EVALUATION OF PRIMERS FOR THE DETECTION OF HIV-1 VIRUSES FROM KELANTAN BY USING REVERSE TRANSCRIPTION POLYMERASE CHAIN REACTION (RT-PCR)

by

NOOR HAMIZAH BINTI MINAL

Dissertation submitted in partial fulfillment of the requirements for the degree of

Bachelor of Health Sciences (Biomedicine)

May 2013

CERTIFICATE

This is to certify that the dissertation entitle "EVALUATION OF THE PRIMERS FOR THE DETECTION OF HIV-1 VIRUSES FROM KELANTAN BY USING REVERSE TRANSCRIPTION POLYMERASE CHAIN REACTION (RT-PCR)" is the bonafide record of research work done by Ms. Noor Hamizah binti Minal during the period from September 2012 to June 2013 under my supervision.

Supervisor,

....

Dr. Maizan Mohamed, PhD
Researcher & Lecturer
Institute for Research in Molecular Medicine (INFORMM)
Universiti Sains Malaysia
Health Campus
16150 Kubang Kerian
Kelantan, Malaysia
Date: 12/6/2013

ACKNOWLEDGEMENT

Praise to God for blessing me with patience and perseverance to conduct this research well.

I would like to take this wonderful opportunity to express my thanks and deepest gratitude for my supervisor, Dr. Maizan bt Haji Mohamed for all her guidance, valuable advice, constructive criticism and scientific discussion. I also would like to thank Miss Che Wan Salma Che Wan Zalati, Miss Eliana Saffie, Miss Julia Abdullah and Miss Fadhilah Usuki, who have taught and assisted me in molecular work and also intellectual and interesting conversation that we had together during this period. Not to forget, Dr. Wan Nazirah Wan Yusof for providing the samples.

I also want to express my deepest gratitude to members of INFORMM Health Campus for their great support and cooperation. I also would like to thank Prof. Norazmi Mohd Nor, Director of INFORMM for accepting me to complete my final year project at INFORMM. Also thanks to INFORMM administration and laboratory staff for their cooperation in allowing me to use their equipment and performing lab work at INFORMM.

Not to forget, I would like to express my heartful thanks to my beloved family and friend, especially my parents for their blessings, wishes, and encouragements for the successful of this final year project.

Last but not least, thanks to all of those who have supported me in any aspects during the completion of this dissertation especially all my classmates.

TABLE OF CONTENTS

Acknowledgement i
Table of Contents ii
List of Tables iv
List of Figures v
Symbols and Abbreviations vi
Abstract vii

CHAPTER 1: INTRODUCTION 1
1.1 Acquired Immunodeficiency Syndrome (AIDS) 1
1.2 Problem & Rationale of The Study 4
1.3 Objectives 6
CHAPTER 2: LITERATURE REVIEW
2.1 Etiology – HIV 7
2.1.1 Classification - HIV 8
2.1.2 Properties and Structure 9
2.1.3 Long Terminal Repeat (LTR) region 12
2.2 Replication Cycle 15
2.2.1 HIV-Entry 15
2.2.2 HIV-1 Reverse Transcription 17
2.2.3 HIV-1 DNA Integration 20
2.2.4 HIV-1 Assembly, Budding and Maturation 20
2.3 Diagnosis 23
2.3.1 Enzyme Immunoassay (EIA) 23
2.3.2 Molecular Diagnosis 27
CHAPTER 3: MATERIALS AND METHODS
3.1 Materials 29
3.2 Methodology 32
3.2.1 Conceptual Frame Work 32
3.2.2 Samples and Data Collection 33
3.2.3 Primers 35
3.2.4 Viral Extraction 36

3.2.5 Complimentary DNA (cDNA) synthesis 36
3.2.6 Polymerase Chain Reaction (PCR) 37
3.2.7 Agarose Gel Electrophoresis 37
3.2.8 Real-Time PCR 38
CHAPTER 4: RESULT
4.1 Primer Design 39
4.2 Primer Evaluation 45
4.3 Optimization of PCR Parameters 49
4.3.1 Optimization of Annealing Temperature 49
4.3.2 Optimization of $MgCl_2$ 49
4.3.3 Optimization of dNTP and *Taq* DNA polymerase 50
4.4 Evaluation of Sensitivity and Specificity of Primers 50
4.5 Detection of HIV-1 viruses in 30 Clinical Samples from Kelantan 56
4.6 Evaluation of HIV LTRF2&HIV LTR2 primers for Real-Time PCR 57

CHAPTER 5: DISCUSSION
5.1 Primer Design 59
5.2 Primers Evaluation 61
5.3 Optimization of PCR Parameters 61
5.3.1 Optimization of Annealing Temperature 62
5.3.2 Optimization of $MgCl_2$ 62
5.3.3 Optimization of dNTP and *Taq* DNA Polymerase 63
5.4 Evaluation Sensitivity and Specificity of Primers 64
5.5 Detection of HIV-1 Virus in Clinical Samples 64
5.6 Evaluation of Primers for Real-Time PCR (rtPCR) 65
CHAPTER 6: CONCLUSION 66
REFERENCES 67

LIST OF TABLES

Table 3.1	List of chemicals and reagents used in this study	29
Table 3.2	List of commercial kit used in this study	30
Table 3.3	List of equipments used in this study	30
Table 3.4	List of bioinformatic tools used in this study	31
Table 3.5	List of consumable items used in this study	31
Table 3.6	List of HIV-1 samples used in this study	34
Table 3.7	Primers used for amplification of HIV-1 viruses	35
Table 4.1	Result of Ct values of real-time PCR on 6 HIV-1 viruses from Kelantan using HIV LTRF2 and HIV LTR2 as primers.	58

