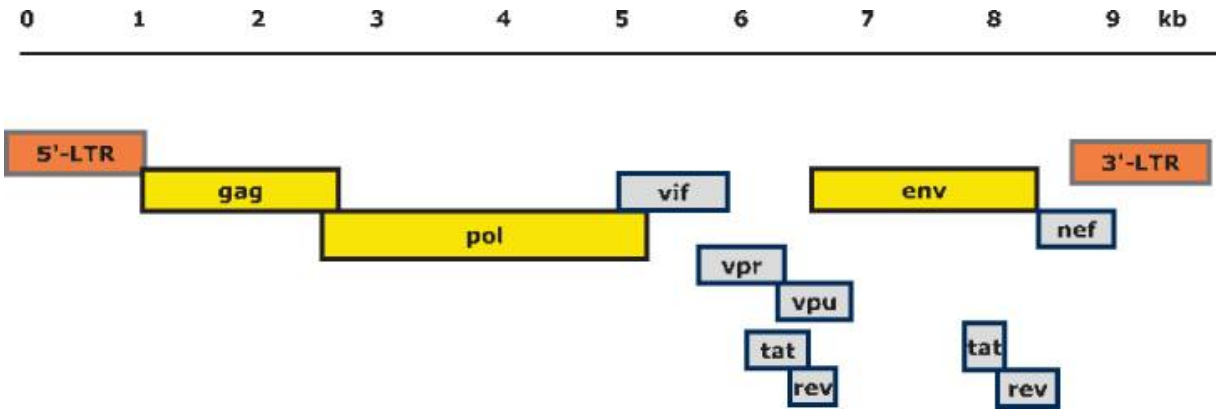

Figure 4. Organization and structure of the HIV type 1 genome. All the reading frames of the different gens which lays within the genome are shown. The ones coding for structural and regulatory proteins are: *LTR*= Long Terminal Repeat, *gag*=group-specific antigen, *pol*= polymerase and *env*= envelope. The other genes, coding for regulatory proteins: *vif, vpr, vpu*, *tat, rev* and *nef*. It is known that tat and rev are made up out of two regions in the genome. The whole genome is composed of around 9 200 nucleotides (Seitz, 2016. P.204).

2.1.7 A HIGHLY VIRULENT VIRUS - THE INFECTION

For infection, HIV particle uses its surface glycoprotein gp120 in the envelope to bind the CD4 receptor of the host cell (Vivès et al., 2005). The CD 4+ cells are T helper cells, dendritic cells, macrophages and astrocytes which play an important role in the immune system. All these cells are highly vulnerable for HIV. Upon attachment of the HIV to the CD 4 receptor, conformational change takes place of the CD 4 receptor. Due to this, an additional site to gp120 occurs, which enables binding to a co-receptor e.g. chemokine receptor 5. The binding of the combination of gp120 and the co-receptor to the CD4 -receptor activates another, further conformational change in the site of binding (Kwong et al., 1998). Due to the attachments to gp41 and another important glycoprotein in the HIV cells envelope. The change of the gp41 makes its N-terminus present in the viral membrane to forms a channel because of its high hydrophobicity. Through the channel the virus can fuse into the targets cell´s plasma membrane, and the host cell is infected (Seitz, 2016).

The infection of HIV is divided into different stages, which indicates the severity of the disease and its spread in the body. The first stage is often seen as the primary infection where the person may get asymptomatic or acute retroviral syndrome, as fever rashes and muscle pain among others. It often takes from between one to approximately four weeks after infection before one is getting the symptoms of the first stage (Jolles et al., 1996). At this stage, the body is responding to the infection of HIV and the immune system is doing everything it can to fight it. As a response it produces antibodies towards the HIV which now are circulating through the

body. This triggering response of the immune system is often called seroconversion which the infected person may experience months after the infection occurred .The amount of virus particles within the body at this stage is often tremendously high, by that one is very infectious and transmitting the virus to someone else is likely (Cardo et al., 1997). The second stage is often asymptomatic and can take up to ten or even fifteen years, all depending on health status and genetics if no treatment is given. While maybe not having any symptoms, the virus is still active and replicating as well as infecting new cells of the host body. Indications of different respiratory tract infections, skin infections such as herpes etc. and smaller unexplained weight loss among others, are often signs of HIV, even if they occur slowly and sporadic over years (Hollingsworth et al., 2008). The third stage of infection often gets heavily symptomatic for the person in question. Now is the damage of the immune system severe. It is often very likely to get ill symptoms of otherwise normal and often easy-going infections, such as a normal cold. Usual symptoms are often distinct weight loss, sustained cough, severe unexplained diarrhea etc. If the disease develops further, the symptoms gets severe and often needs medical treatment in hospital. It is common at this stage that other infections become dreadful for the patient if untreated. At this stage, in clinical medical terms, it is often referred to as AIDS. Development and outcome of HIV research has demonstrated that the infected person needs to get medical treatment in an early stage in order to achieve the best effects of possible treatments (WHO, 2005).

2.1.8 A HIGHLY VIRULENT VIRUS - THE REPLICATION CYCLE

Upon infection of a host cell, the viral core enters the cytoplasm. Here the genetic material in the form of cDNA is released. The cDNA also known as proviral DNA is generated by the HIV reverse transcriptase from the RNA of the virus. The virulent cDNA travels to the nucleus of the host cell accompanied by the viral integrase and the Vpr protein. The cDNA is further integrated into the human DNA. By the infection, activation of the CD4 T cells occurs, which leads to an activation response of transcription factors (Freed E., 2001). Due to the response, host transcriptional factors and viral proteins bind to the proviral *LTR* sequences and the transcription of the viral HIV genome can further be initiated. The first product of the viral transcription is enormously processed, since many spliced mRNAs are created including the coding sequences of all the regulatory proteins. Among these, the gene encoding *tat* is processed, which enhance transcription and stabilizes the mRNA´s upon binding to its transcripts. Rev on the other hand transport the RNA transcripts to the cytosol of the host cell. As the quantity of Revs increases, less un-spliced as well as less significant viral transcripts are transported to the cytosol. The un-spliced transcripts coding for structural proteins are kept within the nucleus. They are used as the new viral genome as well as some of the transcripts are being packed to form new virus particles which are further released to infect new cells (Murphy K. and Weaver C., 2017). The infection as well as the replication cycle are illustrated in Figure 5.

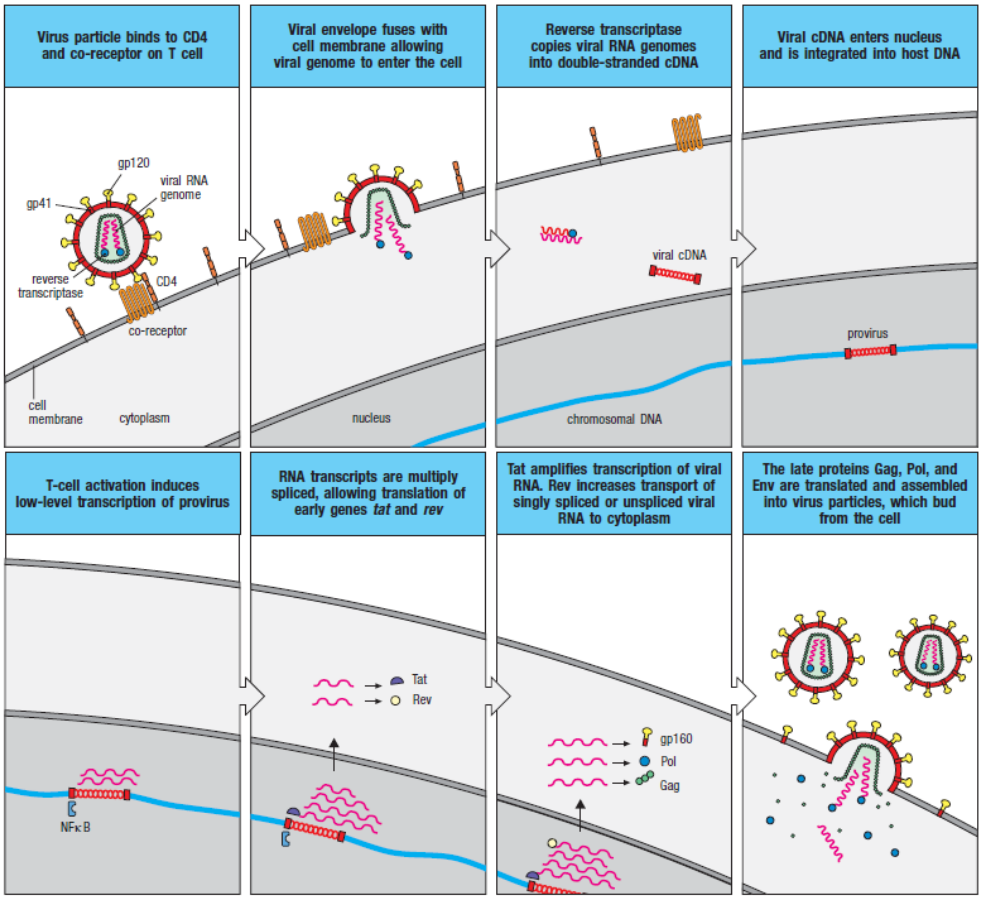

Figure 5. The replication cycle of HIV. In the upper row the cycle stars with attachment to CD4-recptor and co-receptor to the virus particle. Then the genetic material within the virus is released into the cytosol of the host cell. Further is the cDNA conducted in the cell by the reverse transcriptase travels to the nucleus of the host cell. The lower row describes that the fusion of the cDNA into the host cell´s DNA gives a respond for CD4 T cells to be activated and by that an immune response is triggered. The immune response activates the transcription at the *LTR* region, and it is initiated. The sequences of tat induce replication to progress and the transcripts are stabilized. Upon the activation of the sequence rev, the unspliced and structural protein coding mRNA are kept in the nucleus in order to be folded into new virus particles which then are released outside the cell, with a purpose to infect new cells (Murphy K. and Weaver C., 2017, p.577).

2.2 Quantitative Polymerase Chain reaction

2.2.1 Real-Time quantitative Polymerase Chain Reaction

The laboratory technique of Real-Time quantitative Polymerase Chain Reaction (RT qPCR) has its foundation in the method of Real-time Polymerase Chain Reaction (RT PCR). The technology of RT PCR amplifies a specific target in a DNA sequence. Extension of primers, also known as short single-stranded oligonucleotides are possible due to the function of the Taq polymerase (Heid et al., 1996). Cycles of heat denaturation are repeated under which primers are annealed and extended. In order to specify the target sequence of the DNA, the primers are designed to fit that specific region. The target DNA sequence is used as the template when the Taq polymerase performs the extension of the primer. The reaction synthesizes new DNA and by that more templates after each cycle are generated at an exponential rate. The amount of target DNA will, by that be doubled after each cycle, until one of the elements becomes limited. Then the reaction will stand-still and reaches a plateau. The quantitative form of PCR will be

enabled by a specific or non-specific chemistry that quantifies the amplified target. The detection at a specific level of the reaction is directly proportional to the quantity of the target within the sample. The signal given at detection is most often generated using fluorescent techniques. qPCR is often used for detection or quantification of pathogens, analysis of gene expression and microRNA quantification (Vandesompele J., n.d).

Salt ions are important in order to make the PCR run smooth and more reliable. KCl in the PCR buffer reduces the repulsion of the negatively charged DNA, by that stabilizing the binding of primer and template. Its neutralizing property is significant to assure a proper annealing of the primers, which is a crucial step in the PCR since otherwise the polymerase cannot bind and start the amplification (Datta and LiCata, 2002). An increase in K^+ concentration has shown to result in longer DNA templates, which denaturate at a slower rate than shorter DNA strands. Due to this, it has been suggested that longer products are amplified with higher efficiency at lower concentrations of KCl (Cheng et al., 1995). Mg^{2+}, often from the salt $MgCl_2$, is a crucial cofactor for the assay. It acts by enabling integration of dNTPs in the polymerization. The cations catalyze the phosphodiester bond formation, in the polymerase´s active site, connecting the 3´end of a primer and the phosphate group of the dNTP. Mg^{2+} holds similarity to K^+ in the manner of its possibilities of stabilizing the negatively charged DNA templates, and by that facilitates the annealing of primers and DNA (Steitz TA., 1998).

