

# **Investigating the fitness benefit of reverse transcriptase (RT) mutation A62V when co-occurring with M184V and K65R in HIV-1 subtype C**

by

**Duncan Tazvinzwa Njenda**

*Thesis presented in fulfilment of the requirements for the degree of Masters of Sciences in Medical sciences (Medical Virology) Faculty of Medicine and Health Sciences at Stellenbosch University*



Supervisor: Professor Gert Uves Van Zyl  
Co-supervisor: Professor Susan Engelbrecht  
Co-supervisor: Dr Graeme Brendon Jacobs

March 2016

## Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own original work, that I am the authorship owner thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: Duncan Tazvinzwa Njenda

Date: March 2016

## Abstract

### Background and Aims

Tenofovir disoproxil fumarate (TDF) and lamivudine (3TC) or emtricitabine (FTC) combined with efavirenz is the predominant first-line antiretroviral regimen in the Southern African region. Resistance to TDF and 3TC/FTC is largely through the occurrence of the drug resistance mutations (DRMs) K65R and M184V, respectively. Preliminary data from a large laboratory-based dataset of HIV drug resistance that showed a high prevalence of these mutations in patients who received the TDF regimen also revealed a significant co-occurrence of A62V with M184V and K65R. The aim of this study was to investigate the functional interaction and effect on viral fitness that A62V has when it co-occurs with M184V and K65R reverse transcriptase mutations in HIV-1 Subtype C.

### Materials and Methods

Using Infusion™ cloning and site-directed mutagenesis techniques, eight full-length genome infectious clones containing the HIV-1 subtype C polymerase gene were synthesised having all combinations of DRMs - A62V, M184V and K65R, either being present or absent. The mutations in these constructs were verified by sequencing. The constructs were transfected into 293T cells for virion production and harvested virus was infected in the TZM-bl cell line in head to head growth competition experiments and assayed for growth kinetics using an allele-specific quantitative real-time polymerase chain reaction (PCR) assay.

### Results

The growth competition experiment between two viruses (A62V+K65R+M184V vs K65R+M184V) evaluated by taking the mean of 3 biological replicates in the assay in the absence of antiretroviral drugs, revealed that A62V mutation has no significant impact on fitness (Wilcoxon signed rank test  $p$ -value = 0.56). The overall coefficient of variation (CV) in the experiment was 12.8% indicating the high reproducibility of the growth competition assay using real-time PCR measurement of relative growth.

### Conclusion and recommendations

A62V mutation has no effect on fitness when it co-occurs with M184V and K65R. The co-occurrence with M184V and K65R remains unexplained and might be due to an effect on TDF resistance in combination with K65R. This requires investigation in future studies as TDF regimens are part of 1<sup>st</sup> line therapy in many Sub-Saharan countries in the treatment of HIV-1. Finally, the cloning and mutagenesis techniques used coupled with a very sensitive and reproducible real-time

quantitative PCR assay provide an efficient system for detection of mutation fitness interactions and can be used in any future work to study HIV mutation fitness interactions.

## Opsomming

### Agtergrond en doel

Tenofovir disoproksil fumaraat (TDF) en lamivudien (3TC) of emitricitabine (FTC) gekombineer met efavirenz is die oorheersende eerste lyn antiretrovirale regimen in die Suider-Afrikaanse streek. Weerstand teen TDF en 3TC / FTC is grootliks deur die voorkoms van die middel--weerstandigheid-mutasies (DRM's) K65R en M184V, onderskeidelik. Voorlopige data van 'n groot laboratorium- gebaseerde dataset van MIV middel-weerstand het bewys dat daar 'n hoë voorkoms van hierdie mutasies is in pasiënte wat die tenofovir regimen ontvang en het ook 'n beduidende mede-voorkoms van A62V met M184V en K65R onthul. Die doel van hierdie studie was om die funksionele interaksie en invloed op virale fiksheid te ondersoek wat A62V het wanneer dit voorkom saam met M184V en K65R trutranskriptase mutasies in MIV-1 Subtype C.

### Materiale en metodes

Infusion™ kloning en mutagenese tegnieke is gebruik om 8 vollengte MIV rekombinante klone te verkry met alle kombinasies van DRM's, A62V, M184V en K65R, hetsy teenwoordig of afwesig en die mutasies in hierdie konstrakte is bevestig deur nukleinsuurbasisbaarloosvolgordebepaling. Die konstrakte was getransfekteer in 293T selle vir virion produksie. Virus supernatant is geoes en die TZM-bl sellyn is geïnfecteer en gebruik in kop-aan-kop groei kompetisie eksperimente; virusproduksie is getoets met behulp van 'n alleel-spesifieke in-huis kwantitatiewe reële-tyd polimerase kettingreaksie (PKR) om die groei kinetika te ondersoek.

### Resultate

Die groei kompetisie eksperiment tussen twee virusse (A62V + K65R + M184V vs K65R + M184V) wat geëvalueer was deur die neem van die gemiddelde van 3 biologiese herhalings in die teenwoordigheid van anti-retrovirale middels, het getoon dat A62V mutasie het geen beduidende impak op fiksheid ("Wilcoxon signed rank" toets p-waarde = 0.56). Die algehele koëffisiënt van variasie (KV) in die eksperiment was 12,8% wat aandui op die hoë herhaalbaarheid van die groei-kompetisie eksperiment met reële tyd PKR kwantifisering van relatiewe groeikoers.

### Gevolgtrekking en aanbevelings

A62V mutasie het geen effek op fiksheid wanneer dit voorkom met M184V en K65R nie. Die gesamentlike voorkoms van A62V met M184V en K65R bly onverklaarbaar, maar dit mag weens verhoogde weerstandigheid teen TDF wees wanneer A62V met K65R voorkom. Om dit te verklaar benodig verdere studies aangesien TDF- gebaseerde kombinasie-terapie in baie Sub-Saharan lande gebruik word vir 1ste-lyn behandeling van MIV-1. Ten slotte, die kloning en mutageniese

tegnieke, wat gebruik is, tesame met 'n baie sensitiewe en herhaalbare reële-tyd kwantitatiewe PKR toets verleen 'n doeltreffende benadering om fiksheids interaksies te ondersoek.

## Acknowledgements

I have encountered a lot of challenges working on this project but I believe the total of all the obstacles I have faced, have enhanced my knowledge and skills. The same challenges have also allowed me to become a more mature scientist. In particular, I would like to acknowledge the following individuals and organisations for their contributions and support towards the completion of this thesis:

I am grateful to my supervisor **Professor Gert Uves Van Zyl** for providing me with an opportunity to learn and to develop skills that are essential in doing research. His knowledge, mentorship and supervision have all been of great value for the success of this research project and thesis. I have no regrets working under his supervision and I am delighted to have been part of his research group.

I wish to thank **Professor Susan Engelbrecht** and **Dr Graeme Brendon Jacobs** for their invaluable contribution to this project as dedicated co-supervisors. Your advice, knowledge, practical assistance and experience in the lab, form the cornerstone of the development of my practical skills as an upcoming researcher. Words alone could never express how I cherish the support, humility and patience you have had for me in your instruction and even in you listening. I am forever truly grateful and thankful for your support.

I wish to thank all the staff in the Division of Medical Virology, thank you for your moral support throughout the duration of my master's degree. To the following people, especially, **Hilory Munhuweyi, Tongai Maponga, Tatum Lopes, Shaheida Isaacs, Mary-Grace Katusiime, Don Matshazi, Bronwyn Kleinhans, Tanya Kerr, Heleen Loots, Ndapewa Ithete, Amanda Moleich(National Health Laboratory Service- NHLS) and Randall Fisher**, I mindful to have met you and cherish your assistance and friendship dearly. To **Hilory**, your untimely passing away deeply saddened me. You were a brilliant student with a lot of potential and it's unfair for you to have left so soon and now you cannot observe the outcome of your love, support and encouragement towards me. Your friendship will forever be a memorial in my heart.

I would also want to sincerely thank the following organisations for funding: **Poliomyelitis Research Foundation** and the **Harry Crossley Foundation**. Without their support this project would not have taken off. I am truly thankful and appreciate the support and opportunity to advance research.

Finally, I would like to sincerely thank my Family. Special mention to my **Mum (Lydia Zhanje)** you have been a pillar in my darkest times and an unwavering sparkle of light to guide and love me. To **Emily**, my beautiful wife, your love and support for me is a blessing to my life that I will forever hold in my heart.

To my wonderful **God**, the radiance of your love and Spirit is life and living reality that I am forever grateful to be part of and have. I am truly grateful for the love, support and friendship I received in development of my Masters thesis and for the pursuit of my studies.

*“The meaning of life is adaptability the value of life is its progressability...”*

***Urantia papers, paper 130: 4.7***

## Presentation at Meetings

- **Njenda D.T, Jacobs G.B, Engelbrecht S, Van Zyl G. U.** Investigating the fitness of reverse transcriptase(RT) A62V mutation when co-occurring with M184V and K65R in HIV-1 Subtype C. **58 annual Academic day, 2014, Faculty of Medicine and Health Sciences, University of Stellenbosch, Cape Town, South Africa - Poster presentation.**
  
- **Njenda D.T, Jacobs G.B, Engelbrecht S, Van Zyl G. U.** Investigating the fitness of reverse transcriptase(RT) A62V mutation when co-occurring with M184V and K65R in HIV-1 Subtype C. **18<sup>th</sup> International conference on AIDS and STIs (ICASA), 2015, Harare, Zimbabwe – Oral presentation**

## Table of contents

Descriptor	Content	Page
	Thesis title	
	Declaration	I
	Abstract	II-III
	Opsomming	IV-V
	Acknowledgments	VI-VII
	Presentation at meetings	VIII
	Table of Contents	X-XI
	Table of Figures	XII
	Table of Tables and Table of Plates	XIII
	List of Abbreviations	XIV
<b>Literature review</b>	<b>Chapter 1 – Introduction</b>	
<b>1.1</b>	Introduction to HIV drug resistance and viral fitness	<b>1-2</b>
<b>1.2</b>	HIV structure, replication and pathogenesis	<b>3</b>
1.2.1	HIV structure and replication	<b>3-5</b>
1.2.2	HIV target cell infection and pathogenesis	<b>6</b>
1.2.3	HIV transmission routes	<b>7</b>
<b>1.3</b>	Antiretroviral therapy and drug resistance	<b>8-9</b>
1.3.1	HIV drug resistance mechanisms, evolution and viral fitness	<b>10</b>
1.3.2.	Molecular mechanisms of Nucleos(t)ide <i>reverse transcriptase</i> Resistance	<b>10-11</b>
1.3.3	Evolution of acquired antiretroviral resistance (therapy-induced) and transmitted antiretroviral drug resistance	<b>12</b>
1.3.4	Fitness and the role of secondary (compensatory or accessory mutations) in virus evolution	<b>13</b>
<b>1.4</b>	Antiretroviral resistance and fitness testing and <i>ex-vivo</i> mutagenesis	<b>14</b>
1.4.1	Antiretroviral resistance assays: genotypic and phenotypic resistance assays	<b>14</b>
1.4.2	Fitness assays	<b>15-16</b>
1.4.3	<i>Ex-vivo</i> mutagenesis assays: Methods and research/clinical usefulness	<b>17-19</b>

1.4.4	Preliminary Data for Study	19-20
	<b>Research Hypothesis and Aims of study</b>	<b>21</b>
	<b>Chapter 2 – Materials and Methods</b>	
	<b>Introduction</b>	<b>22-23</b>
<b>2.1</b>	Ethical considerations	<b>24</b>
<b>2.2</b>	<b>Section 1</b> -DNA amplification and cloning of HIV-1 <i>gag</i> -->start of <i>env</i> gene	<b>24</b>
2.2.1	DNA amplification of HIV-1 <i>gag</i> -->start of <i>env</i> gene	<b>24</b>
2.2.2	Cloning of HIV-1 <i>gag</i> -->start of <i>env</i> gene	<b>25-26</b>
<b>2.3</b>	<b>Section 2</b> –Site Directed Mutagenesis, HIV-1C vector synthesis and cell culture	<b>28</b>
2.3.1	Site-directed mutagenesis using methylation screening	<b>28-29</b>
2.3.2	Site-directed mutagenesis by Overlap-extension using PCR	<b>30</b>
<b>2.4</b>	HIV-1 subtype C (TV1_FLG) vector back bone construction	<b>31-33</b>
<b>2.5</b>	Cell culture methods to test infectivity of TV1-FLG	<b>34</b>
2.5.1	Seeding cells and Transfection	<b>34-35</b>
2.5.2	Cell culture end-point assays to check infectivity of HIV-1C clone (TV1_FLG)	<b>36-37</b>
<b>2.6</b>	<b>Section 3</b> –Alternative backbone to assess mutations in HIV-1 subtype B reporter Plasmid (NL4.3 eGFP)	<b>38</b>
2.6.1	Synthesis of recombinant constructs using infusion™ cloning method	<b>38-40</b>
2.6.2	Synthesis of mutants by Q5™ Mutagenesis protocol	<b>41</b>
2.6.3	Assays to test infectivity and quantification of replication-competent virus	<b>43</b>
2.6.3.1	Fluorescence imaging assay	<b>42</b>
2.6.3.2	TCID <sub>50</sub> end point titration assays ( <b>A62V+K65R+M184V</b> and <b>K65R+M184V</b> )	<b>43-44</b>
<b>2.6.4</b>	Allele-specific real-time PCR assay design	<b>45-46</b>
<b>2.6.5</b>	Growth competition experiment ( <b>A62V+K65R+M184V</b> vs <b>K65R+M184V</b> )	<b>47-48</b>

	<b>Chapter 3 – Results</b>	<b>49</b>
	<b>Introduction</b>	<b>49</b>
<b>3.1</b>	Sequencing results for TV1-2kb and TV1-2.5kb wild-type and mutant subclones	<b>49-53</b>
<b>3.2</b>	TV1-FLG results	<b>54</b>
3.2.1	Ligation and restriction enzyme digestion results	<b>54</b>
3.2.2	Sequencing results of TV1- FLG	<b>55-56</b>
3.2.3	HIV <i>gag</i> p24 ELISA and <i>X-gal</i> staining results	<b>57-58</b>
<b>3.3</b>	Results from plan to subclone in NL4.3 Subtype B infectious Plasmid	<b>59</b>
3.3.1	Infusion™ cloning and Q5™ mutagenesis results	<b>59-61</b>
3.3.2	Fluorescence imaging results	<b>62</b>
3.3.3	TCID <sub>50</sub> end point titration results for <b>A62V+K65R+M184V</b> and <b>K65R+M184V</b>	<b>63</b>
3.3.4	Allele-specific real-time PCR optimization results	<b>64-66</b>
3.3.5	Growth competition experiment results for <b>A62V+K65R+M184V</b> vs <b>K65R+M184V</b>	<b>67-69</b>
	<b>Chapter 4 – Discussion</b>	
<b>4.1</b>	Summary of main findings	<b>70</b>
<b>4.2</b>	Other important findings/ Technical lessons learned	<b>71-72</b>
<b>4.3</b>	Strengths and Limitations of study	<b>73</b>
<b>4.4</b>	Further investigations to consider for future studies	<b>74</b>
	<b>Chapter 5 – Conclusion</b>	<b>75</b>
	<b>Addenda</b>	<b>75-87</b>
	<b>References</b>	<b>88-99</b>

## Table of Figures

<b>Figure #</b>	<b>Content</b>	<b>Chapter(CH) →sub-title</b>	<b>Page</b>
Figure 1a, 1b	HIV genomic and particle structure	Ch1 – 1.2.1	<b>5</b>
Figure 2	Summary of fitness assays	Ch1 – 1.4.2	<b>16</b>
Figure 3	Preliminary Data summary	Ch1 – 1.4.3	<b>20</b>
Figure 4	Overview of Methodology for study	Ch2 – Introduction	<b>22</b>
Figure 5	Summary of Initial cloning strategy for study	Ch2 – Introduction	<b>23</b>
Figure 6	PCR amplification of HIV-1 RT gene(TV1 -4.5kb)	Ch2 – 2.2.1	<b>24</b>
Figure 7	Restriction enzyme digest of (TV1 -4.5kb)	Ch2 – 2.2.2	<b>28</b>
Figure 8a, 8b	TV1_S-clone and MJ4/TV1 plasmid maps	Ch2 – 2.4	<b>32-33</b>
Figure 9	Automated cell counter Analysis	Ch2 – 2.5.1	<b>35</b>
Figure 10	New cloning strategy (Infusion™ Cloning)	Ch2 – 2.61	<b>39</b>
Figure 11	TCID <sub>50</sub> end-point titration assay plate layout	Ch2 – 2.6.3.2	<b>44</b>
Figure 12	Overview of allele-specific real-time qPCR design	Ch2 – 2.6.4	<b>45</b>
Figure 13	Sequence results for TV1-2kb	Ch3 – 3.1	<b>50</b>
Figure 14	Sequence results for TV1-2.5kb	Ch3 – 3.1	<b>51</b>
Figure 15	Site-directed mutagenesis results	Ch3 – 3.1	<b>52</b>
Figure 16	Site-directed mutagenesis results	Ch3 – 3.1	<b>53</b>
Figure 17a,17b	Restriction enzyme digest results (TV1-FLG)	Ch3 – 3.2.1	<b>54</b>
Figure 18	Sequencing primer binding positions in TV1-FLG	Ch3 – 3.2.1	<b>55</b>
Figure 19	Sequence results for TV1-FLG	Ch3 – 3.2.1	<b>56</b>
Figure 20	Sequence results for TV1-FLG	Ch3 – 3.2.1	<b>57</b>
Figure 21	ELISA results for TV1-FLG	Ch3 – 3.2.2	<b>57</b>
Figure22	X-gal staining results for TV1-FLG	Ch3 – 3.2.1	<b>58</b>
Figure 23	Infusion™ cloning and Q5™ mutagenesis results	Ch3 – 3.3.1	<b>59</b>
Figure 24a-d	Sequence results (A62V+K65R+M184V; K65R+M184V)	Ch3 – 3.3.1	<b>60-61</b>
Figure 25	Fluorescence imaging results	Ch3 – 3.3.2	<b>62</b>
Figure 26a-d	qPCR results	Ch3 – 3.3.4	<b>64-67</b>
Figure 27	Growth competition results (A62V+K65R+M184V vs K65R+M184V)	Ch3 – 3.3.5	<b>68</b>

Figure 28	Fitness analysis results	Ch3 – 3.3.5	69
-----------	--------------------------	-------------	----

## Table of Tables

Table	Content	Chapter/ Heading	Page
<b>Table A</b>	Summary of FDA approved ARVs	Ch1 – 1.3	9
<b>Table B</b>	Go- <i>Taq</i> long-range protocol	Ch2 – 2.2.1	25
<b>Table C</b>	Sequencing primers for TV1_2/2.5kb	Ch2 – 2.2.2	27
<b>Table D</b>	SDM mutagenesis primers	Ch2 – 2.3.1	29
<b>Table E</b>	SDM Overlap-extension PCR primers	Ch2 – 2.3.2	31
<b>Table F</b>	X-Gal staining solution	Ch2 – 2.6.1	40
<b>Table G</b>	Infusion™ cloning primers	Ch2 – 2.6.1	40
<b>Table H</b>	Infusion™ cloning protocol	Ch2 – 2.6.1	40
<b>Table I</b>	Q5™ Mutagenesis primers	Ch2 – 2.6.2	42
<b>Table J</b>	H2H growth competition graphical data	Ch3 – 3.3.5	68
<b>Table K</b>	H2H growth competition raw experimental data	Ch3 – 3.3.5	70

## Table of Plates

Plate	Content	Chapter/Heading	Page
<b>Plate A</b>	TCID <sub>50</sub> titration results for A62V+K65R+M184V virus	Ch3 – 3.3.3	63
<b>Plate B</b>	TCID <sub>50</sub> titration results for K65R+M184V virus	Ch3 – 3.3.3	63

## List of Abbreviations

**ABC** – Abacavir

**AIDS** – Acquired immunodeficiency Syndrome

**ART** – Antiretroviral therapy

**ARV** – Antiretroviral

**AZT** – azidothymidine (Zidovudine)

**cART** – Combination antiretroviral therapy

**ddi** – Didanosine

**DNA** – Deoxyribonucleic acid

**DRM** – Drug resistance mutation

**DRMV** – Drug resistant minority variant

**DRV** – Drug resistant variant

**dsDNA** – Double stranded DNA

**EFV** – Efavirenz

**eGFP** – enhanced Green fluorescent protein

**FTC** – Emtricitabine

**FLG** – Full-length genome

**HIV** – Human immunodeficiency virus

**HIV-1C** – Human immunodeficiency virus type 1 Subtype C

**HIV-1BC** – Human immunodeficiency virus type1 BC recombinant

**HRP** – Horseradish peroxidase

**INSTI** – Integrase strand transfer inhibitor

**MDR** – Multi-class drug resistance

**MTCT** – Mother to child transmission

**NF- $\kappa$ B** – Nuclear factor-kappa Beta

**NRTI** – Nucleos(t)ide reverse transcriptase inhibitor

**NNRTI** – Non-Nucleoside reverse transcriptase inhibitor

**NVP** – Nevirapine

**PIC** – pre-integration complex

**PI** – Protease Inhibitor

**qPCR** – quantitative real-time Polymerase Chain Reaction

**RAM** – Resistant associated mutation(s)

**RNA** – Ribonucleic acid

**RT** – *Reverse transcriptase*

**RRE** – Rev responsive element

**ssDNA** – single stranded DNA

**TAM** – Thymidine analogue mutation(s)

**Tat** – Transcriptional transactivator gene

**TCID<sub>50</sub>** – Tissue culture infectious dose value of a defined inoculum of a pathogenic agent that will produce a pathological change in 50% of cell cultures inoculated

**TDF** – Tenofovir disoproxil fumarate

**3TC** – Lamivudine

**WT** – Wild-type

## Literature Review

### Chapter 1- Introduction

#### **1.1 Introduction to HIV drug resistance and viral fitness**

HIV (Human Immunodeficiency virus) the causative agent of Acquired Immunodeficiency Syndrome (AIDS) undergoes rapid mutation of its viral genome (Mansky and Temin, 1995). The high mutation rate, a function of the lack of proofreading of HIV reverse transcriptase and high replication rate, contributes to the evolution of drug resistance mutations (DRMs). Viral species harbouring DRMs (referred to as drug-resistant variants (DRVs) throughout this review) are often found in infected patients with elevated viral loads while receiving combination antiretroviral therapy (cART). The emergence of drug resistance in patients is multifactorial and encompasses pharmacological, host and adherence factors (Cortez and Malderelli, 2011).

Before the availability of cART regimens, HIV treatment was based on monotherapy (Fischl *et al.*, 1987). Zidovudine (AZT) was the first licenced antiretroviral drug and was Federal Drug Authority (FDA) approved in 1987 (P.H.S., 1987). As monotherapy, it did not result in sustained therapy success and in 1991/1992 antiretroviral resistance was shown to be the cause of failure (Wainberg *et al.*, 1991). Only when cART became the standard of care in 1996/1997, therapy achieved long-term success in the majority of patients (Egger *et al.*, 1997; Gulick *et al.*, 1997). This was initially referred to as highly active antiretroviral therapy (HAART) and consisted of a combination of at least three antiretroviral drugs, either two nucleoside *reverse transcriptase* inhibitors (NRTIs) that were analogues of different native nucleosides combined with either a protease inhibitor or a non-nucleoside *reverse transcriptase* inhibitor (NNRTI) (Arts *et al.*, 2012). There are several classes of antiretroviral (ARV) drugs which target various stages of HIV-1 replication: nucleoside *reverse transcriptase* inhibitors (NRTIs); non-nucleoside *reverse transcriptase* inhibitors (NNRTIs); protease inhibitors (PIs); entry inhibitors; inhibitors of co-receptors and integrase inhibitors (Chen *et al.*, 2007).

In South Africa, where the dominant strain of the virus is HIV-1 subtype C, an estimated 6.8 million individuals are infected with the virus. HIV infected individuals in South Africa represent approximately 20% of the number infected worldwide and approximately 28% of the number of people living with the virus in Sub-Saharan Africa (UNAIDS, 2014a). South Africa started its national ARV rollout program in 2004 (D.o.H., 2004) and presently, an estimate of 50% of the number of individuals in South Africa living with HIV are on ARVs (UNAIDS, 2014b).

The current standardized regimens for 1<sup>st</sup> line therapy for adults and adolescents include tenofovir disoproxil fumarate (TDF), lamivudine (3TC) or emtricitabine (FTC), and efavirenz (EFV) or nevirapine (NVP) (D.o.H., 2010). Despite the success of cART in arresting viral replication to undetectable levels (below 50 copies/ml as detected by viral load assays) and conferring partial immune reconstitution (Autran *et al.*, 1997), development of drug resistance is likely due to a low genetic barrier of the 1<sup>st</sup> line therapy regimen.

South Africa as a low to middle-income country faces clinical care and laboratory monitoring of more than three million people on cART, and the situation places the country with an enormous service-delivery challenge. Continued therapy success is dependent on continuous drug supply, patient-friendly health services and early detection of the failure in patients with inadequate adherence to prevent drug resistance (Stadel and Richman, 2013). However, in the case of systemic health care failure or individual non-adherence, patients who have virological failure are at high risk of development of DRMs (Machado *et al.*, 2009; Hosseinipour *et al.*, 2009; Neogi *et al.*, 2011).

The nomenclature for describing any particular DRM is according to the International Union of Pure and Applied Chemistry (IUPAC). Mutations are numbered according to the amino acid position within the relevant protein, with the native amino acid abbreviation preceding the number and the mutated amino acid following the number. For example, M184V indicates a substitution of Methionine (M) with Valine (V) at position 184 of HIV *reverse transcriptase* (RT) enzyme (Stott *et al.*, 2012).

DRMs are usually non-polymorphic (rarely observed in untreated patients). DRMs that are present could either be a result of transmission from one individual to another individual (transmitted drug resistance) or be acquired as a result of drug resistance selection or HIV evolution during therapy failure (acquired drug resistance) (Pennings, 2013 Kuritzkes, 2004). In Sub-Saharan Africa, acquired drug-resistance ranges from 53%-90% for patients failing 1<sup>st</sup> line therapy within the first 12 months after initiation of ARV therapy (Gupta *et al.*, 2012) compared with a prevalence of approximately 7-17% of transmitted drug resistance (Pennings, 2013). Patients who fail therapy with acquired drug-resistance usually have dual-class resistance whereas multi-class drug resistance (MDR) to NRTIs, NNRTI and PIs is still relatively rare in Sub-Saharan Africa (Pennings, 2013).

## 1.2 HIV-1 structure, replication and Pathogenesis

A brief review of the structure, replication and pathogenesis of HIV is given below. This section is followed by a section reviewing concepts related to antiretroviral therapy, mechanism of HIV drug resistance, measures for viral fitness and drug resistance testing.

### 1.2.1 HIV-1 Structure and replication

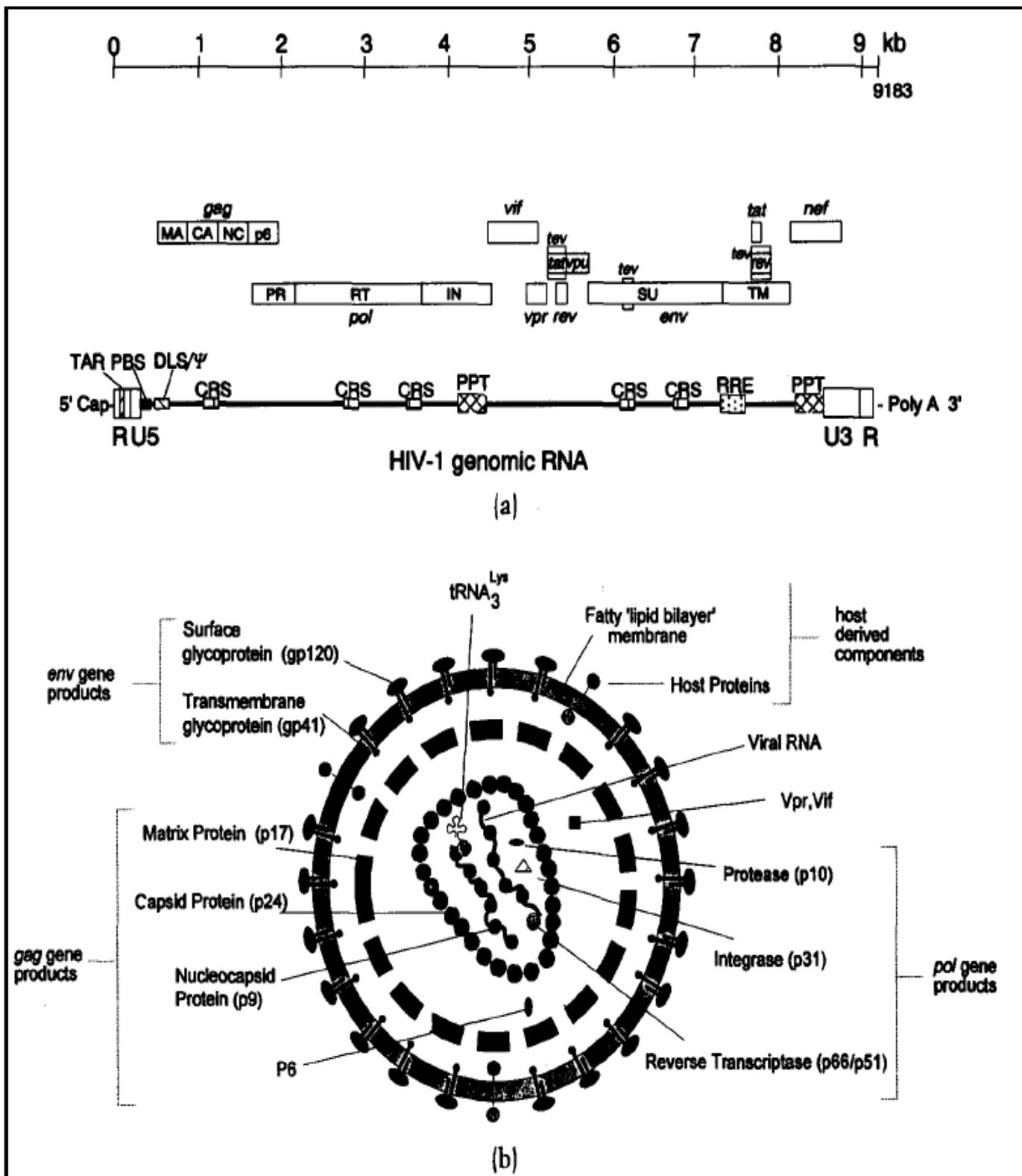
HIV-1 is classified under the *Retroviridae* family and sub-classified under the Lentivirus genus. There are two HIV species, HIV-1 and HIV-2. HIV-1 is divided into four groups M (Main), O (Outlier), N (New) and group P (Vallari *et al.*, 2009). For HIV-2, there are eight groups (A-H) (Etienne *et al.*, 2011). Group M genetic diversity has been well characterised and includes nine subtypes (designated A - D, F - H and J - K), an estimated 54 circulating recombinant forms (CRFs) and emerging Unique recombinant forms (URFs) (Parczewski, 2013)

HIV-1 mature virions are approximately 100-120nm in diameter and have a spiked lipid bilayer membrane, which includes embedded structural envelope proteins: gp120 and gp41. Internal structural proteins include the matrix protein (p17); the capsid protein (p24) and a nucleocapsid protein (p7). Enclosed in the capsid are virus-encoded enzymes; protease, a homodimer; *reverse transcriptase*, a heterodimer and integrase, all essential for viral replication; and two copies of its ~9.2kb positive-stranded RNA genome (Sierra *et al.*, 2005). HIV's genomic features are a common characteristic of most lentiviruses and these features (**Figure 1a**), include structural genes: the group antigen (*gag*), the envelope (*env*) and the polymerase (*pol*) gene. Non-structural regulatory genes include *tat* (trans-activator of transcription); *rev* (regulator of viral expression) and non-structural accessory genes include: *Nef* (negative factor); *vpr* (viral protein r) and *vpu* (viral protein u). The HIV genome also includes non-coding sequences (**Figure 1a**): The Unique 5' (U5), Unique 3' (U3) and Repeat (R) regions which collectively define the 5' and 3' long terminal repeats (LTRs) at the ends of the viral genome (Joshi and Joshi, 1996). Other non-coding sequences segments include the primer binding site (PBS), transactivation response element (TAR), packaging signal ( $\Psi$ ), and cis-acting repressive sequences (CRSs), polypurine tract (PPT) (Wurtzer *et al.*, 2005), Rev-responsive element (RRE) and the dimer linkage structure (DLS) (Paillart *et al.*, 1996).

HIV-1 replication cycle initiates after the concurrent binding of the viral envelope gp120 trimer with the CD4 and CCR5/ CXCR4 cell receptors (Michael, 1999). The CCR5/CXCR4 functions as a co-receptor in the binding of the virus allowing the viral gp41 transmembrane to fuse the viral envelope with the membrane of the cell in a pH-dependent manner. An invagination reaction then follows leading to the subsequent docking of the nucleocapsid into the cytoplasm and uncoating of the viral genome. Uncoating occurs under the influence of cellular factors (Sierra *et al.*, 2005; Joshi and Joshi 1996).

The viral genome is reverse transcribed by the HIV RT enzyme into double-stranded DNA (dsDNA) in a series of strand displacement reactions that creates an overlap termed central DNA flap (Zennou *et al.*, 2000) on one of the DNA strands which is essential for nuclear import. The central DNA flap facilitates formation of the pre-integration complex (PIC) which includes *vpr*, HIV integrase, matrix protein and host proteins (Zennou *et al.*, 2000). *Vpr* promotes the nuclear importation of the PIC via the nuclear pore and the viral DNA is preferentially integrated into active cellular genes by HIV integrase (Schroder *et al.*, 2002). Integrated proviral DNA is transcribed by cellular DNA-dependent RNA polymerase II. Unintegrated proviral DNA can exist within the cell nucleus as 2-LTR circular, 1-LTR circular and linear forms of DNA and provide the initial expression of *tat*, *nef* and *rev* proteins until their full expression by integrated proviral DNA (Wu, 2004).

The expression of *tat* coupled with the presence of transcription factors such as NF- $\kappa$ B, EBP-1, TFIIIP and SP1 within the host cell, initiates downstream expression of the integrated viral genome from the 5' LTR end and lead to the expression of viral enzymes and proteins (Joshi and Joshi, 1996). However, in the absence of transcription factors, HIV establishes latency (Shan *et al.*, 2012; Eisele and Siliciano, 2012) which contributes to the persistence of viral strains carrying DRMs harboured in long-lived cells such as CD4 memory T lymphocytes (Volberding and Deeks, 2010). If transcription of the proviral DNA occurs, three classes of mRNA transcripts, multiply spliced, singly spliced and full-length transcripts, are synthesized (Wu, 2004). Briefly, the *gag*, *pol* and *env* mRNA transcripts are singly spliced mRNA transcripts which when translated give rise to the *gag* precursor (p55), *gag-pol* precursor (p160) and membrane precursor (gp160) proteins respectively but are preceded by singly spliced *nef* and multiply spliced *tat*, *rev*, *vpr* and *vpu* mRNA transcripts (Joshi and Joshi, 1996). Increased *rev* expression later in the viral life cycle results in increased nuclear export of singly spliced and full-length mRNA transcripts into the cytoplasm of the cell, where ribosomal translation of mRNA to proteins take place (Sierra *et al.*, 2005). *Gag* and *env* derived structural proteins then undergo post-translation modification via proteolytic cleavage with protease and are eventually assembled to form the structural components of the virion particle (**Figure 1b**). The virion is also packaged with its primer (tRNA<sub>3</sub><sup>Lys</sup>) and two copies of its RNA genome and *gag-pol* derived viral enzymes (**Figure 1b**). The virion is subsequently enclosed in a vesicle and transported to the cell membrane where it buds off (deriving its lipid bilayer membrane from the host's membrane) to begin another round of infection. The membrane is embedded with *env* derived proteins that include gp41 transmembrane glycoprotein stalks and gp120 surface protein (**Figure 1b**). The gp41 stalks are estimated to be 72 in total per particle (Gelderblom *et al.*, 1987).



**Figure 1- (a) HIV genomic structure. (b) HIV cross-sectional map of fully assembled infectious viral protein (Photo adapted and reproduced with permission (see **Addendum A1**) from Joshi and Joshi, 1996)**

### **1.2.2 HIV-1 target cell infection and pathogenesis**

HIV-1 targets both immune and non-immune cells that can either be dividing and non-dividing cells of the human body. CD4 T-cells are the primary targets of HIV-1 (Bowers *et al.*, 1997; Pattanapanyasat and Thakur 2005). Among the list of immune dividing cells are thymic lymphocyte precursors, follicular dendritic cells and B-cells. Non-dividing immune cells include Langerhans cells and monocytes or macrophages (Aldrovandi, 1997).

In the absence of antiretroviral therapy, the course of HIV pathogenesis follows three stages (Naif, 2013). The first is the acute phase that is characterised by a high viraemia (2-8 weeks), followed by a clinically latent or asymptomatic stage (2-7 years). The last stage is a symptomatic stage which ends in progression to AIDS (Volberding and Deeks, 2010). In the progression to AIDS, there is a rebound of viraemia and the virus changes from CCR5 co-receptor usage to CXCR4 in approximately 50% of infected patients, leaving most treatment naïve patients with a median survival time of less than ten years (Arien *et al.*, 2006).

HIV-1 infection is associated with chronic immune activation and dysregulation of the immune system (Ipp *et al.*, 2013; Freguja *et al.*, 2012). Following successful infection of -permissive cells (such as dendritic cells that migrate to lymph nodes), HIV-1 affects the GALT (gut associated lymphoid tissue) and extensive damage to this tissue leads to translocation of microbial products into the blood stream causing the release of redox species that affect red blood cells (HIV-associated anaemia) and the release of proinflammatory cytokines such as IL 7, IL 10 and IFN- $\gamma$ . IL 10 is known to induce the apoptotic receptor PD1 on macrophages and can inhibit T-cell function by engaging the T-cell receptor PD1 during antigen presentation (Reuter *et al.*, 2012). IFN- $\gamma$ , though reported to have antiviral effects in the early stages of acute HIV infection, leads to an increased expression of apoptotic receptors and an increased expression of the enzyme 2, 3 dioxygenase (IDO) which catalytically breaks down tryptophan - an essential metabolite for T cells (Catalfamo *et al.*, 2012). IL7 plays a role in T-cell homeostasis and may play a role in persistent activation of T-cells making them susceptible to HIV infection. Persistent immune activation is a predictive factor for persistent low-level viraemia (PLLV) observed in some patients receiving therapy (who have achieved viral load suppression below the detection limit of 20 -50 copies/ ml) (Catalfamo *et al.*, 2012).

There are also bystander mechanisms that result in the killing of uninfected CD4+ T-cells. These include Caspase 3 and 1, TRAIL and FAS-mediated signalling pathways. The drop in the number of functional CD4+ T-cells occurs at an average rate of 16 cells per mm<sup>3</sup> per month in non-treated patients (Tebas *et al.*, 2002). Opportunistic infections that include tuberculosis, HIV-associated malignancies such as Kaposi sarcoma and dermatological diseases such as papular pruritic eruptions (PPE) are descriptive of the AIDS-defining symptom complex (Navarini *et al.*, 2010) and are clinical markers of HIV-1 disease progression.

### **1.2.3 HIV transmission routes**

HIV can spread horizontally via needle stick injuries, sexual intercourse and contact with contaminated blood fluids. It can also spread vertically via mother to child transmission (MTCT) (intra-uterine, prenatally or breastfeeding transmission), but heterosexual transmission, is the main mode of transmission and accounts for 85% of new infections worldwide (Simon *et al.*, 2006). In industrialised countries, however, the main high-risk groups associated with HIV transmission are men who have sex with other men (MSM), Intravenous drug users (IDUs), and commercial sex workers.

A recent meta-analysis of 39 studies spanning a period of about 4 years (2011 – 2015) on Sub-Saharan Africa (Adetokunboh and Oluwasanu, 2015) showed that although cART has shown success in curbing HIV transmission rates HIV MTCT is still a major concern. HIV is genetically diverse and different subtypes may have different transmission efficiencies. For example, subtype A has been associated with a higher heterosexual transmission than subtype D but with MTCT, subtype C may be more transmissible than subtype A and D but however, less transmissible when compared to CRFs (Renjifo *et al.*, 2004).

