COMBINATORIAL CRISPR/CAS9 FOR SUPPRESSION OF LATENT HIV-1 PROVIRUS IN VITRO

NITHYA A/P RAVICHANTAR

UNIVERSITI SAINS MALAYSIA

2020

COMBINATORIAL CRISPR/CAS9 FOR SUPPRESSION OF LATENT HIV-1 PROVIRUS IN VITRO

by/oleh

NITHYA A/P RAVICHANTAR

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

November 2020

ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisor, Dr. Kumitaa Theva Das for giving me the opportunity to work on this project, for her constructive criticism, continuous guidance and support that have enabled me to complete this project. I would like to thank my co-supervisors: Dr. Muhammad Amir Yunus and Professor Dr. Narazah Mohd Yusoff and field supervisor: Dr. Alwi Muhd Besari for sharing their knowledge, expertise and for providing guidance and motivation. I would also like to thank Dr. Ida Shazrina Ismail, Dr. Muhammad Rafeezul Mohamed, Dr. Hasni Arsad and Dr. Emmanuel Jairaj Moses for sharing their valuable input, feedback and expertise. I would like to thank my friends, lab members and lab staffs at AMDI, Ishola Deji, Laina, Benjy, Nalini, Nurul Islam, Faizah, Sakinah, Khirun, Azali, Sattam, Precela, Adam Azlan, Rajayaashini, Khor Kang Xi, Wahida, Salwani, Maheswaran, Pn. Hikmah and En. Zul for sharing their knowledge, giving moral support and assistance whenever needed. Next, a very big thanks to my mother and siblings for their constant care, love, support, patience and sacrifices. To my dad whom I lost during this journey, not a day goes by that I do not miss you and thank you for watching over me from heaven. Also, I would like to thank MyPhD and USM Graduate Assistant Scholarship for sponsoring my studies. Most importantly, I would like to thank the Almighty God for His guidance and 'companionship' throughout this journey.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS AND UNITS

%	Percent
±	Plus-minus
+	Positive
-	Negative
µg	Microgram
µL	Microliter
µm	Micrometer
µM	Micromolar
bp	Base pair
g	Gram
h	Hour
kb	Kilobase
L	Liter
M	Molar
mg	Milligram
min	Minute
mL	Milliliter
mM	Millimolar
ng	Nanogram
nm	Nanometer

LIST OF ABBREVIATIONS

Amp	Ampicillin
bGH-polyA	Bovine growth hormone polyadenylation
CD4	Cluster of Differentiation 4
CD8	Cluster of Differentiation 8
CMV	Cytomegalovirus
CCR5	C-C Chemokine receptor 5
CXCR4	C-X chemokine receptor 4
CO_2	Carbon Dioxide
dH_2O	Distilled water
DSB	Double strand break
E. coli	*Escherichia coli*
EtOH	Ethanol
FBS	Fetal Bovine Serum
Gag	Group-specific antigen
HIV	Human Immunodeficiency Virus
HIC	High income countries
IL	Interleukin
IFN	Interferon
LMIC	Low- and middle-income countries
LTR	Long Terminal Repeat
MHC	Major Histocompatibility Complex
Nef	Negative Regulatory Factor
NF-κB	Nuclear Factor Kappa B
PBMCs	Peripheral Blood Mononuclear Cells

PCR	Polymerase Chain Reaction
PHA	Phytohemagglutinin
PLWH	People living with HIV-1
Pol	Polymerase
Rev	Regulatory of expression of viral protein
Tat	Trans-activator of transcription
TNFα	Tumor Necrosis Factor Alpha
Vpr	Viral Protein R

KOMBINASI CRISPR/CAS9 BAGI PENINDASAN PROVIRUS HIV-1 PENDAM SECARA IN VITRO

ABSTRAK

Terapi ubatan HAART mampu menyekat replikasi HIV-1, tetapi keberkesanannya terhad kepada subjenis HIV-1 bukan-B, ketidakupayaan untuk menghapus takungan HIV-1 pendam, rawatan sepanjang hayat dan masalah kerintangan virus. Kaedah pengubahsuaian genom terkini iaitu CRISPR/Cas9, berpotensi mengatasi masalah-masalah tersebut. Dengan mensasarkan jujukan LTR (promoter) pada HIV-1, CRISPR/Cas9 mampu menghapuskan takungan HIV-1 pendam. Dalam kajian ini, sistem CRISPR/Cas9 yang sedia ada ditambahbaik untuk berfungsi sebagai CRISPR/Cas9 gabungan yang mensasarkan ubahsuaian beberapa gen HIV-1 termasuk, gen struktur (Pol dan Gag), gen pengawal (Rev and Tat) dan gen aksesori (Vif) secara serentak. Kajian ini dibentuk bagi menguji keberkesanan kaedah berkenaan sebagai terapi gen untuk subjenis HIV-1 yang berbeza dalam sesebuah populasi. Data awal kajian menunjukkan bahawa CRISPR/Cas9 gabungan mampu menurunkan kandungan virus tanpa menyebabkan ketoksikan kepada sel-sel manusia. Subjenis virus di dalam sampel darah pesakit HIV-1, yang tidak dirawat dengan terapi ubatan HAART, dikenapasti dengan mengunakan keadah PCR nested. Sel mononuklear darah periferi (PBMCs) dirawat dengan CRISPR/Cas9 gabungan untuk menentukan keberkesanannya dalam menurunkan ekspresi gen-gen HIV-1 yang disasarkan dan kekerapan mutasi dalam jujukan sasaran dan kesannya terhadap struktur protein selepas rawatan telah diramal. Keberkesanan menyeluruh keadah ini dalam penurunan kandungan virus, merujuk kepada ekspresi gen p24, dan kesannya terhadap pemulihan ekspresi CD4 juga ditentukan. Kaitan antara keberkesanan

kaedah ini dengan subjenis virus dalam sampel pesakit serta jantina dan umur pesakit turut di analisa. Keadah ini menunjukan keberkesanan yang lebih tinggi dalam menurunkan ekspresi HIV-1 Tat dengan kekerapan mutasi yang tinggi dalam jujukan dan perubahan ketara dalam struktur protein. Keadah ini juga menunjukkan keberkesanan yang lebih tinggi, iaitu 66.19% dalam menurunkan kandungan HIV-1 dan meningkatkan bilangan CD4 sebanyak 77.40% dalam sampel darah pesakit lelaki berbanding pesakit perempuan. Bagi perbandingan keberkesanan dalam kalangan sampel pesakit lelaki yang lebih muda, peningkatan ekspresi CD4 mencecah 94.72%. Kesimpulannya, CRISPR/Cas9 gabungan berpontensi dalam penindasan virus dan mengaruh tindak balas imun yang positif dalam semua pesakit HIV-1, dengan potensi keberkesanan yang lebih baik terhadap HIV-1 subjenis B dan dalam sampel pesakit lelaki yang muda. Kami berjaya menunjukan potensi CRISPR/Cas9 gabungan sebagai terapi alternatif yang mampu menghapuskan HIV-1 dalam sampel pesakit dan membaikpulih sistem imun pesakit. Seterusnya, kami berhasrat untuk menyiasat kecekapan CRISPR/Cas9 gabungan dengan terapi-terapi yang lain dan dalam sampel pesakit-pesakit HIV-1 yang sedang menerima rawatan HAART.