2.2.2. Real-Time TaqMan qPCR

An easy and straight forward modification is the use of TaqMan probes, trademarked by Roche, with the technology of RT qPCR. The TaqMan probe is marked with an attached dye label of fluorescent character at the 5´end and with a quencher at the 3´end. The quencher acts as a blocker for the fluorescent signal, when the two are attached to the probe due to the short distance between the two. On both sides of the probe, the method requires two independent and unlabeled primers, one on each end of the target sequence (Taylor et al., 2010). During the reaction of RT qPCR, the denaturation of the double-stranded DNA occurs due to an increase in temperature. Further is the temperature decreased in order to allow annealing of primers and probe to the template. At this stage, the Taq polymerase binds to the primers and starts to synthesis a new template. During the synthesis, the Taq polymerase eventually reaches the TaqMan probe and cleaves it, due to its nuclease activity of the 5´end. The cleavage separates the quencher and the fluorescent dye, making the blocking properties of the quencher ineffective towards the fluorescent signal (Butler J., 2012). By that the signal of the fluorescent dye can be detected. At each cycle run, the amplification will be exponential, resulting in an increase in the signal of detection for the desired DNA sequence.

The signal intensity is directly correlated to the amount of target templates (ThermoFisher (A), n.d). The RT qPCR with TaqMan probe is described in Figure 6.

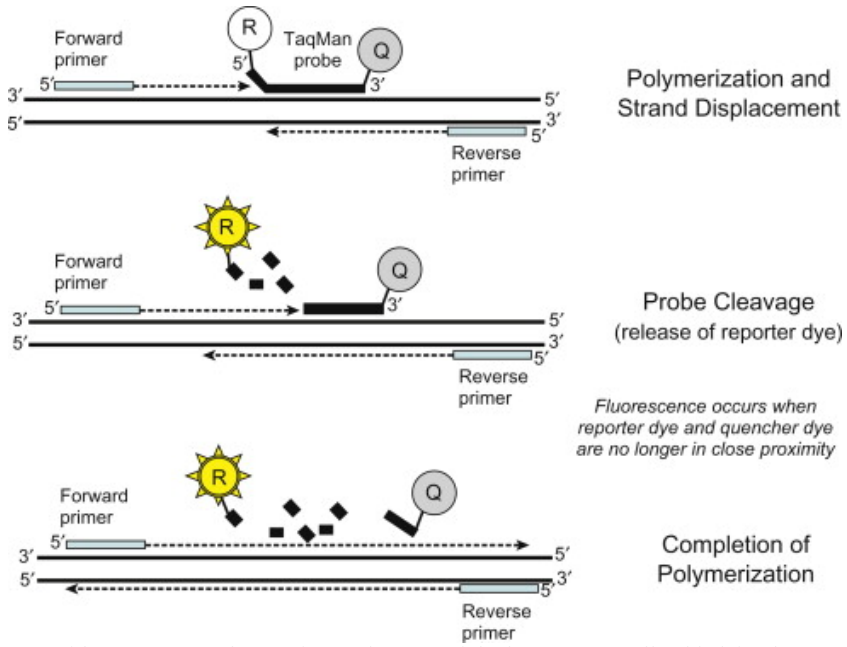

Figure 6. Real-Time qPCR with TaqMan probe. Before: The essential elements are all added in the reaction mixture. Denaturation of double-stranded DNA into single-stranded DNA occurs due to an increase in temperature. Polymerization and strand displacement: Temperature is lowered and allows the primers and probe to bind to the DNA template. DNA synthesis takes place and Taq polymerase start generating new DNA templates by adding nucleotides. Probe cleavage: When the polymerase reaches the probe it is cleaved, letting the reporter dye (fluorophore) and quencher to be cleaved off, due to its nuclease activity at the 5´end. The cleavage decreases the blocking capacity of the quencher on the probe, due to the no longer close attachment to the probe of both reporter dye and quencher. By that letting the reporter dye to give a detectable signal. The Completion of polymerization: Taq polymerase continues to synthesize the target sequence until it has created the finished product. During the amplification of the target sequence, at each cycle each template is doubled. Due to this exponential increase in numbers of templates is possible and the signal intensity increases due to that after each cycle (Butler J., 2012, p.49-67).

2.2.3. QUENCHER AND FLUOROPHORE

The probe used has both a fluorophore dye and a quencher attached to its ends. One common fluorophore to use as a tracer agent is the 6-carboxyfluorescenin (6-FAM) (Stults J et al., 2001). The FAM molecule has a carboxyl group attached to a fluorescein. It emits cyan colored light due to its emission wavelength at around 500 nm. The fluorophore has been widely used with promising results as a tracer dye in analysis for liposomes in the body and cell division tracer among others.

The quencher Black Hole Quencher (BHQ) dye holds no fluorescents signal, often referred to as dark quencher, in comparison to another popular quencher called TAMRA dye. The TAMRA quencher dye (TetraMethylRhodAmine) is a potent quencher for fluorophores with a maximum emission at 560nm. TAMRA has intrinsic fluorescence, which can be a downside since it may complicate analysis, because of the crosstalk between data channels. Due to this BHQ would be preferable, it has been used with results showing that it simplifies the design of qPCR assay but also the implementation as well as the interpretation of the method. The BHQ has a large spectrum of absorption, due to its different types: BHQ-1, BH-2 and BHQ-3. They cover

wavelength from 480-730 nm and are often used to enable larger arrays in multiplexed assays (Biosearch technologies, n.d).

Another promising dark quencher to use is Iowa Black FQ quencher (IBFQ). This is mainly due to its wide range of absorbance, 420-620 nm as well as its stability over a broad spectrum of conditions, where it exhibits great robustness even if the pH or temperature changes. The quencher has the advance of having a hydrophobic handle which enables purification using HPLC. Many of these properties distinguish IBFQ from others and making it to the optimal candidate for the use in fluorescence probes. Further, it has been evaluated that the IBFQ also can make the assay more sensitive and decrease background noise, by that giving more distinct results even when detecting at very low concentrations is needed (Integrated DNA Technologies, n.d).

2.2.4 Primers and Probes

To make the RT qPCR work properly and delivering desired results, the primers and probes needs to operate together suitably. By that, the primers and probe should ideally match together in their properties. This is of great importance in order to have highly efficient and specific designed primers and probes. A parameter to consider upon designing primers is the melting temperature (T_m), since this is crucial for the method (an optimal T_m should be 60-64 °C, the ideal temperature lays in the middle at 62 °C). This T_m is modified from the average cycling, normal reaction conditions as well as the optimal temperature for the PCR enzyme used. The temperature for annealing of the primers also needs to be considered. The annealing temperature (T_a) depends directly on the length of the primer, the composition of oligonucleotides may affect as well (Rychlik et al., 1990). T_a should desirably maximum be 5 °C below the T_m. If one having a too low T_a, the annealing might occur on non-desired places on the template since part of the annealing or mismatching of internal base pairs may be accepted. If the T_a is too high, the primer might not anneal at all. Due to that, the efficiency may be decreased drastically. If T_a do not lay within the optimal range, it may lead to few or nonspecific PCR amplifications which will reduce the yield of the product. The length of the primer is another parameter to consider, which often lays between 18 and 30 bases. If having more or less than the optimal amount of a primer will make it unspecific. A longer chain will have a higher T_m since it takes a higher temperature to denature the primer. An unspecific primer is often correlated to a too short one due to its ability to bind on more than the target. A too long primer might have a too high T_a as well, and by that it may not anneal properly which will result in poor yields since amplification cannot occur if the primer has not bind (Sipos et al., 2006).

The ideal probe should be around 20 to 30 oligonucleotides and be located between the forward and reverse primer. Both the T_m and T_a are important for the probes as well. The T_m should preferably hold a 6-8 °C higher than the one of the primers. If the T_m is lower than the optimal, the amount of bound probe to the target sequence will be low. If so, the amplification will occur at optimum, if primers allow so, but a poor signal will be given, since few probes annealed to the template and can therefore be cleaved and sending out a signal. This will end up in poor detection with a result not representing the true number of templates. The T_a for the probe

should lay in the same range as for the primers. By that, around 5 °C below the T_m of the matching primers (Proudnikov et al., 2003).

Another parameter to have in mind upon designing the primers as well as probe is the amount of Guanine (G) and Cytosine (C) bases, so called GC content (%). The GC should ideally, theoretically lay between 40 – 60 % for both the primers and probe. A much lower GC will result in that the primer or probe might be too unstable. The explanation being that the binding between G and C is stronger than the binding between Thymine (T) and Adenine (A), since it has one more hydrogen bond. More than 3 out of 5 bases, in a row at the 3´end of the primers and probe, of GC bindings should be avoided if possible, since this might cause higher possibilities of secondary structure formations (Watts et al., 2009).

These are the main parameters to consider upon designing sets of primers and probes and the range of the different ones are guidelines of where to start to design and optimize. The parameters should be considered after findings and knowledge of the sequence of the target genome. By that, the primers and probe should theoretically bind properly to the target template in mind of the virus (Prediger E., 2013).

2.3 BIOINFORMATIC

Basic Local Alignment Search Tool, often referred to as BLAST is an easy-accessible bioinformatic algorithm commonly used. The tool identifies homologous similarities between two sequences of targets that are of interest. Using the tool of BLAST allows the researcher to compare sequences and identify them towards its library and by that, get more genetic and theoretical information regarding these specific genes or sequences. The program itself is developed for comparing elementary biological information found within sequences such as amino acid of proteins, nucleotides of RNA or DNA among others. The algorithm is created in a manner where it primarily finds the shortest match between the two sequences and does not consider the whole space of the sequence. Due to this, it might generate many possible as well as un-specific hits, but it helps the researcher to get good understanding of the target sequence and a hint of what to dig deeper into, for further investigation (Bergman NH., 2007).

2.4 CROSS-DIMER TOOL

The tool of Multiple Primer Analyzer evaluates cross-dimer hybridization as well as self-dimer formation for primers and probes theoretically. The tool can be used to predict if the possibility of creating cross-dimer is possible. In that way one gets precious information when designing and optimizing primers and probes within different assays. The tool works by identifying possibilities upon binding and hybridization of the primers and probes added for evaluation. The foundation of the tool lays within the theory of bioinformatics, but in a simpler way. The tool is built up in a manner to identify binding possibilities within the data you add, and not within a data bank (ThermoFisher (B), n.d).

2.5 Extraction of RNA

When performing molecular assays as digital PCR and RT qPCR among others, a highly important prerequisite is to obtain proper quality RNA. In order to isolate and purify the desired RNA of the virus particle, which lay within the blood, extraction is necessary. There are many available extraction methods, all of which are more or less suitable together with different analysis methods. Due to this, the extraction method used should be well adequate for a specific analysis technique. The purification of the RNA is an important step in order to generate highly relevant as well as most sensitive result for the assay (Zeringer et al., 2013). A used extraction method uses magnetic particles to purify the RNA. The method is built on magnetic attractions between a magnetic field and the added magnetic granules. The small magnetic granules hold a paramagnetic nucleus which is coated with a modified layer in order to attract some parts of the lysed virus particle. The paramagnetic granules will migrate when being exposed to a magnetic field. The granules hold minimal magnetic memory, and by that will not be affected if not being exposed to a magnetic field. Due to this, the granules can interact with relevant fragments of the virus. The granules are then migrating and isolate by the exposed magnetic field. By lysing the samples in a RNase inhibitor solution, allows the granules and the undesired fragments of the lysed virus to attach to one another upon magnetic field exposure. Upon numerous rounds of release of magnetic granules, the RNA suspension is resuspended in washing solution. The RNA is further put into an elution solution, TE buffer. Leaving the RNA pure in the solution due to the paramagnetic granules with attachments are being removed (ThermoFisher (C), n.d). In the present project the instrument NucliSENS® easyMAG (BioMérieux) was used allowing the easy automated extraction of RNA via adsorption of undesired cell elements to micron-sized magnetic silica particles (BioMérieux, 2009).

3 Methods

3.1 The design of primers and probes

A literature study was conducted as a starting point for the project. The use of the bioinformatic tool BLAST and HIV Sequence Compendium indicated which genes within the genome of HIV-1 are highly conserved (Abfalterer, 2013). The identified target regions chosen were within the region of Long Terminal Repeat (*LTR*) both on the side of the 5´end and the 3´end, as well as within the *Pol* gene (Eberle and Gürtler, 2012). Due to the knowledge of conserved genes of HIV-1 genome, articles where found within the subject of these regions. Suggestions from the publications were selected as the target sequences. Here previously studied and found promising sets of primers and probes where selected for further investigation. The primers and probes selected had all been studied previously and suggested to be suitable for detection of HIV-1 with the use of PCR assays. The argument of using conserved parts as target sequences, lays within the fact these will less likely mutate as frequent as other sequences. The opinion of which is the best target sequence of HIV-1 among scientist is divided (Lusic and Siliciano, 2017).

Different sets were selected for further investigation due to the widely found promising properties within these sets of primers and probes, as well as the ambition of targeting more

than one sequence of the HIV-1. The *Pol* gene, coding for the polymerase of the virus and has previously been suggested as a highly conserved part of the genome and sets targeting this was selected (Eberle and Gürtler, 2012). The gene of *LTR* which do not code for any essential proteins for the virus, has been consider as a promising conserved sequence. Due to its tolerance for mutations caused by selection pressure of the immune system (Lavezzo et al., 2017).