Group M HIV-1 subtype C has emerged as the world's most prevalent and dominant subtype and is responsible for approximately 48% of the world's HIV-1 infections and an estimated 2 - 3 million new infections per year in the Sub-Saharan region (Hemelaar, 2013). However Sub-Saharan Africa, particularly Southern Africa, holds the highest HIV-1 subtype C prevalence and with more and more countries scaling up the roll out of ARVs, there is a growing concern for the management of patients and the monitoring of drug resistance (Hamers *et al.*, 2012).

### 1.3 Antiretroviral therapy drug classes and mechanism of action

Antiretroviral therapy has greatly changed the management of HIV infection globally and has led to the reduction of mortality and morbidity rates in many countries that have managed to successfully implement a national ARV rollout programme. Antiretroviral compounds used in therapy for long-lasting viral suppression are classified into six different drug classes (Tang and Shafer, 2012). These classes include nucleoside/nucleotide *reverse transcriptase* inhibitors (NRTIs), non-nucleoside *reverse transcriptase* inhibitors (NNRTIs), protease inhibitors (PIs), entry inhibitors (fusion inhibitors and inhibitors of co-receptors) and integrase inhibitors (**Table A**). The classification system is derived from the viral life cycle step they inhibit, and each drug class has a specific mechanism of action.

NRTIs limit HIV replication by acting as DNA chain terminators during reverse transcription of the HIV-1 RNA genome to complementary DNA (cDNA). The pro-drugs (**Table A**) are first phosphorylated to active triphosphate forms that compete with the cells naturally produced dNTPs. Insertion of the active triphosphorylated forms of the drugs prevents the further incorporation of another dNTP molecule as the inhibitors lack a 3' hydroxyl group that is essential for the formation of a phosphodiester bond during cDNA synthesis by HIV-1 RT enzyme (Das and Arnold, 2013). Nucleotide *reverse transcriptase* inhibitors are already mono-phosphorylated and require only two cellular phosphorylation steps. NNRTIs allosterically inhibit reverse transcriptase by binding to the hydrophobic region that is proximal to the RT active site (Jiao *et al.*, 2011). The binding of the NNRTI molecule causes a conformational change in the catalytic core of the enzyme, inhibiting the release of the captured dNTP molecule by the enzyme.

The same concept of enzyme competitive inhibition applies to protease inhibitors. Inhibitors of this class bind to the protease active site and this prevents the enzyme from accessing the gag and gag/pol polyprotein precursors that are catalysed by the enzyme to produce mature infectious viral particles (Tang and Shafer, 2012).

Integrase strand transfer inhibitors (INSTIs) target the active site of HIV-1 Integrase and cause a conformational change in the enzyme that leads to the disruption of the correct positioning of the viral cDNA during strand transfer of the viral cDNA from the pre-integration complex to the open regions of the host chromosome (Xu and Kuritzkes, 2010). In addition to displacing the viral genome relative to the enzyme active site INSTIs also prevent interaction with magnesium ions required as cofactors for the Integrase enzyme. The strand transfer reaction is prevented, and this then blocks integration (Tang and Shafer, 2012).

Attachment and fusion inhibitors prevent viral entry: Inhibitors of CCR5 co-receptors allosterically bind to a pocket region of the host CCR5 co-receptor, changing its structure and inhibit HIV attachment by preventing the interaction of HIV-1 gp120 with CCR5. Binding of HIV-1 gp120 to the CCR5 receptor increases the proximity of HIV-1 gp41 to the host membrane to allow for subsequent fusion of the viral envelope and the host membrane; hence inhibitors targeting this step essentially prevent viral entry: After gp120 attachment HIV-1 gp41 gains access to the cell membrane and forms a hairpin by folding of its stalk. The shortened transmembrane protein brings the viral and cell membranes together promoting fusion of the two. In this regard, fusion inhibitors interfere with the hairpin formation of HIV-1 gp41, thereby preventing entry (Matos *et al.*, 2010) that CCR5 inhibitors are only of benefit if the virus is CCR5 tropic and in cases where patients have CXCR4 or dual tropic viruses or a mixture of viruses with CCR5 and CXCR4 tropism CCR5 inhibitors are not clinically useful.

**Table A – Summary of FDA approved ARVS**

Drug Class	Integrase Inhibitors	Entry Inhibitors		Reverse Transcriptase Inhibitors			Protease Inhibitors
		Fusion Inhibitor	CCR5 Antagonists	Nucleoside (NRTI)	Nucleotide (NRTI)	Non-nucleoside (NNRTI)	
Single compound tablets	Raltegravir Dolutegravir Elvitegravir	Enfuvirtide	Maraviroc	Abacavir (ABC)	Tenofovir (TDF)	Delviridine Efavirenz Nevirapine	(Fos)Amprenavir
				Didanosine (ddI)			Atazanavir
				Lamivudine (3TC)			Darunavir
							Indinavir
							Ritonavir
							Saquinavir
							Tripanavir
							Lopinavir/ritonavir
Fixed dose combinations				Abacavir/Lamivudine (Epzicom)			
				Zidovudine/Lamivudine (Combivir)			
				Tenofovir/Emtricitabine (Truvada)			
				Abacavir/ Lamivudine/Zidovudine(Trizivir)			
				Tenofovir/emtricitabine/efavirenz (Atripla)			
				Elvitegravir/cobicistat/Emtricitabine/Tenofovir (Stribild)			

### **1.3.1 HIV drug resistance mechanisms, evolution and viral fitness**

The high mutation rate of HIV-1 is an important biological factor contributing to HIV drug resistance. Approximately 1 - 3 mutations occur in each RNA genome target per round of replication (Mansky, 1995). This is due to the error prone HIV-1 RT enzyme that lacks a 3'→5' proofreading activity. Another contributing factor is that HIV-1 has a high recombination rate when co-infection of a cell occurs with more than one variant (Tang and Shafer, 2012). The high recombination rate is as a result of the mechanism (known as Copy choice) by which HIV-RT switches from one template to another during viral cDNA synthesis (Ramirez *et al.*, 2008).

HIV -1 replicates as a pool of diverse genomic sequences (known as quasispecies) containing both mutated and wild type strains. The quasispecies theory of molecular evolution describes HIV-1 replication at viral population level (Perales *et al.*, 2012). Replication proceeds under complementation and interference genomic interactions. Complementation occurs when gene products necessary for the replication of a particular virus is provided by another viral genome whereas interference is when replication of one virus inhibits the replication of another strain (being the opposite of complementation). HIV-1 quasispecies interact by complementation and interference and the high diversity among replicating viruses also provide a genetic reservoir (termed quasispecies memory) (Perales *et al.*, 2012). The genetic 'reservoir' allows replication-competent DRVs first to emerge as drug-resistant minority strains and subsequently predominate as major variants if their genetic makeup provides a replication advantage in the presence of ARVs. The precise molecular mechanisms for drug resistance vary for each ARV class and the detailed discussion of drug resistance below is limited to NRTIs as the study focused on NRTI drug resistance and mutation fitness interactions.

### **1.3.2 Molecular mechanisms of Nucleos(t)ide Reverse transcriptase resistance**

Molecular resistance with NRTIs is the result of the accumulation of mutations in HIV-1 *pol* gene (Johnson *et al.*, 2003). These mutations can be classified as discriminatory or primer unblocking mutations based on the location within or near to the coding region of HIV-1 RT active site respectively (Tang and Shafer, 2012).

Discriminatory mutations have been shown to have an effect on nucleoside incorporation, and common examples include K65R, M184V and Q151M. Discriminatory mutations derive their classification from the fact that their presence in HIV-1 RT enzyme enables the enzyme to differentiate better or discriminate the NRTI from the cell's natural dNTPs, and this mechanism is a result of the alteration caused by the mutation on the conformation of HIV-1 RT active site (Curr *et al.*, 2005; Garforth *et al.*, 2007).

Discriminatory mutations such as M184V/I, K65R, L74V and Y115F located within the active site of the RT enzyme reduce viral fitness and range in their effect of causing intermediate to high-level clinical resistance to the NRTIs, which selects for them. For example, M184V causes high-level resistance to 3TC and FTC (Gotte, 2012) and low-level resistance to ABC and DDI but shows increased susceptibility to AZT and TDF (Tang and Shafer, 2012). Another example, K65R, which is selected by TDF and to some extent by ABC, D4T and ddI, shows intermediate resistance to all clinically approved NRTIs except AZT (Feng *et al.*, 2005). The fact that these discriminatory mutations result in reduced fitness is an observation explained by the virus's need to conserve energy at the molecular level during viral cDNA synthesis. This evolutionary mechanism suggests that reduced fitness is associated with increased replication fidelity and implies a decrease in the mutation rate of HIV-1 (Dapp *et al.*, 2012).

Primer unblocking mutations (also known as thymidine analogue mutations – TAMs) have been shown to have an effect on ATP-mediated phosphorolytic excision of nucleosides (Richman, 2006). TAMs are selected by thymidine analogues such as AZT and D4T, promote phosphorolytic excision of an NRTI that has been incorporated during chain elongation (Tang and Schafer, 2012). These mutations occur proximal to the RT active site coding area. This mechanism of phosphorolytic excision is based on the discovery that nucleotide incorporation by *reverse transcriptase* is a reversible process in the presence of ATP (Richman, 2006). Two evolutionary pathways have been described for the acquisition of TAMs: type I or type II (Shafer and Schapiro, 2008). The type I TAMs include M41L, L210W, T215Y show increased high-level of resistance to TDF, ABC and ddI, whereas the type II TAMs, which include D67N, K70R, T215F, K219Q/E are described for intermediate resistance to AZT and D4T. The combination of TAMs and discriminatory mutations causes high-level resistance to almost all NRTIs (Martinez-Cajas and Wainberg, 2008).

Two rare mutations associated with multiple NRTI drug resistance are insertions or deletions at codon 69 (known as T69ins/ T69del) and the Q151M (Johnson *et al.*, 2003). The T69ins is known to co-occur with TAMs and causes high level clinical resistance to all NRTIs (Cases Gonzalez and Franco, 2007). The Q151M mutation is known to usually co-occur with accessory mutations that tend to promote the efficiency of the phosphorolytic excision reaction of NRTIs (Shirasaka *et al.*, 1995) and also known to be structurally located in HIV RT active site as the first amino acid in contact with a dNTP molecule and probably has an effect on dNTP insertion fidelity (Rezende *et al.*, 1998). As part of the precursor of the multi-nucleoside resistance (MNR) complex that include accessory mutations F75I, F77L, F116Y and A62V, Q151M causes high-level clinical resistance to AZT, D4T, ddI and ABC and low – intermediate resistance to TDF, 3TC and FTC (Ueno *et al.*, 1995).

Q151M and T69ins mutations are rarely observed to occur, but when they appear severely limit therapy options as they incur resistance to all NRTIs (Deval *et al.*, 2002; Saini *et al.*, 2012).

### **1.3.3 Evolution of acquired antiretroviral resistance (therapy-induced) and transmitted antiretroviral drug resistance**

DRMs can be classified also as either arising from transmission of a drug resistant strain from one individual to another (transmitted drug resistance) or acquired (therapy-induced or secondary) resistance (Stott *et al.*, 2012). Transmitted DRMs results from infection with a strain not fully susceptible to ARVs and varies in prevalence (from 5%-15%) depending on specific geographic locations (Jordan *et al.*, 2012), whereas acquired resistance develops over time in patients receiving treatment and is due to selection pressures emanating from ARV use (Stott *et al.*, 2012)

DRMs can revert to wild type forms in untreated individuals over time (Kuritzkes, 2004) and this observation partially explains low prevalence rates observed before therapy initiation in many clinical settings. A random model of genetic bottlenecks has been postulated as a factor limiting the transmission of DRMs (Wagner *et al.*, 2012). The model has a few assumptions which include: the probability of HIV-1 being transmitted in one single sex act as ranging from 0.001-0.003; time taken for reversion to occur for specific DRVs to wild-type forms ranging from 2 months – 8 years; and that reversion is due to the relative loss in fitness of the DRV compared to the wild-type. The smaller the fitness loss, the higher the probability that the variant will survive a genetic bottleneck and get transmitted and the more time needed for reversion to occur (Wagner *et al.*, 2012). The dynamics of transmitted drug resistance remains complicated by a number of factors such as viral load, mode of transmission, level of adherence to ARVs and possible subtype differences which all could play a role in determining survival of DRVs exposed to genetic bottlenecks (Wagner *et al.*, 2012).

The risk of acquired resistance in clinical practice is influenced by the genetic barrier of the ARVs used in treatment. A genetic barrier in the context of antiretroviral therapy refers to the number of mutations required to achieve resistance to any given ARV and ranges from low to high genetic barrier to resistance. For example, NNTRIs have a low genetic barrier as a single point mutation is sufficient to acquire high-level resistance to these drugs. PIs (when boosted with low dose ritonavir), in contrast, have high genetic barriers to resistance as three or more mutations are needed to acquire high level resistance (Tang and Shafer, 2012). The high genetic barrier of PI therapy as part of HAART contributed to the success of these regimens when introduced in patients who were therapy experienced as seen in a Swiss cohort study (Scherrer *et al.*, 2012). The success of cART is also due to the different mechanisms of resistance to the different drug components.

The action of two or more drugs working in combination in suppressing viral replication decreases the probability of resistance-associated mutations (RAMs) being acquired by the virus to all drugs used in the treatment regimen (Rosenbloom *et al.*, 2012). The observed patterns of acquired resistance have changed over time as a function of the change in recommended regimens according WHO guidelines (Van Zyl *et al.*, 2013) and will likely continue to change as novel ARV drugs become available.

#### **1.3.4 Fitness and the role of secondary (compensatory or accessory) mutations in virus evolution**

Fitness is a relative measure of the replicative capacity of a given viral variant when compared to a wild-type strain (taken as a reference sequence) (Perales *et al.*, 2012). In a particular study (Dapp *et al.*, 2012), the relationship between fitness and replication fidelity was investigated. The study suggested a mutation rate close to HIV-1 wild-type strain represents the optimum for viral evolutionary adaptive forces. To add on, the study proposed a model that assumes that high replication fidelity implied evolution of HIV-1 to a lower mutation rate and factors such as changes in enzyme processivity and increased time for base recognition would require high energy demands that would inevitably result in a reduction of viral fitness. This could possibly explain why discriminatory mutations in HIV-1 pol gene result in less fit viral mutants (Tang and Shafer, 2012). However, the presence of secondary mutations in this gene, change enzyme kinetics in a way that overcomes the energy constraints needed for correct base incorporation during viral synthesis and this compensates for fitness loss (Dapp *et al.*, 2012).

Major NRTI mutations are known to decrease drug susceptibility of the virus at the expense of reduced viral fitness whereas the general accepted dogma for secondary (accessory or compensatory) mutations is that these mutations maintain relative effect of decreased drug susceptibility imposed by major mutations but will enhance viral fitness (Xu and Kuritzkes, 2011)

Secondary/compensatory mutations have been observed in some studies (Huang *et al.*, 2011; Hu and Kuritzkes, 2011) playing a role in restoring viral fitness. Computer mathematical models have been developed to try and distinguish compensatory mutations from one another and to predict whether they have a role in viral fitness or drug resistance. Though controversial, one such model (Chen and Lee, 2006) takes into consideration the conditional parametric constant  $K_a/K_s$ , which is a measure of the selection pressure force on amino acid changes at one site in comparison to amino acid changes on another site. The conditional  $K_a/K_s$  gives a score to indicate negative selection (low score) or positive selection (high score) and was shown to be reproducible among four independent data sets and could efficiently be used to differentiate between primary and compensatory mutations. Its assumptions were argued to be reasonable given that HIV-1 is under strong selection pressure from both the immune system and ARV drugs.

Selection forces result in a high turnover of base transversions and transitions at the molecular level in the viral genome that gives the emergence of compensatory mutations that enable the virus to escape immune and drug neutralization while at the same time remaining fit (Quinones-Mateu and Arts, 2002).

#### **1.4 Antiretroviral resistance and fitness testing and ex-vivo mutagenesis**

##### **1.4.1 Antiretroviral resistance assays: genotypic and phenotypic drug resistance tests**

Antiretroviral drug resistance tests can be classified as genotypic and phenotypic tests (Stott *et al.*, 2012). Standard genotypic testing is done by performing PCR followed by standard DNA Sanger sequencing using primers that target relevant HIV-1 genes. Sequences obtained are analysed using either one of ten standard interpretation algorithms (five that are proprietary and five that are publicly available online) (Tang and Shafer, 2012). These interpretation algorithms compare the sample sequence to a wild-type strain (aligned to it as a reference) by first converting both genomic sequences to amino acid sequences and finding amino acid residues in the sample sequence that are different from the wild-type sequence. The amino acids that differ from the wild-type are designated as mutations and are compared to a known list of clinically relevant DRMs before being reported to a specialist physician. Two commercial kits for HIV-1 genotypic testing and interpretation assays have been developed and these include the TRUGENE (Siemens, USA) and Celera ViroSeq® (Tang and Shafer, 2012). Genotypic assays have the following advantages over phenotypic tests: they are less expensive, are easier to perform, have short turnaround times, can detect evolving resistance in a presence of a mixture of viral genome sequences and can be used to quickly identify primary and compensatory mutations (Stott *et al.*, 2012). However, they have the disadvantage of being less predictive of the effect of complex mutation combinations on drug resistance and viral fitness (Quinones-Mateu and Arts, 2002). To add on, the sensitivity of genotypic assays ranges from 100 – 1000 plasma HIV-1 RNA copies/ml and standard genotypic testing cannot detect DRVs below 20% of circulating virus population. However next generation sequencing technologies can detect minority drug-resistant variants as low as 1% of circulating virus population but their use is largely restricted to research settings only (Chang *et al.*, 2013).

Phenotypic tests detect DRMs by measuring viral replication of DRVs in cell culture assays in the presence of serial dilutions of the ARV drug. They are currently the gold standard for detecting the potency of novel ARV drugs and use two parametric measurements that determine cut off values for ARVs (Tang and Shafer, 2012). The cut-off values include the fold resistance cut off which is a measure of the minimal antiviral activity of the drug based on clinically relevant drug concentrations and the fold decrease in susceptibility based on  $IC_{50}$  values (inhibitory drug concentration required to suppress viral replication by 50%) (Tang and Shafer, 2012).

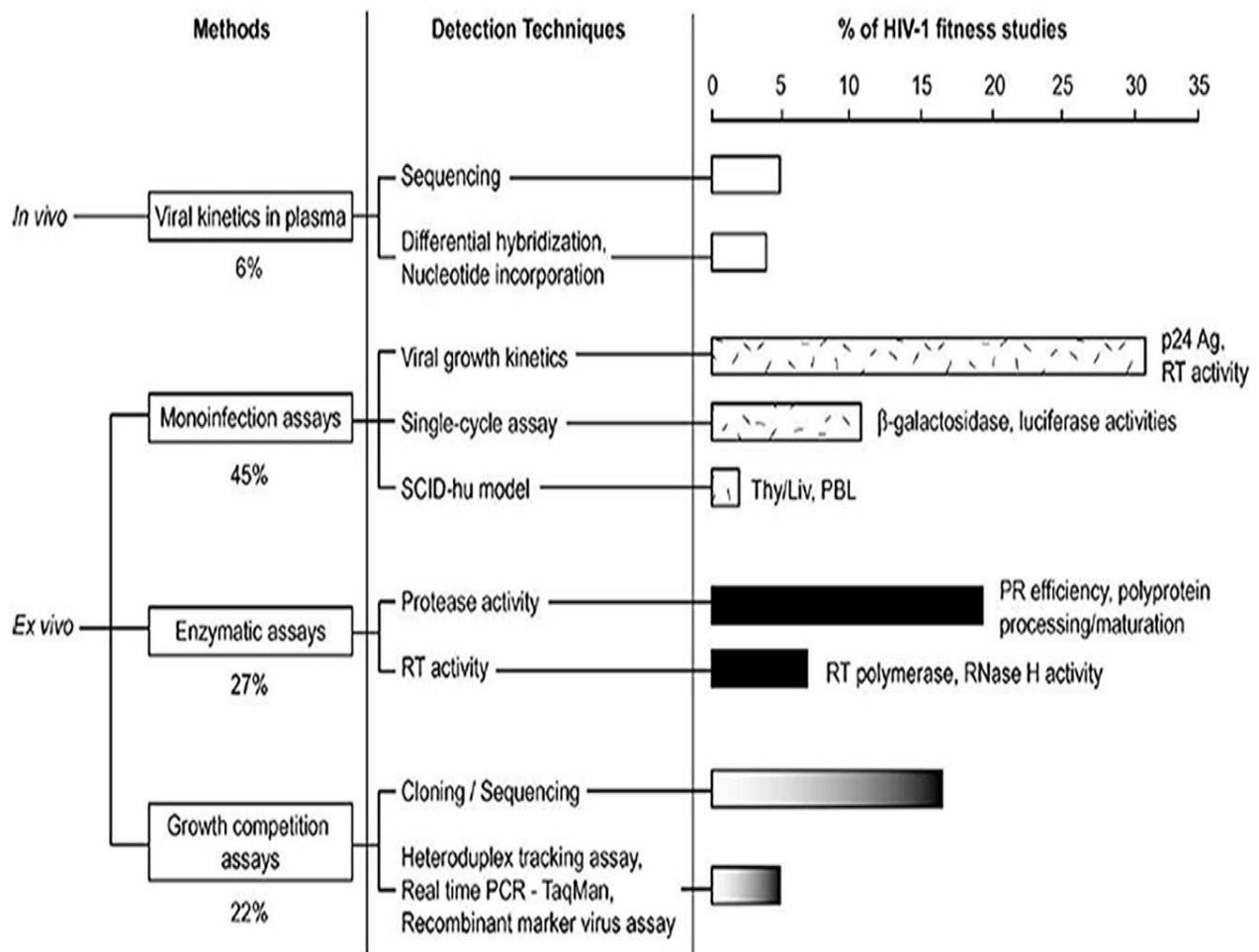
Phenotypic assays are more sensitive than standard genotypic assays in detecting minority drug-resistant variants and in detecting fitness interactions of DRMs (Stott *et al.*, 2012).

Phenotypic assays are also used to accurately determine the tropism of viral variants. HIV-1, co-receptor usage of patient viral quasispecies can be determined by amplifying the complete *env* gene, cloning the amplicons in mammalian expression vectors and transfecting the resultant plasmids in reporter cell lines expressing CCR5 or CXCR4 co-receptors. Though extensively used in research practices there are a few available commercial phenotypic assays to determine ARV potency and these include the PhenoSense (Virologic, Inc.) and AntiVirogram. The Trofile® Manogram is the commercial kit available to assess HIV-1 co-receptor usage (Tang and Shafer, 2012).

Phenotypic assays have a significant disadvantage in that many of the assays require extensive optimization for them to be validated in order to obtain results accurately (Weber *et al.*, 2011). They are time consuming because of cloning steps that are essential in the synthesis of plasmid constructs before transfection in cell lines. They also require multiple controls for experimental validity and the plasmid constructs in turn should be quantified and assessed to display optimal fitness in cell culture (Quinones-Mateu and Arts, 2006). The assay setup also requires various optimisation steps before implementation in a laboratory.

#### **1.4.2 Fitness assays**

Assays for measuring viral fitness require reporter infectious plasmids to accurately measure viral replication in target cells. Examples of common reporter systems utilized in the study of HIV-1 pathogenesis have plasmid constructs that use eGFP (enhanced green fluorescence protein), Luciferase or membrane-bound proteins such as HSA (Heat stable antigen) and PLAP (placental alkaline phosphatase) that are targeted by fluorescence-labelled antibodies for detection of infected cells (Imbeault *et al.*, 2009). However, most HIV-1 reporter constructs, have the reporter systems mentioned above replacing accessory genes (*nef*, *vpu* or *vpr*) and even in *env* compromising the fitness status of the plasmid when transfected in primary or immortal cell lines (Imbeault *et al.*, 2009). Viral fitness as a measure of the replicative capacity of a given viral isolate in a particular environment is hard to measure because of environment fluctuations and interactions of viral quasispecies that complicate its estimation (Quinones-Mateu and Arts, 2006). Various methods used for quantification of viral fitness (**figure 2**) have been developed and are classified into *in-vivo* or *in vitro* based assays.



**Figure 2.** Summary of HIV-1 fitness assays and their relative proportion as used in fitness studies (Picture adapted with permission (see **Addendum A2**) from Quinones-Mateu and Arts, 2002)

*In vivo* based fitness assays (performed in model host organisms) assess HIV-1 kinetics in plasma (by PCR amplification and sequencing or differential hybridization techniques of extracted viral isolates from target cells). However, they present a challenge in that differences in host factors, host-pathogen interactions vary from individual to individual and viral fitness cannot be assumed to be constant in such scenarios (Quinones-Mateu and Arts, 2006). On the other hand, *ex vivo/in vitro* assays are independent of genetic differences between individuals and measure fitness on standardized systems. *Ex vivo* assays are categorised into mono-infection assays (where a single isolate is used to infect cell lines in a single round of replication or in SCID mice models or in viral growth kinetics) They are also classified as enzymatic assays that detect HIV-1 protease, RT activity (Malmsten, 2005) and growth competition assays (which assess co-infection of two or more viral isolates in cell lines and the proportion of each isolate in the mixture is determined after several viral passages).

Growth competition assays are the gold standard for measuring fitness as they model the *in vivo* quasispecies evolutionary dynamics more accurately (Picardo and Martinez, 2008). Various detection systems for growth competition systems have been developed and these include FACS-based analysis (Brockmann *et al.*, 2005); qPCR (Hoffmann *et al.*, 2011); steady-state kinetic analyses measuring HIV p24 Antigen or RT activity (Van der Kuyl *et al.*, 2010) and heteroduplex tracking assays that use fluorescence labelled specific probes that target a specific viral isolate (Van der Kuyl *et al.*, 2010). A more sensitive variation for growth competition assays uses allele-specific qPCR detection (Bergroth *et al.*, 2005) and in such a format primers are specifically designed to detect point mutations with increased sensitivity and specificity. The major challenge with this format of the assay is that multiple primers would have to be designed to detect all relevant mutations of interest and that it will be labour intensive to perform head to head growth competition experiments of all mutant variants with the wild type. Regardless of the technical and labour-intensive challenges growth competition assays impose, they are still the main advantage over other assays and give a more accurate measure of fitness by direct comparison between the outgrowth of one viral isolate from another under constant and internally controlled conditions (Quinones-Mateu and Arts, 2002).

The majority of studies, calculate fitness as a relative measure of the proportion of one isolate to another and the basic formula described by Holland and colleagues (Koval *et al.*, 2006) given as:

$$s = \ln[(Mt / M0) / (Rt / R0)] / t,$$

Where *s* is the fitness measurement value obtained when the natural logarithm of the ratio of *Mt* and *M0* (mutant proportions at time *t* and time 0 respectively) to the ratio of *Rt* and *R0* (reference strain/wild-type proportions at time *t* and time 0 respectively) are computed.

The formula can also be extended by a number of variables that take into account the growth and death rate of isolates that are in comparison to each other (Miao *et al.*, 2008; Ma *et al.*, 2010).

#### **1.4.3 Ex vivo mutagenesis assays: Methods and research/clinical usefulness**

The technique of ex vivo mutagenesis/ site-directed mutagenesis can be used to assess whether a mutation can be considered to be clinically relevant to be adequately defined as a DRM.

Site directed mutagenesis (SDM) allows the change of a nucleotide sequence to represent that of the desired mutant sequence thereby allowing evaluation of gene function in phenotypic assays (Wu *et al.*, 2013). Methods employed in site directed mutagenesis include PCR and selection-based methods.

PCR-based methods have many variations but the general method requires two sets of a pair of primers designed to introduce base substitutions, deletions or insertions at the target site of the sequence (Ho *et al.*, 1989). In the first round of PCR amplification of the target sequence, each pair of primers generates amplicons that overlap each other at the mutagenic site

The amplicons are then allowed to anneal to each other in a second round of PCR amplification generating a linear fragment that has the intended base modification(s) (Landt *et al.*, 1990). The linear mutated fragment can be cloned into a suitable backbone vector and quantified in downstream experiments. Cloning of the fragment is dependent on unique restriction sites that are incorporated into the primers and are introduced in the resultant amplicon after successful rounds of PCR amplification.

The same unique restriction sites can then be used to check for correct insertion of the mutated fragment via restriction enzyme digestion and gel electrophoresis analysis of the products of digestion. Standard DNA Sanger sequencing can then be used for verification of the insert, confirmation of desired mutations and additional checks to verify the absence of random undesired mutations that would have been generated by PCR (Ho *et al.*, 1989).

Selection based methods include the use of phosphorothioate-modified nucleotides that are treated in the desired sequence intended for base modification (preventing the sequence from restriction enzyme digestion and allowing selection from the untreated wild-type sequence) (Lui and Lomonosoff, 2006). Selection-based methods include ampicillin or tetracycline resistance genes that are expressed on plasmids subcloned with the mutagenic fragment but absent on wild-type plasmids and are used for selection during cloning techniques (Jian *et al.*, 1991). Finally, selection-based methods include deletion of unique restriction sites and *dpnI* treatment (Lui and Lomonosoff, 2006). *DpnI* treatment has become the most popular method for SDM and is commercialised in most SDM kits such as the Quick Change XL SDM kit (Stratagene, USA) and Q5™ mutagenesis kit (New England Biolabs (NEB), USA). The selection based method for *DpnI* treatment takes advantage of the property of methylation (Weiner *et al.*, 1994). The wild-type plasmid is chemically transformed in *dam*<sup>+</sup> bacterial competent cells that result in the methylation of the sequence. The wild-type methylated sequence is then used as a template for mutagenesis in a PCR that includes mutagenic primers overlapping at the target site for base modification. The products generated by PCR are linear and partially nicked unmethylated sequences. These can then be distinguished from the methylated wild-type using *DpnI* restriction enzyme that will catalyse the digestion of the methylated or hemimethylated sequence (Weiner *et al.*, 1994). The partial nick in the mutant sequence is then rectified by the action of DNA ligase that is added to the post-PCR reaction or repaired by DNA repair systems (Goodman, 2012) of the competent bacterial cells upon subsequent transformation.

The use of SDM has been shown as an invaluable tool for the study of gene and protein function and in trying to establish the determinants of the mechanisms of various pathogens in causing diseases (Seyfang and Jin, 2003). Particularly, SDM assays in HIV-1 research have been used to identify critical determinants of gene function and structure and examples include RNase H studies (Mizrahi *et al.*, 1990) HIV-1 RT enzyme studies (Garforth *et al.*, 2007) and research on accessory proteins such as *vpu* (Nomaguchi *et al.*, 2010). To add on, application of SDM assays in HIV-1 research has been extensive in the classification of DRMs, fitness interactions and lethal mutagenesis (Dapp *et al.*, 2013).

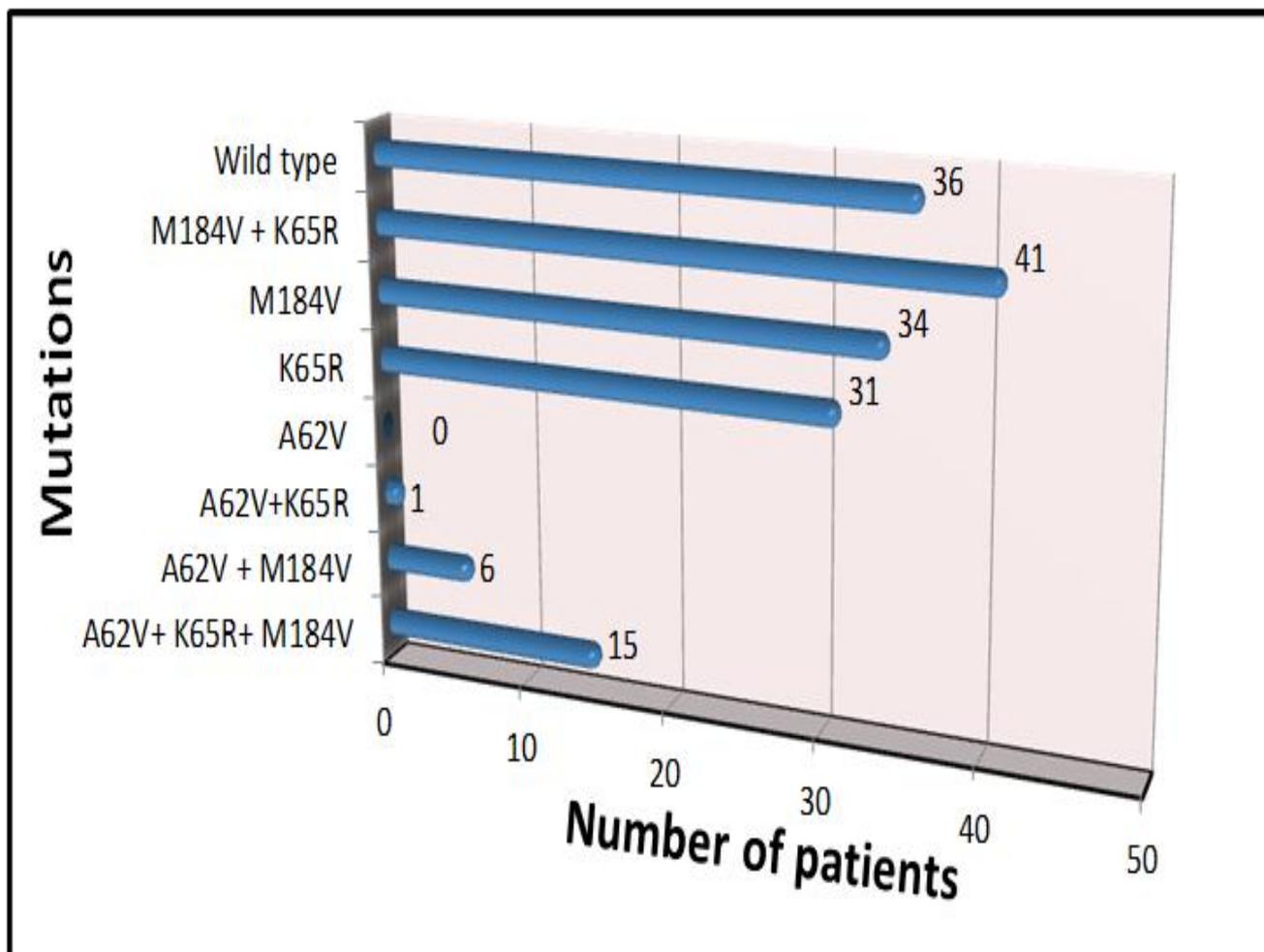
By introducing various point mutations in the gene sequence of HIV-1 RT and subcloning the gene in an HIV-1 backbone vector, mutant derivatives synthesized can be used to determine the effect of a point mutation that affects drug resistance or viral fitness.

More recent work utilizing SDM techniques on the A62V mutation, describe patients having this mutation experiencing failure on most of all the approved NRTIs except TDF. The A62V is part of both multi-nucleoside resistance complex (MNR) mutation pathways. Its biological role has been hypothesized to cause increased excision of NRTIs, when present in MNR complex containing insertions at position 69 of RT (Svarovskaia *et al.*, 2008). Another report describes A62V to be associated with conformational changes in HIV 1 RT that causes it to shift between two states (pre-translocational and post-translocational). The A62V mutation promotes the pre-translocational state that encourages the ATP or pyrophosphate mediated excision of NRTIs, which enables HIV-1 RT to resume chain elongation (Scarth *et al.*, 2011).

#### **1.4.4 Preliminary data for study**

Preliminary data (**Figure 3**) from a large laboratory-based dataset of HIV drug resistance comprising 164 patients (from Tygerberg Hospital, South Africa) on the TDF regimen involving either EFV or NVP as an NNRTI backbone, revealed that A62V occurred in ~36.6% of the total number of patients who had the K65R and M184V RT mutations, whereas the rate of co-occurrence with M184V or K65R alone was low. From the observation, the exact role of A62V mutation is not entirely understood and whether it acts as a compensatory or accessory mutation is the primary focus of this study.

164 HIV-1 patients tested at Tygerberg Hospital on TDF/3TC or FTC/EFV or NVP



**Figure 3:** Preliminary Data for study showing that A62V occurs more frequently ( $P < 0.01$  fisher exact test) with M184V and K65R.

## Research Hypothesis and Aims

This study is a descriptive experiment design study.

The **research hypothesis** is as follows:

**HIV-1 RT A62V mutation has a compensatory role when it co-occurs with M184V and K65R in HIV-1 Subtype C, restoring viral fitness.**

The **null hypothesis** is:

**There is no significant change in viral fitness when A62V co-occurs with M184V and K65R HIV-1 RT mutations in HIV-1 Subtype C.**

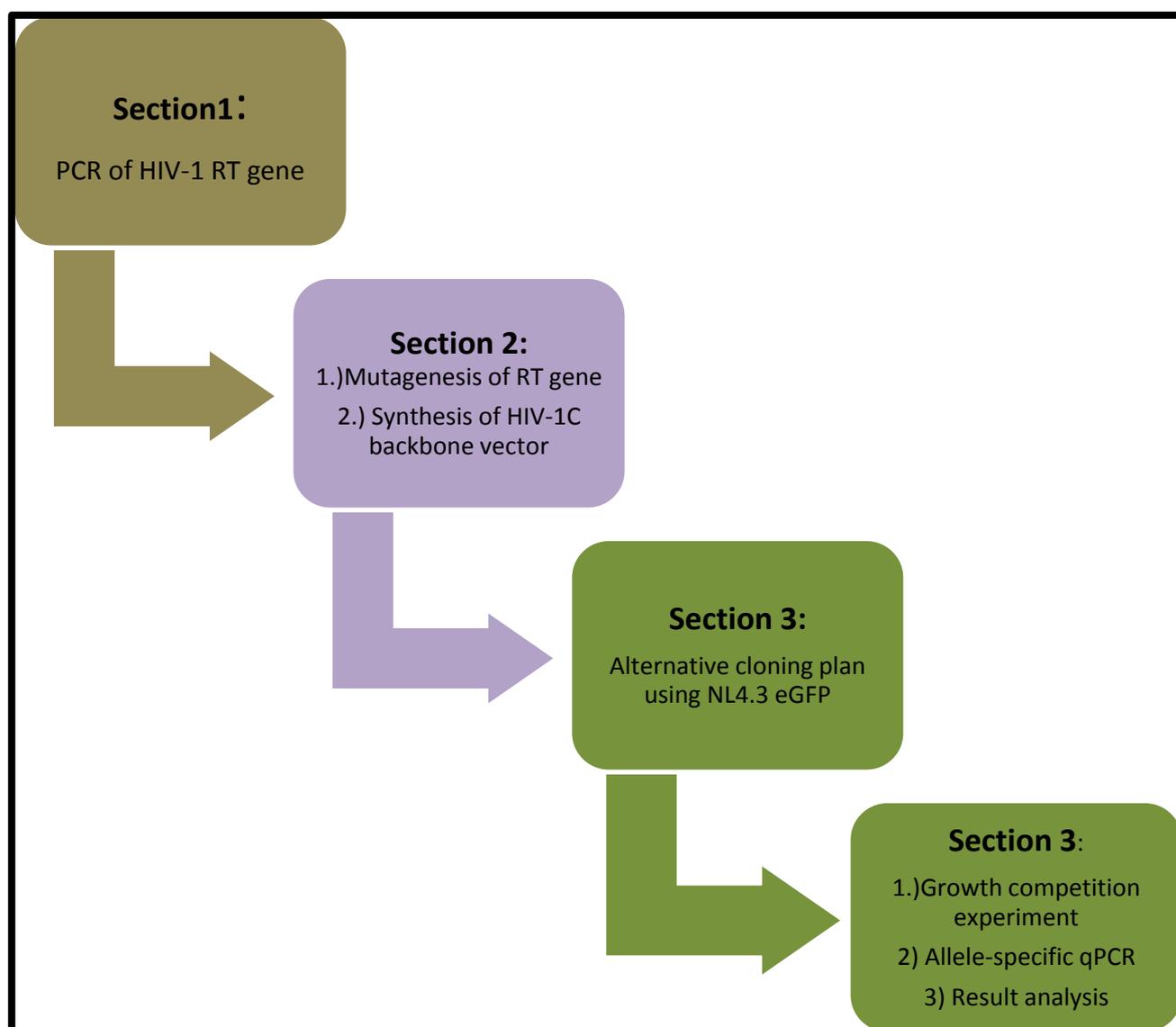
The specific aim of the study is to design and optimize a phenotypic growth competition assay that will investigate this in HIV-1 Subtype C. This includes the following specific objectives:

1. Synthesis of the following full-length genome (FLG) plasmid constructs harbouring these mutations:
  - A62V;
  - M184V,
  - K65R,
  - A62V+M184V,
  - K65R+ M184V;
  - A62V+ K65R,
  - A62V+K65R+M184V
2. Testing the relative fitness of A62V+K65R+M184V vs M184V+K65R clones in a head to head in-house allele-specific quantitative PCR (qPCR)-based growth competition assay.