COMBINATORIAL CRISPR/CAS9 FOR SUPPRESSION OF LATENT HIV-1 IN VITRO

ABSTRACT

Although HAART effectively suppresses HIV-1 replication, its compromised effectiveness against non-B subtypes, the challenge of eliminating latent proviruses, life-long treatment, and viral resistance complicates the cure for HIV-1. CRISPR/Cas9, the latest genome editing tool, can overcome the limitations seen with HAART. By targeting the promoter of HIV-1 (LTR), CRISPR/Cas9 can disrupt the latent reservoirs. Here, we designed an improved CRISPR/Cas9 system, combinatorial CRISPR/Cas9 to concurrently knockdown multiple HIV-1 genes: structural (Pol and Gag), regulatory (Rev and Tat) and accessory genes (Vif). We investigated the efficacy of this tool as a therapy against different viral subtypes in a subset population. Our preliminary data, showed a huge decrease in viral load and we verified the safety of CRISPR/Cas9 in human cell lines. Next, we subtyped the virus from HIV-1 positive HAART naïve patients' samples via nested PCR. Infected peripheral blood mononuclear cells (PBMCs) were treated with combinatorial CRISPR/Cas9. Mutation frequency in targeted genes was determined and its corresponding changes in protein structure were predicted. The overall efficacy of the tool in reducing the viral load with p24 measurements and its effect on cell proliferation with CD4 quantification was also determined. Additionally, we determined the correlation of the efficacy of the tool with subtype, gender and age. Combinatorial CRISPR/Cas9 effectively knocked down targeted HIV-1 genes with highest suppression in HIV-1 Tat mRNA expression, corresponding with occurrence of highest mutation frequency and protein structure distortion after treatment. The tool shown greater efficacy in suppressing the virus

production by 66.19% and improving the CD4 T cell count by 77.40% in elderly male patient sample compared to elderly female. A 94.72% increase in CD4 T cell count was observed in younger male patient sample compared to elderly male patient. Thus, the combinatorial CRISPR/Cas9 facilitates an efficient viral suppression and positive immune response in all HIV-1 patient samples, with higher efficiency against subtype B, in younger male patient. This study suggests combinatorial CRISPR/Cas9 could potentially serve as an alternative therapy to control the morbidity of HIV-1. Next, we will investigate the efficacy of CRISPR in combination with other currently existing therapies, and in patients on HAART.

CHAPTER 1

INTRODUCTION

1.1 Research Overview

AIDS is caused by Human Immunodeficiency Virus 1 (HIV-1) and it has remained a global concern since its discovery in 1980s. HIV-1 infects the immune cells, primarily T cells that express CD4 surface proteins (known as CD4+ T cells) (Dragic et al., 1996; Vidya Vijayan et al., 2017) . The infection will cause a gradual loss of T cells which weaken the immunity and thus increases vulnerability to various opportunistic infections such as tuberculosis, pneumonia, and candidiasis (Kim et al., 2016). Since HIV-1 infection is primarily transmitted through sexual intercourse, sharing of needles, and mother-to-child (Shaw and Hunter, 2012), this illness impacts the society, as well as the economy of the country due to the need to provide medical and welfare to the people living with HIV-1 (PLWH). The only treatment that is currently available is HAART (Highly Active Anti-Retroviral Therapy) and it has changed the recognition of AIDS from a 'death sentence' to a chronic but manageable infection. However, a cure has yet to be found. The challenges include the cost, compromised effectiveness of HAART against non-B subtypes, inability to eliminate latent proviruses which require the patients to be on life-long treatment, and prolonged exposure that causes many side-effects and viral resistance (Desai et al., 2012; Lu et al., 2017). These issues highlight the importance for an alternate method to control morbidity and mortality of HIV-1. The latest genome editing tool, CRISPR/Cas9 which uses an RNA-guided Cas9 as a nuclease, acts like a molecular scissors that cleaves the targeted gene by causing double strand break (DSB). The DSB will be

repaired endogenously by a cell repair mechanism known as non-homologous end joining (NHEJ) via random insertion and deletions, that eventually disrupts the function of the gene (Perez et al., 2008; Mussolino et al., 2014; Kistler et al., 2015). Several studies have reported CRISPR's capability in disrupting the genome of HIV-1, including the latent reservoirs, by targeting the promoter region of the virus, LTR (Ebina et al., 2013; Hu et al., 2014; Kaminski et al., 2016; Huang and Nair, 2017). However, the emergence of viral escape variants due to a single gene targeting has made CRISPR a 'double-edged sword' (Wang et al., 2016). In addition, its simple mode of binding and cutting has raised some concerns whether it could cause off-target binding (Fu et al., 2013).

Therefore, in this study, we used CRISPR to target LTR, Gag, Pol, Tat, Rev and Vif genes of HIV-1, genes that play important roles in structural, regulatory and virulence functions. We compared the efficiency and safety of CRISPR with its mutated variant known as nickase, which nicks the DNA strand instead of causing a DSB as the wild type (nuclease variant) does. We achieved encouraging results in the *in vitro* study of assessing CRISPR's capability in reducing viral production without compromising its safety in human cells with the nuclease variant. Although the nickases were found to achieve similar efficiency in reducing viral production as the nucleases, they were toxic to human cells, and were therefore excluded from downstream experiment. Next, we explored the potentiality of this tool in HIV-1 patient samples in Malaysia, specifically in Kelantan. We subtyped the virus isolated from patient samples to determine the current subtype prevalence and to compare the efficiency of treatment across different viral strains. CRISPR showed a decrease in viral production across different subtypes and positive immune responses in patient samples. Also, we found greater decrease in viral production and better immune

recovery in patients with B viral subtype, male and younger patients. Therefore, this preliminary pilot study in HIV-1 patient samples suggests that CRISPR could serve as an alternative approach to HAART to control the morbidity of HIV-1.

1.2 Hypothesis

CRISPR/Cas9 can serve as a potential gene therapy tool to control different HIV-1 subtypes in Malaysia may serve an alternate to HAART. CRISPR/Cas9 will be able to cause greater viral load reduction and enable better immune response after treatment.

1.3 General objectives

To develop CRISPR/Cas9 as a potential gene therapy tool to reduce viral load and elicit positive immune response in patient samples in Malaysia with different HIV-1.

1.4 Objectives

- To design and establish CRISPR nuclease versus nickase system targeting HIV-1 genes LTR, Pol, Gag, Tat, Rev and Vif.

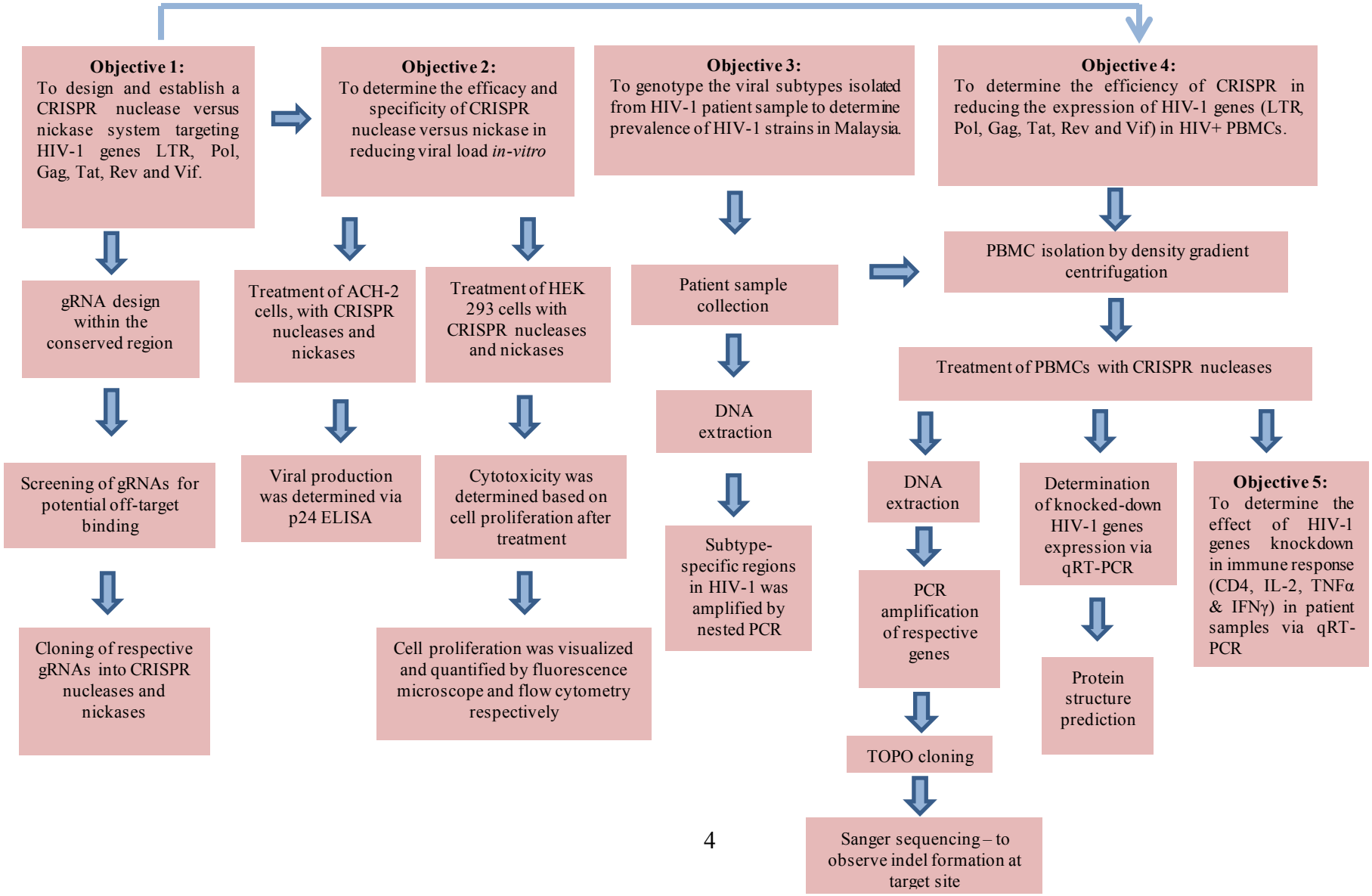
- To determine the efficacy and specificity of CRISPR nuclease versus nickase in reducing HIV-1 viral load *in-vitro*.

