Genetic and Functional Diversity of Central Nervous System (CNS) Derived Human Immunodeficiency Virus Type 1 (HIV-1) *tat* from Tuberculous Meningitis (TBM) Patients

Submitted by : Jenine Ramruthan

Supervised by : Dr Paradise Madlala

In fulfilment of the requirements for the degree: Master of Medical Science (Virology)

School of Laboratory Medicine and Medical Science

University of KwaZulu-Natal

2018

PREFACE

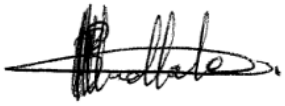
The experimental work described in this thesis was conducted at the HIV Pathogenesis Programme Hasso Platner Research Laboratory, Doris Duke Medical Research Institute, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa, from February 2017 to November 2018, under the supervision of Dr Paradise Madlala.

This work has not been submitted in any form for any degree or diploma to any tertiary institution, where use has been made of the work of others, it is duly acknowledged in the text.

J. Ramruthan Date 10 December 2018

Dr Paradise Madlala Date 10 December 2018

As the candidates supervisor I agree to the submission of this thesis:

Dr Paradise Madlala  Date 10 December 2018

DECLARATION

I Jenine Ramruthan declare that:

1. The research reported in this dissertation, except where otherwise indicated, is my original work.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This data does not contain other person's data, graphs, pictures, or other information, unless specifically acknowledged as being sourced from other persons.

4. This dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

    a. Their words have been re-written but the general information attributed to them has been referenced.

    b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.

    c. This dissertation does not contain text, graphics or tables copied and pasted from the internet unless specifically acknowledged, and the source being detailed in the thesis and reference section.

Signed  Date 10 December 2018

PUBLICATIONS AND PRESENTATIONS

Conference Presentations

Ramruthan, J., Patel, V., Ndung'u, T and Madlala, P. Genetic and Functional Diversity of Central Nervous System (CNS) Derived Human Immunodeficiency Virus Type 1 (HIV-1) Tat from Tuberculous Meningitis (TBM) Positive Patients. College of Health Sciences' Research Symposium, 5 – 6 October 2017, Durban, South Africa.

Ramruthan, J., Ndung'u, T and Madlala, P. Effect of CNS derived Tat diversity on HIV-1 disease in TB Meningitis Patients. Frontiers of Retrovirology, 11 – 14 September 2018, Leuven, Belgium.

This thesis is dedicated to my mother, without your unfailing love and sacrifice I would not be where I am today. I love you and appreciate everything that you have done for me.

“You are the wind beneath my wings”

ACKNOWLEDGMENTS

I would like to thank the following people and organisations who have supported me throughout my Masters degree:

- Dr. Madlala for his supervision, motivation, and attention to detail.
- Prof. Ndung'u his unwavering support and insight throughout the project.
- Dr. Mann for her kindness and willingness to help me throughout my project
- Dr. Gordon for all her advice and encouragement.
- Staff and students at the HIV Pathogenesis Programme, Hasso-Plattner Research Laboratory in particular: Keshni Hiramen, Nothemba Nontala, Doty Achieng, Katlego Sojane and Msizi Mbele
- Dr. Patel and Inkosi Albert Luthuli Central Hospital for providing access to the Tuberculous meningitis study.
- The National Research Foundation (Thuthuka Grant: 107322), The Poliomyelitis Research Foundation and The College of Health Sciences for financial assistance.

ABSTRACT

INTRODUCTION

Human immunodeficiency virus type 1 (HIV-1) transactivator of transcription (*tat*) is a regulatory gene that encodes the transactivator of transcription Tat protein. The Tat effectively increases the activity of the HIV-1 5' long terminal repeat (5' LTR) viral promoter to transcribe viral genes. The *tat* gene has two exons; the first 72 amino acids of Tat are encoded by the first exon, whilst amino acids 73 – 101 are encoded by the second exon. Exon 1 of Tat is sufficient for the transactivation of the 5' LTR and therefore was the focus of this study. The Tat encoded by exon 1 consists of 5 functional domains these include: the acidic domain (domain I) comprising amino acids 1 – 21, this is a proline rich domain with high sequence variation; the cysteine-rich domain (domain II) comprising amino acids 22–37, is composed of 6 well conserved cysteine residues in subtype C Tat proteins, a mutation at any of the 6 cysteine residue results in loss of Tat activity; the core domain (domain III) comprising amino acids 38–48, is made of a hydrophobic motif and is relatively well conserved. Together, the first 48 amino acids of Tat comprising domains I – III, allow for the transactivation activity of Tat responsible for enhancing viral gene transcription. The basic domain (domain IV) is an RNA-binding domain made up of amino acids 49 – 57 which allows for the binding ability of Tat to the TAR loop structure of the 5' LTR. Lastly the glutamine-rich domain (domain V) comprised of amino acids 58 – 72, also concentrated with basic amino acids, has the highest sequence variation in Tat. During the early stages of infection, HIV-1 enters the central nervous system (CNS) and replicates at marginal levels compared to high viral replication in the periphery. Yet, there is higher HIV-1 RNA levels in the in the cerebrospinal fluid (CSF) compared to plasma of tuberculosis meningitis (TBM) co-infected patients. However, the mechanisms driving the higher viral replication in the CNS of TBM patients are not well understood. Therefore, the major aim of this study is to characterise genetic and functional diversity of CNS and plasma derived Tat from TBM coinfected patients. We hypothesized that TBM coinfected patients will display genetically distinct HIV-1 *tat* variants in the CSF as a driver or consequence of higher viral replication in this compartment compared to plasma.

METHODS

Viral RNA was extracted from matched CSF and plasma samples obtained from 20 HIV-1 chronically infected patients (17 TBM and 3 non-TBM) using the QIAmp viral RNA Mini kit (Qiagen Inc., Valencia, CA, USA). Extracted viral RNA was reverse transcribed into viral DNA using SuperScript IV Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) and amplified using two rounds of (nested) PCR with the Platinum ® Taq DNA Polymerase High Fidelity PCR kit (Thermo Fisher Scientific, Boston, MA, USA). Genetic diversity of plasma and CSF derived isolates was assessed in 19 patients (16 TBM and 3 non-TBM) by sequencing, neighbour-joining phylogenetic analysis and both interpatient and intrapatient diversity analysis. The Tat sequences with previously reported mutations that affect Tat function were selected for downstream functional assays. Twelve *tat* PCR amplicons were cloned into a pTargeT™ expression plasmid (Promega Corporation, Madison, WI). Recombinant pTargeT clones containing patient derived HIV-1 *tat* was propagated using the QIAfilter Plasmid Maxi Kit (Qiagen Inc., Valencia, CA, USA) to transfect the TZM-bl mammalian cells, which contains the luciferase gene *luc* under the control of the LTR promoter. A luciferase assay was done to measure the relative luminescence for each Tat mutant and this was correlated to markers of disease progression such as viral load.

RESULTS

The phylogenetic data from our study show that sequences from plasma and CSF derived HIV-1 *tat* clustered closely per patient. Genetic variation was seen as varying branch lengths between patient clusters. However, our data do not show significant nucleotide differences between the plasma and CSF *tat* sequences with a p-distance of 0.059 and 0.062 respectively (p = ns). Additionally, our data revealed that the amino acid sequences were the same between the CSF and plasma compartments, except in 5% of patients that showed differences in positions that were not previously reported to affect Tat activity. However, Tat diversity was observed to occur in all 5 domains of the first 72 amino acids of Tat namely: V4I, P21A, K24S, H29R, S31C, S46Y, R52W, S57R, P59S and D64G.

The functional data from our study revealed that most patient derived Tat mutations occurred in combination with other previously reported mutations. Interestingly, Tat mutations that occurred together with P21A in five different patients showed a showed strong positive correlation with CSF viral load in the CNS ($p = 0.003$; $r = 0.98$).

CONCLUSION

We reject our hypothesis that CNS specific Tat mutations were responsible for the high viral load in the CNS of patients who have TBM, as the allele frequencies of reported amino acid substitutions were represented in equal proportions within plasma and CSF derived Tat variants. Furthermore, our functional data shows that majority of all Tat variants from the TBM group had a reduced capacity to transactivate the 5' LTR. Whilst we cannot confirm that Tat is responsible for the higher viral replication seen in the CNS of TBM coinfected patients, our data demonstrate that all Tat variants with a P21A mutation significantly correlates to viral replication in the CNS. Future studies should perform site directed mutagenesis to determine the exact mutations that mediate LTR activity.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABRREVIATIONS

%	Percentage
µg	Microgram
µl	Microliter
µm	Micromolar
aa	Amino Acids
AIDS	Acquired Immunodeficiency Syndrome
APD	Average Pairwise Distance
APOBEC 3G	Apolipoprotein B mRNA-Editing Catalytic Polypeptide Like-3G
ART	Antiretroviral Therapy
BBB	Blood–brain barrier
Bp	Base Pairs
BRE	B Recognition Element
CCR5	C-C chemokine receptor type 5
CDC	Centers For Disease Control And Prevention
CDK2	Cyclin-Dependent Kinase 2
CDK9	Cyclin-Dependent Kinase 9
Cells/µl	Cells Per Microlitre
Cells/mm^3	Cells Per Cubic Millimeter
CNS	Central Nervous System
Copies/ml	Copies Per Milliliter
CRF	Circulating Recombinant Form
CSF	Cerebrospinal Fluid
CT	Computerized Tomography
CTD	C-Terminal Domain
CXCR4	C-X-C chemokine receptor type 4
Cyc T1	Cyclin T1
DMEM	Dulbecco Modified Eagle Medium
DNA	Deoxyribonucleic Acid
DPE	Downstream Core Promoter Element

DPEC	Diethyl Pyrocarbonate
DSIF	DRB Sensitivity Inducing Factor
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
E.g.	For example
ELISA	Enzyme-Linked Immunosorbent Assay
EMSA	Electrophoretic Mobility Shift Assay
Env	Envelope Glycoprotein
ESCRT	The Endosomal Sorting Complexes Required For Transport Machinery
FTA-ABS	Fluorescent Treponemal Antibody Absorption Test
Gag	Group Specific Antigen
GM	Growth Media
GRID	Gay-Related Immunodeficiency
HAD	HIV-Associated Dementia
HAND	HIV Associated Neurocognitive Disorder
HAT	Histone Acetyltransferases
HDAC	Histone Deacetylases
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hidi-Formamide	Highly Deionized Formamide
HIV-1	Human Immunodeficiency Virus Type 1
HIV-1B	Human Immunodeficiency Virus Type 1, Subtype B
HIV-1C	Human Immunodeficiency Virus Type 1, Subtype C
HIV-2	Human Immunodeficiency Virus Type 2
HIV/AIDS	Human Immunodeficiency Virus, Acquired Immunodeficiency Syndrome
HIVE	Human Immunodeficiency Virus Encephalopathy
HTLV-III	Human T-Lymphotropic Retrovirus, Type 3
IL-1β	Interleukin-1β
IL-2	Interleukin-2
IL-6	Interleukin-6

IL-8	Interleukin-8
IL-10	Interleukin-10
IFN-γ	Interferon gamma
IN	Integrase
Inr	Initiator
IQR	Interquartile Range
IRIS	Immune reconstitution inflammatory syndrome
KS	Kaposi's Sarcoma
LacZ	B-Galactosidase
LAV	Lymphadenopathy Associated Virus
LB	Lysogeny Broth
LEDGF/p75	Lens Epithelium-derived Growth factor
LTR	Long Terminal Repeat
Luc	Luciferase
M	Molar
M.tb	Mycobacterium Tuberculosis
MEGA	Molecular Evolutionary Genetics Analysis
$MgSO_4$	Magnesium sulfate
Min	Minutes
Ml	Milliliter
Mm	Millimolar
mRNA	Messenger RNA
NaOAc	Sodium Acetate
Nef	Negative Regulation Factor
NELF	Negative Elongation Factor
NF-κB	Nuclear Factor Kappa-Light-Chain-Enhancer Of Activated B Cells
ng	Nanograms
^{O}C	Degrees Celsius
OPTI-MEM	Reduced-Serum Medium An Improved Minimal Essential Medium
ORF	Open Reading Frame
PBS	Phosphate-Buffered Saline

Pcf11	Pre-mRNA 3'-end processing factor
PCP	Pneumocystis Pneumonia
PCR	Polymerase Chain Reaction
PIC	Preinitiation Complex
PID	Patient Identification
Pol	Polymerase
Pol II	RNA Polymerase II
PR	Protease
P-TEFb	Positive Transcription Elongation Factor B
RBD	RNA Binding Domain
Rev	Regulator Of Virion Protein
RGD	Arg-Gly-Asp
RLU	Relative Luminescence Unit
RNA	Ribonucleic acid
RNase	Ribonuclease
RNase H	Ribonuclease H
RNaseOUT	Recombinant Ribonuclease Inhibitor
rpm	Revolutions per minute
RT	Reverse Transcriptase
S/Nt	Number Of Nucleotide Differences Between Two Sequences/Total Nucleotide Sequences
SEM	Standard Error Of The Mean
SGA	Single Genome Amplification
SOC	Super Optimal Broth
SSIV	SuperScript IV Reverse Transcriptase
SU	Surface Protein
TAR	Transactivation Response Element
Tat	Transactivator Of Transcription
TB	Tuberculosis
TBM	Tuberculosis meningitis
TFIIA	Transcription factor II A

TFIIB	Transcription factor II B
TFIID	Transcription factor II D
TFIIE	Transcription factor II E
TFIIF	Transcription factor II F
TFIIH	Transcription factor II H
Th2	T Helper Cell Type 2
TM	Transmembrane Protein
TNF-α	Tumor necrosis factor alpha
TNF -β	Tumor necrosis factor beta
U/µl	Enzyme Unit Per Microliter
U/Rx	Enzyme Unit Per Reaction
UNAIDS	Joint United Nations Programme on HIV/AIDS
VDRL	Venereal Disease Research Laboratory test
VESPA	Viral Epidemiology Signature Pattern Analysis
Vif	Viral Infectivity Factor
Vpr	Viral Protein R
Vpu	Viral Protein U
WHO	World Health Organization
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Chapter 1

1. Introduction

In June 1981, the Centers for Disease Control and Prevention (CDC) had published a weekly report that described five cases of *Pneumocystis carinii* pneumonia (PCP) in homosexual North-American men (CDC, 1981a). The PCP infection is known for causing pulmonary inflammation and fluid buildup in the lungs. The five PCP cases were reported to be from non-related individuals who were all sexually-active homosexual men. Although all 5 men were reported to be using inhalant drugs, they were otherwise healthy (CDC, 1981a). However, PCP is an opportunistic infection, known to infect severely immunosuppressed individuals. At the time the report was published, two of the five men had died (CDC, 1981a). A subsequent report in July 1981, following the PCP incidence had described the occurrence of a rare malignancy known as Kaposi sarcoma (KS), found within 26 homosexual men (CDC, 1981b). The KS was described an aggressive cancer affecting elderly men. The KS cases began to spread, initially infecting 41 homosexual men – eight of which died in less than 2 years after being diagnosed (CDC, 1981b).

In August 1981, a case definition was then established to control the outbreak of these infections. The case definition consisted of two criteria; (1) either persons acquiring opportunistic infections with no known underlying illness or history of immunosuppressive therapy or (2) individuals over 60 years of age with a confirmed biopsy of KS. Although there were a few cases of heterosexual women fitting the case definition, due to majority of the cases being documented amongst homosexual men, the disease was termed Gay-Related Immunodeficiency (GRID) (CDC, 1981b). However during December 1981 and early 1982, due to the growing increase in case reports describing heterosexual populations infected with KS and other opportunistic infections, it became evident that the disease was not limited to the homosexual population and the CDC renamed the condition Acquired Immune Deficiency Syndrome or AIDS (CDC, 1982a, CDC, 1982b). Hence it was only in 1982, when the case definition for AIDS was first described. Whilst the cause of the condition was still unknown the AIDS pattern was characterized by low CD4+ T cell counts (below 200 cells/mm^3), inflamed lymph nodes (lymphadenopathy) and susceptibility to opportunistic infections (CDC, 1982b).

Following the CDC's case definition of AIDS in 1982, two independent research teams identified a novel retrovirus present in individuals with AIDS. In 1983 Montagnier and co-workers isolated the Lymphadenopathy-associated virus (LAV) and characterized it as part of the Lentivirus genus, Retroviridae family (Barré-Sinoussi et al., 1983). Separately, Gallo and co-workers named it human T lymphotropic virus type III (HTLV-III) (Gallo et al., 1984). However in 1986, it was decided by the International Committee on Taxonomy of Viruses that the virus be renamed to Human Immunodeficiency Virus (HIV) (Francki et al., 1991). Subsequently both the Montagnier and Gallo groups provided integral information which revealed that HIV is the causative agent of AIDS. Since then, substantial research has been done on HIV, allowing for the development of laboratory and diagnostic tests used to detect HIV antibodies and biomarkers for HIV progression.

1.1 The Epidemiology of HIV

Since the discovery of HIV as the causative agent of AIDS (Barre-Sinoussi, 1983, Gallo et al., 1984), the HIV/AIDS pandemic has continued to grow on a global scale. According to the recent Joint United Nations Programme on HIV/AIDS (UNAIDS) 2018 global report, there were 36.9 million people living with HIV, 1.8 million new HIV infections and 940 000 AIDS related deaths (UNAIDS, 2018). The HIV virus is categorized into two genetically distinct viral types: HIV type 1 (HIV-1) and HIV type 2 (HIV-2). While HIV-2 is geographically restricted to West Africa, HIV-1 is distributed worldwide (Lemey et al., 2003). The HIV-1 virus is arranged into four distinct phylogenetic groups: Group M (major), group O (outlier), group N (non-major and non-outlier) (Hemelaar et al., 2006) and the recently identified group P in Cameroon (Vallari et al., 2011).

The Group M viruses are further divided into 9 different subtypes: A, B, C, D, F, G, H, J, K and 90 circulating recombinant forms (CRF's) (Foley B, 2017). The CRF's are the product of the recombination of different viral subtypes that in essence form a new strain of the virus (Thomson et al., 2002). However, HIV-1 viral subtypes and CRF's are not evenly distributed throughout the globe (Figure 1.1).

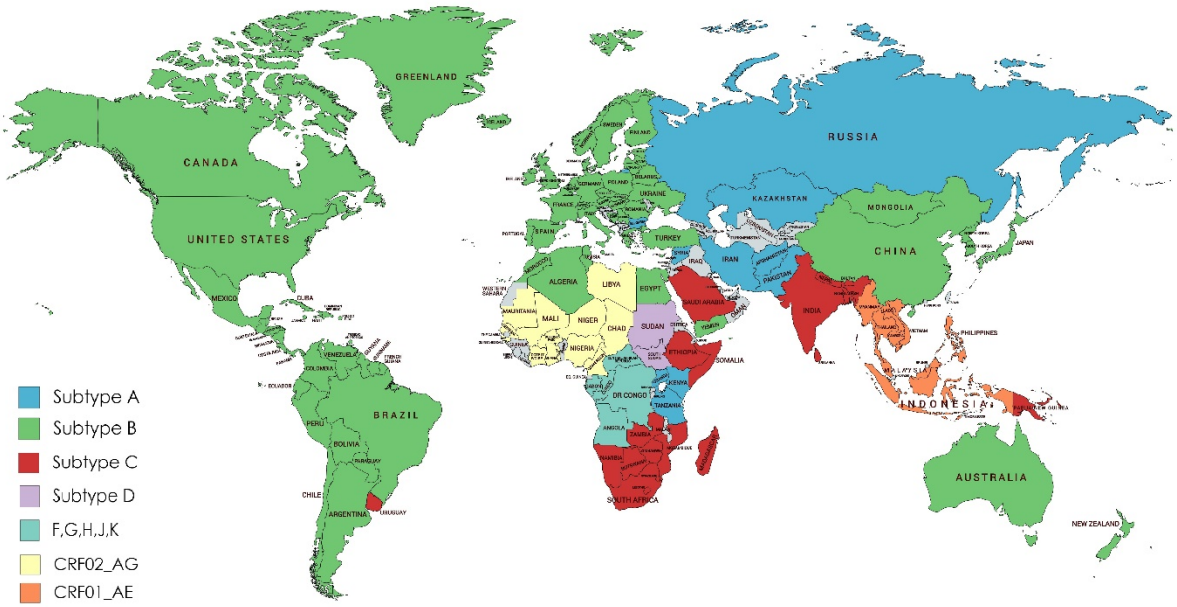


Figure 1:1 Global distribution of HIV-1 subtypes and circulating recombinant forms. The HIV-1 viral subtypes are distributed worldwide. Subtype A (blue) is commonly found in Eastern Europe, Central Asia and Central and Eastern Africa. Subtype B (green), is distributed in the Americas, Western Europe and Australasia. Subtype C (red) is found within India, Eastern and predominantly in Southern Africa. Subtype D (purple) is mainly present in East Africa while subtypes F, G, H, J, K as represented in turquoise are concentrated in West Africa. The CRF's; CRF_01_AE (orange) are found in Southeast Asia and CRF_02_AG (yellow) in West Africa. Source: Adapted from (Santos and Soares, 2010, Hemelaar, 2012).

The HIV-1 subtype A predominates in Eastern Europe, Central Asia and Central and Eastern Africa, with a global prevalence of 12% (Taylor et al., 2008). The HIV-1 subtype B is widely distributed in the Americas, Western Europe and Australasia, however it only accounts for 10% of global infections. The HIV-1 subtype D is mainly present in East Africa and is associated with a rapid disease progression in Uganda, Kenya and Tanzania (Kaleebu et al., 2002). Although HIV-1 subtype D is more virulent it represents only 3% of global infections. Subtype G is mainly concentrated in West Africa, with a global prevalence of 6%. Whilst subtype F, H, J and K are scattered in various countries and represent a minor proportion of <1% of the global spread of HIV-1. The two major CRF's making up a significant proportion of individuals infected are CRF01_AE and CRF02_AG which are found in Southeast Asia and West Africa respectively, and they account for

approximately 5% of the global prevalence of HIV-1. Whilst the African continent contains a heterogeneous mix of viral subtypes, HIV-1 subtype C found within India, Eastern Africa and predominantly in sub-Saharan Africa, is responsible for approximately 50% of global HIV-1 infections (Hemelaar et al., 2011). Hence the sub-Saharan African region bears the brunt of the HIV/AIDS pandemic (Thomson et al., 2002). Therefore, HIV-1 subtype C (HIV-1C) is the focus of this study.

1.2 HIV-1 genome organisation

The HIV-1 virus is a member of the Lentivirus genus, Retroviridae family and Orthoretrovirinae subfamily (Coffin et al., 1986a, Coffin et al., 1986b). The viral genome consists of two single stranded ribonucleic acid (RNA) molecules enclosed within the virus core (Kurth and Bannert, 2010). The integrated viral deoxyribonucleic acid (DNA), known as the provirus, is approximately 9.8 kilobases (kb) in length consisting of 9 genes, flanked by identical long terminal repeat (LTR) on both at the 5' and 3' end of the HIV-1 genome (Figure 1:2)(Freed, 2001). The HIV-1 genes are divided into three groups: structural, accessory and regulatory. There are three structural genes: group specific antigen (*gag*); polymerase (*pol*); and envelope (*env*), four accessory genes: viral protein U (*vpu*); viral protein R (*vpr*); viral infectivity factor (*vif*) and negative regulation factor (*nef*). Lastly two regulatory genes: regulator of virion expression (*rev*) and transactivator of transcription (*tat*) each comprising two exons (Kuiken et al., 1999). Moreover, the first exon of *tat* have open reading frames (ORF) which partially overlap with *vpu* and *rev* (Fernandes et al., 2016).

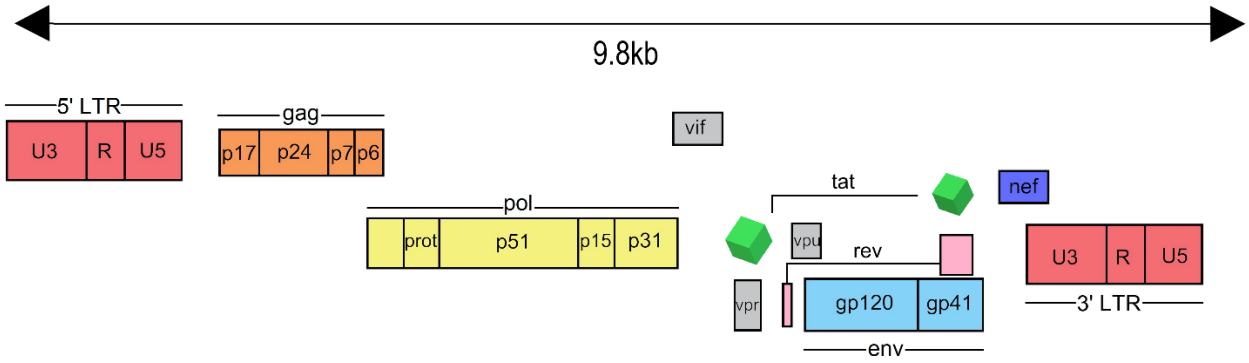


Figure 1:2 HIV-1 genome organisation. The HIV-1 provirus genome is approximately 9.8kb, consisting of 9 genes flanked by long terminal repeats (red) on either end. Structural genes; gag, pol and env (orange, yellow and light blue), accessory genes; vpu, vpr, vif (grey), and nef (dark blue) and two regulatory genes; rev and tat (pink and green) each gene comprising of two exons. Source: Adapted from (Suzuki and Suzuki, 2011, http://www.hiv.lanl.gov, 2018)

The HIV-1 genome has 15 proteins encoded by 9 overlapping genes (Frankel and Young, 1998). The Gag protein encoded by *gag* is a polyprotein encoding the inner structural proteins of the virus – the matrix (p17), capsid (p24) and nucleocapsid (p7) responsible for mediating viral assembly (Frankel and Young, 1998). The Pol polyprotein encoded by *pol* gets cleaved into three constituent viral enzymes; protease (prot), reverse transcriptase (p51) and integrase (p31) (Frankel and Young, 1998). The Env glycoprotein (gp) encoded by *env* is a gp160 protein gets processed into gp120 surface protein (SU) and gp41 transmembrane protein (TM) making up the structural components of the outer viral envelope (Frankel and Young, 1998). In addition, the accessory proteins Vpu, Vpr, Vif and Nef encoded by *vpu, vpr, vif and nef* respectively, are responsible for particle infectivity and release of HIV-1 virions from the infected cell. Lastly two regulatory proteins Rev encoded by *rev* is responsible of nuclear export of viral messenger RNA (mRNA) into the cytoplasm and Tat encoded by *tat* whose function is enhancing viral transcription (Frankel and Young, 1998). The Tat protein is the focus of this study and will be discussed further below.