The sets of primers and probes selected from the literature study where compared with various established genotypes using the HIV Sequence Compendium (Abfalterer et al., 2013).

3.2 ALIQUOTATION

Primers and probers selected from the literature study was ordered in concentrations of 100 µM (from Integrated DNA Technologies). Upon arrival, the primers and probes were diluted into concentrations of 60 µM and 12 µM both corresponding to a concentration of 600nM in the PCR wells. TE buffer was used as dilution media (from Thermo Fisher Scientific). Each primer and probe were further divided into aliquots of 15 µL. All aliquots were further stored at -70 ºC.

3.3 EXTRACTION

Extraction of viral RNA was done using the instrument NucliSENS® easyMAG (BioMérieux).

Blood plasma was received from blood banks and used as the starting material, referred to as the samples. The positive HIV plasma and negative plasma have already been tested on beforehand by well-trained personnel, in order to guarantee presence of absence of virus.

The extraction performed started with the preparation of 5 mL tubes. The tubes were used for the extraction samples, one tube for each sample. 16 samples were prepared, 8 for positive control (PosQ612B) and 8 for negative control (negative plasma). Internal control (IC) was added to each sample to confirm further on that the test has been executed properly. 50 µL of IC virus was added to each one of the 16 tubes. Further 1 mL of positive control and 1 mL of negative control were added to the tubes according to the work list, 8 tubes each. 2 mL of lysis buffer was added to each one of the 16 tubes. All tubes were turned upside down approximately 10 times before placed on a Multi Speed vortex for 1-5 minutes. Thereafter, the samples were placed in an easyMAG sample cassettes for incubation for 20 to 120 minutes. 0.6 mL of a magnetic silica absorbent was added to 2 separate tubes. 550 µL of RNase and DNase free water was added to each one of the two tubes. Further was the magnetic silica mixed by pipetting the solution up and down approximately 10 times. Then, 100 µL of the magnetic silica mix was added to each one of the 16 tubes, 8 tubes with positive control and 8 tubes with negative control. The cassettes with samples as well as compatible pipet tips were put into the easyMAG instrument, and the extraction program was initiated. The product of positive control extracted will be purified RNA of HIV-1. Whereas, negative control extracted will be purified RNA from non-HIV infected samples.

The positive control, Pos Q612A was extracted by well-trained personnel on beforehand, and was stored in -70 ºC. The Pos Q612A holds another titer of virus particles due to the difference

in starting material and procedure of the extractions. For the Pos Q612A control, the starting material was 5 ml blood plasma and was initiated by a centrifugation step, in contrast to the control of Pos Q612B. The positive controls of PosQ612A and PosQ612B origin both from subtype B of HIV-1.

The positive control, Pos Q612A holds a titer of 36 International Units/mL and positive control, Pos Q612B a titer of 219 IU/mL. The International Units is a standard decided by the World Health Organization. It is defined to as an indiscriminate amount of material per volume unit, determined by agreed scientists and doctors (Prescott. G et al, 2017). Where 1 IU/mL equals 5.6 copies of virus /mL (Luft et al., 2011). The sensitivity in the individual PCR well for each sample will therefore be 8.18 IU of Pos Q612A and 9.96 IU for PosQ612B, calculated from Calculation 1 and Calculation 2, receptivity.

$$36\frac{IU}{mL} \times \frac{5mL}{1mL} \times \frac{1mL}{0.11mL} = 1636.4\frac{UI}{mL}$$ Calculation 1

$$1636.4\frac{IU}{mL} \times 0.005\ mL = 8.18\ IU$$

$$219\frac{IU}{mL} \times \frac{1mL}{0.11mL} = 1990.9\ IU/mL$$ Calculation 2

$$1990.9\frac{IU}{mL} \times 0.005\ mL = 9.96\ IU$$

3.4 Real-Time quantitative Polymerase Chain Reaction

3.4.1 The Master mix

The master mix used in the experiment of RT qPCR was made for each PCR assay. The precursors: RNase and DNase free water and TaqMan Fast Virus 1-Step were mixed with the specific primers and probe of a defined set. The preparation of the master mix was made for 15 replicates. The total volume of master mix for one set was 225 µL and consisted of 105 µL RNase and DNase free water, 75 µL TaqMan Fast Virus 1-Step enzyme and 15 µL of each one of the primers and probe. The master mix was shaken on vortex for 2x2 seconds. The master mix was prepared to hold a concentration of 600nM of primers and probe within the PCR wells. Preparation of the master mix was made within a highly cleaned Laminar air flow bench with good air circulation. All the elements where transferred with well calibrated manual pipets in the range suited for the specific amount of each component. The reagents as well as the master mix were kept on cooling blocks during the whole experiment. The primers and probes and the finished master mix were all covered with aluminum foil, to keep the reagents away from light.

3.4.2 Real-Time quantitative Polymerase Chain Reaction analysis

The master mix was shaken on vortex for 2x2 seconds and centrifuged for 10 seconds. The master mix was then transferred in amount of 15 µL to each one of the PCR wells used in a Light Cycler 480® Multiwell Plate, referred to as plate. 10 replicates were performed. The positive control extracted from HIV-1 subtype B (Pos Q612A) was then added in amounts of 5 µL to 8 of the PCR wells with master mix. The last 2 PCR wells with master mix was used as negative control and 5 µL of Buffer 3 was added. The total volume is 20 µL for each PCR well. The plate was covered with a thin plastic film to avoid contamination of the samples as well as the surrounding. Further the plate was centrifuged for 2 minutes at 3000 rpm. All the preparations were executed within a highly cleaned Laminar air flow bench with good air circulation placed in a clean classed room. The reagents and the plate were kept on cooling blocks during the experiment, until the RT qPCR analysis. The plate was then transported and placed into the Light Cycler 480®, performing the RT qPCR analysis. The assay is performed within the pre-programmed program for RNA viruses. The program starts with the reaction of converting RNA into cDNA, by a Reverse Transcriptional stage, which is a pre-requisite for the method to work. It runs for 55 cycles, includes denaturation and annealing/extension in different cycles at temperature 95 ºC and 60 ºC, respectively. The program ends by a cooling stage. The schedule of the program used can be seen in Table 1. The Light Cycler 480® (Roche) does not start to plot data for the fluorescent signal until after the first 10 cyclers have finished (Boisselier, 2012).

Table 1. The schedule of the RT qPCR assay performed by Light Cycler 480® together with TaqMan Fast Virus 1-Step assay mix, used for detection of HIV-1 with novel developed primers and probe.

Step	Target	Acquisition Mode	Hold	Ramp Rate (ºC/s)
Reverse transcriptase-step	50 ºC	None	15 min	4.4
Denaturation/activation	95 ºC	None	2 min	4.4
Amplification	55 cycles (10 pre-cycles + 45 cycles)			
Denaturation	95 ºC	None	10 sec.	4.4
Annealing/extension	60 ºC	Single	35 sec.	2.2
Cooling	37 ºC	None	30 sec.	2.2

3.5 Specificity analysis

A RT qPCR was performed with 8 positive control (Pos Q612A) and 8 negative control (negative plasma), as the eluate. The master mix was prepared as previously described in section 3.4.1 but with volumes of 100 µL TaqMan Fast Virus 1-Step, 140 µL of DNase/RNase free water and 20 µL each of the primers and probe within the set used for detection. The total volume of the master mix for one set of primers and probe was 300 µL, prepared for 20 PCR

wells. The preparation of the plate was further executed and followed by the assay of RT qPCR, as previously described in section 3.4.2.

3.6 Internal control

The experiment was executed with the setup as described in section 3.4.2. The master mix prepared contained the Internal Control (IC-B). IC-B virus Bovine Viral Diarrhea Virus (BVDV), was added to the positive and negative control during extraction, described in section 3.2. The primers and probe for detection of the IC-B virus was added to the master mix during preparation. The primers and probe of IC were in concentrations of 60nM and 30nM respectively. The master mix contained of 100 µL TaqMan fast Virus 1-Step, 136 µL of DNase/RNase free water, 4 µL of IC-B primers and probe mix and 20 µL each of the primers and probe within the set used for detection. The total volume of the master mix for one set of primers and probe is 300 µL, prepared for 20 PCR wells.

3.7 Optimization

3.7.1 Salt influence analysis

Master mix was prepared as in section 3.4.1, with different amount and type of dilution media. Each set of primers and probe was prepared into three different master mixes all contained different amount of salt, but the same concentration of primers and probes of the PCR wells at 600nM. Master mix 1 contained: 50 µL TaqMan Fast Virus 1-Step, 92 µL DNase/RNase free water, 2 µL primers and probe mix of IC and 2 µL each of primers and probe from more concentrated aliquots of 60 µM. Master mix 2 contained: 50 µL TaqMan Fast Virus 1-Step, 68 µL DNase/RNase free water, 2 µL primers and probe mix of IC and 10 µL each of the primers and probe from less concentrated, previously used aliquots of 12 µM. Master mix three contained: 50 µL TaqMan Fast Virus 1-Step, 68 µL of TE buffer, 2 µL primers and probe mix of IC and 10 µL each of the primers and probe from the less concentrated, previously used aliquots of 12 µM. The RT qPCR analysis was then preformed as described in section 3.4.2.

3.7.2 Concentration influence analysis

Master mix was prepared as described in section 3.4.1. Each set of primers and probe was prepared in different concentrations of the primers and probe. The concentrations of the primers were: 600nM, 300nM, 300nM, 150nM. The concentrations of the probes were: 300nM, 300nM, 150nM, 75nM. The different concentrations were matched up as observed in Table 2. Each master mix was prepared for 10 wells and contained: 50 µL TaqMan Fast Virus 1-Step and 2 µL of IC primers and probe. Besides that, master mix 1 contained: 73 µL DNase/RNase free water, 10 µL of each primer and 5 µL probe, from aliquots with a concentration of 12 µM. Master mix 2 also contained. 83 µL DNase/RNase free water, 5 µL each of the primers and probe, from aliquots with a concentration of 12 µM. Master mix 3 also contained: 73 µL DNase/RNase free water, 5 µL of each primer, from aliquots with a concentration of 12 µM

and 15 µL of probe from a diluted aliquot with a concentration of 2 µM. Master mix 4 also contained: 53 µL DNase/RNase free water, 15 µL of each primer from a diluted aliquot with a concentration of 2 µM and 15 µL probe from diluted aliquot with a concentration of 1 µM.

The diluted aliquot with concentration of 2 µM was prepared with 9.9 µL of a primer or probe, from aliquots with a concentration of 12 µM and 50.1 µL of DNase/RNase free water. The diluted aliquot with concentration of 1 µM was prepared with 5.1 µL of a probe from aliquots with a concentration of 12 µM and 54.9 µL of DNase/RNase free water.

Each pair of concentrations were analyzed in replicates of 8. 7 of those were with positive control, Pos Q612B and 1 with negative control, negative plasma. The RT qPCR was then performed as described in section 3.4.2.

Table 2. Different concentration used in the experiment of Concentration influence analysis. Each row shows the match up of concentration between primer and probe tested for each set of primers and probe.

Master mix	**Primer concentration**	**Probe concentration**
1	600nM	300nM
2	300nM	300nM
3	300nM	150nM
4	150nM	75nM

3.7.3 COMPATIBILITY ANALYSIS

Master mix was performed as described in section 3.7.2, with different combinations of concentrations of the primers and probes. Tests were performed for mixes of 2 sets. The different master mixes contained sets of 2 as observed in Table 3. All master mixes were prepared for 10 replicates of 15 µL, 150 µL in total. Master mixes 1,5,9 contained 48 µL DNase/RNase free water. Master mixes 2,6,10 contained 68 µL DNase/RNase free water. Master mixes 3,7,11 contained 48 µL DNase/RNase free water. Master mixes 4,8,12 contained 8 µL DNase/RNase free water. Otherwise the master mixes contained the same ratios of elements are described in section 3.7.2. Each master mix was in replicates of 7 with positive control (Pos Q612B) and 1 with negative control (negative plasma). The RT qPCR assay was performed as described in section 3.4.2.

Table 3. Shows the different combinations of concentrations of primer and probe, in each master mix. Different sets were combined with their different combinations of concentrations of the primers and probe.

Master mix	***LTR 3´*** **Concentration**		***Pol*** **Concentration**		***LTR 5´*** **Concentration**	
	Primers	Probe	Primers	Probe	Primers	probe
1	600nM	300nM	600nM	300nM		
2	300nM	300nM	300nM	300nM		
3	300nM	150nM	300nM	150nM		
4	150nM	75nM	150nM	75nM		
5	600nM	300nM			600nM	300nM
6	300nM	300nM			300nM	300nM
7	300nM	150nM			300nM	150nM
8	150nM	75nM			150nM	75nM
9			600nM	300nM	600nM	300nM
10			300nM	300nM	300nM	300nM
11			300nM	150nM	300nM	150nM
12			150nM	75nM	150nM	75nM

3.7.4 CROSS-DIMER EFFECT ANALYSIS

The primers and probes used in the compatibility test, described in section 3.7.1 were added to the data tool of Oligo Analyzer for DNA and RNA Primers (Thermo Fisher Scientific). Both for evaluation of self-dimer formation and of cross-dimer hybridization between the primers and probes used.