LIST OF FIGURES

Figure 2.1	Illustrated basic structure of human immunodeficiency virus (HIV)	11
Figure 2.2	HIV-1 Long terminal repeat (LTR) structure of proviral DNA.	14
Figure 2.3	Overview of HIV entry.	19
Figure 2.4	Overview of HIV-1 replication cycle.	22
Figure 2.5	The generation of immunoassay in HIV diagnostic.	24
Figure 4.1	Result from BLAST analysis for HIV LTRF2	39
Figure 4.2	Result from BLAST analysis for HIV LTR2	40
Figure 4.3	Result from BLAST analysis for HIV ENV1F	41
Figure 4.4	Result from BLAST analysis for HIV ENVRNew	42
Figure 4.5	Result from BLAST analysis for HIV DRF	43
Figure 4.6	Result from BLAST analysis for HIV RTR2N	44
Figure 4.7	Comparison in primers evaluation between HIV DRF+HIV RTR2N and HIV LTRF2+HIV LTR2	46
Figure 4.8	Primers evaluation in PCR reaction using HIV ENV1F+HIV ENVRNew	47
Figure 4.9	Optimization of annealing temperature for PCR reaction	50
Figure 4.10	Optimization of magnesium concentration for PCR reaction	51
Figure 4.11	Optimization of dNTP and *Taq* DNA polymerase concentration for PCR reaction	52
Figure 4.12	Sensitivity evaluation of the primers by a 10-fold serial dilution.	54
Figure 4.13	Specificity evaluation of the primers on the three RNA of viruses	55
Figure 4.14	Result from PCR reaction in detecting HIV-1 viruses from Kelantan	56
Figure 4.15	A dissociation curve of real-time PCR	57

SYMBOLS AND ABBREVIATIONS

°C	Degree Celcius
µg	Microgram
µL	Microlitre
bp	Base pair
BLAST	Basic Local Alignment Search Tool
cDNA	Complementary deoxyribonucleic acid
Ct	Cycle threshold
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	deoxynucleoside Triphosphate
EDTA	Ethylene Diamine Tetra Acetic acid
EtBr	Ethidium bromide
g	Gravity
HIV	Human immunodeficiency virus
kDa	kiloDalton
LTR	Long terminal repeat
mg	Miligram
$MgCl_2$	Magnesium Chloride
mM	miliMolar
NFW	Nuclease Free Water
ng	Nanogram
PCR	Polymerase Chain Reaction
qPCR	Qualitative Polymerase chain reaction
RNA	Ribonucleic acid
RNase	Ribonuclease
Rpm	Revolution per minute
RT	Reverse Transcriptase
RT-PCR	Reverse Transcription-Polymerase chain reaction
TBE	Tris Borate EDTA
U	Unit
UV	Ultraviolet
ρM	Picomolar

ABSTRACT

Human immunodeficiency virus type 1 (HIV-1) has been recognized as the causative agent of acquired immunodeficiency syndrome (AIDS). Approximately 42 million people carrying the virus at present, but its case fatality rate is close to 100%, making it an infection of devastating ferocity. Since the evolutionary change of the virus is at rapid rate, it imposes a challenge on HIV diagnostic development, vaccine development, antiretroviral drug sensitivity and drug resistance. Rapid and sensitive methods for the detection of HIV-1 viruses would be valuable in controlling this disease. Currently, nucleic acid amplification tests such as reverse transcription polymerase chain reaction (RT-PCR) and real-time PCR (rtPCR) are widely used in detecting HIV-1 viruses. These tests are commonly applied in blood donation screening and early detection of HIV-1 in infants. High sequence variations of the HIV-1 viruses from Kelantan have led to the failure to obtain the amplification product using published primers. Therefore, in this study, a few set of primers were designed based on the sequence alignments of published HIV-1 viruses from Malaysia and Kelantan which were retrieved from NCBI genebank. The primers were evaluated on 30 confirmed HIV-1 positive samples from Kelantan, Malaysia. HIV LTRF2 and HIV LTR2 were chosen as the best primers to detect these viruses since they were successfully amplified all the 30 samples used in this study. The primers were further evaluated their sensitivity and specificity against 3 viruses which are Japanese encephalitis (JE), Chikungunya and Western encephalitis virus (WEE) where no expected band yielded. These primers were sensitive and specific enough in detecting HIV-1 viruses from Kelantan. They were further evaluated for use in the real-time PCR and can be used in this test. Thus, these primers could be potentially use for the future diagnosis of HIV-1 viruses from Kelantan either by conventional or real time PCR.

ABSTRAK

"Human Immunodeficiecy Virus" (HIV) jenis 1 telah dikenal pasti sebagai agen penyebab kepada Sindrom Kurang Daya Tahan Penyakit (AIDS). Sehingga kini dianggarkan hampir 42 juta orang pembawa virus ini, dengan kadar kematiannya hampir kepada 100%, menjadikannya sebagai satu jangkitan yang boleh membunuh. Memandangkan perubahan evolusi virus tersebut pada kadar yang sangat cepat, ini menimbulkan satu cabaran dalam pembangunan diagnostik dan vaksin, sensitiviti ubat antiretroviral dan ketahanan terhadap ubat. Ujian amplifikasi asid nukleik merupakan pendekatan yang digunakan secara meluas dalam mengesan virus HIV-1. Pendekatan ini juga banyak dilaksanakan untuk memeriksa darah yang didermakan dan pengesanan awal HIV-1 dalam bayi. Variasi yang tinggi dalam jujukan virus HIV-1 yang terdapat di Kelantan telah menyebabkan virus tersebut tidak dapat dikesan dalam proses PCR dengan menggunakan primer rujukan yang dihasilkan oleh kajian sebelum ini. Justeru, beberapa set primer telah direka dan dinilai keatas 30 sampel positif HIV-1 dari Kelantan. Dalam kajian ini, HIV LTRF2 dan HIV LTR2 telah dipilih sebagai primer yang terbaik untuk mengesan virus HIV-1 yang terdapat di Kelantan memandangkan primer-primer ini berjaya mengamplifikasikan kesemua 30 sampel yang digunakan dalam kajian ini. Sensitiviti dan spesifisiti primer ini kemudian dinilai keatas tiga jenis virus lain iaitu Japanese encephalitis (JE), Chikungunya dan 'Western equine encephalitis virus' (WEE) di mana tiada amplifikasi yang berlaku. Primer ini didapati sensitif dan spesifik dalam pengesanan virus HIV dari Kelantan. Primer ini seterusnya diuji sama ada ia boleh digunakan dalam 'real-time' PCR. Hasil yang diperolehi menunjukkan primer-primer ini boleh digunakan dalam 'real-time' PCR. Oleh itu, primer ini disahkan boleh untuk digunakan untuk diagnosa virus HIV dari Kelantan pada masa akan datang samada secara konvensional PCR atau 'real-time' PCR.