- To genotype the viral subtypes isolated from HIV-1 patient sample to determine the prevalence of HIV strains in Malaysia.

- To determine the efficiency of CRISPR in reducing the expression of HIV-1 genes (LTR, Pol, Gag, Tat, Rev and Vif) in HIV+ PBMCs.

- To determine the effect of HIV knockdown on immune response (CD4, IL-2, TNFα & IFNγ) in patient samples

1.5 Study flow chart.



CHAPTER 2

LITERATURE REVIEW

2.1 The global burden of HIV-1.

The acquired immunodeficiency syndrome (AIDS) is a serious medical condition in humans that weakens the immune system to fight off infections. AIDS is the end stage of an immune system infection caused by Human Immunodeficiency Virus (HIV-1) that primarily infects CD4 T cells, a major source of immune protection against pathogen in the human body. Depletion of CD4 T cells causes the individual to be vulnerable to various life-threatening infections and eventually leads to death (Douek et al., 2002; Dragic et al., 1996; Vidya Vijayan et al., 2017).

Thus far, HIV/AIDS has infected 75 million individuals and claimed more than 32 million lives since its discovery 30 years ago. Globally, there are 37.9 million infected individuals, 1.7 million new infections, and 770 000 deaths due to AIDS (AVERT, 2018; WHO, 2018) . In Southeast Asia (SEA), with the second largest burden of HIV-1 globally (Figure 2.1), there are 3.5 million individuals living with HIV-1, 180 000 new infections and 130 000 AIDS related death (Pendse et al., 2016). In Malaysia, we have 87 041 individuals living with HIV-1, 3 293 new infection and 42 843 AIDS related deaths (AVERT, 2018; MOH, 2019).

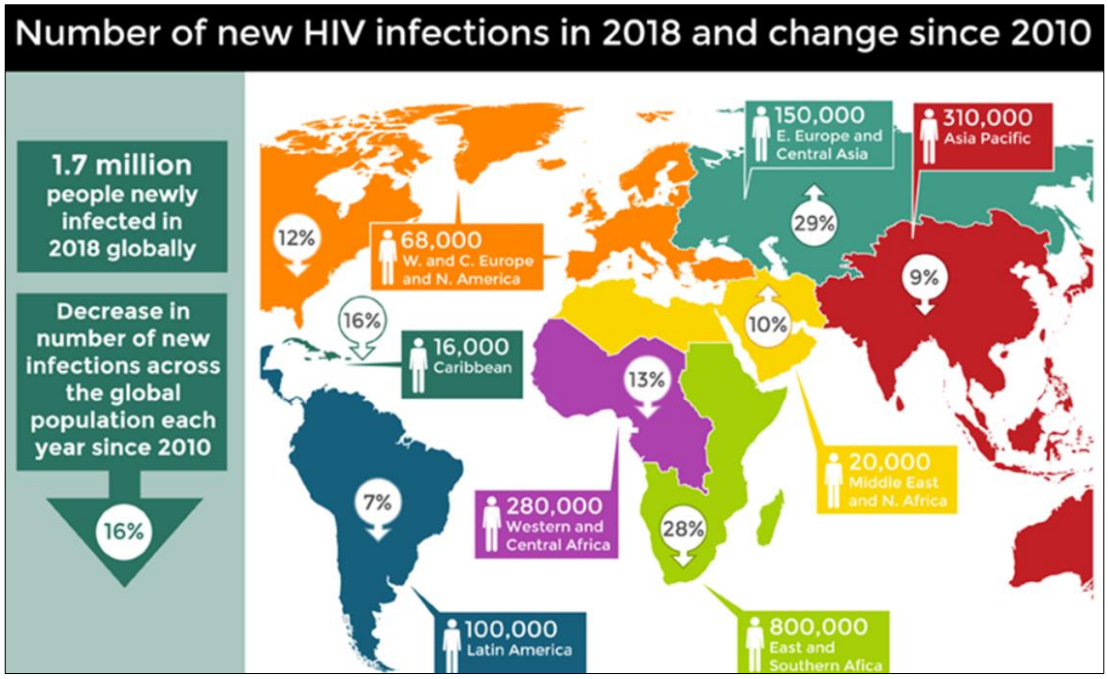


Figure 2.1 The global burden of HIV-1. Asia Pacific region (in red) which includes Southeast Asia (SEA), has second largest burden of HIV/AIDS after East and Southern Africa with 310,000 of new infection. Adapted from the (AVERT, 2018).

2.2 The distribution of HIV-1 subtypes

HIV-1 is a part of the Lentivirus genus and a member of the Retroviridae family. The M group of HIV-1 which is responsible for 90% of global infection can be subdivided into 10 subtypes: A, B, C, D, F, G, H, I, J and L (Figure 2.2) and has more than 100 circulating recombinant forms (CRFs) (Figure 2.2). There is a 25%-30% genetic difference between these subtypes which explains the dissimilarities in transmission rate, natural history, rate of CD4 T cells decline, progression to AIDS, and recovery rate in response to treatment (Taylor et al., 2008).

Among these 10 subtypes, subtype C accounts for most of the global HIV-1 infections (approximately 50%) (AVERT, 2018) . Africa and Southeast-Asia, account for the largest and second largest HIV-1 burden with subtypes C, A, D and CRF07_AG (Bbosa et al., 2019) and C, B and E, now known as CRF01_AE (Pendse et al., 2016)

respectively. Despite the prevalence of other subtypes, subtype B which accounts for only 12% of the infection, is the most researched subtype. Development of HAART (Highly Active Anti-Retroviral therapy), the only drug therapy able to control the infection of HIV-1, was also developed against subtype B, due to its predominance in developed countries such as USA, Japan, Western Europe and Australasia.

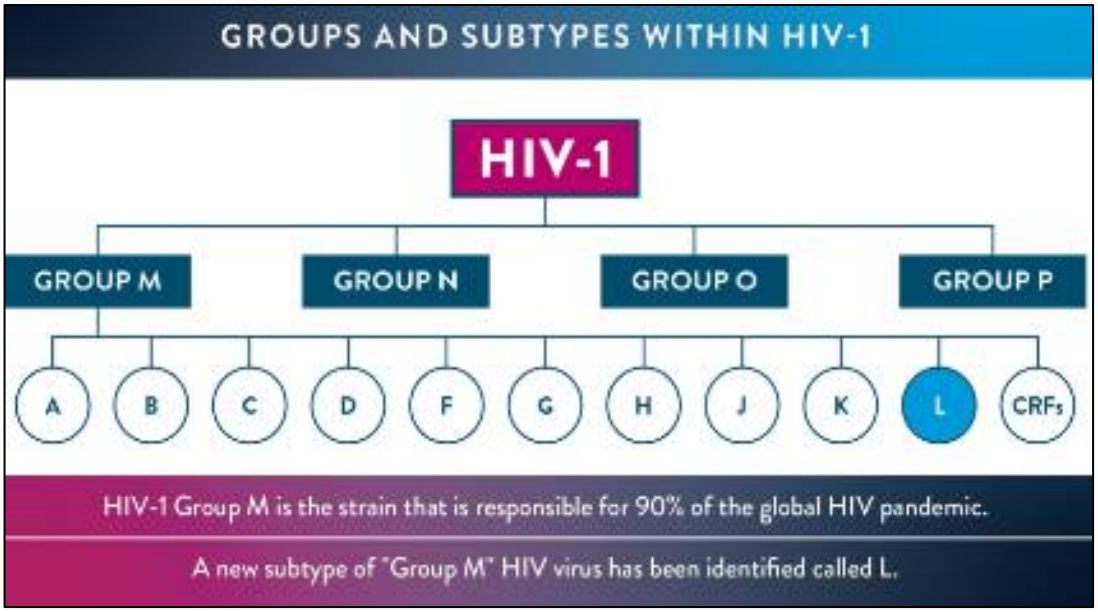


Figure 2.2 Groups and subtypes of HIV-1. There is an addition of a new subtype in group M known as L (highlighted in blue). Adapted from The Star Online, 2019.