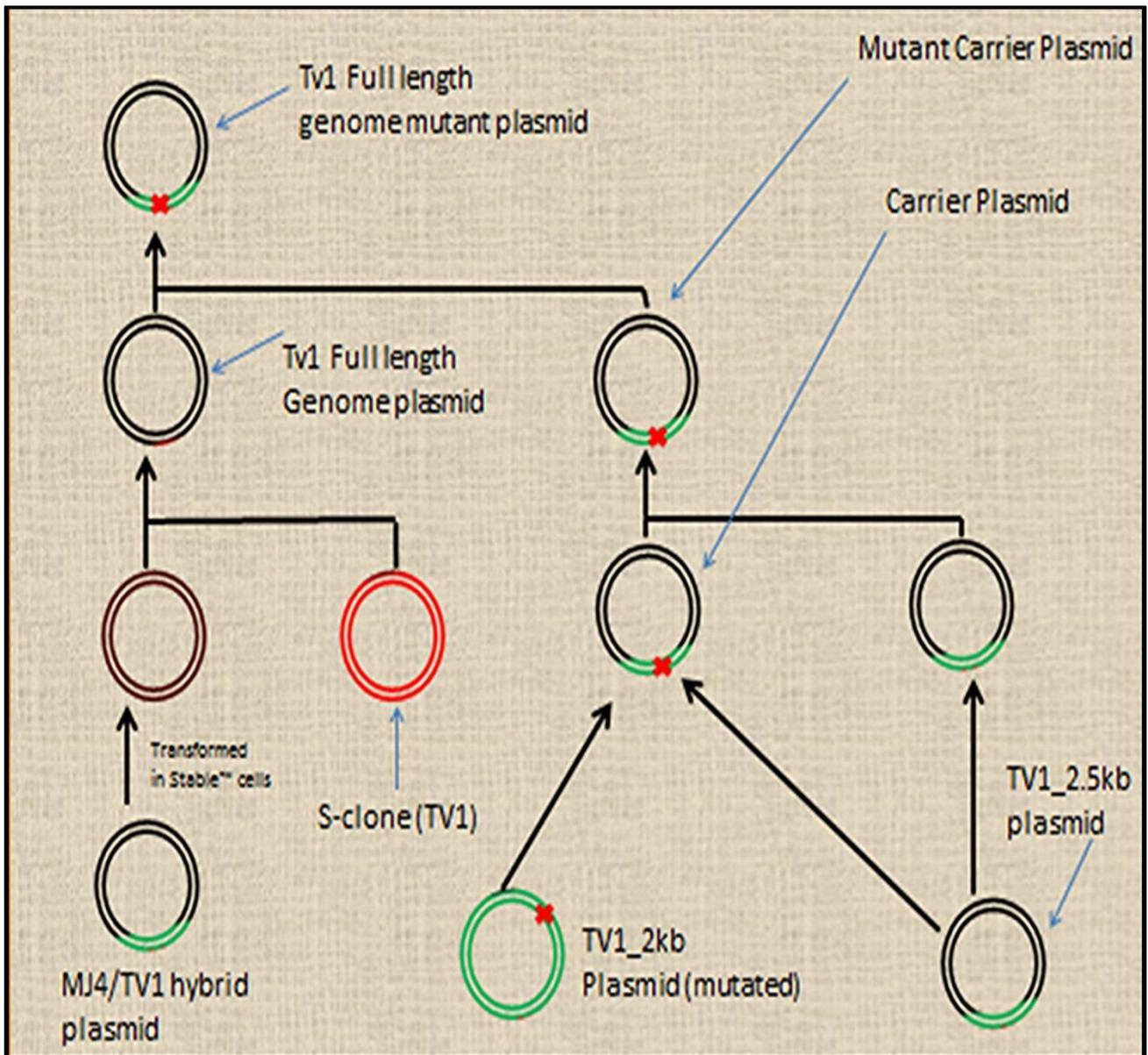
## Chapter 2 - Materials and Methods

### *Introduction*

This study's aim was to test the hypothesis that A62V compensates for viral fitness loss, associated with M184V and K65R, when it co-occurs with M184V and K65R. The methodology used to approach the study objectives under this aim is described under **three sections (Figure 4)** with the third section being a result of the technical challenges encountered with the first two sections. The initial overall plan to synthesize full-length genome (FLG) mutant clones under the first two sections is given in **Figure 5** whereas the alternative plan that employed different cloning methods is described in section 3 of the methodology for the study. A list of the major computer software, equipment, reagents and research kits mentioned in this chapter is presented in **Addendum B**.



**Figure 4.** Overview of methodology plan for study



**Figure 5.** Initial Overall cloning strategy derived from Section 1 and 2 methods. RT gene is amplified from TV1 (HIV-1C) sequence, mutated and subcloned into mini-plasmids (shown as circles) before being finally cloned in the TV1 HIV-1C backbone vector. Green ticks indicate successfully synthesized plasmids whereas red crosses do not.

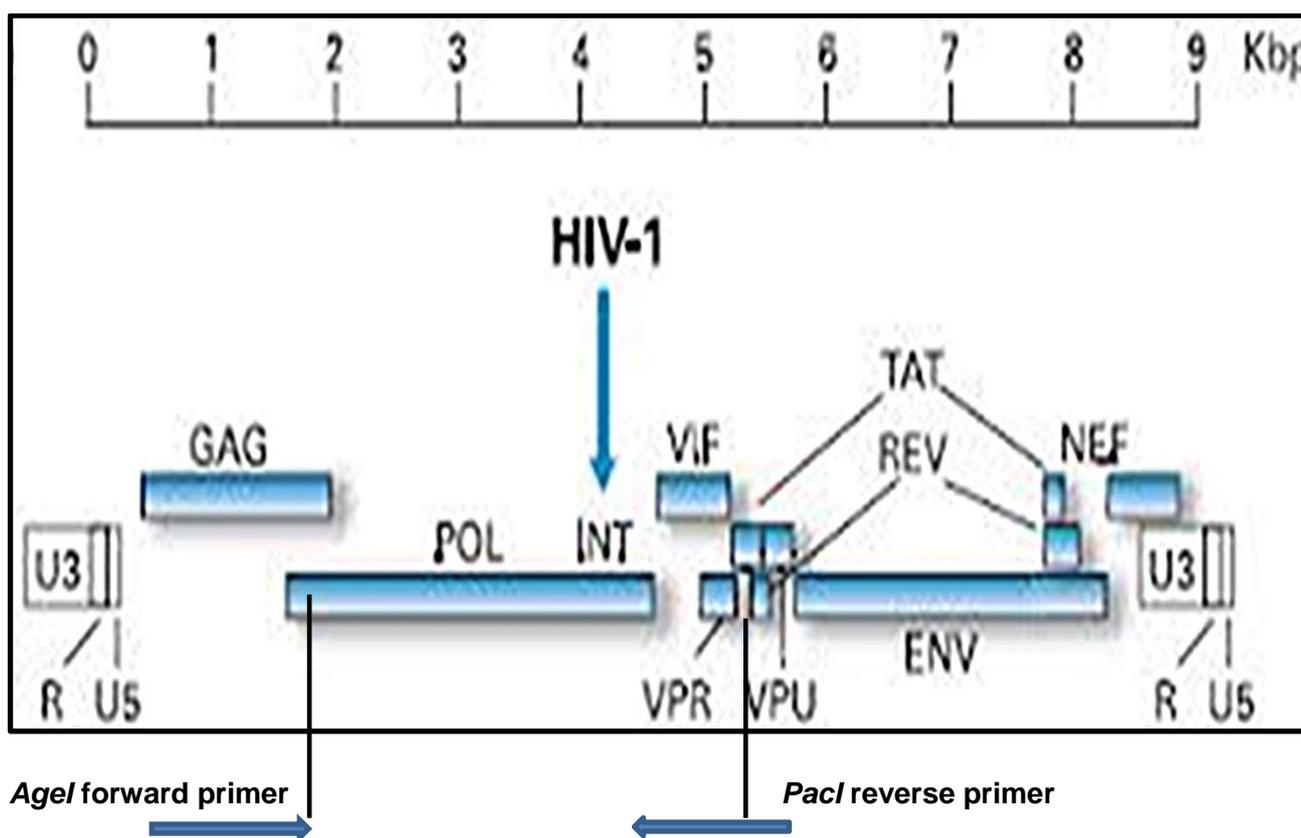
## 2.1 Ethical considerations

Ethical approval for the study was granted by the Health Research Committee of Stellenbosch University, South Africa on the 8<sup>th</sup> of April 2013 (ethics number N13/02/027) for one year and renewed for another year on the 14<sup>th</sup> of February 2014.

## 2.2 Section 1- PCR of HIV-1C RT gene

### 2.2.1-DNA amplification of HIV-1 gag-->start of env gene (HXB2 nucleotide positions 1683 → 6189)

A patient derived HIV-1 Subtype C full genome DNA viral sequence designated **TV1** (Tygerberg Virology sample 1) was PCR amplified from the mid part of HIV-1 *gag* to the start of *env* (HXB2 nucleotide positions 1683 → 6189 - **Figure 6**). PCR with the *Go-Taq* long-range PCR system (Promega, USA) yielded an amplicon that was approximately 4.5kb long (**Addendum C1**). The PCR reaction was validated with a 17.5kb  $\beta$ -globin fragment DNA positive control that is supplied with the kit. The optimized PCR protocol is given in **Table B**.



**Figure 6.** HIV-1 genome map showing positions of primers (*Agel* forward primer and *PacI* reverse primer) that amplify the *gag*--> start of *env* gene (HXB2 nucleotide positions 1683 → 6189)

**Table B. Optimized protocol for PCR amplification of HIV-1C RT gene**

Go <i>Taq</i> _Long range PCR protocol				
Reagents	Stock concentration	Stock Volume	Final concentration	Volume(µl)/50µl
Go <i>Taq</i> _Master Mix	2X	1.25ml	1X	25
<i>AgeI</i> _forward primer (CTATGTAGACCGGTTCTTT)	100nM	1.5ml (10µM)	200nM – 300nM	3
<i>PacI</i> reverse primer (CTCTAATTCTTTTAATTAACCA)	100nM	1.5ml (10µM)	200nM – 300nM	3
Template DNA (TV1)	variable	variable	10ng	2
Double Sterilized Water				17
<b>Total</b>				<b>50</b>
Go <i>Taq</i> _Long range PCR control				
Go <i>Taq</i> _Master Mix	2X	1.25ml	1X	25
PCR master primer pair	10µM each primer pair			1
Human genomic DNA (17.5	30µg		4ng	1/50 of Stock
Nuclease free water to a final volume				<b>50</b>

Thermal cycling conditions (Using the ABI Veriti thermal cycler – Applied Biosystems, CA, USA) were set as follows: initial denaturation at 94°C for 2 minutes for 1 cycle, 30 cycles of denaturation at 94°C for 15 seconds, annealing at 55°C /53°C /52 °C for 30 seconds and elongation at 68°C for 5 minutes 5 sec and with 1 cycle for final elongation at 68°C for 7 min before chilling at 4°C. The correct size PCR product was visualised by gel electrophoresis using Novel Juice™ DNA stain (Promega, USA) and a 1 kb molecular weight marker(Promega, USA) with the remainder stored at -20°C for later use.

### **2.2.2 Cloning of HIV-1 gag-->start of env gene (HXB2 nucleotide positions 1683 → 6189)**

Successful cloning of the HIV-1 *pol* amplicon was achieved via two sequential steps. The first step involved a clean-up of the PCR product using the MinElute enzyme reaction clean-up (Qiagen, Germany) following the manufacturer's protocol (**Addendum C2**). The subsequent reaction involved restriction enzyme digestion at 37°C using 100units per reaction of *Bam*HI restriction enzyme diluted in a final concentration of 1x NEB cutsmart buffer(New England Biolabs, USA) in a final volume of 50 µl. The reaction yielded two fragments of size 2kb and 2.5kb respectively which were again purified using the mini elute clean-up reaction from Qiagen.

The 2kb fragment covered the central part of *gag* to the mid part of *pol* (HXB2 nucleotide positions 1683→3776) and the 2.5kb fragment covered the mid of *pol* to the start of *env* (HXB2 nucleotide positions 3776→ 6189). The second step involved separate ligation of the two individual fragments using T4 DNA ligation system (Promega, USA) into p- GEM T easy vector (Promega, USA). The p- GEM T easy 3015bp vector contains an ampicillin resistant gene. The vector also has an essential feature that can be used to discriminate between desired bacterial colonies that contain the required plasmids and undesired colonies after successful transformation. This feature is a multiple cloning region that overlaps with the  $\beta$ -galactosidase enzyme gene (*LacZ*) contained in the vector. Cloning a fragment in the multiple cloning site results in the disruption of the *LacZ* gene and hence bacterial cells that incorporate this plasmid after chemical transformation generate white colonies when plated out on LB agar ampicillin plates, containing *X-gal* and *IPTG*. Bacterial cells that incorporate a plasmid with a functional *LacZ* gene (specifically if self-ligation of the vector occurred and no fragment has been integrated into the multiple cloning site) appear with a blue phenotype. The blue colour is indicative of the catabolism of the *X-gal* substrate by the expressed  $\beta$ -galactosidase enzyme and hence colonies that show this phenotype are undesired colonies. The ligation reaction conditions in this study using p-GEM T-easy vector, were optimized following the manufacturer's recommendations (see **Addendum D1** for specific reaction conditions and **Addendum D2** for the plasmid maps of the cloned fragments designated TV1-2kb and Tv1 2.5kb plasmids respectively).

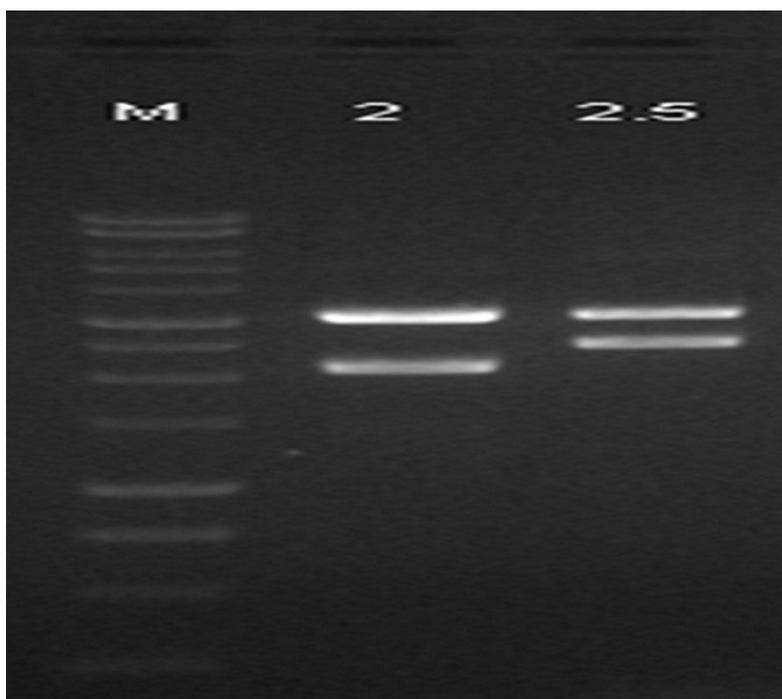
For chemical transformation, 5 $\mu$ l of each ligation reaction for TV1-2kb and TV1-2.5kb to p-GEM T easy vector was added to 50 $\mu$ l of JM109 (Promega, USA) *E. coli* competent cells in a propylene Eppendorf tube and incubated on ice for 30 minutes before being chemically transformed by heat shock at exactly 42°C. The heat shock step activates calcium-gated channels within the bacterial membrane pores, which open up to allow plasmid DNA uptake into the bacterial cytoplasm. The calcium-gated channels are regulated by the concentration difference of calcium ions inside and outside the bacterial cell. After the heat shock step, reactions were further incubated on ice for 2 minutes before 900 $\mu$ l of S.O.C medium was added and after that the reactions were transferred to a shaker incubator (Labcon, USA) for incubation at 37°C and shaking at 225 rpm for 1hr. After the incubation step, bacteria from 100 $\mu$ l of each reaction tube were spread on prepared Lb agar ampicillin plates containing *X-gal* and *IPTG*. These plates were incubated overnight at 37°C, and white colonies were selected from the agar plates the following day. Each white bacterial colony that was expected to represent the successfully cloned fragments were cultured in 5ml of 1 x LB media and incubated for 16hrs in a shaker incubator (Labcon, USA) at 37°C and shaking at 185rpm to generate Miniprep DNA.

Miniprep DNA was prepared from bacterial cultures using the Promega Miniprep Kit (Promega, USA) and DNA plasmids obtained were quantified on the NanoDrop™ ND 1000 Spectrophotometer (ThermoScientific, USA) before subsequent DNA Sanger sequencing using the Big dye cycle sequencing kit v3.1 (Applied Biosystems, CA, USA). Fragments generated from the sequencing PCR reaction were then read on the ABI 3130xL Genetic Analyser. Primers for sequencing inserts, including the standard vector sequencing primers (M13 forward and M13 reverse) and are shown in **Table C**. The primers were designed using the open source bioinformatics program UGENE (Unipro, Novosibirsk, Russia) and were used at a final concentration of 200nmol for the sequencing reaction following the manufacturer's sequencing protocol that was optimized for plasmid miniprep sequencing (**Addendum E**) for the sequencing PCR reaction.

**Table C. Sequencing primers for tv1-2kb and tv1-2.5kb inserts**

<i>Sequencing primers for tv1-2kb and tv1-2.5kb inserts</i>			
<b>Tv1_2.5kb primers (5'-3' direction)</b>		<b>Tv1-2kb primers (5'-3' direction)</b>	
p1	AAGAGCAAGAAATGGAGCCA	p10_compliment	GAACCCCCATTTCTTTGGAT
p3	AGGAATTGGGGGAAATGAAC	p11	AGTGCTAGATGTGGGGGATG
p3 compliment	TTAGCAGGAAGATGGCCAGT	p11_compliment	AAGAACCCCCATTTCTTTGG
p7	GCATTTGGGTCATGGAGTCT	2p4	AAAGGGCTGTTGGAAATGTG
p7_compliment	TCAAGCAGGAAGCTGTCAGA	2p4_compliment	ATGATAGGGGGAATTGGAGG
p1.1	CCCTACAATCCCCAAAGTC	2p7	GGGCTACACCAGTCAACAT
p1.1_compliment	GACCAGCCGAACTACTCTGG	2p7_compliment	AGGAATACCACACCCAGCAG
<b>M13 Forward Primer - GTAAAACGACGGCCAGT</b>			
<b>M13 Reverse Primer - GTTTCCCAGTCACGAC</b>			

Plasmid integrity was further confirmed by restriction enzyme digestion as shown in **Figure 7** following the NEB restriction enzyme digestion protocol (**Addendum F**).



**Figure 7.** Restriction enzyme digestion to confirm insertion of *pol* fragments in *p-Gem T easy* vector. 'M' is the 1kb molecular marker (Promega, USA). Band fragments correspond to the size fragment expected when digested with restriction enzymes for TV1-2kb and TV1-2.5kb which are *AgeI + BamHI* and *PacI + BamHI*, respectively, as shown on the gel photo.

## **2.3 Section 2 –Site-directed mutagenesis, HIV-1 subtype C vector backbone construction and cell culture**

### **2.3.1 Site-Directed Mutagenesis (SDM)**

The TV1-2kb plasmid was the plasmid of choice to introduce mutations of interest as the insert from this plasmid contains the whole coding region of RT enzyme. The mutations A62V, K65R and M184V, were introduced in this clone by using the following methods for site-directed mutagenesis:

- 1) Direct mutagenesis of plasmid using the property of DNA methylation to differentiate wild-type clones from mutant clones
- 2) Overlap extension PCR method (Ho *et al.*, 1989).

In this study, direct mutagenesis of the TV1-2kb plasmids was done using the Quick Change II XL Site-directed mutagenesis kit (Stratagene, USA). SDM primers for this method were designed using the mutagenesis primer tool from Stratagene's website

([http://www.Genomics.agilent.com/primer\\_DesignProgram.jsp](http://www.Genomics.agilent.com/primer_DesignProgram.jsp)) and are indicated in **Table D section A** overleaf.

Table D. Site-directed Mutagenesis plan and primers

Mutations	Base mismatch	Amino acid change	Restriction Sites flanking mutation	Fragment sizes relative to clone after RE digestion
A62V	C = A	Alanine -> Valine	AvrII / HindIII	910 bp
K65R	A = G	Lysine -> Arginine	AvrII / HindIII	910 bp
M184V	A = G	Methionine -> Valine	HindIII/PasI	448bp
A62V+K65R	C = A/ A = G	Alanine -> Valine/ Lysine -> Arginine	AvrII / HindIII	910 bp
A62V + M184V	C = A/ A = G	Alanine -> Valine/ Methionine -> Valine	AvrII / HindIII/PasI	-
K65R + M184V	A = G/ A = G	Lysine -> Arginine/ Methionine -> Valine	AvrII / HindIII/PasI	-
<b>Triple Mutant (A62V+K65R+M184V) is derived from combining M184V construct and A62V+K65R fragment*</b>				
<b>Section A**</b>		<b>Section B**</b>		
<b>Methylation screening method Primers</b>		<b>Overlap PCR method Primers_1<sup>st</sup> round PCR</b>		
<b>A62V + K65R mutant Primers</b>		<b>A62v + K65R primers</b>		
F_ataataactccagctatttgTcataaaaaGgaaggacagctactaagtgag		(AK_F) acactccagctatttgAcataaaaaGgaaggacagctactaa		
R_ctccacttagtactgtccttcCttttatgAcaaatactggagtggtatgg		HindIII_R cctgaagcttcatctaaag		
		(AK_R) ttagtactgtccttcCttttatgTcaaatactggagtg		
		AvrII_F cagggcccctaggaaaaagggctg		
		} H		
		} A		
<b>A62V mutant Primers</b>		<b>A62V primers</b>		
F_tccatataactccagctatttgTcataaaaaagaaggacagctacta		(AV_F) cactccagctatttgAcataaaaaagaaggacagta		
R_tagtactgtccttcctttttatgAcaaatactggagtggtatggga		(AvrII_F)_R cagcccttttcctaggggccctg		
		} HA		
<b>K65R mutant Primers</b>		(AV_R) tactgtccttcctttttatgTcaaatactggagtg		
F_aatccatataactccagctatttgcataaaaaGgaaggacagctactaag		AvrII_F cagggcccctaggaaaaagggctg		
R_cttagtactgtccttcCttttatgcgaaatactggagtggtatggatta		} AA		
<b>M184V mutant Primers</b>		<b>K65R primers</b>		
F_gcaaaaaatccagacatagttatctatcaatatGtggatgactgtatgta		(K5_F) tatttgcataaaaaGgaaggacagctactaagtgaga		
R_tacatacaagtcaccaCatattgatagataactatgtctgatttttgc		HindIII_R cctgaagcttcatctaaag		
		(K5_R) tctccacttagtactgtccttcCttttatggcaata		
		AvrII_F cagggcccctaggaaaaagggctg		
		} A		
		<b>M184V primers</b>		
		(MV_F) gttatctatcaatatGtggatgactgtatgtagga		
		PasI_R cttaatccctgggtaaatctg		
		(MV_R) tctacatacaagtcaccaCatattgatagataac		
		AvrII_F cagggcccctaggaaaaagggctg		
		} P		
		} A		

The reaction conditions for amplification (Using the ABI Veriti thermal cycler – Applied Biosystems, CA, USA) of the plasmid for mutagenesis were followed according to the manufacturer's protocol and these are: a single cycle of initial denaturation at 95°C for 1min, 18 cycles of denaturation at 95°C for 50 sec, annealing at 57°C/62 °C/63°C/65°C ( being primer specific) for 50 sec and elongation at 68°C for 5min 10 sec and a final elongation cycle of 68°C for 7 minutes before chilling at 4°C.

Afterwards, the PCR product obtained was further incubated on ice for 2min before subsequent treatment with *dpnI* enzyme at 37°C for 1hr to digest methylated and hemi-methylated wild type sequences.

The digested product was subsequently transformed into XL1 super competent cells and mutant colonies were distinguished from non-mutant colonies by downstream DNA Sanger sequencing. Normal screening with *X-gal* chemistry could not be applied as these were sub-clones that already had the  $\beta$ -galactosidase gene disrupted upon insertion of the TV1-2kb fragment. Hence, white colonies were expected. Optimization of this method proved technically challenging and time-consuming; and only the M184V mutant subclone construct was successfully synthesized. Five of the remaining six mutant constructs needed for the study were synthesized using the overlap extension PCR method described in the following section.

### **2.3.2 Site-directed mutagenesis by Overlap extension using PCR - method principle and application:**

The method was first described by Ho *et al.*, 1989 and uses two sets two of anti-parallel primers for mutagenesis. In each set of primers, one of the primers contains the base mismatch(es) that will be directed to the target sequence. This primer is overlapping with the complementary primer of the other set, while the other primer of the set contains a restriction enzyme target sequence. The amplicons generated by each of the two primer sets are used in a subsequent second PCR reaction as primers against each other. This second round PCR reaction is conditioned by the fact that the amplicons overlap each other by at least 15bp upstream the 3' end of each amplicon. The resultant dsDNA fragment that is synthesized from the 2<sup>ND</sup> PCR reaction contains the mutation of interest and restriction enzyme sequences that are incorporated at the 3' and 5' end of the fragment. The fragment is subsequently, digested with the appropriate enzymes to allow re-ligation into the original vector. The presence of mutations are confirmed with DNA Sanger sequencing.

In this study, two sets of primers were designed to generate a fragment that contains the mutations of interest for the project using the overlap extension by PCR mutagenesis method. The primers listed in **TABLE D section B** for this method yielded the expected fragments for both rounds of PCR (**Addendum G1**) using the Expand Hi-fidelity PCR System (Roche Diagnostics, Switzerland) following the manufacturer's guidelines and protocol for PCR amplification (**Addendum G2**). Calculations for annealing temperatures of the primers for the first round of PCR were done using the formula adapted from the National Forensic Science Technology Center (NFSTC, USA- [http://www.nfstc.org/pdi/Subject04/pdi\\_s04\\_m01\\_02\\_c.htm](http://www.nfstc.org/pdi/Subject04/pdi_s04_m01_02_c.htm)) below:

$$Ta_{Opt} = 0.3(TmP_1) + 0.7(TmP_2) - 25^{\circ}C,$$

Where  $Ta_{Opt}$  is the desired optimal annealing temperature and  $TmP_1$  and  $TmP_2$  are melting temperatures of primers 1 and 2 respectively.

The second round used a constant annealing temperature of 55°C and resultant fragments were synthesized by the pair of the first round amplicons representing each mutant construct (**Table E**). To complete the first objective of the project three other mutant constructs were synthesized by recombination of the fragments (**Table E section A**) generated in the 2<sup>nd</sup> PCR round via compatible restriction sites with the M184V mutant construct synthesized by Quick Change XL SDM kit.

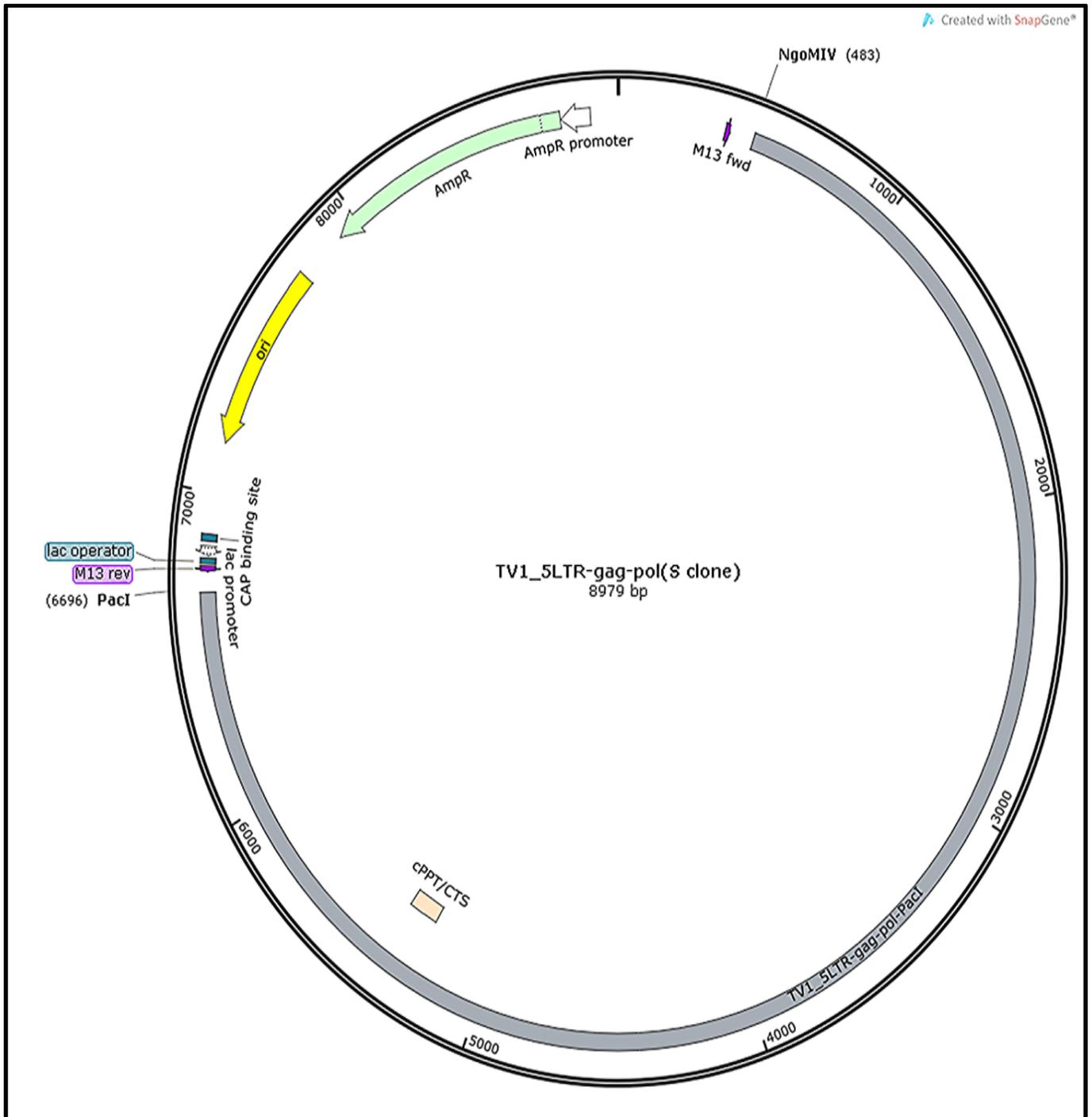
**Table E. Summary of SDM overlap extension PCR generated amplicons**

<i>Mutant Construct</i>	<i>Combinations of Amplicons from 1<sup>st</sup> PCR rnd</i>	<i>Restriction enzymes to be used</i>
A62V + K65R	HK + AK	AvrII + HindIII
A62V	HA + AA	AvrII + HindIII
K65R	H5 + A5	AvrII + HindIII
M184V	PM + AM	HindIII + PstI or AvrII + PstI
<b>Section A</b>		
<i>Other constructs</i>	<i>Fragment(from 2<sup>nd</sup> PCR rnd) ligated to M184V sub-clone Construct</i>	<i>Restriction enzymes to be used</i>
A62V + M184V	(HA + AA)	AvrII + HindIII
M184V + K65R	(H5 + A5)	AvrII + HindIII
A62V + M184V + K65R	(HK + AK)	AvrII + HindIII

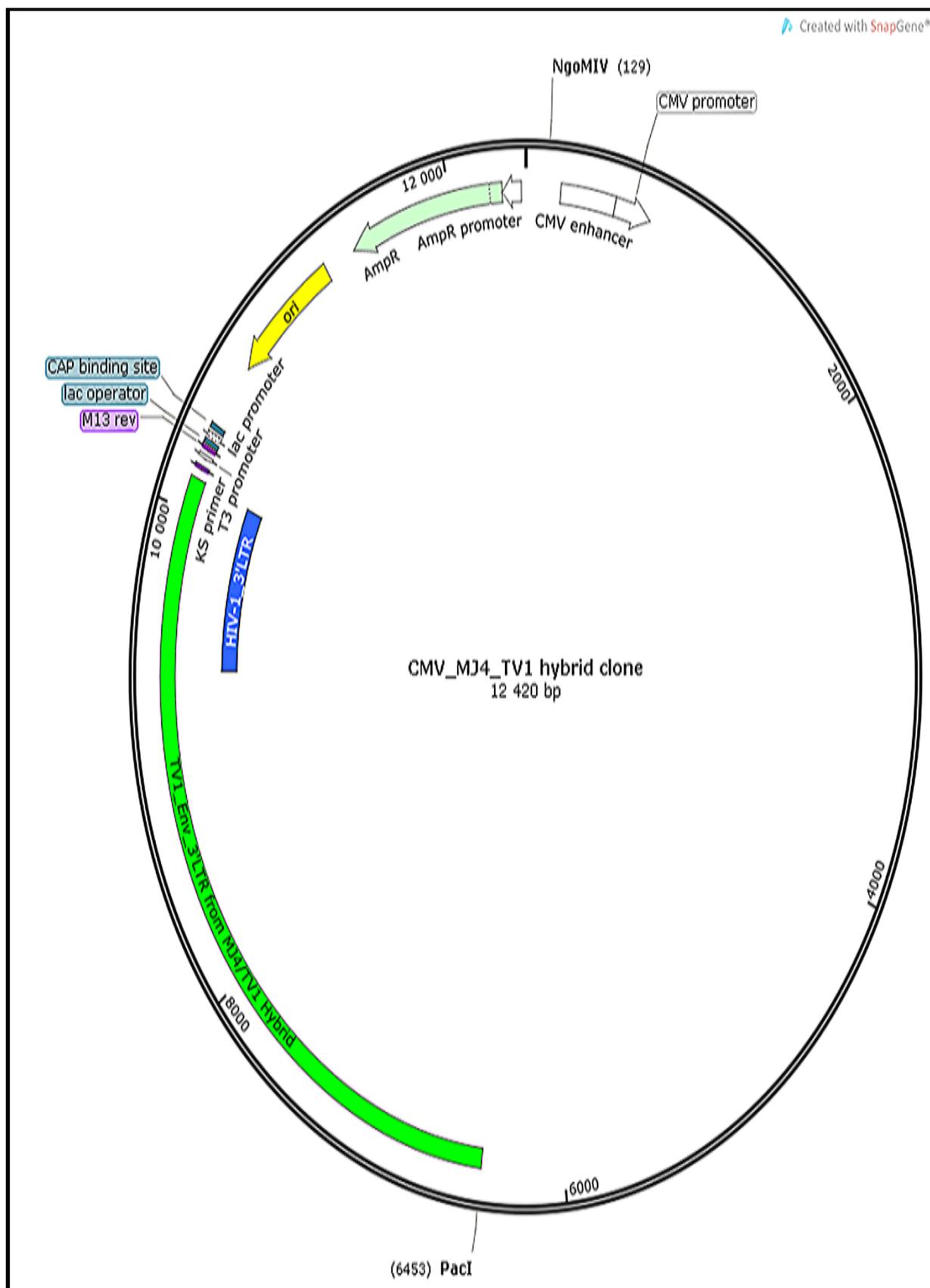
#### **2.4 HIV-1 subtype C backbone vector construction**

The plan for synthesis used two source clones that had the **5' LTR – pol end** (HXB2 nucleotide positions 1 → 6186) and **pol end – 3'LTR** (HXB2 nucleotide positions 6223 → 8795) inserts of the consensus TV1 HIV-1 full genome sequence (Genbank ascension number AY162223). The clones designated in this study as TV1- S-clone and MJ4/TV1 hybrid for the **5' LTR – pol end** and **pol end – 3'LTR** of HIV-1 respectively, (shown in **Figure 8A and 8B**) were each digested with the restriction enzymes *PacI* and *NgoMIV*, following the NEB digest protocol (**Addendum F**). The desired fragments after digestion were size-selected using gel electrophoresis and after that, gel extracted using the QiaxII gel extraction kit (Qiagen, Germany). The gel-extracted fragments were ligated using T4 DNA ligase system (Promega, USA) and transformed into Stable 2 *E. coli* competent cells (Invitrogen, USA).

Colonies were picked from agar plates after successful transformation and were then inoculated in 5ml of 1x LB media for miniprep DNA following the procedure and protocol described in the previous sections. Positive miniprep DNA samples were confirmed by restriction enzyme digestion and sequencing. Plasmid maps and simulations for ligation, cloning and digestion reactions were done using the SnapGene software program (GSL Biotech, USA).



**Figure 8A** – Tv1- S-clone with desired fragment for full-length genome synthesis being the feature labelled and highlighted in grey. Restriction sites PacI and NgoMIV are shown marked in black.



**Figure 8B** – MJ4/TV1 hybrid labelled clone with the desired fragment for full-length genome synthesis highlighted in green. Restriction sites *PacI* and *NgoMIV* are shown marked in black.

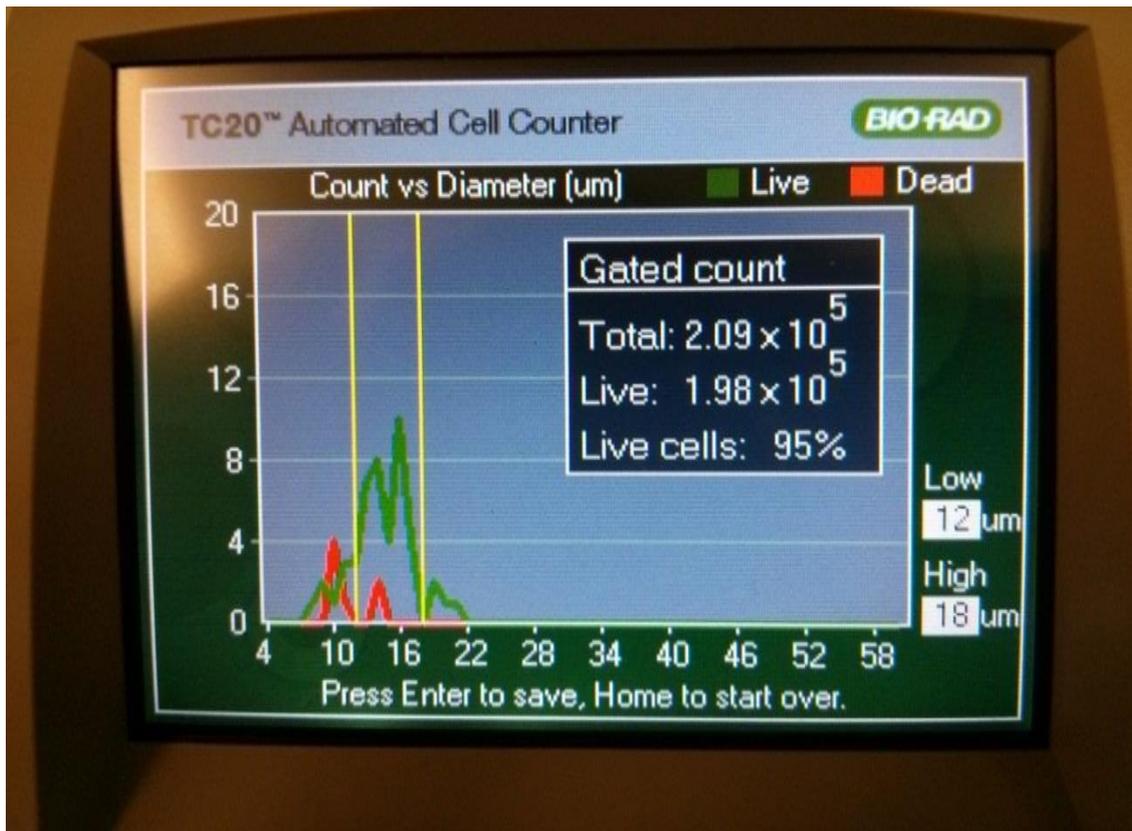
The full-length HIV-1 subtype C construct (TV1-FLG) cloned in the p-Blue script vector after ligation of the two fragments is shown in **Addendum H** with the essential features marked. Ligation of these fragments was successfully achieved by extending the incubation time for ligation to 48hrs at 4°C to promote efficient circularization. Maxiprep stocks were prepared for TV1-FLG using the Pure Yield Maxi prep Kit (Promega, USA) and quantified spectrophotometrically by the NanoDrop™ 1000 (ThermoScientific, USA) before being stored at -20°C for subsequent use in cell culture assays. The average concentration of DNA plasmid obtained was 500ng/μl

## **2.5 Cell culture to test infectivity of TV1-FLG**

### **2.5.1 Seeding cells and transfection**

Infectivity of TV1-FLG was tested by infecting 293T and TZM-bl cells with viral supernatant obtained by the initial transfection of 293T cells and then measuring viral expression and infectivity of 293T and TZM-bl cells using standard HIV p24 ELISA and *X-gal* staining assays respectively. Before infection, 293T and TZM-bl cells were seeded at a concentration of  $2 \times 10^6$  cells in T75cm<sup>2</sup> cell culture flasks and were grown until they achieved a 70 - 80% monolayer. Cells were initially cultured with growth media formulated with 10% (v/v) of fetal calf serum (FCS), DMEM (Dulbecco's minimum essential medium) (Sigma-Aldrich, USA) and a cocktail of antibiotics (containing penicillin and streptomycin- Sigma-Aldrich, USA) before being refreshed every 2 days with maintenance media which had the same formulation as the growth media with an alteration of the FCS used (which was reduced to 5% (v/v)). For subculturing, cells were trypsinized using a 1:10 dilution of ATV solution (Sigma-Aldrich, USA) and then centrifuged at 900rpm for 7 minutes at room temperature. Afterwards, the supernatant was discarded before the cells were resuspended in growth media.