1.3 The HIV-1 Replication cycle

To establish infection the virus follows a replication cycle which includes: binding and entry, reverse transcription, integration, transcription, translation, assembly, budding and virion maturation (Figure 1:3) (Freed, 2001). In summary, following cell entry the viral genome RNA is reverse transcribed to form double-stranded DNA, which is integrated stably into chromosomal DNA also known as a provirus (Fauci, 1988).

Entry of HIV-1 into target cells occurs through the binding of the HIV-1 surface envelope glycoprotein gp120 to the $CD4^+$ antigen receptor protein on the surface of the susceptible host cell (Ferguson et al., 2002). In addition to the $CD4^+$ receptor there is also a co-receptor on the host cell, which differs for host cell types. The C-C chemokine receptor 5 (CCR5) protein commonly found on a subset of helper T lymphocytes, monocyte-macrophages dendritic cells, and Langerhans cells whereas the CXC chemokine receptor 4 (CXCR4) protein found on T cells subsets (Freed, 2001). This high affinity receptor and co-receptor binding results in the fusion of the viral and cellular membranes, mediated by glycoprotein gp41. Thus, allowing the viral core composed of p17 matrix protein and the p24 capsid protein to enter the cytoplasm of the susceptible cell (Freed, 2001).

Subsequent to entry, the uncoating process of the viral core allows for the release of two identical viral RNA copies and viral proteins including the three viral enzymes; reverse transcriptase (RT), integrase (IN) and protease (PR) into the cytoplasm of the infected cell. Once inside the cell, the HIV-associated RT directs the synthesis of single stranded DNA copy from the viral RNA genome in a process known as reverse transcription (Kuiken et al., 1999). During DNA synthesis, the ribonuclease H (RNase H) domain from RT degrades the RNA strand and produces RNA primers which are used to synthesize complimentary DNA, resulting in a double stranded DNA molecule (Boyer et al., 2018).

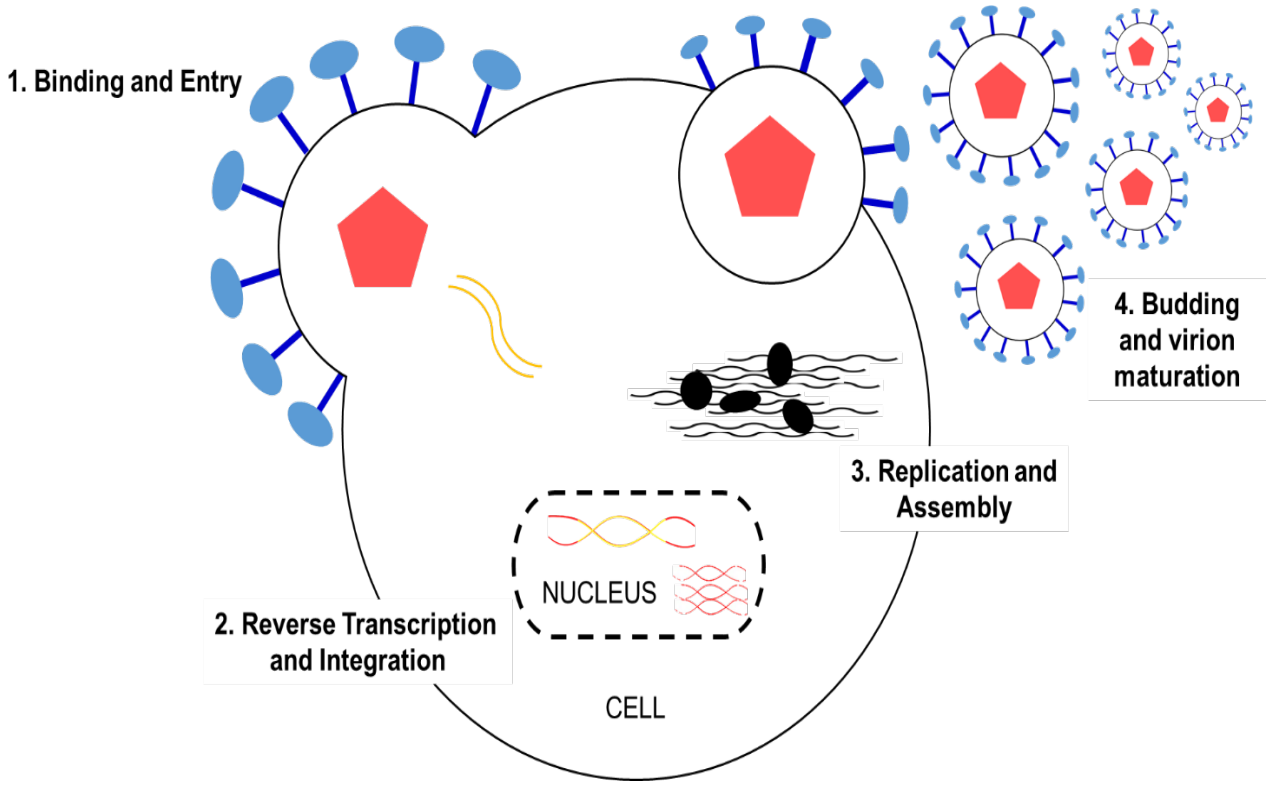


Figure 1:3: HIV-1 replication cycle. The HIV-1 replication cycle described in 4 major steps: Binding and entry whereby HIV-1 enters the cell through receptor-coreceptor binding and fusion of viral and cellular membranes; reverse transcription and integration of the viral RNA into proviral DNA; replication and assembly of mRNA into viral proteins and lastly budding of newly made viruses (intracellular) and maturation of virions (extracellular). Source: Adapted from (Freed, 1998, Kirchhoff, 2013).

The resulting double stranded DNA then moves into the nucleus and is inserted into the host chromosome by integrase (IN) with cellular proteins including but not limited to the Lens Epithelium-derived Growth factor (LEDGF/p75), helping in tethering the HIV-1 DNA and proteins to the chromatin (Christ and Debyser, 2013). Subsequently, the integrated proviral DNA and cellular DNA are replicated through cell division cycles. With the proviral DNA acting as a template for viral gene transcription. Transcription is initiated in the 5' LTR of the provirus. The viral LTR element consists of 3 regions (U3, R and U5) each involved in transcription. Transcription is mediated by the host-cell RNA polymerase II and upregulated by the viral transactivator of transcription protein (Tat) yielding numerous copies of viral genome messenger RNA (mRNA)(Coffin et al., 1997). The transactivation

(enhancing) ability of Tat in HIV-1 transcription will be discussed subsequently in detail as it is the focus of this study.

The newly synthesized viral mRNA is then transported into the cytoplasm where it is spliced and translated, producing viral enzymes and structural proteins (Coffin et al., 1997). Some of the structural proteins are formed by cleavage of a long polyprotein by the PR (Gomez and Hope, 2005). Assembly of the virus commence when core structural proteins; p17 matrix protein and the p24 capsid protein encapsulate the newly formed viral RNA genome with a viral capsid. Viral glycoproteins gp120 and gp41, which will eventually form the viral envelope, are transported to the surface of the cell through cellular secretory mechanisms (Gomez and Hope, 2005).

Budding occurs when the viral capsid extrudes out of the cell membrane of the infected cell. This process is mediated by Gag and the host cellular endosomal sorting complexes required for transport (ESCRT) pathway (Sundquist and Kräusslich, 2012). Briefly, viral budding occurs in two stages: (1) membrane deformation, whereby a plasma membrane is wrapped around the nascent virion and (2) membrane scission, whereby the newly formed bud is detached from the host membrane (Votteler and Sundquist, 2013). The combination of these steps is required for virion release and maturation. Virion maturation is mediated by viral PR, allowing for cleavage of structural proteins and viral enzymes which results in the formation of an infectious virion (Sundquist and Kräusslich, 2012).

1.4 Structure and Function of HIV-1 Tat

As aforementioned, the *tat* gene encodes Tat, whose length varies between 86 – 101 amino acids in length, with the predominant form of Tat being 101 residues (Jeang, 1996). The structure of Tat will be discussed in section 1.4.1. The Tat protein plays a crucial role in the upregulation of viral gene transcription by transactivating the 5’ LTR viral promoter. The transactivation activity of Tat occurs upon binding to its RNA target, the transactivation response element (TAR), located in the 5′ LTR (Dingwall et al., 1989). The

function of Tat will be discussed in section 1.4.2. In addition to its transactivation activity, Tat has both chemotactic and neurotoxic which contribute to its pathogenesis.

The viral Tat functions at an intracellular and extracellular level as reviewed by (Chiozzini and Toschi, 2016). The Tat proteins produced within an HIV infected cell allows for of viral gene expression by transactivating the 5' LTR. In addition, other non-transcriptional functions of intracellular Tat include HIV-1 RNA splicing, capping, translation, and reverse transcription (Das et al., 2011). Whilst, extracellular Tat is involved in modulating cellular gene expression allowing for; apoptosis, cytokine dysregulation, and neurotoxicity all of which contribute to its role in HIV-1 pathogenesis. The neurotoxic activity of Tat stems from its ability to travel along neuronal pathways as a single protein molecule and infect various central nervous system (CNS) cells. Several neuropathogenesis studies have reported the presence of Tat and elevated expression of its mRNA in the CNS of patients with HIV-1 encephalitis (HIVE) and HIV-1-associated dementia (HAD) (Cowley et al., 2011). However, the transactivation ability of Tat will be further described in this chapter as it is the focus of the study.

1.4.1 The structure of HIV-1 Tat

The Tat protein is approximately 86 – 101 amino acids (aa) in length (Jeang, 1996). The predominant form of Tat found in clinical isolates is 101 aa in length (Figure 1:4) (Zhang et al., 2014). The first 72 amino acids of Tat (aa: 1-72) are encoded by the first exon of the HIV-1 *tat* gene and are sufficient for the transactivation of the viral 5' LTR (Jeang, 1996). These 72 amino acids are divided into 5 domains. An N-terminal acidic domain (domain I) comprising aa: 1 - 21 of Tat, this is a proline rich acidic domain with 6 proline residues. This domain has been reported to tolerate mutations well, with only Tryptophan 11 (W11) that is conserved (Campbell and Loret, 2009). A cysteine rich domain (domain II) consisting of aa: 22–37, with seven well conserved cysteine amino acids at positions 22, 25, 27, 30, 31, 34 and 37, except in subtype C where a C31S mutation is more frequently observed (Campbell and Loret, 2009). Aside from the C31S mutation, a mutation at any one of the cysteine sites will result in a non-functional Tat (Kameoka et

al., 2002). Domain III with aa: 38–48 is the core domain, containing a conserved Phenylalanine (F38) and amino acid sequence LGISYG (Campbell and Loret, 2009). Hence, the first three domains, composed of the first 48 amino acids of Tat form the Tat activation domain that is responsible for the transcriptional elongation and binding to Cyclin T1 (Tahirov et al., 2010).

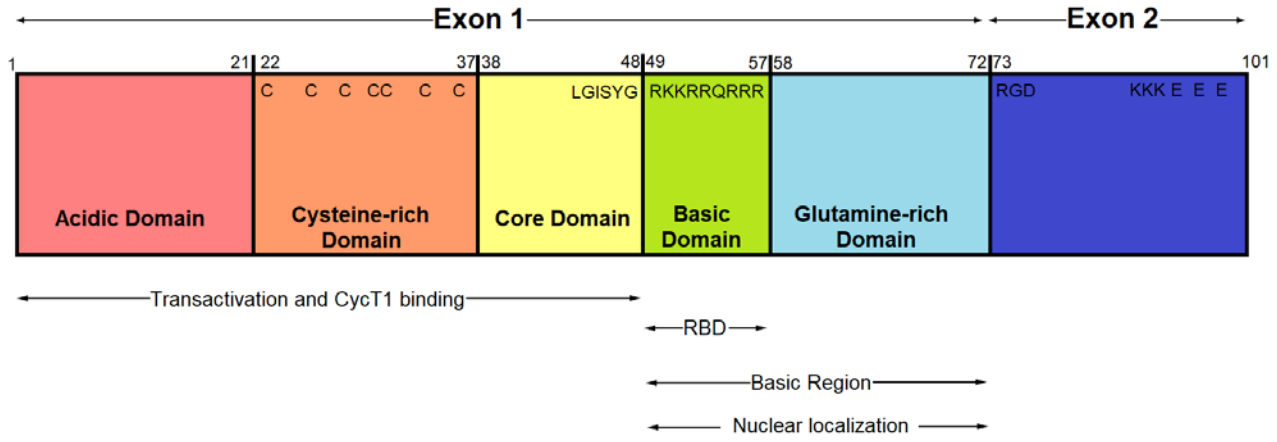


Figure 1:4: The structure and functional domains of Tat. The Tat protein is approximately 86-101 amino acids in length and consists of 5 functional domains that are encoded by exon 1. These include the acidic domain (red), cysteine-rich domain (orange), core domain (yellow) made of a hydrophobic motif, basic domain (green) is an RNA-binding domain (RBD) and a glutamine-rich domain (light blue). The first 48 amino acids allow for the transactivation activity of Tat via Cyclin T1 (CycT1) binding which is responsible for enhancing viral transcription. The basic domain made up of amino acids 49 – 57 is the RNA binding domain (RBD) which allows for the binding ability of Tat to the TAR loop structure of the 5' LTR allowing for transactivation of viral transcription. Both the basic and glutamine-rich domains are concentrated with basic amino acids, this basic region is also responsible for nuclear localization of Tat. Domain VI which is encoded in exon 2 (dark blue) is comprised amino acids 73 – 101 which allow for integrin binding of Tat. Adapted from: (Li et al., 2012, Jeang, 1996).

Domain IV, is the basic domain, made of aa: 49–57 which has a well conserved RKKRRQRR basic amino acid sequence. This domain is also known as the RNA binding domain (RBD), responsible for the binding ability of Tat to the transactivation response (TAR) element found in the 5' LTR, viral promoter (Bagashev and Sawaya, 2013). Domain V, consisting of aa: 60–72, forms the glutamine-rich domain – which also has high

sequence variation. The second exon of Tat (Domain VI), composed of amino acids 73 – 101 in the C-terminus of Tat is not well studied. Although the function of the second exon of Tat is still unclear, it has been implicated to contribute towards overall optimal transactivation and transrepression of Tat allowing for increased infectivity (Jeang, 1996). Furthermore exon 2 contains an arginine-glycine-aspartic acid (RGD) motif which supports integrin binding and signalling allowing for increased cellular update of Tat (Vogel et al., 1993).

1.4.2 Mechanism and Function of Tat

The primary function of Tat is to enhance viral gene transcription by binding to the TAR element located at the 5' LTR, viral promoter. The binding of Tat to the TAR element allows for the elongation of RNA polymerase II (Pol II) through the recruitment of various transcription factors that will in turn increase the processivity of transcription (Lu et al., 2013). In the absence of Tat, viral gene transcription is significantly inhibited due the loss of Pol II processivity allowing mRNA to separate from the DNA template prematurely (Karn, 2000).

1.4.2.1 Transcription initiation from the HIV-1 core promoter region

Transcription is initiated at the first nucleotide (+1) in the R region of the HIV-1 5’ LTR (5’ LTR). The 5’ LTR alone is capable of initiating and driving basal gene transcription – albeit a low level in the absence of Tat, resulting in short abortive viral mRNA transcripts (Lassen et al., 2004). Transactivation of transcription is primarily due to the presence of Tat (Karn, 2000). Briefly, initiation of HIV-1 transcription occurs when Pol II binds to the core promoter region of the 5’ LTR, which is inclusive of: the TATA box, initiator (Inr) at the transcription start site +1, downstream core promoter element (DPE) and transcription factor B (TFIIB) that is bound to the B recognition element (BRE) (Butler and Kadonaga, 2002). These core control elements allow for the assembly of basal transcription proteins. A preinitiation complex (PIC) is formed , through the binding of the transcription factor II D (TFIID) to the TATA box in the core promoter region (Ne et al., 2018). This is followed

by five more TF's (TFIIA, TFIIB, TFIIE, TFIIF, and TFIIH) binding to the Pol II. The combined complex is able to facilitate transcription initiation (Ne et al., 2018).

Whilst basal transcription (in the absence of Tat) is only able to effectively initiate transcription, elongation is defective (Karn, 2000). Consequently, a Tat-mediated transcription allows for an effective elongation, following viral transcription initiation. Thus, ensuring that the full-length viral mRNA is transcribed and processed to make the necessary proteins. Hence Tat resumes transcriptional control solely at the elongation stage of HIV-1 transcription.

1.4.2.2 A Tat mediated elongation

As aforementioned in the absence of Tat, viral mRNA transcription is effectively initiated but Pol II only processes short abortive mRNA transcripts, this is due to the pause in transcription that is induced by the DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) soon after the transcriptional start site (+1) (Figure 1:5) (Zhou and Rana, 2002). This negative feedback loop occurs in basal conditions, preventing Pol II to proceed into elongation (Nilson and Price, 2011). Moreover, the pre-mRNA 3'-end processing factor (Pcf11) reinforces this block in elongation by prematurely terminating the transcription of the short nascent viral mRNA products – resulting in abortive mRNA transcripts (Natarajan et al., 2013). Histone deacetylases (HDAC) are also recruited to the paused complex to further reinforce a transcriptionally repressed chromatin state (Zhou and Rana, 2002).

This transcriptionally repressed state is relieved upon Tat binding to the TAR element in the in the R region of the 5' LTR, allowing for the elongation of Pol II (Lu et al., 2013). Hence, Tat specifically recognizes the TAR RNA loop and binds to a uridine-rich (U-rich) bulge in the apex of the loop (Dingwall et al., 1989).

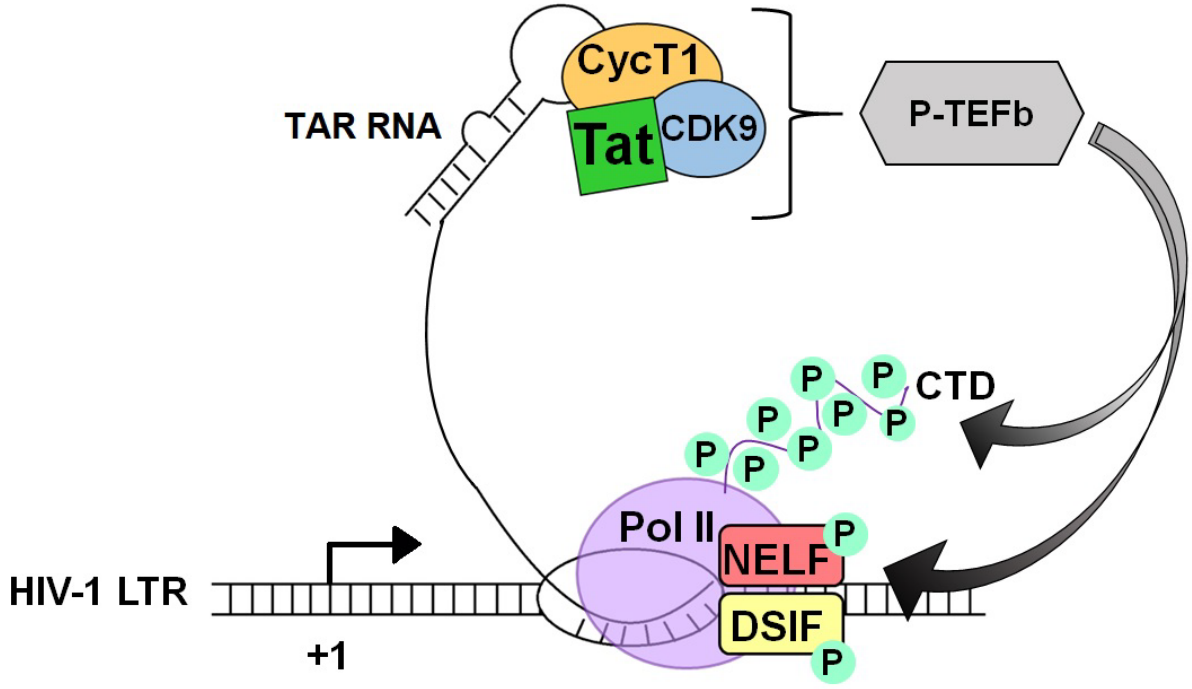


Figure 1:5 Model for Tat mediated HIV-1 LTR activation of RNA polymerase II. Transcription is initiated at the promoter region of HIV-1 5' LTR. The 5' LTR is capable of driving a basal gene transcription, however during basal conditions Pol II is only processed a short distance downstream from the transcriptional start site when the DRB sensitivity-inducing factor (DSIF; yellow) and negative elongation factor (NELF; red) induce a pause in transcription – creating a negative feedback loop. This transcriptionally repressed state is relieved upon the binding of the Tat protein to the 5' LTR. Tat establishes a positive feedback loop, upon binding to the transactivation response (TAR) element. Once Tat is bound to the TAR stem loop structure, it recruits positive transcription elongation factor (P-TEFb; grey), composed of CDK9 (blue) and Cyc T1 (orange), which drives the phosphorylation of the CTD allowing for transcriptional activation from the HIV-1 LTR. Tat mediates hyperphosphorylation of the CTD relieving the negative feedback loop. Hyperphosphorylation of the CTD allows for enhanced transcription. Source: Adapted from (Zhou and Yik, 2006, Liu et al., 2014).

Once Tat is bound to the TAR element, it mediates the recruitment of positive transcription elongation factor (P-TEFb). The P-TEFb ternary complex is composed of the catalytic subunit cyclin-dependent kinase 9 (CDK9), and the regulatory subunit Cyclin T1 (Cyc T1), that drive the phosphorylation of the C-terminal domain (CTD) of Pol II, allowing for activation of transcriptional elongation from the HIV-1 LTR (Karn and Stoltzfus, 2012).

Hence phosphorylation of the Pol II CTD are controlled by kinases (CDK9, Cyc T1) which allow for regulation of viral transcription (Francis et al., 2011). The CTD of Pol II is the combination of multiple heptad repeats of amino acids (tyrosine 1 – serine 2 – proline 3 – threonine 4 – serine 5 – proline 6 – serine 7), whose phosphorylated state is crucial to its activity, and therefore maintained until transcription termination (Ne et al., 2018).

Once P-TEFb mediates hyperphosphorylation of the CTD of Pol II, the negative feedback loop (DSIF and NELF) is relieved from the complex (Karn and Stoltzfus, 2012). This allows for the recruitment of chromatin remodeling machinery such as histone acetyltransferases (HATs) which displace blocked nucleosomes and facilitates acetylation allowing nucleosomes to become more accessible (Ne et al., 2018). Ultimately, supporting transcription elongation of Pol II.

1.5 Genetic variation in Tat

Following infection, HIV-1 is characterized by a great degree of genetic diversity – due to a high mutation rate caused by error prone viral reverse transcriptase (RT), recombination and host factor proteins such as APOBEC 3G (Sanjuán and Domingo-Calap, 2016). Additionally a high replication rate, lifelong infection and large numbers of infected individuals all contribute to the circulation of different strains of the virus (Lidsky et al., 2017). Previous studies have shown that Tat variation within subtypes have significantly affected its transactivation ability (Kurosu et al., 2002, Roy et al., 2015, Rossenkhan et al., 2013). Consequently, functional variation of the protein leads to varying degrees of pathogenesis. Whilst HIV-1C is responsible for over 50% of global infections, there is limited information available for the transactivation abilities of subtype C Tat isolates.

One of the major HIV-1C studies showed that genetic variation within Tat was responsible for a significant change in the transcriptional activity of the viral LTR (Rossenkhan et al., 2013). The Essex group were able to show that HIV-1C Tat diversity within aa: 1 – 72 results in varying transactivation potential during primary infection which positively correlated with correlated with viral load (Rossenkhan et al., 2013). Whilst minimal genetic diversity and high homogeneity was observed in their phylogenetic analysis, they further

explored a mixed-effects model of evolution (MEME) to determine if any specific Tat residues were under positive selection. A preceding study that performed a signature pattern analysis identified five amino acid positions in Tat (21A, 24N, 29K, 40K, and 60Q) that contained signature residues unique for Indian HIV-1C (Neogi et al., 2012). However, the MEME analysis by the Essex group showed that residues; 3P, 4V, 21A, 24N, 29K, 39I, and 68L of Tat exon 1 were positively selected in African HIV-1C. Furthermore, site directed mutagenesis in the first exon of Tat was performed to determine the functional effects of these residues undergoing positive selection. As mentioned previously, *tat* exon 1 has an overlapping ORF with genes *vpu* and *rev*. Whilst there is paucity of data regarding the overlapping nature of non-structural genes, it should be noted that the MEME analysis done by the Essex group showed no evidence of positive selection mutations as a result of overlapping within Tat. Interestingly, the majority of the mutations were associated with a decrease in LTR transcription activity of Tat namely; E2D, V4I, P21A, K24S, L35Q, Q39L, P59S and P68S (Rossenkhan et al., 2013). Additionally, they examined patient derived Tat, due to the relative compensatory mutations that can occur when individual deleterious mutations occur in a natural infection. In the patient derived Tat isolates, the largest decrease (80%) in Tat activity was seen in a patient who had both E2D and P68L mutations. Conversely the largest increase was seen in a patient having both with a V4I and H29R that significantly upregulated LTR activation by 73%. Therefore, compensatory mutations in patient derived Tat could allow for a greater alteration in the functional capacity of Tat.