In Malaysia, among the 15,100 HIV cases identified in 1990's, the patients were mostly infected with HIV-1 subtypes B, C and CRF01_AE (Brown et al., 1996; Saraswathy et al., 2000). Individuals infected with CRF01_AE (40.9%) is the predominant subtype, followed by CRF33_01B (20.5%), subtype B (10.1%) and new CRFs and URFs (Chow et al., 2016). The shift in prevalent subtypes is mostly due to the change in transmission, from intravenous drug use (IVDU) to sexual contact which accounts for 3% and 94 % of HIV-1 infection in Malaysia respectively (Figure 2.3). Thus, the diverse distribution and change of subtypes in the span of 20 years emphasizes

the need to have a treatment that would combat genetically diverse HIV-1 subtypes, and not just HIV-1 subtype B.

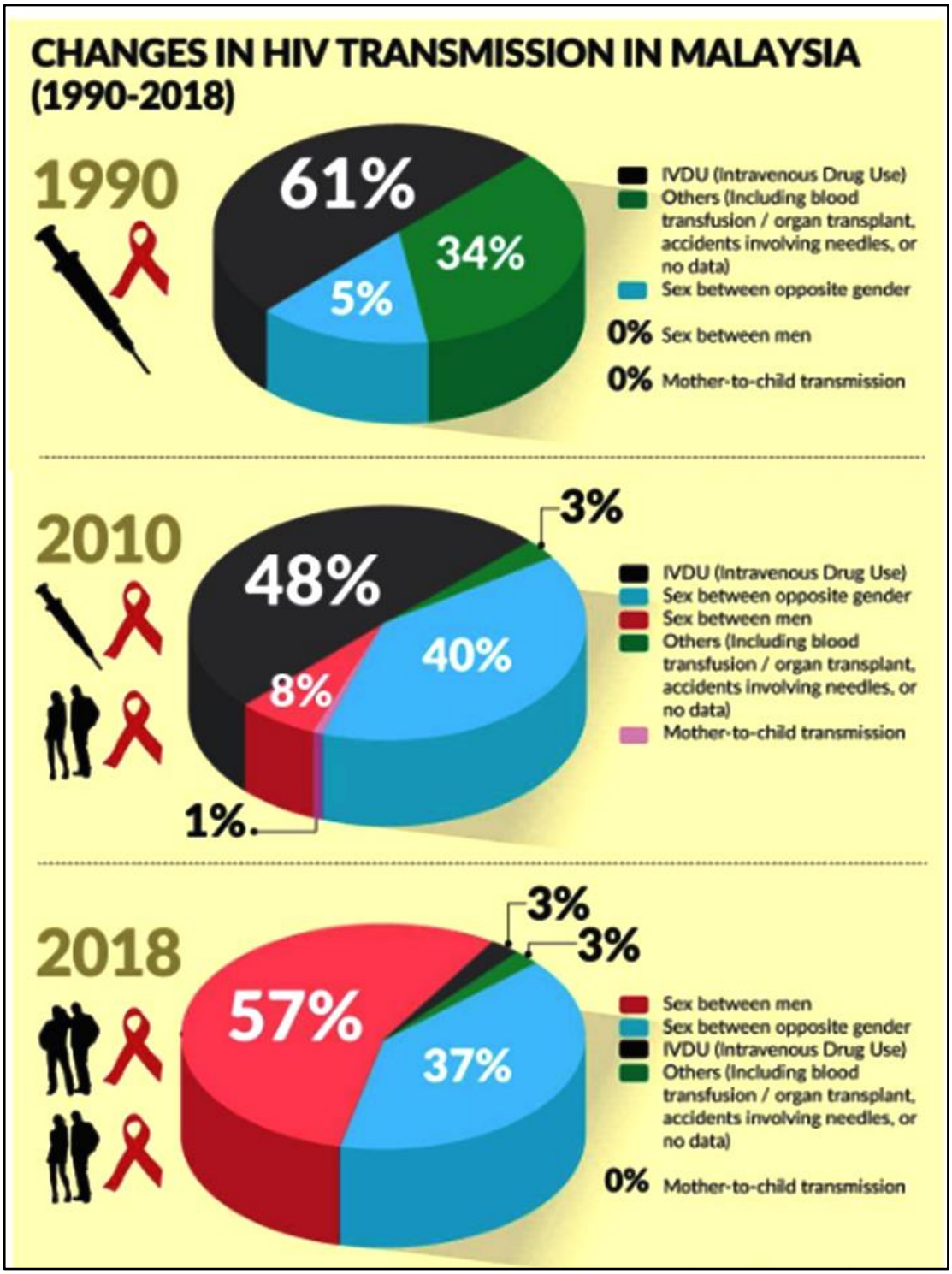


Figure 2.3 The changes in HIV-1 transmission in Malaysia from 1990-2018. The changes in trend of transmission from increased intravenous drug use (IVDU) and decreased sexual contacts in 1990s and decreased IVDU and increased sexual contacts between men is described. Adapted from The Star Online, 2019.

2.3 HIV-1 genome

HIV-1 is a single-stranded, positive-sense, RNA virus. The 9 kb genome of the virus is flanked by LTR (Long Terminal Repeats), which acts as a promoter and terminator at both ends respectively. The 5' to 3' reading frame of structural gene Gag (precursor of virus inner structural proteins) codes for matrix protein (p17: forms an inner membrane layer), capsid protein (p24: forms a conical capsid) nucleocapsid (p7: nucleic acid stabilizer) and p6 (releases the virus release from infected cells). This is followed by the second structural gene, Pol (precursor for viral enzyme) which codes for essential enzymes: protease (p12: facilitates proteolytic cleavage), reverse transcriptase (p51: facilitates transcription of HIV-1 RNA to proviral DNA), integrase (p32: facilitates integration of proviral DNA into host genome) and RNase H (p15: facilitates RNA degradation in RNA/DNA replication complex). At the end of Pol, the third structural gene, Env (precursor of envelope proteins) is present, which codes for surface proteins (gp120: attachment of virus to host cell) and transmembrane glycoproteins (gp40: allows fusion of virus and host membrane). There are also regulatory genes, Tat (p14: viral trans-activating protein) and Rev (p19: RNA splicing regulator) which modulate the splicing, transcribe viral genes and export viral mRNA. Accessory genes such as Nef, negative regulatory factor (p27: enhances viral infectivity and downregulate CD4 expression on target cells), Vif, viral infectivity factor (p23: blocks the antiviral activity of cellular protein, APOBEG3G), Vpu, viral protein unique (p16: allows efficient virus release and contains nuclear localization signal, thus is involved in intracellular trafficking) and Vpr, viral protein R (p15: interacts with p6, involved in viral infectivity and cell cycle) (Figure 2.4). Thus, each of these genes plays

crucial roles in the virus' life cycle (Levy, 2011; Sauter et al., 2012; Schmalen et al., 2018).

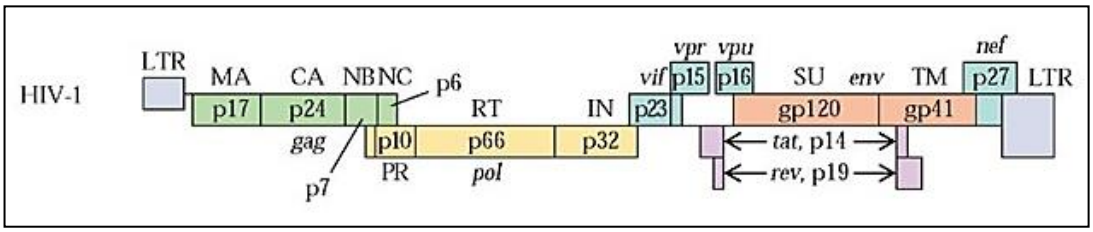


Figure 2.4 The linear genome of HIV-1. The genome consists of three structural genes and six regulatory genes. Adapted from : http://what-when-how.com/acp-medicine/human-retroviral-infections-part-1/.