Chapter 1: INTRODUCTION

1.1 Acquired Immunodeficiency Syndrome (AIDS)

The acquired immunodeficiency syndrome or commonly known as AIDS is a disease caused by the human immunodeficiency virus type 1 (HIV-1). AIDS is defined by a depletion of CD4 T lymphocytes (< 200cells/ µL) and high susceptibility of opportunistic infections (Champoux and Lawrence, 2004). According to International HIV & AIDS Charity (AVERT), AIDS was first recognized in the United States in the early of 1980s, where it was found among gay and most of them were injection drug user (IDU). It is followed by an increase in the incidence commonly due to rare opportunistic infections such as *Pneumonia* in homosexual men (Rambaut *et al.*, 2004). AIDS is commonly transmitted through sexual activities, including homo- and heterosexual (Castro-Nallar *et al.*, 2012), mother to child (Fischetti *et al.*, 2005) and needle sharing among injecting drug users (IDUs) (Thomson, 2009).

Human immunodeficiency virus type 1 (HIV-1) is a type of retrovirus that has been recognized as the causative agent for this devastating infectious disease (Barre-Sinoussi *et al.*, 1983; Gallo *et al.*, 1984; Popovic *et al.*, 1984). According to 2012 Global Report of the Joint United Nations Program on HIV/AIDS, about 34.0 million people were living with HIV at the end of 2011. An estimated 0.8% of adults aged between 15-49 years worldwide are living with HIV. Sub-Saharan Africa remains the most severely affected region with nearly 1 in every 20 adults (4.9%) living with HIV and accounting for 69% of people with HIV worldwide. About 1.6 million people are

estimated to be living with HIV in Latin America and the Caribbean combined, including 96,000 newly infected in 2011. Caribbean with the adult HIV prevalence rate of 1% is the second hardest hit region in the world after sub-Saharan Africa. More than 5 million people are living with HIV in South, South-East and East Asia combined. Malaysia is home to one of the fastest growing AIDS epidemics in the East Asia and Pacific region. According to Ministry of Health, Malaysia, since the first case was detected in Malaysia in 1986 until 2010, 91,362 men, women and children have been recognized living with HIV while 12, 943 have died from AIDS (Ministry of Health Malaysia., 2010).

HIV infection/AIDS is a complex vulnerable disease since it appeared, the outcome is invariably fatal and there is no vaccine available to treat this disease to date. It is important to understand that laboratory diagnostic is the only way to establish the status of HIV infection of the patients. HIV screening is important in order to identify the asymptomatic as well as symptomatic HIV patients. It is also required for blood transfusion to assure that blood and any blood-related product are safe to be transfused. Furthermore, by screening for HIV, it will prevent further transmission from HIV patients by giving counseling to them. Many countries used this step as a guideline to monitor the HIV epidemic in their countries (Pant Pai *et al.*, 2013).

Current routine laboratory diagnosis is mainly based on the detection of specific anti-HIV antibodies. Whole blood, serum or plasma is the preferred specimen for the detection of anti-HIV antibodies. Those antibodies usually can be detected from 3 to 6 weeks after infection. However, antibodies would not be detected within window period even though the infected person is highly infectious. Window period refers to the period

between the onset of HIV infection and the appearance of detectable antibodies to the virus. The window period is variable in individuals. However, undetectable antibodies more than three months after infections are rare (Gaines *et al.*, 1987; Horsburgh Jr *et al.*, 1989; Yılmaz, 2001).

Enzyme immunoassay and agglutination test are the most widely used basic screening serological method. However, all reactive results should be confirmed with Western blot or immunoblot. Since the specificity and sensitivity of the screening test is limited, two classical methods were used for direct detection of viral infection; p24 viral antigen detection or HIV culture. Nowadays, both classical methods have been replaced with Polymerase Chain Reaction (PCR) and other amplification methods for direct detection of HIV infection. The viral load of HIV, which is measured by quantitative amplification methods will be used as prognostic marker of disease progression, for monitoring antiretroviral treatment and for diagnosis of acute HIV infection (before seroconversion) (S. Luft, 2004).

1.2 Problems and Rationale of Study

Variation of HIV-1 viruses create challenges in the development of diagnostic, virulence, emergence of drug resistance and vaccine development (Rambaut *et al.*, 2004). Despite the global effort to control the AIDS pandemic, HIV infections continue to spread relatively unabated in many parts of the world. While some countries have showed a significant success in reducing the HIV infections, others still find ways to identify infected person, control the transmission and implement effective treatment. Enzyme immunoassay and Western Blot are widely accepted as a 'gold standard' in detecting HIV infections. However, it is not helpful on infants due to the persistence of maternal antibodies for up to first 15 months of life (Sohn *et al.*, 2009). Thus, uninfected-infants could not be distinguished as they also have anti-HIV antibodies.

In order to determine the HIV status during window period and to monitor the progression of HIV-infected individuals included on infants, a direct detection of HIV is required. Different amplification methods are the most widely used approach for this purpose (Yılmaz, 2001). The discovery of a persistently high level of viral turnover in all stages of disease, even during the period of clinical latency, suggests that sensitive measurement of viral load may be useful in assessing HIV disease stage and progression, response to treatment and risk of transmission (Cassol *et al.*, 1994; Brown *et al.*, 1997; Yılmaz, 2001). PCR and Real-Time PCR can give a positive result in the earlier infection (Daar *et al.*, 2001).

Real-Time PCR is a quantitative assay to determine the viral load of the HIV patient. It is used in order to know the stage of infection. However, currently used viral load assay is based on a commercial kit without revealing the primer sequences and is

very expensive. Furthermore, the primers established based on HIV-1 sequences of manufacturer's country and cannot be used for Malaysian HIV viruses due to high sequence variations.

Currently we are having problem with published primers for detecting all the Kelantan HIV viruses due to high variation of the genes. Therefore, in this study, we are going to design and evaluate a few set of primers using RT-PCR for the specific detection of HIV viruses from Kelantan. These primers will be further use in Real-Time PCR for the viral load assay.

1.3 Objectives

The objectives of this study are:

1. To design few set of primers for the detection HIV-1 viruses from Kelantan by RT-PCR.
2. To optimize the PCR parameters for the detection of HIV-1 viruses.
3. To evaluate the sensitivity and specificity of the primers.
4. To evaluate the primers for the detection HIV-1 sample from Kelantan by using RT-PCR.
5. To further evaluate these primers for use in real time-PCR

Chapter 2: LITERATURE REVIEW

2.1 Etiology - HIV

Human immunodeficiency virus (HIV) is known as an etiological agent for acquired immunodeficicency syndrome (AIDS), a condition in which the progression of this disease can cause failure to the immune system which allows life-threatening opportunistic infection (Weiss, 1993 and Douek *et al.*, 2009). HIV is a Lentivirus, a member of the *retroviridae* family. It is characterized by a long incubation period. Lentiviruses can deliver a significant amount of viral RNA into the DNA of the host cell and have the unique ability among retroviruses of being able to infect non-dividing cells. They are different than other *retroviridae* members due to their unusual morphological characteristics such as a cylindrical or cone-shaped nucleoid in the mature virion. They also differed by several genes that are absent in other retroviral genomes such as *Tat* and *Rev* (Freed and Mouland, 2006).