1.6 HIV-1 Tat in the Central Nervous System

The HIV-1 virus enters the CNS early in infection, by migrating across the blood brain barrier (BBB)(An et al., 1999). The mechanism of entry has been widely described as the ‘Trojan horse’ model, since the CNS is tricked into allowing HIV-1 infected monocytes, macrophages, lymphocytes, endothelial cells as well as free virions passed the BBB (González-Scarano and Martín-García, 2005). However, viral replication is marginally lower in the CNS due to the cellular tropism of HIV-1 target cells (Martín-García et al., 2006). Since neurons are not susceptible to HIV-1 infection the virus actively replicates in macrophages, whilst astrocytes that harbor the virus aren’t able to actively replicate

(Russell et al., 2017). Although a previous report showed that astrocyte activation can be triggered by neurotoxic proteins like Tat, allowing for the in vitro expression of Tat from astrocytes (Zhou et al., 2004). Additionally the virus is able to persist latently within the long lived cells (e.g. astrocytes) of the CNS whose half-lives range from months to years and even a lifetime (Gray et al., 2016). Furthermore the $CD4^+$ T helper cells (Th) found in the CNS are very few and do not actively replenish as the removal of large amounts of HIV-1 infected cells from the CNS may cause further neurological damage (Gray et al., 2016).

Although there is evidence that during mono-infection HIV-1 is able to evolve independently from non-CNS compartments, this finding is highlighted in HIV associated CNS disorders (Ene, 2018). Thus a repertoire of CNS disorders including; opportunistic infections, immune reconstitution inflammatory syndrome (IRIS) and HIV-associated neurocognitive disorders (HAND) (Ene, 2018) has resulted in sequence compartmentalization between the CNS and non-CNS compartments at the level of HIV-1 Env, Nef and Tat as reviewed by Gray et al., 2016. However, it is unclear if the source of compartmentalized HIV-1 is the result of an early seeded virus population or independent viral evolution in the CNS.

A study by Cowley et al., (2011) found distinctive sequence variations when comparing HIV-1 *tat* nucleotide sequences derived from CNS and non-CNS compartments of patients with HIV-1 associated dementia (HAD). This group further demonstrated that Tat amino acid variation predominated within the Tat activation domain and the basic domain (Cowley et al., 2011). This Tat variation could possibly explain the heterogeneity in the ability of different Tat variants to transactivate the viral promoter, 5' LTR (Cowley et al., 2011). Viral variants causing the distinct genetic compartmentalization in the CNS suggest that differing selection pressures are acting on HIV-1 within and outside the CNS (Ohagen et al., 2003).

Many studies suggest that immune selection pressures in the CNS are the result of upregulated inflammatory cytokine responses mediated by Tat through various signaling

cascades (Rappaport et al., 1999, Deshmane et al., 2009, Cupp et al., 1993). As previously mentioned, Tat activates many cellular genes in addition to the HIV-1 LTR, these include: tumor necrosis factor alpha (TNF-α), transforming growth factor beta 1 (TGFβ-1), interleukin-2 (IL-2), interleukin-6 (IL-6) and interleukin-10 (IL-10) in cells of the CNS leading to dysregulation of cytokine expression (Sawaya et al., 1998). During CNS inflammation, pro-inflammatory cytokine TNF-α levels are elevated once macrophages and monocytes are exposed to Tat (Haij et al., 2015). Furthermore the upregulation of TNF-α expression increases the permeability of the BBB creating a paracellular route for HIV-1 infected macrophages and monocytes to invade the CNS (Brabers and Nottet, 2006, Rappaport et al., 1999). In addition Tat is able to synergistically work with TNF-α to enhance the T Helper Cell Type 2 (Th2) cytokine, IL-6 secretion and activate CNS endothelial cells (Bagashev and Sawaya, 2013, Chen et al., 1997). Elevated levels of IL-6 can be correlated with stages of HIV disease progression (Li et al., 2010). Another Th2 cytokine induced by Tat is IL-10 (an anti-inflammatory cytokine) which is expressed in microglial and astrocytes cells during bacterial or viral invasion (Brabers and Nottet, 2006). Usually IL-10 is produced to repair damages caused by excessive inflammation, high levels of IL-10 allow for inhibition of HIV replication (Naicker et al., 2009). On the contrary IL-10 single nucleotide polymorphisms in the proximal promoter region is associated with lower IL-10 production and a faster progression to AIDS (Naicker et al., 2009). Hence IL-10 serves as a feedback regulator to stop overproduction of cytokines; IL-6 and TNF-α, interleukin-8 (IL-8), interleukin-12 (IL-12), or even IL-10 itself which aids in the development of chronic inflammation. However during CNS inflammation a state of immune dysregulation rises and Tat induces the extensive production of both IL-6 and IL-10 (Li et al., 2010).

1.7 HIV-1 and tuberculosis meningitis (TBM) co-infection

Tuberculosis (TB) caused by *Mycobacterium tuberculosis* (*M. tb*) is the most acquired opportunistic infection in HIV/AIDS patients (Bruchfeld et al., 2015). The burden of the HIV and TB co-epidemic and low socioeconomic status in sub-Saharan Africa – make HIV vaccine development a priority. Furthermore, in hyperendemic settings such as the

sub-Saharan Africa region, disease progression to AIDS is often accompanied by opportunistic infections including but not limited to TB, which increase the morbidity and mortality in HIV-1 infected individuals. According to the recent World Health Organization (WHO) epidemic update, an estimated 40% of AIDS deaths were due to TB (WHO, 2018). Additionally, HIV-1 infected individuals with active TB have a high risk of extra pulmonary forms of tuberculosis (Thwaites et al., 2005).

Tuberculosis meningitis (TBM) is the most severe extrapulmonary manifestation of TB, caused by the haematogenous spread of TB from the primary source of infection, usually the lungs (García-Moncó and Rodriguez-Sainz, 2018). As a result TBM is associated with a high mortality rate and severe neurological consequences (Faksri et al., 2018). Whilst the incidence of TBM in sub-Saharan Africa is still uncertain due to lack of diagnostic tools, a report by Patel et al., (2013) observed that approximately 10% of TB patients experience TBM in KwaZulu-Natal, South Africa (Patel et al., 2013).

Furthermore, in an HIV-1 and TB co-infection both diseases become exacerbated, especially when TB is disseminated to the CNS (Be et al., 2009). Whilst the cause of TB in the CNS is not well described, the high HIV-1 replication in the CNS results in viral diversity, which impacts its replication capacity (Seipone et al., 2018). To the best of our knowledge, the effect of HIV-1 *tat* genetic variation on the modulation of viral transcription in the CNS of TBM patients has not been studied. Thus, understanding the effects of HIV-1 *tat* genetic variation on viral gene transcription in matched lymphoid and CNS compartments of TBM versus non-TBM patients will enhance our understanding of HIV-1 replication in the CNS. Moreover, understanding the mechanisms that drive high HIV-1 replication *in vivo* could provide valuable information which could guide the design of interventions aimed at HIV-1 cure and [or] remission.

1.8 Hypothesis

We hypothesize that CNS derived HIV-1 Tat may display CNS specific mutations that may mediate higher viral replication in the CNS compartment of TBM co-infected patients.

1.9 Aim

The purpose of this project was to characterize the effect of CNS derived HIV-1 *tat* genetic variation and functional activity on viral replication in this compartment from TBM and non-TBM patients.

1.10 Objectives

1.10.1 To amplify and sequence CNS and plasma derived HIV-1 *tat* obtained from TBM and non-TBM patients in order to study genetic variation between these compartments.

1.10.2 To perform a Tat transactivation assay to study the transactivation ability of different Tat variants.

1.10.3 To correlate Tat mutations with markers of disease progression such as viral load in the CNS.

Chapter 2

2. Methodology

2.1. Study Population

A total of 20 study participants were selected from a larger prospective cohort of 235 TBM patients, as previously described (Patel et al., 2013). Using non-probability (purposive) sampling, 17 of 108 HIV+ samples with a high CSF viral load was taken from the TBM group. Whilst 3 out of 70 HIV+ samples with a low viral load in the CSF was taken from the non-TBM group.

Briefly, all patients in the cohort were clinically assessed for tuberculosis meningitis; by having a computerized tomography (CT) scan done to remove contraindications to lumbar puncture. Plasma samples taken from patients underwent routine laboratory testing including: diagnostic syphilis test by fluorescent treponemal antibody absorption (FTA-ABS), venereal disease research laboratory test and screening for HIV by enzyme-linked immunosorbent assay (ELISA) and $CD4^+$ count. Additionally, 15ml of cerebrospinal fluid (CSF) samples collected by lumbar puncture, were subjected to routine tests such as: testing for acid-fast bacilli by gram staining, fluorescent staining using auramine, Mycobacteria testing using BD BactecMGIT 960, detection of Mycobacterium tuberculosis (*M.tb*) using the Roche Amplicor Mycobacterium Tuberculosis PCR Test, and fungal culture testing using a cryptococcal latex agglutination test. Subsequently, the CSF samples were analyzed by routine chemistry tests for protein, glucose, chloride levels, and viral PCR assay for cytomegalovirus, varicella zoster virus, and herpes simplex. Lastly CSF samples were subjected to a diagnostic syphilis test by venereal disease research laboratory test (VDRL), fluorescent treponemal antibody absorption (FTA-ABS) and cysticercus antibodies tests. Each sample of CSF was subsequently biobanked using Xpert MTB/RIF analysis.

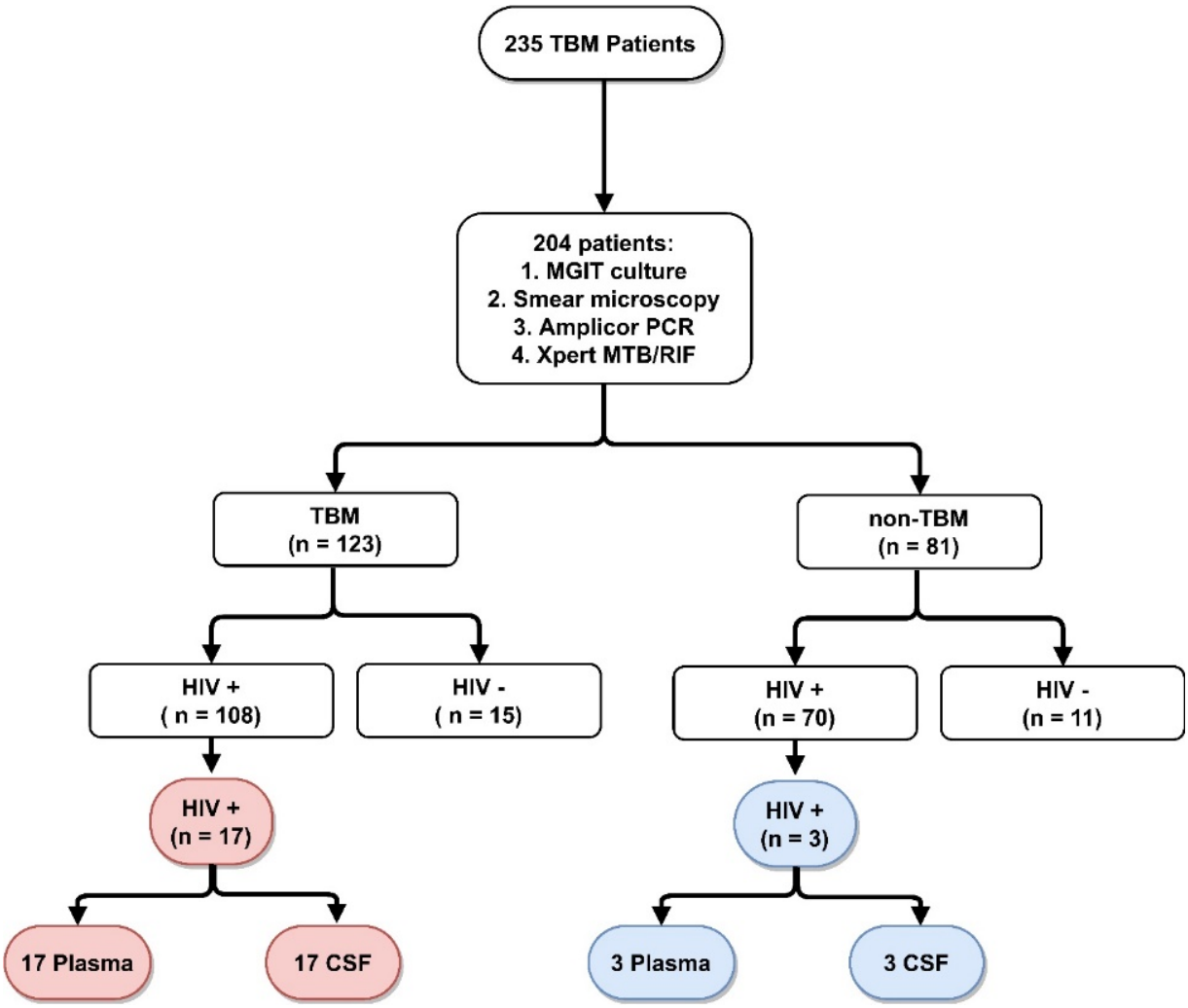


Figure 2:1 Summary flow chart of cohort design and patient selection. Initially 235 TBM patients with a meningeal like illness were recruited, 31 patients were excluded, and the remainder were categorised, based on published diagnostic criteria, as TBM and non-TBM. These patients were further divided into HIV positive and HIV negative. Both plasma and CSF samples were extracted from each HIV positive patient. Adapted from: (Patel et al., 2013).

Clinical information recorded included: patient demographics, extent of symptoms, patients treated with either anti-tuberculous or steroid therapy, HIV status, history of TB, and history of TB contact. The test results obtained were used to categorize patients into three groups namely: definite TBM, probable-TBM and non-TBM based on standardized published diagnostic criteria (Thwaites et al., 2009). Hence, patients with positive CSF

M.tb culture and/or amplicor PCR were regarded as definite TBM, whilst patients that did not meet the standards for being definite TBM but were treated empirically with anti-TB drugs categorized as probable-TBM. Lastly, patients with a CSF *M.tb*. culture negative had an additional confirmatory alternate diagnosis and responded to therapy in the absence of anti-TB treatment were classified as non-TBM.

For the purposes of this study all 20 (n=40) chronically infected, antiretroviral therapy (ART) naïve study participants were made available. Out of 20, 17 were defined as the TBM group, and 3 as the non-TBM control group. Each patient had a matched plasma and CSF sample recorded by a respective patient identification (PID). All together 17 TBM (n=34 samples) and 3 non-TBM (n=6 samples) participants, making a total of 20 PID's (n=40 samples) were analysed (Figure 2:1).

2.1.1. Baseline Demographic and Clinical Characteristics

The demographic data for each study participant recorded at baseline included the median age and sex percentage (%) male (Table 2:1). In the TBM group the median age was 28 with an interquartile range [IQR] of 20 – 40 years, with 53% (majority) of study participants being males. The median age of the non-TBM group was 34 [IQR, 29-38] with majority female (66%).

The baseline clinical data shows a median plasma $CD4^+$ count of 117 cells/µl [IQR, 14 – 677] and 113 cells/µl [IQR, 36 – 532] for the TBM and non-TBM groups is respectively (Table 2:1). Whilst, the median lymphocyte count in the CSF were 6 cells/µl [IQR, 2 – 122] and 2 cells/µl [IQR, 0.24 – 6] for the TBM and non-TBM groups respectively. Comparing plasma viral loads, the TBM median viral load was 390 650 copies/ml [IQR, 289 217 – 1 710 603] which was higher than the non-TBM group with a median plasma viral load of 102 025 copies/ml [IQR, 4317 – 213 260]. Consistent with a previous report (Morris et al., 1998), the median CSF viral loads of the TBM group was significantly higher 630 291 copies/ml [IQR, 8133 – 10 000 000] compared to the non-TBM group 10 631 copies/ml [IQR, 5109 – 14 230])($p = 0.020$).

Table 2:1 Comparison of Demographic and Clinical Characteristics of study participants from TBM and Non-TBM groups

Characteristics		TBM	Non-TBM	P – value
Demographic	**No. of study participants**	**17**	**3**	N/A
	Sex % male	**53%**	**33%**	**0.006**[#]
	Median Age (IQR) years	**28** (20 – 40)	**34** (29 – 38)	0.221*
Clinical	**Median CD4 count (IQR) cells/µl in Plasma**	**117** (14 – 677)	**113** (36 – 532)	0.695*
	Median lymphocyte count (IQR) cells/µl in CSF	**6** (2 – 112)	**2** (0.24 – 6)	0.112*
	Median Plasma Viral Load (IQR) copies/ml	**390 650** (289217 – 1 710 603)	**102 025** (4317 – 213 260)	0.169*
	Median CSF Viral Load (IQR) copies/ml	**630 291** (8133 – 10 000 000)	**10 631** (5109 – 14 230)	**0.020***

N/A: Not applicable. IQR: Interquartile range. Significant differences between the groups are shown in bold P values
#Fisher's exact test
*Kruskal–Wallis test

2.2. Ethics

Ethical clearance was obtained from the Biomedical Research Ethics Committee of the University of KwaZulu-Natal (Reference: BE201/18). Patient identifiers were removed and written informed consent was obtained from the study participants. The patients remained anonymous and recorded using patient ID (PID). The study was carried out in the HIV Pathogenesis Programme (HPP) Department at the University of KwaZulu-Natal's Nelson R. Mandela School of Medicine Campus.

2.3. RNA Extraction

Viral RNA was extracted from CSF and plasma samples obtained from TBM and non-TBM patients using QIAmp viral RNA Mini kit (Qiagen Inc., Valencia, CA, USA) as per manufacturer's instructions. Briefly, 140 µl of either CSF or plasma sample was added in 560 µl of Buffer AVL (viral lysis buffer) containing carrier RNA and incubated at room

temperature for 10 minutes to ensure complete lysis. To this lysate, 560 µl of ethanol was added to precipitate nucleic acids and enhance efficiency of viral RNA binding to the silica gel membrane of the QIAamp Mini column in the subsequent step. Approximately 630 µl of the solution made in the prior step was applied directly to the silica gel membrane of the QIAamp Mini column that was inserted in a 2ml collection tube, centrifuged at 8000 rpm for 1 minute and the filtrate together with the collection tube was discarded. This step was repeated twice to make sure that all the solution containing viral RNA passed through the silica gel membrane to the QIAamp Mini column. A volume of 500 µl of wash Buffer AW1 was added to the QIAamp Mini column to wash away the debris or contaminants in the QIAamp Mini column, this was centrifuged at 8000 rpm for 1 minute, and the filtrate together with the collection tube was discarded. Washing was repeated using 500 µl of wash Buffer AW2, the QIAamp Mini column containing the sample was centrifuged at 14000 rpm for 3 minutes, the filtrate together with the collection tube was discarded. Lastly, the column containing viral RNA bound to its silica gel membrane was placed in a sterile 1.5 ml tube, 50 µl of elution Buffer AVE was added directly to the silica membrane without touching the silica membrane or wetting the rim and incubated at room temperature for 1 minute. The column was then centrifuged at 8000 rpm for 1 minute resulting in a sample of pure viral RNA being eluted into a sterile 1.5ml tube. Extracted viral RNA samples from CSF and plasma were stored at -20^{o}C until use.

2.4. cDNA Synthesis

The extracted viral RNA was reverse transcribed to complementary DNA (cDNA) using SuperScript IV Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA), as per manufacturer's instructions. Briefly, a reaction mixture to anneal the primer to the template viral RNA was prepared in a 13 µl reaction volume, made of 5 µl of either extracted CSF or plasma RNA, 0.2 mM dNTP mix, 0.1 µM gene specific primer (TatR1) and 6 µl of Diethyl pyrocarbonate (DPEC) water. The reaction mixture was incubated at 65^{o}C for 5 minutes. Subsequently a 7 µl reverse transcription mix was prepared by mixing 1X SSIV buffer, 5 mM DTT, 2 U/µl RNaseOUT recombinant RNase inhibitor and 10U/µl SuperScript IV reverse transcriptase. Next the reverse transcription mix was added to the

annealed primer. This mixture was then incubated at 50ºC for 10 minutes and the reaction terminated at 85ºC for 5 min. Extracted reverse transcribed samples from CSF and plasma was stored at -20ºC until use.

2.5. Polymerase chain reaction (PCR)

The 216 nucleotides of *tat* exon 1 were amplified through two rounds of polymerase chain reaction (PCR), known as nested PCR using the Platinum ® Taq DNA Polymerase High Fidelity PCR kit (Thermo Fisher Scientific, Boston, MA, USA). The first round of PCR was carried out in a 25µl reaction volume containing 1X High Fidelity PCR Buffer, 2.0 mM $MgSO_4$, 0.2 µM dNTP mix, 0.2 µM forward primer (Tat F1 5'-CATTTCAGAATTGGG-3'), 0.2 µM reverse primer (Tat R1 5′-TTGCTATTATTATTGCTAC-3′)(Cowley et al., 2011) and 1 U/rxn Platinum® Taq DNA Polymerase High Fidelity enzyme, to this solution a 2.5 µl cDNA template was added and DPEC water was used to make up a final reaction volume of 25 µl. The second round of PCR (Nested PCR) was performed in a 25 µl reaction using 0.2 µM forward TatF3 (5′-AGACGACAGAGAAGAGCAAG-3′) and reverse TatR3 (5′-CATTACTATTTTGTACTTACTGC -3′) primers (Cowley et al., 2011). To this solution a 2 µl volume of DNA from the first round of PCR was added and DPEC water was used to make up a final reaction volume of 25 µl. Both rounds of PCR were carried out under the following thermalcycler conditions: denaturation at 94°C for 2 minutes; 30 cycles at 94°C for 15 seconds; 50°C for 30 seconds and 68°C for 1 minute; and a final extension step at 68 °C for 5 minutes. The presence of PCR products was verified on a 2% agarose gel and stained with GelRed and visualized under ultraviolet light.

Subsequently PCR products with a band size of approximately 255 bp were subjected to a PCR clean up using QIAquick PCR Purification Kit (Qiagen Inc., Valencia, CA, USA) as per manufactures instructions. Briefly, 100 µl of Buffer PB was added to 20 µl of remaining PCR product, the mixture was transferred into a QIAquick spin column collection tube and centrifuged for 1 minute at 13 000 rpm to allow for efficient binding of the DNA to the membrane. Afterwards, the filtrate in the collection tube was discarded and 750 µl of Wash Buffer PE was added to the spin column. The sample was then centrifuged for 1 minute at 13 000 rpm to remove any remaining contaminants. The DNA was eluted from

the spin column into a sterile 1.5ml tube using 10mM EB buffer (pH 8.5). The yield of the purified DNA was quantified using the NanoDrop™ 2000 spectrophotometer (Thermo Scientific, Delaware, USA) and stored at -20 °C.

2.6. DNA Sequencing and Phylogenetic Analysis

Nucleotide sequences were obtained by direct sequencing of the HIV-1 purified *tat* amplicons. The DNA sequencing was performed in a 96 well plate (Applied Biosystems) using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, Boston, MA, USA) as per manufacturer's instructions. Briefly a 10 µl sequencing reaction was separately prepared for each nested PCR primer (TatF3 and TatR3). Each sequencing reaction contained 3.2 µM primer, 1X Sequencing Buffer, 0.4 µl of Big Dye Terminator v3.1 Ready Reaction Mix, 20 ng of purified PCR products and DPEC water to make a final volume of 10 µl. The sequencing reactions were then subjected to an initial denaturation at 96°C for 1 minute; followed by 25 cycles of denaturation at 96°C for 10 seconds, annealing at 50°C for 10 seconds, and extension at 60°C for 4 minutes.

Following the cycle-sequencing reaction, sequences underwent a sequencing reaction clean up. Briefly 1ul of 125 mM ethylenediaminetetraacetic acid (EDTA) (pH 8.0) was added into each well, followed by the addition of a 26 µl solution made up of; 1ul of sodium acetate (NaOAc) (3M, pH5.2), and 25 µl of ice cold 100% ethanol (previously stored at -20^{o}C) to precipitate DNA sequences and this was followed by centrifugation at 4000 rpm for 20 minutes (Eppendorf centrifuge 5810R, Merck, Germany). The supernatants were removed by inverting the plate and centrifuging at 880 rpm for 1 minute which was followed by the addition of freshly made ice cold 70% ethanol to each well. The plate was then centrifuged again at 4000 rpm for 5 minutes and inverted 880 rpm for 1 minute to remove the supernatant. Lastly, the sequencing reactions were then dried at 50°C for 5 minutes and stored at -20^{o}C until analyzed using the ABI 3130xl Genetic Analyzer (Applied Biosystems). To run plates on the ABI 3130xl Genetic Analyzer, sequencing products were re-suspended in 10µl of highly deionized formamide (HiDi-formamide),

mixed and denatured at 95ºC for 3 minutes and cooled to 4ºC for 3 minutes in a thermalcycler.