2.4 HIV-1 life cycle

The stages in the HIV-1 life cycle are a) Entry and fusion of viral particle to host co-receptor, b) reverse transcription of RNA to cDNA, c) integration of cDNA into nucleus, d) replication, e) assembly and f) budding. Fusion takes place when the virus binds to co-receptors; CCR5 or CXCR4 of CD4 expressed on T cells. This process is mediated by envelope (Env) which is a heavily glycosylated trimer of gp120 and gp40 heterodimers. The gp120 contains five conserved domain (C1-C5) and five variable domains (V1-V5), which are genetically heterogeneous. These variable loops play an important role in immune evasion and co-receptor binding, especially the V3 loop. Post-fusion, this is followed by the release of HIV-1 genetic materials (RNA) into host cytoplasm and conversion of RNA to double stranded (ds) dsDNA by reverse transcriptase. dsDNA is transported into the host nuclease in a multiprotein complex form, pre-integration complex (PIC), which comprises of dsDNA, nucleo-capsid, viral matrix protein, Vpr. and host proteins such as p75 (which forms a bridge from PIC to nucleus). The integration would result in two possible events; a) integration into open

chromatin (where DNA is loosely packed and accessible by RNA polymerase and transcription factors) which would allow viral gene expression, thus facilitating productive infection or b) integration into the heterochromatin region (where DNA is firmly packed and not accessible by RNA polymerase and transcription factors) that would repress the transcription, thus facilitating latency establishment, where the virus resides for years until it is transcriptionally activated in the presence of άα, optimal concentration of Tat, and PKC modulators (which increase the level of necessary transcription factors and de-compact chromatin).

HIV-1 genes transcription is initiated when host's RNA polymerase II (RNA Pol II) binds to LTR. However, after initiation, Pol II pauses at the 5' promoter proximal region due to formation of stem loop structure, TAR. Transcription will be elongated by HIV-1 Tat, by binding to TAR which attracts host cofactors such as P-TEFb to release Pol II, allowing for the continuation of transcription and the production of viral mRNA. Then, the mRNA will undergo splicing to remove the non-coding regions (introns). The mature mRNAs will be transported to cytoplasm by Rev and translated into proteins which will then be cleaved by protease (an enzyme encoded by Pol) to produce functional proteins. With the help of HIV-1 Gag, the viral progeny with functional proteins then assemble in the cytoplasm and buds out the cell, to infect new cells, and the cycle continues.

However, the virulence of virus depends on the expression of its accessory proteins, Vif, Vpu and Nef. During the later stage of replication, Vif blocks the entry of host cellular proteins, APOBC3G by targeting it for ubiquitination, which would otherwise cause hypermutation in HIV-1 and kill them on arrival in the next host cell. Vpu causes degradation of CD4 T cells, thereby decreasing expression of CD4 T cells

surface receptor, while Nef downregulates MHC II expression which activates the adaptive immune system (Nisole and Saïb, 2004; Gélinas et al., 2018; Vansant et al., 2019).

2.5 Immune response to HIV-1 infection

Antigen presenting cells such as dendritic cells and macrophages that express CD4 receptors on their cell surfaces, recognize the presence of HIV-1 upon its entry. The MHC II of APC interacts with HIV-1 antigen forming a peptide complex. This peptide complex will be presented to T cells (which consist of helper CD4+ T cells and cytotoxic CD8+ T cells) in the lymph nodes. HIV-1 gains access to CD4 T cells and use the host cell machinery to multiply and lyse the cell. At this stage, there will be an increase in viremia and a decline of CD4 T cells. On the other hand, CD8 T cells that recognize the presence of HIV-1, secrete degradation enzymes such as granzyme B (protease which cleaves viral protein) and perforin (pore-forming protein) to target infected CD4 T cells for apoptosis, and release cytokines and chemokines such as INFγ, TNFα, MIP-1α, MIP-1β and Rantes which attract other immune cells such as monocytes and neutrophils to the site of infection. This promotes proliferation of T cells (some of the proliferated T cells differentiates into CD4 T cells but not to pre-infection level) and blocks the viral entry, which causes a slight decrease in viremia which continues to multiply in low levels. At this point, CD4 T cells function as usual by secreting cytokines (IFNγ and IL-2) to signal the movement of other immune cells to the site of infection. This causes B cells to undergo clonal expansion (differentiate into memory B cells) and create HIV-1 specific antibodies that recognize and mark HIV-1 for destruction. When the multiplied virus rebounds due to high expression of viral and host transcriptional regulators such as HIV-1 Tat and TNFα, it causes an extreme immune

activation and excessive T cell replication. This chronic antigen stimulation causes T cell exhaustion and some T cells become replicative incompetent while the rest are lysed by the virus. This will result in a peaked viremia and excessive decline in CD4 T cells, causing the immune system to be compromised by secondary infections, establishing the end stage of the infection, AIDS (Mogensen et al., 2010; Vajpayee et al., 2013; Perdomo-Celis et al., 2019) (Figure 2.5).

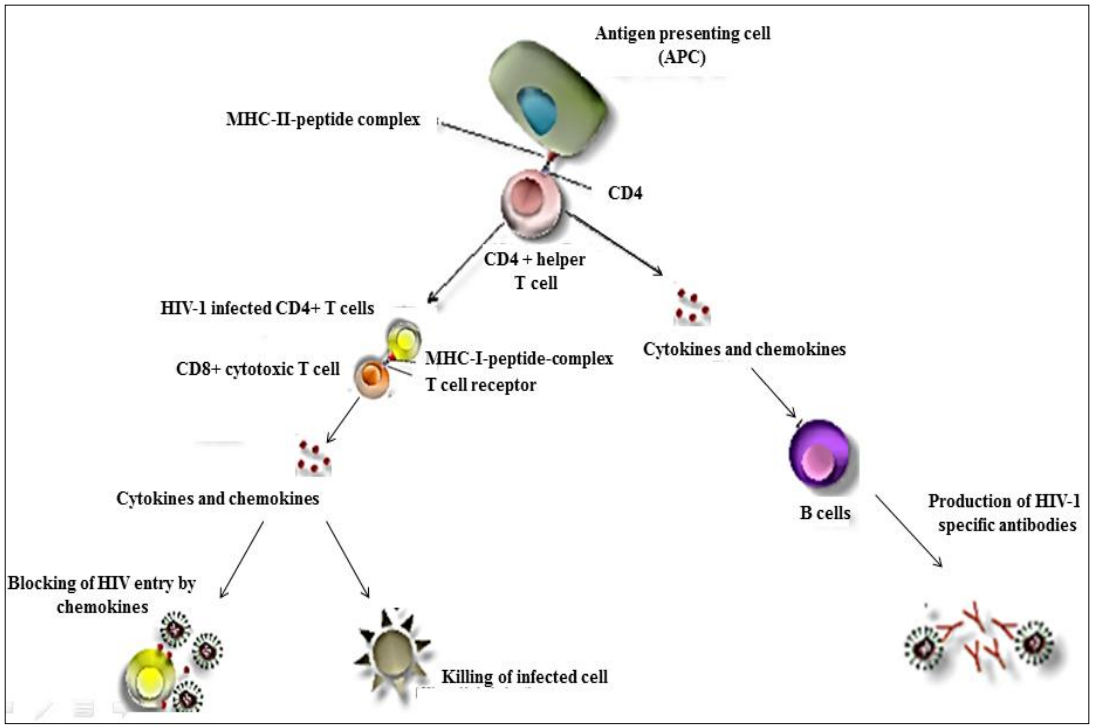


Figure 2.5 The immune response to HIV-1 infection. HIV-1 is first recognized by antigen presenting cells (APCs). APCs will capture the HIV-1 antigen and present it to T cells in lymph nodes. HIV-1 reaches its primary target, CD4+ T cells, and uses its machinery to multiply. Meanwhile, CD8+ T cells exert cytotoxic response by releasing degradation enzymes, cytokines and chemokines to kill the infected CD4+ T cells, block the entry of HIV-1 into host cells, and increase the proliferation of T cells. This will cause a decrease in viremia and increase in CD4+ T cells. CD4+ T cells release cytokines to attract other immune cells to the site of infection and instruct B cells to create HIV-1 specific antibodies to target the virus for destruction. Adapted fromhttps://www.immunology.org/publicinformation/bitesizedimmunology/pathogens-and-disease/human-immunodeficiency-virus-hiv.