There are two types of HIV, HIV-1 and HIV-2. HIV-1 is more virulent and caused the majority of HIV infections globally (Shankarappa *et al.*, 1999). In 1986, HIV-2 was isolated from AIDS patients in West Africa. However, HIV-2 is less virulent and is confined largely to countries in West Africa such as Senegal, Liberia, Ghana and the Ivory Coast (De Cock Km and et al., 1993; Olaleye *et al.*, 1993; Ankrah *et al.*, 1994). Both HIV-1 and HIV-2 are the causative agents with the same modes of transmission and shared many biologic and virologic characteristics (Kanki *et al.*, 2002).

Due to its modest genomic size which is less than 10kb (Simon *et al.*, 2006) with a few genes, HIV-1 is able in taking advantages of cellular pathways while preventing itself from being recognised by the components of immune system (Barré-Sinoussi, 1996; Emerman and Malim, 1998).

2.1.1 Classification of HIV-1 Variants

The phenotype of HIV-1 is defined by the cells that undergo *in vitro* replication. However, these phenotypes can change *in vivo* with profound implications for viral transmission, pathogenesis and disease progression (Berger *et al.*, 1998). Based on their genetic make-up, HIV-1 viruses are classified into three groups which are M (Main), O (Outlier) and N (New or non M/ non O) group (Shankarappa *et al.*, 1998).

The HIV-1 group M is responsible for the most of the global pandemic. The diversity present within the group is extensive when compared to other rapidly evolving viral genomics such as influenza (Korber *et al.*, 2001). From a phylogenetic tree strains, group M strains form another nine clusters, which are A, B, C, D, F, G, H, J, and K (Robertson *et al.*, 2000). They are consistent in their phylogenetic topology in relation to each other regardless of the section in their genome (Archer and Robertson, 2007). Those subtypes are separated from each other by equivalent genetic distances and differ in their geographic distribution. Subtype B is predominant in West Europe and the United States of America (USA). In Africa, there are several subtypes exist such as subtype A, B, C, D, E, F, G, H and O (Murphy *et al.*, 1993; Myers *et al.*, 1995). In Thailand, two common subtype have been isolated which are subtype B and subtype E. Subtype B is suspected to have been there mainly through intravenous drug users (IDU) while subtype E mainly through sexual transmission (Weniger *et al.*, 1994). Only 1 year

after Thailand epidemic wave among intravenous drug users in 1988, the virus then spread to the adjacent countries including Myanmar and Malaysia.

Group O was first identified as a new HIV-1 group due to the discovery of a distinct strain from a French woman (Charneau *et al.*, 1994). HIV-1 group O infection have been principally found in Africa. They are highly divergent from other strains constituting group M (Charneau *et al.*, 1994). This group has 30-50% sequence divergence from group M depending on the genes being compared (Yamaguchi *et al.*, 2003). Till now, HIV-1 group N is the rarest form of the virus and has only been seen in Cameroon (Ayouba *et al.*, 2000).

2.1.2 Properties and Structure

The structure of HIV is similar to the other retrovirus family member. HIV-1 is an enveloped virus which consists of two copies on non-covalently linked positive single stranded RNA (Bukrinsky *et al.*, 1993). The RNA surrounded by a conical capsid comprising the viral protein p24 which is very common in lentiviruses. The capsid consists of approximately 2000 copies of the p24 viral protein (Turner and Summers, 1999). It is surrounded by the viral envelope. The viral envelope is composed of a lipid bilayer membrane, formed from the cellular membrane of the host cell during budding of the newly formed virus particle (Barré-Sinoussi, 1996). Each envelope subunit consists of two non-covalently linked membrane proteins, glycoprotein 120 (gp120) and glycoprotein 41 (gp41) (Weiss, 1993). Gp120 is the outer envelope protein while gp41 is the transmembrane protein. The envelope protein is the most variable component of HIV. Gp120 facilitates viral entry by binding to the CD4, the main cellular receptor for all primate lentiviruses (Barré-Sinoussi, 1996). Gp41 is involved in the fusion of viral envelope with cellular membrane (Coffin, 1992; Levy, 1993). The nucleocapsid protein

is a low molecular weight and basic protein that conserved in all replication-competent retroviruses (Henderson *et al.*, 1981). The mature nucleocapsid protein on a mature virion is p7 (Henderson *et al.*, 1988; Tritch *et al.*, 1991). The first open reading frame (ORF) of the viral proteins, *gag,* encoded three proteins which are the matrix protein (p17), the major capsid protein (p24), and the nucleocapsid protein (p7). Whereas for the second ORF which is *pol* encoded protease, reverse transcriptase and integrase. Protease is needed for proteolytic processing from the *gag* and *gag/pol* precursor proteins (Coffin, 1992; Levy, 1993).

Other six virally encoded accessory proteins include *tat, rev, vif, vpr, vpu,* and *nef.* These proteins are not found in other classes of retroviruses impart novel levels of complexity to lentiviral replication. The function of *vif* is to suppress host factors that inhibit infection. *Vpr* enhances post cell entry infectivity while *tat* involves in activation of viral transcription. The *tat* protein up-regulates transcription from the viral promoter, the U3 region of the long terminal repeat (LTR) region by binding to nascent RNA transcripts. *Nef* down-regulates CD4 or major histocompatibility (MHC) expression (Cullen and Greene, 1990; Kestier Iii *et al.*, 1991; Park and Sodroski, 1995), whereas *rev* induces nuclear export of viral RNA from the translation of *gag/pol* proteins and template for envelope proteins. Without *rev,* no structural proteins are made. *Tat, rev,* and *nef* are the first viral components produced from multiply spliced viral mRNA. *Vpu* is the highlighted point for HIV-1 as it is differed by *vpx* in HIV-2 (Coffin *et al.*, 1986; Coffin *et al.*, 1997). The basic structure of HIV virus is illustrated as figure 2.1.