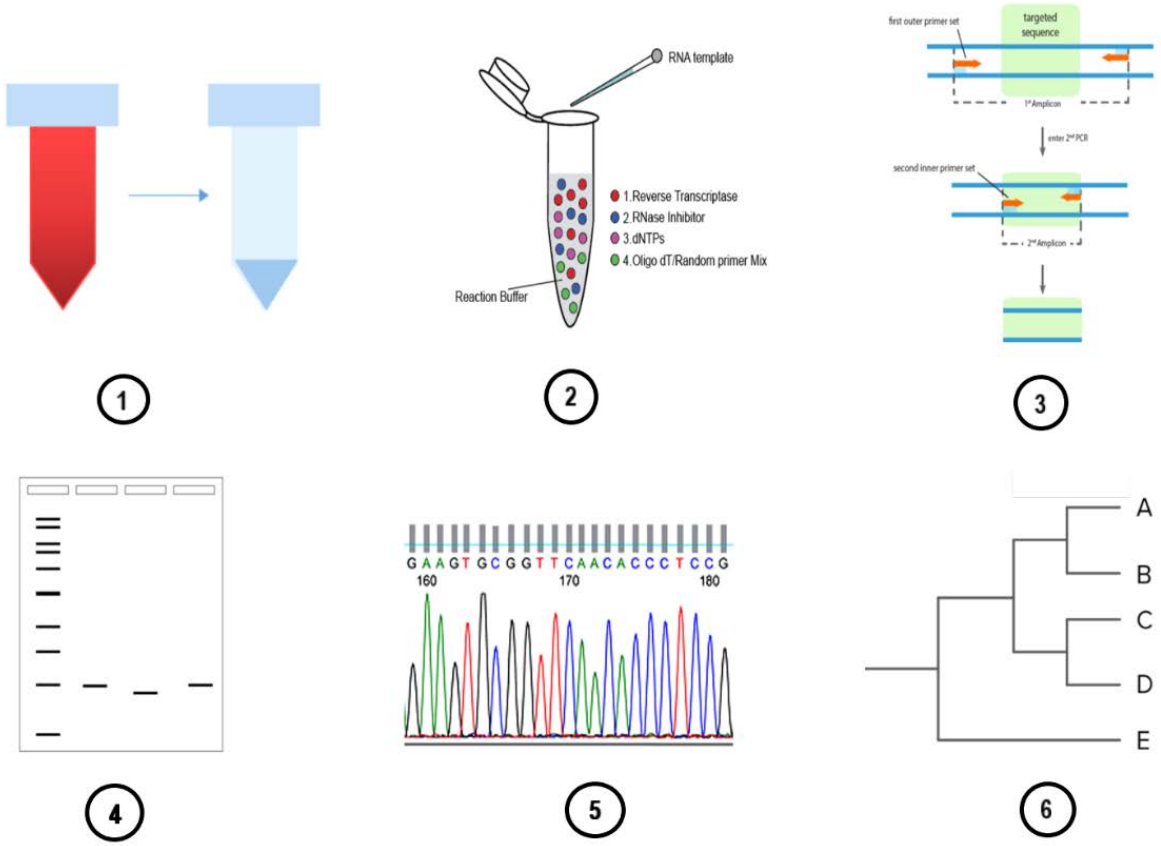


Figure 2:2 Overview of molecular methods use to amplify and sequence HIV-1 *tat*. (1) Viral RNA was extracted from all plasma and CSF samples. (2) Extracted RNA was reverse transcribed into cDNA. (3) All cDNA products were amplified using two rounds (nested) of PCR. (4) Amplified DNA was run on a gel to confirm size. (5) All *tat* amplicons were subjected to Sanger sequencing. (5) Sanger sequences were analysed, and a phylogenetic tree was drawn.

The *tat* sequences were edited and analyzed using the Sequencher 5.1 software (Gene Codes Corporation, Ann Arbor, USA). Sequences were then aligned to a *tat* subtype C consensus sequence (GenBank accession No. EU293450.1) from the Los Alamos HIV Sequence Database (https://www.hiv.lanl.gov/) using the Clustal W (Thompson et al., 1994) multiple-sequence alignment program. Subsequently, a neighbor-joining phylogenetic tree (Saitou and Nei, 1987) was then generated using the Geneious

software v10.2.3 (Biomatters, Ltd., Auckland, New Zealand). A diversity analysis based on divergence was done using the codon-aligned nucleotide sequences with MEGA X (Kumar et al., 2018). The model of substitution used was nucleotide and the method used was p-distance. The codon-aligned nucleotide sequences of *tat* exon-1 variants were then translated into amino acid sequences using the online software Gene Cutter (www.hiv.lanl.gov) and visualized using BioEdit v.7.0.9. 8 (Hall, 1999) for further analysis. A Tat amino acid sequence logo was generated by using Weblogo 3.1 (http://weblogo.threeplusone.com). Subsequently a Viral Epidemiologic Signature Pattern Analysis (VESPA) (www.hiv.lanl.gov) was carried out to determine the presence of signature amino acid substitutions.

2.7. Cloning into pTargeT™ Mammalian Expression Vector System

Out of the 38 PCR products generated (32 TBM and 6 non-TBM), twelve sequences were chosen for downstream functional assays. Hence one representative sequence was chosen from 10 patients (except X205 where both plasma and CSF derived Tat were selected). Therefore, a total of 12 samples with key mutations reported to affect Tat activity were selected for cloning. These 12 PCR products were cloned into a T-tailed vector pTargeT™ expression plasmid (Promega Corporation, Madison, WI) according to the manufacturer's instructions. In the pTargeT™ expression vector, the HIV-1 *tat* exon 1 target sequence is under the control of T7 and CMV promoter.

In the ligation reaction, up to 2 µl of the nested PCR product was added into the pTargeT vector in a standard reaction volume of 10 µl. For the positive control, 2 µl of control DNA was ligated into pTargeT vector, whilst the background control contained an empty pTargeT vector with no insert. Ligation reaction tubes were incubated overnight at 4^{o}C. Transformation of each sample was performed by adding 2 µl of each ligation reaction into a 1.5ml tube containing 50 µl of JM109 high efficiency competent cells (Promega Corporation, Madison, WI) for each sample, whilst the control tubes contained 100 µl of JM109 cells. Tubes were subjected to heat shock for 50 seconds at 42^{o}C in a water bath. A total volume of 950 µl and 900 µl of SOC medium was added into the sample tube and

control tube respectively to reach a final volume of 1ml. Tubes were then incubated for 1 hour and 30 minutes at 37^{o}C with shaking at 150 rpm. After which 100 µl of each transformation reaction were plated on ampicillin agar plates and incubated overnight at 37^{o}C. Blue/white screening was used for confirmation of *tat* clones using X-gal (Thermo Fisher Scientific, Boston, MA, USA).

Colony PCR was employed for screening clones generated. Briefly, colonies grown on ampicillin plates were picked and boiled at 95^{o}C for 10 minutes. Afterwards each clone was amplified using nested PCR as previously described. The presence of PCR products was verified on a 2% agarose gel and stained with GelRed and visualized under ultraviolet light. All positive clones with a band size of approximately 255bp, were inoculated into a 3ml of Lysogeny broth (LB), containing 3 µl of ampicillin. The culture was then grown at 37^{o}C in a 230rpm shaking incubator (Infors HT, Bottmingen, Switzerland) for 16 hours. Purified plasmid DNA were extracted using the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, Boston, MA, USA). All purified plasmid DNA were subjected to a restriction digest using EcoRI (New England Biolabs, Hitchin, UK) at 37^{o}C for 1 hour. The confirmation of each *tat* insert (approximately 255bp) that was successfully cut from plasmid (approximately 5670bp) was verified on a 1% agarose gel. All plasmid DNA with a confirmed *tat* insert after restriction were sequenced using Sanger sequencing as previously described. This was followed by sequence analysis to determine if the clonal sequence matched the bulk sequence previously amplified.

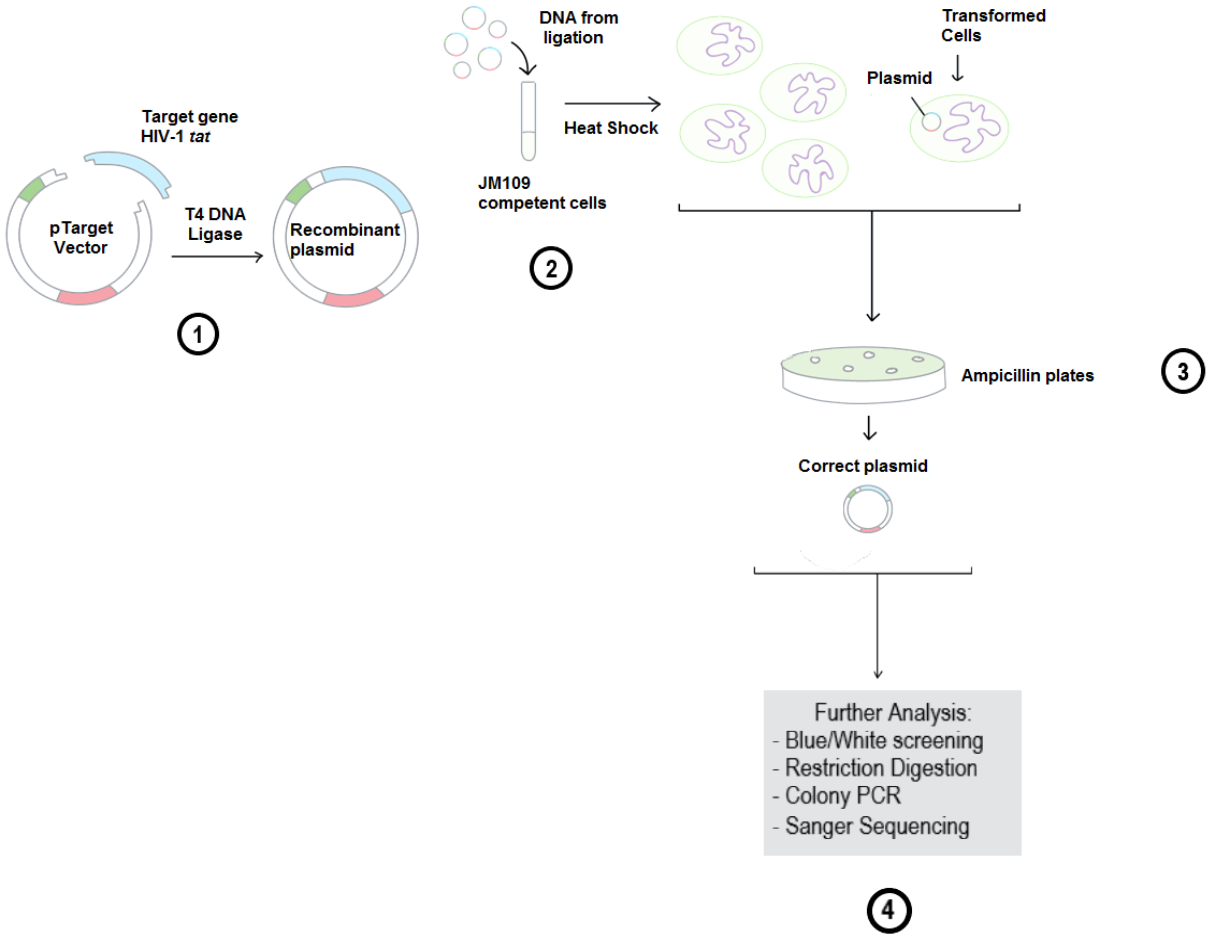


Figure 2:3 Summary outline of cloning procedure. (1) PCR purified *tat* amplicons with interesting mutations were ligated into a pTargeT vector with T4 DNA ligase, creating a recombinant plasmid. (2) Recombinant plasmids were heat shocked into JM109 competent cells. (3) Transformed JM109 cells were grown on ampicillin agar. (4) The correct plasmid was chosen via colony screening methods such as, blue/white screening, restriction digestion, colony PCR and Sanger sequencing to confirm the insert of *tat*. (Source: Adapted from http://www.khanacademy.org).

Once the insert containing specific mutations were confirmed, plasmid DNA was propagated in 100ml of LB broth containing 100µg/ml ampicillin and incubated overnight at 37^{o}C in a shaking incubator (Infors HT, Bottmingen, Switzerland). Plasmid DNA was then purified using the QIAfilter Plasmid Maxi Kit (Qiagen Inc., Valencia, CA, USA) and concentrations of approximately 1.5µg were quantified using a NanoDrop spectrophotometer (Thermo Scientific, Delaware, USA) and stored in a -80°C ultrafreezer until use.

2.8. Tat Transactivation Assay

To assess the functional diversity of selected HIV-1 *tat* variants, a Tat Transactivation Assay was performed. This assay is divided into 3 major parts: culturing the cells, transfecting pTargeT *tat* variants into TZM-bl cells and measuring the luminescence of Tat variants with a luciferase assay. The TZM-bl cells were chosen as host cells for transfection because the cell line contains an integrated reporter gene for firefly luciferase (*luc*) driven by the LTR promoter. Therefore, the cells had grown relative to the attenuation of the viral *tat* and luciferase activity was directly proportional to transactivation activity of the Tat mutant.

2.8.1 Maintenance and Preparation of TZM-bl cells

As previously mentioned the transactivation assay was carried out using TZM-bl cells (also called JC53BL-13) obtained from the NIH AIDS Research and Reference Reagent Program (Cat. No. 8129). The TZM-bl cells are adherent HeLa cell derivatives that express reporter genes; galactosidase (*LacZ*) and firefly luciferase (*luc*) under the control of a subtype B HIV-1 LTR (Sarzotti-Kelsoe et al., 2014). The cells were cultured in a growth medium (GM) consisting of Dulbecco modified Eagle medium (DMEM) (Gibco BRL, Life Technologies, Cergy Pontoise, France) supplemented with 50ml of 10% heat inactivated fetal bovine serum (FBS) (Gibco BRL, Life Technologies, Cergy Pontoise, France), 125 ug/ml Gentamicin (Sigma–Aldrich, St. Louis, MO, USA) and 10 mM of HEPES (Gibco BRL, Life Technologies, Cergy Pontoise, France).

To plate the cells, a 1ml vial of cells previously stored in liquid nitrogen freezer were thawed by swirling the vial in a 37^{o}C water bath for 5 minutes. Thereafter the thawed cells were transferred into a 50 ml conical tube supplemented with 30 ml of pre-warmed GM. The 50ml tube containing the cells and GM were centrifuged for 10 minutes at 1000 rpm (Eppendorf centrifuge 5810R, Merck, Germany). The supernatant was discarded, and cells were suspended in 15 ml of fresh GM, transferred into a T-75 culture flask and incubated for 48 hours at 37^{o}C. After 48 hours of incubation, culture flask containing cells

were then viewed on the Zoe® fluorescent imager (Bio-Rad) to assess for a 90-100% confluency. Thereafter cells were split by decanting the supernatant and rinsing the monolayers with 5 ml of sterile Phosphate-Buffered Saline (PBS) (pH7.2) (GIBCO Laboratories, Grand Island, NY). Subsequently cells were trypsonized with 3 ml of 0.25% Trypsin-EDTA (GIBCO BRL, Life Technologies, Cergy Pontoise, France). This was followed by incubation at room temperature for 30 seconds. After which the trypsin solution was removed, and cells were further incubated at 37^{o}C in 5% CO_2 for 4 minutes. The flasks were then suspended with 10 ml of GM and seeded in new T-75 culture flasks with approximately 1 x 10^6 cells in 10 ml GM. The flasks were then incubated at 37^{o}C in a 5% CO_2 and 95% air environment.

To determine the number of cells per ml of GM, flasks containing cells that were 90-100% confluent were split as previously mentioned and 10 µl of cells were thoroughly mixed with 10 µl of Tryphan blue (BIO-RAD Laboratories, Hercules, CA) and counted using the TC20™ Automated Cell Counter (BIO-RAD Laboratories, Hercules, CA).

A 24 well plate was used to plate 4 x 10^4 cells into each well (supplemented with 500 µl of GM) that would be required for transfection. Thereafter T-75 flasks with remaining cells were maintained daily until a passage number of 15 was reached. After which time a new viral of stored TZM-bl cells were used.

2.8.1 Transfection using Lipofectamine 2000

A 24 well plate containing approximately 4 x 10^4 cells per well at a 70 – 80% confluency were prepared as previously explained. Briefly, two 2-ml Eppendorf tubes were prepared per each well, 1 tube to contain 1.5 µg of plasmid DNA and 250 µl of OPTI-MEM I reduced-serum medium (Gibco-BRL, Paisley, UK) and the other with 10 µl of Lipofectamine 2000 (Invitrogen, CA, USA) and 250 µl of OPTI-MEM I reduced-serum medium (Gibco-BRL, Paisley, UK). After 5 minutes of incubation at room temperature both tubes were combined to form a Lipofectamine 2000 transfection complex, followed by incubation for 25 minutes at room temperature. During this interval, the 24-well plate containing cells were prepared by removal of old GM and each well was replenished with

1.5ml of fresh GM. After the 25 minute incubation the Lipofectamine 2000 transfection complex was then added to each corresponding well and rocked gently back and forth to distribute the Lipofectamine. Each sample was done in triplicate. A patient derived Tat sample (PID A219) that was highly similar to the Tat subtype C consensus sequence (96%) was used as the consensus C control, whilst a pTargeT vector without any *tat* insert was used as a negative control. Both controls were similarly transfected into TZM-bl cells. Each plate was then incubated at 37^{o}C in a 5% CO_2 and 95% air environment for 24 hours. Following a 24 hour incubation, 1.5 ml of supernatant was removed from each well and replenished with 2 ml of pre-warmed GM and incubated at 37^{o}C in a 5% CO_2 and 95% air environment for an additional 24 hours.

2.8.2 Luciferase Assay using Bright-Glo

After 24 hours, approximately 1.9 ml of supernatant was removed from each well leaving a 100 µl layer of supernatant to cover the cells in each well. The cells were lysed with 100 µl of Bright-Glo™ Luciferase Assay System (Promega, Madison, WI, USA) and were incubated at room temperature for 2 min, in the dark. Following incubation, the contents of the pate were mixed twice by pipetting, thereafter 150 µl of the lysate was transferred to corresponding wells of a 96 well black solid bottom microplate (Greiner Bio-One, Monroe, NC).

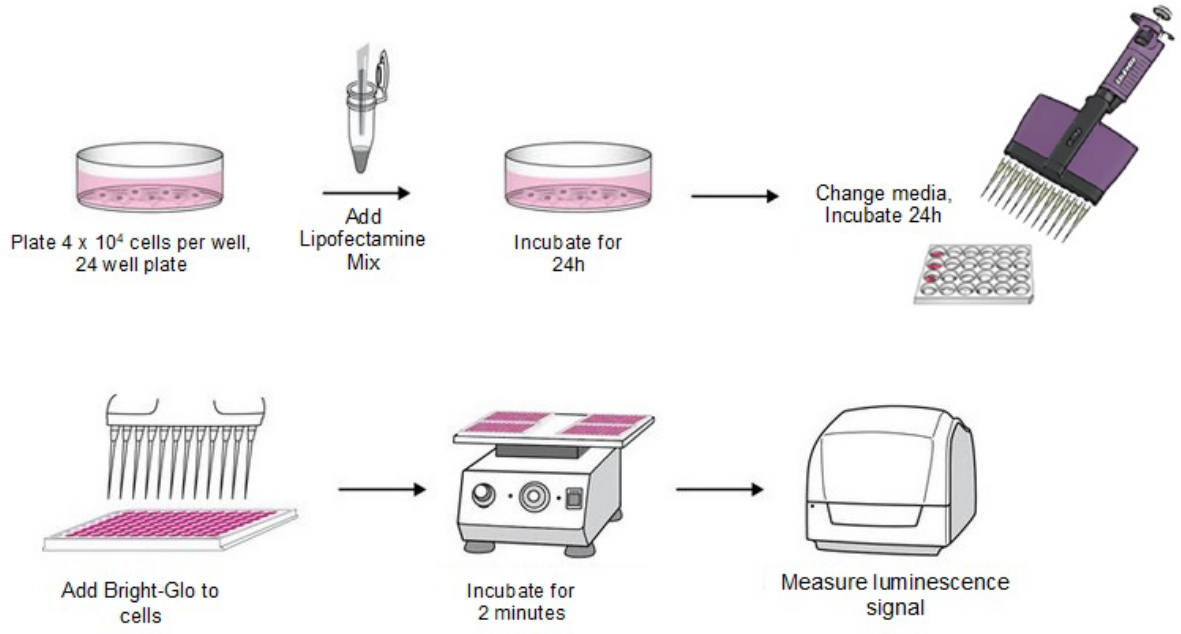


Figure 2:4 Summary outline of Tat transactivation assay procedures. (Source: Adapted from https://www.thermofisher.com)

The luciferase activity was determined using a VICTOR Nivo Multimode Microplate Reader (Perkin-Elmer Life Sciences, Shelton, CT) and the output data was expressed as relative light units (RLUs). Firefly luciferase values measured in relative luminescence units (RLU) were normalized against a cell control containing TZM-bl cells only. The relative fold change for each sample were calculated relative to the consensus. A Student's *t* test was performed on each patient derived Tat against the consensus C Tat. The Pearson correlation coefficient and linear regression was used to measure the correlation between Tat transactivation and patients' viral load and lymphocyte counts.

Chapter 3

3. Results

3.1. Phylogenetic Analysis Reveals High Genetic Diversity of Tat Exon-1

Initially a total of 20 PID's (n=40 [20 plasma and 20 CSF] samples) were chosen, however we were unable to amplify HIV-1 *tat* from the CSF and plasma of 1 TBM patient. Therefore, a phylogenetic tree was constructed using 38 HIV-1 *tat* sequences obtained from 19 patients (16 TBM and 3 non-TBM). For each patient ID (PID), we used one *tat* sequence composed of 216 nucleotides from each compartment. The tree was rooted to a subtype c *tat* consensus (GenBank accession No. EU293450.1).

Our data demonstrated that matched plasma and CSF *tat* sequences from each patient clustered together. This indicates they were isolated from the same patient, more so there were no cross-contamination between PID's (Figure 3:1). Intermingling between paired clusters of the TBM and non-TBM patients were observed and no distinct compartmentalization was seen within these two groups, which could be due to a small sample size of the non-TBM group. Interestingly PID A129 (non-TBM) paired closely with the subtype C *tat* consensus, suggesting that viruses from this PID were closely related to the consensus sequence.

Although three major clusters were formed, the non-TBM patients all grouped within one cluster whilst the TBM were dispersed. Within each cluster, each subpopulation showed variation in the *tat* quasispecies (Figure 3:1). Distinct differences in branch lengths between CSF and plasma sequences of PID's: A96, A105, X215, X205, B98 (TBM) and X235 (non-TBM) were observed, suggesting diversity differences within the two compartments. Conversely, some patients had little or almost no differences between their respective plasma and CSF sequences, including X201, A104, X195, X223 (TBM), A129 and X229 (non-TBM), suggesting that the same viral strains were dominant in the two compartments.

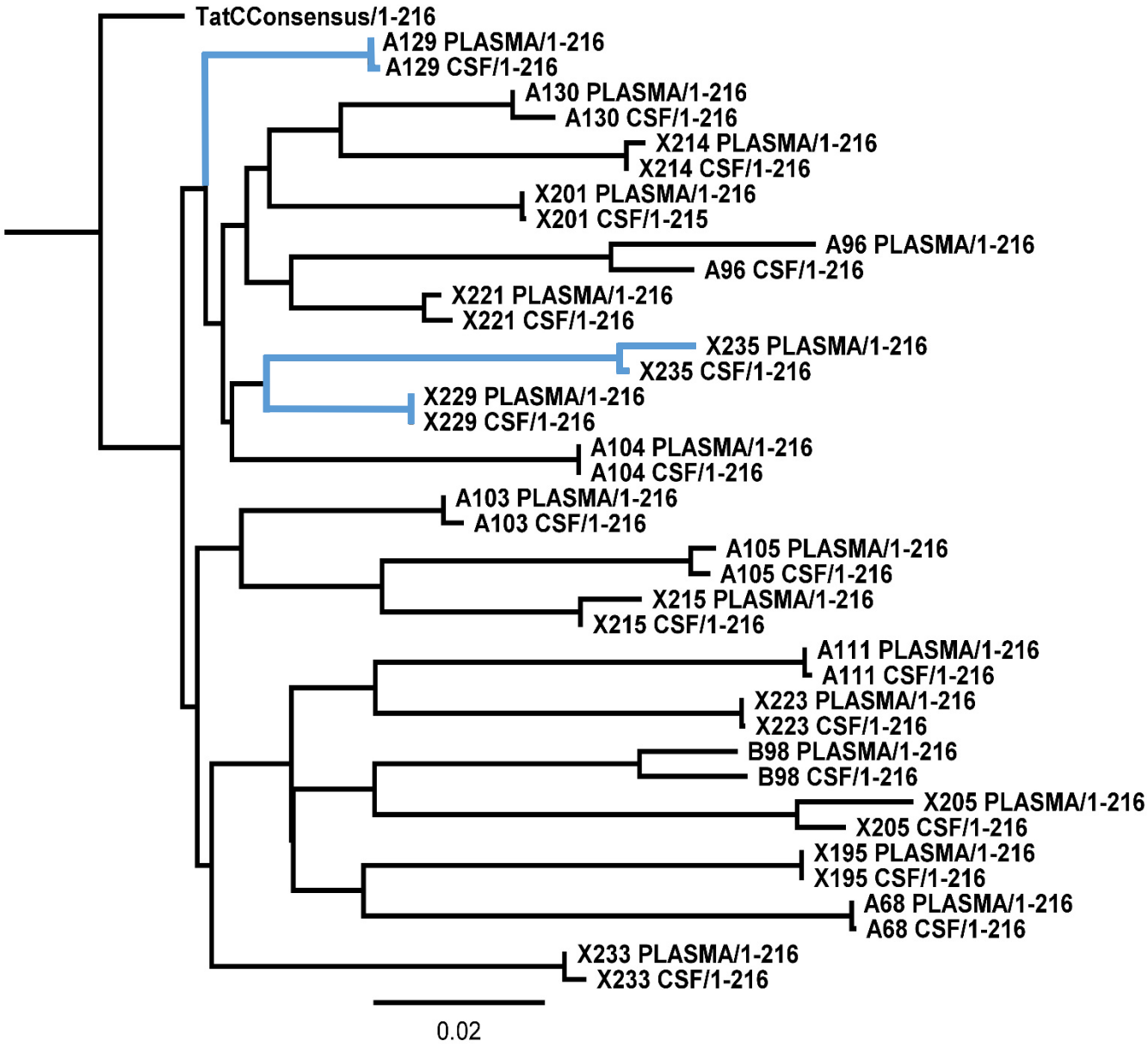


Figure 3:1 Neighbor joining phylogenetic tree representing 19 matched plasma and CSF derived *tat* sequences. All sequences in the tree were 216 nucleotides in length. The tree was rooted to a subtype C *tat* consensus sequence (GenBank accession No. EU293450.1), and each plasma sequence paired with its respective CSF sequence from the same patient. All TBM patients were marked in black whilst all non-TBM patients marked in blue. There was no cross contamination between sequences. All sequences were related to the subtype C consensus and two clusters formed. A scale of 0.02 represents 20% nucleotide differences between sequences.