2.5.1 Stages of HIV-1 infection

Stages or severity of the infection and efficiency of treatment in HIV-1 patients is determined by the number of viral load and CD4 T cell count, which are inversely correlated. There are 3 stages of infection associated with HIV-1.

a) Acute stage:

The primary stage of infection develops within 2-6 weeks of exposure. The contracted individual will present symptoms such as flu, fever, headache, and rash. HIV-1 multiplies rapidly and spreads throughout the body. At this point, the level of HIV-1 in the blood is very high, thus, possesses a high transmission risk. This stage does not last long due to CD8 T cells activation, which reduces the viral load and slightly increases CD4 T cells.

b) Chronic stage:

The virus that survives the CD8 T cells attack continues to multiply at low levels causing a gradual decrease in CD4 T cells. At this stage, patients may not have any HIV-1 related symptoms, therefore this stage is known as clinical latency. The duration of clinical latency is usually dependent on the remaining virus (known as viral load set point) which determines the rate of disease progression. However, it is still possible to transmit HIV to others at this stage, unless the patient is on HAART with undetectable viral load. This stage can last between 7 to10 years.

c) AIDS:

AIDS is the end stage of the infection, with very high viral load, extremely low CD4 T count (less than 200 cells/mm3) and is accompanied by opportunistic infections. Due

to the high viral load, infected individuals are able to transmit HIV very easily to others. Without treatment, the survival rate is typically about 3 years (Grossman et al., 2002; Fauci, 2007) (Figure 2.6).

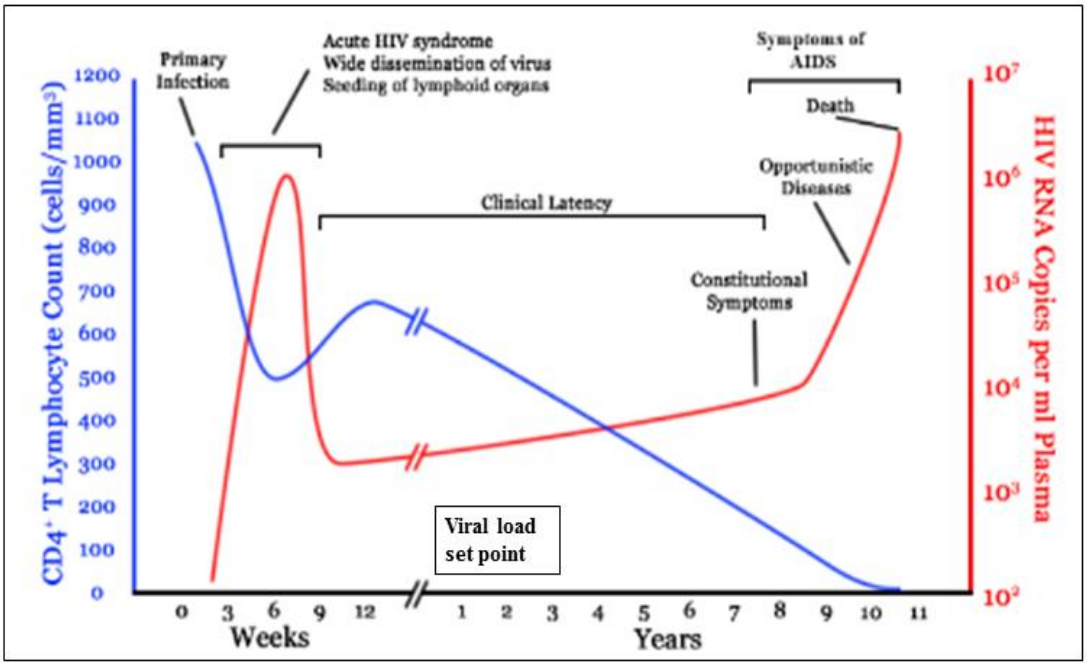


Figure 2.6 Natural course of HIV-1 infection. There are 3 stages in HIV-1 infection: acute, chronic and end stage. The viral load set point (the degree of viremia that determines the rate of infection progression) under clinical latency stage has also been illustrated here. Adapted from (Selinger and Katze, 2013).

2.6 Highly Active Anti-Retroviral Therapy (HAART)

The only drug therapy that manages the morbidity and mortality of HIV-1 infection is HAART, which is administered in a combination of three to four drugs (Agosto et al., 2014). The combination of multiple drugs inhibits the viral replication at different stages of the viral life cycle (Figure 2.7) including:

- Entry inhibitors: Enfuvirtide and maraviroc
- Reverse transcriptase inhibitors
    a) Nucleoside/nucleotide inhibitors: Abacavir, emtricitabine and tenofovir
    b) Non-nucleoside inhibitors: Efavirenz, etravirine and nevirapine
- Protease inhibitors: Atazanavir, darunavir and ritonavir
- Integrase inhibitors: Dolutegravir and raltegravir

HAART is highly effective in reducing viral loads below the detectable level and increasing the CD4 T cells count, provided that the patient strictly adheres to a life-long HAART regimen. This is because HAART does not eliminate latent proviruses, thus non-compliance could revoke the latent proviruses that would eventually be resistant to HAART. Despite HAART being the only effective treatment, HAART is expensive and not available in some low-income African countries, which has the largest HIV-1 burden (Ndashimye and Arts, 2019). Since HAART is developed against subtype B, its effectiveness to non-B subtypes could be compromised (Gatell, 2011). Side effects such as lipodystrophy, hyperglycemia and liver failure due to drug toxicity could worsen the complication faced by patients. Therefore, although HAART changes HIV-1 infection from deadly to a manageable chronic infection, a functional cure without being on HAART is still not possible. Thus, an alternative search for a functional cure that

permanently controls HIV-1 infection, even after cessation of HAART has been the priority in current HIV-1 research (Bartlett and Shao, 2009; Lu et al., 2017).