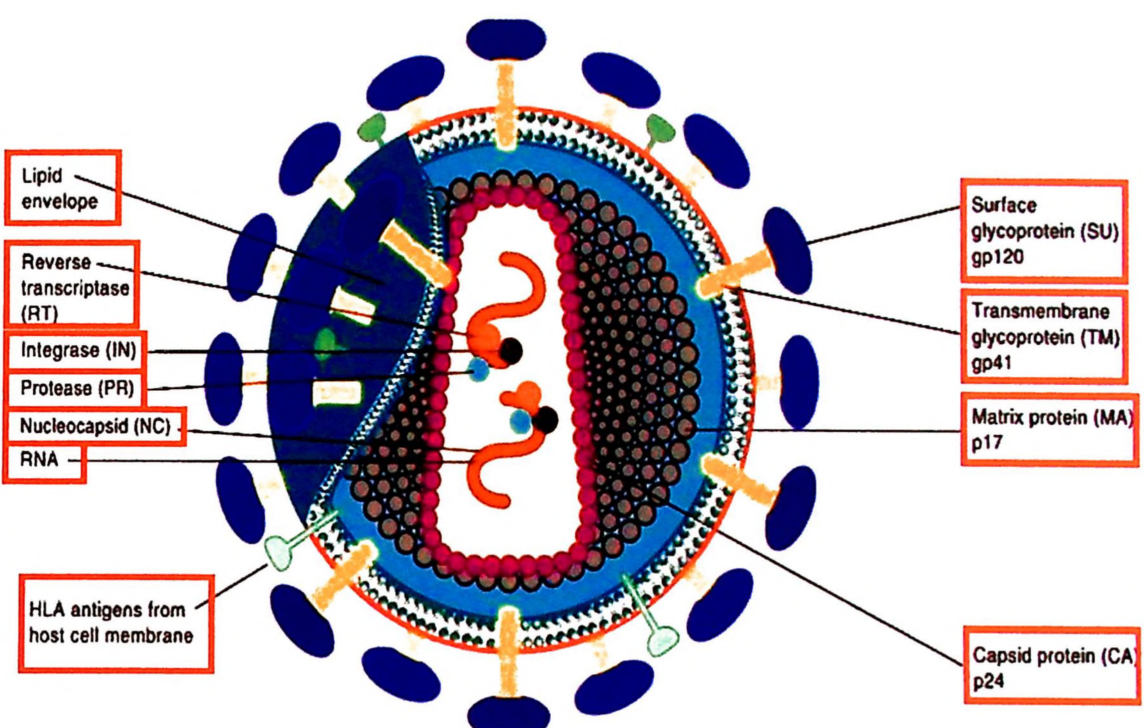


Figure 2.1 Basic structure of human immunodeficiency virus (HIV). Figure is adapted from The McGraw-Hill Companies.

2.1.3 Long Terminal Repeat (LTR) region

Besides the genes that encoding structure proteins, regulatory proteins and accessory proteins, the HIV-1 genome also contains two viral long terminal repeats (LTRs) regions at both ends of the proviral DNA when the provirus is integrated into the host genome (Engelman and Cherepanov, 2012). They are necessary for proviral DNA synthesis, integration into the host genome, transcription of the viral genes and modulating gene expression (Hiebenthal-Millow *et al.*, 2003; Romanchikova *et al.*, 2003). Proviral gene expression of HIV-1 is tightly regulated by the binding of cellular host proteins to a variety of *cis*-acting DNA sequences in the long terminal repeat (LTR) region of the viral genome (Gaynor, 1992). Each LTR is approximately 640 base pair (bp) in length and divided into three regions which are U3, R and U5 (Pereira *et al.*, 2000) (Figure 2.2).

U3 has been further divided into the modulatory, enhancer and basal/core promoter functional regions (Pereira *et al.*, 2000). These contain four functional regions that are important in the regulation of the viral replication as well as assembly and releasing infectious progeny of the virus. A modulatory region (nt -454 to -104) (Gaynor, 1992) is comprised of sequences upstream of the nuclear factor-kappa B (NF-κB) sites, which contains binding sites for numerous factors, including CCAAT/enhancer binding protein (C/EBP) factors (Tesmer *et al.*, 1993; Henderson *et al.*, 1995), activating transcription factor/cyclic AMP response element binding protein (ATF/CREB) factors (Krebs *et al.*, 1997; Krebs *et al.*, 1998), lymphocyte enhancer factor (LEF-1), and nuclear factor of activated T cells (NF-AT) (Shaw *et al.*, 1988). Early studies have shown that deletion within this region increased HIV-1 LTR-directed transcription and viral replication (Rosen *et al.*, 1985; Siekevitz *et al.*, 1987). It is

known as a negative regulatory element (NRE) which is between nt -340 and -184. The enhancer (nt -105 to -79) is located directly upstream of the core promoter and the Special protein (Sp) binding sites. It is defined primarily by the presence of two copies of the 10bp binding site for NF-κB and related factors (Nabel and Baltimore, 1987). The core promoter region is composed by the TATAA box and Specific protein (Sp) sites. The TATAA box is located 29 to 24 nucleotides of the transcriptional site whereas Sp binding sites refer to the three tandem GC-rich binding site (-45 to -77) interacting with specificity protein transcription factors (Sp1-Sp4). In order to initiate the transcription process, the TATAA box will bind with TATAA binding protein (TBP) and a few of other proteins which involve RNA polymerase II (*pol II*) transcription complex before elongation process takes place (Huh *et al.*, 1999; Ping and Rana, 2001).

The R region of HIV-1 LTR is consists of the transactivation response (TAR) element (nt +1 to +60) (Pereira *et al.*, 2000). The transcription starts at the beginning of R region and proceeds through U5 and the rest of the provirus. The process usually terminated by the addition of a poly A tract just after the R sequence in the 3' end of LTR. The U5 region contains other binding site for cellular transcription factors such as activating protein-1 (AP-1) that regulates gene expression like-activating protein 3, which is corresponding to the nuclear factor of activated T cells (NF-AT) site, and Sp1, which is very important for virus infectivity (Munier *et al.*, 2005).