A trypan blue exclusion assay was done to determine the number of viable cells, which were then counted by a TC20 automated cell counter (Bio-Rad, USA). 10μl of the suspended sample was taken and mixed (1:1) with 10μl of trypan blue solution (Bio-Rad, USA) before being added to a slide that the machine could read. The trypan dye solution can penetrate cells only if their cell membranes are disrupted indicating that the cell is non-viable. **Figure 9** shows the gating strategy used on the TC20 automated counter to select the desirable cell size for 293T and TZM-bl cells for sub-culturing.



**Figure 9.** Image was taken from the TC20 automated cell counter showing the concentration of viable cells and the gating strategy used (setting the lower gate at 12µm – higher gate at 18µm) to allow the selection of mature and functional cells based on diameter cell size.

The *X-treme* gene HD transfection reagent (Roche Diagnostics, Switzerland) was used to incorporate 2-5µg of plasmid DNA in 293T cells following the manufacturer's protocol. The protocol followed, was done in a BSL3 (Biosafety Level 3 access lab in the Division of Medical Virology, Stellenbosch University) under a biosafety class II cabinet without need for optimization and required that the 293T cells be seeded a day before transfection in 6 well culture plates at a concentration of  $5 \times 10^5$  cells per well and incubated at 37°C with a constant 5% CO<sub>2</sub> level in a CO<sub>2</sub> incubator (Nuair, USA). After successful transfection of 293T cells, viral supernatant was then harvested after three days, centrifuged at 1200 rpm for 10 minutes at room temperature and the supernatant obtained was either cryopreserved at -20°C or -80°C or directly used to infect TZM-bl cells. The TZM-bl cells trypsinized from maintained cultures were seeded in a 96 well plate at a concentration of  $1 \times 10^5$  cells per well a day before they were infected with 100µl of viral supernatant. The viral supernatant derived from the transfection of 293T cells included TV1-FLG and control plasmids MJ4/TV1 subtype C hybrid plasmid (Jacobs *et al.*, unpublished data), pZAC subtype C plasmid (Jacobs *et al.*, 2012) and NL4.3 subtype B plasmid (AIDS Reagent Program, NIH, USA). The viral inoculums were stored -20°C or -80°C or directly used to infect TZM-bl cells for the *X-gal* staining/ B- galactosidase assay.

## **2.5.2 Cell culture end point assays to check infectivity of TV1\_FLG**

### **HIV p24 ELISA assay**

The QuickTiter™ HIV Lentivirus Quantification kit (Cell Biolabs, Inc., San Diego USA) was used to measure *gag* p24 expression of HIV-1 in infected 293T cells: Anti-HIV p24 monoclonal coating antibody adsorbed onto a microtiter plate captures p24 antigen present in the sample.

FITC-conjugated mouse anti-p24 antibody is added to the sample to bind p24 antigen captured by the first antibody. After several incubation and wash steps, a HRP-conjugated mouse anti-FITC antibody is added and binds to the FITC conjugated anti-p24. The unbound HRP-conjugated mouse anti-FITC antibody is removed in a final wash step that is followed by addition of a substrate solution reactive with HRP which results in the formation of a chromogenic product that correlates with the amount of p24 antigen present in the. The reaction is then terminated by addition of acid and absorbance is measured at 450 nm using an ELISA plate reader. A standard curve prepared from recombinant HIV-1 p24 protein standards is used to determine the concentration of the sample. In this study, eight recombinant p24 standards were generated from the QuickTiter™ HIV Lentivirus Quantification kit and used to determine the concentration of HIV-1 *gag* p24 expression in infected 293T cells.

The assay was set up with NL4.3 viral supernatant being added to wells as a positive control and supernatant from pure cultures being added to wells as negative controls in a 96 well p24 monoclonal antibody coated microtiter plate. TV1-FLG and MJ4/TV1 viral supernatant were also added, and the assay was then run following the manufacturer's guidelines with each sample including standards being added in duplicate on the microtiter plate. Blank controls for normalization with the ELx800™ ELISA plate reader (Bio-Tek, USA) were also included to allow standardisation and accurate spectrophotometric readings of test samples. The ELx800 ELISA plate reader was configured for a system self-test to check whether the filters for the different absorbance wavelengths were functional before running the assay. The assay was only run after confirmation of a system test pass.

### **HIV-1 X-gal Staining assay**

The TZM-bl cells plasmid (AIDS Reagent Program, National Institute of Health (NIH), USA) used in this study are a HeLa-derived cell line expressing CD4, CCR5 and CXCR4. The cells contain Luciferase and  $\beta$ -galactosidase genes cloned in-frame and downstream the HIV LTR unit (Derdeyn *et al.*, 2000). Upon being infected with HIV-1 virus particles the HIV-1 *tat* accessory protein can bind to the LTR unit which induces Luciferase and  $\beta$ -galactosidase genes downstream expression in these cells and can be used as a marker to show infection with HIV-1.

In this study, determination of viral infectivity of TV1-FLG was done by analysing the expression of  $\beta$ -galactosidase in TZM-bl cells using X-gal staining solution. 100 $\mu$ l of serially diluted stocks of TV1-FLG, MJ4/TV1, pZAC and NL4.3 were added to individual wells seeded with TZM-bl cells in triplicate on a 96 well culture plate and the plate was then incubated at 37°C with a constant 5% CO<sub>2</sub> level in a CO<sub>2</sub> incubator (Nuair, USA) for 48hrs.

Afterwards, the cells were fixed with a solution of methanol and acetone (1:1) for 5 minutes, washed 3 times with PBS and stained with 100  $\mu$ l of X-gal staining buffer(**Table F**) per well which had the following formulation for a total volume of 5ml:

**Table F: Xgal staining buffer formulation**

<i>Reagents</i>	<i>Volumes required</i>
Xgal (50mg/ml in DMF- dimethyl fomamide)(Promega, USA)	100 $\mu$ l
100mM MgCl <sub>2</sub> (Sigma-Aldrich)	50 $\mu$ l
100 mM potassium ferricyanide (K <sub>3</sub> [Fe(CN) <sub>6</sub> ]) (Sigma-Aldrich)	250 $\mu$ l
100mM potassium ferrocyanide (K <sub>4</sub> [Fe(CN) <sub>6</sub> ]) (Sigma-Aldrich)	250 $\mu$ l
1xPhosphate buffer solution (PBS)(Lonza, Switzerland)	4.35ml
<b>Total</b>	<b>5ml</b>

The stained plate was incubated for 24hrs, and image results of stained cells were captured using a light microscope after washing the plate with 1x PBS and damp drying on a lab sterile tissue towel.

## **2.6 Section 3 –Alternative backbone to assess mutations in HIV-1 subtype B reporter Plasmid**

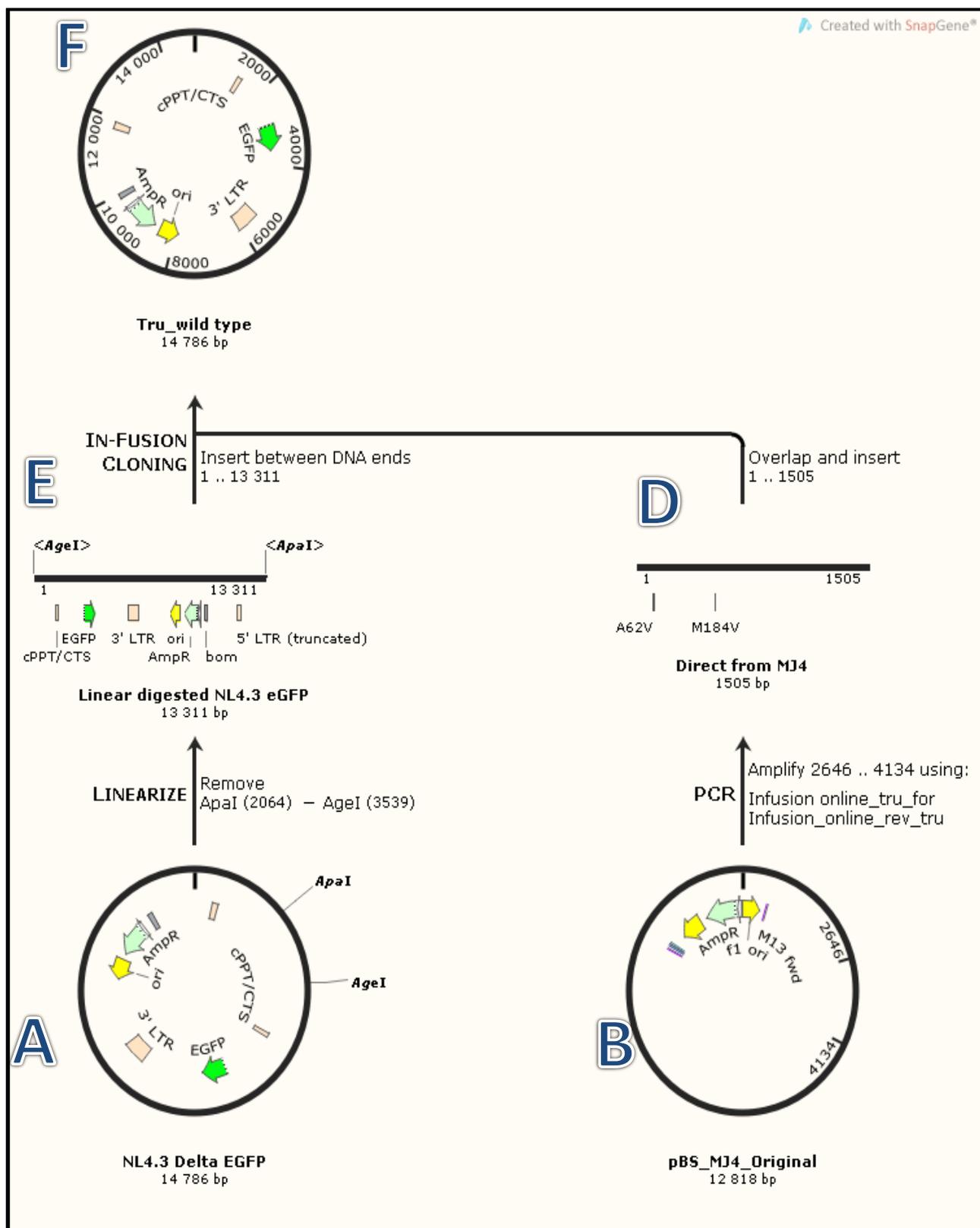
The difficulties in constructing a pure HIV-1 infectious subtype C infectious clone led to a contingency plan to analyse the mutations of interest by ligation of the HIV-1 subtype C *reverse transcriptase* fragment into the lab adapted NL4.3 eGFP (enhanced green fluorescent protein) subtype B reporter plasmid (Aids Reagent Program, NIH, USA). The plasmid has the eGFP gene cloned in-frame in the region where the *nef* gene of HIV-1 has been deleted. Expression of the eGFP gene in cell cultures can be used as a marker to quantify infected cells in downstream assays such as FACS based assay or for qualitative assessment using fluorescence microscopy.

The methods described in the following sub-sections include efficient methods for cloning and mutagenesis of the desired constructs required by the first objective of this study and the design and optimization of the allele-specific qPCR assay used for the growth competition experiment specifically involving the A62V+K65R+M184V and K65R+M184V viruses to determine whether A62V compensates for K65R and M184V associated fitness loss.

### **2.6.1 Synthesis of recombinant constructs using the infusion™ cloning method**

The preferred cloning method employed in this section was developed by Clontech, USA and is termed Infusion™ cloning. The basic components required for the cloning reaction include a PCR product (desired insert) that has been amplified by infusion primers, a suitable vector, Clontech's proprietary Infusion HD cloning enzyme master mix and bacterial competent cells. The infusion primers are custom designed to have 15bp ends that are homologues to the ends of a suitable vector such that when used to amplify the desired insert, result in a PCR product with 15bp homologous ends to the vector. The vector itself is prepared by either restriction enzyme digestion (exposing overhangs with 15bp homology to the Insert) or PCR amplified with infusion primers partially complementary to the vector that result in the vector having the same 15bp homology with the Insert. The prepared insert and vector are then fused together in a ligation-independent reaction that includes the Infusion cloning HD enzyme master mix and is incubated at 50°C for 15-30 minutes resulting in a circularized plasmid ready for transformation in bacterial competent cells.

In this study, the vector was prepared from NL4.3 eGFP plasmid that was digested using the *ApaI* and *AgeI* restriction enzymes and the overall plan incorporating infusion cloning is shown in **figure 10**. The insert was derived from MJ4 HIV -1 subtype C plasmid. The region amplified was the ***gag-pol* region** corresponding to *ApaI*- *AgeI* restriction sites in NL4.3 eGFP (HXB2 nucleotide positions 2010 -3493). The infusion primers used to amplify the MJ4 are indicated in **Table G** and the 15bp homologous ends are color-coded in red.



**Figure 10.** Overall plan for Infusion cloning using NL4.3eGFP and the PCR insert derived from the MJ4 subtype C clone. Steps **A** and **B** are executed in parallel to obtain fragments at step **E** and **D** respectively. The fragments at **E** and **D** are the fused using Infusion cloning to give the product at step **F**

**Table G: - Infusion cloning primers**

<i>Infusion primers for Insert</i>	<i>Primer Sequence</i>
Forward primer( <b>Infusion_online_tru_for</b> )	AAAAATTGCAGGGCCCCCTAGGAAAAAGGGCTG
Reverse primer( <b>Infusion_online_tru_rev</b> )	CACTCCATGTACCGGTGTACTGTTCTTTTAGAATTT

The complete workflow for the cloning plan was initiated by amplification of the MJ4 insert using the Q5™ hot start Hi-fidelity PCR protocol (NEB, USA) and the forward and reverse infusion primers for the MJ4 *gag-pol* region. The resultant PCR product (~1505bp) was cleaned up using the MiniElute clean-up kit (Qiagen, Germany). The clean-up reaction was done in parallel with the digestion of NL4.3 eGFP with *ApaI* and *AgeI* restriction enzymes following the restriction enzyme digest protocol from NEB (**Addendum F**). The linearized vector was gel extracted and purified using the QiaxII Gel extraction kit (Qiagen, Germany). The vector and PCR product were then quantified on the NanoDrop™ ND 1000 spectrophotometer and confirmed for DNA purity as well as the optimal concentration required for the Infusion cloning reaction. The optimized protocol for the infusion cloning reaction that resulted in an HIV-1 BC full-length genome (FLG) recombinant clone (~14.7kb) is indicated below in **Table H**

**Table H: - Infusion™ cloning reaction protocol**

<i>Reaction Component</i>	<i>Reaction</i>
Vector(203ng/μl)_Linearized NL4.3 eGFP	2μl
Insert(46ng/μl)_ MJ4 <i>gag-pol</i>	3μl
HD enzyme mix	2μl
ddH <sub>2</sub> O	3μl
Total Volume	10μl
Time of incubation @50°C	15 min

SDM using overlap extension PCR described in **Section 2.3.1** was used to mutate the MJ4 Insert to generate mutated fragments containing A62V, K65R, M184V and their different combinations yielding seven mutated fragments in total. However, these fragments proved to be technically challenging to clone using the Infusion™ cloning method because of the resultant DNA yield obtained when the SDM protocol was applied. The yield proved to be insufficient to use in the infusion cloning reaction. An alternative strategy was used an optimized protocol adopted from Q5™ Mutagenesis protocol (NEB, USA) was then used to introduce mutations in the wild-type recombinant clone that had been successfully cloned by the infusion cloning technique.

### **2.6.2 Synthesis of mutants by Q5™ Mutagenesis protocol**

The Q5™ Mutagenesis protocol (NEB, USA) is based on the SDM technique described previously in section **2.3.2** which uses discrimination of the wild-type and mutated constructs by *dpnI* digestion of methylated or hemi-methylated products that remain after PCR amplification of the wild-type template with mutagenic primers. Briefly, for the Q5™ mutagenesis protocol, the protocol is modified in a way to combine the ligation and *dpnI* enzymatic treatment in 1 step efficiently. Ligation in the reaction is first achieved by two mechanisms which include phosphorylation of the linear amplified PCR products by the action of a kinase enzyme followed by subsequent ligation of the phosphorylated ends by DNA ligase. The result after this process are circularized plasmids that are then discriminated for selection by *dpnI* enzyme, which in its mechanism of action will target the wild type methylated and hemi-methylated plasmids leaving the desired mutated constructs. The entire process that includes phosphorylation by the kinase enzyme, ligation by DNA ligase and digestion by *dpnI* is referred to as the **KLD** reaction (NEB, USA). After PCR amplification with the mutagenic primers, followed by the KLD reaction, samples are then transformed into bacterial competent cells and colonies obtained can be selected based on their ability to grow on antibiotic resistance plates. These colonies are indicative of successful transformation with plasmids that carry antibiotic resistance genes. The plasmids are then verified using standard DNA sequencing for integrity and to check for incorporation of desired mutations.

In this study, Q5™ mutagenesis protocol was successfully used to mutate RT gene in the 14.7kb HIV-1 BC FLG recombinant clone obtained by infusion™ cloning. Reverse and forward mutagenic primers were designed using the Snapgene program (GSL Biotech, USA). The primers are indicated in **Table I**, and the names designated for the primers are parenthesized below the primer sequence, with the mutagenic bases indicated in small cap letters. Primers were designed in a way to introduce substitution mutations (according to guidelines supplied by NEB in the Q5™ mutagenesis protocol) to synthesis A62V, K65R, M184V, A62V+K65R, A62V+M184V+K65R, K65R+M184V and A62V+M184V FLG constructs.

The manufacturer specifies that the protocol is designed to mutate fragments up to 14.3kb in length; however I modified the protocol by replacing the Q5™ hot start enzyme master mix used for PCR amplification and used Clontech's Hi Fi PCR master mix and this resulted in an improved PCR efficiency of the wild- type recombinant clone. The KLD reaction components were used unmodified but the resultant mutated constructs were then transformed in *JM109 mix and go* competent cells (Zymo Research, CA, USA) which were competent cells observed to be valuable in shortening the time for the transformation reaction (~ 1min) compared with ~2hrs using the NEB 5- alpha competent cells.

**Table I: - Q5™ mutagenesis primers**

<b>Q5™ Mutagenesis designed primers</b>			
<b>Target Mutation</b>	<b>Plasmid template for PCR amplification</b>	<b>Forward Primer</b>	<b>Reverse Primer</b>
A62V	Wild type recombinant clone	CTCCAGTATTTGtAATAAAAAAGAAAGACAGTAC <b>(A_4wd)</b>	TATTATATGGATTTTCAGGCCCAATTTTGTG <b>(Q5_mut_rev)</b>
K65R	Wild type recombinant clone	CTCCAGTATTTGCAATAAAAAgGAAAGACAGTAC <b>(K_4wd)</b>	TATTATATGGATTTTCAGGCCCAATTTTGTG <b>(Q5_mut_rev)</b>
M184V	Wild type recombinant clone	GTCATCTATCAATATgTGGATGATTTGTATG <b>(M_4wd)</b>	TATTTCTGGATTTTGTGTCTAAAGGGCTC <b>(Q5_mv_rev)</b>
K65R+M184V	K65R positive clone	GTCATCTATCAATATgTGGATGATTTGTATG <b>(M_4wd)</b>	TATTTCTGGATTTTGTGTCTAAAGGGCTC <b>(Q5_mv_rev)</b>
A62V+K65R	K65R positive clone	CTCCAGTATTTGtAATAAAAAAGAAAGACAGTAC <b>(A_4wd)</b>	TATTATATGGATTTTCAGGCCCAATTTTGTG <b>(Q5_mut_rev)</b>
A62V+M184V	A62V positive clone	GTCATCTATCAATATgTGGATGATTTGTATG <b>(M_4wd)</b>	TATTTCTGGATTTTGTGTCTAAAGGGCTC <b>(Q5_mv_rev)</b>
A62V+M184V+K65R	K65R+M184V positive clone	CTCCAGTATTTGtAATAAAAAAGAAAGACAGTAC <b>(A_4wd)</b>	TATTATATGGATTTTCAGGCCCAATTTTGTG <b>(Q5_mut_rev)</b>

The 14.7kb FLG constructs derived were verified by standard Sanger DNA sequencing using primers whose names and sequences are AK10 (TYCCCACTAAYTTCTGTATRTC) and AK11(GTACCAGTAAAATTAARCCAG) following the Big dye cycle sequencing kit v3.1 (Applied Biosystems) optimized protocol (**Addendum E**). Once the constructs were obtained a maxi prep stock was made for each plasmid construct using the Maxiprep kit (Promega, USA) and stored in 200µl constructs with an average concentration of 400ng/µl.

Afterwards, each plasmid construct was transfected into 293T cells for virion production using the *X-treme* gene (Roche Diagnostics, South Africa) protocol described in **section 2.5.1**. The efficiency of transfection was confirmed using a fluorescence imaging assay described in the next section. The harvested virus of the ***A62V+K65R+M184V*** and ***K65R+M184V*** constructs, obtained, post-transfection, was then quantified using a TCID<sub>50</sub> end point titration assay and using this the results of this assay a measure of the replication-competent infectious virion particles was then calculated and used in preparation for the growth competition assay.

### **2.6.3 Assays to test infectivity and quantification of replication competent virus**

#### **2.6.3.1 Fluorescence imaging assay**

To optimize the harvesting of virus, post-transfection, a fluorescence imaging assay targeting the expression of eGFP protein that is expressed by successfully transfected cells incorporating the recombinant plasmids was done. Setup of the assay involved growing 293T cells on 20 x 20 mm glass coverslips in a 6-well culture plate. The cover slips were first sterilised before use with a solution of 57% ethanol and 10% sodium hydroxide in petri dish following a prescribed protocol (Redig, 2013). Afterwards, these were rinsed with distilled water before further sterilization with UV light for 30 min under a UV cabinet hood (Nuair, USA). In addition to UV sterilization the coverslips were then coated with 1mg/ml poly-Lysine (Sigma- Aldrich, USA) to allow attachment of 293T-cells. Following transfection of the recombinant plasmids into each well on a 6 culture well plate, coverslips were extracted using sterile forceps and fixed with 4% paraformaldehyde for 10 minutes before being rinsed with once with 1 x PBS and visualized for eGFP using a standard fluorescence microscope. The maximum peak in eGFP expression was then observed between 18 - 48hrs to determine the optimal harvest point and to qualitatively establish the efficiency of the transfection experiment. After the optimal conditions for transfection had been identified, the transfection experiment was repeated to generate viral stocks for the recombinant plasmids constructs. Post-transfection, viruses were obtained by harvesting cell culture supernatant, followed by 10-minute centrifugation at room temperature at 1200rpm and afterwards storing the supernatant obtained in aliquots of 1 - 2ml at -20°C or -80°C for later use.

#### **2.6.3.2 TCID<sub>50</sub> end point titration assays (*A62V+K65R+M184V* and *K65R+M184V*)**

TZM-bl cells were used to set up the TCID<sub>50</sub> titration assay as they permit indirect quantification of replication competent virus by assaying for B-galactosidase enzyme activity for HIV-1 infected cells. Serial dilutions ranging from 10<sup>2</sup>-10<sup>6</sup> for were prepared in 1.5ml Eppendorf tubes for the viral aliquots of the *A62V+K65R+M184V* and *K65R+M184V* viruses that were to be used in the growth competition experiment setup. Briefly, 900µl was pipetted into 1.5ml Eppendorf tubes marked from 10<sup>-1</sup>-10<sup>-6</sup> for each virus. 100µl of each viral aliquot was then added to the tubes labelled with the ten-fold dilution. From there on, 100µl from the 10-fold dilution tube was then transferred to the tubes

marked  $10^2$ . The procedure was repeated for each successive dilution in the dilution series up to the tube marked  $10^6$  to achieve a dynamic range of dilution from 100 – 1000000 times). TZM-bl cells were then seeded in a 96 flat bottomed cell culture plate as indicated in **Figure 11**. A total of two 96 flat bottomed cell culture plates were used – one for each virus and a total of 12 replicates for each dilution in the dilution series for each virus were used in the assay to attain a statistically accurate measure of the TCID<sub>50</sub> value. The Spearman-Karber formula was applied to calculate TCID<sub>50</sub> values for both viral mutants (A62V+K65R+M184V and K65R+M184V). The Spearman – Karber formula is described with the following equation: Spearman-Karber Formula =  $x_k + d [0.5 - (1/n) r]$  - Where  $x_k$  = dose of highest dilution;  $d$  = space between dilutions;  $n$  = number of replicates per dilution and  $r$  = total number of negative wells excluding negative controls. For the conversion of the TCID<sub>50</sub> value to plaque forming units (PFUs), a conversion factor of 0.56 was used. This factor is a standard mathematical conversion factor to get the equivalent of replication-competent infectious virion particles as measured by the plaque forming assay. The results obtained were then used to normalize for the starting concentration of both viruses in the growth competition assay that would have an overall Multiplicity of infection (M.O.I) of 0.01.

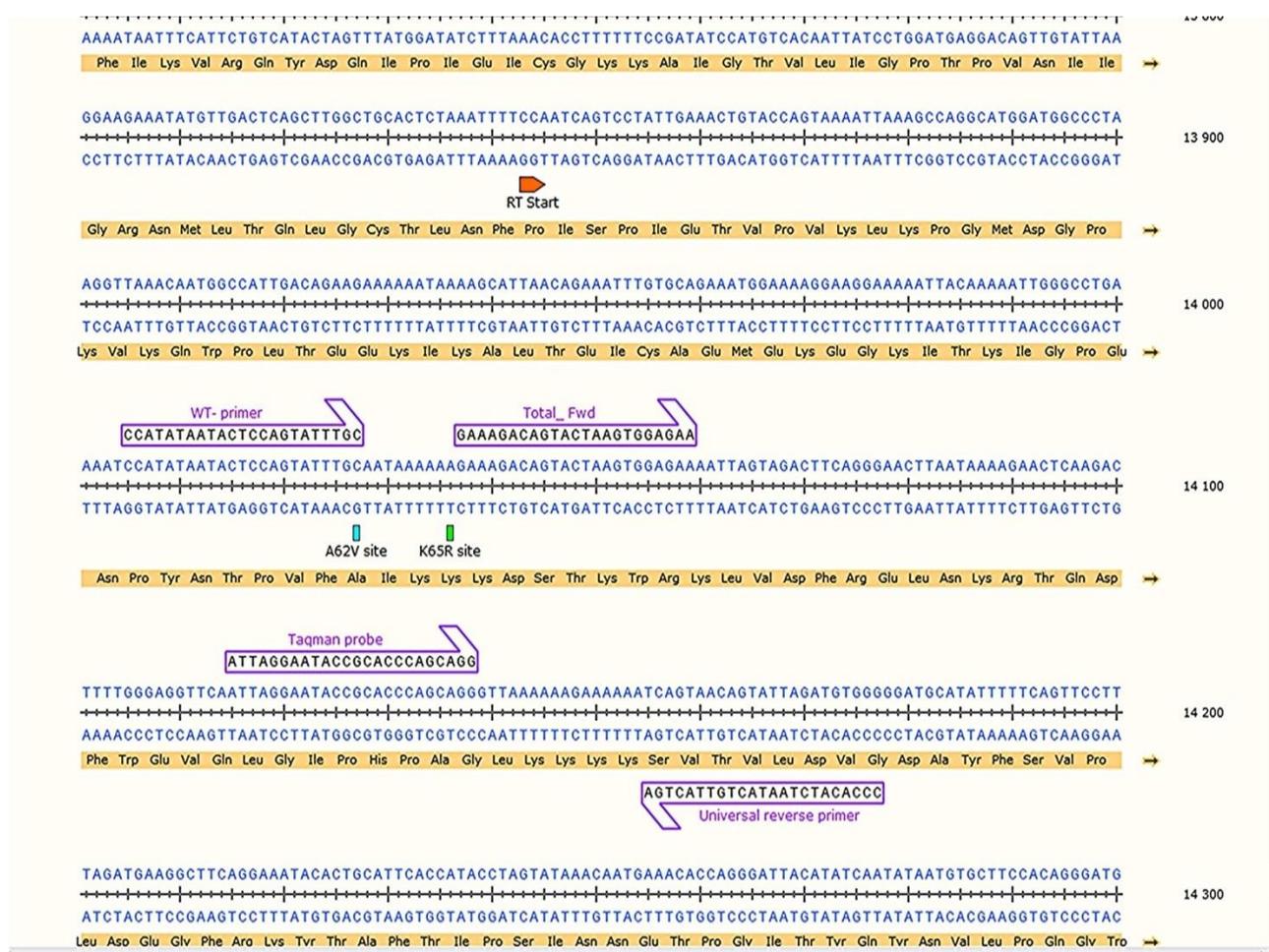
	1	2	3	4	5	6	7	8	9	10	11	12
A	$10^2$	$10^2$	$10^3$	$10^3$	$10^4$	$10^4$	$10^5$	$10^5$	$10^6$	$10^6$	Negative	
B	$10^2$	$10^2$	$10^3$	$10^3$	$10^4$	$10^4$	$10^5$	$10^5$	$10^6$	$10^6$	Negative	
C	$10^2$	$10^2$	$10^3$	$10^3$	$10^4$	$10^4$	$10^5$	$10^5$	$10^6$	$10^6$	Negative	
D	$10^2$	$10^2$	$10^3$	$10^3$	$10^4$	$10^4$	$10^5$	$10^5$	$10^6$	$10^6$	Negative	
E	$10^2$	$10^2$	$10^3$	$10^3$	$10^4$	$10^4$	$10^5$	$10^5$	$10^6$	$10^6$	Negative	
F	$10^2$	$10^2$	$10^3$	$10^3$	$10^4$	$10^4$	$10^5$	$10^5$	$10^6$	$10^6$	Negative	
G												
H	Neat	Neat										
Key:												
	Negative controls (Uninfected wells)											
	Inoculated wells											
Number of TZM_bl cells required per well = $0.5 \times 10^5$ cells												
Total number of cells required for plate = Number of wells x $0.5 \times 10^5$ = $3.4 \times 10^6$												
Total number of cells required for 2 plates = 2 x Number of wells x $0.5 \times 10^5$ = $6.8 \times 10^6$												

**Figure 11.** Layout for TCID<sub>50</sub> end point titration assay for A62V+K65R+M184V and K65R+M184V with an indication of the number of cells seeded in each well for the assay

### 2.6.4 Allele-specific real-time qPCR assay design

To accurately detect and differentiate the growth of A62V+K65R+M184V and K65R+M184V mutant HIV-1 BC recombinant viruses in the growth competition experiment, an allele-specific and probe-based real-time qPCR assay was designed.

The qPCR assay used forward and reverse primers to amplify a ~140 - 170bp fragment targeting the reverse transcriptase gene region in which A62V and K65R mutations were introduced. Two forward primers were specifically designed. The first primer designated Total forward (T-Fwd) primer was designed to be able to detect the total virus that was to be assayed during the Growth competition experiment. This primer was different from another forward primer designated the Wild-type specific primer (WT-primer) that was specifically designed to detect viruses that did not contain the A62V mutation (A62V negative). This primer had the wild-type DNA nucleotide base in the position of the mutated base (A62V), making it specific for A62V-negative viruses (see **Figure 12** for an overview of the design).



**Figure 12.** Overview of allele-specific real-time qPCR design. Primers for assay are indicated in purple arrows.

The difference between the total virus and the A62V-negative virus deduced from the combination of using the two primers in the qPCR assay will allow for the quantification of the A62V positive virus and further analysis of the growth kinetics of both A62V-negative and A62V-positive viruses. The WT-primer would specifically differentiate between A62V-negative and positive viruses by causing inefficient amplification of the A62V-positive virus due to the mismatched base, and this then gives a delay in the cycle threshold (Ct) value obtained during qPCR. The greater the Ct value delay, the greater the primer specificity and the greater the statistical confidence the primer has in the detection of wild-type (A62V-negative and also perfect matched template). For example, a Ct delay of 6 cycles during qPCR would mean that the assay is 66-fold more sensitive for the wild type target than the mutant as at 100% efficiency a target is amplified 10-fold each 3.3 cycles. Therefore 6 cycles would equate to  $10^{(6/3.3)} = 65.79$ . At 50% wild-type and mutant concentration, less than  $0.5 \times 1/66 \sim 1\%$  of the wild-type signal could therefore be ascribed to a false positive (mutant template) signal. Using the T-forward primer (detecting wild-type and mutant) and wild-type assay in parallel would therefore allow the accurate differentiation of variants harbouring the mutant vs wild-type by subtracting the wild-type signal from the total signal to obtain the mutant concentration as long as the Ct difference is large and the expected wild-type and mutant concentrations are the same range (as the false signal would also be affected by the proportion of the 'false' template, with a higher discrimination needed when the 'false' template is in excess relative to the 'true' template)(Bergroth *et al.*,2005).

Setup of the assay was completed by the addition of DNA standards derived from the plasmid wild type and A62V-positive recombinant clones. DNA standards were prepared by using a short fragment from the plasmids amplified using the infusion primers and Q5™ hot start PCR protocol (NEB,USA) described in section 2.6.1. The ~1505bp amplicon obtained was PCR cleaned up using the MiniElute reaction clean-up kit (Qiagen, Germany) and after that quantified and confirmed for purity on the NanoDrop™ ND 1000 spectrophotometer. After quantification of both wild-type and A62V, DNA amplicon stocks were confirmed to be at the same starting concentration regarding DNA copy numbers using data acquired from NanoDrop™ ND 1000 spectrophotometer and a DNA copy number calculator (Staroscik, 2004). Dilution series of both wild-type and A62V specific standards from  $10^2$ - $10^6$  were made using wild-type and A62V plasmid amplicon stocks.

Serial dilutions of the standards were done following the same procedure for making 10 fold factor serial dilutions described in section 2.6.3.2 but with modifications which include using a volume to volume mix ratio of 20µl (plasmid amplicon solution) : 180µl (nuclease free water preheated to 56°C) mixed by pipetting up and down 20 times. After preparation, the standards were stored at -20°C for later use.

Optimization of the qPCR assay was completed by multiple runs done using master mixes specifically made up with either T-Fwd or WT-specific primers. In addition to the primers the master mix consisted of a *TaqMan* probe (using FAM as a fluorophore and BHQ™ as the quencher), 1 Step Lo - Rox PCRbio real time master mix (PCR biosystems, USA) and the template that was either wild type or A62V mutant standards. The total volume of the reaction was 20 µl and the cycling parameters on the CFX96 real time PCR system (Bio-Rad, USA) were as follows: Reverse transcriptase (RT) incubation step at 50°C for 10 min, Initial denaturation at 95°C for 2 minutes and 39 cycles of denaturation at 95°C for 10 seconds followed by elongation and extension at 60°C for 20 seconds. The standards were assayed in triplicate. The 1 Step Lo-Rox PCRbio real time master mix can be used for detection of either RNA or DNA templates because of the addition RT in the mixture and this has the direct advantage of quantifying viral RNA extracted from the cell culture growth competition experiment.

### **2.6.5 Growth competition experiment (A62V+K65R+M184V vs K65R+M184V)**

The TCID<sub>50</sub> values obtained using the TCID<sub>50</sub> end point titration assay were used to calculate the amount of replication-competent A62V+K65R+M184V and K65R+M184V mutant viruses and the results were used to normalize the starting concentration of these viruses for the head to head growth competition experiment. The head to head growth competition experiment set up was finalized by seeding 5 x 10<sup>5</sup> TZM-bl cells per well in a standard 6-well culture plate, followed by 24hr pre- incubation of the 6-well culture plate in a humidified, 5% CO<sub>2</sub> and 37°C incubator (Nuaire, USA). Three wells were used as biological replicates for the growth competition experiment and labelled accordingly.

Two wells were used as positive controls that were inoculated with A62V+K65R+M184V and K65R+M184V viral aliquots respectively. The inoculum used represented an M.O.I. of 0.01. Viral supernatant from the growth competition experiment was harvested at time 0, 48 and 72hr following co-inoculation of the mutant viruses in the biological replicate and control wells marked on the plate and the supernatant was stored at -20°C for later extraction with the Qiamp Viral RNA extraction kit (Qiagen, Germany). Viral RNA was then quantified using the allele-specific real-time qPCR assay optimized for running on the CFX96 real-time PCR system (Bio-Rad, USA)

For quality control purposes, two technical replicates were assessed in parallel for the same biological replicate's RNA extract in the qPCR experiment. Three replicates for No template controls (NTCs) and cell culture negative controls harvested from the growth competition experiment were also included.

Analysis of qPCR experiment results was done using the standard curves method to assess reaction efficiency and the correlation coefficient for the assay. For the WT-primer allele-specific qPCR assay, mutant standards were included as false standards to check the specificity of the WT-primer. Statistical analysis with the Wilcoxon signed rank test in the program R v3.2.2 (R Development Core Team, 2008) for comparison of the relative fitness of A62V+K65R+M184V and K65R+M184V mutant viruses and the overall coefficient of variation (CV) was calculated to assess the assay's reproducibility. Results also include growth curves for the two viruses and are presented in the next chapter.