3.2. Genetic diversity analysis

In order to assess *tat* nucleotide diversity within compartments of TBM and non-TBM patients as observed in Figure 3:1, nucleotide alignments were generated and used to calculate both interpatient and intrapatient genetic diversity analyses. This would allow for a more accurate measure of the displayed relationships in the phylogenetic tree. Average pairwise distance (APD) is a statistic tool commonly used in published methods of phylogenetic analysis. P-distance is a measure of APD and defined as the number of nucleotide differences between two sequences/total nucleotide sequences (s/nt). In most HIV-1 diversity studies, the p-distance determinations are relatively small (<0.03) (Kumar and Nei, 2000).

3.2.1 Interpatient P-distance Analysis

The interpatient p-distance analysis calculated the nucleotide divergence, as a measure of the genetic distance between plasma and CSF sequences in each PID. In this analysis we hypothesized that TBM would drive a high viral diversity. To perform this analysis, aligned *tat* sequences were separated into two groups (TBM and non-TBM) using MEGA X. The average mean distance (s/nt) for TBM and non-TBM patients were 0.003 and 0.002 respectively. As seen in Figure 3:2, the TBM group had a greater amount of diversity compared to the non-TBM group although this was not significant. The highest diversity was seen in PID X205 and A96 when comparing their CSF sequence to their respective plasma sequence. Furthermore A105, X215 and B98 were all represented with the same p-distance and this was comparable to X235 of the non-TBM group.

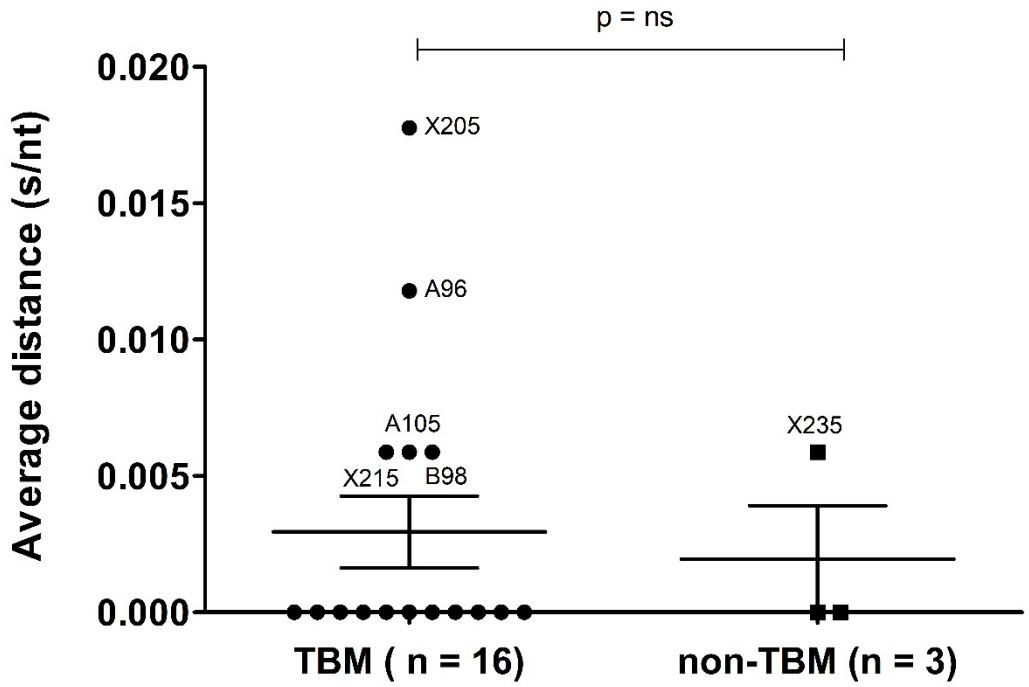


Figure 3:2: Interpatient diversity analysis of TBM and non-TBM study participants. The above plot shows the mean and SEM of both the TBM and non-TBM groups. Each dot represents the distance whereby the plasma and CSF sequence of each PID being compared are different. It is obtained by dividing the number of nucleotide differences by the total number of nucleotides compared.

Whilst majority of the participants had a p-distance of 0.00 this does not regard them as having no diversity – rather there is no difference in the p-distance between their plasma and CSF sequence. Overall there was no significant differences in *tat* diversity between TBM and non-TBM participants.

3.2.2 Intrapatient P-distance Analysis

An intrapatient p-distance analysis was conducted to assess the difference between the plasma and CSF *tat* sequences within TBM coinfected patients. We hypothesized that the higher viral load in the CSF compartment would result in a higher CSF *tat* diversity compared to plasma. Briefly, aligned *tat* sequences were put into groups (plasma and CSF) and compared to a Tat consensus C sequence (GenBank accession No.

EU293450.1) using MEGA X. This comparison would indicate; how much *tat* has evolved from the transmitter consensus like strain and whether the CSF compartment has a higher viral *tat* diversity than the plasma compartment. Whilst the mean p-distance seen in the CSF compartment is slightly higher than the plasma (p = ns) (Figure 3:3), the calculated overall average mean distance (s/nt) generated for plasma and CSF from TBM patients had a negligible difference of 0.059 and 0.062 respectively.

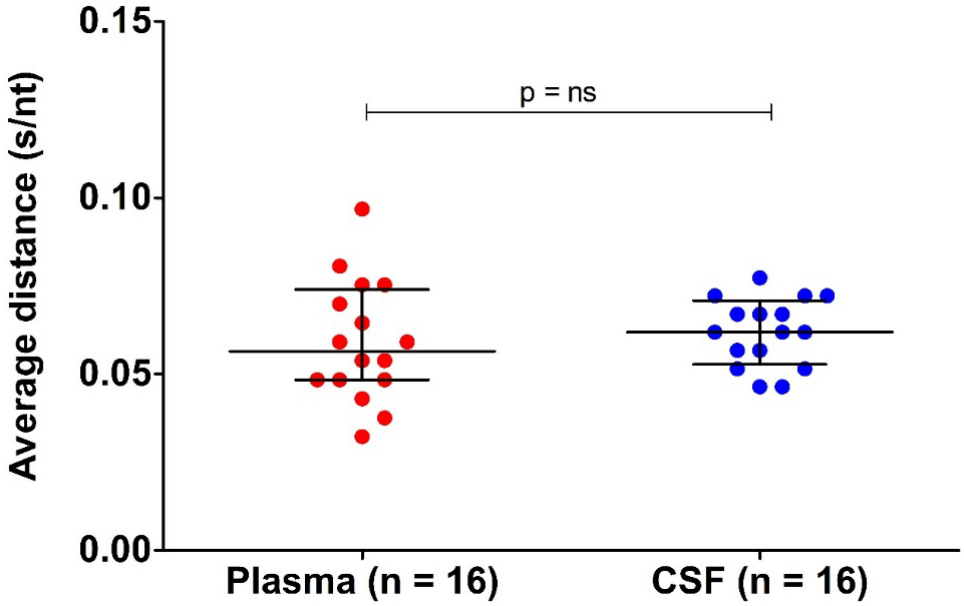


Figure 3:3: Intrapatient diversity analysis of plasma and CSF derived *tat* sequences from TBM study participants. The above plot shows the mean and SEM of both the plasma sequences (red) and CSF sequences (blue). Each dot represents the distance whereby the plasma and CSF sequences are different to a consensus C sequence. It is obtained by dividing the number of nucleotide differences by the total number of nucleotides compared.

Overall there was no significant differences in diversity between plasma and CSF derived *tat* sequences in patients with TBM. Since TBM viruses are the focus of this study, all further analyses will be based on their respective plasma and CSF HIV-1 sequences.

3.3. Amino acid analysis of Tat exon-1 variants

A total of 38 HIV-1C *tat* sequences (obtained from 19 CSF and 19 plasma samples) were translated to amino acids and aligned for analysis. As previously mentioned the HIV-1C Tat protein structure is divided into five domains encoded by the first exon. Domain I is composed of amino acids (aa): 1–21 also known as the acidic domain. Domain II, aa: 21–37 known as the cysteine-rich domain. Domain III (aa: 38–48) is the core domain, Domain IV (aa: 49–57; basic domain) and Domain V (aa: 58–72) is the glutamine domain. Amino acids 1–48 are efficient for Tat transactivation. Whilst aa: 49–57 is the RNA binding domain (RBD) (Figure 1:4).

3.3.1 Conservation of Tat amino acid sequences

The Tat amino acid logo generated, graphically represents the patterns within a multiple sequence alignment (Figure 3:4). The height of each amino acid indicates frequency of the amino acid at that position (Crooks et al., 2004).

The acidic domain of Tat was proline rich and relatively well conserved, with two major variable positions 7 and 19 (Figure 3:4). In comparison the cysteine region of Tat seemed highly diverse with variable positions 23, 24 and 29. Consistent with the previous studies the 6 cysteine residues reported in HIV-1C remained well conserved (Rao et al., 2013). Additionally, we identified the previously reported di-cysteine motif C30C31 in our cohort (Figure 3:4), however S31 was more prevalent as commonly seen in subtype C variants (Ranga et al., 2004, Wong et al., 2010, Campbell et al., 2011).

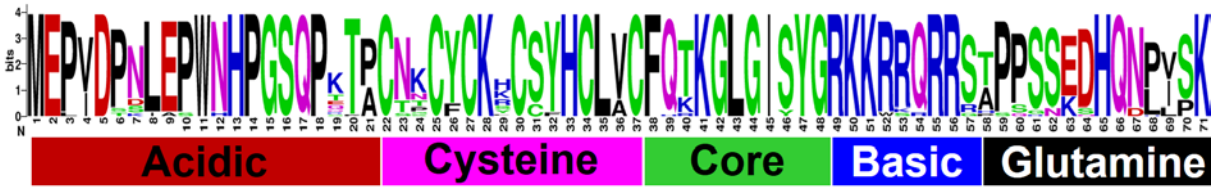


Figure 3:4 : Tat amino acid sequence logo. All 38 Tat sequences were aligned to create a Weblogo. The x-axis shows the position of each amino acid, y-axis shows the frequency of amino acids measured in bits. The five regions of *tat* exon 1 are highlighted: An N-terminal acidic domain comprising the first 21 amino acids of Tat, with a highly conserved Tryptophan 11 (W11), a cysteine-rich domain with six well conserved cysteine amino acids at positions in subtype C, the core domain has a conserved Phenylalanine (F38) and amino acid sequence LGISYG. The basic domain made of amino acids 49–57 and has a well conserved RKKRRQRR basic amino acid sequence. Lastly the glutamine-rich domain with high sequence variation.

The most important functional domains in Tat are the core and basic domains, with conserved motifs; LGISYG (43 – 48) and RKKRRQRR (49 – 56). The core domain motif remained 83% conserved with only a S46Y substitution in 3 samples (Figure 3:4). Whilst the basic domain motif was only 63% conserved and many low frequency mutations including R52W, R53S, R53K and Q54R were observed (Figure 3:4). The extent of this variability seen in Tat from HIV-1 chronically infected patients who have TBM will be further analyzed. Lastly the glutamine domain had the most sequence variation (80%), with many non-conservative mutations (Campbell et al., 2004).

A subtype C amino acid consensus was created from the Tat amino acid logo generated and used a means of comparison to identify patient derived Tat mutations (Figure 3:4). It should be noted that our consensus contained a 21A, whilst the wildtype amino acid in Tat variants is a 21P (Rossenkhan et al., 2013).

3.3.2 Mutations in Tat

A multiple sequence alignment was performed to determine the presence of specific residues that distinguish CNS and plasma Tat proteins (Figure 3:5). Whilst no major difference between CSF and plasma derived Tat were observed, the presence of amino acid substitutions within domains previously reported to affect Tat function, were observed in both compartments and assessed.

Mutations previously described to significantly affect LTR-mediated transcriptional activity were observed in all domains of Tat. A list of 10 Tat mutations collated from existing literature (Table 3:1) was used as a reference to identify mutations in Tat that have been associated with an increase or decrease in LTR-mediated transcriptional activity. Eight out of 10 Tat sequences had mutations known to be associated with a detrimental Tat activity were detected (Table 3:1 and Figure 3:5). These included: V4I, P21A, K24S, S31C, S46Y, R52W, P59S and D64N which will be subsequently discussed. The remaining 2 of the 10 Tat sequences had mutations: H29R and S57R, which were reported to be associated with enhancing Tat activity.

Mutations in TBM patients with a detrimental impact on LTR activation (Figure 3:5):

(i) Within the acidic domain, 48% of amino acids were highly conserved, whilst 52% had sequence variation. Although there were contradictory reports regarding the importance of the acidic domain, with one study suggesting that high sequence variation within this domain does not impact Tat activity (Li et al., 2012). Whilst another showing the mutations V4I and P21A significantly altered the activity of Tat (Rossenkhan et al., 2013). Our data demonstrate that mutations V4I as seen in 5 of 16 TBM patients and P21A as seen in 7 of 16 TBM patients have been associated with a lower Tat transactivation activity, thus confirming the latter study (Rossenkhan et al., 2013). Furthermore, the acidic domain is a proline-rich domain and many low frequency mutations at proline residues namely; P2L, P6T and P10S were observed in this cohort. However, mutations in proline residues other

than P21A affecting Tat function is still unknown. Therefore, the data from this study warrants the further characterization of mutations: P2L, P6T and P10S.

(ii) The cysteine domain had 56% of sequence variation in this cohort. Although high sequence variation in this domain is uncommon, all 6 cysteine residues linked to Tat function remained highly conserved. In addition the C31S and dicystine motif C30C31 was seen in 84% and 16% respectively – although the C31S mutation is most frequently observed in subtype C (Rao et al., 2013). In addition, the functional capability of C31S versus C31C is still unclear. Position 24 was one of the most highly diverse amino acid positions and mutation K24S, previously reported with altered Tat activity was observed in PID A105 (Figure 3:4) (Rossenkhan et al., 2013).

(iii) Whilst the core domain of Tat are reportedly highly conserved (Jeang, 1996), 27% diversity was observed in this region (Figure 3:4). At position 46, 13% of plasma sequences and 6% of CSF sequences had the S46Y substitution. Position 46S is a site of phosphorylation and amino acid substitutions at this site is associated with reduced transactivation (Cowley et al., 2011). In addition 19% of patients had a T40K mutation, which was identified as a signature residue for Indian HIV-1C (Neogi et al., 2012).

(iv) The basic domain is responsible for Tat binding to the TAR element. While this region is reportedly well conserved – 44% of sequence variation was observed in our cohort. Although mutations that abolish the Tat-TAR interaction have not yet been reported. An R52W mutation was observed in PID 68 and this mutation was previously associated with high attenuation of the viral transactivation activity (Sivakumaran et al., 2007). In addition, unreported mutations including R53S could allow for a greater attenuation of Tat activity and needs to be further studied. The mutations in this domain warrants bind affinity assays such as electrophoretic mobility shift assay (EMSA) in order to determine their impact on Tat binding to the TAR element.

(v) Lastly, the glutamine domain had fairly high sequence variation with only 3 positions being fully conserved. This is consistent with previous reports suggesting that the glutamine domain has the greatest degree of variability within Tat (Li et al., 2012). With such high sequence variation, it is unclear which residues correlate to impaired Tat activity. However the P59S mutation observed in PID X233 was previously associated to decrease in LTR activity (Rossenkhan et al., 2013).

Mutations in TBM patients with an enhancing effect on LTR activation (Figure 3:5);

(i) Position 29 is highly variable within the cysteine domain. Mutation frequencies were as follows: H29K (31%), H29R (19%), H29Y (13%), H29S (13%), H29C (6%), and H29F (6%). Of all the amino acid substitutions seen in this position, only the H29R mutation has been reported to significantly upregulate LTR-luciferase expression (Rossenkhan et al., 2013).Therefore, the other mutations at this position necessitate the characterization of their impact on Tat activity.

(ii) Position 57 in the basic domain had two low frequency mutations. Which included S57G and S57R which observed in 6% and 13% of matched plasma and CSF sequences respectively. Whilst the functional activity of the S57G mutation is not known, the S57R has been previously reported to enhance LTR transactivation activity (Endo-Munoz et al., 2005). These data suggest the need to characterize the function of S57G mutation.

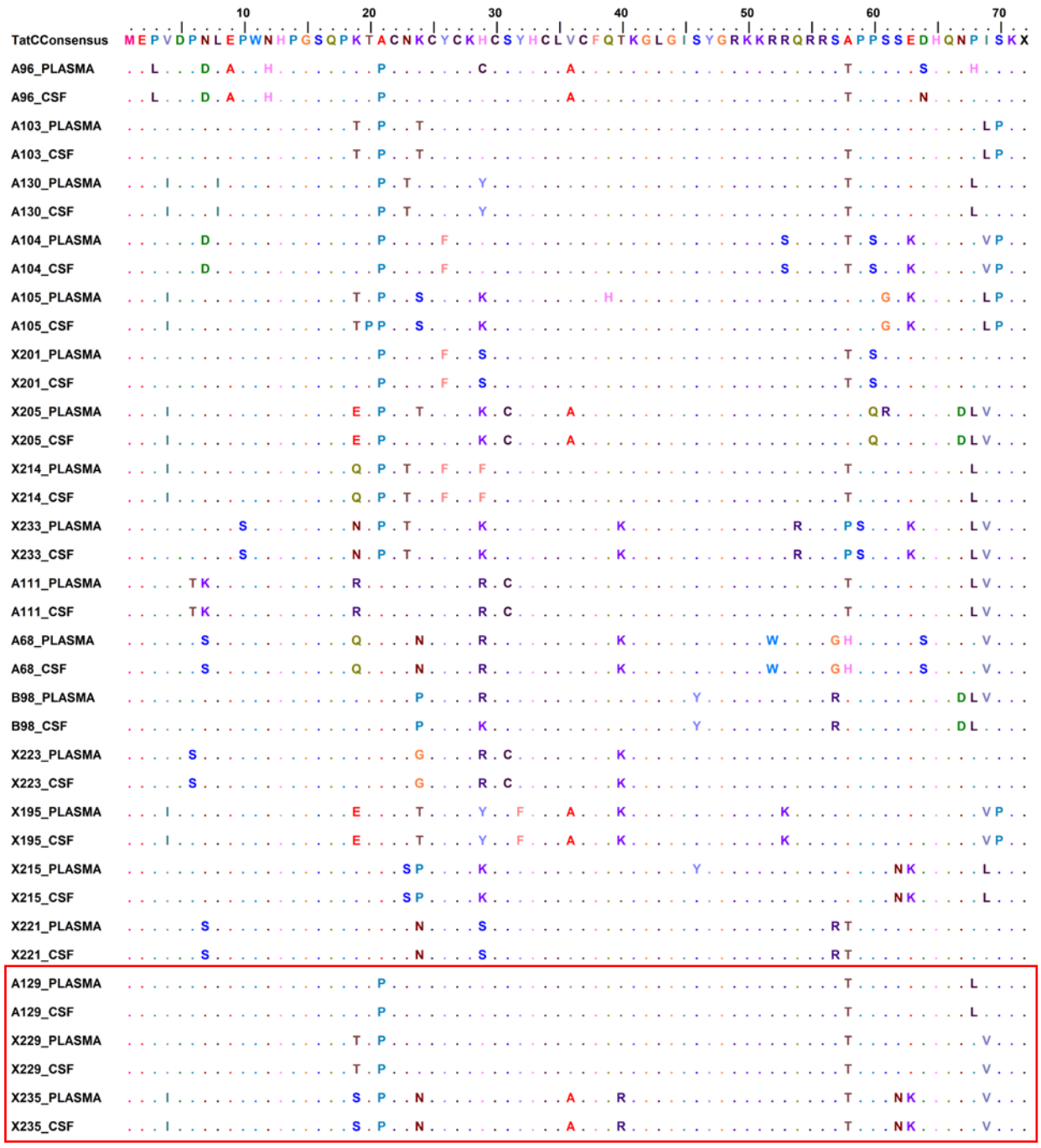


Figure 3:5 Amino acid sequence alignment of subtype C consensus sequence and 38 full length patient derived Tat sequences. Schematic presentation of the 71 amino acids within Tat exon 1: domain I (residues 1 to 21, acidic domain); domain II (residues 22 to 37, cysteine domain); domain III (residues 38 to 48, core domain); domain IV (residues 49 –57, basic domain) and; domain V (residues 58–71, glutamine domain). All sequences are aligned to subtype C consensus in the first line of the alignment, the dots represent residues identical to the consensus. Outlined in red is all non-TBM study participants.

Table 3:1 Allele frequency of previously reported Tat mutations in the CSF and plasma compartments of all 16 TBM study participants.

Functional Domain	Mutations	Allele Frequency (%)	
		PLASMA	CSF
Acidic Domain	V4I	31	31
Acidic Domain	P21A	43	43
Cysteine-rich Domain	K24S	6	6
Cysteine-rich Domain	H29R	25	19
Cysteine-rich Domain	S31C	18	18
Core Domain	S46Y	13	6
Basic Domain	R52W	6	6
Basic Domain	S57R	13	13
Glutamine Rich Domain	P59S	6	6
Glutamine Rich Domain	D64N	-	6

Additionally, a VESPA analysis to detect CNS signature residues was conducted on all 16 TBM plasma and CSF sequences to determine the difference in viral signature patterns. Although no signature residues were found, a further codon by codon analysis showed that patients; A105, B98 and A96 harboured additional mutations in the CSF that were absent from the plasma compartment these include T20P, H29K and D64N respectively. The effect D64N mutation on transactivation activity of Tat, was subsequently assessed.

3.4. Tat Transactivation activity

To determine the impact of specific Tat variants on HIV-1 LTR transcription, a Tat Transactivation assay was carried out using TZM-bl cells as previously described (section 2.8). Since the amino acids of reported mutations, were seen in almost equal proportions within plasma and CSF compartments (Figure 3:3, Table 3:1), one representative sequence was chosen from each study participant (except X205 where both plasma and CSF derived Tat were selected). The basis of this was to identify amino acid substitutions that allowed for altered transactivation abilities of Tat. Briefly patient derived Tat samples

containing reported mutations were selected. These Tat variants were compared to a patient derived Tat sequence (A129C) that highly represented (96%) the subtype C consensus sequence previously described with only 3 amino acids changes: 21P, 58T and 68L. Henceforth A129C will be referred to as the subtype C consensus.

Overall two analyses were performed:

(i) The relative change in LTR activity of patient derived Tat
(ii) The relative change in LTR activity of a single polymorphism

Within the patient derived Tat, heterogeneity was observed in the ability of CSF and plasma derived Tat variants to transactivate the LTR (Figure 3:6). The transactivation ability of Tat was determined by the relative change in LTR activity. Overall patient derived Tat from 5 patients (A130C, B98C, X195C, X205P and A111C) had significantly restricted capacity to transactivate the LTR.

Patient A130C with mutation V4I had significantly downregulated LTR activity by 37% (± 6%) relative to the subtype C consensus ($p = 0.01$). A previous report showed that the V4I mutation had significantly downregulated LTR activity through site directed mutagenesis (Rossenkhan et al., 2013). In our study, we did not perform site directed mutagenesis, however the A130C patient derived Tat that contained other mutations together with the V4I mutation showed a decreased LTR activity. In addition the V4I mutation coupled with P21A as seen in X195C resulted in only a 34% (± 7%) reduction in Tat activity ($p = 0.02$), although the dual effort of both mutations has not been previously reported, the single effect of V4I and P21A is known to lower LTR activation (Rossenkhan et al., 2013). Another double mutation P21A and S46Y seen in patient B98C significantly, decreased LTR activity by 46% (± 6%) ($p = 0.007$). Whilst the P21A mutation is a naturally occurring polymorphism, position 46 is a site of phosphorylation could explain the reduction in Tat activity by almost 50%. Although we cannot assume which mutation had a more dominant effect.

Whilst the V4I mutation was seen in X205 plasma and CSF derived Tat, the plasma derived Tat from this patient harbored additional mutations; K24T, and S61R (Figure 3:5) which has been associated with a reduction in Tat activity. The presence of K24T and S61R mutations in X205P may be the cause of diversity differences within the plasma and CSF Tat in this subject (Figure 3:2). In addition, X205P had a significantly reduced LTR transactivation activity by 57% (± 11%) when compared to the consensus (Figure 3:6) ($p = 0.04$). Whilst it is not clear which mutation in X205P had the more pronounced effect, this patient derived Tat sample synergistically reduced LTR activity. Furthermore, a large difference in LTR activity was observed between X205P and X205C (although not significant), however site directed mutagenesis is needed to target which mutation is responsible for this antagonistic effect of Tat between compartments.

The biggest decrease in LTR activity across all patients was seen in A111C, that contained a triple mutation P21A, H29R and S31C, which reduced LTR activity by 61% (± 4%) ($P = 0.002$). Whilst H29R is previously reported to enhance relative LTR activity by site directed mutagenesis, the combination of H29R (enhancing), P21A (detrimental) and S31C (detrimental) resulted in a reduction of LTR activity. Which could imply that H29R is less dominant than the two detrimental mutations.