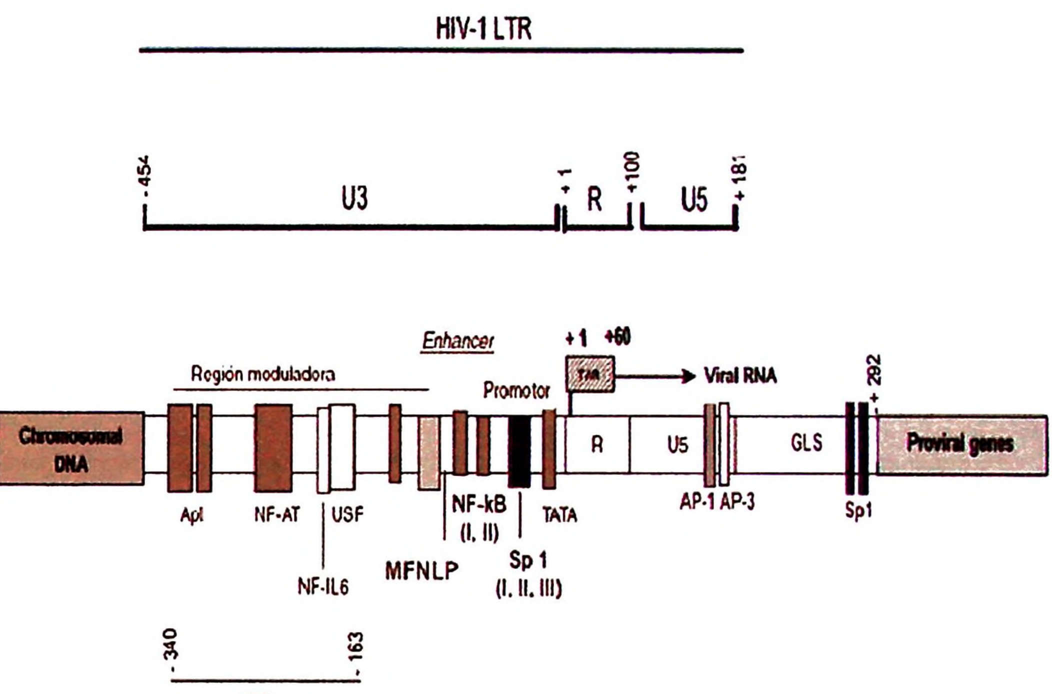


Figure 2.2 HIV-1 LTR structure of proviral DNA and binding sites for cellular transcription factors in the promoter, enhancer and modulatory regons of U3, R, and U5. This figure is adapted from Ramírez de Arellano et al., 2007.

2.2 Replication cycle

A virus requires a living cell as host that will provide energy and important components necessary for viral replication. The host for HIV is the T-cell or CD4 cell. Once attached to the cell, HIV injects its proteins into cytoplasm of the cell. This causes a fusion of the cell membrane to the outer envelope of HIV. Once in the cell, the single stranded RNA of the HIV must be converted to the double stranded DNA with the aid of reverse transcriptase enzyme. In order for the cell to reproduce, it must integrate the newly formed DNA into the cell nucleus. All the materials then must be separated and assembled into new HIV. Finally, the newly formed HIV pinches off and enters into the circulation and ready to start the whole process again.

2.2.1 HIV-1 Entry

HIV undergoes a complex series of steps to deliver its genome, into the host cell cytoplasm and at the same time evading the host immune response. HIV entry starts with the attachment of the virus to the host cell and ends up with the fusion of the cell and viral membranes with subsequent delivery of the viral components into the cytoplasm. Virus entry starts with the binding to the target cell mediated by either the viral envelope (Env) protein or the host cell membrane proteins incorporated into the virion with any one of a number of various cell attachment factors. The viral envelope (*Env*) might be binding to the negatively charged cell surface heparin sulphate proteoglycans of the host cell (Saphire *et al.*, 2001), or interacting with α4β7 integrin (Arthos *et al.*, 2008; Cicala *et al.*, 2009), or pattern recognition receptors such as a dendritic cell specific HIV-1 binding protein (DC-SIGN) that enhances trans-infection of T cells (Geijtenbeek *et al.*, 2000) which then would bring the *Env* closely to the viral

receptor CD4 and co-receptor, thus increasing the efficiency of infection (Figure 2.3) (Orloff *et al.*, 1991).

The second step of virus entry is binding of the HIV-1 *Env* to the CD4. The *Env* is heavily glycosylated with gp120 and gp41 heterodimers (Fenouillet and Jones, 1995). The gp120 contains five relatively conserved domains (C1-C5) and five variable loops (V1-V5) which is very important for receptor binding. With the exception of V5, each variable region consists of a loop structure. This loops lie at the surface of gp120 and play critical roles in immune evasion and co receptor binding especially V3 loop (Hartley *et al.*, 2005; Sirois *et al.*, 2005). CD4 normally functions to enhance T cell receptor (TCR)-mediated signalling. *Env* interacts with the CD4 binding site in gp120 (Kwong *et al.*, 1998). Once *Env* binds to the CD4, rearrangements of V1/V2 occurred and subsequently V3. It is then leads to formation of the bridging sheet, a four-stranded β sheet consist of two double stranded β sheets that are spatially separated in the unliganded state (Kwong *et al.*, 1998; Chen *et al.*, 2005).

Formation of bridging sheet and rearrangements of V3 is important for the third step of viral entry which is co-receptor binding. Next is the movement of the virus particle to the place where membrane fusion occurs. HIV love to use host cell machinery to reach sites where membrane fusion occurs (Lehmann *et al.*, 2005; Sherer *et al.*, 2010). HIV must control the host cell's endocytic machinery and be internalized by it in order for the membrane fusion to occur (Miyauchi *et al.*, 2009).

The fifth which is the final step of virus entry is membrane fusion mediated by *Env*. Co-receptor binding from previous step expose the hydrophobic gp41 fusion

peptide, which then inserts into the host cell membranes to connect it with the virus. The fusion peptide folds at the hinge region, where an amino-terminal helical region (HR-N) and a carboxyl-terminal helical region (HR-C) from each gp41 subunit combining together to form a six-helix bundle (6HB) (Weissenhorn *et al.*, 1997). Formation of the 6HB resulting in the formation of a fusion pore (Melikyan, 2008). When the membrane fusion pore is opened and stabilized, the viral contents can be delivered into the host cell cytoplasm.