## Chapter 3 – Results

### **Introduction**

The results are presented on a section by section basis following the same outline of the methodology described in the previous chapter and these include: -

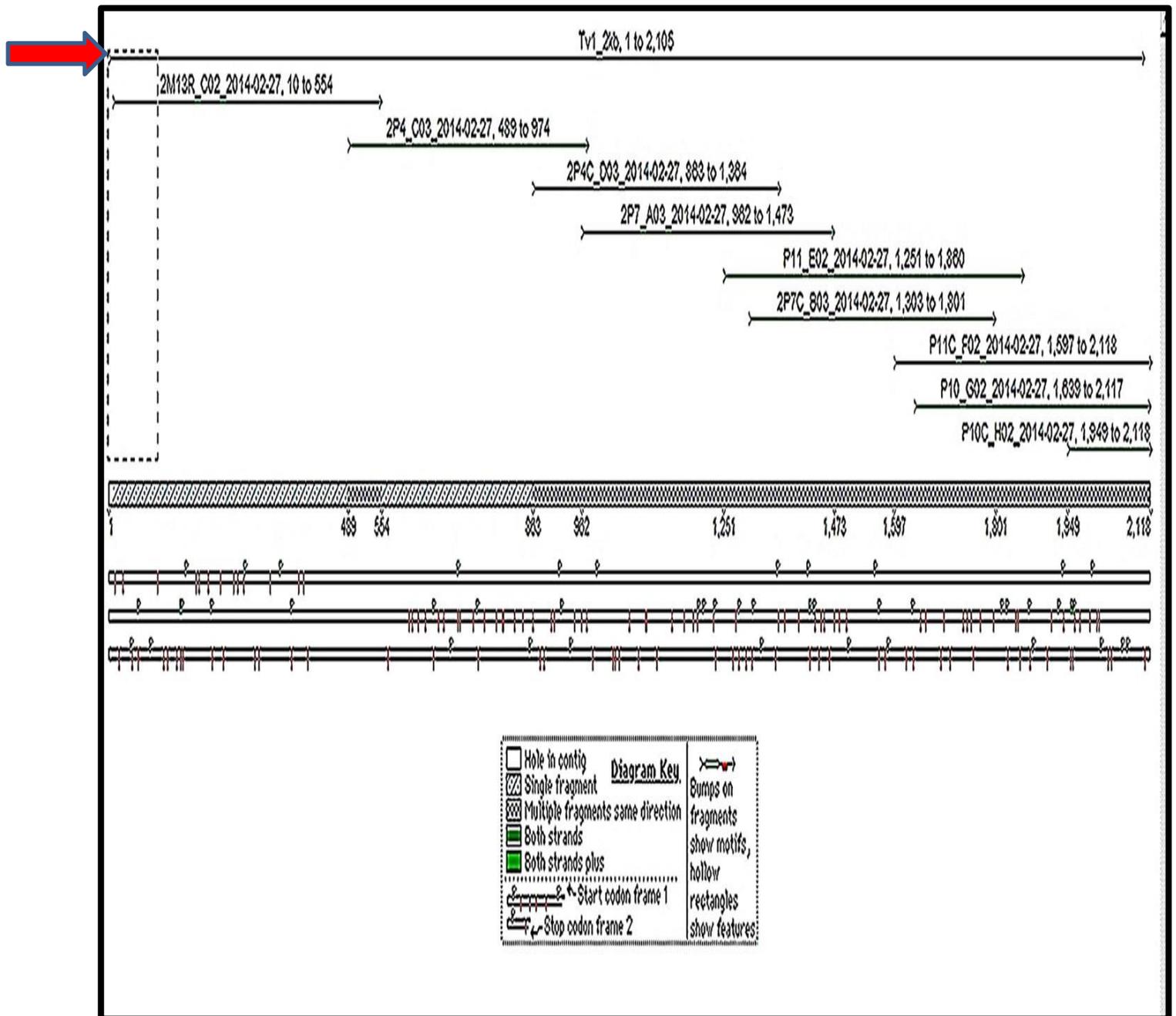
**Section 1-2** - Sequencing results for TV1-2kb and TV1-2.5kb wild-type and mutant subclones, and TV1-FLG; *gag* p24 HIV ELISA and *X-gal* staining for TV1-FLG results that were essential in testing the infectivity of the clone and to assess its potential to be used as a backbone vector to harbour the following mutation combinations:

- A62V,
- K65R,
- M184V
- A62V+K65R
- A62V+M184V
- K65R+M184V
- A62V+K65R+M184V

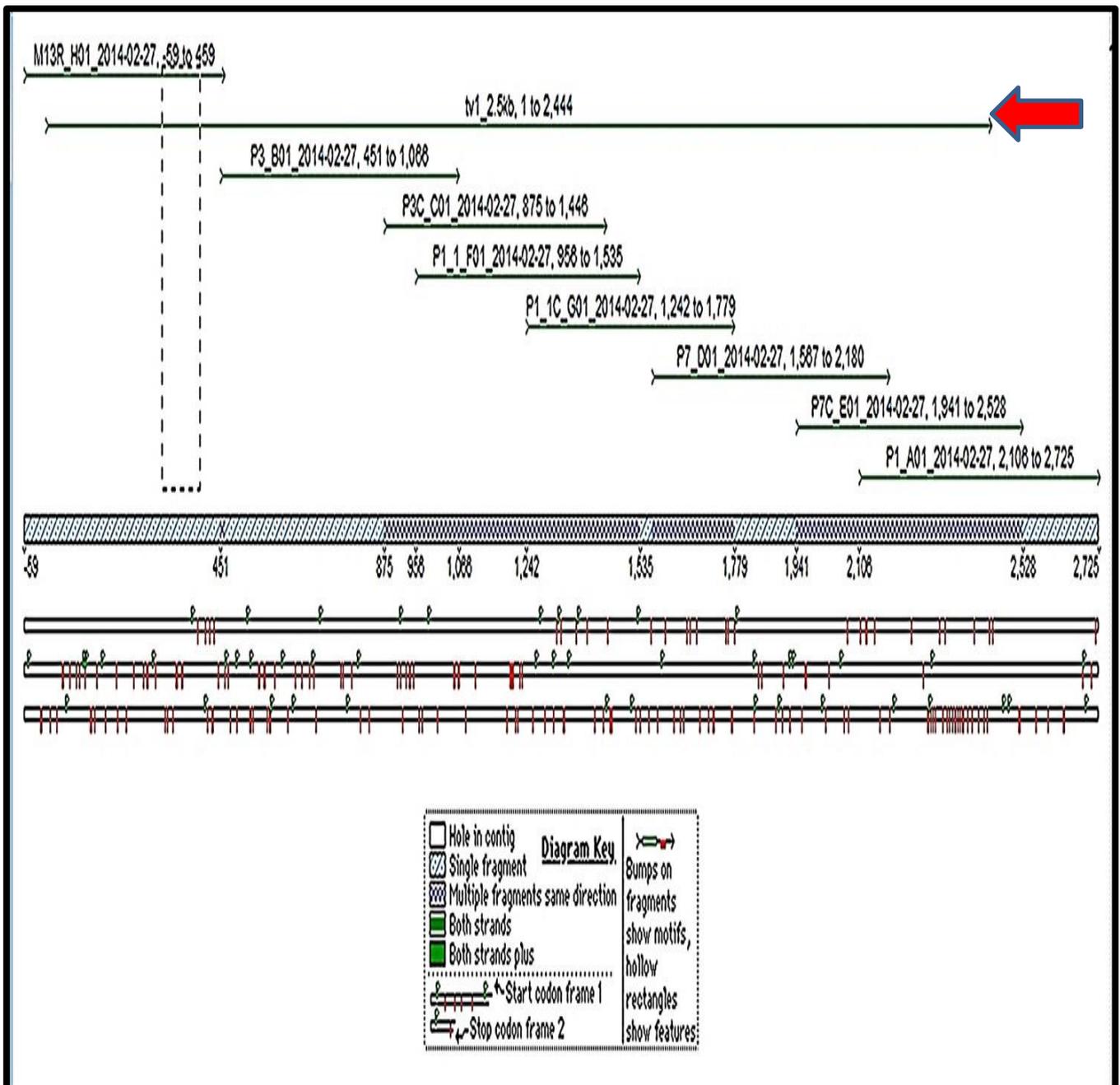
**Section 3** - Infusion™ cloning and Q5™ mutagenesis results; fluorescence imaging and TCID<sub>50</sub> results; qPCR optimization and the growth competition results between A62V+K65R+M184V and K65R+M184V mutant viruses

### **3.1 Sequencing results for TV1-2kb and TV1-2.5kb wild-type and mutant subclones**

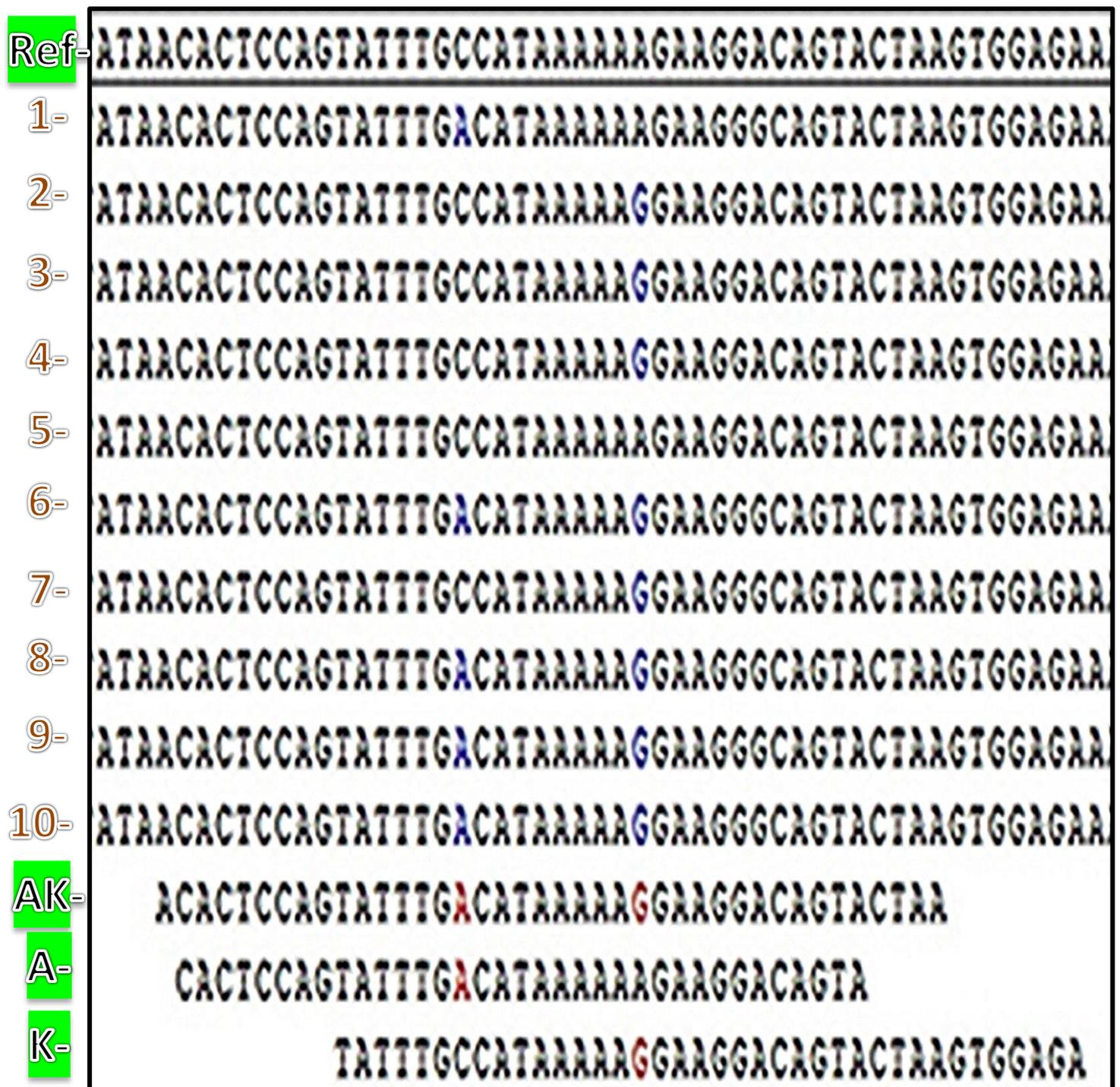
Standard DNA Sanger sequencing was utilized to verify successful cloning of inserts (TV1-2kb and TV1-2.5kb) fragments as shown in **Figure 13** and **Figure 14**, respectively. Sequences generated using the Big dye cycle sequencing kit (Applied Biosystems, USA) and run on the ABI 3130xL Genetic Analyser were analysed by Sequencher v5 (Gene Codes Corporation, Ann Arbor, MI USA) software. The read lengths of each sequence were an average of 485bp and the sequences were trimmed from the edges to remove low-confidence base calls and were about 98% quality (**Addendum I**). Verification of the presence of the desired mutations in the subclones (**Figure 15** and **Figure 16**) was confirmed using the Sequencher v5 program.



**Figure 13.** Contig tree drawn using Sequencher v 5. The top left red arrow indicates the reference TV1-2kb wild type sequence aligned to the primer generated sequences (line arrows shown below) that overlap each other and verify the presence of the insert in p-Gem T easy Vector.



**Figure 14.** Contig tree drawn using Sequencher v 5. The top right **red** arrow indicates the reference TV1-2.5kb wild type sequence aligned to the primer generated sequences that overlap each other and verify the presence of the insert in p-Gem T easy Vector.



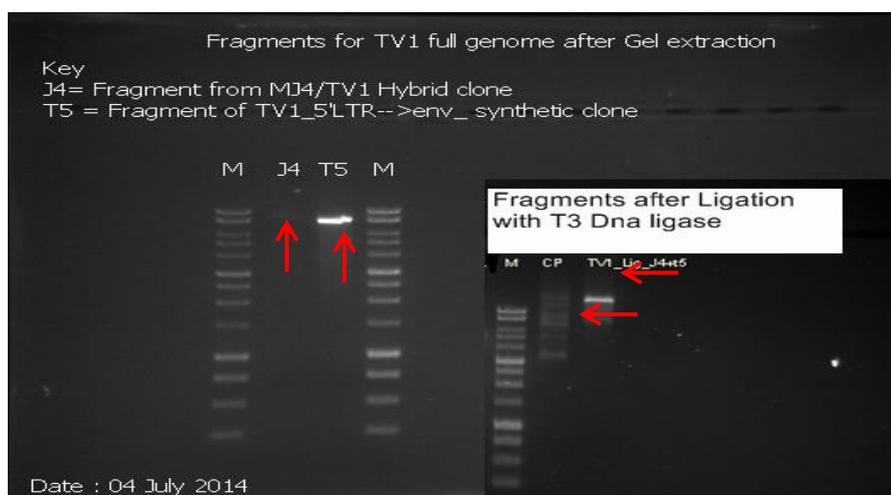
**Figure 15.** Image cropped from Sequencher v5 showing base change confirmations (from C--> A highlighted in **blue**) for A62V mutation on A62V (1), A62V+ K65R (6), A62V+K65R+M184V (8-10), sample clones. Base change confirmation (from A--> G highlighted in **blue**) of K65R mutation on A62V+ K65R (6), A62V+K65R+M184V (8-10), K65R (2-4) and K65R+M184V (7) mutant clones. The wild-type reference sequence is indicated at the top and primer sequences aligned, are indicated at the bottom. (5) M184V sample clone.



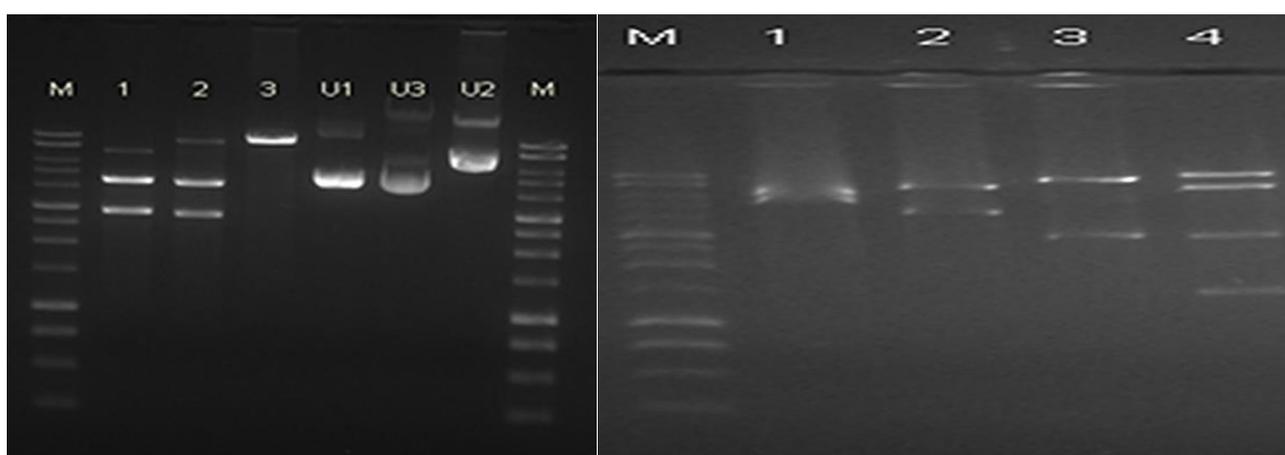
### 3.2 TV1-FLG

#### 3.2.1 Ligation and restriction enzyme digestion results

**Figure 17A** below indicates gel photo of the two fragments needed for TV1-FLG synthesis prior to ligation and after ligation respectively. **Figure 17B** indicates restriction enzyme digestion confirmation of TV1-FLG cloning. The TV1-FLG HIV-1 subtype C constructs synthesized in this study was the final ligation product of the 5' LTR-> start of *env* fragment (**T5**) from the consensus TV1 HIV-1 full genome sequence enzymatically digested and gel purified from the S clone and the *env*-> 3' LTR of TV1 (**J4**) derived from the MJ4/TV1 hybrid clone.



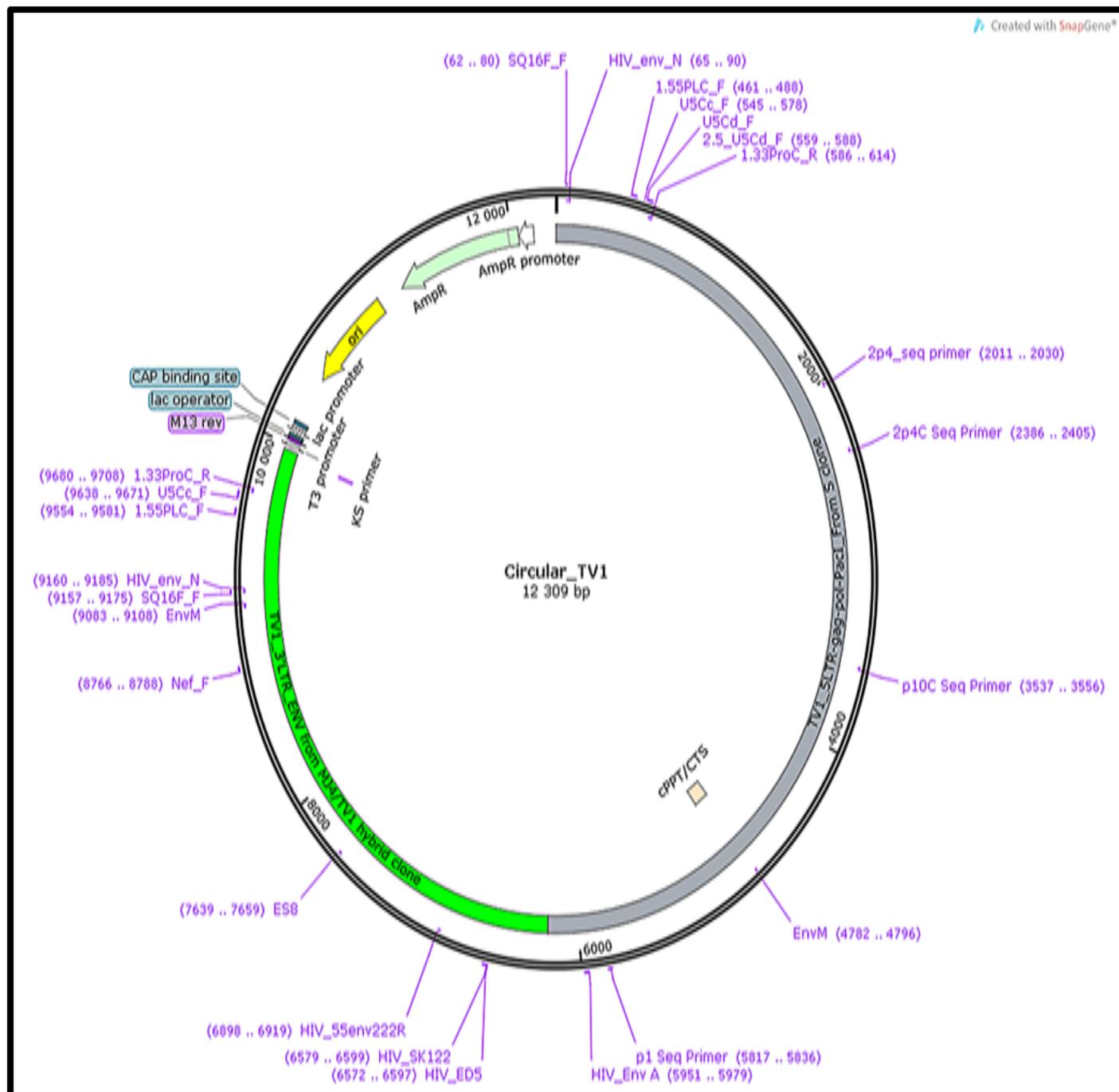
**Figure 17A.** Gel photos showing the linear fragments used for synthesis of TV1-FLG plasmid – **J4** (from MJ4/TV1 hybrid) and **T5** (from S-clone). Right image shows confirmation of ligation for both the ligation control and TV1-FLG (**TV1\_Lig\_j4+t5**).



**Figure 17B** Gel photo showing restriction enzyme digestion confirmation of TV1-FLG. TV1-FLG plasmid is digested with a panel of restriction enzymes according to the expected map of TV1-FLG (shown in **Addendum H**). U1, U3 and U2 are uncut controls, digest samples 1-4 M = 1KB molecular weight marker (Promega, USA)

### 3.2.1 Sequencing results of TV1- FLG

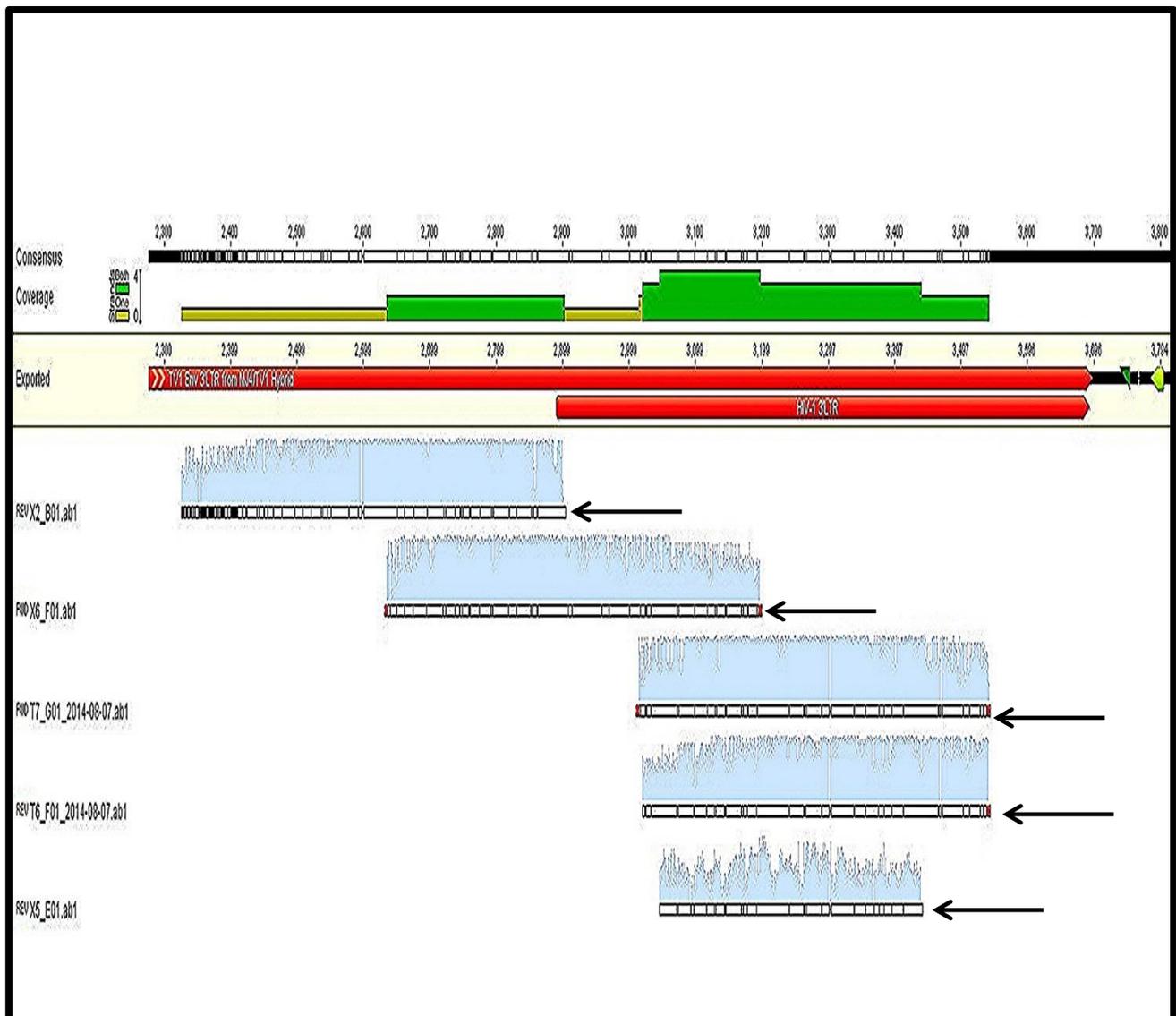
The TV1-FLG plasmid was partially sequenced (**Figure 18**) by primers targeting *env* and *nef* overlapping regions (HXB2 map positions 8526-9157 and 8766-9375 on TV1-FLG), 3'LTR end of HIV and the 5'LTR -> end of *pol*



**Figure 18** Image taken from Snap Gene software (GSL biotech, USA) showing binding positions of primers used for sequencing TV1-FLGs (labelled in purple) annealing to the TV1-FLG genome

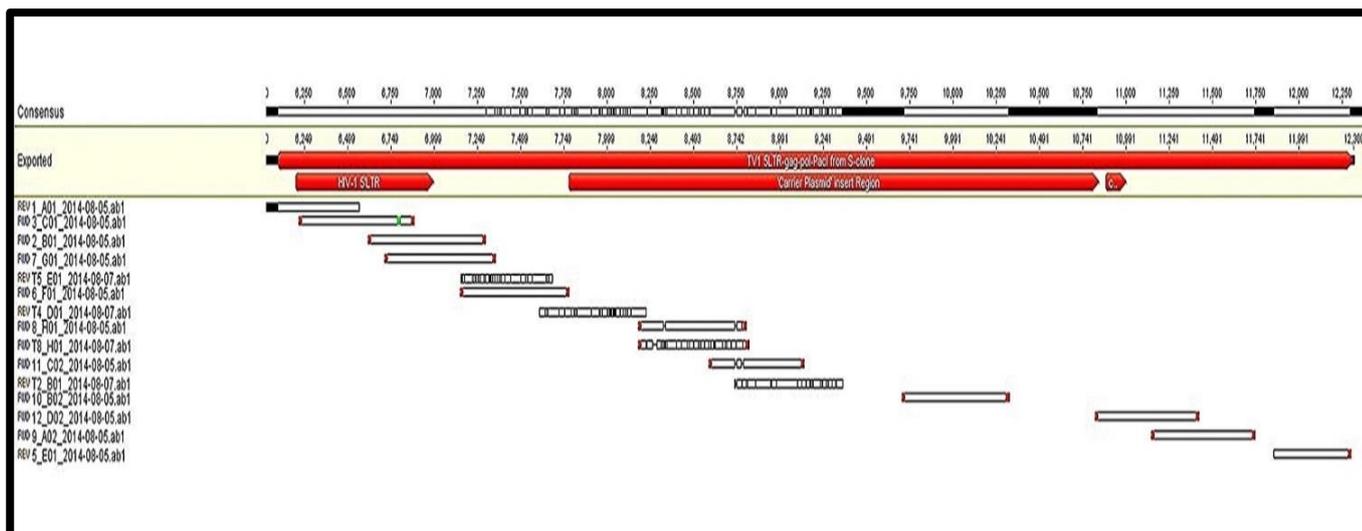
Only one primer (Nef\_f) yielded an *env* region sequence (**Figure 19**). Sequencing reactions using the other *env* –targeting primers were not successful, even after optimising conditions.

Nevertheless, sequences were obtained for the 5'LTR – end of *pol* region of TV1-FLG (**Figure 20**).



**Figure. 19** Image taken from Geneious R6 showing sequences (indicated by black arrows) generated by primers (From top to bottom - Nef\_f, U5Cc\_F, 1.33ProC, 1.55PLC\_f and M13 rev) annealing to the TV1-FLG 3'LTR and *env* and *nef* overlapping regions shown in red above.

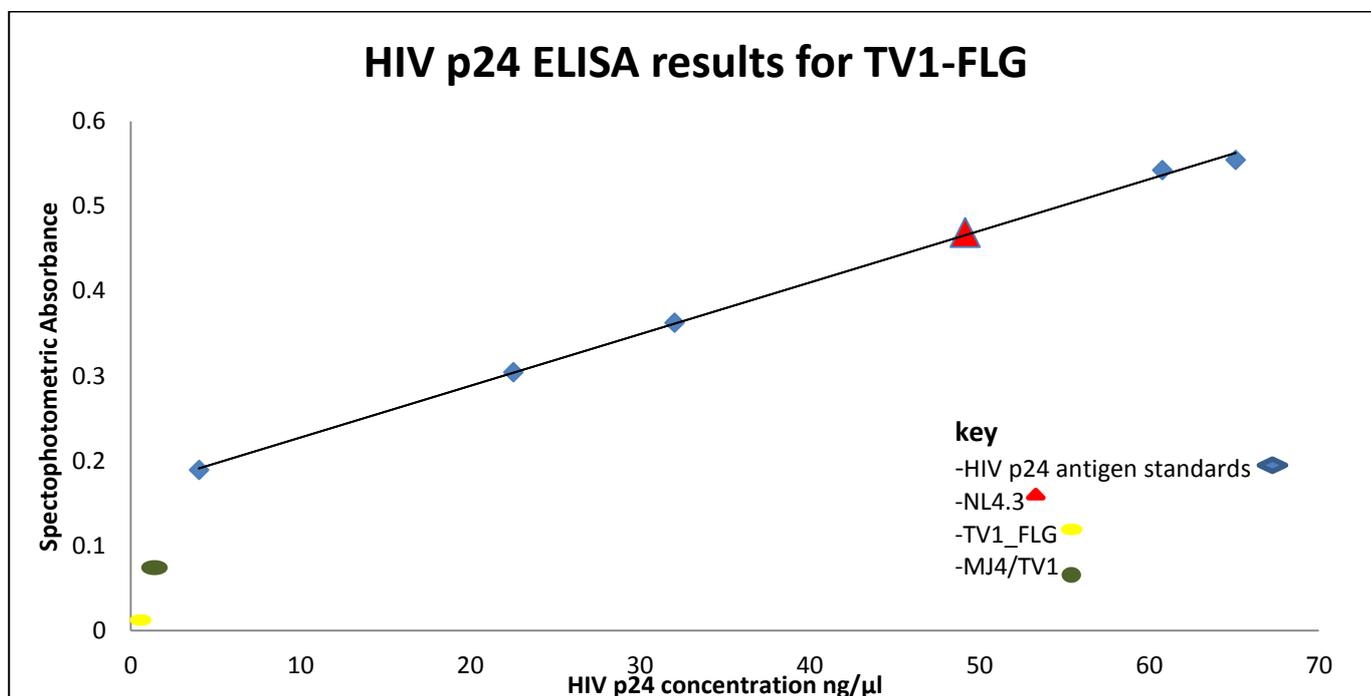
In each sequencing result for TV1-FLG, the primer sequences generated are also indicated with internal black markings that indicate mismatches or ambiguities in the sequence when cross referenced with the TV1 consensus sequence. Despite not being able to obtain a full genome sequence, the results suggest successful cloning of the target region into the TV-1 backbone and a logical move to proceed to test for infectivity in cell culture using HIV- p24 antigen or *X-gal* staining assays.



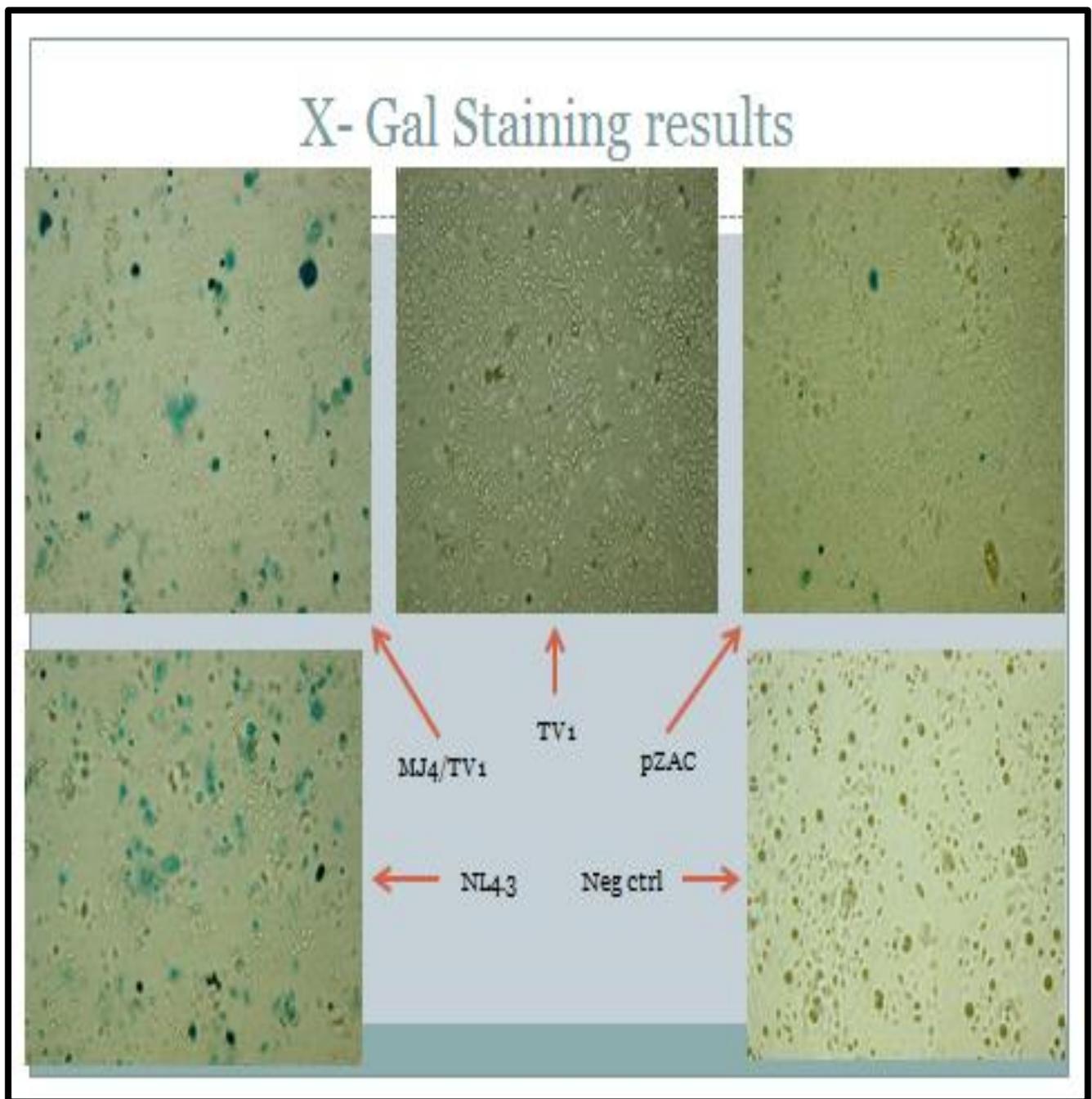
**Figure.20** Image taken from Geneious R6 showing sequences generated by primers annealing to the TV1-FLG 5'LTR – end of pol shown in red above.

### 3.2.2 HIV gag p24 ELISA and X-gal staining results for TV1-FLG

HIV p24 ELISA results (**Figure 21**) using the Quick Titer™ HIV Lentivirus Quantification kit (Cell Biolabs, Inc., San Diego USA) showed TV1-FLG to be no infectious in HEK 293T cells and this was further confirmed by X-gal staining in TZM- bl cells (**Figure 22**)



**Figure 21.** TV1-FLG HIV-p24 concentration estimation using linear regression with known p24 standards. The figure legend indicates the NL4.3 (positive control) plasmid, HIV p24 antigen standards, and TV1-FLG and MJ4/TV1 plasmids.



**Figure 22.** Image taken by a light microscope showing results for TV1-FLG X-gal staining. Blue cells are indicative of HIV-1 infection

X-gal staining was negative for TV1-FLG but positive for NL4.3, pZAC and MJ4/TV1 hybrid plasmids. The image results represent the moderate staining in each of the respective wells that contained TZM-bl infected cells and images shown in figure 22 are the wells which were concluded subjectively to be the best representation of image analysis.

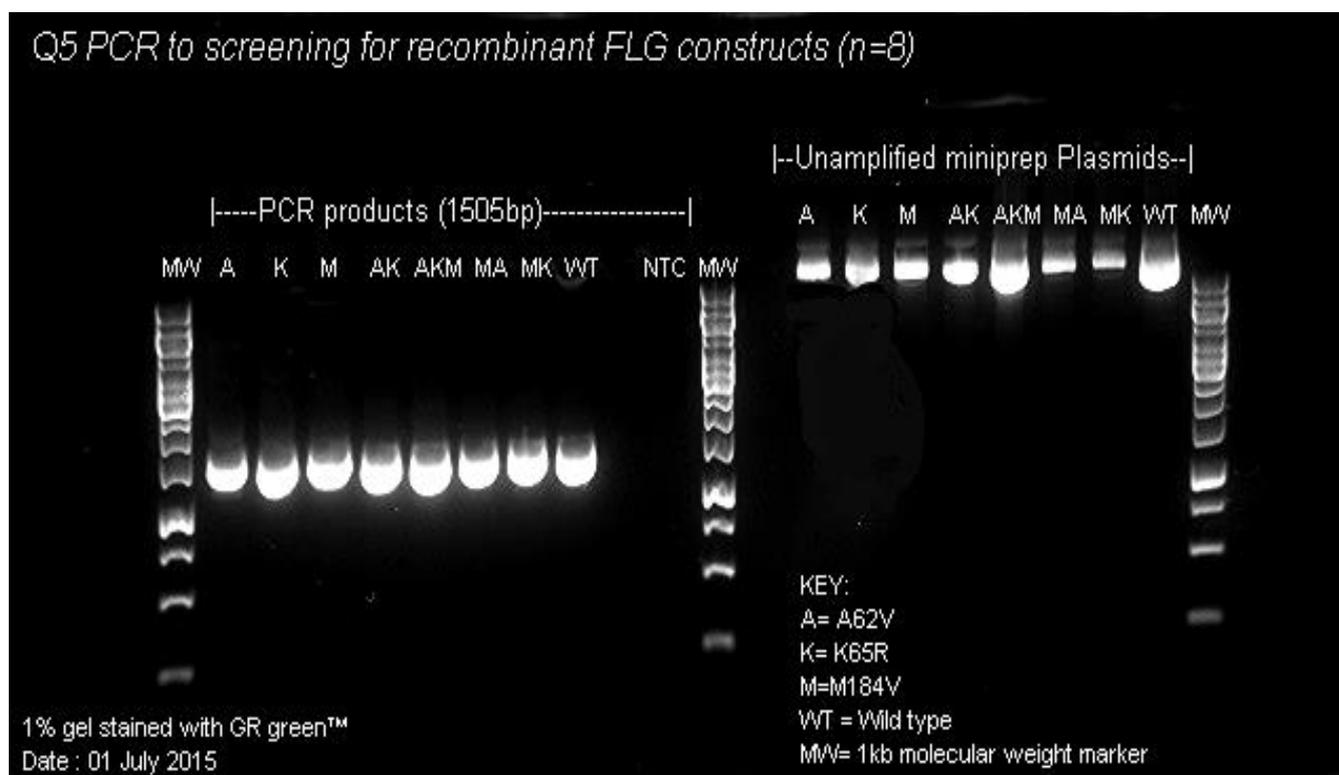
### 3.3 Results from plan to subclone in NL4.3 Subtype B infectious Plasmid

Cloning and mutagenesis experiments resulted in HIV-1 BC recombinants. HIV-1C constructs with the gag-pol gene were derived from the MJ4 HIV-1C clone and could be used to assess HIV-1C mutation interactions.