3.8 SUBTYPE ANALYSIS

Extraction was performed as described in section 3.3 for each of the subtypes available: A-D, AE and F-H, as well as for negative control (negative plasma) and positive control (Pos Q612B). Master mixes was performed as described in section 3.6 with concentrations of 600nM of primers and probe for the set targeting *LTR3´*. Concentrations of 600nM of primers and 300nM of the probe for the set targeting *Pol.* Concentrations of 300nM each of primers and probe for the set targeting *LTR5´*. Master mixes of any 300nM contained 12.5 µL of each element and 600nM contained 25 µL of each. All master mixes contained 25 replicates, in total prepared for 375 µL. Each subtype was tested in duplicates, the positive control (Pos Q612B) in 4 replicates and the negative control (negative plasma) in 1 replicate, for each set of primers and probe. The RT qPCR assay was performed as described in section 3.4.2.

4 RESULTS

4.1 THE PRIMER AND PROBE DESIGN

Primers and probes chosen after literature study were all tested as described in section 3.4 as a first evaluation. The experiment conducted will reveal if the selected primers and probes are specific enough as well as if they can work properly. Therefore, it will be discovered if the primers and probes are able to detect the desired target within the HIV genome, or not. Hence, significance will be put in evaluation in the quantity of the signal.

The first assays tested the sets found within the *LTR3´*region. The results, which can be seen in Figure 7, shows 3 different sets within the *LTR3´*. Negative control, as buffer 3 fills the wells of column 2, 4, and 6 within the plate. Column 1, 3 and 5 contained positive control (Pos Q612A). As the results reveal, one can observe that the first set with positive control, in column 1 does not give any positive signal in none of the wells. The second set with positive control, in column 3 gives a positive signal of 3 out of 8 wells. The last set, in column 5 with positive control gives positive signal in all wells. The positive signal given indicates a presence of the HIV virus in amounts within the sample, which can be detected. Since all the odd columns hold HIV virus in the form of positive control, the ability of detecting the virus can be evaluated from the number of positive signals given. The primers and probe set of the positive control of the 5th column was further selected, due to its prominent positive signal in all wells of positive control.

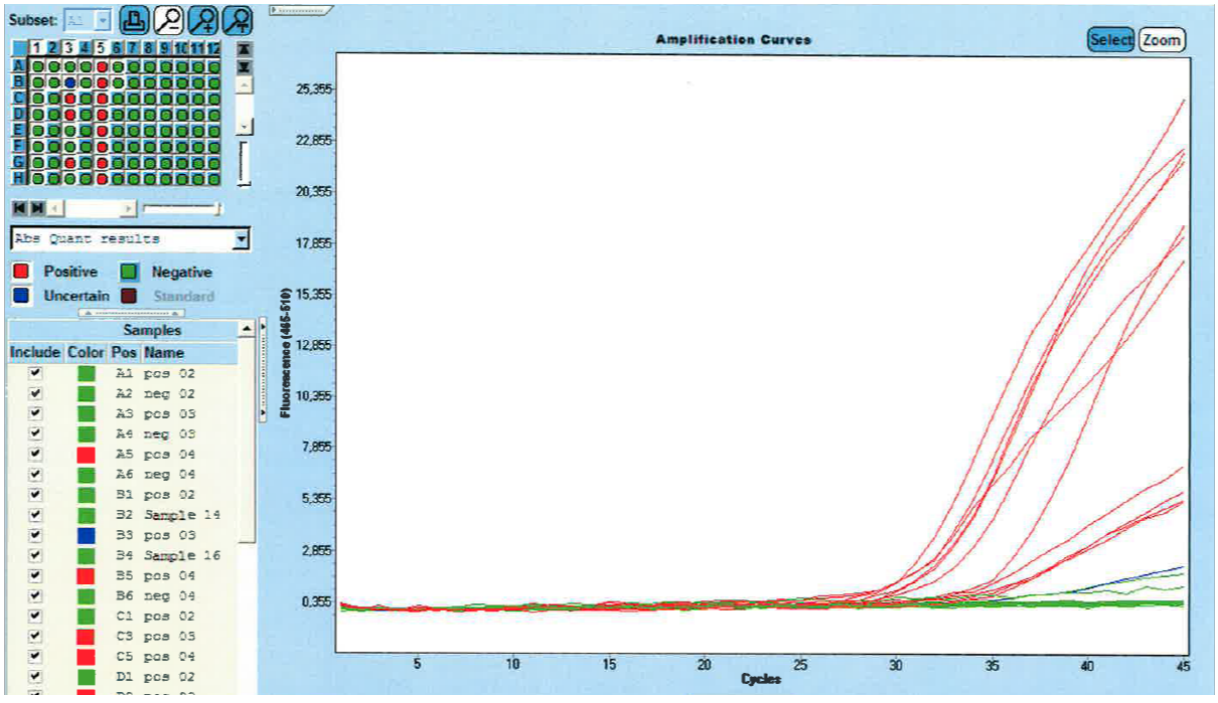

Figure 7. The graph and the plate set up show the data regarding the results of the first evaluated region of the HIV-1 genome, targeting *LTR3´* with RT qPCR assay. Each set of primers and probe was examined in two columns, with positive control as PosQ612A in one and negative control as buffer 3 in one. The fluorescence signal shows on the y-axis and the cycle value of the assay shows on the x-axis.

The second assay tested sets within the middle *Pol* region of the HIV-1 genome. 4 different sets of primers and probe were tested. The execution of the lineup of the assay was the same as described for the first assay, with positive control (PosQ612A) in odd columns and negative control as buffer 3 in even columns. The first set of *Pol* in column 1 with positive control showed positive signal in 4 out of 8 wells. The second set in column 3 with positive control showed positive signal in 5 out of 8 wells. The third set in column 5 with positive control showed positive signal in 5 out of 8 wells. The last set in column 7 with positive control showed positive signal in 6 out of 8 wells. The results can be seen in Figure 8. All investigated sets of primers and probe gave some signals. The set within the 6th column gave the most signals; therefore, it was selected for further studies.

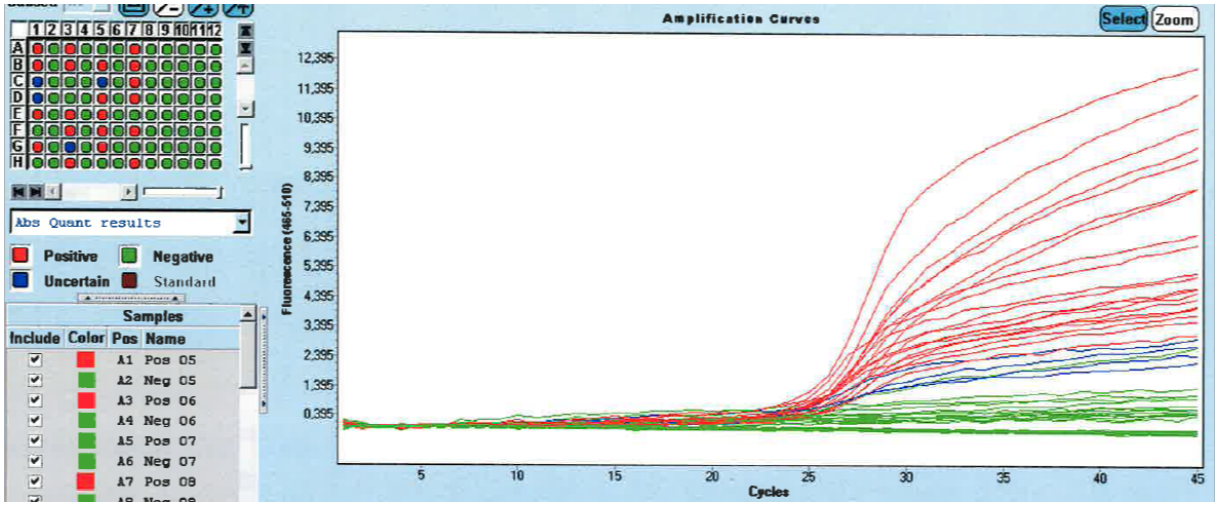

Figure 8. The graph and the plate set up show the data regarding the results of the first evaluation of the sets within the middle *Pol* region of HIV-1 genome with RT qPCR assay. Each set of primers and probe was examined in two columns, with positive control as PosQ612A in one and negative control as buffer 3 in one. The fluorescence signal shows on the y-axis and the cycle value of the assay shows on the x-axis.

The third assay examined the *LTR5'* region. Positive control (Pos Q612A) and the first set in column 1 showed no positive signal in none of the wells. In column 3, positive control (PosQ612A) and the second set tested showed positive signal in 7 out of 8 wells. The third set and positive control (PosQ612A) in column 5 showed positive signal in 5 out of 8 wells. The results can be seen in Figure 9. The primers and probes within the 3rd and 5th columns gave results with positive signal in most of the wells. The set of the 3rd column holds the best quantitative results from this experiment and was therefore selected with the previously elected sets of primers and probe. 3 sets of primers and probe were elected from the first evaluation upon quantitative data from the presented results of this experiment, for further investigation.

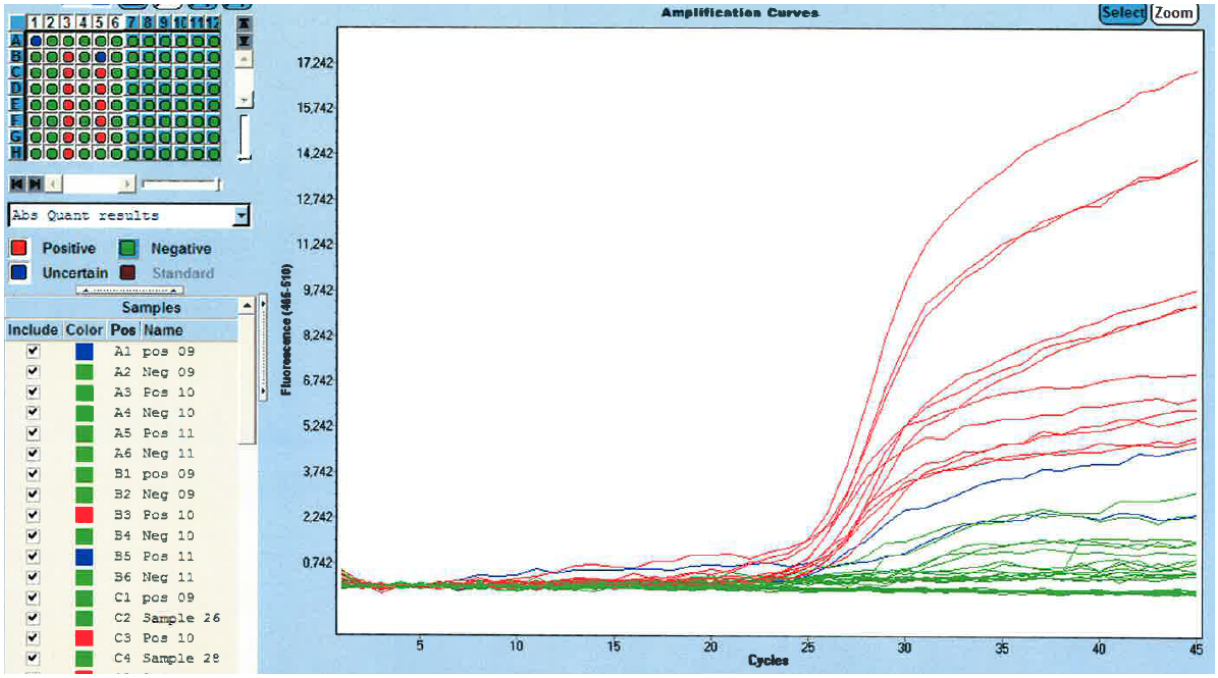

Figure 9. The graph and the plate set up show the data regarding the results of the first evaluation of the sets within the end region of *Pol* in column 1 and 2, with positive control as PosQ612A and negative control as buffer 3, respectively. The 3-6 columns contained sets within the LTR5´region, two different sets divided in 2 columns each. The region was examined with RT qPCR assay. The fluorescence signal shows on the y-axis and the cycle value on the x-axis.

4.2 Specificity test

The best sets of primers and probe within each PCR assay were chosen for further investigation. The three sets were: the first set of the first assay, targeting the LTR 3´region, the fourth set of the second assay, targeting the *Pol* region and the second set of the third assay targeting the *LTR5´*region.