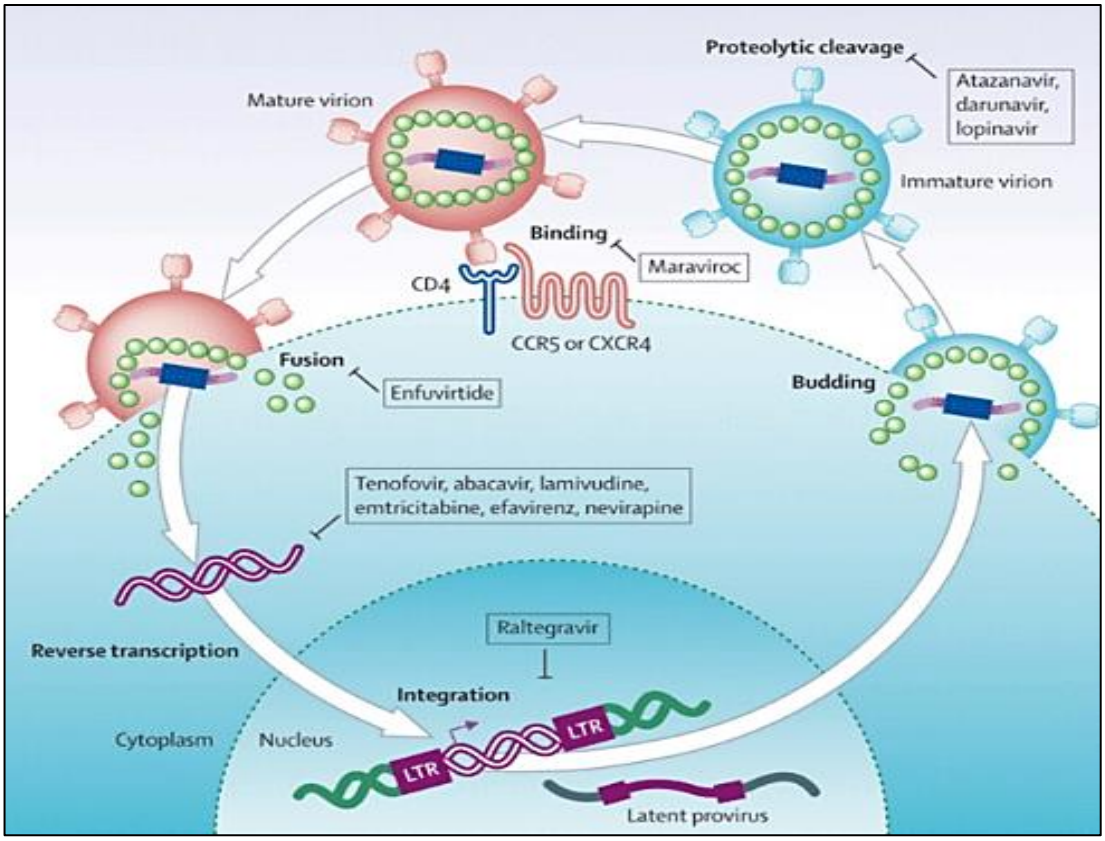


Figure 2.7 Interaction of HAART drugs at different stages of HIV-1 infection. The virus enters host cells by binding to CD4 T cells and co-receptors CCR5 or CXCR4. fusion occurs, followed by release of viral core into the cytoplasm. The viral genome, RNA is reverse transcribed (by reverse transcriptase) into double stranded proviral DNA which is integrated (by integrase) into the host cells genome. The virus utilizes the host cell's machineries to transcribe and translate the genome to yield viral RNA and polyproteins, which are then assembled into new immature viral particles. The viral particles will be cleaved by protease to form functional (infectious) proteins and released into circulation by budding. The boxes show the site of action of currently available antiretroviral drugs such as enfuvirtide and maraviroc inhibits the fusion and binding involves in the HIV-1 entry respectively, Tenofovir, abacavir, lamivudine, emtricitabine, efavirenz and nevirapine inhibits the activity of reverse trancriptase. Raltegravir inhibits the activity of integrase. Atazanavir, darunavir and ritonavir inhibits the activity of protease. Adapted from: (Volberding and Deeks, 2010).

2.7 Gene therapy for HIV-1 treatment

In 2009, the groundbreaking news of the first human cured of HIV-1, 'The Berlin Patient', was published. This patient was diagnosed with Acute Myeloid Leukemia (AML) 10 years after he was on HAART. After a few rounds of chemotherapy, the AML rebound in 2006 which required him to undergo hematopoietic stem cell transplantation (HSC) to survive. The Berlin patient obtained a matched donor who was naturally resistant to HIV-1 with the homozygous CCR5 delta32 mutation, rendering the CCR5 non-functional for the fusion and entry of HIV-1. This transplantation reconstituted the Berlin patient's immune system with CCR5 deficient cells, which caused prolonged-viral suppression, without being on HAART (Manjunath et al., 2013; Brown, 2015). Following similar procedure, another patient known as 'London patient' had an allogeneic HSC transplantation in the presence of opportunistic infection lymphoma. The treatment was reported to be successful as the plasma viral load remain undetectable up to 30 months of treatment and in the absence of HAART as of March 2020 (Gupta et al., 2020) . Thus, HSC transplantation gained tremendous attention, and many tried to replicate this study in hopes of achieving a global HIV-1 cure. However, engraftment efficiency, identification of matched donor with HIV-1 resistant gene, exposure to high risk of mortality due to graft versus host disease and immunosuppressive therapy have been major drawbacks in replicating this success (Moy et al., 2017; Vangelista and Vento, 2018).

The HIV-1 cure strategy was expanded to protein-based agents such as dominant–negative inhibitory proteins (uses a mutant form of HIV-1 Rev to block viral RNA export to cytoplasm) (Reddy and Dasgupta, 1992), intrabodies and intrakines (binds to viral proteins and sequester to proteasome for degradation) (Schroers et al.,

2002; Steinberger et al., 2000), fusion inhibitors (blocks viral fusion and entry into cell by binding to HIV gp41) (Lu et al., 2014) and viral restriction factors (host restriction factor such as TRIM5α which binds to HIV-1 capsid in the cytoplasm and interferes with un-coating process, thus protecting the cells from infection) (Shi and Aiken, 2006). Although these methods have antiviral effects *in vitro*, insufficient expression of the protein and *in vivo* immunogenicity are principle limitations of these approaches (Bunnell and Morgan, 1998; Sharma et al., 2014).

On the other hand, nucleic acid-based agents such as antisense RNA (which complementarily binds to HIV-1 and cleaves mRNA), aptamers (RNA molecules that binds to any target proteins with high affinity and renders it non-functional), RNA interference (small and double-stranded RNAs which modulate homology dependent degradation of target mRNA) and RNA decoys (RNA nucleotides mimics TAR sequence, which tricks Tat to bind to it rather than the original TAR region, thus inhibiting viral transcription) (Smythe and Symonds, 1995; Dorman and Lever, 2001; Zeller and Kumar, 2011). However, this approach activates the innate immune system, which triggers IFN α and β cytokines signalling that leads to activation of 2'-5' oligoadenylate synthase. This leads to non-specific degradation of cellular mRNA resulting in apoptosis (Hoxie and June, 2012; Scherer et al., 2007; Younan et al., 2013)

The need for an efficient and safe HIV cure caused researchers to divert their focus to genome editing tools (functions like a molecular scissors which can permanently edit any DNA or RNA sequence) as a potential alternative to combat and eradicate HIV-1.

2.7.1 Genome editing tools

The first generation of genome editing tool, meganuclease, was found in *Saccharomyces cerevisiae* as a homing endonuclease, and functions in splicing. The dimeric homing endonuclease recognizes 14-40bp of DNA sequences, therefore requires 6 to 9 amino acid residues to bind and cleave each strand of DNA, resulting in a double stranded break (DSB) (Paques and Duchateau, 2007; Silva et al., 2011) (Figure 2.8).

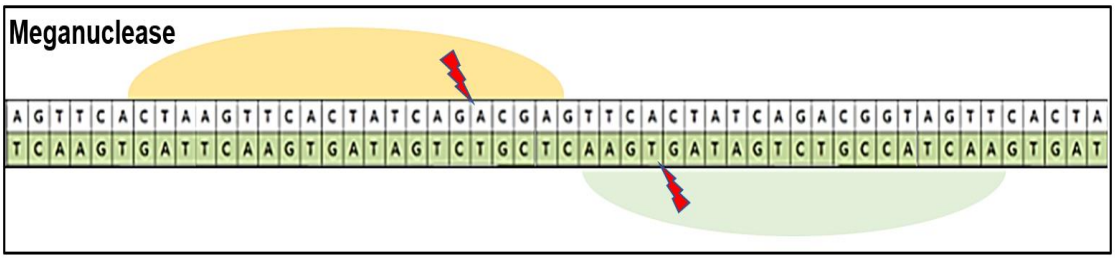


Figure 2.8 Schematic representation of meganucleases (MN). The MN dimers consisting 6-9 amino acid residues, binds and cleave 14-40 base pair DNA sequences to cause DNA double strand break (DSB).

The DSB will activate the cell repair mechanism - non-homologous end-joining (NHEJ) or homology directed repair (HDR). NHEJ fixes the broken DNA strands in an error-prone manner by introducing nucleotide insertions or deletions that would result in a frameshift mutation, rendering a non-functional protein. On the other hand, HDR repairs the DSB faithfully in the presence of a donor template, thus the resulting sequence will resemble the donor template sequence, and therefore making it suitable for base editing rather than knockout (Figure 2.9).

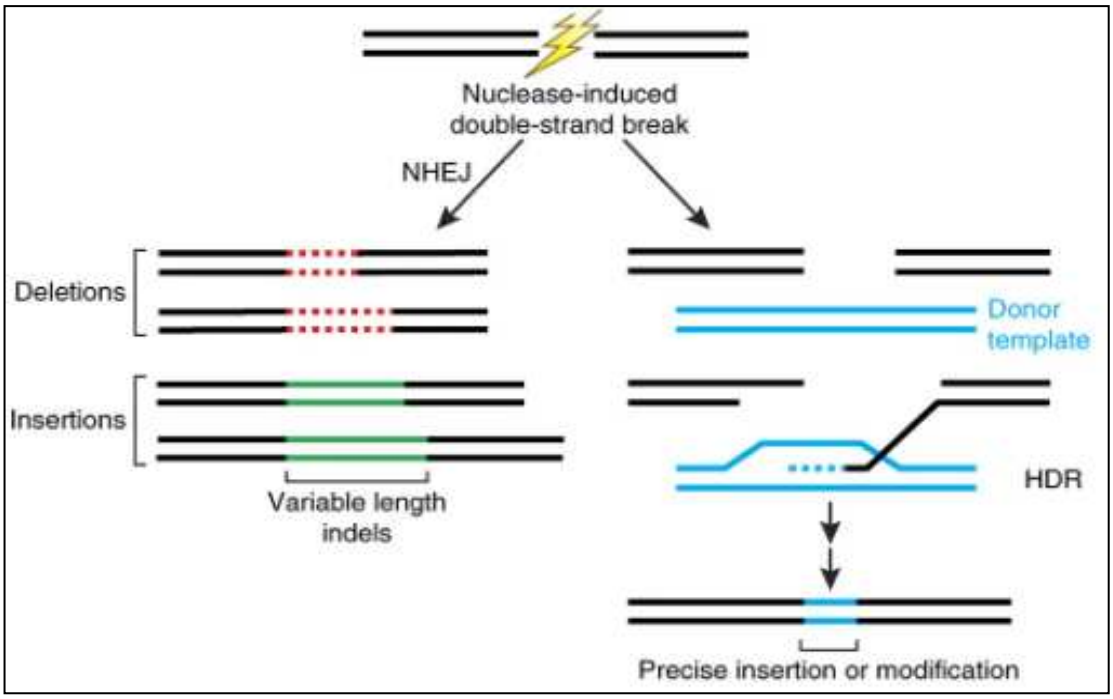


Figure 2.9 Mechanism of double strand break (DSB) repair. Non-homologous end joining (NHEJ) introduces insertions and deletions resulting in a disrupted sequence. Homology-directed repair (HDR) utilizes a template DNA and offers precise genome modification. Figure adapted from (Sander and Joung, 2014).