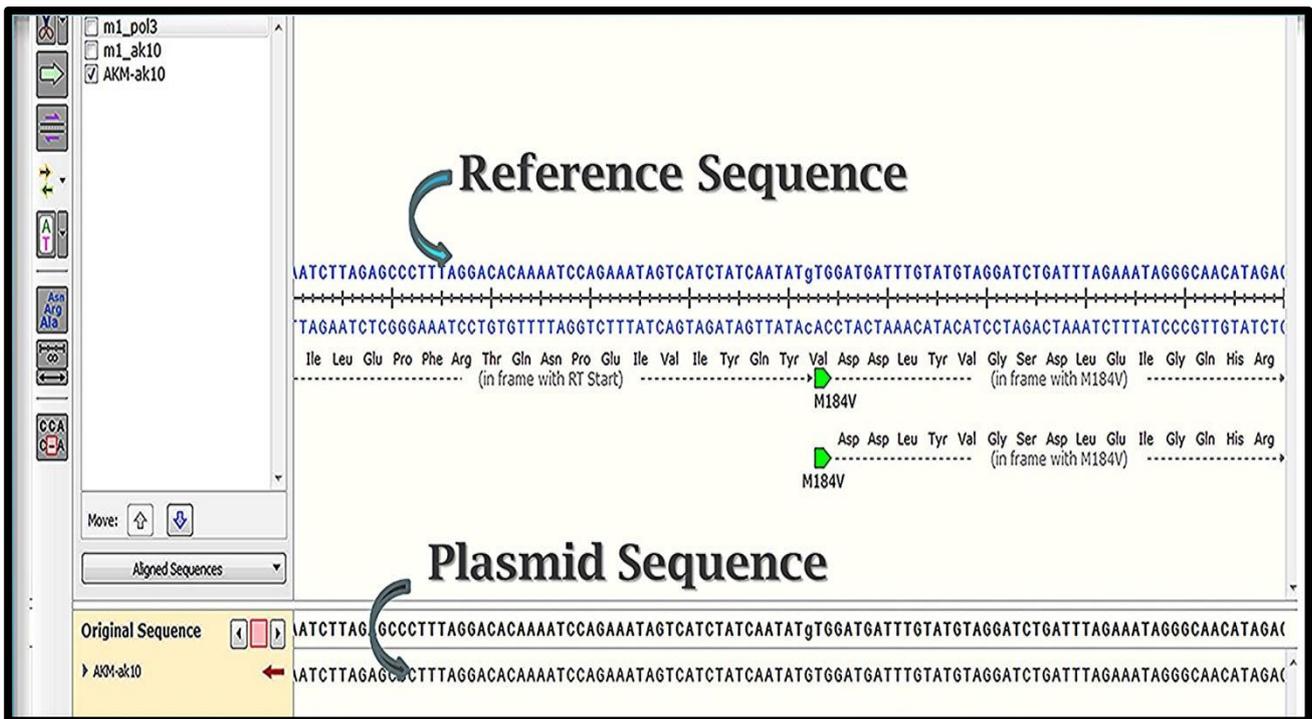
#### 3.3.1 Infusion™ cloning and Q5™ mutagenesis results

The MJ4 gag-pol gene (nucleotide positions 2010 -3493 on the HXB2 map) was successfully fused to the NL4.3 eGFP prepared vector and used as the starting template to derive all the full-length genome (FLG) mutant recombinants required by the first objective of this study. In summary, a total of eight HIV-1BC recombinants were synthesized and confirmation of the constructs were obtained via PCR amplification of the inserts using the infusion primers designed for the study and DNA Sanger sequencing.

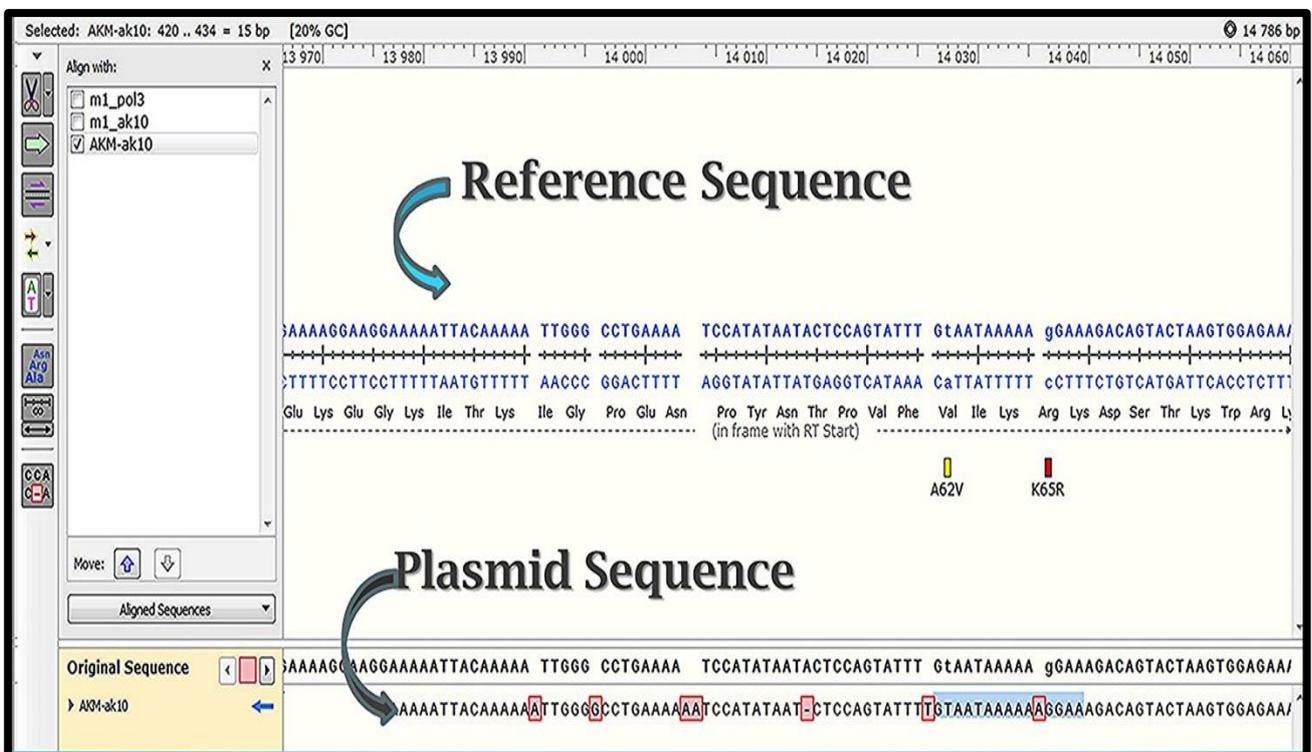
Gel electrophoresis analysis after PCR amplification and DNA Sanger sequencing (specifically for M184V+K65R and A62V+K65R+M184V used in the growth competition experiment) results are shown in **Figure 23** and **Figure 24 (A-D)**. Images showing sequence results were taken using SnapGene program (GSL Biotech, USA)



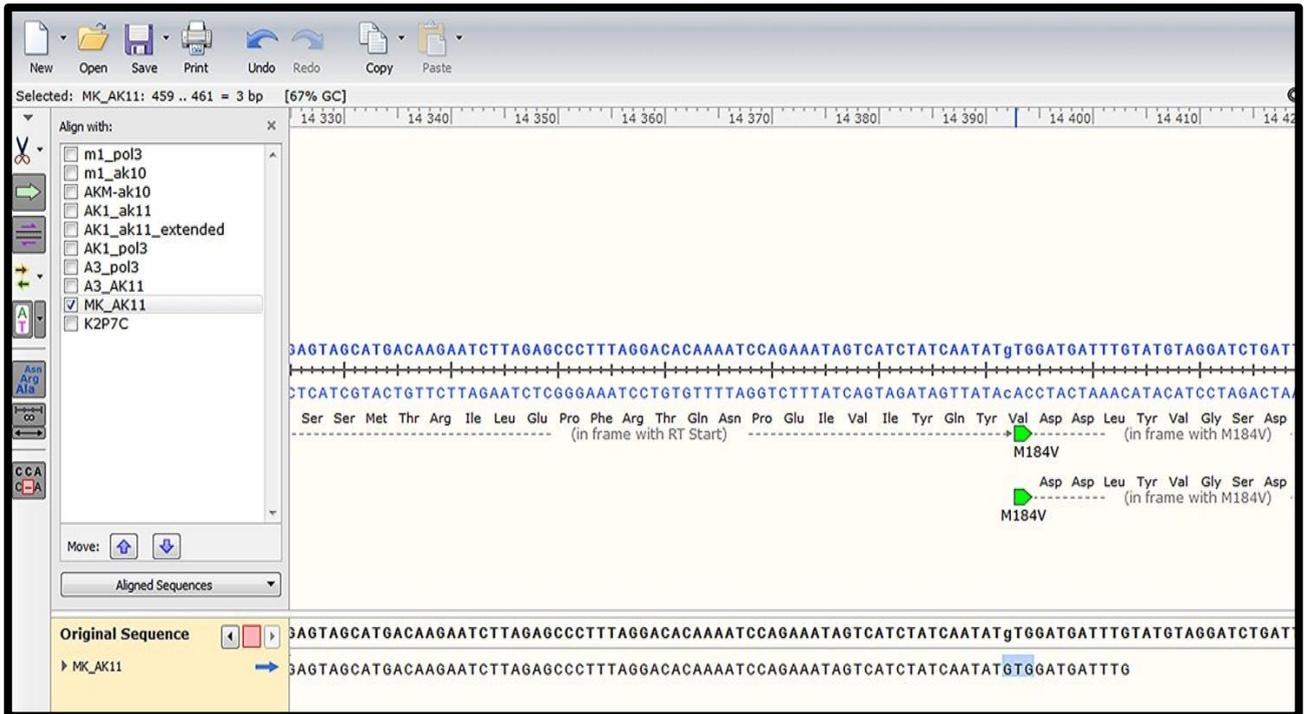
**Figure 23.** Gel photo showing PCR confirmation of HIV-1BC recombinants. The unamplified plasmids are also shown on the gel and serve as controls to verify the size of the full length genome (FLG) constructs.



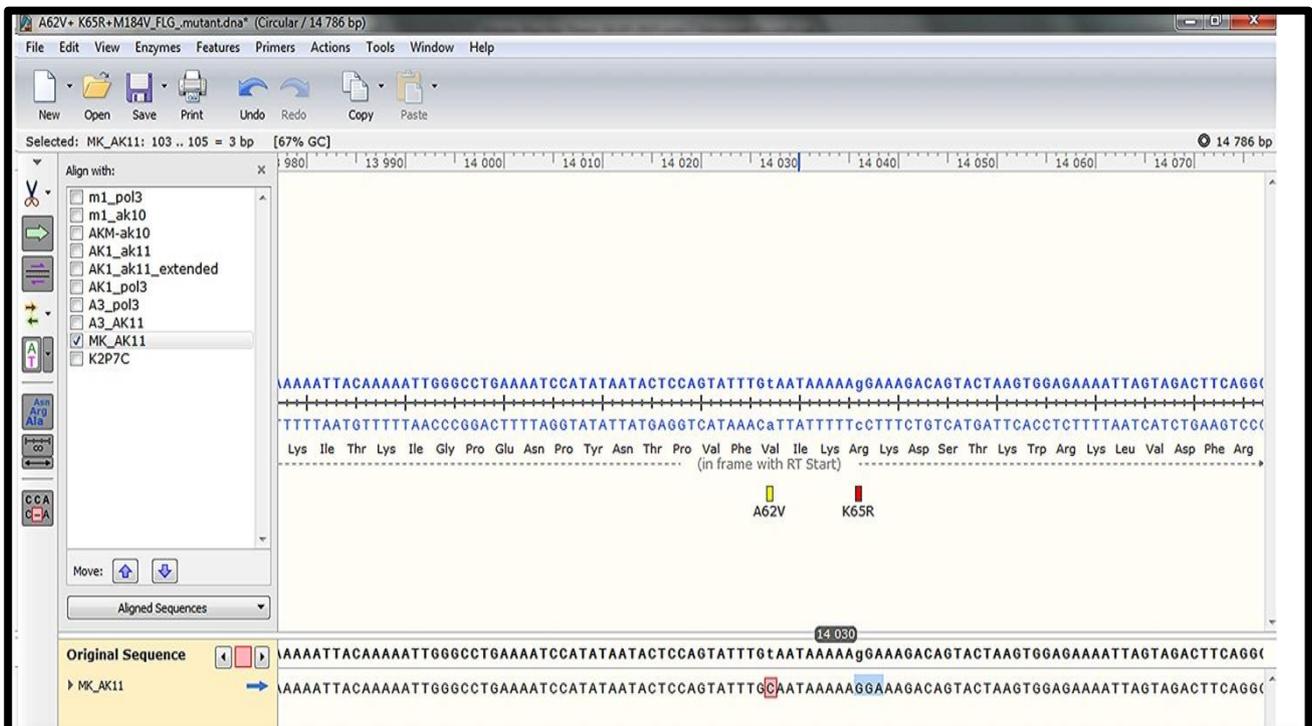
**Figure 24A.** DNA Sangers sequencing results to show proof of insertion of M184V mutation on **A62V+K65R+M184V** construct using Q5™ Mutagenesis technique.



**Figure 24B.** DNA Sanger sequencing results to show proof of insertion of K65R +A62V mutations on **A62V+K65R+M184V** construct using Q5™ Mutagenesis technique.



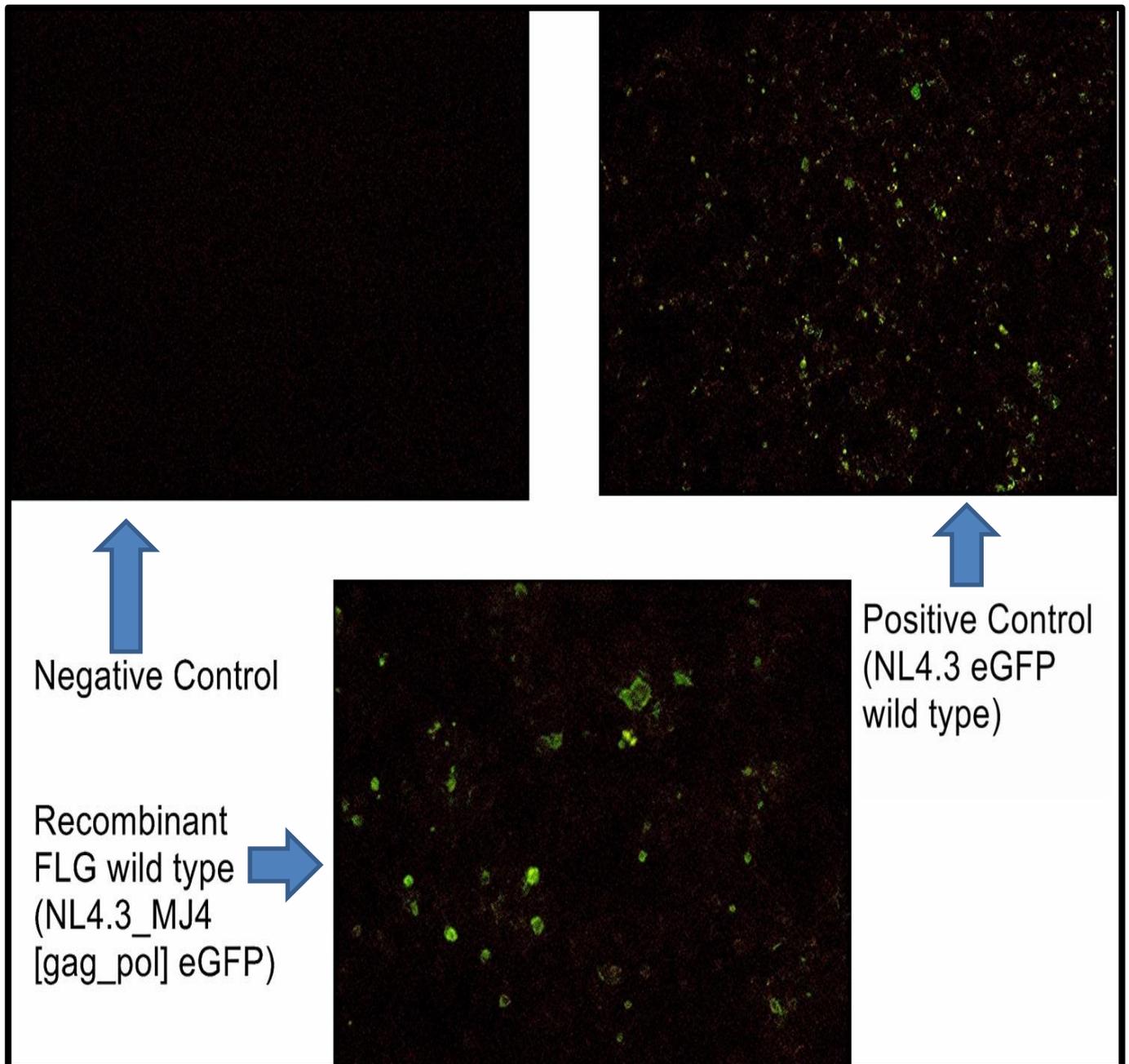
**Figure 24C.** DNA Sanger sequencing results showing proof of insertion of M184V mutation on K65R+M184V construct using Q5™ Mutagenesis technique. Reference sequence highlighted in blue and plasmid sequence highlighted in black



**Figure 24D.** DNA Sanger sequencing results showing proof of insertion of K65R but not A62V mutation on K65R+M184V construct using Q5™ Mutagenesis technique. Reference sequence highlighted in blue and plasmid sequence highlighted in black

### 3.3.2 Fluorescence imaging results

The optimal harvest point and validation of the transfection experiment for the HIV-1 BC recombinants was determined to be 22.5hrs post-transfection and fluorescence image results of the wild-type recombinant virus synthesised in comparison to pure Subtype B NL4.3 eGFP infectious virus are shown in **Figure 25** below:



**Figure 25.** Fluorescence image results to confirm validation of transfection experiment using synthesized recombinant constructs and lab adapted NL4.3 eGFP as a control. Green fluorescence expression of eGFP is indicative of successfully transfected cells. Images acquired 22.5hrs post-transfection.

### 3.3.3 TCID<sub>50</sub> end point titration results for A62V+K65R+M184V and K65R+M184V

The TCID<sub>50</sub> end point titration assay employing X Gal staining/ β- galactosidase detection in TZM-bl cells for A62V+K65R+M184V and K65R+M184V mutant viruses showed K65R+M184V virus to be having a high concentration of replication competent virus. The results for both viruses were used to normalize for the starting concentration of the growth competition experiment between the two viruses. The plate results for both viruses obtained after the experiment are shown below:

**Plate A: - A62V+K65R+M184V X-gal staining/β- galactosidase detection in TZM-bl cells results**

	1	2	3	4	5	6	7	8	9	10	11	12
A	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
B	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
C	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
D	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
E	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
F	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
G												
H	Neat(AKM)	Neat(AKM)										

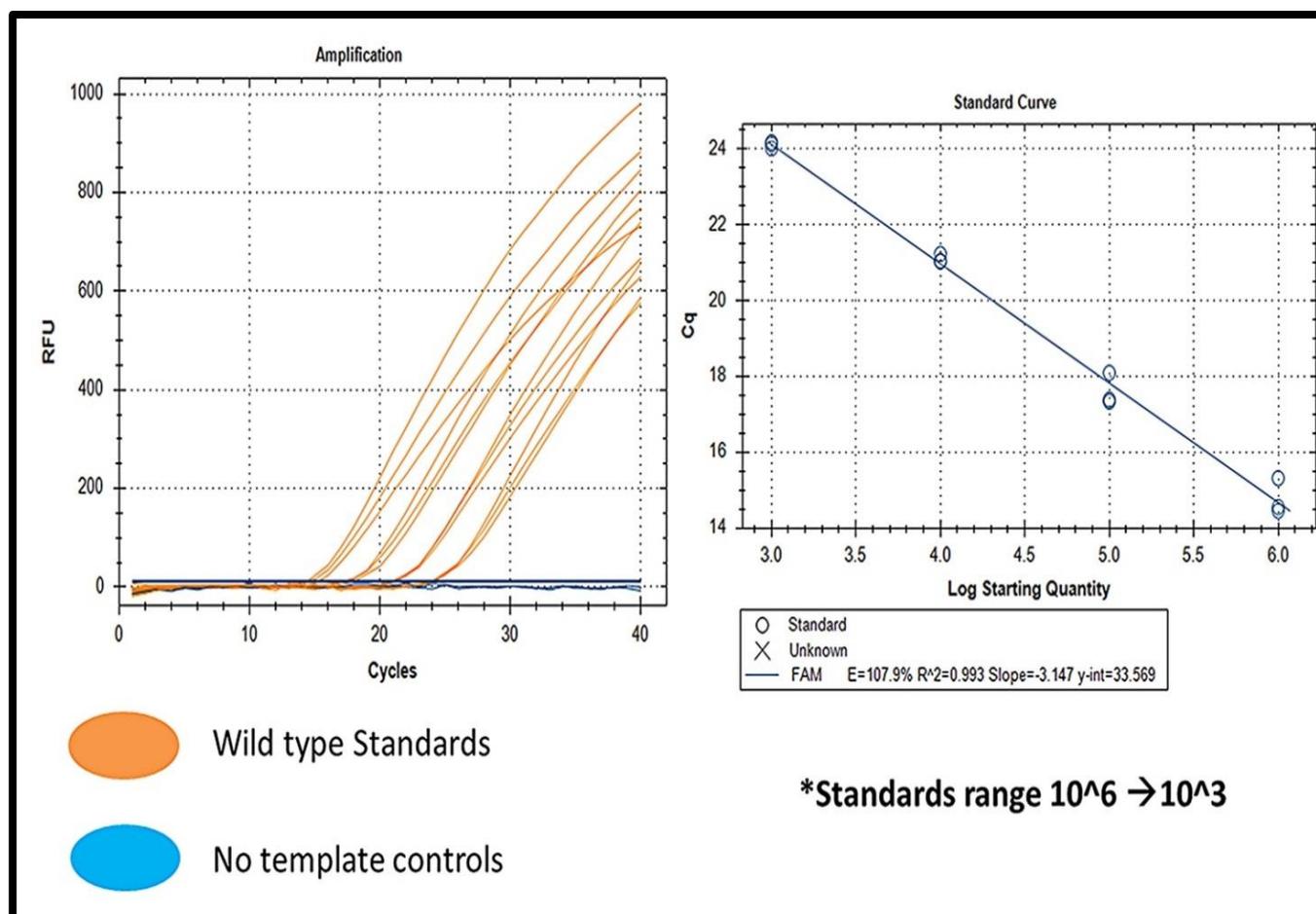
**Plate B: - K65R+M184V X-gal staining/β- galactosidase detection in TZM-bl cells results**

	1	2	3	4	5	6	7	8	9	10	11	12
A	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
B	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
C	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
D	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
E	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
F	10 <sup>2</sup>	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>5</sup>	10 <sup>5</sup>	10 <sup>6</sup>	10 <sup>6</sup>	Negative	
G												
H	Neat(MK)	Neat(MK)										
	Positive											
	Negative											

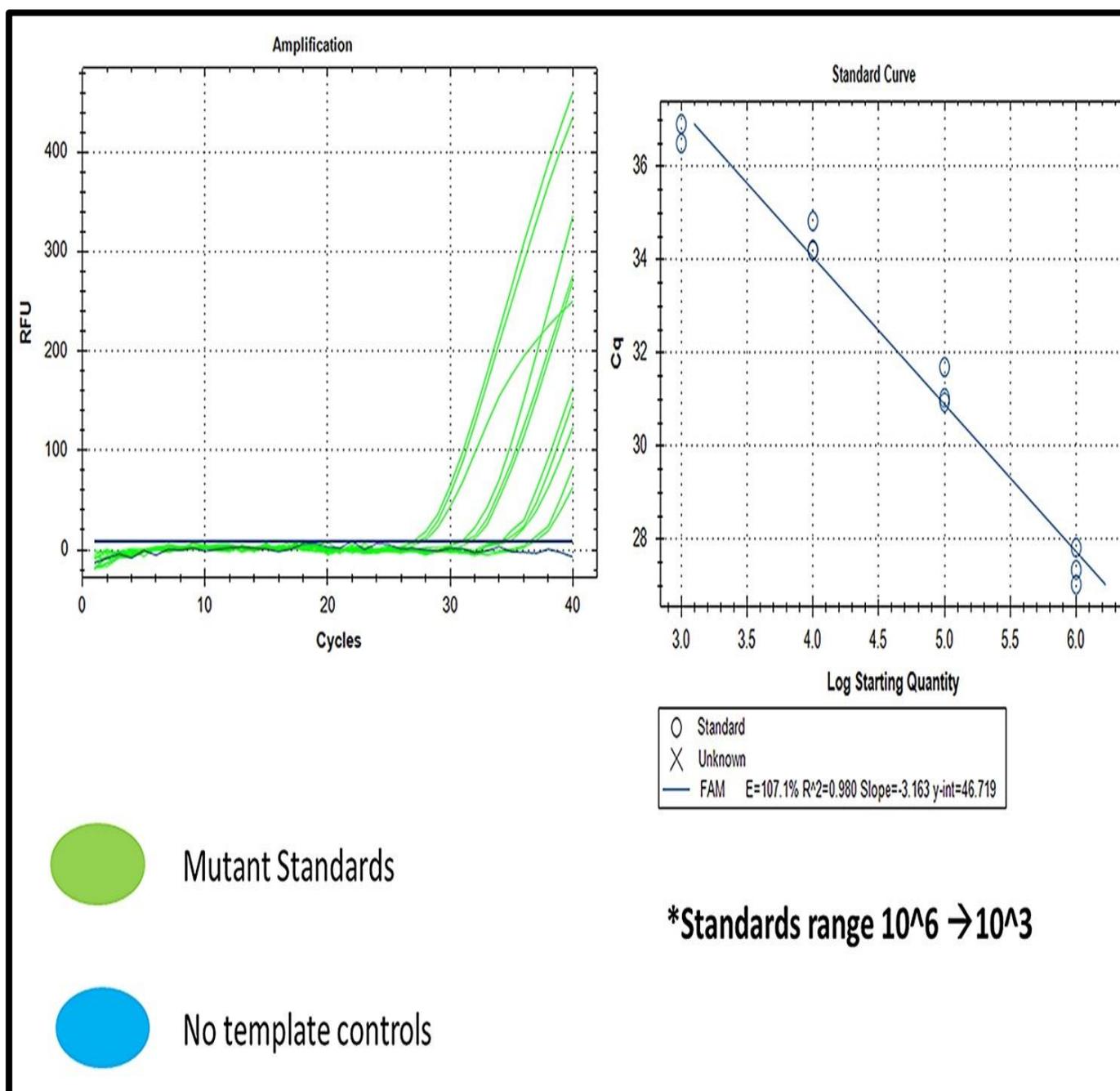
The TCID<sub>50</sub> values for both viruses were calculated using the Spearman-Kärber formula described in **section 2.6.3.2** with data acquired from the plates and were  $2.42E+03$  and  $6.32E+03$  TCID<sub>50</sub>/ml and this is equivalent to  $1.36E+03$  and  $3.54E+03$  replication-competent copies per ml for A62V+K65R+M184V and K65R+M184V viruses respectively. The values were used to adjust the viral inoculums for the head to head growth competition experiment.

### 3.3.4 Allele-specific real-time PCR optimization results

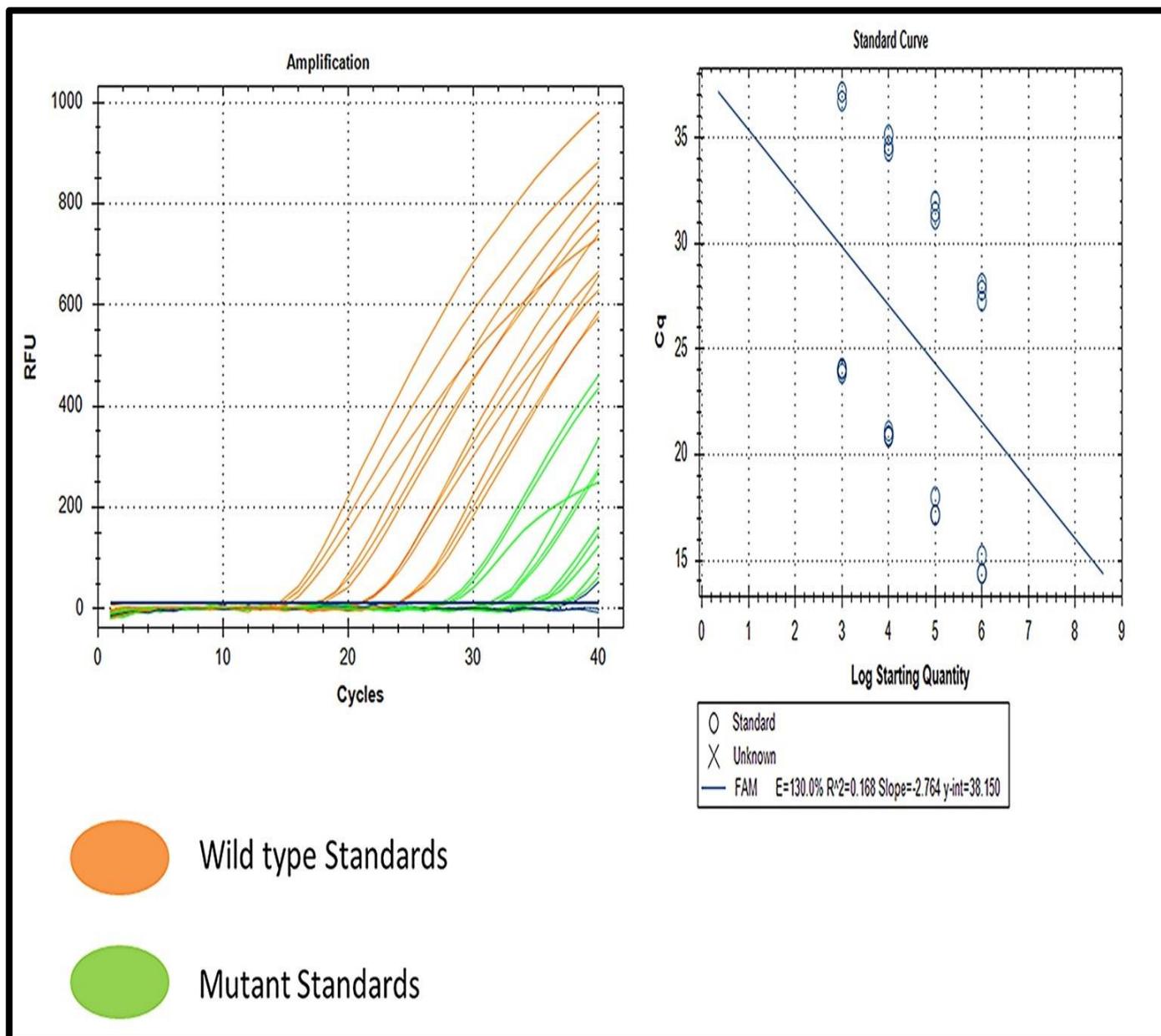
After optimization of the allele-specific real-time PCR assay, WT-primer detected specific WT product approximately 13 Cts before the mutant plasmid when assayed at the same concentration (whereas the total assay using the T-fwd primer quantified them equally). This provided a high specificity for the WT allele-specific assay as 13 Cts equates to a  $\sim 10^4$ -fold difference in detection of the 'true' wild type signal vs the 'false' mutant signal. When used in parallel, the total assay could therefore be used to quantify reliably WT and mutant (by subtracting wild-type from the Total virus) concentrations. The qPCR assay was analysed using the Relative Standard Curve method with the Bio-Rad CFX manager software v3.1 (Bio-Rad, USA). The amplification plots and standard curves for the WT-primer are shown below (**Figure 26 A-C**) and include an individual and overall analysis of both the wild type and mutant standards to show a 13 Ct difference. Real time PCR results of the T-forward primer on the wild type and mutant standards is shown in **Figure 26 D**. The assay was in the accepted standard range for qPCR experiments with an average reaction efficiency of 107.5% and an  $R^2$  value of  $\geq 0.98$  and average slope of -3.2. No-template controls (NTCs) included did not give a signal ruling out the possibility of contamination in the experiments.



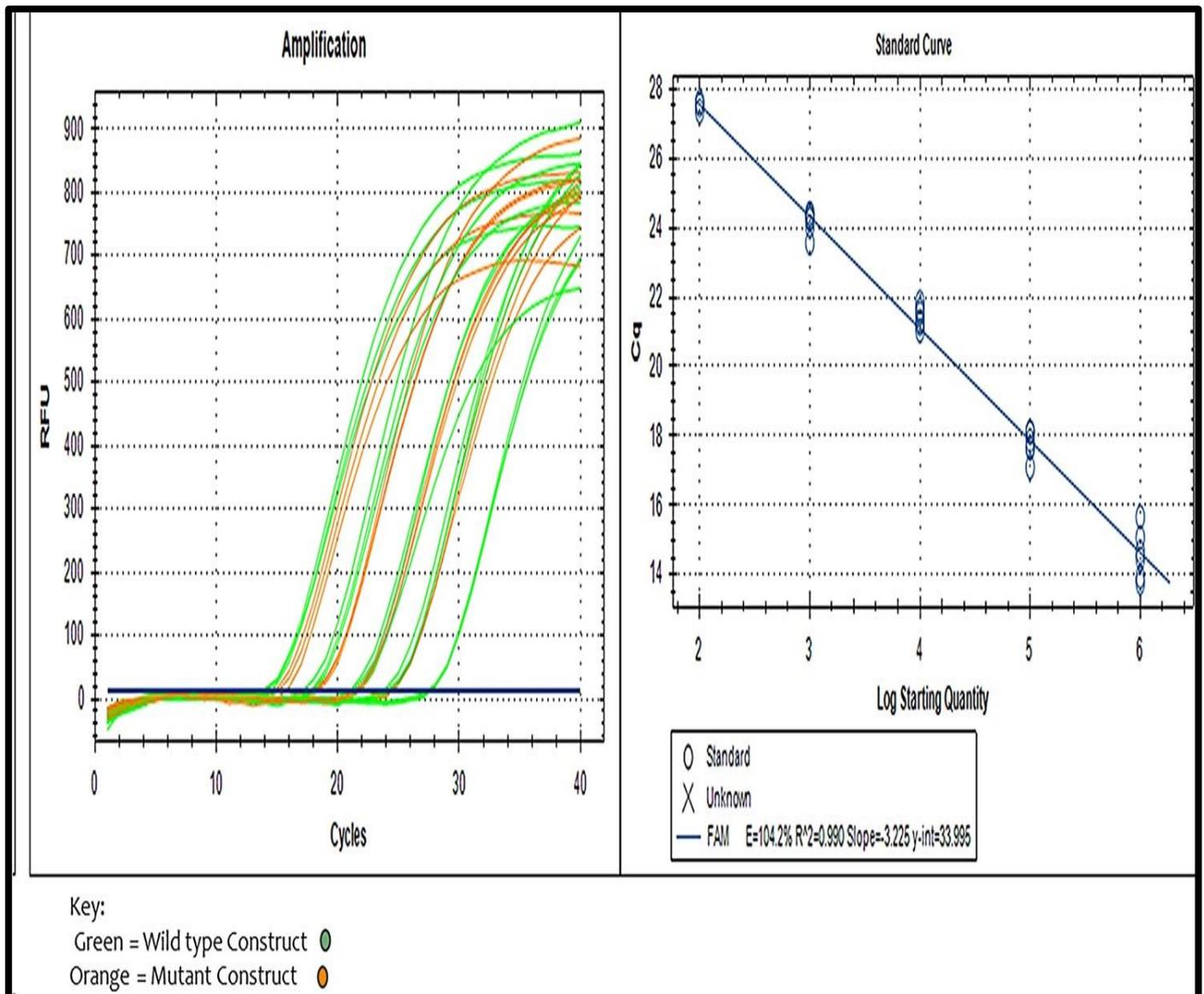
**Figure26A.** WT primer analysis on wild type standards



**Figure26B.** WT-primer analysis on Mutant (A62V) standards



**Figure26C.** Overall analysis of WT-primer on Wild type and Mutant (A62V) standards



**Figure 26D.** *T-fwd primer Wild type and Mutant (A62V) standards optimization* The wild type and mutant amplicons are both detected at the same Ct value confirming that there are at the same concentration and non-specificity of the T-forward primer.

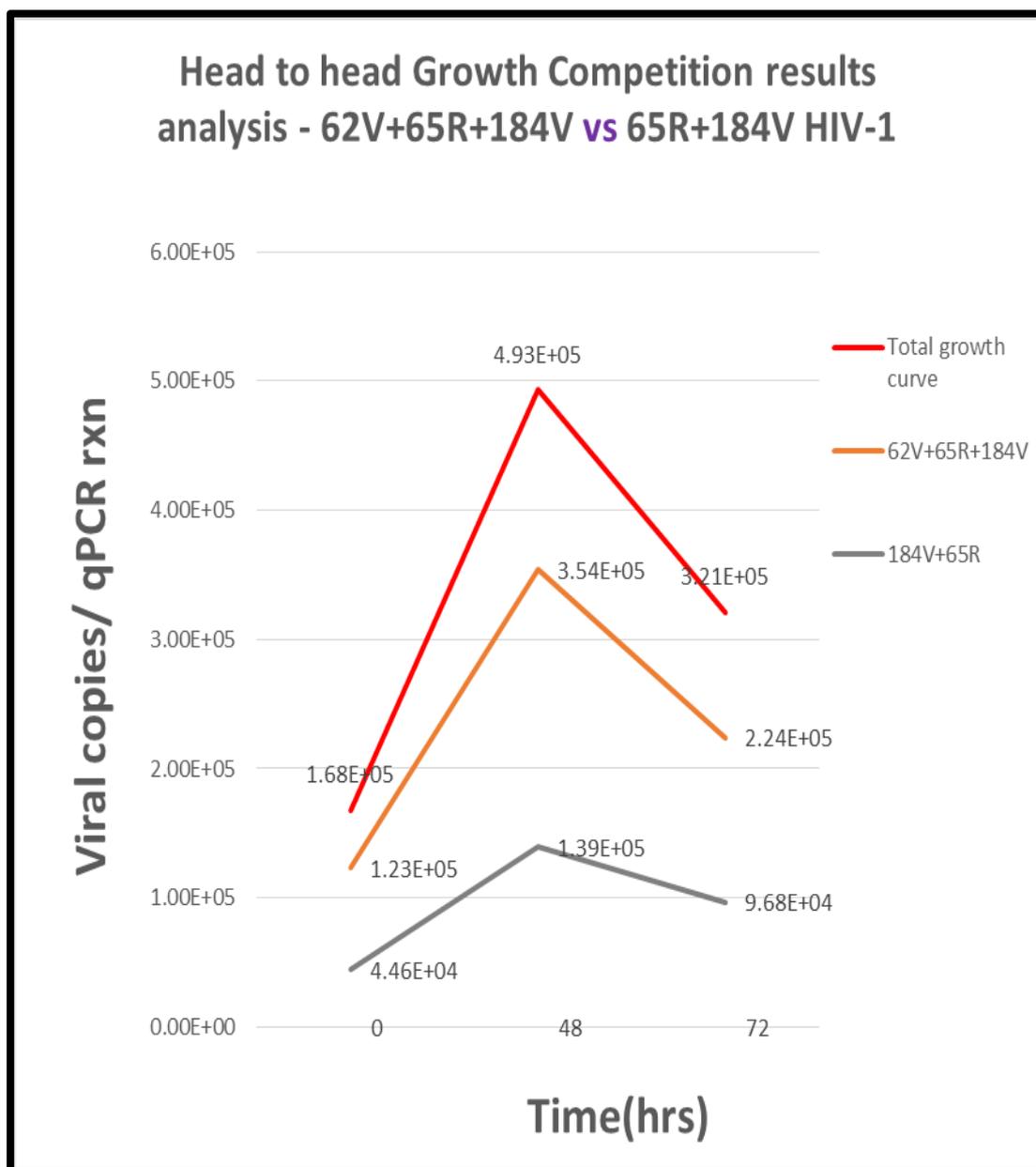
### 3.3.5 Growth competition experiment results for A62V+K65R+M184V vs K65R+M184V

The results obtained after qPCR analysis of *A62V+K65R+M184V* vs. *K65R+M184V* growth competition samples were analysed with Bio-Rad CFX manager software v3.1 (Bio-Rad, USA) before being exported to an excel file for graphical analysis and finally to program R v3.2.2 (R Development Core Team, 2008) for statistical analysis. The quantification data for the experiment is represented in **Table J** and the values are used to generate graphical results shown in **Figure 27**. **Table K** shows the overall data and the raw values obtained after the qPCR experiment.