To reassure no false positive signals were present the negative control was changed from Buffer 3 to extracted negative plasma. The lineup of the plate was conducted by the manner that all odd numbers of columns hold the negative control of extracted negative plasma and the even numbers of columns contained the positive control (Pos Q612A). Columns 1 and 2 represents the set targeting *LTR3´*region, with negative control and positive control, respectively. The following two columns of number 3 and 4, represents the set targeting the *Pol* region with negative control and positive control, respectively. The last set targeting the *LTR5´*region, represents in the 5 and 6 columns with negative control and positive control, respectively. As can be observed in Figure 10, all the odd columns containing extracted negative plasma does not show any signs of a positive signal. Indicating that no false positive signal has been detected. Therefore, all results given can be trusted to represent the signal of the detection of the desired target, not anything else within the matrix.

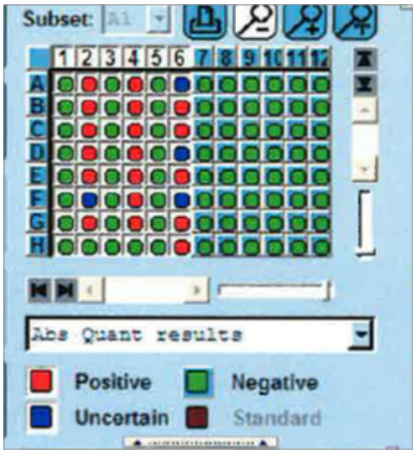

Figure 10. The lineup of the PCR plate as well as the results of the different possible signals of the PCR assay. Samples were within columns of 1-6. Columns of numbers 1,3 and 5 holds extracted negative plasma as negative control and shows no positive signal in any wells. The columns of 2,4 and 6 holds positive control of PosQ612A and shows signal of positive signal as well as some uncertain and 2 of negative character of signal. In columns 1 and 2 the set targeting the *LTR3´*region was examined. In columns 3 and 4 set targeting *Pol* region was examined. In column 5 and 6 set targeting the *LTR5´*region was examined.

4.3 INTERNAL CONTROL

Internal control was investigated by comparing the same setup of the three best primers and probe set from the first evaluation of screening. The experiment was performed as described in section 3.4 and 3.6. The experiments performed of section 3.4, with no IC added to the master mix showed some indication of IC signal, within that range of detection for all three used sets. The hint of amplification curves can be observed in the upper row of Figure 11, for all graphs with no IC. Most of the curves remains green in the computer program due to the amplification did not reach a prominent value of fluorescence signal. However, the graphs still indicate some type of amplification due to their sigmoidal appearance.

Upon addition of IC to the master mixes as described in section 3.6, observation of positive IC signal is prominent within its range of detection. All primers and probe sets tested showed a nicely appearance of sigmoidal curves with prominent fluorescence signals. The results from the addition of IC to master mix can be observed in Figure 11, in the lower row for all graphs with IC.

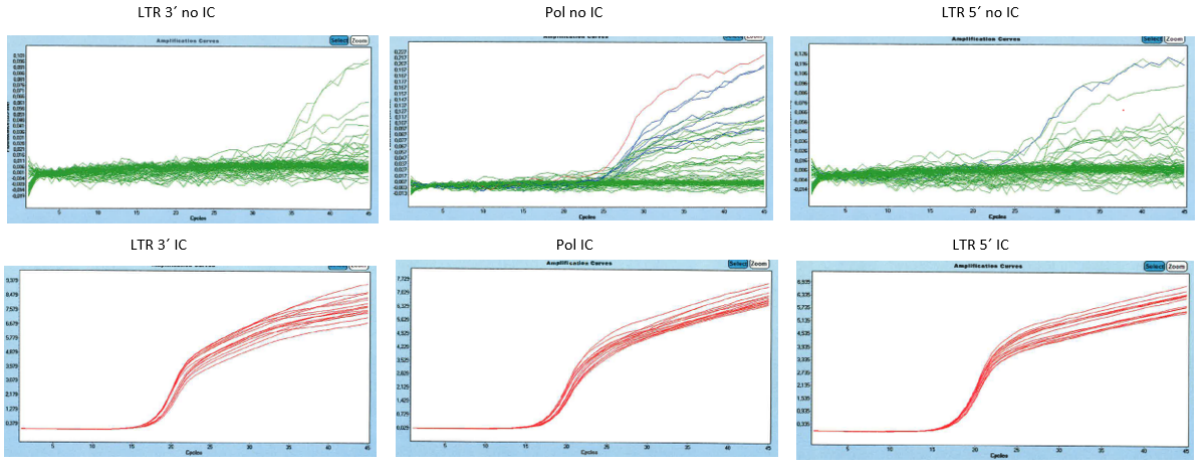

Figure 11. The three sets of primers and probes, targeting *LTR3´, Pol* and *LTR5´*region, respectively. The upper row shows the signal of IC, without the primers and probe of IC added. The lower row shows the same set of primers and probe but with added primers and probe for IC in the IC signal channel. The fluorescence signal shows on the y-axis and the cycle value of the assay shows on the x-axis.

For further investigation of the small appearance of sigmoidal curves within the samples with no added IC, a linear regression was conducted for all three sets with data from experiment described in section 3.4 and 3.6. The result of the linear regression analysis can be observed in Figure 12. The y axis holds the values of IC signal and the x axis holds the values of fluorescence signal of positive control. In the upper row of graphs, from the screening experiment illustrate the three used sets with no added IC to the master mixes. In the lower row of graphs, IC was added to all the master mixes. The interpretation of the linear regression analysis indicates that there is no correlation between the samples of no added IC and added IC, in the range of IC detection. The slope of the linear regression holds a different value in the two compared cases as well as the values of y axis are far from the same range.

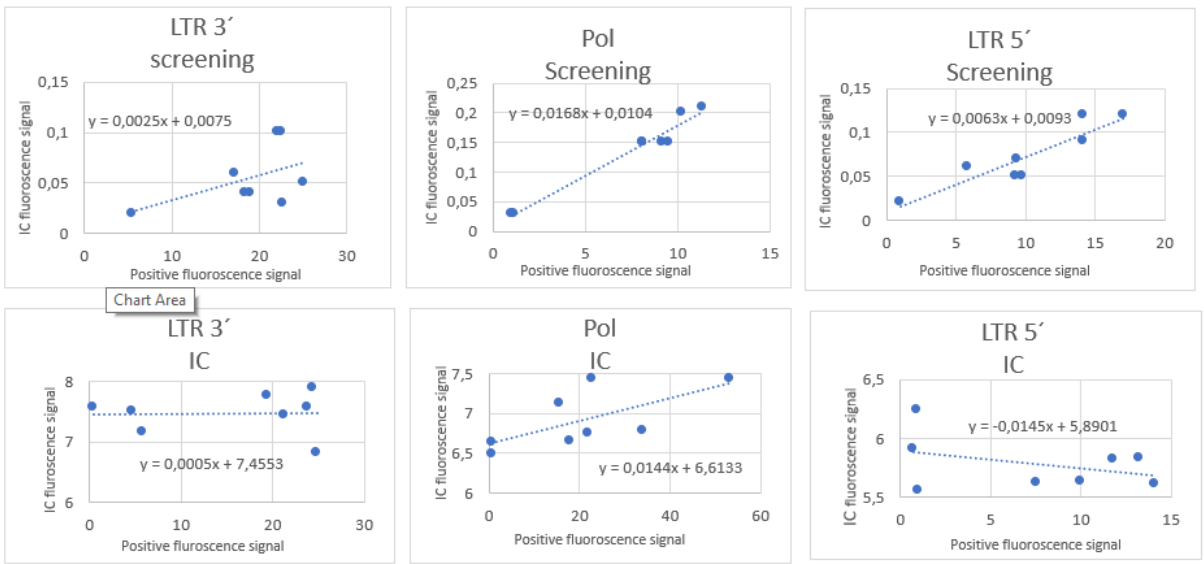

Figure 12. The data from the results in figure 11. The upper row shows linear regression of the data from experiment without addition of IC primers and probe, for the three sets targeting the *LTR3´*, *Pol* and *LTR5´*region, respectively. The lower row shows the linear regression of the data from the experiment with addition of IC primers and probe. The y-axis shows the values of the fluoresce signal and the x-axis shows the cycling values of the assay.

4.3 Optimization

4.3.1 Salt influence test

The influence of salt from the TE buffer used during preparation of aliquots was investigated. The results are shown graphically in Figure 13. The y axis holds the values of positive signals in percentage and the x axis holds the discrete values of intervals referring to 0.4%, 4.4% and 38.4% of TE buffer in the master mix. The data set in blue represents the primers and probe set targeting the *LTR3´*region, which has a slightly decreasing look in percentage of positive signals upon higher amount of salt. The primers and probe set targeting the *Pol* region, represented in the orange data shows a decrease followed by an increase in percentage of positive signal. The *LTR5´*region, represented in the grey data holds the same percentage throughout all the different amount of TE tested. The results therefore suggest that salt does not interfere in a negative way for the primers and probes targeting the *LTR5´*and *Pol* regions. Whereas the *LTR3´*data reveals that salt might have an undesirable effect upon higher salt concentrations, and thereby is less robust to salt.

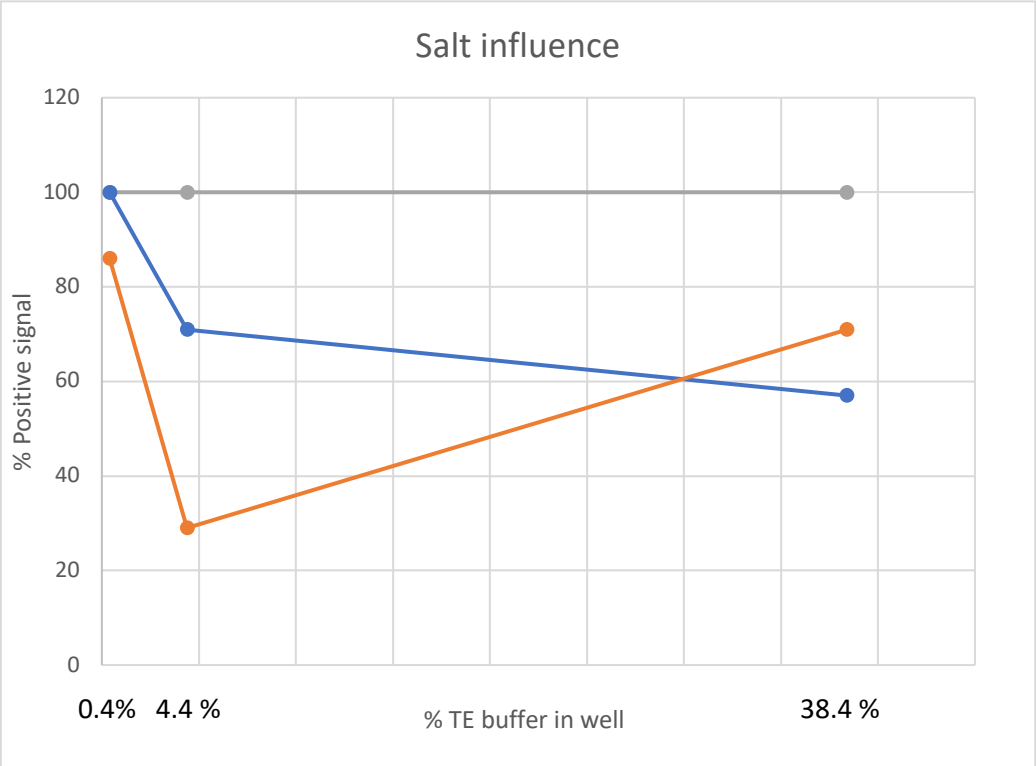

Figure 13. Data from salt influence experiment. The y-axis holds the values of percentage of positive signal of all the positive control samples, the x-axis holds the discrete values of percentage intervals of the amount of TE buffer used in the different wells. The grey data shows the result of the primers and probe set targeting the *LTR5´*region. The blue data shows the result of the primers and probe set targeting the *LTR3´*region. The orange data shows the result of the primers and probe targeting the *Pol* region.

4.3.2 Concentration influence test

Concentrations of primers and probe were combined within each set. The combinations of the different concentrations can be observed in Table 2, in section 3.7.2. The results of positive signals in percentage of each concentration combination, described as a relationship of the two can be observed in Figure 14. The y axis holds the values of positive signals in percentage. The

concentration combinations of primers and probes: 600nM + 300nM, 300nM + 300nM, 300nM + 150nM, 150nM + 75nM, respectively describes the intervals on the x axis in Figure 14. In Figure 14: the orange data describes the set targeting *LTR3´*. The yellow data describes the set targeting *Pol* and the grey data describes the set targeting *LTR5´*. As the observations reveals, the concentration combination of 600nM + 300nM was the most optimal tested for the sets of *LTR3´*and *Pol*. Whereas the concentration combination 300nM + 300nM was the most optimal tested for *LTR5´*. Due to the observation reveals that at this point the graphical data reaches its maximum within the investigated interval for each of the tested sets of primers and probe.