Conversely, the greatest increase in Tat activity (although not statically significant) was seen in X214C that enhanced LTR activity by 18% (± 9%). Although X214C contained the V4I mutation, we cannot say for certain that the higher Tat activity in this patient is due to this particular mutation – more so that the opposite effect has been reported for the V4I mutation. This highlights the point that the synergistic effect of the combined mutations: V4I, K19Q, A21P, N23T, Y26F, H29F, A58T and P68L as seen within X214C (although not previously reported) could be driving higher LTR activity in this patient. Furthermore, single amino acid substitutions, followed by transactivation of each mutation will be needed to demonstrate synergy. Similarly, patient A96C had the V4I and D64N mutation that was only present in its CSF derived Tat, which also displayed an enhanced Tat ability to transactivate the LTR (not significant).

Lastly X233C with mutation P59S and A105C both had similarity in inducing the LTR by 1% relative to the consensus, although no statistically significant. To summarize, approximately 66% of the samples were associated with a downregulation of LTR activity, 17% had no difference and 17% upregulated LTR activity.

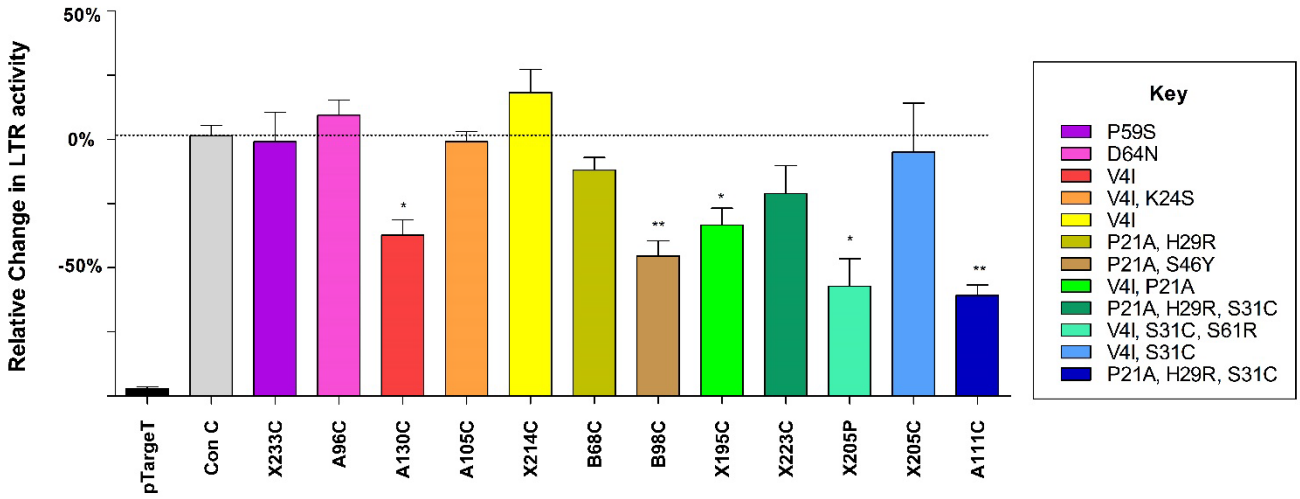


Figure 3:6 Relative changes in LTR activity of patient derived Tat. The change in patient-derived LTR activity were determined relative to the HIV-1 subtype C consensus sequence (Con C) measured 24 hours after transfection. The pTargeT empty vector was included as a negative control. Values represented are the average from each sample performed in triplicate, and the error bars indicate the standard error of the mean (SEM). Statistical significance is indicated as follows: *, $P < 0.05$; **, $P < 0.01$. The colour encoded key represents the different Tat mutations found within patient the Tat derived. Patient ID were labeled either CSF derived (C) or plasma derived (P).

Albeit, the reported mutations mentioned above does not necessarily prove that LTR activity was affected by those specific mutations alone. Rather the combination of these mutations within each Tat variant had resulted in reduced or enhanced LTR activity. Although this analysis included previously reported mutations, it remains to be determined if other mutations also contribute to the combined detrimental or enhancing activity of Tat and further site directed mutagenesis studies are needed for that.

To determine the impact of specific Tat mutations on HIV-1 LTR transactivation, we selected 5/12 (42%) of Tat sequences that contained the P21A mutation (amongst other mutations) for the Tat Transactivation assay. The relative change in LTR activity was calculated using the mean RLU of all patients with that mutation compared to the subtype consensus C. Aforementioned the P21A polymorphism was reported to be associated with attenuated transactivation activity (Rossenkhan et al., 2013). Therefore, to determine the effect of the Tat variants with the P21A mutation to LTR activity in our cohort, five patient derived Tat with Alanine at position 21 (Ala^{21}) and 7 Tat variants with Proline at position 21 (Pro^{21}) irrespective of other mutations, were compared to a subtype C consensus. Our data shows all Tat variants with Ala^{21}, were associated with significantly reduced Tat activity by 40% (± 8%) compared to the Consensus C ($p = 0.009$) (Figure 3:7). Although, the relative change in LTR activity in variants with a Pro^{21} was greater than variants with an Ala^{21}, this trend was not statically significant (Figure 3:7). To increase the statistical significance of this result a greater sample size is warranted.

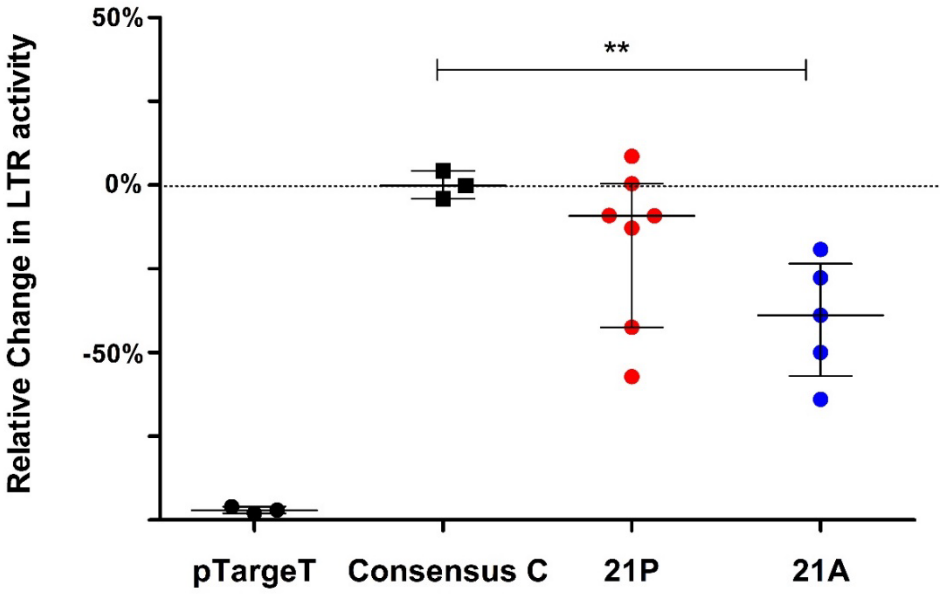


Figure 3:7 Relative changes in LTR activity of a single Tat mutation. The change in LTR activity of the P21A mutation were determined relative to the HIV-1 subtype C consensus sequence (Consensus C) measured 24 hours after transfection. The pTargeT empty vector was included as a negative control. Values represented are the average taken from independent samples performed in triplicate, and the error bars indicate the standard error of the mean (SEM). Statistical significance is indicated as follows: *, $P < 0.05$; **, $P < 0.01$.

Hence all individuals with a wildtype Pro^{21} had a higher LTR activity than individuals with an Ala^{21} which could implicate that mutation P21A irrespective of other mutations drastically reduces LTR activity.

3.5. Tat LTR Activity correlated to markers of disease progression

Lastly, to assess the relationship between Tat-mediated LTR activity and markers of disease progression such as viral load – a Pearson correlation coefficient and linear regression was done. Overall 3 analyses were done as seen in Figure 3:8: (A) correlating the relative LTR activity of all 12 patient derived Tat variants to viral load, (B) correlation to show the relationship of Tat variants with the Ala^{21} mutation to CSF viral load and (C) correlation to show the relationship of Tat variants with the wildtype Pro^{21} mutation to viral load.

As observed in Figure 3:8A, all twelve patient derived Tat variants and their corresponding mutations were correlated to viral load (Table 6:1). Our data shows a positive yet non-significant trend of Tat LTR activity and viral load in all 12 patient derived Tat. We next wanted to determine the correlation between the Ala^{21} and Pro^{21} mutation to viral load. Our results show that individuals with an Ala^{21} mutation showed a strong positive correlation with CSF viral load ($r = 0.98$; $p = 0.003$) (Figure 3:8B). However, this correlation does not show the effect of the Ala^{21} mutation on viral load, rather shows the relationship between transactivation ability in variants with the Ala^{21} mutation to patient viral load. Conversely, a weak positive correlation between Tat-mediated LTR activity and Tat variants with a Pro^{21} was observed ($p = 0.39$; $r = 0.38$) (Figure 3:8C). Hence the Pearson's correlation showed no concordance between data, indicating that Pro^{21} did not influence Tat mediated replication capacity.

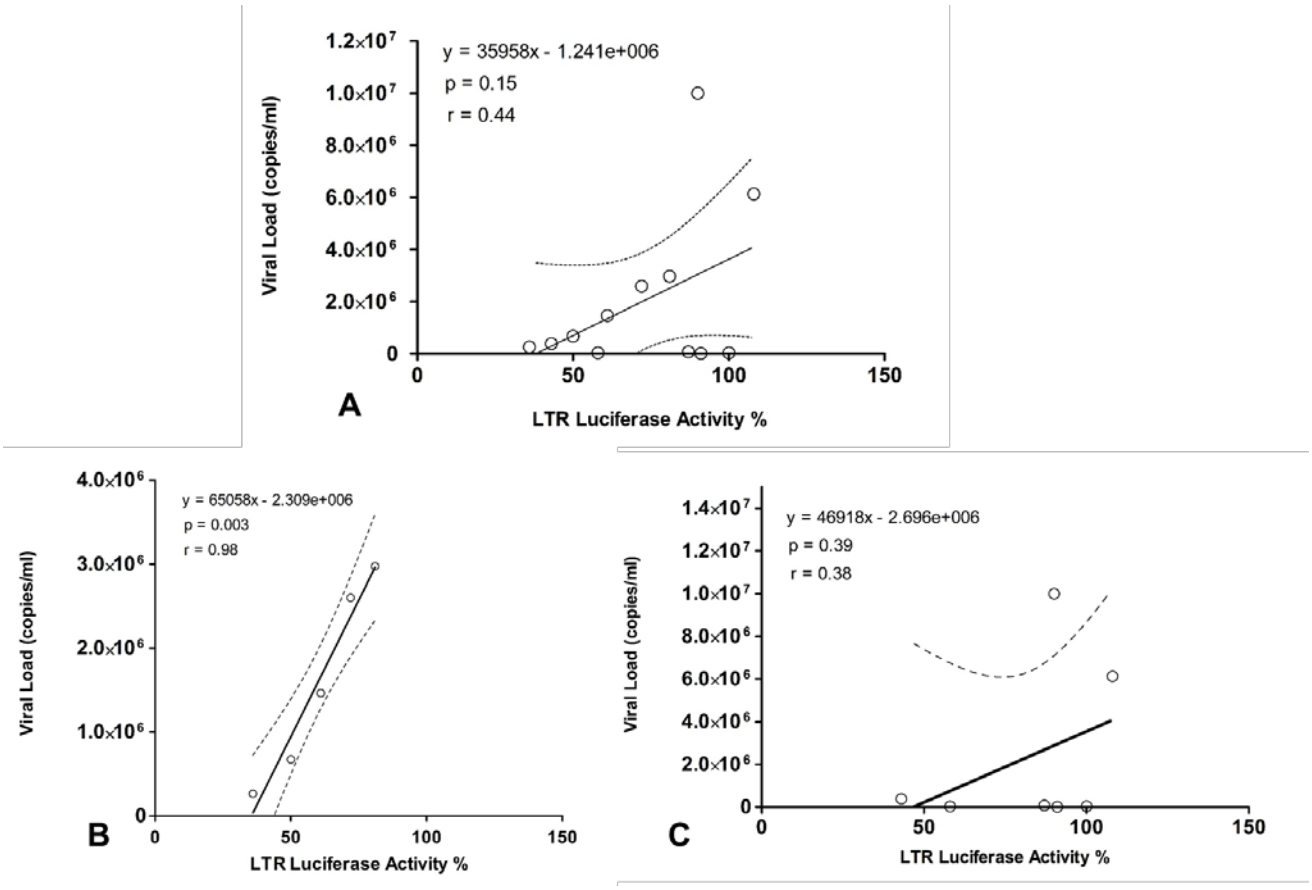


Figure 3:8 Correlation of P21A Tat activity with viral load. The strength of association between Tat-mediated LTR activity (%) and viral load (copies/ml) were measured by a Pearson correlation coefficient and linear regression (solid line) for all twelve patient derived Tat (A), Tat variants with a 21A mutation (B) and Tat variants with a 21P mutation (C). The dotted lines represent the 95% confidence intervals.

Whilst the linear relationship between the Tat-mediated luciferase expression and patient viral loads showed a positive trend and statistical significance for the Ala^{21} mutation, there are two major limitations; number of variants used in the correlation was small and greater a sample size would allow for a much more precise trend. Secondly, the trend that when Ala^{21} mutation is held constant, viral load was linked to LTR transactivation could be confounded. Since Tat variants with the Ala^{21} mutation was occurring together with other mutations, therefore we cannot necessarily rule out that it is solely the mutation Ala^{21} that is responsible for the observed effect.

Chapter 4

4. Discussion

In our study we hypothesized that CNS specific Tat mutations were responsible for the high viral replication seen in the CNS of TBM patients. To test this hypothesis, we genetically and functionally characterized HIV-1 *tat* from 19 patients with TBM. Plasma and CSF derived *tat* sequences were successfully amplified and sequenced in all 19 patients to determine genetic variation, our data showed that only 5 patients had differences between *tat* sequences in each compartment, whilst majority showed no difference between plasma and CSF sequences in each patient. Next, 12 different Tat variants, representative of reported mutations were cloned into a mammalian expression vector (pTargeT) and transfected into TZM-bl cells in order to determine the functional diversity of CNS derived Tat from TBM co-infected patients. Although our study confirmed that the previously reported Tat mutations impact its activity, none of these mutations were CSF specific. Taken together the data from our study rejects the hypothesis.

The major aim was to characterize the effect of HIV-1 *tat* genetic variation and functional diversity on viral replication in the CNS of TBM and non-TBM patients. In 1998, a study reported that patients with TBM had the highest amount of HIV-1 RNA copies/ml in the CSF compared to patients infected with other types of meningitis (Morris et al., 1998). Twenty years later the same observation was described with plasma and CSF HIV-1 viral load compared to other types of meningitis (Seipone et al., 2018). Hence the cause of what drives higher viral replication in the CNS of TBM patients is still unknown.

Data from the phylogenetic analysis revealed that *tat* sequence pairs obtained from matched CSF and plasma of each patient clustered closely together with substantial genetic variation between compartments as was reflected by branch length differences. Although no tests were done for viral compartmentalisation between these two compartments, a prior study observed minimal compartmentalisation of *tat* sequences between CSF and plasma derived viruses in patients with HIV-associated dementia (HAD) (Cowley et al., 2011). Additionally, the unpublished data from our group have described similar results for *env* sequences from patient X195 in the TBM study, however

their data were unable to show sequence segregation between compartments of patients B98 and X223. Whilst we did not test for compartmentalisation, our data show that 2 patients: X229 and A104 had no nucleotide differences between their respective plasma and CSF *tat* sequences, while 58% had evident differentiation between compartment which could indicate interpatient diversity. Furthermore, intermingling between TBM and non-TBM study participants were observed. Therefore, we reason that an interpatient diversity analysis would determine the precise diversity between sequences in the TBM and non-TBM groups (Figure 3:2). The analysis showed that TBM patients X205 and A96 had the highest diversity when comparing their CSF to plasma sequences. Large differences in diversity between these groups were expected, since the viral loads in the TBM group were notably higher for plasma and CSF ($p= 0.02$) than the non-TBM group. However, overall there was no significant difference in diversity between the TBM and non-TBM groups. One of the limiting factors in this analysis could be due to the small sample size in the non-TBM group.

Next, we reasoned that Tat differences could be more pronounced within a patient (intrapatient). Focusing on the TBM group an intrapatient diversity analysis was conducted to determine if *tat* sequences were more diverse in the CSF compartment than plasma of the same TBM patients. While CSF-derived *tat* sequences showed a slightly higher diversity compared to plasma there was no significant difference in *tat* sequences between these compartments. This was further confirmed when the allele frequencies of reported Tat mutations were represented in almost equal proportions within plasma and CSF Tat sequences (Table 3:1). Although HIV-1 compartmentalization has been previously observed between plasma and CSF from patients with opportunistic infections and HIV associated neurocognitive disorders (HAND), one plausible reason for the lack segregation within Tat variants in our current study could be due to pleocytosis of the blood brain barrier (BBB). Based on the published criteria for pleocytosis; normal values of CSF lymphocyte counts are 0 to 5 cells/µl in HIV-1 infected patients, while pleocytosis is designated as CSF lymphocyte counts >5 cells/µl (Fishman, 1992). In lieu of this, approximately 63% of the 16 individuals in the TBM group are undergoing lymphocytic pleocytosis (Table 6:1). Due to the large of amount of lymphocyte trafficking through the

BBB, the introduction of plasma derived viruses into the CNS or visa versa maybe the cause of no differentiation in sequences between plasma and CSF derived Tat.

Whilst there was no significant difference between plasma and CSF derived Tat sequences, we assumed that mutations observed in our cohort may be responsible for the higher viral replication in TBM co-infected patients. The residues vital for Tat function are highly conserved these include; K28, K41, K50 and K51 (for acetylation); R56 and R57 (for TAR interaction); C22, C31, C34 and Y47 (for LTR transactivation), S16 and S46 (for phosphorylation) (Ronsard et al., 2017). Almost all these residues remained highly conserved within our cohort. Interestingly our data show two subtype C specific mutations; a S57 (84%) is more prevalent than a R57 (11%) as seen in subtype B, this is in line with other findings that suggest S57 coupled with E63 contribute to the higher activity of HIV-1C Tat (Kurosu et al., 2002). Likewise the C31S mutation has also been associated to HIV-1C Tat, however Tat activity between S31 and C31 are not significantly different (Kurosu et al., 2002).

Our data shows that majority of amino acid substitutions observed in this cohort were between amino acids (aa) 1 – 48, which is in agreement with a previous report (Rossenkhan et al., 2013). Hence, aa: 1 – 48, comprise the first three domains of Tat making up the Tat activation domain, which is responsible for transactivation activity and CycT1 binding (Li et al., 2012). Although high variability in Tat is expected during a chronic infection, amino acid changes within the Tat activation domain (aa: 1 – 48) could impact the Tat mediated elongation, affecting the transactivation ability of Tat variants in our cohort.

In the acidic domain both the V4I and P21A mutation were represented in the highest frequencies of 31% and 43% in TBM patients respectively, this could imply that they are naturally occurring polymorphisms. In agreement, two consecutive studies found the P21A mutation to be undergoing positive selection and unique to HIV-1C Tat (Neogi et al., 2012, Rossenkhan et al., 2013). Although, positive selection is used to detect amino

acid mutations increase reproductive fitness (Pan et al., 2006) , the 21A mutation had shown the greatest impact on reducing LTR activity by 88% (Rossenkhan et al., 2013).

Additionally, sites 19 and 24 in the acidic domain and 29 in the cysteine domain, had high frequency variation at these positions. Although it is still unclear what causes the high variation at these particular positions, mutations K24S and H29R were previously confirmed by site directed mutagenesis to have a detrimental (K24S) and enhancing (H29R) effect on LTR activity (Rossenkhan et al., 2013). Hence further site directed mutagenesis needs to be done at mutations K19T, K24N and H29K which had high occurrence in our cohort.

Within the cysteine domain, the reported S31C was observed in 19% of TBM patients. The C31S polymorphism is seen in approximately 90% of individuals with HIV-1C and is associated with reduced cytokine expression of monocytes circulating in the periphery (Rao et al., 2013). Thus, the reversion mutation S31C that restores the C30C31 dicysteine motif in HIV-1C as seen in our cohort, could impact Tat function. This is based on a previous report showing that the C30C31 motif allows for high cognitive impairment in patients with HAND (Mishra et al., 2008).

In the core domain, the S46Y was observed in 13% and 6% of plasma and CSF sequences respectively, since position 46 has been reported to aid in phosphorylation of Tat by cyclin-dependent kinase 2 (CDK2), mutations in this position attenuate Tat transactivation (Ammosova et al., 2006). The RNA binding domain (RBD) (composed of amino acids 49 – 57) contains nuclear localization signals that allow for efficient Tat-TAR binding. Therefore, amino acid changes observed in this region could have functional consequences on LTR activity. Mutations in the RBD include R52W, P59S and S61R, which were all previously associated with reduced Tat function, due to inefficient Tat-TAR binding resulting in a drastic decrease of viral transcription (Rossenkhan et al., 2013, Sivakumaran et al., 2007). Whilst there was limited segregation between CNS and plasma derived Tat, three patients A105, B98 and A96 did show CNS specific mutations including: T20P, H29K and D64N respectively. However since the D64N in PID A96 has

been reported to reduce TAR binding (Endo-Munoz et al., 2005), this Tat variant was selected for downstream analysis.

Since Tat sequences demonstrated substantial sequence diversity in the Tat activation domain and low frequency mutations in the RBD, we sought to assess the functional effects of this sequence variation. In particular, we were interested in the functional capacity of previously reported mutations within Tat.

The functional variation was measured using 10 different CSF Tat variants and 1 matched (plasma and CSF) sample, hence a total of 12 samples each representing a one or more reported mutations. Our findings suggest that there is heterogeneity in the ability of CNS derived Tat to transactivate the LTR, whilst 66% of the samples were associated with a downregulation of LTR activity, 17% had no difference – relative to the subtype C consensus and 17% upregulated LTR activity. Although majority of TBM patients demonstrated a restricted ability to transactivate the LTR, a functional comparison of CSF derived Tat from non-TBM patients would have been ideal to test our hypothesis. As prior research showed that CNS derived Tat had defective transactivation abilities when compared to non-HAD patients (Boven et al., 2007).

In agreement with previous studies mutations: S46Y (detrimental) and S57R (enhancing) as seen in patient B98C had significantly reduced ability to transactivate the LTR by 46% (± 6%) ($p = 0.007$)(Ruiz, 2016, Sivakumaran et al., 2007). Whilst the functional capacity of each individual mutation within the B98C is not clear, the S46Y mutation which was previously associated with high attenuation of Tat activity exerted a more dominant effect.

However, the biggest reduction in transactivation activity was seen in A111C which downregulated LTR activity by 61% (± 4%) ($p = 0.002$). The A111C sample was a triple mutant with P21A, H29R and S31C mutations. Although the A111C had the H29R mutation, reported to upregulate Tat activity, upon further analysis A111C also harbored an additional mutation K19R , previously described to impair transactivation activity (Kamori and Ueno, 2017). Therefore 3 out of the 4 mutations were associated with

detrimental Tat activity and could be responsible for the significant decrease in Tat transactivation. Additionally, K19T, P21A and S31C are also reported detrimental mutations. Although A111C and other patient derived Tat clones had a downregulated Tat activity – further mutagenesis studies are required to identify the specific residues involved.

Aforementioned, patient A96 with the CNS specific D64N mutation had enhanced LTR activity by 9% (although not significant). This patient also had a reducing V4I mutation, which could imply that the D64N mutation exerted a more dominant effect, resulting in the upregulation of CNS derived Tat from this patient. Albeit, both site directed mutagenesis and a functional comparison against the A96 plasma compartment is needed to answer this question.

To determine the effect of a single amino acid mutation, we measured the transactivation ability of the P21A mutation, observed in 5 patient derived Tat variants with Ala^{21} and 7 patient derived Tat variants with Pro^{21}, although it should be noted that these variants also contained other mutations as expected in during HIV-1C chronic infection. Consequently, without performing site directed mutagenesis we assessed the antagonistic effect of all patient derived Tat variants with the Ala^{21} mutation and variants with the Pro^{21} mutation (Figure 3:7). In agreement with a prior study our data shows that patient derived Tat variants with an Ala^{21}, were associated with significantly reduced Tat activity by 40% (± 8%) compared to the Consensus C ($p = 0.009$). Furthermore, the relative Tat activity of variants with the wildtype Pro^{21} was higher by 29% when compared to Tat variants with an Ala^{21} (although not statically significant). Since the P21A mutation is found within the Tat activation domain, it could suggest that variants with the mutation have impaired transcriptional elongation leading to inefficient gene expression and a low viral replication. Although this is the opposite of what we expected to find in the CNS of TBM patients – as the higher viral load seen in the CSF of TBM patients is indicative of a faster disease progression. Albeit a slower disease progression could imply a virulence trade-off for the virus to allow for longer survival in the host (Brown et al., 2012).

We next examined the relationship between Tat mediated LTR activity and viral load. In the first analysis (Figure 3:8A), our data shows a moderate positive, yet non-significant correlation between Tat transactivation activity and viral load. Next, reasoned that Tat variants with the Ala^{21} mutation and Pro^{21} mutation would show a clearer trend. Our data indicates a strong positive correlation ($p = 0.003$; $r = 0.98$) between Tat LTR activity and viral load, in all patients with the Ala^{21} mutation, which could suggest that the P21A mutation is a significant contributor to CSF viral load during chronic infection (Figure 3:8B). However, to fully test this hypothesis, the matched plasma samples from these patients should also be correlated with viral load. In addition, the Pro^{21} mutation which had high transactivation activity in our cohort – showed no concordance with patient viral load (Figure 3:8C). This result could account for the moderate trend observed in Figure 3:8A.