2.2.2 HIV-1 Reverse Transcription

HIV-1 reverse transcriptase is produced from an interaction between *Gag* and DNA polymerase II (*Pol*) which is then formed *Gag-Pol* polyprotein by cleavage with the viral protease (PR) (Hu and Hughes, 2012). After fusion of the viral and cellular membrane, the virion undergo uncoating process in the cytoplasm (Wilen *et al.*, 2012). Contents of the virion is released into the cytoplasm. The viral RNA is then converted to a double stranded cDNA. Genomic RNA is plus-stranded which is both genome and messages are copied from the same DNA strand (Hu and Hughes, 2012). Reverse transcription is begin with the annealing of Lys3 as primers to the primer binding site (PBS) (Hu and Hughes, 2012). Based on several studies, once the first five or six deoxyribonucleotides have been added to the 3' end of the Lys3, the DNA synthesis accelerating (Isel *et al.*, 1996; Lanchy *et al.*, 1998). This process creates an RNA-DNA duplex, which is a substrate for RNase H which then will degrade RNA (Hu and Hughes, 2012). RNase H degradation removes the 5' end of the viral RNA, exposing the newly synthesized minus-strand DNA. It is then be transferred to the 3' end of the viral RNA by direct repeats called R. After this transfer, minus-strand synthesis can proceeds as well as RNase H degradation (Hu and Hughes, 2012). Each end of the new

synthesis viral DNA consists of the same sequence, U3-R-U5 which is longer than the original RNA. U3-R-U5 is called long terminal repeats (LTRs) that will be the end of the provirus, immediately after integration process (Hu and Hughes, 2012).

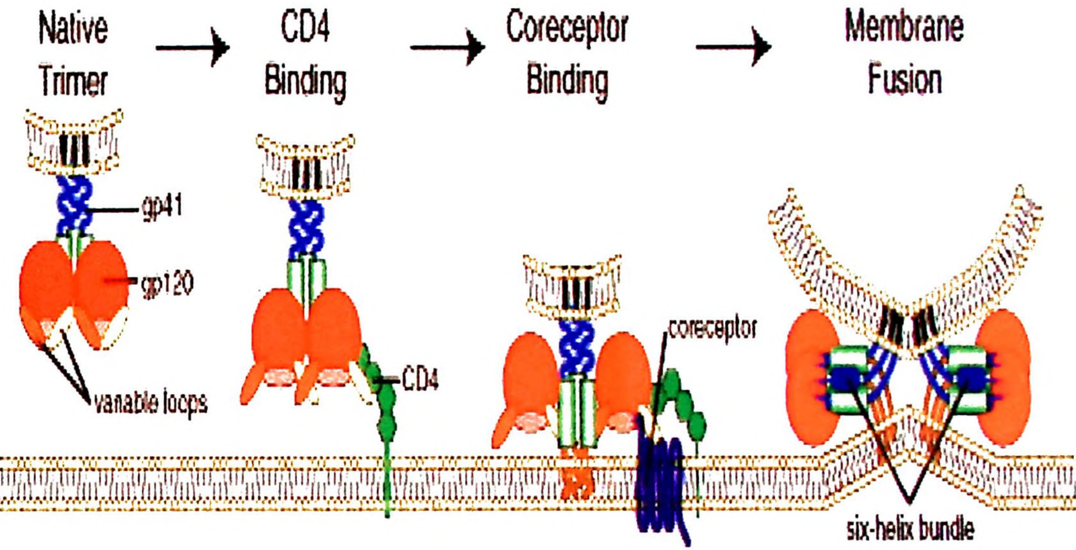


Figure 2.3 Overview of HIV entry. Starts with binding to the host cell which is to CD4. Conformational changes in Env, allow the co-receptor binding mediated by the V3 loop of Env. When the fusion peptides of gp41 inserts into the target membrane, six-helix bundle is formed and membrane fusion process is completed. This figure is adapted from (Tilton and Doms, 2010).

2.2.3 HIV-1 DNA Integration

After the viral DNA is synthesized by reverse transcription process in cytoplasm, the DNA is integrated by the virus-encoded IN (viral integrase) protein (Craigie and Bushman, 2012). The viral DNA is stably associated with IN and other proteins as a high-molecular-weight nucleoprotein complex which is then transported to the nucleus for integration (Craigie and Bushman, 2012). In nucleus, nucleus transcription factor NF-κB binds with NF-κB binding sites within HIV-1 LTR region (Knysz *et al.*, 2007). At first, *tat* and *rev* are transcribed and then followed by transcription of *gag, pol,* and *env* genes. The replication process is completed by the transcription of proviral DNA using host RNA polymerase II (Hu and Hughes, 2012).

2.2.4 HIV-1 Assembly, Budding and Maturation

In order to spread the virus successfully, the site of virus assembly needs to be precisely regulated. Defects in this process could either inhibit efficient production of infectious virions or block transmission to another host cell. HIV-1 virion assembly normally occurs at the plasma membrane, within specialized membrane microdomains. Retrovirus particle production is mediated by the HIV-1 Gag polyprotein. This polyprotein is important in virion assembly as it mediates the binding of the virus to the plasma membrane, making the protein-protein interaction, concentrating the viral Env protein and packaging the genomic RNA (Sundquist and Kräusslich, 2012) Gag is synthesized as a precursor polyprotein $Pr55^{Gag}$. Upon virus release, the viral protease cleaves $Pr55^{Gag}$ and gives rise to mature Gag protein (Ono, 2009). Virus particle production is a multistep process that includes Gag targeting the site of virus assembly, binding of Gag to plasma membrane, Gag multimerization and budding and release of

nascent virus particles. Gag binding with the plasma membrane is mediated by the N-terminal myristate moiety and a highly basic region in matrix, whereas Gag multimerization is mediated by both capsid-capsid interaction and RNA and nonspecific RNA binding as well as promoting Gag association is mediated by nucleocapsid (Gorelick *et al.*, 1993; Lingappa *et al.*, 2006). The final step of the viral replication cycle is called budding where the genetic materials of the virus are pushed away and a new outer coat is made from the host $CD4^+$ cells membrane (Simon *et al.*, 2006). The budding and release event is mediated by the host endosomal sorting complexes required for transport (ESCRT) machinery recruited by p6 to the site of virus assembly (Demirov and Freed, 2004; Morita and Sundquist, 2004; Bieniasz, 2006). The overview of HIV-1 replication process is illustrated in Figure 2.4.