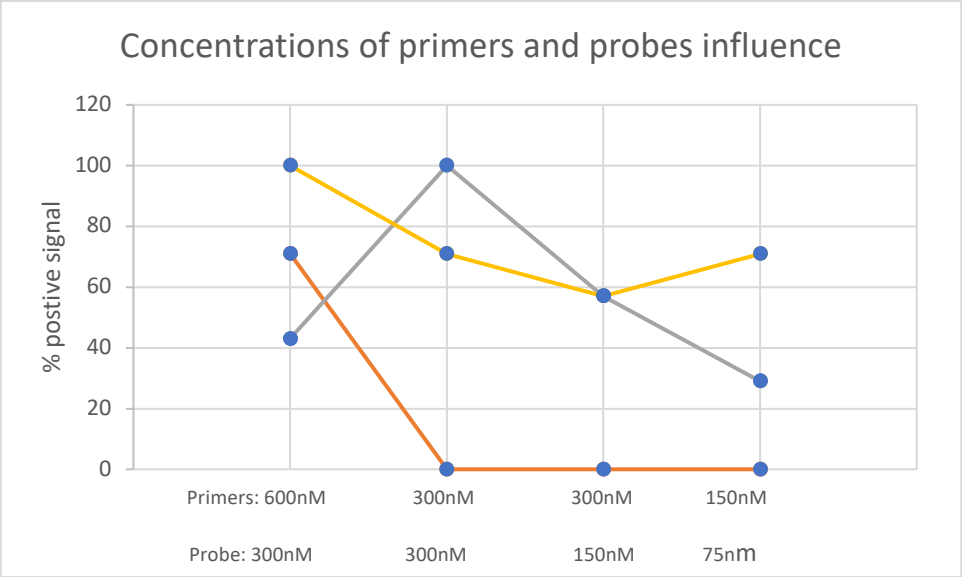

Figure 14. The relationship between different combinations of concentrations for the sets targeting: *LTR3´*in orange, *LTR5´* in grey and *Pol* in yellow. The different concentration combinations were: 600nM + 300nM, 300nM + 300nM, 300nM + 150nM and 150nM 75nM, of primers and probe, respectively. All the concentration combinations can be observed in Table 2.

4.3.3 Compatibility test

To investigate the compatibility between different sets, the sets: *LTR3´*, *Pol* and *LTR5´*were mixed in pairs as described in section 3.7.3. The positive signals given from the assay are presented in Table 4 for each of the mixes, and their concentrations, respectively. The results reveal that the compatibility of the sets is not promising, at this stage since only two concentration mixes out of twelve gave a positive response. The first of two samples giving a signal of detection were the combination of *LTR3´*and *Pol* of 150nM and 75nM of primers and probes, respectively. The second sample was the combination of *LTR3´*and *LTR5´*of 300nM of each primers and probes. None of the two were able to detect the positive control in all possible wells given. The best, with 4 out of 7 replicates lays within the combination of *LTR3´*and *LTR5´*.

Table 4. The master mixes tested in compatibility test with concentrations of the different combinations of sets. The results are to be observed, of the positive signals given for the different combinations in the column to the right.

Master mix	***LTR3´* Concentration**		***Pol* Concentration**		***LTR5´* Concentration**		**Positive signal response**
	Primers	Probe	Primers	Probe	Primers	probe	-
1	600nM	300nM	600nM	300nM			-
2	300nM	300nM	300nM	300nM			-
3	300nM	150nM	300nM	150nM			-
4	150nM	75nM	150nM	75nM			1/7
5	600nM	300nM			600nM	300nM	-
6	300nM	300nM			300nM	300nM	4/7
7	300nM	150nM			300nM	150nM	-
8	150nM	75nM			150nM	75nM	-
9			600nM	300nM	600nM	300nM	-
10			300nM	300nM	300nM	300nM	-
11			300nM	150nM	300nM	150nM	-
12			150nM	75nM	150nM	75nM	-

4.3.4 CROSS-DIMER EFFECT TEST

Cross-dimer possibilities between the different primers and probe sets were investigated as described in section 3.7.4. The results showed some possible cross-dimer formation between 2 different primers and probes. The data generated can be observed in Table 5. The binding between the reverse primer of the *LTR3´*set and the forward primer of the *LTR5´*set, holds a possible binding between 10 bases. The probes of the *LTR3´*set and the *LTR5´*set holds a possible binding of 10 bases. The maximum number of bases to bind is 20 bases in both cross-dimer possibilities found. Due to the length is 20 bases of the shortest primer or probe within the binding pair and by that is the limiting factor. The results are hard to interpret since the possible cross-dimer formations lays within the combination of sets of primers and probes giving the most signals of detection of the combability test.

Table 5. Investigation of cross-dimer effect with the tool of Multiple Primer Analyzer from Thermo Fisher. The possible cross-dimers to occur, in theory, are from the binding between the *LTR3´*reverse primer and the *LTR5´*forward primer as well as from the binding of the probes of *LTR3´*and *LTR5´*. Both holds a maximum binding potential of 20 bases and binds 10 of these.

Possible cross-dimer	**Number of binding bases**	**Number of maximum bases to bind**
LTR3´ reverse primer ⋮ *LTR5´*forward primer	10	20
*LTR3´*probe ⋮ *LTR5´*probe	10	20

Data regarding the properties was given from the analysis of cross-dimer and can be observed in Table 6. The data presents properties regarding T_m, content of G-C bindings (GC%) and numbers of bases of each primer and probe. The GC% of each set of *LTR3´, Pol* and *LTR5´* was summarized to a value of 38.4%, 33.3% and 56.5%, respectively. The results were evaluated by correlate the data of the properties to the number of positive signals for each set of primers

and probe. This data is represented in Table 6 in the right section. 24 replicates have been evaluated with the correlation of GC% of each set of primers and probe. The replicates represent the positive signals detected from screening, sensitivity and IC analysis. All these experiments were performed the same regarding the positive control (Pos Q612A) samples. The positive signals were summarized for each set of *LTR3´*, *Pol* and *LTR5´*to numbers of 17, 18, 19. The data indicates that a lower amount of GC%, out of the possibilities of GC% evaluated, is more feasible in order to receive more positive signals.

Table 6. Shows data from primer and probe evaluation analysis and data correlated between GC% and numbers of positive signals given in 24 replicates. The replicates are collected from experiments regarding the screening, selectivity and IC. All these experiments were performed the same regarding the positive control, by that the positive signals.

Primers and probes	**GC (%)**	**T_m (ºC)**	**Number of bases**	**Positive signals**
*LTR3´*forward primer	47.8	65.7	23	-
*LTR3´*reverse primer	16.7	66.1	30	-
*LTR3´*probe	60.0	61.4	20	-
Set of *LTR3´*	38.4			17/24
Pol forward primer	29.6	62.3	27	-
Pol reverse primer	20.7	62.7	29	-
Pol probe	57.9	68.4	19	-
Set of *Pol*	33.3			19/24
*LTR5´*forward primer	41.7	67.6	24	-
*LTR5´*reverse primer	68.4	70.3	19	-
*LTR5´*probe	50.0	65.6	26	-
Set of *LTR5´*	56.5			18/24

4.3.5 SUBTYPE TEST

The subtypes of the genotype M were investigated as described in section 3.8. The results presented as number of the replicates out of positive signals given from the assay are presented in Table 7. As to be observed, none of the genes suggested as target regions were able to detect all subtypes tested. The most promising of the three genes seems to be the *Pol* gene*,* regarding its possibilities to detect the most subtypes, even if not all. When combining the sets, detection of all subtypes is possible with a probability of: A:2/12, B:3/12, C:6/12, D:1/12, AE:2/12, F:4/12, G:8/12, H:1/12. Thereby, suggesting that the results is not desired at this stage of the primers and probes used in the sense of their specificity.

Table 7. The result from experiment regarding the subtypes of genotype M of HIV. The subtypes tested is represented in the column to the left, with corresponding ratio of detection for each set targeting *LTR3´*, *Pol* and *LTR5´*tested.

Subtype	***LTR3´***	***Pol***	***LTR5´***	**Sum**
A	0/4	2/4	0/4	2/12
B	2/4	1/4	0/4	3/12
C	3/4	1/4	2/4	6/12
D	0/4	1/4	0/4	1/12
AE	0/4	1/4	1/4	2/12
F	1/4	2/4	1/4	4/12
G	1/4	4/4	3/4	8/12
H	0/4	0/4	1/4	1/12

5 DISCUSSION

Findings from the conducted literature study gave a proper foundation to start regarding desired properties as well as targets for the primers and probes developed. Sequences from published research work within the conserved parts of the genome of HIV were screened with the RT qPCR assay, previously described in section 3.1 and 3.4 (Eberle and Gürtler, 2012). From these results, three promising targets were further evaluated with analysis in the wet lab.

The stock solution and aliquots of primers and probes were kept at -70 ºC and only the needed number of aliquots were unfrozen during master mix preparation and leftovers were discharged. The aliquots as well as the master mixes were kept on ice during the laboratory work. In this manner, any possibility of decreasing efficiency of the primers and probes was eliminated. Therefore, the unsuccessful data of some results cannot be explained by the incorrect handling of the primers and probes nor their shelf life (Integrated DNA Technologies, n.d).

The concentrations of primers and probes were in a range of suggested concentrations from published articles (Green and Sambrook, 2019). The 600nM has been suggested to be proper for evaluation of the possibility to work as desired. Upon optimising the primers and probes, different concentrations were investigated. Due to some published data indicated that cross-dimer problems may arise with to high concentrations (Brownie et al., 1997).

Negative plasma as a negative control was used to eliminate any possibilities of false positive signal. These might arise if the primers and probes are not specific and able to bind to other human originated sequences within the matrix other than the target sequence. As the results shows in section 4.2 no negative control showed any positive response. This eliminates any doubt of false positive signals (Drobniewski et al., 2020).

Internal control was used to certify that the enzymes of TaqMan Fast Virus 1-Step and Reverse Transcriptase worked properly, and the experiment had been carried out correctly. The IC signal of samples with and without addition of IC, were used to investigate if any bleed through of fluorescence signal did affect the detection. Since amplification appearance has been observed although no IC was added. If the fluorochromes of IC and the fluorochromes of the method used until this stage interfered, they could not be used together anymore. The data was investigated by linear regression and could be evaluated that the small bleed through noticed

did not affect the detection of the IC nor the positive signal. Because conclusions of no pattern can be drawn from the linear regression data. The range of the axis holds widely difference in signal appearance of the two compared experiments. By that, the fluorochromes of the different probes could be used and combined upon detection, since they do not interfere with each other's range of detection (Lütgenhetmann et al., 2020). The IC showed proper results for all assays, indicated that each reagent worked properly as well as execution of the experiments was correct.

The Crossing Point value (CP) for all RT qPCR assays was calculated to a mean value of 27 cycles, indicating the primers are well designed. Because the amplification occurs as desired and amplifications arise within the time lap of similar qPCR methods (Jansen et al., 2011).

The sensitivity of the method was desired to be high, in order to capture and detect all virus. Extracted positive control (Pos Q612A) with a titer as low as 8.18 IU/well was used (FDA, 2010). This indicates that the method with all three sets of primers and probe holds a high sensitivity of detection (Hemelaar J., 2012). Published sensitivity studies have reported prominent results with a titer of 10^4 IU/mL, resulting in a value of 100 IU/well if comparing with this projects results (Yang et al., 2013).

The set of primer and probe targeting the *LTR3´*region showed promising result during screening with a high fluorescence signal and 8 out of 8 positive signals. Further evaluation of the set indicated that salt effected the primers and probe slightly negative, due to the small decreasing signal of increasing concentration of TE Buffer in its master mix. The investigation of concentrations revealed that the set worked better with higher concentrations of both primers and probe. Since its maximum was located at the highest concentrations evaluated. The most optimal concentrations tested were therefore found to be 600nM of the primers and 300nM of the probe, which was further used in the subtype evaluation. The *LTR3´* targeting set showed detection of 4 out of 8 different subtypes, with a best detection of 75% of the replicates. The subtypes were B, C, F and G. In previous studies sets in this target area has only been able to detect subtype B (Shan. L et al, 2013). Therefore, the results are promising, as the range of subtypes being detected is greater than previously known. The fluorescence signal is prominent of all positive signals detected, indicating that the probe is properly designed.