To the best of our knowledge, this is the first study to look at CNS derived Tat from patients co-infected with TBM. Although there is paucity of data on the genetic diversity of the *tat* gene and its association with disease progression during chronic infection. Our data suggests that variation within CNS derived *tat* from patients co-infected with TBM may affect the functional capacity. Differences in LTR transactivation activity, was observed in different mutations within the 5 domains of *tat* exon 1. Whilst we expected that sequence variation within *tat* would result in a higher viral replication as seen in the CSF of patients with TBM, most of the amino acid differences resulted in a significantly reduced ability of Tat to transactivate the LTR. More so we observed that the detrimental mutation P21A was significantly correlated to viral load in the CNS compartment. Although we reject our hypothesis, we can still prove that Tat transactivation activity may contribute to viral load in chronic infection.

Whilst measuring the genetic and functional diversity of Tat, did not fully correlate with the high viral production in the CNS, as implied by the baseline clinical data. There are various possible explanations that could account for the high viral load in the CSF of TBM patients. Firstly, high viral load in the CNS could be due to the elevated levels of cytokines interleukin-1β (IL-1β), IL-6, IL-8, IL-10, interferon gamma (IFN-γ) and TNF-α (Patel et al.,

2011). In a recent study, identifying immunological biomarkers in the CSF between TBM and non-TBM infected individuals, elevated levels of IL-6, IL-8 and IL-10 were found to be significantly higher in the CSF of TBM patients compared to plasma (Seipone et al., 2018). Therefore, it is possible that Tat could play an integral role in the activating and enhancing the immune response of TBM in HIV-1 co-infected persons leading to increased viral replication. Hence a better understanding of Tat immune dysregulation in the CNS could potentially inform therapeutic interventions and AIDS related pathologies. Secondly, since subtype C Tat that is well adapted for a subtype C LTR and the TZM-bl cell line used in our transfection assays contain a subtype B LTR, the resultant effect of some mutations could possibly be masked. While there are no TZM-bl cell lines available that contain an integrated subtype C LTR, this system to test the activity of our subtype C derived Tat. Thirdly, cells of the CNS including macrophages/monocytes, microglia, and astrocytes all produce their own transcription factors, hence high viral load in the CNS could be due to the result of cell specific transcription regulation and highly active transcription factors – regardless of the attenuation observed in Tat.

The present study shows the importance of *tat* diversity in study participants chronically infected with TBM. However, there are few limitations that could be improved as well as future work that needs to be done.

One of the biggest limitations in this study, as previously mentioned is the small sample size of the non-TBM group. Albeit, additional samples from non-TBM patients is difficult to obtain due to the invasive lumbar puncture procedure required for extraction of CSF. In addition, all *tat* plasma and CSF variants from both TBM and non-TBM participants should have been functionally assayed for a more thorough comparison as the representative sample set chosen could not accurately answer our question. Secondly, single genome amplification (SGA) would be a more precise measure to test for genetic variation and compartmentalization in *tat* sequences from the CSF and plasma of TBM patients as bulk PCR could account for potential bias from recombination and resampling of DNA (Butler et al., 2009). Moreover, this would have allowed for stronger comparisons of previously reported and novel mutations to be identified. Lastly, whilst TZM-bl cells are

an adherent cell line and could be transfected and easily measured due to *luc* driven by the LTR promoter, it doesn't necessarily represent the population of infected cells that are able to cross the blood brain barrier. Hence transfection into cells of the monocytes/macrophage lineage, endothelial cells and SVG astrocytes could be used to investigate functional activity of Tat. In addition, there was disparity is fully understanding the transactivation data as each sequence had more than one mutation, therefore we cannot say for certain that a particular mutation causes and decrease/increase in LTR activity. Further site directed mutagenesis will be needed to fully answer that question.

Future work recommendations include: electrophoretic mobility shift assay (EMSA) to further study the Tat-TAR interaction and relative protein expression of Tat variants with reported mutations as identified in this cohort. This study provides evidence that Tat is a key regulator of viral gene transcription and ultimately disease progression. In essence, modern practical approaches that target the cellular transcription factors including p-TEFb elongation and NF-kB LTR activation could provide valuable clues to guide research on vaccine development and viral eradication.

References

5. References

CDC 1981a. Pneumocytis Pneumonia - Los Angeles. *Morbidity and Mortality Weekly Report***,** 250 - 252.

CDC 1981b. Kaposi's sarcoma and Pneumocystis Pneumonia among homosexual men—New York City and California. *Morbidity and Mortality Weekly Report***,** 305-307.

CDC 1982a. Update on the Acquired Immune Deficiency Syndrome (AIDS)—United States. *Morbidity and Mortality Weekly Report***,** 507-514.

CDC 1982b. Current trends update on Acquired Immune Deficiency Syndrome (AIDS)—United States Weekly. *Morbidity and Mortality Weekly Report***,** 507-508, 513-514.

BARRÉ-SINOUSSI, F., CHERMANN, J.-C., REY, F., NUGEYRE, M. T., CHAMARET, S., GRUEST, J., DAUGUET, C., AXLER-BLIN, C., VÉZINET-BRUN, F. & ROUZIOUX, C. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science,* 220**,** 868-871.

GALLO, R. C., SALAHUDDIN, S. Z., POPOVIC, M., SHEARER, G. M., KAPLAN, M., HAYNES, B. F., PALKER, T. J., REDFIELD, R., OLESKE, J. & SAFAI, B. 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *science,* 224**,** 500-503.

FRANCKI, R. I. B., FAUQUET, C., KNUDSON, D. L. & BROWN, F. 1991. Classification and Nomenclature of Viruses : Fifth Report of the International Committee on Taxonomy of Viruses. *Springer-Verlag Wien New York***,** 140-144, 290-299.

BARRE-SINOUSSI, F. 1983. Chermann 1-C, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vénizet-Brun F, Rouzioux C, Rozenbaum W, and MontagnierL: Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science,* 220**,** 868-871.

UNAIDS 2018. United Nations Programme on HIV/AIDS.

LEMEY, P., PYBUS, O. G., WANG, B., SAKSENA, N. K., SALEMI, M. & VANDAMME, A.-M. 2003. Tracing the origin and history of the HIV-2 epidemic. *Proceedings of the National Academy of Sciences,* 100**,** 6588-6592.

HEMELAAR, J., GOUWS, E., GHYS, P. D. & OSMANOV, S. 2006. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *Aids,* 20**,** W13-W23.

VALLARI, A., HOLZMAYER, V., HARRIS, B., YAMAGUCHI, J., NGANSOP, C., MAKAMCHE, F., MBANYA, D., KAPTUÉ, L., NDEMBI, N. & GÜRTLER, L. 2011. Confirmation of putative HIV-1 group P in Cameroon. *Journal of virology,* 85**,** 1403-1407.

FOLEY B, L. T., APETREI C, HAHN B, MIZRACHI I, MULLINS J, RAMBAUT A, WOLINSKY S, AND KORBER B, EDS 2017. HIV Sequence Compendium 2017. *Published by Theoretical Biology and Biophysics Group, Los Alamos National Laboratory,* NM, LA-UR 17-25240.

THOMSON, M. M., PÉREZ-ÁLVAREZ, L. & NÁJERA, R. 2002. Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. *The Lancet infectious diseases,* 2**,** 461-471.

SANTOS, A. F. & SOARES, M. A. 2010. HIV genetic diversity and drug resistance. *Viruses,* 2**,** 503-531.

HEMELAAR, J. 2012. The origin and diversity of the HIV-1 pandemic. *Trends in molecular medicine,* 18**,** 182-192.

TAYLOR, B. S., SOBIESZCZYK, M. E., MCCUTCHAN, F. E. & HAMMER, S. M. 2008. The challenge of HIV-1 subtype diversity. *New England Journal of Medicine,* 358**,** 1590-1602.

KALEEBU, P., FRENCH, N., MAHE, C., YIRRELL, D., WATERA, C., LYAGOBA, F., NAKIYINGI, J., RUTEBEMBERWA, A., MORGAN, D. & WEBER, J. 2002. Effect of human immunodeficiency virus (HIV) type 1 envelope subtypes A and D on disease progression in a large cohort of HIV-1—positive persons in Uganda. *The Journal of infectious diseases,* 185**,** 1244-1250.

HEMELAAR, J., GOUWS, E., GHYS, P. D. & OSMANOV, S. 2011. Global trends in molecular epidemiology of HIV-1 during 2000–2007. *AIDS (London, England),* 25**,** 679.

COFFIN, J., HAASE, A., LEVY, J. A., MONTAGNIER, L., OROSZLAN, S., TEICH, N., TEMIN, H., TOYOSHIMA, K., VARMUS, H. & VOGT, P. 1986a. Human immunodeficiency viruses. *Science,* 232**,** 697.

COFFIN, J., HAASE, A., LEVY, J. A., MONTAGNIER, L., OROSZLAN, S., TEICH, N., TEMIN, H., TOYOSHIMA, K., VARMUS, H. & VOGT, P. 1986b. What to call the AIDS virus? *Nature,* 321**,** 10-10.

KURTH, R. & BANNERT, N. 2010. *Retroviruses: molecular biology, genomics and pathogenesis*, Horizon Scientific Press.

FREED, E. O. 2001. HIV-1 replication. *Somatic cell and molecular genetics,* 26**,** 13-33.

KUIKEN, C., FOLEY, B., HAHN, B., KORBER, B., MCCUTCHAN, F., MARX, P., MELLORS, J., MULLINS, J., SODROSKI, J. & WOLINKSY, S. 1999. Human retroviruses and AIDS 1999: a compilation and analysis of nucleic acid and amino acid sequences. *Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, N. Mex*.

FERNANDES, J. D., FAUST, T. B., STRAULI, N. B., SMITH, C., CROSBY, D. C., NAKAMURA, R. L., HERNANDEZ, R. D. & FRANKEL, A. D. 2016. Functional segregation of overlapping genes in HIV. *Cell,* 167**,** 1762-1773. e12.

SUZUKI, Y. & SUZUKI, Y. 2011. Gene regulatable lentiviral vector system. *Viral gene therapy.* InTech.

HTTP://WWW.HIV.LANL.GOV 2018. Los Alamos National Laboratory. HIV Sequence Database.

FRANKEL, A. D. & YOUNG, J. A. 1998. HIV-1: fifteen proteins and an RNA. *Annual review of biochemistry,* 67**,** 1-25.

FAUCI, A. S. 1988. The human immunodeficiency virus: infectivity and mechanisms of pathogenesis. *Science,* 239**,** 617-622.

FERGUSON, M. R., ROJO, D. R., VON LINDERN, J. J. & O'BRIEN, W. A. 2002. HIV-1 replication cycle. *Clinics in laboratory medicine,* 22**,** 611-635.

BOYER, P. L., SMITH, S. J., ZHAO, X. Z., DAS, K., GRUBER, K., ARNOLD, E., BURKE, T. R. & HUGHES, S. H. 2018. Developing and evaluating inhibitors against the RNase H active site of HIV-1 RT. *Journal of virology***,** JVI. 02203-17.

FREED, E. O. 1998. HIV-1 gag proteins: diverse functions in the virus life cycle. *Virology,* 251**,** 1-15.

KIRCHHOFF, F. 2013. HIV life cycle: overview. *Encyclopedia of AIDS***,** 1-9.

CHRIST, F. & DEBYSER, Z. 2013. The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy. *Virology,* 435**,** 102-109.

COFFIN, J. M., HUGHES, S. H. & VARMUS, H. E. 1997. *The interactions of retroviruses and their hosts*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).

GOMEZ, C. & HOPE, T. J. 2005. The ins and outs of HIV replication. *Cellular microbiology,* 7**,** 621-626.

SUNDQUIST, W. I. & KRÄUSSLICH, H.-G. 2012. HIV-1 assembly, budding, and maturation. *Cold Spring Harbor perspectives in medicine***,** a006924.

VOTTELER, J. & SUNDQUIST, W. I. 2013. Virus budding and the ESCRT pathway. *Cell host & microbe,* 14**,** 232-241.

JEANG, K.-T. 1996. HIV-1 Tat: structure and function. *Human Retroviruses and AIDS 1996: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences***,** 11-26.

DINGWALL, C., ERNBERG, I., GAIT, M. J., GREEN, S. M., HEAPHY, S., KARN, J., LOWE, A. D., SINGH, M., SKINNER, M. A. & VALERIO, R. 1989. Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro. *Proceedings of the National Academy of Sciences,* 86**,** 6925-6929.

CHIOZZINI, C. & TOSCHI, E. 2016. HIV-1 TAT and immune dysregulation in AIDS pathogenesis: a therapeutic target. *Current drug targets,* 17**,** 33-45.

DAS, A. T., HARWIG, A. & BERKHOUT, B. 2011. The HIV-1 Tat protein has a versatile role in activating viral transcription. *Journal of virology,* 85**,** 9506-9516.

COWLEY, D., GRAY, L. R., WESSELINGH, S. L., GORRY, P. R. & CHURCHILL, M. J. 2011. Genetic and functional heterogeneity of CNS-derived tat alleles from patients with HIV-associated dementia. *Journal of neurovirology,* 17**,** 70-81.

ZHANG, L., QIN, J., LI, Y., WANG, J., HE, Q., ZHOU, J., LIU, M. & LI, D. 2014. Modulation of the stability and activities of HIV-1 Tat by its ubiquitination and carboxyl-terminal region. *Cell & bioscience,* 4**,** 61.

CAMPBELL, G. R. & LORET, E. P. 2009. What does the structure-function relationship of the HIV-1 Tat protein teach us about developing an AIDS vaccine? *Retrovirology,* 6**,** 50.

KAMEOKA, M., MORGAN, M., BINETTE, M., RUSSELL, R. S., RONG, L., GUO, X., MOULAND, A., KLEIMAN, L., LIANG, C. & WAINBERG, M. A. 2002. The Tat protein of human immunodeficiency virus type 1 (HIV-1) can promote placement of tRNA primer onto viral RNA and suppress later DNA polymerization in HIV-1 reverse transcription. *Journal of virology,* 76**,** 3637-3645.

TAHIROV, T. H., BABAYEVA, N. D., VARZAVAND, K., COOPER, J. J., SEDORE, S. C. & PRICE, D. H. 2010. Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature,* 465**,** 747-751.

LI, L., DAHIYA, S., KORTAGERE, S., AIAMKITSUMRIT, B., CUNNINGHAM, D., PIRRONE, V., NONNEMACHER, M. R. & WIGDAHL, B. 2012. Impact of Tat genetic variation on HIV-1 disease. *Advances in virology,* 2012.

BAGASHEV, A. & SAWAYA, B. E. 2013. Roles and functions of HIV-1 Tat protein in the CNS: an overview. *Virology journal,* 10**,** 358.

VOGEL, B. E., LEE, S.-J., HILDEBRAND, A., CRAIG, W., PIERSCHBACHER, M. D., WONG-STAAL, F. & RUOSLAHTI, E. 1993. A novel integrin specificity exemplified by binding of the alpha v beta 5 integrin to the basic domain of the HIV Tat protein and vitronectin. *The Journal of Cell Biology,* 121**,** 461-468.

LU, H., LI, Z., XUE, Y. & ZHOU, Q. 2013. Viral–host interactions that control HIV-1 transcriptional elongation. *Chemical reviews,* 113**,** 8567-8582.

KARN, J. 2000. Tat, a novel regulator of HIV transcription and latency. *HIV sequence compendium***,** 2-18.

LASSEN, K. G., BAILEY, J. R. & SILICIANO, R. F. 2004. Analysis of human immunodeficiency virus type 1 transcriptional elongation in resting CD4+ T cells in vivo. *Journal of virology,* 78**,** 9105-9114.

BUTLER, J. E. & KADONAGA, J. T. 2002. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes & development,* 16**,** 2583-2592.

NE, E., PALSTRA, R.-J. & MAHMOUDI, T. 2018. Transcription: Insights From the HIV-1 Promoter. *International review of cell and molecular biology.* Elsevier.

ZHOU, C. & RANA, T. M. 2002. A bimolecular mechanism of HIV-1 Tat protein interaction with RNA polymerase II transcription elongation complexes. *Journal of molecular biology,* 320**,** 925-942.

NILSON, K. A. & PRICE, D. H. 2011. The role of RNA polymerase II elongation control in HIV-1 gene expression, replication, and latency. *Genetics research international,* 2011.

NATARAJAN, M., LESTER, G. M. S., LEE, C., MISSRA, A., WASSERMAN, G. A., STEFFEN, M., GILMOUR, D. S. & HENDERSON, A. J. 2013. Negative elongation factor (NELF) coordinates RNA polymerase II pausing, premature termination, and chromatin remodeling to regulate HIV transcription. *Journal of Biological Chemistry,* 288**,** 25995-26003.

ZHOU, Q. & YIK, J. H. 2006. The Yin and Yang of P-TEFb regulation: implications for human immunodeficiency virus gene expression and global control of cell growth and differentiation. *Microbiology and Molecular Biology Reviews,* 70**,** 646-659.

LIU, P.-Y., SHI, Z.-Y., TUNG, K.-C., SHYU, C.-L., CHAN, K.-W., LIU, J.-W., WU, Z.-Y., KAO, C.-C., HUANG, Y.-C. & LIN, C.-F. 2014. Antimicrobial resistance to cefotaxime and ertapenem in Enterobacteriaceae: the effects of altering clinical breakpoints. *The Journal of Infection in Developing Countries,* 8**,** 289-296.

KARN, J. & STOLTZFUS, C. M. 2012. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harbor perspectives in medicine,* 2**,** a006916.

FRANCIS, A., DI PRIMIO, C., ALLOUCH, A. & CERESETO, A. 2011. Role of phosphorylation in the nuclear biology of HIV-1. *Current medicinal chemistry,* 18**,** 2904-2912.

SANJUÁN, R. & DOMINGO-CALAP, P. 2016. Mechanisms of viral mutation. *Cellular and molecular life sciences,* 73**,** 4433-4448.

LIDSKY, P. V., ANDINO, R. & ROUZINE, I. M. 2017. Variability in viral pathogenesis: Modeling the dynamic of acute and persistent infections. *Current opinion in virology,* 23**,** 120-124.

KUROSU, T., MUKAI, T., KOMOTO, S., IBRAHIM, M. S., LI, Y. G., KOBAYASHI, T., TSUJI, S. & IKUTA, K. 2002. Human immunodeficiency virus type 1 subtype C exhibits higher transactivation activity of Tat than subtypes B and E. *Microbiology and immunology,* 46**,** 787-799.

ROY, C. N., KHANDAKER, I., FURUSE, Y. & OSHITANI, H. 2015. Molecular characterization of full-length Tat in HIV-1 subtypes B and C. *Bioinformation,* 11**,** 151.

ROSSENKHAN, R., MACLEOD, I. J., SEBUNYA, T. K., CASTRO-NALLAR, E., MCLANE, M. F., MUSONDA, R., GASHE, B. A., NOVITSKY, V. & ESSEX, M.

2013. tat Exon 1 exhibits functional diversity during HIV-1 subtype C primary infection. *Journal of virology,* 87**,** 5732-5745.

NEOGI, U., GUPTA, S., SAHOO, P. N., SHET, A., RAO, S. D., RANGA, U. & PRASAD, V. R. 2012. Genetic characterization of HIV type 1 Tat exon 1 from a southern Indian clinical cohort: identification of unique epidemiological signature residues. *AIDS research and human retroviruses,* 28**,** 1152-1156.

AN, S. F., GROVES, M., GRAY, F. & SCARAVILLI, F. 1999. Early entry and widespread cellular involvement of HIV-1 DNA in brains of HIV-1 positive asymptomatic individuals. *Journal of neuropathology and experimental neurology,* 58**,** 1156-1162.

GONZÁLEZ-SCARANO, F. & MARTÍN-GARCÍA, J. 2005. The neuropathogenesis of AIDS. *Nature reviews immunology,* 5**,** 69.

MARTÍN-GARCÍA, J., CAO, W., VARELA-ROHENA, A., PLASSMEYER, M. L. & GONZÁLEZ-SCARANO, F. 2006. HIV-1 tropism for the central nervous system: Brain-derived envelope glycoproteins with lower CD4 dependence and reduced sensitivity to a fusion inhibitor. *Virology,* 346**,** 169-179.

RUSSELL, R. A., CHOJNACKI, J., JONES, D. M., JOHNSON, E., DO, T., EGGELING, C., PADILLA-PARRA, S. & SATTENTAU, Q. J. 2017. Astrocytes resist HIV-1 fusion but engulf infected macrophage material. *Cell reports,* 18**,** 1473-1483.

ZHOU, B. Y., LIU, Y., OH KIM, B., XIAO, Y. & HE, J. 2004. Astrocyte activation and dysfunction and neuron death by HIV-1 Tat expression in astrocytes. *Molecular and Cellular Neuroscience,* 27**,** 296-305.

GRAY, L. R., BREW, B. J. & CHURCHILL, M. J. 2016. Strategies to target HIV-1 in the central nervous system. *Current Opinion in HIV and AIDS,* 11**,** 371-375.

ENE, L. 2018. Human Immunodeficiency Virus in the Brain—Culprit or Facilitator? *Infectious Diseases: Research and Treatment,* 11**,** 1178633717752687.

OHAGEN, A., DEVITT, A., KUNSTMAN, K. J., GORRY, P. R., ROSE, P. P., KORBER, B., TAYLOR, J., LEVY, R., MURPHY, R. L. & WOLINSKY, S. M. 2003. Genetic and functional analysis of full-length human immunodeficiency virus type 1 env genes derived from brain and blood of patients with AIDS. *Journal of virology,* 77**,** 12336-12345.

RAPPAPORT, J., JOSEPH, J., CROUL, S., ALEXANDER, G., VALLE, L. D., AMINI, S. & KHALILI, K. 1999. Molecular pathway involved in HIV-1-induced CNS pathology: role of viral regulatory protein, Tat. *Journal of leukocyte biology,* 65**,** 458-465.

DESHMANE, S. L., KREMLEV, S., AMINI, S. & SAWAYA, B. E. 2009. Monocyte chemoattractant protein-1 (MCP-1): an overview. *Journal of interferon & cytokine research,* 29**,** 313-326.

CUPP, C., TAYLOR, J., KHALILI, K. & AMINI, S. 1993. Evidence for stimulation of the transforming growth factor beta 1 promoter by HIV-1 Tat in cells derived from CNS. *Oncogene,* 8**,** 2231-2236.

SAWAYA, B. E., THATIKUNTA, P., DENISOVA, L., BRADY, J., KHALILI, K. & AMINI, S. 1998. Regulation of TNFα and TGFβ-1 gene transcription by HIV-1 Tat in CNS cells. *Journal of neuroimmunology,* 87**,** 33-42.

HAIJ, N. B., PLANÈS, R., LEGHMARI, K., SERRERO, M., DELOBEL, P., IZOPET, J., BENMOHAMED, L. & BAHRAOUI, E. 2015. HIV-1 Tat protein induces production of proinflammatory cytokines by human dendritic cells and monocytes/macrophages through engagement of TLR4-MD2-CD14 complex and activation of NF-κB pathway. *PloS one,* 10**,** e0129425.

BRABERS, N. & NOTTET, H. 2006. Role of the pro-inflammatory cytokines TNF-α and IL-1β in HIV-associated dementia. *European journal of clinical investigation,* 36**,** 447-458.

CHEN, P., MAYNE, M., POWER, C. & NATH, A. 1997. The Tat Protein of HIV-1 Induces Tumor Necrosis Factor-α Production IMPLICATIONS FOR HIV-1-ASSOCIATED NEUROLOGICAL DISEASES. *Journal of Biological Chemistry,* 272**,** 22385-22388.

LI, J. C.-B., YIM, H. C.-H. & LAU, A. S. 2010. Role of HIV-1 Tat in AIDS pathogenesis: its effects on cytokine dysregulation and contributions to the pathogenesis of opportunistic infection. *Aids,* 24**,** 1609-1623.

NAICKER, D. D., WERNER, L., KORMUTH, E., PASSMORE, J.-A., MLISANA, K., KARIM, S. A., NDUNG'U, T. & TEAM, C. A. I. S. 2009. Interleukin-10 promoter polymorphisms influence HIV-1 susceptibility and primary HIV-1 pathogenesis. *Journal of Infectious Diseases,* 200**,** 448-452.

BRUCHFELD, J., CORREIA-NEVES, M. & KÄLLENIUS, G. 2015. Tuberculosis and HIV coinfection. *Cold Spring Harbor perspectives in medicine***,** a017871.

WHO 2018. Tuberculosis Fact Sheet

THWAITES, G. E., DUC BANG, N., HUY DUNG, N., THI QUY, H., THI TUONG OANH, D., THI CAM THOA, N., QUANG HIEN, N., TRI THUC, N., NGOC HAI, N. & THI NGOC LAN, N. 2005. The influence of HIV infection on clinical presentation, response to treatment, and outcome in adults with tuberculous meningitis. *The Journal of infectious diseases,* 192**,** 2134-2141.

GARCÍA-MONCÓ, J. C. & RODRIGUEZ-SAINZ, A. 2018. CNS Tuberculosis and Other Mycobacterial Infections. *CNS Infections.* Springer.

FAKSRI, K., XIA, E., ONG, R. T.-H., TAN, J. H., NONGHANPHITHAK, D., MAKHAO, N., THAMNONGDEE, N., THANORMCHAT, A., PHURATTANAKORNKUL, A. & RATTANARANGSEE, S. 2018. Comparative whole-genome sequence analysis of Mycobacterium tuberculosis isolated from tuberculous meningitis and pulmonary tuberculosis patients. *Scientific reports,* 8**,** 4910.

PATEL, V. B., THERON, G., LENDERS, L., MATINYENA, B., CONNOLLY, C., SINGH, R., COOVADIA, Y., NDUNG'U, T. & DHEDA, K. 2013. Diagnostic accuracy of quantitative PCR (Xpert MTB/RIF) for tuberculous meningitis in a high burden setting: a prospective study. *PLoS medicine,* 10**,** e1001536.