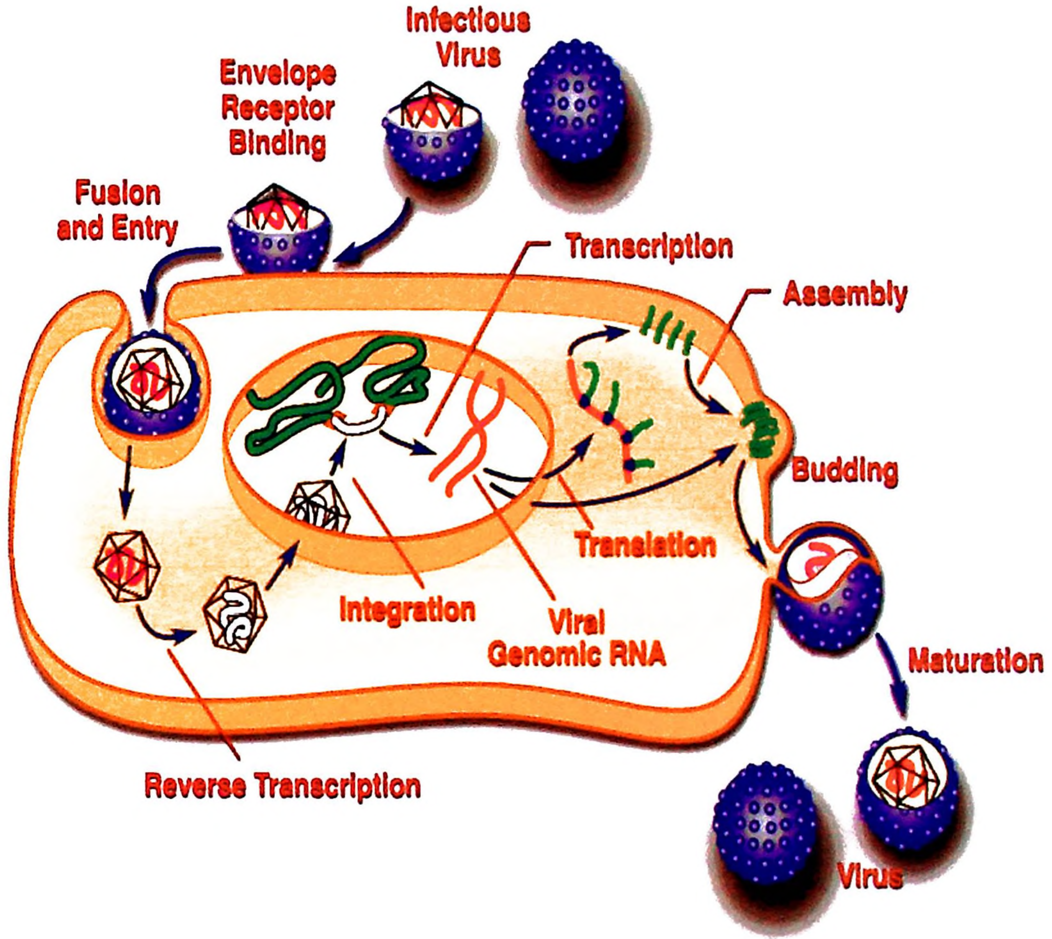
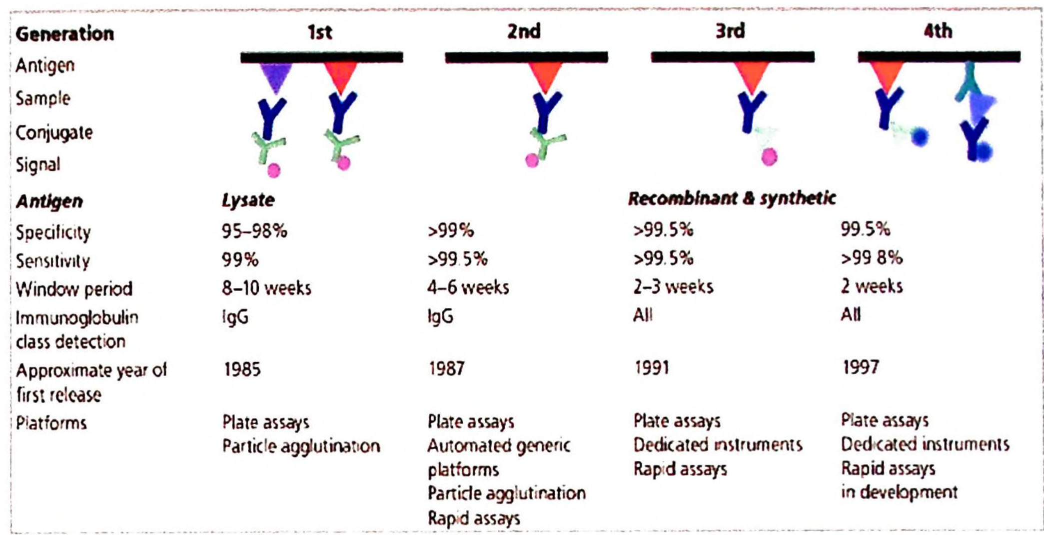


Figure 2.4 The figure above shows an overview of HIV-1 replication cycle. Replication process started with the viral entry, reverse transcription process, followed by integration in the nucleus and translation. It is ended up with the budding and maturation of the virus. Figure is adapted from The RCAS System.

2.3 Diagnosis

2.3.1 Enzyme Immunoassay (EIA)

The principles of immunoassays are based on the specific interaction of particular proteins or peptides with antibody raises against them. There have been four successive generations of HIV enzyme immunoassays for the purpose of screening and diagnosis. Each generation came with advanced immunoassay especially reduction of window period, thus increasing the effectiveness in detecting early infection of HIV. Each generation would be much better than previous generation. The performance of each generation is summarized in figure 2.5. Enzyme immunoassays (EIAs) remain the cornerstone in HIV diagnostic since the first HIV enzyme immunoassay was introduced 28 years ago (Branson, 2010). They are widely available, fastest and accurate.



Generation	1st	2nd	3rd	4th
Antigen				
Sample				
Conjugate				
Signal				
Antigen	***Lysate***		***Recombinant & synthetic***	
Specificity	95–98%	>99%	>99.5%	99.5%
Sensitivity	99%	>99.5%	>99.5%	>99.8%
Window period	8–10 weeks	4–6 weeks	2–3 weeks	2 weeks
Immunoglobulin class detection	IgG	IgG	All	All
Approximate year of first release	1985	1987	1991	1997
Platforms	Plate assays Particle agglutination	Plate assays Automated generic platforms Particle agglutination Rapid assays	Plate assays Dedicated instruments Rapid assays	Plate assays Dedicated instruments Rapid assays in development

Figure 2.5 The generation of immunoassay in HIV diagnostic (Chappel *et al.*, 2009).