The results of the screening regarding the sets targeting the *Pol* region, looked promising for one set with 6 out of 8 positive signals. The appearance of the signals was with a broad range of signal amplitudes, indicating different titers in some of the extracted eluate tubes used. This is most likely the case since the eluate from different extraction wells was not pooled together before stored in freezer. The salt concentration does not indicate to affect the set, in general based on the data in Figure 13, which show a decrease in positive signals when lowering the concentration of salt until a minimum value, then increasing again (Green and Sambrook, 2018). The data from the evaluated concentration analysis showed that the most optimal signals were obtained by the combination of 600nM of primers and 300nM of probe, within the range tested. As to be observed as the yellow data in Figure 14, the highest number of positive signals is given at this concentration combination. The set targeting *Pol* showed most promising so far for detecting the subtypes. It was able to detect 7 out of 8 subtypes, only leaving out H. The only set possible to detect at frequency 100% was the set targeting *Pol* for detecting subtype G.

The data from the subtype results is shown in Table 6. The results of the subtype analysis are highly positive, since reported and published data indicates that targets in this region has previously only been able to detect subtype B (Strain et al., 2013).

The screening of the targets within the *LTR5´*region was found highly promising for one set, holding 7 out of 8 positive signals. The fluorescence signal was prominent, but with a wide range of amplitudes indicating that the extracted HIV-1 RNA, might have been inhomogeneous during the experiment. The concentration of salt did not seem to affect the set, neither positive nor negative. The detection gave signal in all possible samples with 100% detection level each time. The concentration of primers and probe on the other hand showed indications of a lower concentration would be more optimal, than the previously used. The data gave results suggesting that 300nM of each primers and probe would be the most optimal concentrations, observed as the grey data in Figure 14. The detection of possible subtypes within the genotype M, showed signal for 5 out of 8 possible ones. The set was able to detect subtype C, AE, F, G and H. This result was dejected, since targets in this region has previously been proven to be able to detect all subtypes within genotype M (Drosten et al., 2006).

The compatibility between the sets was investigated in order to get more targets within one assay, which could lead to detection of more subtypes. The results from the compatibility test in section 4.3.3 reveals that the sets do not work properly together. A positive signal was only given in 2 out of 12 different possibilities of concentration and set combinations. This data was a disappointment since the sets had shown promising results separately and did not either show any significant problem in theory, such as cross-dimer (Das et al., 1999). The cross-dimer possibilities were investigated as described in 3.7.4 and showed only two possibilities of cross-dimer formation, presented in Table 5. It is possible that some possible cross-dimer formation may occur between forward and reverse primer of *LTR5´*and *LTR3´*, respectively, as well as the probes between these two sets. If cross-dimer formation is the major problem when mixing the sets, no positive signals would be given. Especially surprising is the fact that the combination of sets targeting *LTR3´*and *LTR5´*, is the combination giving the best signal in the compatibility test. Due to this, it is hard to state that this is a reason why the sets of primers and probes does not work together proper. At this stage, the problem with mixing the sets is due to an unexplained factor. The bioinformatic tool, Multiple Primer Analyzer from Thermo Fisher, revealed indications of possible self-dimer formation as well as secondary structure formation, this was not further investigated due to it was not considered as an issue. Because the sets have shown results to work well separately and if any formations of self-dimer or secondary structure would occur, it would have been a problem during the test of the individual sets as well, not only in the tests regarding the combinations of sets. A possibility of detecting all subtypes but without mixing them and by that limiting the possible interferences of one another, is to use the targeting sets in the same assay but in separately PCR wells within the plate. By that, one can detect possibly all subtypes. At the same time and only one analysis is necessary.

6 CONCLUSION

The findings reveal that the investigated parameters are critical for the method to work as optimized as possible, especially crucial is the target sequence. The target sequence must be highly specific in order to anneal proper and generate the exact product desired, both by region and variations within regions possible. Parameters such as concentrations of the primers and probe as well as salt has indicated to be important in order to achieve a most optimal method possible, nevertheless it is not crucial for the method to work. Conserved target sequences are suggested to be found within the regions of *LTR´*s as well as *Pol* of the genome of HIV-1 (Eberle and Gürtler, 2012). The observations reveal that none of the sets investigated within the promising regions, can detect all subtypes at all time individually. Regarding this, the results suggest the best potential within the primers and probe set targeting the *Pol* region of HIV-1. The data generated in the current project indicates that the *Pol* set is most promising. In order to make it proper enough to meet the standards for an implementable novel method it needs further optimization. Possibilities of subtype detection is the most prominent area to investigate further and optimize in the future. As future work, I suggest changing nucleotides which indicates of being less conservative than others within the target sequence. Investigation of the different possibilities of variations may be done with multiplex PCR assay. In order to evaluate the outcome of the changed nucleotides in comparison between each other as well as from the original sequence, suggested here. This should be conducted for both primers and the probe within the set of *Pol* in order to gain a better detection of the subtypes of M, as well as genotype O and N further on.

7 ABBREVIATION LIST

A – Adenine

AIDS – Acquired Immunodeficiency Syndrome

BHQ – Black Hole Quencher

BLAST – Basic Local Alignment Search Tool

BVDV – Bovine Viral Diarrhea Virus

C – Cytosine

cDNA – Complementary DNA

CD4 – Cluster of Differentiation 4

CP – Crossing Point

CRFs – Circulating Recombinant Forms

DNA – Deoxyribonucleic Acid

DNase – Deoxyribonuclease

Env – Gene codes for envelope proteins

FAM - Carboxyfluorescenin

FDA – Food and Drug Administration

G – Guanine

Gag - Gene codes for

GC - Guanine and Cytosine bindings

GMP – Good Manufacturing Practice

Gp41 – Glycoprotein 41

Gp120 – Glycoprotein 120

HIV – Human Immunodeficiency Virus

HPLC – High-Performance Liquid Chromatography

IBFQ – Iowa Black Quencher

IC – Internal Control

IU – International Units

LTR3′ - Long Terminal Repeat gene at 3′end

LTR5′ - Long Terminal Repeat gene at 5′end

mRNA – Messenger RNA

Nef - Gene codes for negative regulating factor

PCR – Polymerase Chain Reaction

Pol – Gene codes for DNA polymerase

Pos Q612A – Extracted positive control A

Pos Q612B - Extracted positive control B

qPCR – quantivative Polymerase Chain Reaction

Rev - Gene codes for RNA splicing regulator

RNA – Ribonucleic Acid

RNase – Ribonuclease

RT qPCR - Real-Time quantitative PCR

SIV – Simian Immunodeficiency Virus

T -Thymine

Tat - Gene codes for trans activator protein

T_a – Annealing temperature

TE – Tris EDTA

T_m – Melting temperature

Vif - Gene codes for viral infectivity factor

Vpr - Gene codes for virus protein r

WHO – World Health Organization

8 REFERENCES

Abfalterer. 2014. *HIV Sequence Compendium.* Theoretical Biology and Biophysics. New Mexico, USA.

Albert J. 2018. Interviewed by Cecilia Odlind for Medicinsk Vetenskap:4. Available at: https://ki.se/en/research/hiv-is-an-infectious-disease-like-any-other (accessed: 2020-11-20)

Avert. 2019. *HIV strains and types.* Available at: https://www.avert.org/professionals/hiv-science/types-strains (accessed: 2020-10-07)

Bailes, E. Gao, F. Bibollet-Ruche, F. Courgnaud, V. Peeters, M. Marx, P. Hahn, B. Sharp, P. 2003. *Hybrid Origin of SIV in Chimpanzees.* Science 300 (5626):1713

Bergman NH, editor. 2007. *Comparative genomics: Volumes 1 and 2.* Totowa (NJ): Humana Press, chapter 9.

BioMéreux. 2009. *NucleiSens – easyMAG USER MANUAL.* BioMérieux S.A. Lyon, France.

Boisselier. 2012. *Genotyping using digital PCR with the Light Cycler 1536 Real-Time PCR System to detect the IDH1 mutation in plasma samples.* Roche.

Briggs, J. Simon, M. Gross, I. Kräusslich, H-G. Fuller, S. Vogt, V. Johnson, M. 2004. *The stoichiometry of Gag protein in HIV-1.* Nature Structural & Molecular Biology. 11:672-675.

Brownie, J. Shawcross, S. Theaker, J. Whitcombe, D. Ferrie, R. Newton, C. Little, S. 1997. *The elimination of primer-dimer accumulation in PCR.* Nucleic Acids Research. 25:16: 3235-3241.

Butler J. 2012. *Advanced Topics in Forensic DNA Typing: Methodology.* Academic press. London.

Cardo, D. Culer, D. Ciesielski, C. Srivastava, P. Marcus, R. Abiteboul, D. Patch, M. Ippolito, G. Lot, F. McKibben, P. Bell, D. 1997. *A Case-Control Study of HIV Seroconversion in Health Care Workers after Percutaneous Exposure.* The New England Journal of Medicine. 337:1485-1490.

Cheng, S. Chen, Y. Monforte, J. Higuchi, R. Van Houten, B. 1995. *Template integrity is essential for PCR amplification 20-30kb for genomic DNA.*PCR Methods Appl. 4:294-298.

Cuevas, J. Geller, R. Garjio, R. López-Aldeguer, J. Sanjuán, R. 2015. *Extremely High Mutation Rate of HIV-1 in vivo.* PLoS Biology. 13:9.

Das, S. Mohapatra, S. Hsu, J. 1999. *Studies on primer-dimer formation in polymerase chain reaction (PCR).* Biotechnology Techniques. 13 :643-646.

Datta, V. LiCata, V. 2002. *Salt Dependence of DNA binding by Thermus aquaticus and Escherichia coli DNA Polymerase.* Journal of Biology Chemistry. 278:8:5694-5701.

Drobniewski, F. Nikolayevskyy, V. Surkova, E. 2020. *False-positive COVID-19 results: hidden problems and costs.* The Lancet. 8:2:1167-1168.

Drosten, C. Panning, M. Drexler, J. Hänsel, F. Pedroso, C. Yeats, J. Kleber de Souza Luna, L. Samuel, M. Liedigk, B. Lippert, U. Stürmer, M. Doerr, H W. Brites, C. Preiser, W. 2006. *Ultrasensitive Monitoring of HIV-1 Viral Load by a Low-Cost Real-Time Reverse Transcription-PCR Assay with Internal Control for the 5´Long Terminal Repeat Domain.* Clinical Chemistry 52:7:1258-1266.

Eberle, J. Gürtler, L. 2013. *HIV Types, Groups, Subtypes and Recombinant Forms: Errors in Replication, Selection Pressure and Quasispecies.* Intervirology. 55:79-83.

Food and Drug Administration (FDA). 2010. *Guidance for Industry – Nucleic Acid Testing (NAT) for Human Immunodeficiency Virus Type 1 (HIV-1) and Hepatitis C Virus (HCV): Testing, Production Disposition, and Donor Deferral and Reentry.* US Department of Health and Human Services: Center of Biologics Evolutions and Research.

Flint S. 2000. *Principles of Virology, Molecular Biology, Pathogenesis and Control.* ASM Press, Washington D.C. 8: 630-660.

Freed E. 2001. *HIV-1 Replication.* Somatic Cell and Molecular Genetics. 26:13-33.

Gao, F. Bailes, E. Roberson, D. Chen, Y. Rodenburg, C M. Michael, S F. Cummins, L B. Arthur, L O. Peeters, M. Sharp, P M. Hahn, B H. 1999. *Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes.* Nature 397 (6718):436-441.

Green, MR. Sambrook, J. 2018. *Optimizing Primer and Probe Concentrations for Use in Real-Time Polymerase Chain Reaction (PCR) Assays.* Cold Spring Harbor Laboratory press.

Haase A. 1986. *Pathogenesis of lentivirus infections.* Nature. 322 :130-136.

Haleem, R. Salem, M Y. Fatahallah, F A. Abdelfattag, L E. 2015. *Quality in the pharmaceutical industry – A literature review.* Saudi Pharmaceutical Journal. 23:5:463-469.

Heid, C A. Stevens, J. Livak, K J. Williams, P M. 1996. *Real Time Quantitative PCR.* Genome Methods. 6: 986-994.

HIV Virus: Structure and Function. Study.com, 21 July 2016, https://study.com/academy/lesson/hiv-virus-structure-and-function.html

Hemelaar J. 2012. *The origin and diversity of the HIV-1 pandemic.* Trends in Molecular Medicine 18:3:182-192.

Hollingsworth, D. Anderson, R M. Fraser, C. 2008. *HIV-1 Transmission, by Stage of Infection.* Journal of Infectious Diseases. 198:5:687-693.