The engineered meganuclease, such as I-Sec I, worked efficiently in mammalian cell cultures (Haber et al., 1995; Smith et al., 1995; Poeut et al., 1994; Choulika et al., 1994). However, it was technically time consuming and difficult to construct meganuclease residues to target DNA sequences, as it involved several amino acid constructs to recognize 14-40bp long target DNA. This limits the application of meganuclease in clinical setting and causes scientist to investigate another genome editing tool, Zink Finger Nuclease (ZFN).

Zinc Finger Nuclease (ZFN), which was first discovered in transcription factors that bound to DNA in eukaryotes. ZFN consists of zinc finger proteins (ZFP), a DNA binding protein, and FokI, a nuclease domain with selective genomic cleavage. ZFP consists of 3-6 domains linked together, with each domain capable of recognizing a codon. Thus, each ZFN can identify between 9-18 nucleotide lengths of target DNA. FokI functions as an obligate heterodimer, where a pair of FokI binds to nucleotides on opposite strands of the DNA sequence (Figure 2.10). Dimerization of FokI allows genomic cleavage and causes a DSB (Figure 2.9) (Urnov et al., 2005; Chou et al., 2012).

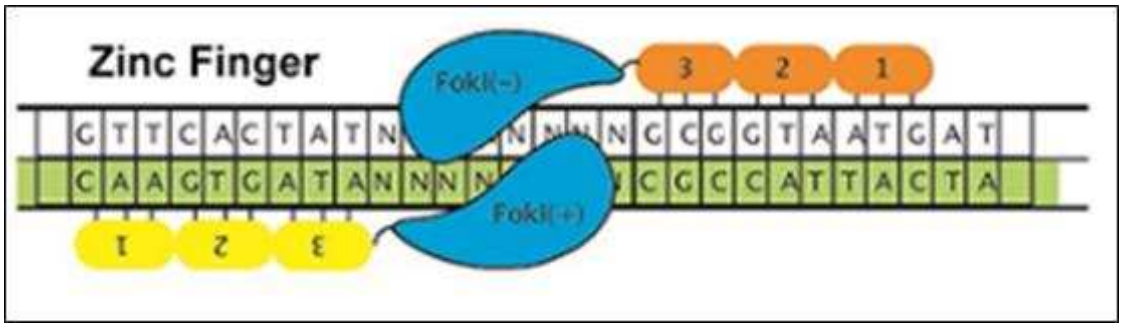


Figure 2.10 Schematic representation of zinc finger nucleases (ZFN). Each ZFN usually contains 3-6 zinc finger proteins (ZFP), which recognizes 3 base pair sequence each. When two halves of the FokI nuclease form a dimer, DNA cleavage occurs. Adapted from https://www.addgene.org/genome-engineering/.

The ZFN had been widely used to disrupt CCR5 expression in various *in-vitro* studies. It successfully entered phase 2 clinical trials modifying the CCR5 of autologous CD4 T cells for infusion into HIV-1 infected individuals (Perez et al., 2008; Manjunath et al., 2013; Hütter et al., 2015). However, ZFN is tedious to design, requires optimization (as it requires each domain to recognize 3bp nucleotide) and causes off-target activity (Kwarteng et al., 2017; Xiao et al., 2019).

The second generation of genome editing tool, TALENs, (Transcription Activation like Effector Nuclease) derived from a plant bacterial pathogen, Xanthomonas, which was used to alter transcription of genes in plants. TALEN involves the fusion of TALE repeats, the DNA binding domain, and FokI restriction enzyme, the cleavage domain. A pair of TALE units is required to bind to opposite sides of target DNA to allow FokI to dimerize and enable genome editing by inducing target DSB (Figure 2.11). The differences between ZFN and TALENs are each TALE unit binds to a single base pairs of target DNA, thus requires fifteen to twenty pair of TALENs to identify similar bp of target DNA, unlike ZFN where each domain binds to 3bp of DNA (Boch et al., 2009; Moscou and Bogdanove, 2009). Although this increases the specificity of TALENS, TALENs are relatively larger compared to ZFN, thus making it difficult to deliver them into the cells of interest using common viral vectors (Manjunath et al., 2013; Mussolino et al., 2014; Hütter et al., 2015). This limitation has shifted the attention from TALENs to CRISPR/Cas9.

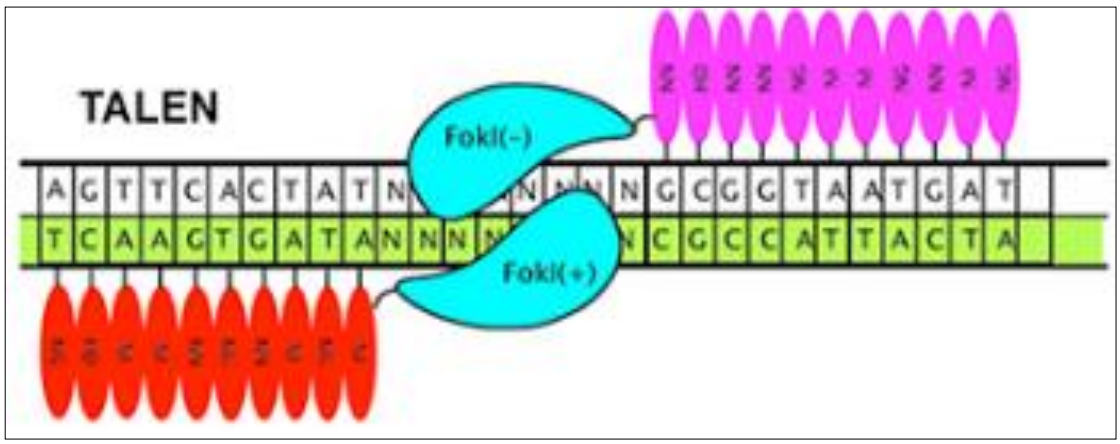


Figure 2.11 Schematic representation of TALENs. Each TALE unit identifies a single nucleotide and one TALEN usually identifies 14-20 nucleotides. Similar to ZFNs, TALENs are designed to bind to nucleotides on opposite sides of the target DNA to bring the Fok1 domains together to perform DNA cleavage. Adapted from https://www.addgene.org/genome-engineering/.

2.7.1(a) CRISPR/Cas9

CRISPR, which stands for Cluster Regularly Interspaced Short Palindromic Repeats, is an adaptive immune system found in prokaryotes. The CRISPR pathway involves capturing a short segment of foreign DNA immediately proceeded by a protospacer adjacent motif sequence (PAM) usually, NGG or NCC (to differentiate between the self and foreign DNA) and incorporates them into CRISPR locus as spacers. Then, cells will transcribe this spacer to form a CRISPR RNA (cRNA) which has a complementary sequence to the invading foreign DNA, thus creating a memory. The constitutively expressed trans-activating RNA (tracRNA) will interact with cRNA to form guide RNA (gRNA), The gRNA is responsible for chaperoning the Cas9 to form an active gRNA/Cas9 complex which works like an antibody. Upon encountering similar foreign DNA again, gRNA will bind to it, marking it as a target for Cas9 to cleave, resulting in an inactivated foreign DNA (Burmistrz and Pyrć, 2015; Rath et al., 2015; Hille and Charpentier, 2016) (Figure 2.12).