**Table J. Quantification data for head to head growth competition experiment using WT and T- Forward primers to detect M184V+K65R and both viruses respectively.**

Key: - **blue** = 1<sup>st</sup> replicate; **lime green** = 2<sup>nd</sup> replicate; **grey** = 3<sup>rd</sup> replicate; **yellow** = Geometric mean of 3 biological replicates at a given sampling time; **Orange** = values used to plot overall growth curves

<i>WT- primer Assay</i>		<i>T-Forward - primer Assay</i>	
Replicates/time (hrs)	Viral copies/qPCR rxn	Replicates/time (hrs)	Viral copies/qPCR rxn
<b>G1</b>	Geometric mean of 2 technical replicates per sampling time	<b>G1</b>	Geometric mean of 2 technical replicates per sampling time
0	3.42E+04	0	1.72E+05
48	1.52E+05	48	5.03E+05
72	1.04E+05	72	3.49E+05
<b>G2</b>		<b>G2</b>	
0	5.97E+04	0	2.40E+05
48	1.32E+05	48	6.49E+05
72	8.41E+04	72	4.20E+05
<b>G3</b>		<b>G3</b>	
0	4.36E+04	0	1.14E+05
48	1.55E+05	48	3.67E+05
72	9.27E+04	72	2.26E+05
	<b>Geomean (G1+G2+G3)</b>		<b>Geomean (G1+G2+G3)</b>
0	4.46E+04	0	1.68E+05
48	1.39E+05	48	4.93E+05
72	9.68E+04	72	3.21E+05



**Figure 27.** Average growth curve plot for A62V+K65R+M184V vs. K65R+ M184V viruses. A62V+K65R+M184V growth curve is derived from the subtraction of K65R+M184V curve from the total growth curve generated from the T-Fwd primer qPCR assay. The growth curve values used to plot the curves are the arithmetic mean of the 3 biological replicates used in the head to head growth competition assay for the two viruses. Raw data is shown below in **Table J**.

The experiment was validated for reproducibility by a calculation of the Coefficient of Variation (CV) which was found to be 13%. The CV-value given is computed using qPCR quantification data from the three biological replicates used in the growth competition experiment and is shown in the **table K**

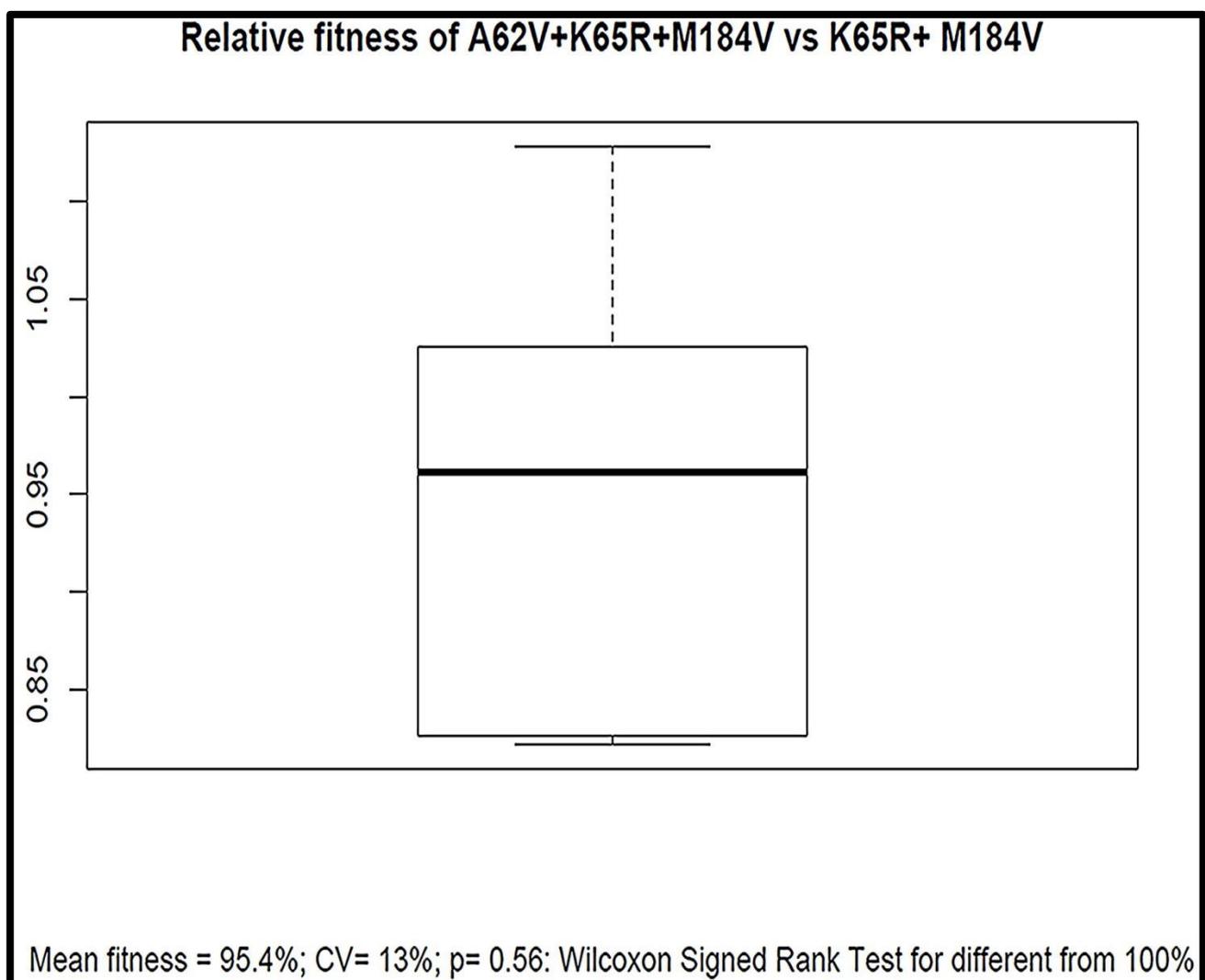
**Table K: qPCR quantification data for head to head Growth competition experiment for 2 technical replicates for each of the 3 biological replicates**

. Key: - blue = time 0 data; lime green = time 48 data; grey = time 72 data; yellow = Geometric mean of two technical replicates of each biological replicate at a given time

Sample	Time	Technical Replicate	Time_Assay	Ct	Starting Quantity (viral copies/ qPCR rxn)	Geo_Mean
G1_0	0	1	0_Total	18.18	1.91E+05	
G1_0	0	1	0_Total	18.50	1.56E+05	1.72E+05
G2_0	0	2	0_Total	17.79	2.45E+05	
G2_0	0	2	0_Total	17.86	2.35E+05	2.40E+05
G3_0	0	3	0_Total	18.85	1.24E+05	
G3_0	0	3	0_Total	19.11	1.05E+05	1.14E+05
G1_0	0	1	0_WT	19.79	4.57E+04	
G1_0	0	1	0_WT	20.64	2.56E+04	3.42E+04
G2_0	0	2	0_WT	19.26	6.54E+04	
G2_0	0	2	0_WT	19.53	5.44E+04	5.97E+04
G3_0	0	3	0_WT	19.35	6.15E+04	
G3_0	0	3	0_WT	20.36	3.09E+04	4.36E+04
G1_48	48	1	48_Total	16.52	5.53E+05	
G1_48	48	1	48_Total	16.82	4.57E+05	5.03E+05
G2_48	48	2	48_Total	16.19	6.86E+05	
G2_48	48	2	48_Total	16.36	6.14E+05	6.49E+05
G3_48	48	3	48_Total	17.12	3.76E+05	
G3_48	48	3	48_Total	17.20	3.58E+05	3.67E+05
G1_48	48	1	48_WT	18.01	1.54E+05	
G1_48	48	1	48_WT	18.48	1.12E+05	1.31E+05
G2_48	48	2	48_WT	18.10	1.45E+05	
G2_48	48	2	48_WT	18.37	1.21E+05	1.32E+05
G3_48	48	3	48_WT	17.95	1.61E+05	
G3_48	48	3	48_WT	18.04	1.50E+05	1.55E+05
G1_72	72	1	72_Total	17.03	4.01E+05	
G1_72	72	1	72_Total	17.46	3.04E+05	3.49E+05
G2_72	72	2	72_Total	16.77	4.73E+05	
G2_72	72	2	72_Total	17.14	3.73E+05	4.20E+05
G3_72	72	3	72_Total	17.92	2.26E+05	
G3_72	72	3	72_Total	17.93	2.25E+05	2.26E+05
G1_72	72	1	72_WT	18.37	1.21E+05	
G1_72	72	1	72_WT	18.48	1.12E+05	1.16E+05
G2_72	72	2	72_WT	18.72	9.47E+04	

Table J continued						
G2_72	72	2	72_WT	19.07	7.47E+04	8.41E+04
G3_72	72	3	72_WT	18.61	1.03E+05	
G3_72	72	3	72_WT	18.90	8.38E+04	9.27E+04

The geometric mean of two technical replicates obtained from each biological replicate in the growth competition experiment was used as variables to calculate the relative fitness of each virus using the formula described in Chapter 1 section 1.4.2. A relative fitness Box and whisker plot (**Figure 28**) derived from the statistical analysis of the two viruses using the Wilcoxon signed rank test showed that there is no significant difference in the relative fitness of *A62V+K65R+M184V* compared to *K65R+M184V* mutant viruses.



**Figure 28.** Relative fitness plot of *A62V+K65R+M184V* vs *K65R+M184V*

Results are discussed in the next chapter.

## Chapter 4 – Discussion

### 4.1 Summary and significance of major findings

The major finding of this study is that A62V mutation does not have a significant effect on viral fitness when it co-occurs with M184V and K65R. This is unexpected as previous work suggested that A62V restored fitness loss conferred by K65R in the presence of S68G by increasing the rate of incorporation of natural dNTPs during HIV-1 reverse transcription (Svarovskaia *et al.*, 2008). There are a few differences between the reported findings and our setting. 1) A62V and S68G occurred in patients treated with stavudine, abacavir or didanosine (Roge *et al.*, 2003); 2) In our setting A62V was observed to co-occur with K65R and M184V; 3) Previous reports were in predominantly HIV-1 subtype B populations, whereas our patient population is predominantly HIV-1 subtype C. M184V is a primary drug resistant mutation and like K65R causes reduced susceptibility of 3TC, FTC, ddi, ABC and diminished RT processivity and natural dNTP incorporation (Wainberg, 2004). M184V in combination with K65R reduces viral fitness more than any of the mutations on their own. It is possible that A62V may have had an effect on the fitness of K65R without M184V, which was negated by an interaction with M184V but this would not explain the co-occurrence of the three mutations. Other explanations are that A62V contributes to TDF resistance or that the fitness benefit is dependent on other accessory mutations being present.

In an earlier study (Maeda *et al.*, 1998), the A62V mutation was inferred to compensate for viral fitness when it occurred in concert with Q151M, F116Y, F77L and V75L mutations, however the study compared mutant viruses using *in vitro* kinetic assays that do not model the relative fitness accurately compared to growth competition experiments and may have incorrectly overestimated the extent to which A62V acts as a compensatory mutation. In comparison, the head to head growth competition experiment in this study between A62V+K65R+M184V and K65R+M184V allowed a direct comparison of the two viruses under similar conditions. The high sensitivity of a growth competition experiment coupled with the high precision for quantification of growth kinetics with real-time PCR; make it unlikely that the negative finding on A62V is due to analytic error. However there is a possibility that A62V fitness interaction might not have been sensitively detected as major drug resistance mutations like M184V have a fitness loss of only 6% (Wainberg, 2004) and compensation in fitness loss by A62V may not be sensitively detected by the qPCR assay employed in this study.

With reference to TDF resistance the Stanford University HIV drug resistance database (Stanford HIVDB - <http://hivdb.stanford.edu/>) resistance interpretation algorithm assigns a score of 5 to the A62V mutation when it occurs alone and an additional score of 10, with a net score of 15 when it co-occurs with K65R.

Although no peer-reviewed publication provides evidence of the effect of A62V on TDF resistance, the Stanford DB scoring suggests that it may have a synergistic effect on drug resistance when co-occurring with K65R. Phenotypic data on its effect on TDF resistance in HIV-1 subtype C would be informative.

In the presence of K65R, TDF has been reported to retain activity, probably due to the moderate effect of K65R on drug resistance and its effect on viral fitness impairs viral replication (Grant *et al.*, 2010). In many African settings therapy monitoring relies on clinical and or immunological monitoring and virologic failure could remain undiagnosed. In these patients secondary mutations following M184V and K65R, such as A62V could accumulate. If secondary mutations would contribute to higher levels of TDF resistance or fitness compensation, the result could lead to higher levels of viral replication and clinical or immunological deterioration. Such accumulation of secondary drug resistance mutations has been reported for patients receiving thymidine analogue- and protease inhibitor-containing regimens but the role of A62V in reducing the benefit of TDF therapy in patients with failure is unknown.

#### **4.2 Other important findings/ Technical lessons learned**

Other results obtained in this study indicate the technical challenges that should be considered when developing a growth competition assay for the study of viral fitness. The wild-type and mutant subclones obtained as a result of section 1 and 2 of the methodology of the study represent the essential components of a phenotypic assay. The logical workflow for the synthesis of the mutant clones was based on the initial use of traditional- cloning methods that allow synthesis of full-genome mutant constructs by subcloning. The technique of subcloning small fragments to eventually derived large plasmids has been widely used in some studies (Louder *et al.*, 2005; Rodriguez *et al.*, 2006). The advantage of using a subclone is that it provides a simple target template for efficient site-directed mutagenesis and it also provides the use of unique restriction enzyme sites that can be used to introduce the mutated sequence in a backbone vector construct harbouring the same restriction sites. The template size of an ideal subclone has to be relatively small to prevent non-specific binding of mutagenic primers (Carter, 1986). A small template also decreases the probability of the number of methylation sites that are known to direct mismatch repair systems in bacterial competent cells which will abrogate the intended mutation (Weiner *et al.*, 1994). These factors explain in part why in this study, SDM on the TV1-2kb (~5122bp) clone using the Quick Change XL SDM kit (Stratagene, USA) was only successful to produce the M184V subclone and failed to synthesise the other mutant constructs regardless of efforts to optimize the experiment. This suggests that perhaps the optimal size of a subclone for SDM using the Quick Change XL kit (Stratagene, USA) is below the size of the TV1-2kb subclone. However other factors such as primer design, fidelity and DNA processivity of the enzyme used during the SDM reaction could have also contributed to technical challenges experienced in trying to optimize the reaction.

Particularly, enzyme processivity was observed to significantly impact the outcome of the mutagenesis reaction. The observation was specifically experienced when synthesizing full-length recombinant constructs using the Q5™ mutagenesis kit (NEB, UK) in section 3 of the methodology. Section 3 of the methodology highlights another technical lesson learnt on the use of traditional based cloning methods using restriction enzymes compared to seamless advanced cloning methods such as the Infusion™ cloning (Clontech, USA). Infusion™ cloning minimized the technical challenges of cloning the large (1505bp) into the NL4.3 eGFP vector and compared to the previous approach using subclones, there was fast turnaround time for cloning the fragment (that is ~30 minutes compared to two days using traditional-based methods). More than 95% precision was observed in obtaining the recombinant clone and less time was spent on optimisation and troubleshooting.

The use of real-time PCR for the growth competition experiment between *A62V+K65R+M184V* and *K65R+ M184V* viruses also simplified the growth competition setup as it did not require different reporter inserts and allowed a high precision as evident from the reproducibility of the assay. (Dykes *et al.*, 2006; Wang *et al.*, 2010). The high specificity obtained for the allele-specific qPCR assay in this study (~13 Ct difference) and sensitivity (detection limit of 100 copies/ qPCR reaction) allows direct comparison of the findings in this study and for them to be comparable to other studies (Bergroth *et al.*, 2005; Svarovskaia *et al.*, 2008) employing similar approaches

Another important finding was on the TV1-FLG plasmid (which was the initial backbone vector to introduce different combinations of A62V, M184V, K65R) turned out to be non-infectious. The probable cause of non-infectivity in cell culture (as validated by a negative result obtained in *X-gal* staining of TZM-bl cells and no detectable p24 antigen content with the HIV p24 ELISA assay) is possibly due to mutations in the *gag/pol* genes (shown in the sequence results obtained). The mutations in the *gag/pol* region were defined as such because of base mismatches with the TV1 consensus sequence. The TV1 *env* consensus sequence which was shown to code a for functional *env* gene (Jacobs unpublished data) and partially proven in this study by the growth of the MJ4/TV1 hybrid plasmid clone in TZM-bl cells also provided supporting evidence that non-infectivity was probably caused by *gag/pol* mutations as the *env* was derived directly from MJ4/TV1 plasmid. Although mutations were also found in the TV1- FLG *env*, difficulty in sequencing this region with the primers that were available makes it difficult to conclude whether these mutations had a role in non – infectivity of the plasmid in cell culture. However, the possibility still remains that they could have had an effect as deleterious *env* mutations could hinder HIV-1 receptor and co-receptor binding rendering cells immune to HIV-1 infection (Louder *et al.*, 2005).

The mutations of TV1-FLG plasmid found in the protease and RT coding genes could have a deleterious effect, resulting in non-functional proteins being expressed which in turn result in defective virions with poor infectivity in cell culture and the explanation is also supported by literature (Kozisek *et al.*, 2012).

Had TV1 – FLG proven to be infectious, it would have been a suitable HIV-1 subtype C backbone vector and mutations could have been introduced via homologous recombination and ligation from an intermediate plasmid (designated ‘carrier plasmid’ in the study) or direct ligation of PCR amplified mutated fragments from the TV1-2kb mutant sub clones. However, the technique of directly ligating PCR amplified fragments generated with homologous restriction sites to the vector is only a convenient way to synthesize full-length genome constructs, only if suitable unique restriction enzyme sites are present in the amplicon and vector, otherwise a subclone would be required. There was also a strong motivation to have a subtype C infectious clone. Given that there are only a limited number of HIV-1C plasmid constructs that have been published and these include five that are naturally derived (Mochizuki *et al.*, 1999; Ndungu *et al.*, 2000; Grisson *et al.*, 2004; Dash *et al.*, 2008; Jacobs *et al.*, 2012) and a one that is synthetically made (Nauwelaers *et al.*, 2011). Naturally derived wild-type HIV-1C molecular clones have been reported to have poor growth in cell culture (Jacobs, 2013 personal communication) and may not be ideal representatives to study viral fitness interactions. Although there is a report of *in-vitro* synthesis of an HIV-1C infectious clone (Nauwelaers *et al.*, 2011) with comparable growth kinetics to HIV-1B and HIV-1C clones, synthetic chemical processes lack the means to introduce certain natural intrinsic properties associated with DNA such as methylation. DNA Methylation affects the epigenetic control of genes regulated by transcription factors *in vitro* or *in vivo* (Lim and Maher, 2010; Moore *et al.*, 2013). DNA methylation could have played a part in the non-infectivity of TV1-FLG in this study as half of the genome was derived from a synthetically non-methylated construct (TV1\_S-clone).

#### **4.3 Strengths and Limitations of study**

The key strength(s) of the study include significant advances in technical expertise and the understanding of the requirements needed to construct full-genome plasmid constructs for phenotypic drug resistance profiling. Such advances are reflected by achievements in troubleshooting cloning reactions; optimization for rapid ligation of large fragments; determining limiting factors for most commercially used SDM kits and the knowledge of more efficient cloning methods. The use of a highly reproducible allele-specific qPCR assay and the simple approach taken to set up a growth competition experiment proved to be a critical strength in this study assay, however, significant time was lost at the beginning of the study using traditional cloning methods and the need to synthesize a pure HIV-1C backbone vector to evaluate fitness interactions.

#### **4.4 Further investigations to consider for future studies.**

Our results did not explain the co-occurrence of M184V, K65R and A62V. Further investigations would be required to investigate the fitness effect of the individual mutations and all the possible mutation interactions: A62V vs. K65R; A62V vs. M184V; A62V vs wild-type viruses; A62V vs. A62V+K65R; A62V vs. A62V+M184V; A62V+K65R+M184V vs. K65R+M184V and A62V+K65R+M184V vs. wild-type in the presence of TDF. In addition, experiments to investigate the effect of A62V alone and in combination with K65R and/ or M184V on phenotypic drug resistance would assess whether an increased drug resistance conferred by A62V would be the basis for its selection in these patients.

As genomic interactions outside the gene of interest (e.g. protease inhibitor resistance mutations and *gag*) could influence the net effect of a mutation, an infectious HIV-1C clone would be valuable for future experiments. This would allow assessing these mutation interactions in a backbone that represents the current circulating HIV-1 subtype C strains in the southern African region. These further experiments fell outside the scope of the study (based on time limitations and funding).

## Chapter 5 - Conclusion

A62V *reverse transcriptase* mutation in HIV-1 has no significant fitness compensation effect when it co-occurs with M184V and K65R. However, further investigations are needed to characterise the mutation in the context of other accessory or secondary mutations commonly occurring in HIV-1 subtype C to further explore fitness interactions. The effect of A62V when occurring alone or in combination with K65R and /or M184V also requires investigation. The investigation would be greatly aided by an infectious naturally derived pure HIV-1C molecular clone. The use of a highly reproducible qPCR assay to characterise growth kinetics of different constructs used in a growth competition experiment would facilitate such work. As an increasing number of patients receive TDF-based regimens it would remain a priority to study the accumulation of mutations on this regimen and their effect on fitness and drug resistance.

**Addenda*****Addendum A1 – License details for Figure 1***

<b>Supplier</b>	<b>Elsevier Limited</b> <b>The Boulevard, Langford Lane</b> <b>Kidlington,Oxford,OX5 1GB,UK</b>
Registered Company Number	1982084
<b>Customer name</b>	<b>Duncan Njenda</b>
Customer address	Division of Medical Virology Cape Town
<b>License number</b>	<b>3623560890960</b>
License date	May 07, 2015
Licensed content publisher	Elsevier
Licensed content publication	Transfusion Science
Licensed content title	Molecular biology of human immunodeficiency virus Type-1
Licensed content author	Sadhna Joshi and Rajiv L. Joshi
Licensed content date	September 1996
Licensed content volume number	17
Licensed content issue number	3
Number of pages	28
Start Page	351
End Page	378
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Format	both print and electronic

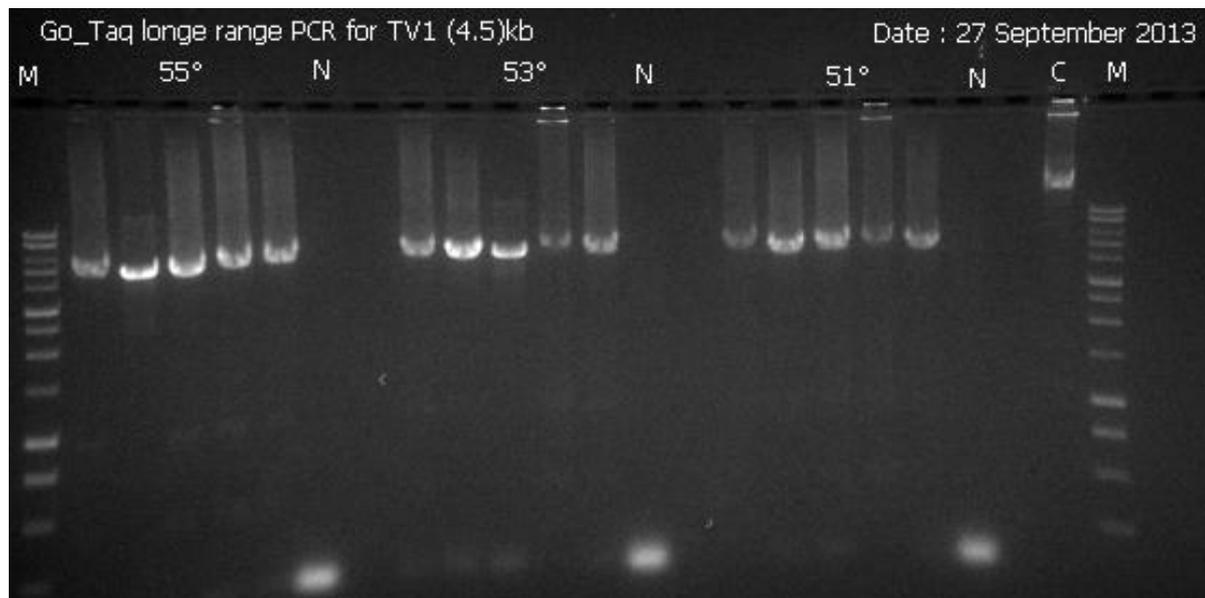
**Addendum A2 – License details for Figure 2**

<b>Supplier</b>	<b>Elsevier Limited</b> <b>The Boulevard, Langford Lane</b> <b>Kidlington,Oxford,OX5 1GB,UK</b>
Registered Company Number	1982084
<b>Customer name</b>	<b>Duncan Njenda</b>
Customer address	Division of Medical Virology Cape Town, None 7505
<b>License number</b>	<b>3737000502569</b>
License date	Oct 27, 2015
Licensed content publisher	Elsevier
Licensed content publication	Drug Resistance Updates
Licensed content title	Fitness of drug resistant HIV-1: methodology and clinical implications
Licensed content author	Miguel E Quiñones-Mateu and Eric J Arts
Licensed content date	December 2002
Licensed content volume number	5
Licensed content issue number	6
Number of pages	10
Start Page	224
End Page	233
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic

**Addendum B – Summary of major reagents, software and equipment used in the study**

<b>Material</b>	<b>Application/ Methods</b>	<b>Supplier</b>	<b>Local Distributer (South Africa)</b>
<b>Reagents</b>			
Lab consumables – PCR Eppendorf tubes, pipette tips, culture flasks	Various Laboratory standard techniques		WhitSci (Whitehead Scientific_Pty_Ltd)
Qiamp Viral RNA extraction Kit	Viral RNA extraction	Qiagen, Germany	WhitSci (Whitehead Scientific_Pty_Ltd)
Go-Taq Long-range PCR system	PCR experiments	Promega, USA	Anatech Instruments_Pty_Ltd
Restriction enzymes	Restriction enzyme digestion and cloning	New England Biolabs, USA	Inqaba Biotec
Oligonucleotides	qPCR, PCR, Infusion™ cloning and Mutagenesis	Inqaba Biotec	Inqaba Biotec
MiniElute reaction clean-up	PCR amplicon and gel extraction clean-up	Qiagen, Germany	WhitSci (Whitehead Scientific_Pty_Ltd)
QiaxII gel extraction kit	PCR amplicon and gel extraction clean-up	Qiagen, Germany	WhitSci (Whitehead Scientific_Pty_Ltd)
Infusion™ Cloning Kit	Cloning	Takara-Clontech, USA	Separations, South Africa
Q5™ Mutagenesis kit	Mutagenesis	New England Biolabs, USA	Inqaba Biotec
Maxiprep™ System kit	Preparation of plasmid maxiprep Stocks	Promega, USA	Anatech Instruments_Pty_Ltd
Miniprep™ System kit	Preparation of plasmid maxiprep Stocks	Promega, USA	Anatech Instruments_Pty_Ltd
1 Step Lo-Rox PCRbio real-time master mix	qPCR experiments	PCR Biosystems, USA	Biocom Biotec
Q5™ hot start PCR system	PCR experiments	New England Biolabs, USA	Inqaba Biotec
X-treme gene HD Transfection reagent	Transfection in cell culture assays	Roche diagnostics, Switzerland	Roche, South Africa
Expand Hi-fi PCR System	PCR experiments	Roche diagnostics, Switzerland	Roche, South Africa
Competent cells (NEB-5 alpha)	Cloning	New England Biolabs, USA	Inqaba Biotec
Competent cells (JM109)	Cloning	Promega, USA	Anatech Instruments_Pty_Ltd
Competent cells (Stable™ 2)	cloning	Invitrogen, USA	Life Technologies, South Africa
<b>Major Equipment</b>			
NanoDrop™ ND 1000	Spectrophotometric measurements of DNA or RNA	ThermoScientific, USA	
UV- ITEC Prochem Gel Dock System	Gel visualization and imaging	Cambridge, UK	
CFX96 Real- time PCR	qPCR experiments	Bio-Rad, USA	Bio-Rad, South Africa

System			
Applied Biosystems Veriti® PCR thermal cycler	PCR experiments	ThermoFisher, USA	Life Technologies, South Africa
TC20 Automated cell counter	Cell counting in cell culture assays	Bio-Rad, USA	Bio-Rad, South Africa
ELx800™ ELISA plate reader	Enzyme Linked Immunosorbent assays	Bio-Tek, USA	
ABI 3130xL Genetic Analyser	Sequencing	Applied Biosystems, CA, USA	
CO <sub>2</sub> Incubator	Cell culture assays	Nuaire, USA	
CO <sub>2</sub> Shaker incubator	Bacterial cultures	Labcon, USA	
<b>Software</b>			
SnapGene	Molecular analysis	GSL biotech, USA	
Sequencher v 5	Sequence analysis	Gene Codes Corporation, Ann Arbor, MI USA	
Ugene	Cloning analysis	Unipro, Novosibirsk, Russia	
Bio- Rad CFX Manager	qPCR analysis	Bio-Rad, USA	Bio-Rad, South Africa
R program v 3.2.2	Statistical analysis	R Development Core Team ( <a href="https://www.r-project.org/">https://www.r-project.org/</a> )	

**Addendum C1 –TV1 4.5kb proof**

*Key: - C – Control, N = No template control, M = 1kb molecular weight marker. 4.5kb PCR products shown for different temperatures (°C) used*

**Addendum C2 – MinElute Reaction Protocol (Qiagen, Germany)**

Procedure:

The protocol was followed by addition of 300µl of Buffer ERC to 20µl per enzymatic reaction volume followed by mixing. If the color of the mixture was orange or violet, 10µl of 3 M sodium acetate, pH 5.0 until the color of the mixture would turn to yellow. This was then followed by inverting the tube 4-5 times to mix the contents. A MinElute column was then placed in a 2 ml collection tube in a suitable rack and the sample was then added to the column before subsequent centrifuging\* for 1min. To obtain maximal recovery, all traces of sample were transferred to the spin column and the procedure repeated before the flow through was discarded. The MinElute column was then placed back into the same 1.5ml Eppendorf tube. Washing was done by addition of 750µl Buffer PE to the MinElute column and centrifuged for 1 min. The flow through was discarded and the tube was centrifuged again to remove any residual wash buffer. Afterwards the column was then placed in a new Eppendorf tube and 10µl elution buffer (preheated to 50°C)

- \*All steps for centrifugation were carried out at 13000rpm

***Addendum D1 –T4 DNA Ligation protocol***

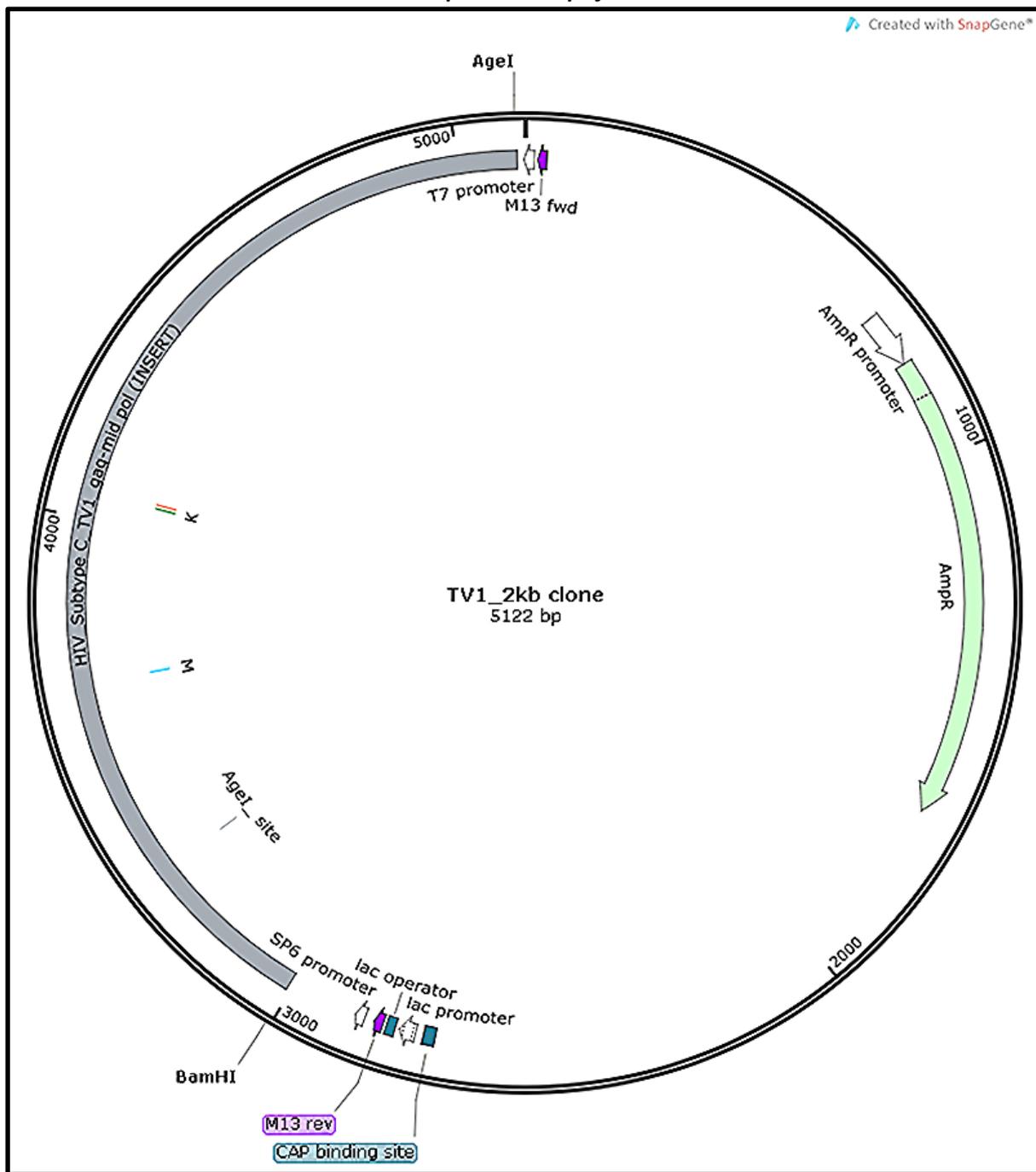
1. Reaction assembly in a sterile microcentrifuge tube:

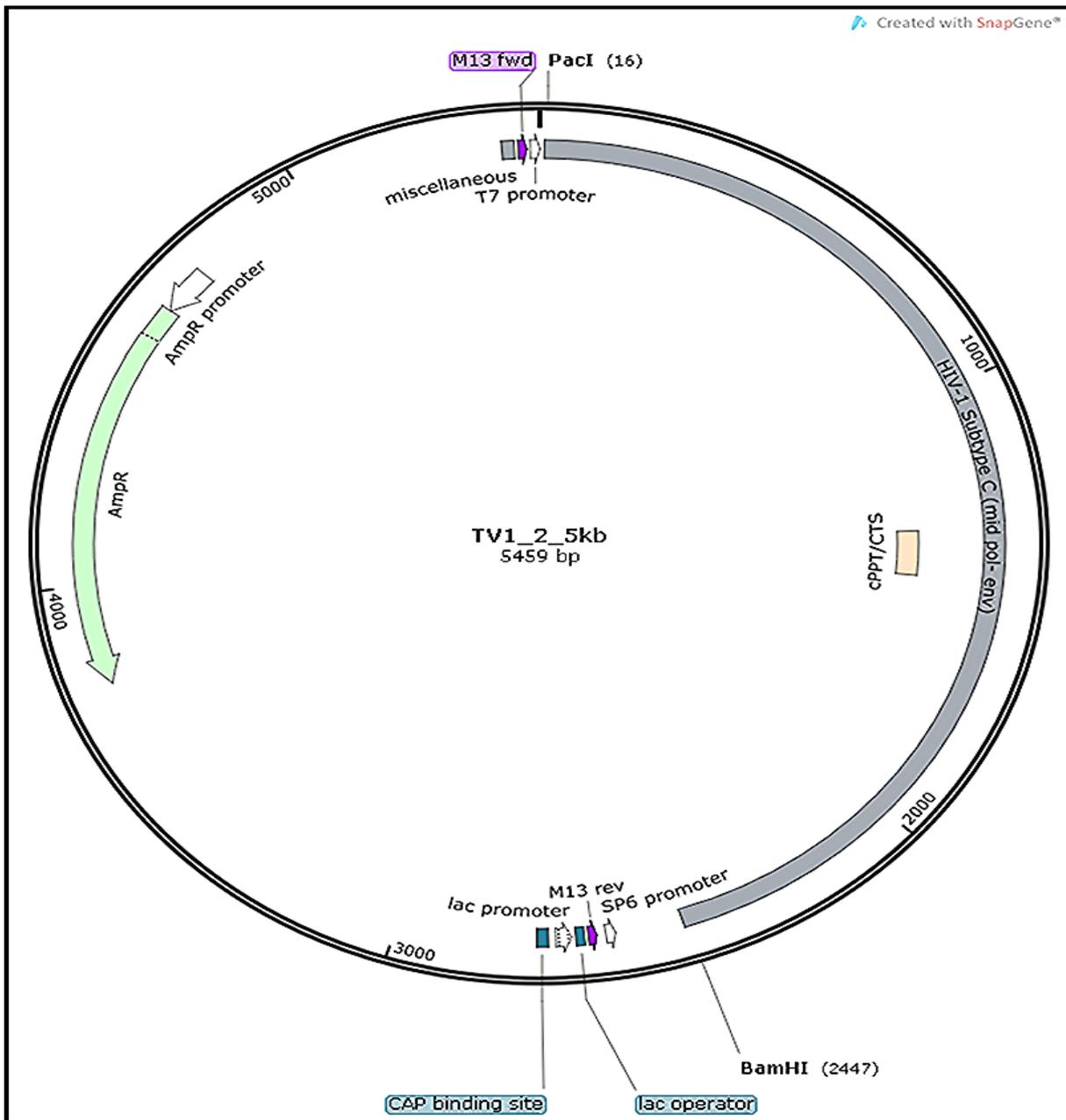
Vector DNA	100ng
Insert DNA	17ng
Ligase 10X Buffer	1µl
T4 DNA Ligase (Weiss units)	<u>0.1–1u</u>
Nuclease-Free Water to final volume of	<u>10µl</u>

2. Incubate time for reaction:

Room temperature for 3 hours, or 4°C overnight, or 15°C for 4–18 hours.