Integrated DNA Technologies. n.d. *Iowa Black® Dark Quenchers Modifications.* Available at: https://eu.idtdna.com/site/Catalog/Modifications/Category/4 (accessed: 2020-10-13)

Integrated DNA Technologies. n.d. *Oligonucleotides Stability Study.* Available at: https://sfvideo.blob.core.windows.net/sitefinity/docs/default-source/technical-report/stability-of-oligos.pdf?sfvrsn=c6483407_10 (accessed: 2020-12-03)

Jansen, R R. Schinkel, J. Koekkoek, S. Pajkrt, D. Beld, M. de Jong, M D. Molenkamp, R. 2011. *Development and evaluation of a four-tube real time multiples PCR assay covering fourteen respiratory viruses, and comparison to its corresponding single target counterparts.* Journal of Clinical Virology. 51:3:179-185.

Jolles, S. Kinloch De Loes, S. Johnson, M A. Janossy, G. 1996. *Primary HIV-1 infection: a new medical emergency? Recognition of this initial illness may permit early diagnosis and treatment.* British medical Journal. 312:7041:1243-1244.

Kwong, p D. Wyatt, r. Robinson, J. Sweet, R W. Sodroski, J. Hendrickson, W A. 1998. *Structure of an HIV gp120 envelope glycoprotein in complex with the CD¤ receptor and a neutralizing human antibody.* Nature. 393:6686:648.

Lavezzo, E. Perrone, R. Palù, G. Richter, S N. 2017. *Conserved presence of G-quadruplex forming sequences in Long Terminal Repeat Promoter of Lentiviruses.* Sci Rep 7, 2018 (2017).

LGC Biosearch Technologies. 2020. *FAQ: What is the difference between Black Hole Quencher (BHQ) dye and TAMRA?* LGC Biosearch Technologies. Available at: https://www.biosearchtech.com/support/faqs/fluorogenic-probe-and-primers/what-is-the-difference-between-black-hole-quencher-bhq-dye-and-tamra (accessed. 2020-10-13)

Li, R. Weisel, R. 2014. *Cardiac Regeneration and Repair.* Cambridge: Woodhead Publishing.

Li, Y. Hui, H. Burgess, C J. Price R W. Sharp, P M. Hahn, B H. Shaw, G M. 1992. *Complete Nucleotide Sequence, Genome Organization, and Biological Properties of Human Immunodeficiency Virus Type 1 In Vivo: Evidence for Limited Defectiveness and Complementation.* Journal of Virology. 66:11:6587-6600.

Lopez-Vergés, S. Camus, G. Blot, G. Beauvoir, R. Benarous, R. Berlioz-Torrent, C. 2006. *Tail-interacting protein TIP47 is a connector between Hah and Env and is required for Env incorporation into HIV-1 virions.* PNAS. 103:40:14947-14952.

Luft, L M. Gill, M J. Church, D L. 2011. *HIV-1 viral diversity and its implications for viral load testing: review of current platforms.* International Journal of Infectious Diseases. 15:661-670.

Lusic, M. Siliciano, R. 2017. *Nuclear landscape of HIV-1 infection and integration.* Nature rev Microbiology. 15:69–82.

Lütgenhetmann, M. Nörz, D. Reucher, S. Pfefferle, S. 2020. *Evaluation of a quantitative RT-PCR assay for the detection of the emerging coronavirus SARS-CoV-2 using a high throughput system.* Eurosurveillance. 25:9.

Manning, M. Chou, D K. Murphy, B M. Payne, R W. Katayama, D S. 2010. *Stability of Protein Pharmaceuticals: An update.* Pharmaceutical Research. 27:544-575.

Murphy, K. Weaver, C. 2017. *Janeway´s Immunobiology.* 9th edition. Garland Science, Taylor & Francis Group. USA.

O´Brien, T. 1994. *HIV-2 Transmission: Implications for Spread of HIV-1.* JAMA. 271:12:903–904.

Octapharma Nordic AB. 2020. *Plasmabaserad tillverkning*. Available at: https://www.octapharma.se/produkter/plasmabaserad-tillverkning/ (accessed: 2020-10-02)

Osmanov. S et al. 2002. *Estimated Global Distribution and Regional Spread of HIV-1 Gentetc Subtypes in the Year 2000.* Journal of Acquired Immune Deficiency Syndrom. 29:2:184-190.

Pebody, R. 2009. *HIV-1 Subtypes.* Aidsmap. England. Available at: https://www.aidsmap.com/about-hiv/hiv-1-subtypes (accessed: 2020-10-19)

Petti, C A. Polage, C R. Quinn, T C. Ronald, A R. Sande, M A. 2006. *Laboratory Medicine in Africa: A barrier to Effective Health Care.* Clinical Infectious Diseases. 42:3:377-382.

Prediger, E. 2013. *How to design primer and probes for PCR and qPCR.* Integrated DNA Technologies, USA. Available at: https://eu.idtdna.com/pages/education/decoded/article/designing-pcr-primers-and-probes (accessed: 2020-10-15)

Prescott, G. Hockley, J. Atkinson, E. Rigsby, P. Morris, C. 2017. *International Collaborative Study to Establish the 4th WHO International Standard for HIV-1 NAT Assays.* Expert committee on biological standardization. World Health Organization. Geneva.

Proudnikov, D. Yuferov, V. Zhou, Y. LaForge, K S. Ho, A. Kreek, M J. 2003. *Optimizing primer-probe design for fluorescent PCR.* Journal of Neuroscience Methods. 123:1:31-45.

Ren, C. Xiong, W. Peng, D. He, Y. Zhou, P. Li, J. Li, B. 2018. *Effects of thermal sterilization on soy protein isolate/polyphenol complexes: Aspects of structure, in vitro digestibility and antioxidant activity.* Food Research International. 112:284-290.

Richetta, C. Thierry, S. Thierry, E. Lesbats, P. Lapaillerie, D. Munir, S. Subra, F. Leh, H. Deprez, E. Parissi, V. Delelis, O. 2019. *Two-long terminal repeat (LTR) DNA circles are a substrate for HIV-1 integrase.* Journal of Biological Chemistry. 294:8286-8295.

Robertson, D L. Anderson, J P. bradac, J A. Carr, J K. Foley, B. Funkhouser, R K. Gao, F. Hahn, B H. Kuiken, C. Learn, G H. Leitner, T. McCutchan, F. Osmanov. S. peeter, M. Pieniazek, D. Kalish, M L. Salminen, M. Sharp, P. Wolinsky, S. Korber, B. 2000. *HIV-1 nomenclature proposal.* Science. 288 :5463 :55-55.

Rychlik, W. Spencer, W J. Rhoads, R E. 1990. *Optimization of the annealing temperature for DNA amplification in vitro.* Nucleic Acids res 18(21):6409-6412.

Seitz, R. 2016. *Human Immunodeficiency Virus (HIV).* Transfusion Med Hemother. 43: 203–222.

Shang, L. Rabi, S A. Laird, G M. Eisele, E E. Zhang, H. Magrolick, J B. Siliciano, R F. 2013. *A Novel PCR Assay for Quantification of HIV-1 RNA.* Journal of Virology. 86:11:6521-6525.

Simon, V. Ho, D D. Karim Q A. 2006. *HIV/AIDS epidemiology, pathogenesis, prevention, and treatment.* The Lancet. 368:9534:489-504.

Sipos, R. Székely, A J. Palatinszky, M. Révész, S. Márialigeti, K. Nikolsusz, M. 2006. *Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targeting bacterial community analysis.* Federation of European Microbiological Societies. 60:341-350.

Steitz, TA. 1998. *A mechanism for all polymerases.* Nature. 391(6664) :231-232.

Strain, M C. Lada, S M. Luong, T. Rought, S E. Gianella, S. Terry, V H. Spuna C A. Woelk, C H. Richman, D D. 2013. *Highly Precise Measurement of HIV DNA by Droplet Digital PCR.* PLoS ONE 8:4.

Stults, J. Snoeyenbos-West, O. Methe, B. Lovley, D R. Chandler, D P. 2001. *Application of the 5′Fluorogenic Exonuclease Assay (TaqMan) for Quantitative Ribosomal DNA and rRNA Analysis in Sediments.* Applied and Environmental Microbiology. 67:6:2781-2789.

Taylor, B S. Sobieszczyk, M E. McCutchan, F E. Hammer, S M. 2008. *The challenge of HIV-1 Subtype Diversity.* New England Journal of Medicine. 358:1590-1602.

Taylor, S. Wakem, M. Dijkman, G. Alsarraj, M. Nguyen, M. 2010. *A practical approach to RT-qPCR-Publishing data that conform the MIQE guidelines.* Methods. 50:4:1-5.

The AIDS Institute. n.d. *where did HIV come from?* The AIDS Institute. Available at: https://www.theaidsinstitute.org/education/aids-101/where-did-hiv-come-0 (accessed: 2020-10-06)

ThermoFisher (A). n.d. *How TaqMan Assays Work.* Available at: https://www.thermofisher.com/se/en/home/life-science/pcr/real-time-pcr/real-time-pcr-learning-center/real-time-pcr-basics/how-taqman-assays-work.html (accessed: 2020-10-09)

ThermoFisher (B). n.d. *Multiple Primer Analyzer*. Available at: https://www.thermofisher.com/se/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/multiple-primer-analyzer.html (accessed: 2020-11-09)

ThermoFisher (C). n.d. *The Basics: RNA isolation.* Available at: https://www.thermofisher.com/se/en/home/references/ambion-tech-support/rna-isolation/general-articles/the-basics-rna-isolation.html (accessed: 2020-10-28)

Toes, R. Ossendorp, F. Offringa R. Melief, C. 1999. *CD4 T Cells and Their Role in Antitumor Immune Response.* Journal of Experimental Medicine. 189:5:753-756.

UNAIDS. 2020. *90-90-90: An ambitious treatment target to help end to AIDS pandemic.* Available at: https://www.unaids.org/en/resources/909090 (accessed: 2020-10-02)

Vanden Haesevelde, M. Peeters, M. Jannies, G. Janssens, W. Van Der Groen, G. Sharp, P. Saman, E. 1996. *Sequence Analysis of a Highly Divergent HIV-1-Related Lentivirus Isolated from a Wild Captured Chimpanzee.* Virology. 211:346-350.

Vandesompele, J. n.d. *qPCR guide.* Available at : https://www.gene-quantification.de/eurogentec-qPCR-guide.pdf (accessed: 2020-10-09)

Vivès, R. Imberty, A. Sattentau, Q J. Lortat-Jacob, H. 2005. *Heparan Sulfate targets the HIV-1 Envelope Glycoprotein gp120 Coreceptor Binding Site.* Journal of Biological Chemistry. 280:21353-21357.

Watts, J. Dang. K K. Gorelick, R J. Leonard, C W. Bess, J W. Swanstrom, R. Burch, C L. Weeks, K M. 2009. *Architecture and secondary structure of an entire HIV-1 RNA genome.* Nature. 460:6.

Weber, J. Rangel, H R. Chakraborty, B. Marotta, M L. Valdez, H. Fransen, K. Florence, E. Connick, E. Smith, K Y. Colebunders, R L. Landay, A. Kuritzkes, D R. Lederman, M M. Vanham, G. Quinones-Mateu, M. 2003. *Role of Baseline pol Genotype in HIV-1 Fitness Evolution.* Journal of Acquired Immune Deficiency Syndromes. 33:448-460.

World Health Organization. 2005. *Interim WHO clinical staging of HIV/AIDS and HIV/AIDS case definitions for surveillance.* 2005:2

World Health Organization. 2020. *HIV/AIDS.* World Health Organization. Available at: https://www.who.int/news-room/fact-sheets/detail/hiv-aids (accessed: 2020-10-02)

Worobey, M. gemmel, M. Teuwn, D E. Haselkron, T. Kunstman, K. Bunce, M. Muyembe, J-J. Kabongo, J-M M. Kalengayi, R M. Van March, E. Gilbert, T P. Wolinsky, S M. 2008. *Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960.* Nature. 455:661-664.

Worobey, M. Watts, T D. McKay, W R. Suchard, M A. Granade, T. Teuwen, D E. Koblin, B A. Heneine, W. Lemey, P. Jaffe, H W. 2016. *1970s and 'Patient 0' HIV-1 genomes illuminate early HIV/AIDS history in North America.* Nature. 539:98-101.

Yahaya, L. Jimoh, A A G. Balogun, O R. 2010. *Factors Hindering Acceptance of HIV/AIDS Voluntary Counseling and Testing (VCT) among Youth in Kwara State, Nigeria.* African Journal of Reproductive Health. 14:3.

Yang, Y-C. Yang-Chih Shih, D. Tsai, M-H. Cheng, C-H. Cheng, H-F. Lo, C-F. Wang, D-Y. 2013. *A collaborative study to establish the first National Standard for HIV-1 RNA nucleic acid amplification techniques (NAT) in Taiwan.* Journal of Virological Methods. 191:122-127.

Zeringer, E. Li, M. Barta, T. Schageman, J. Winther Pedersen, K. Neurauter, A. Magdaleno, S. Setterquist, R. Vlassov, A V. 2013. *Methods for the extraction and RNA profiling of exosomes.* World Journal of Methodology. 3:1:11-18.