**Addendum D2 – plasmid maps for Tv1 subclones**





**Addendum E – Sequencing protocol****Big dye cycle sequencing kit v3.1 (Applied Biosystems) optimized protocol**

Sequencing reaction setup in a sterile 1.5ml Eppendorf tube		Control reaction	
DNA Sample*	10µl	DNA Sample*	1µl
X terminator	1µl	X terminator	1µl
Double distilled H <sub>2</sub> O	3µl	Double distilled H <sub>2</sub> O	4µl
5 x Buffer	4µl	5 x Buffer	3µl
Primer	2µl	Primer	1µl
<b>Total</b>	<b>20µl</b>	<b>Total</b>	<b>10µl</b>

\*DNA samples had varying concentrations but the final volume of 10µl was made to a minimum of 400ng for miniprep plasmid DNA

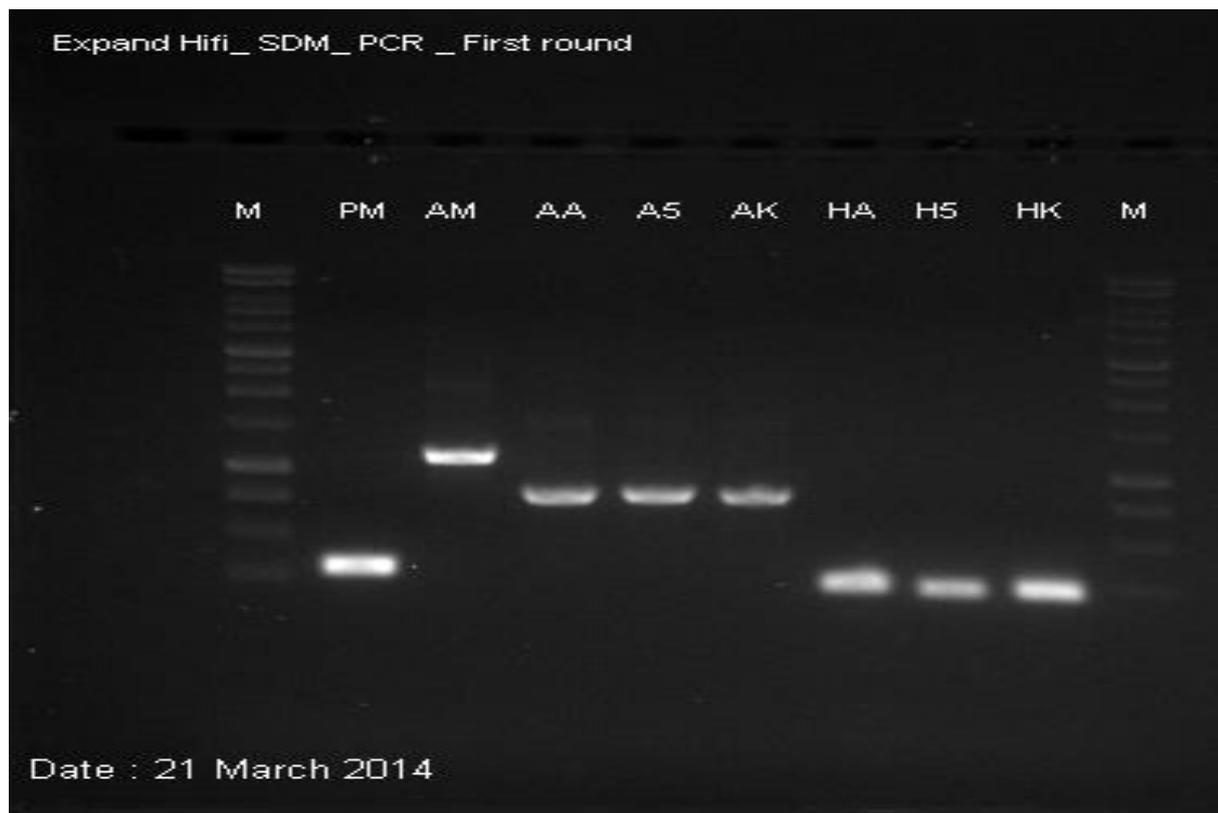
**Addendum F – Restriction enzymes digest protocol**

NEB Double digest Protocol (New England Biolabs, USA)

DNA Sample (1µg)	2µl
Nuclease free H <sub>2</sub> O	41µl
Enzyme 1 (10 units)	1µl
Enzyme 2 (10 units)	1µl
5 x Cutsmart buffer	5µl
<b>Total</b>	<b>50µl</b>

Incubation at 37°C for 1hr

**Addendum G1 - SDM - PCR round 1**

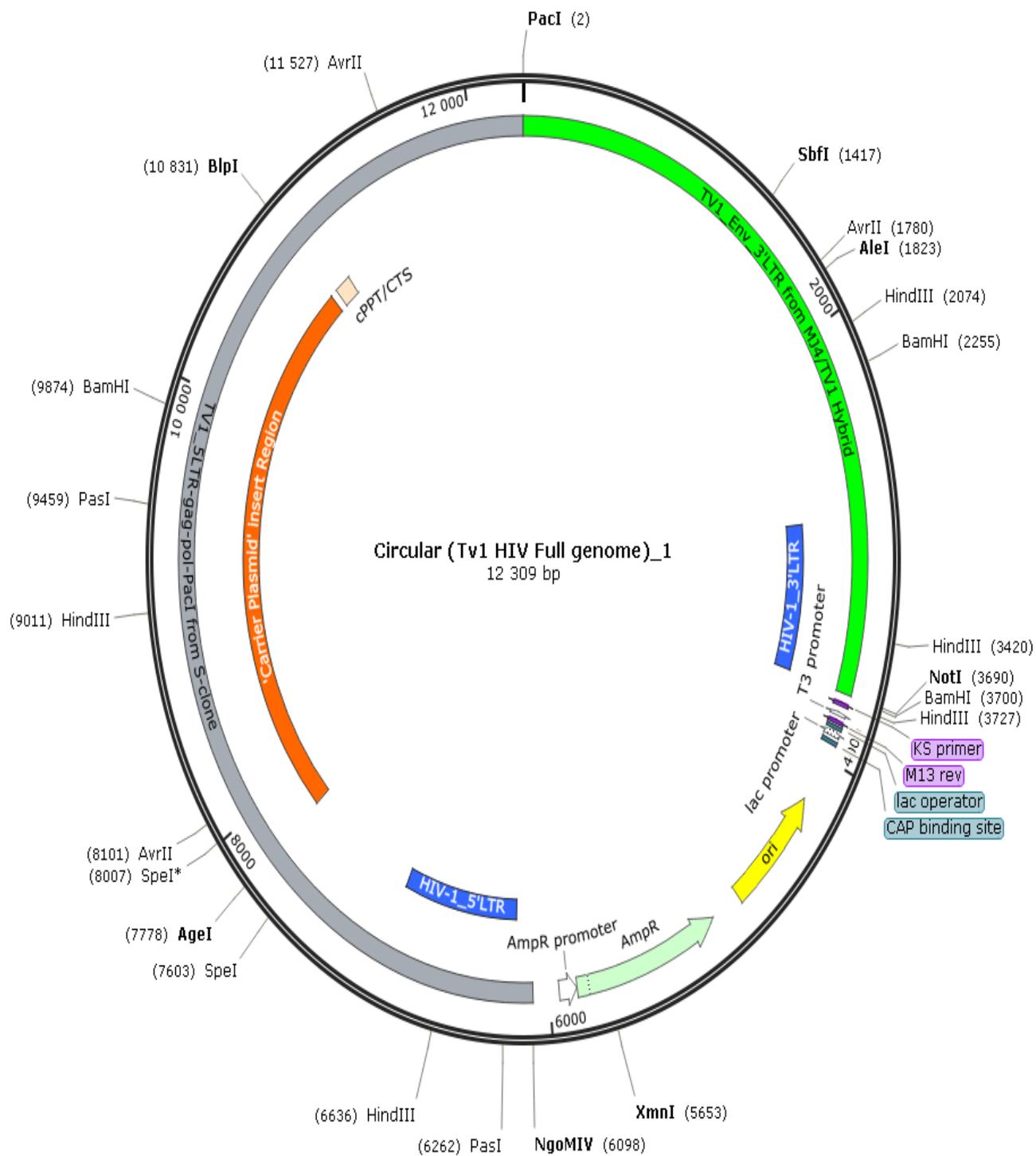


**Addendum G2 - SDM- PCR round 2**



**Addendum H – TV1\_FLG map**

Created with SnapGene®



**Addendum I \_Sequence quality proof**

Sequencher - [Mutagenesis1.SPF]

File Edit Select Assemble Contig Sequence View Window Help

Assembly Parameters AbN Assemble Automatically Assemble Interactively Assemble to Reference

Parameters: (Dirty Data, With ReAligner, 3' gap placement): Min Overlap = 20, Min Match = 85%

Name	Size	Quality	Kind	Label
A82V + K85R primer	40 BPs		DNA Fragment	-
A82V primer	35 BPs		DNA Fragment	-
A82V_sample 1	491 BPs	99.4%	AutoSeq Frag....	-
AKM_D01_2014-05-07	440 BPs	100%	AutoSeq Frag....	-
AKR_F03_2014-05-07	477 BPs	99.8%	AutoSeq Frag....	-
C_C02_2014-04-02	449 BPs	100%	AutoSeq Frag....	-
C_E02_2014-05-13	518 BPs	100%	AutoSeq Frag....	-
Double mutant A82V + K8...	528 BPs	100%	AutoSeq Frag....	-
Double Mutant K85R + M1...	807 BPs	100%	AutoSeq Frag....	-
K85R primer	38 BPs		DNA Fragment	-
K85R sample 1	544 BPs	99.6%	AutoSeq Frag....	-
K85R sample 2	413 BPs	98.1%	AutoSeq Frag....	-
K85R sample 3	573 BPs	100%	AutoSeq Frag....	-
K85R sample 4	576 BPs	99.7%	AutoSeq Frag....	-
KF_F02_2014-05-07	568 BPs	99.3%	AutoSeq Frag....	-
KMc_A01_2014-05-07	579 BPs	100%	AutoSeq Frag....	-
KMF_H02_2014-05-07	538 BPs	99.8%	AutoSeq Frag....	-
KMF_H02_2014-05-07	538 BPs	99.8%	AutoSeq Frag....	-
KMR_G03_2014-05-07	584 BPs	99.8%	AutoSeq Frag....	-
KMX_C04_2014-05-13	251 BPs	99.6%	AutoSeq Frag....	-
KM_B01_2014-05-07	532 BPs	99.8%	AutoSeq Frag....	-
KR_E03_2014-05-07	533 BPs	100%	AutoSeq Frag....	-
M184V primer	38 BPs		DNA Fragment	-
M184V Sample	821 BPs	100%	AutoSeq Frag....	-
M4_A04_2014-05-13	485 BPs	99.8%	AutoSeq Frag....	-
MR_B03_2014-05-07	441 BPs	99.3%	AutoSeq Frag....	-
Mx_E01_2014-05-07	587 BPs	99.5%	AutoSeq Frag....	-
M_F01_2014-05-07	582 BPs	99.6%	AutoSeq Frag....	-
Triple mutant ( A82V + K8...	519 BPs	99.8%	AutoSeq Frag....	-
Triple Mutant ( A82V+ K85...	515 BPs	99.6%	AutoSeq Frag....	-
Triple mutant (A82V + k85...	580 BPs	100%	AutoSeq Frag....	-
TR_H03_2014-05-07	405 BPs	98.8%	AutoSeq Frag....	-
Tv1_2kb_map_reference...	5120 BPs		Ref: DNA Fra...	-

## References

- Adetokunboh, O. & Oluwasanu, M., 2015.** Eliminating mother-to-child transmission of the human immunodeficiency virus in Sub-Saharan Africa: The journey so far and what remains to be done. *Journal of Infection and Public Health*.
- Aldrovandi, G.M., 1997.** Pathogenesis of HIV infection in children. *Complications of AIDS in Infants*, 7(1), pp.19–31.
- Arien, K.K., Gali, Y., El-Abdellati, A., Heyndrickx, L., Janssens, W. & Vanham, G., 2006.** Replicative fitness of CCR5-using and CXCR4-using human immunodeficiency virus type 1 biological clones. *Virology*, 347(1), pp. 65-74
- Arnott, A., Darren J., Kim W., Paul R. G., Kate M., Patricia G., Matthew G. L, Dax E. M., Kelleher A.D., Smith E.D., McPhee D. A. & Pulse Study Team, 2010.** High Viral Fitness during Acute HIV-1 Infection. *PloS one* 5, no. 9: 12.
- Arts, E.J. & Hazuda, D.J., 2012.** HIV-1 antiretroviral drug therapy. *Cold Spring Harbor perspectives in medicine*, 2(4), p.a007161.
- Autran, B., Carcelain G., Li T. S., Blanc C., Mathez D., Tubiana R., Katlama C., Debre P. & Leibowitch J., 1997.** Positive effects of combined antiretroviral therapy on CD4+ T cell homeostasis and function in advanced HIV disease. *Science* 277, no. 5322: 112-116.
- Bansode, V., Zuzanna J D., Simon A Travers, Emmanuel B., Molesworth A., Crampin A., Ngwira B., French N., Glynn J. R. & McCormack G.P., 2011.** Drug resistance mutations in drug-naive HIV type 1 subtype C-infected individuals from rural Malawi. *AIDS research and human retroviruses* 27, no. 4: 439-44.
- Bergroth, T., Sönnernborg, A. & Yun, Z., 2005.** Discrimination of lamivudine resistant minor HIV-1 variants by selective real-time PCR. *Journal of virological methods*, 127(1), p.100-7.
- Bowers, K; Pitcher, C. & Marsh, MI; (1997).** CD4: A co-receptor in the immune response and HIV infection. *The International Journal of Biochemistry & Cell Biology* 29, 871-875.
- Brockman, M., Giancarlo O. T., Bruce D. W., & Todd M. A., 2006.** “Use of a Novel GFP Reporter Cell Line to Examine Replication Capacity of CXCR4- and CCR5-Tropic HIV-1 by Flow Cytometry.” *Journal of Virological Methods* 131 (2): 134–42.
- Carter, P., 1986.** Site-directed mutagenesis. *The Biochemical journal*, 237(1), pp.1–7.

- Cases-González & Sandra F., 2007.** Mutational Patterns Associated with the 69 Insertion complex in Multi-Drug-Resistant HIV-1 Reverse Transcriptase that confer increased excision activity and high-Level resistance to zidovudine. *Journal of Molecular Biology*, 365(2), p.298-309
- Catalfamo, M., Le Saout, C., & Lane, H. C., 2012.** The role of cytokines in the pathogenesis and treatment of HIV infection. *Cytokine & Growth Factor Reviews* 23, 207-214.
- Chang, M. W., Oliveira, G., Yuan, J., Okulicz, J. F., Levy, S., & Torbett, B. E., 2013.** Rapid deep sequencing of patient-derived HIV with ion semiconductor technology. *Journal of Virological Methods* 189, 232-234.
- Chen, L. & Christopher L., 2006.** Distinguishing HIV-1 Drug Resistance, Accessory, and Viral Fitness Mutations Using Conditional Selection Pressure Analysis of Treated versus Untreated Patient Samples. *Biology Direct* 1 (January): 14.
- Chen, L.F., Hoy, J. & Lewin, S.R., 2007.** Ten years of highly active antiretroviral therapy for HIV infection. *The Medical journal of Australia*, 186(3), pp.146–151.
- Cortez, K.J. & Maldarelli, F., 2011.** Clinical management of HIV drug resistance. *Viruses*, 3(4), pp.347–78.
- Curr, K., Snehlata T., Johan L., Larder B.A. & Vinayaka R. P., 2006.** Influence of naturally occurring insertions in the fingers subdomain of human immunodeficiency virus type 1 reverse transcriptase on polymerase fidelity and mutation frequencies in vitro. *Journal of General Virology* 87, no. 2: 419-428
- Dapp, M.J., Heineman, R.H. & Mansky, L.M., 2012.** Interrelationship between HIV-1 fitness and mutation rate. *Journal of molecular biology*, 425(1), pp.41–53.
- Dapp, M.J., Patterson, S.E. & Mansky, L.M., 2013.** Back to the future: revisiting HIV-1 lethal mutagenesis. *Trends in Microbiology*, 21(2), p.56-62.
- Dash, P.K. et al., 2008.** Exceptional molecular and coreceptor-requirement properties of molecular clones isolated from an Human Immunodeficiency Virus Type-1 subtype C infection. *Retrovirology*, 5, p.25
- De Luca, A., 2006.** The impact of resistance on viral fitness and its clinical implications. In A. M. Geretti, ed. *Antiretroviral Resistance in Clinical Practice*. Mediscript
- Department of Health, Republic of South Africa, 2010.** The South African Antiretroviral treatment guidelines 2010b Available online at [http://www.sanac.org.za/documents/2010\\_Adult ART Guidelines.pdf](http://www.sanac.org.za/documents/2010_Adult_ART_Guidelines.pdf). [Accessed 16 July 2013].

- Derdeyn, C. A., J. M. Decker, J. N. Sfakianos, X. Wu, W. A. O'Brien, L. Ratner, J. C. Kappes, G. M. Shaw, and E. Hunter, 2000.** Sensitivity of human immunodeficiency virus type 1 to the fusion inhibitor T-20 is modulated by coreceptor specificity defined by the V3 loop of gp120. *Journal of virology* 74, no. 18: 8358-8367
- Deval, J., Boulbaba S., Joëlle B., Marie P. E., Guerreiro C., Sarfati S. & Bruno C., 2002.** The molecular mechanism of multidrug resistance by the Q151M human immunodeficiency virus type 1 reverse transcriptase and its suppression using  $\alpha$ -boranophosphate nucleotide analogues. *Journal of Biological Chemistry* 277, no. 44: 42097-42104
- Dykes, C., Jiong W., Xia J., Vicente P., Dong S.A., Tallo A., Yangxin H., Wu H., & Demeter L.M., 2006.** Evaluation of a multiple-cycle, recombinant virus, growth competition assay that uses flow cytometry to measure replication efficiency of human immunodeficiency virus type 1 in cell culture. *Journal of Clinical Microbiology* 44, no. 6: 1930-1943
- Egger, M., B. Hirschel, P. Francioli, P. Sudre, M. Wirz, M. Flepp, M. Rickenbach, R. Malinverni, P. Vernazza & M Battegay. 1997.** Impact of new antiretroviral combination therapies in HIV infected patients in Switzerland: prospective multicentre study. Swiss HIV Cohort Study. *BMJ (Clinical research ed.)* 315, no. 7117: 1194-9.
- Eisele, E. & Siliciano, R. F., 2012.** Redefining the Viral Reservoirs that Prevent HIV-1 Eradication. *Immunity* 37, 377-388
- Etiebet, M. A., James S., Nowak G.R., Man C., Chang H., Ajayi S., Elegba O., Ndembiet N., Abimikual A., Carr J. K., Eyzaguirre L.M & Blattner W.A., 2013.** Tenofovir-based regimens associated with less drug resistance in HIV-1-infected Nigerians failing first-line antiretroviral therapy. *AIDS (London, England)* 27, no. 4: 553-61.
- Feng, J. Y., Florence T. M., Nicolas A. M., Gilbert B. M., Laurence R., Katyna B., Boulbaba S. & Bruno C., 2006.** Virologic and enzymatic studies revealing the mechanism of K65R- and Q151M-associated HIV-1 drug resistance towards emtricitabine and lamivudine. *Nucleosides, nucleotides & nucleic acids* 25, no. 1: 89-107
- Fischl, M. A., D. D. Richman, M.H. Grieco, M. S. Gottlieb, P. A. Volberding, O. L. Laskin, J. M. Leedom, J. E. Groopman, D. Mildvan, & R. T. Schooley. 1987.** The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. *The New England Journal of Medicine* 317, no. 4: 185-91.

- Freguja, R., Ketty G., Marisa Z. & Anita D., 2012.** Cross-talk between virus and host innate immunity in pediatric HIV-1 infection and disease progression. *The new microbiologica* 35, no. 3: 249-57.
- Garforth, S.J., Tae W.K., Parniak M., Kool E. T. & Vinayaka R. P., 2007.** Site-Directed Mutagenesis in the Fingers Subdomain of HIV-1 Reverse Transcriptase reveals a specific role for the beta3-beta4 Hairpin Loop in dNTP Selection." *Journal of Molecular Biology* 365 (1): 38–49.
- Gelderblom, H. R., E. H. Hausmann, M.Ozel, G. Pauli, & M. A. Koch. 1987.** Fine structure of human immunodeficiency virus (HIV) and immunolocalization of structural proteins. *Virology* 156, no. 1: 171-6.
- Goodman, M.F., 2012.** An accidental biochemist. *DNA Repair*, 11(6), pp.527–536.
- Götte, M., 2012.** The distinct contributions of fitness and genetic barrier to the development of antiviral drug resistance. *Current Opinion in Virology* 2, no. 5: 644-650
- Grant, P.M. et al., 2010.** International cohort analysis of the antiviral activities of zidovudine and tenofovir in the presence of the K65R mutation in reverse transcriptase. *Antimicrobial agents and chemotherapy*, 54(4), pp.1520–5
- Grison, R. D., Agnès-Laurence C., Lan-Yu Y., Jun H., Charles W., Ganapati J. B., Weidong X., Chipepo K., and Ruth M. R., 2004.** Infectious molecular clone of a recently transmitted pediatric human immunodeficiency virus clade C isolate from Africa: evidence of intraclade recombination. *Journal of virology* 78, no. 24: 14066-14069
- Gulick, R. M., Mellors J.W., D. Havlir, J. J. Eron, C. Gonzalez, D. McMahon, D. D. Richman, Valentine F.T., Jonas L., Meibohm A., Emini E.A., & Chodakewitz J.A., 1997.** Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *The New England journal of medicine*, 337(11), p.734-739.
- Gupta, R. K., Michael R. J., Binta J. S., Andrew H., Daniel H. J. D., John G., Anthony W. S., Raph L. H. Nicaise N., Deenan P. & Silvia B., 2012.** Global trends in antiretroviral resistance in treatment-naive individuals with HIV after rollout of antiretroviral treatment in resource-limited settings: A global collaborative study and meta-regression analysis. *The Lancet* 380, no. 9849: 1250-1258
- Hamers, R. L., C. Kityo, J. M. A. Lange, T. F. Rinke & P. Mugenyi. 2012.** Global threat from drug resistant HIV in Sub-Saharan Africa. *Bmj* 344: e4159-e4159.

- Hemelaar, J., 2013.** Implications of HIV diversity for the HIV-1 pandemic. *The Journal of infection*, 66(5), pp.391–400.
- Ho S.N., Henry D. H., Robert M. H., Jeffrey K. P.& Larry R. P., 1989.** Site-Directed Mutagenesis by Overlap Extension Using the Polymerase Chain Reaction. *Gene* 77, no. 1: 51-59. 77: 51–59.
- Hoffmann, D., Albert D. G., P. R. Harrigan, Ian C. D. J., Tadashi N., J. G. García-Lerma & Walid H. 2011.** Measuring enzymatic HIV-1 susceptibility to two reverse transcriptase inhibitors as a rapid and simple approach to HIV-1 drug-resistance testing. *PLoS ONE* 6, no. 7
- Hosseinipour, M. C., Joep J. G. O., Ralf W., Sam P., Debbie K., Neil P., Susan A. F., Julie A. E. N., Joseph J E., & Johnstone K., 2009.** The public health approach to identify antiretroviral therapy failure: high-level nucleoside reverse transcriptase inhibitor resistance among Malawians failing first-line antiretroviral therapy. *AIDS (London, England)* 23, no. November 2008: 1127-1134
- Hu, Z. & Kuritzkes D.R.,2011.** Interaction of Reverse Transcriptase (RT) Mutations Conferring Resistance to Lamivudine and Etravirine: Effects on Fitness and RT Activity of Human Immunodeficiency Virus Type 1.” *Journal of Virology* 85 (21): 11309–14.
- Huang, K., Hsiang G., Dominique G., Jonathan M. C., Brockman M.A., Mishra S., Zabrina L. B., Stephen H., et al; 2011.** Progression to AIDS in South Africa is associated with both reverting and compensatory viral mutations. *PLoS ONE* 6, no. 4
- Imbeault, M., Robert L., Michel O. & Michel J. T., 2009.** Efficient magnetic bead-based separation of HIV-1-infected cells using an improved reporter virus system reveals that p53 up-regulation occurs exclusively in the virus-expressing cell population. *Virology* 393, no. 1: 160-167
- Ipp, H., Annalise E. Z., Glashoff R. H., Johan V. W., Naadira V., Reid T. & Linda G. B., 2013.** Serum adenosine deaminase and total immunoglobulin G correlate with markers of immune activation and inversely with CD4 counts in asymptomatic, treatment-naive HIV infection. *Journal of Clinical Immunology* 33, no. 3: 605-612
- Jacobs, G. B., Bock, S., Schuch, A., Moschall, R., Schrom, E.-M., Zahn, J., Reuter, C., Preiser, W., Rethwilm, A., Engelbrecht, S., et al., 2012.** Construction of a High Titer Infectious HIV-1 Subtype C Proviral Clone from South Africa. *Viruses*, 4(9), pp.1830–1843.
- Jiao, L., Li, H., Li, L., Zhuang, D., Liu, Y., Bao, Z., Liu, S., & Li, J; 2012.** Impact of novel resistance profiles in HIV-1 reverse transcriptase on phenotypic resistance to NVP. *AIDS Research and Treatment* 2012

**Johnson, V. A., Françoise B., Bonaventura C., Brian C., Richard T. D., Lisa M. D., Kuritzkes D. R., et al. 2003.** Drug resistance mutations in HIV-1. *Topics in HIV medicine: A publication of the International AIDS Society, USA*; 6: 215-21.

**Jordan, M. R., D. E. Bennett, M. A. Wainberg, D. Havlir, S. Hammer, C. Yang, L. Morris, et al. 2012.** Update on World Health Organization HIV drug resistance prevention and assessment strategy: 2004-2011. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* 54 Suppl 4:S245 -9.

**Joshi, S. & Joshi, R.L., 1996.** Molecular biology of human immunodeficiency virus type-1. *Transfusion science*, 17(3), pp.351–378.

**Kozísek, M., Sandra H., Klára G. S., G. B. Jacobs, Anita S., Bernd B., V. Müller, et al., 2012.** Mutations in HIV-1 Gag and Pol Compensate for the Loss of Viral Fitness Caused by a Highly Mutated Protease. *Antimicrobial Agents and Chemotherapy* 56 (8): 4320–30.

**Kuritzkes, D.R., 2004.** Preventing and managing antiretroviral drug resistance. *AIDS Patient Care and STDs*, 18(5), p.259-273.

**Landt, O., Grunert, H. & Hahn, U., 1990.** A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene*, 96(1), pp.125–128.

**Lim, D. H. K. and Maher, E. R., 2010.** DNA methylation: a form of epigenetic control of gene expression. *The Obstetrician & Gynaecologist*, 12: 37–42. doi: 10.1576/toag.12.1.037.27556

**Liu, L. & Lomonossoff, G., 2006.** A site-directed mutagenesis method utilising large double-stranded DNA templates for the simultaneous introduction of multiple changes and sequential multiple rounds of mutation: Application to the study of whole viral genomes. *Journal of virological methods*, 137(1), pp.63–71.

**Louder, M.K. et al., 2005.** HIV-1 envelope pseudotyped viral vectors and infectious molecular clones expressing the same envelope glycoprotein have a similar neutralization phenotype, but culture in peripheral blood mononuclear cells is associated with decreased neutralization sensi. *Virology*, 339(2), p.226-238.

**Machado, E. S., Adriana O. A., Dwight V. N., Philippe L., Silvia M. C., Ricardo H. O., & Marcelo A. S., 2009.** Emergency of primary NNRTI resistance mutations without antiretroviral selective pressure in a HAART-treated child. *PLoS ONE* 4, no. 3: 1-6.

- Maeda, Y., Venzon, D.J. & Mitsuya, H., 1998.** Altered drug sensitivity, fitness, and evolution of human immunodeficiency virus type 1 with pol gene mutations conferring multi-dideoxynucleoside resistance. *The Journal of infectious diseases*, 177(5), p.1207-13
- Malmsten, A., 2005.** Reverse Transcriptase Activity Assays for Retrovirus Quantitation and Characterization. *Acta Universitatis Upsaliensis, Sweden. Digital Comprehensive summaries of Uppsala Dissertations from the Faculty of Medicine 1.84pp. Uppsala. ISBN91-554-6127-1*
- Mansky, L., 1996.** The Mutation Rate of Human Immunodeficiency Virus Type 1 Is Influenced by the vpr Gene. *Virology*, 400, pp.391–400
- Mansky, L.M. & Temin, H.M., 1995.** Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *Journal of virology*, 69(8), p.5087-5094..
- Martinez-Cajas, J.L. & Wainberg, M. a, 2008.** Antiretroviral therapy: optimal sequencing of therapy to avoid resistance. *Drugs*, 68(1), p.43-7
- Martinez-Picado, J. & Martínez, M.A., 2008.** HIV-1 reverse transcriptase inhibitor resistance mutations and fitness: A view from the clinic and ex vivo. *Virus Research*, 134, p.104-123
- Matos, P.M., Castanho, M. a R.B. & Santos, N.C., 2010.** HIV-1 fusion inhibitor peptides enfuvirtide and T-1249 interact with erythrocyte and lymphocyte membranes. *PloS one*, 5(3), p.e9830.
- Miao, H., Carrie D., L. M. Demeter, J. Cavanaugh, S. Y. Park, A. S. Perelson & Hulin Wu. 2008.** Modeling and estimation of kinetic parameters and replicative fitness of HIV-1 from flow-cytometry-based growth competition experiments. *Bulletin of Mathematical Biology* 70, no. 6: 1749-1771.
- Ma, J., Carrie D., T. Wu, Y. Huang, L. Demeter, & H. Wu., 2010.** vFitness: a web-based computing tool for improving estimation of in vitro HIV-1 fitness experiments. *BMC bioinformatics* 11: 261
- Mochizuki, N., N. Otsuka, K. Matsuo, T. Shiino, A. Kojima, T Kurata., K Sakai, et al., 1999.** An infectious DNA clone of HIV type 1 subtype C. *AIDS research and human retroviruses* 15, no. 14: 1321-1324.
- Moore, L.D., Le, T. & Fan, G., 2013.** DNA methylation and its basic function. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, 38(1), p.23-38

- Naif, H. M. 2013.** Pathogenesis of HIV Infection. *Infectious Disease Reports* 5 (Suppl 1): e6.
- Navarini, A.A., Stoeckle, M., Navarini, S., Mossdorf, E., Jullu, B.S., Mchomvu, R., Mbata, M., Kibatala, P., Tanner, M., Hatz, C. & Schmid-Grendelmeier, P; 2010.** Antihistamines are superior to topical steroids in managing human immunodeficiency virus (HIV)-associated papular pruritic eruption. *International journal of dermatology*, 49(1), pp. 83-86.
- Nauwelaers D., Margriet H., Bart W., Kim S., Kurt B., Ellen C., Mimi Z., et al., 2011.** A synthetic HIV-1 subtype C backbone generates comparable PR and RT resistance profiles to a subtype B backbone in a recombinant virus assay. *PLoS ONE* 6, no. 5.
- Ndungu, T., B. Renjifo and M. Essex, 2001.** Construction and analysis of an infectious human Immunodeficiency virus type 1 subtype C molecular clone. *Journal of virology* 75, no. 11: 4964-4972
- Neogi, U., A. Shet, R Shamsundar & M L Ekstrand. 2011.** Selection of nonnucleoside reverse transcriptase inhibitor-associated mutations in HIV-1 subtype C: Evidence of etravirine cross-resistance. *AIDS* 25: 1123-1125
- Paillart, J C., L Berthoux, M. Ottmann. J. L Darlix, R. Marquet, B. Ehresmann & C Ehresmann, 1996.** A dual role of the putative RNA dimerization initiation site of human immunodeficiency virus type 1 in genomic RNA packaging and proviral DNA synthesis. *Journal of virology* 70, no. 12: 8348-8354.
- Pattanapanyasat, K & M. R. Thakar, 2005.** CD4+ T cell count as a tool to monitor HIV progression & anti-retroviral therapy. *Indian J.Med.Res.* 121, no. 4: 539-549.
- Pennings, P.S., 2013.** HIV drug resistance: Problems and perspectives. *Infectious Disease Reports*, 5(SUPPL.1), pp.21–25.
- Perales, C., Iranzo, J., Manrubia, S. C., & Domingo, E. (2012).** The impact of quasispecies dynamics on the use of therapeutics. *Trends in Microbiology* 20, 595-603
- Public Health Service, USA, 1987, 20<sup>th</sup> March.** – Available online from [http://www.avert.org/history-aids-1987-1992.htm#footnote11\\_2hs5coc](http://www.avert.org/history-aids-1987-1992.htm#footnote11_2hs5coc) [Accessed online on 1st December 2014]
- Quiñones-Mateu, M. E. & Arts, E. J., 2002.** Fitness of drug resistant HIV-1: Methodology and clinical implications. *Drug Resistance Updates* 5, 224-233
- Quiñones-Mateu, M. E. & E. J. Arts., 2006.** Virus fitness: concept, quantification, and application to HIV population dynamics. *Current topics in microbiology and immunology* 299: 83-140.

- Ramirez, B. C., Etienne S., Roman G. & Matteo N., 2008.** Implications of recombination for HIV diversity. *Virus Research* 134, no. 1-2: 64-73
- Redig J., 2013** How to fix adherent cells for microscopy and imaging- Available online from <http://bitesizebio.com/13460/how-to-fix-adherent-cells-for-microscopy-and-imaging> [Accessed online 1 July 2013]
- Rezende, L. F., K. Curr, T. Ueno, H. Mitsuya & V. R. Prasad. 1998.** The impact of multidideoxynucleoside resistance-conferring mutations in human immunodeficiency virus type 1 reverse transcriptase on polymerase fidelity and error specificity. *Journal of virology* 72, no. 4: 2890-2895
- Richman, D.D., 2006.** Antiviral drug resistance. *Antiviral research*, 71(2-3), pp.117–21
- Renjifo, B., Peter G., Beth C., Gernard M., Davis M., Wafaie F & Max E., 2004.** Preferential in-Utero Transmission of HIV-1 Subtype C as Compared to HIV-1 Subtype A or D.” *Aids* 18 (12): 1629–36.
- Reuter, M. A., Carolina P., & Michael R. B., 2012.** Cytokine production and dysregulation in HIV pathogenesis: Lessons for development of therapeutics and vaccines. *Cytokine and Growth Factor Reviews* 23, no. 4-5: 181-191.
- Rodriguez, M. A. et al., 2006.** Construction and characterization of an infectious molecular clone of HIV-1 subtype A of Indian origin. *Virology*, 345(2), p.328-36.
- Roge B. T., et al. 2003.** K65R with or without S68: a new resistance profile in vivo detected in most patients failing abacavir, didanosine, and stavudine. *Antivir. Ther.* 8:173–182
- Rosenbloom, D. I. S., Alison L H., S. A. Rabi, R. F. Siliciano & Martin A. N., 2012.** Antiretroviral dynamics determines HIV evolution and predicts therapy outcome. *Nature Medicine* 18 (9). Nature Publishing Group: 1378–85.
- Saini, S., P. Bhalla, H. Gautam, U. K. Baveja, S. T. Pasha, & R. Dewan. 2012.** Resistance-Associated Mutations in HIV-1 among Patients Failing First-Line Antiretroviral Therapy. *Journal of the International Association of Physicians in AIDS Care (JIAPAC)*.
- Scarth, B., McCormick, S. & Götte, M., 2011.** Effects of mutations F61A and A62V in the fingers subdomain of HIV-1 reverse transcriptase on the translocational equilibrium. *Journal of molecular biology*, 405(2), pp.349–60.

- Scherrer, A. U., Jürg B., Sabine Y., Thomas K., Vincent A., Hansjakob F., Alexandra C., 2012.** Long-lasting protection of activity of nucleoside reverse transcriptase inhibitors and protease inhibitors (PI) by boosted PI containing regimens. *PloS One* 7 (11): e50307.
- Schröder, A. R. W., Paul S., Huaming C., Charles B., Joseph R. E. & F. Bushman. 2002.** HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* 110, no. 4: 521-529.
- Seyfang A. & Jean H. J., 2004.** Multiple Site-Directed Mutagenesis of more than 10 sites simultaneously and in a single Round. *Analytical Biochemistry* 324 (2): 285–91.
- Shafer, R.W. & Schapiro, J.M., 2008.** HIV-1 drug resistance mutations: An updated framework for the second decade of HAART. *AIDS Reviews*, 10(2), p.67-84
- Shan, L., Kai D., Neeta S. S., Christine M. D., S. A. Rabi, Hung C. Y., H. Zhang, Joseph B. M., Joel N. B. & R. F. Siliciano. 2012.** Stimulation of HIV-1-Specific Cytolytic T Lymphocytes facilitates elimination of Latent Viral Reservoir after virus reactivation. *Immunity* 36, no. 3: 491-501
- Shen, T.J., Zhu, L.Q. & Sun, X., 1991.** A marker-coupled method for site-directed mutagenesis. *Gene*, 103(1), p.73-77.
- Sierra, S., Kupfer, B. & Kaiser, R., 2005.** Basics of the virology of HIV-1 and its replication. *Journal of Clinical Virology*, 34(4), p.233-244
- Shirasaka, T., M. F. Kavlick, T. Ueno, W. Y. Gao, E. Kojima, M. L. Alcaide, S. Chokekijchai, B. M. Roy, E. Arnold, & R Yarchoan; 1995.** Emergence of human immunodeficiency virus type 1 variants with resistance to multiple dideoxynucleosides in patients receiving therapy with dideoxynucleosides. *Proceedings of the National Academy of Sciences of the United States of America* 92, no. 6: 2398-
- Simon, V., David D. H. & Q. A. Karim., 2006.** Seminar HIV / AIDS epidemiology, pathogenesis, prevention, and treatment *Lancet* 2006 368: 489-504
- Stadeli, K.M. & Richman, D.D., 2013.** Rates of emergence of HIV drug resistance in resource-limited settings: a systematic review. *Antiviral therapy*, 18(1), p.115-23
- Staroscik A., 2004.** Calculator for determining the number of copies of template. *URI genomics and Sequencing center* – Available online from <http://cels.uri.edu/gsc/cndna.html> [Accessed on 1 January 2014]
- Stott, K., Janet M. & T. De Oliveira; 2012.** Principles of HIV drug resistance. *HIV Nursing Matters* 1: 46-51

- Sunpath, H., Baohua W., Michelle G., Jane H., Brent J., Mahomed-Yunus S. M., Claudia O., D. R. Kuritzkes, and Vincent C M., 2012.** High rate of K65R for ART naive patients with subtype C HIV infection failing a TDF-containing first-line regimen in South Africa. *AIDS Publish Ah*.
- Svarovskaia, E. S., J. Y. Feng, Nicolas A. M., Florence M., Derrick G., John K. L., Kirsten L. W., et al. 2008.** The A62V and S68G mutations in HIV-1 reverse transcriptase partially restore the replication defect associated with the K65R mutation. *Journal of acquired immune deficiency syndromes (1999)* 48, no. 4: 428-36.
- Tang, M.W. & Shafer, R.W., 2012.** HIV-1 antiretroviral resistance: scientific principles and clinical applications. *Drugs*, 72(9), pp.e1–25.
- Tebas, P., Keith H., Kristin M., Steven D., Herman V., Cal C. & William G P., 2002.** Effect of Prolonged Discontinuation of Successful Antiretroviral Therapy on CD4+ T Cell Decline in Human Immunodeficiency Virus–Infected Patients: Implications for Intermittent Therapeutic Strategies.” *Journal of Infectious Diseases* 186 (6): 851–54.
- Ueno, T., T. Shirasaka, & H. Mitsuya. 1995.** Enzymatic Characterization of Human Immunodeficiency Virus Type 1 Reverse Transcriptase Resistant to Multiple 2',3'-Dideoxynucleoside 5'-Triphosphates. *Journal of Biological Chemistry* 270 (40): 23605–11.
- Van der Kuyl., Antoinette C., Karolina K., Kevin K A., Youssef G., Victoria R B., Stefan J. D., Fokla Z., Guido V., Ben B. & Marion C., 2010.** Analysis of Infectious Virus Clones from Two HIV-1 Superinfection Cases Suggests that the Primary Strains Have Lower Fitness. *Retrovirology* 7 (January): 60
- Van Zyl, G.U., Tommy F. L., Mathilda C., Susan E., T. de Oliveira; W. Preiser; Natasha T W., S. Travers & R. W. Shafer, 2013.** Trends in Genotypic HIV-1 Antiretroviral Resistance between 2006 and 2012 in South African Patients Receiving First- and Second-Line Antiretroviral Treatment Regimens.” *PloS One* 8 (6): e67188.
- Volberding, P. A. & Deeks, S.G., 2010.** Antiretroviral therapy and management of HIV infection. *The Lancet*, 376(9734), p.49-62.
- Wagner, B.G., Garcia-Lerma, J.G. & Blower, S., 2012.** Factors limiting the transmission of HIV mutations conferring drug resistance: fitness costs and genetic bottlenecks. *Scientific Reports*, 2, p.1-10
- Wainburg, M.A., 2004.** The impact of the M184V substitution on drug resistance and viral fitness. *Expert Review of Anti-infective Therapy*, 2(1), p.147-151.

**Wang, J., Robert A. B., L. M. Demeter & C. Dykes; 2010.** Reduced fitness in cell culture of HIV-1 with nonnucleoside reverse transcriptase inhibitor-resistant mutations correlates with relative levels of reverse transcriptase content and RNase H activity in virions. *Journal of virology* 84, no. 18: 9377-9389

**Wang, H. et al., 2011.** An efficient approach for site-directed mutagenesis using central overlapping primers. *Analytical biochemistry*, 418(2), pp.304–6.

**Weiner, M. P., Gina L. C., Warren S., Janice C., Eric M. & John C. B., 1994.** Site-Directed Mutagenesis of double-stranded chain reaction DNA by the polymerase 151: 119–23.

**Wu, D., Xuewu G., Jun L., Xi Sun., Feng L., Yefu C. & D. Xiao, 2013.** A Rapid and Efficient One-Step Site-directed deletion, insertion, and substitution mutagenesis Protocol. *Analytical Biochemistry* 434 (2): 254–58.

**Wu, Y., 2004.** HIV-1 gene expression: lessons from provirus and non-integrated DNA. *Retrovirology*, 1, p.13.

**Xu, H.T., Eugene L. A., Maureen O., Peter K. Q., Yudong Q., B. G. Brenner and M. A. Wainberg, 2011.** Compensation by the E138K mutation in HIV-1 reverse transcriptase for deficits in viral replication capacity and enzyme processivity associated with the M184I/V mutations. *Journal of virology*, 85(21), pp.11300–8.

**Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L. & Charneau, P., 2000.** HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell* 101, 173-185.