BE, N. A., KIM, K. S., BISHAI, W. R. & JAIN, S. K. 2009. Pathogenesis of central nervous system tuberculosis. *Current molecular medicine,* 9**,** 94-99.

SEIPONE, I. D., SINGH, R., PATEL, V. B., SINGH, A., GORDON, M. L., MUEMA, D. M., DHEDA, K. & NDUNG'U, T. 2018. Tuberculous meningitis is associated with higher cerebrospinal HIV-1 viral loads compared to other HIV-1-associated meningitides. *PloS one,* 13**,** e0192060.

THWAITES, G., FISHER, M., HEMINGWAY, C., SCOTT, G., SOLOMON, T. & INNES, J. 2009. British Infection Society guidelines for the diagnosis and treatment of tuberculosis of the central nervous system in adults and children. *Journal of Infection,* 59**,** 167-187.

MORRIS, L., SILBER, E., SONNENBERG, P., EINTRACHT, S., NYOKA, S., LYONS, S. F., SAFFER, D., KOORNHOF, H. & MARTIN, D. J. 1998. High human immunodeficiency virus type 1 RNA load in the cerebrospinal fluid from patients with lymphocytic meningitis. *Journal of infectious diseases,* 177**,** 473-476.

THOMPSON, J. D., HIGGINS, D. G. & GIBSON, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research,* 22**,** 4673-4680.

SAITOU, N. & NEI, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution,* 4**,** 406-425.

KUMAR, S., STECHER, G., LI, M., KNYAZ, C. & TAMURA, K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular biology and evolution,* 35**,** 1547-1549.

HALL, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic acids symposium series, 1999. [London]: Information Retrieval Ltd., c1979-c2000., 95-98.

SARZOTTI-KELSOE, M., BAILER, R. T., TURK, E., LIN, C.-L., BILSKA, M., GREENE, K. M., GAO, H., TODD, C. A., OZAKI, D. A. & SEAMAN, M. S. 2014. Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. *Journal of immunological methods,* 409**,** 131-146.

KUMAR, S. & NEI, M. 2000. *Molecular evolution and phylogenetics*, Oxford University Press New York.

CROOKS, G. E., HON, G., CHANDONIA, J.-M. & BRENNER, S. E. 2004. WebLogo: a sequence logo generator. *Genome research,* 14**,** 1188-1190.

RAO, V. R., NEOGI, U., TALBOOM, J. S., PADILLA, L., RAHMAN, M., FRITZ-FRENCH, C., GONZALEZ-RAMIREZ, S., VERMA, A., WOOD, C. & RUPRECHT, R. M. 2013. Clade C HIV-1 isolates circulating in Southern Africa exhibit a greater frequency of dicysteine motif-containing Tat variants than those in Southeast Asia and cause increased neurovirulence. *Retrovirology,* 10**,** 61.

RANGA, U., SHANKARAPPA, R., SIDDAPPA, N. B., RAMAKRISHNA, L., NAGENDRAN, R., MAHALINGAM, M., MAHADEVAN, A., JAYASURYAN, N., SATISHCHANDRA, P. & SHANKAR, S. K. 2004. Tat protein of human immunodeficiency virus type 1 subtype C strains is a defective chemokine. *Journal of virology,* 78**,** 2586-2590.

WONG, J. K., CAMPBELL, G. R. & SPECTOR, S. A. 2010. Differential induction of interleukin-10 in monocytes by HIV-1 clade B and clade C Tat proteins. *Journal of Biological Chemistry***,** jbc. M110. 120840.

CAMPBELL, G. R., WATKINS, J. D., LORET, E. P. & SPECTOR, S. A. 2011. Differential induction of rat neuronal excitotoxic cell death by human immunodeficiency virus type 1 clade B and C tat proteins. *AIDS research and human retroviruses,* 27**,** 647-654.

CAMPBELL, G. R., PASQUIER, E., WATKINS, J., BOURGAREL-REY, V., PEYROT, V., ESQUIEU, D., BARBIER, P., DE MAREUIL, J., BRAGUER, D. & KALEEBU, P. 2004. The glutamine-rich region of HIV-1 Tat protein Involved in T-cell apoptosis. *Journal of Biological Chemistry*.

SIVAKUMARAN, H., WANG, B., GILL, M. J., BECKHOLDT, B., SAKSENA, N. K. & HARRICH, D. 2007. Functional relevance of nonsynonymous mutations in the HIV-1 tat gene within an epidemiologically-linked transmission cohort. *Virology journal,* 4**,** 107.

ENDO-MUNOZ, L., WARBY, T., HARRICH, D. & MCMILLAN, N. A. 2005. Phosphorylation of HIV Tat by PKR increases interaction with TAR RNA and enhances transcription. *Virology journal,* 2**,** 17.

FISHMAN, R. A. 1992. *Cerebrospinal fluid in diseases of the nervous system*, Saunders Philadelphia.

RONSARD, L., GANGULI, N., SINGH, V. K., MOHANKUMAR, K., RAI, T., SRIDHARAN, S., PAJANIRADJE, S., KUMAR, B., RAI, D. & CHAUDHURI, S. 2017. Impact of Genetic Variations in HIV-1 Tat on LTR-Mediated Transcription via TAR RNA Interaction. *Frontiers in microbiology,* 8**,** 706.

PAN, C., KIM, J., CHEN, L., WANG, Q. & LEE, C. 2006. The HIV positive selection mutation database. *Nucleic acids research,* 35**,** D371-D375.

MISHRA, M., VETRIVEL, S., SIDDAPPA, N. B., RANGA, U. & SETH, P. 2008. Clade-specific differences in neurotoxicity of human immunodeficiency virus-1 B and C Tat of human neurons: significance of dicysteine C30C31 motif. *Annals of neurology,* 63**,** 366-376.

AMMOSOVA, T., BERRO, R., JEREBTSOVA, M., JACKSON, A., CHARLES, S., KLASE, Z., SOUTHERLAND, W., GORDEUK, V. R., KASHANCHI, F. & NEKHAI, S. 2006. Phosphorylation of HIV-1 Tat by CDK2 in HIV-1 transcription. *Retrovirology,* 3**,** 78.

BOVEN, L. A., NOORBAKHSH, F., BOUMA, G., VAN DER ZEE, R., VARGAS, D. L., PARDO, C., MCARTHUR, J. C., NOTTET, H. S. & POWER, C. 2007. Brain-derived human immunodeficiency virus-1 Tat exerts differential effects on LTR transactivation and neuroimmune activation. *Journal of neurovirology,* 13**,** 173-184.

RUIZ, A. P. 2016. *A Naturally occurring polymorphism in the HIV-1 Tat protein basic domain modulates its uptake by bystander cells and their subsequent inflammatory output*, Yeshiva University.

KAMORI, D. & UENO, T. 2017. HIV-1 Tat and viral latency: What we can learn from naturally occurring sequence variations. *Frontiers in microbiology,* 8**,** 80.

BROWN, S. P., CORNFORTH, D. M. & MIDEO, N. 2012. Evolution of virulence in opportunistic pathogens: generalism, plasticity, and control. *Trends in microbiology,* 20**,** 336-342.

PATEL, V. B., SINGH, R., CONNOLLY, C., KASPROWICZ, V., NDUNG'U, T. & DHEDA, K. 2011. Comparative utility of cytokine levels and quantitative RD-1-specific T cell responses for the rapid immunodiagnosis of TB meningitis. *Journal of clinical microbiology***,** JCM. 01128-11.

BUTLER, D. M., PACOLD, M. E., JORDAN, P. S., RICHMAN, D. D. & SMITH, D. M. 2009. The efficiency of single genome amplification and sequencing is improved by quantitation and use of a bioinformatics tool. *Journal of virological methods,* 162**,** 280-283.

Appendix

6. Appendix

6.1 Baseline Clinical Data

Table 6:1 Baseline Clinical Data of TBM and non-TBM study participants

TBM				
ARV NAÏVE				
LAB NO:	Plasma Viral load (copies/ml)	CD4+ Count (copies/µl)	CSF Viral load (copies/ml)	Lymphocyte Count (copies/µl)
B98	1489556	71	672812	111,6
X214	943314	94	6129219	10,72
X223	1710603	78	2599081	32
X233	163590	32	> 10000000	6,4
X195	1008150	200	1462385	2,8
X201	99078	14	713295	2
B103	38409	50	8133	4
X215	94407	274	1264491	5,68
X221	243480	72	630291	2,4
A96	109677	441	35766	6,56
A104	866178	347	53640	4,8
A105	28921	677	11454	22,08
A68	66041	85	2975556	13,44
A111	960561	199	262777	4,8
A130	519029	158	30175	7,44
X205	390650	117	75258	11,44
Non-TBM				
ARV NAÏVE				
LAB NO:	Plasma Viral load (copies/ml)	CD4+ Count (copies/µl)	CSF Viral load (copies/ml)	Lymphocyte Count (copies/µl)
X229	102025	36	14230	0,24
B129	213260	532	5109	5,92
X235	4317	113	10631	2,16

6.2 HIV-1C DNA sequences from TBM cohort

>TatCConsensus_1-216
ATGGAGCCAGTAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAAAACT
GCTTGTAATAAGTGTTATTGTAAACACTGTAGCTATCATTGTCTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCGCTCCTCCA
AGCAGTGAGGATCATCAAAATCCTATATCAAAGCAA

>A96_PLASMA_1-216
ATGGAGCTAGTAGATCCTGACCTAGCGCCCTGGCACCATCCAGGAAGTCAGCCTAAAACT
CCTTGCAATAAATGCTATTGTAAATGCTGTAGTTATCATTGTTTAGCTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAACGACGAAGCACTCCTCCA
AGCAGTGAGAGCCATCAAAATCATATATCAAAGCAG

>A96_CSF_1-216
ATGGAGCTAGTAGACCCTGACCTAGCGCCCTGGCACCATCCAGGAAGTCAGCCTAAAACT
CCTTGCAATAAATGCTATTGTAAACACTGTAGCTATCATTGTTTAGCTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAACGACGAAGCACTCCTCCA
AGCAGTGAGAATCATCAAAATCCTATATCAAAGCAG

>A103_PLASMA
ATGGAGCCAGTAGATCYTAACYTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTACAACT
CCTTGTAATACGTGMTACTGTAAACAYTGCAGCTACCATTGTCTAGTTTGCTTTCAGACC
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCRCTCCTCCA
AGCAGTGAGGATCATCAAAATCCTCTACCAAASCAG

>A103_CSF
ATGGAGCCAGTAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTACAACT
CCTTGTAATACGTGCTACTGTAAACACTGCAGCTACCATTGTCTAGTTTGCTTTCAGACC
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCACTCCTCCA
AGCAGTGAGGATCATCAAAATCCTCTACCAAAGCAG

>A111_PLASMA
ATGGAGCCAGTAGATACTAAACTAGAGCCATGGAATCACCCAGGGAGTCAGCCTAGAACT
GCTTGTAATAAGTGTTATTGCAAAAGATGYTGCTACCACTGTTTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCACTCCTCCA
AGCAGTGAGGATCATCAAAATCTTGTATCAAAGCAG

>A111_CSF
ATGGAGCCAGTAGATACTAAACTAGAGCCATGGAATCACCCAGGGAGTCAGCCTAGAACT
GCTTGTAATAAGTGTTATTGCAAAAGATGYTGCTACCACTGTTTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACRAASCACTCCTCCA
AGCAGTGAGGATCATCAAAATCTTGTATCAAAGCAG

>A130_PLASMA
ATGGAGCCAATAGATCCTAACMTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAAAACC
CCTTGTACTAAATGCTATTGTAAATACTGTAGCTACCACTGTCTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACARCGACGAAGCACTCCTCCA
AGCAGTGAGGATCATCARAATCTTATATCAAAGCAG

>A130_CSF
ATGGAGCCAATAGATCCTAACATAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAAAACC
CCTTGTACTAAATGCTATTGTAAATACTGTAGCTACCACTGTCTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACARCGACGAAGCACTCCTCCA
AGCAGTGAGGATCATCAGAATCTTATATCAAAGCAG

>A68_PLASMA
ATGGAGCCAGTAGAYCCTAGCCTAGAGCCCTGGAAMCAYCCAGGAAGTCAGCCTCAAACT
GCTTGTAATAATTGTTATTGTAARAGATGTAGCTATCATTGTCTAGTTTGCTTTCAGAAA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGTGGAGACAGCGACGAGGACATCCTCCG
AGCAGTGAGAGCCATCAAAATCYTGTATCAAAGCAG

>A68_CSF
ATGGAGCCAGTAGAYCCTAGCCTAGAGCCCTGGAAMCAYCCAGGAAGTCAGCCTCAAACT
GCTTGTAATAATTGTTATTGTAARAGATGTAGCTATCATTGTCTAGTTWGCTTTCAGAAA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGTGGAGACAGCGACGAGGACATCCTCCG
AGCAGTGAGAGCCATCAAAATCYTGTATCAAAGCAG

>B98_PLASMA
ATGGAGCCAGTAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAAAACT
GCTTGTAATCCATGCTATTGTAAACGTTGTAGCTATCATTGTTTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTACTATGGCAGGAAGAAGCGGAGACAACGACGCAGAGCTCCTCCG
AGCAGTGAGGATCATCAAGATCTTGTATCAAAGCAG

>B98_CSF
ATGGAGCCAGTAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAACCTAAAACT
GCTTGTAATCCATGCTATTGTAAAAAATGTAGCTATCATTGTTTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTACTATGGCAGGAAGAAGCGGAGACAACGACGCAGAGCTCCTCCG
AGCAGTGAGGATCATCAAGATCTTATATCAAAGCAG

>A104_PLASMA
ATGGAGCCAGTAGATCCTGACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAAAACT
CCTTGTAATAAGTGCTTTTGTAAACACTGTAGTTATCATTGTCTAGTTTGCTTYCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGTCAGCGACGAAGCACTCCTTCA
AGCAGTAAAGATCATCAAAATCCTGTACCAAAGCAG

>A104_CSF
ATGGAGCCAGTAGATCCTGACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAAAACT
CCTTGTAATAAGTGCTTTTGTAAACACTGTAGTTATCATTGTCTAGTTTGCTTYCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGTCAGCGACGAAGCACTCCTTCA
AGCAGTAAAGATCATCAAAATCCTGTACCAAAGCAG

>A105_PLASMA
ATGGAGCCAATAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTACAWCT
CCTTGTAATTCTTGTTATTGCAAAAAGTGTAGCTATCATTGTCTAGTTTGCTTTCATACG
AAAGGCTTAGGCATTTCCTATGSCAGGAAGAAGCGGAGACAGCGACGAAGCGCTCCTCCA
GGCAGTAAGGATCATCAAAATCCTCTACCAAAGCAG

>A105_CSF
ATGGAGSCAATAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTACACCT
CCTTGTAATTCTTGTTATTGCAAAAAGTGTAGCTATCATTGTCTAGTTTGCTTTCAGACG
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCGCTCCTCCA
GGCAGTAAGGATCATCAAAATCCTCTACCAAAGCAG

>A129_PLASMA
ATGGAGCCAGTAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCCAAAACT
CCTTGTAATAAGTGCTATTGTAAACACTGCAGCTATCATTGTTTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCACTCCTCCA
AGCAGTGAGGATCATCAAAATCTTATATCAAAGCAG

>A129_CSF
ATGGAGCCAGTAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCCAAAWCT
CCTTGTAATAAGTGCTATTGTAAACACTGCAGCTATCATTGTTTAGTTTGYTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGMGGAGACAGCGACGAAGCACTCCTCCA
AGCAGTGAGGATCATCAAAATCTTATATCAAAGCAG

>X195_PLASMA
ATGGAGCCAATAGATCCTAAYCTAGAGCCCTGGAACCACCCAGGRAGTCAGCCTGAAACT
GCTTGTAATACTTGTTATTGTAAATACTGTAGTTTCCATTGTCTAGCTTGCTTTCAGAAA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAAACAGCGACGAAGCGCTCCTCCT
AGCAGTGAGGACCATCAAAATCCTGTACCAAAGCAG

>X195_CSF
ATGGAGCCAATAGATCCTAACCTAGAGCCCTGGAACCACCCAGGRAGTCAGCCTGAAACT
GCTTGTAATACTTGTTATTGTAAATACTGTAGTTTCCATTGTCTAGCTTGCTTTCAGAAA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAARCAGCGACGAAGCGCTCCTCCT
AGCAGTGAGGACCATCAAAATCCTGTACCAAAGCAG

>X201_PLASMA
ATGGAGCCAGTAGATCCTAACCTAGAGCCCTGGAATCATCCAGGAAGTCAGCCTAAAACT
CCTTGTAATAAATGCTTTTGTAAAAGCTGTAGCTACCACTGTCTAGTTTGCTTTCAGACG
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCACTCCTTCA
AGCAGTGAGGATCATCAAAATCCTATATCAAAGCAG

>X201_CSF
ATGGAGCCAGTAGATCCTAACCTAGAGCCCTGGAATCATCCAGGAAGTCAGCCTAAAACT
CCTTGTAATAAATGCTTTTGTAAAAGCTGTAGCTACCACTGTCTAGTTTGCTTTCAGACG
AAAGGCTTAGGCTTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCACTCCTTCAA
GCAGTGAGGATCATCAAAATCCTATATCAAAGCAG

>X205_PLASMA
ATGGAGCCAATAGATCCTAACCTAGAACCCTGGAATCATCCAGGAAGTCAGCCTGAAACT
CCTTGTAATACATGTTATTGCAAAAAATGTTGCTATCATTGCTTAGCTTGCTTTCAGACG
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCGCTCCTCAG
CGCAGTGAGGATCATCAAGATCTTGTATCAAAGCAG

>X205_CSF
ATGGAGCCAATAGATCCTAACCTAGAACCCTGGAATCATCCAGGAAGTCAGCCTGAAACT
CCTTGTAATAAATGTTATTGTAAAAAATGCTGCTATCATTGCTTAGCTTGCTTTCAGACG
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCGCTCCTCAG
AGCAGTGAGGATCATCAAGATCTTGTATCAAAGCAG

>X214_PLASMA
ATGGAGCCAATAGATCCTAAYTTAGAACCTTGGAACCACCCAGGAAGTCAGCCTCAAACT
CCTTGTACTAAGTGCTTTTGTAAATTCTGTAGCTATCATTGTCTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCACTCCTCCA
AGCAGTGAGGATCATCAAAATCTTATATCAAAGCAG

>X214_CSF
ATGGAGCCADTAGATCCTAACTTAGAACCTTGGAACCACCCAGGAAGTCAGCCTCAAACT
CCTTGTACTAAGTGCTTTTGTAAATTCTGTAGCTATCATTGTCTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCACTCCTCCA
AGCAGTGAGGATCATCAAAATCTTATATCAAAGCAG

>X215_PLASMA
ATGGAGCCAGTAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAARACT
GCTTGTAGTCCGTGTTATTGCAAAAAATGTAGCTATCATTGTCTAGTTTGCTTTCAGACA

AAAGGCTTAGGCATTTACTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCGCTCCTCCA
AGCAATAAGGACCATCAKAATCCTCTATCAAAGCAG

>X215_CSF
ATGGAGCCAGTAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAAGACT
GCTTGTAGTCCGTGTTATTGCAAAAAATGTAGCTATCATTGTCTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCGCTCCTCCA
AGCARTAARGAYCATCAAAATCCTCTATCAAAGCAG

>X221_PLASMA
ATGGAGCCAGTAGATCCTAGCCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAAAACT
GCTTGCAATAATTGCTATTGTAAAWGCTGTAGCTACCATTGTCTAGTTTGCTTTCAGAMA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAACGAMGAMGCACTCCTCCA
AGYRGTGAGGATCATCAAAATCYTATATCAAAGCAG

>X221_CSF
ATGGAGCCAGTAGATCCTAGCCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAAAACT
GCTTGCAATAATTGCTATTGTAAATYCTGTAGCTACCATTGTCTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGACGCACTCCTCCA
AGCAGTGAGGATCATCAAAATCYTATATCAAAGCAG

>X223_PLASMA
ATGGAGCCYGTAGATTCTAACCTAGAGCCCTGGAATCATCCAGGAAGTCAGCCAAAAACT
GCTTGTAACGGTTGCTATTGTAAAAGATGCTGCTACCATTGTCTAGTTTGCTTTCAGAAA
AAAGGCTTARGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCGCTCCTCCA
AGCAGKGAGGATCATCAAAATCCTATATCAAAGCAG

>X223_CSF
ATGGAGCCCGTAGATTCTAACCTAGAGCCCTGGAATCATCCAGGAAGTCAGCCAAAAACT
GCTTGTAACGGTTGCTATTGTAAAAGATGCTGCTACCATTGTYTAGTTTGCTTTCAGAAA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCGCTCCTCCA
AGCAGTGAGGATCATCAAAATCCTRTATCAAAGCAG

>X229_PLASMA
ATGGAGCCAGTAGATCCTAACTTAGAGCCCTGGAACCATCCAGGAAGTCAGCCCACAACT
CCTTGTAATAAATGCTATTGTAAACACTGTAGCTACCATTGTCTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCACTCCTCCA
AGCAGTGAGGATCATCAAAATCCTGTATCAAAGCAG

>X229_CSF
ATGGAGCCAGTAGATCCTAACTTAGAGCCCTGGAACCATCCAGGAAGTCAGCCCACAACT
CCTTGTAATAAATGCTATTGTAAACACTGTAGCTACCATTGTCTAGTTTGCTTTCAGACA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCACTCCTCCA
AGCAGTGAGGATCATCAAAATCCTGTATCAAAGCAG

>X233_PLASMA
ATGGAGCCAGTAGATCCTAACCTAGAGYCCTGGAACCATCCAGGAAGTCAGCCTAACACT
CCTTGTACTAAGTGCTATTGTAAAAAATGTAGCTATCATTGTCTAGTTTGCTTTCAGAAA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACGGCGACGAAGCCCTTCTCCA
AGCAGTAAGGATCATCAAAATCTTGTATCAAAGCAG

>X233_CSF
ATGGAGCCAGTAGATCCTAACCTAGAGTCCTGGAACCATCCAGGAAGTCAGCCTAACACT
CCTTGTACTAAGTGCTATTGTAAAAAATGTAGCTATCATTGTCTAGTTTGCTTTCAGAAA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACGGCGACGAAGCCCTTCTCCA
AGCAGTAAGGAYCATCAAAATCTTGTATCAAAGCAG

>X235_PLASMA
ATGGAGCCAATAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAGTACT
CCTTGCAATAACCGCTATTGTAAACACTGTAGCTACCATTGTTTAGCTTGCTTTCAGAGA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGCAGCACTCCTCCA
AGCAATAAGGATCATCAAAATCCTGTATCAAAGCAG

>X235_CSF
ATGGAGCCAATAGATCCTAACCTAGAGCCCTGGAACCATCCAGGAAGTCAGCCTAGTACT
CCTTGCAATAACTGCTATTGTAAACACTGTAGCTACCATTGTCTAGCTTGCTTTCAGAGA
AAAGGCTTAGGCATTTCCTATGGCAGGAAGAAGCGGAGACAGCGACGCAGCACTCCTCCA
AGCAATAAGGATCATCAAAATCCTGTATCAAAGCAG

Turnitin Report

Genetic and functional diversity of central nervous system (CNS) derived human immunodeficiency virus type 1 (HIV-1) tat from tuberculous meningitis (TBM) patients

ORIGINALITY REPORT

15%	11%	14%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Rossenkhan, R., I. J. MacLeod, T. K. Sebunya, E. Castro-Nallar, M. F. McLane, R. Musonda, B. A. Gashe, V. Novitsky, and M. Essex. "tat Exon 1 Exhibits Functional Diversity during HIV-1 Subtype C Primary Infection", Journal of Virology, 2013. Publication	1%
2	"Proceedings of the Frontiers in Retrovirology Conference 2018", Retrovirology, 2018 Publication	1%
3	journals.plos.org Internet Source	1%
4	Ikanyeng D. Seipone, Ravesh Singh, Vinod B. Patel, Avashna Singh et al. "Tuberculous meningitis is associated with higher cerebrospinal HIV-1 viral loads compared to other HIV-1-associated meningitides", PLOS ONE, 2018 Publication	<1%

5	146.230.128.141 Internet Source	<1%
6	Social Disparities in Health and Health Care, 2016. Publication	<1%
7	www.mdpi.com Internet Source	<1%
8	Larance Ronsard, Sneh Lata, Jyotsna Singh, Vishnampettai G. Ramachandran, Shukla Das, Akhil C. Banerjea. "Molecular and Genetic Characterization of Natural HIV-1 Tat Exon-1 Variants from North India and Their Functional Implications", PLoS ONE, 2014 Publication	<1%
9	www.pubmedcentral.nih.gov Internet Source	<1%
10	file.scirp.org Internet Source	<1%
11	Schiralli Lester, Gillian M., and Andrew J. Henderson. "Mechanisms of HIV Transcriptional Regulation and Their Contribution to Latency", Molecular Biology International, 2012. Publication	<1%
12	paduaresearch.cab.unipd.it Internet Source	<1%

13	orca.cf.ac.uk Internet Source	<1%
14	www.biomedcentral.com Internet Source	<1%
15	scholar.sun.ac.za Internet Source	<1%
16	Monique R Ferguson, Daniel R Rojo, Jana J von Lindern, William A O'Brien. "HIV-1 replication cycle", Clinics in Laboratory Medicine, 2002 Publication	<1%
17	"Handbook of Neurotoxicity", Springer Nature America, Inc, 2014 Publication	<1%
18	www.diss.fu-berlin.de Internet Source	<1%
19	bioinformation.net Internet Source	<1%
20	tel.archives-ouvertes.fr Internet Source	<1%
21	James Chun-bong Li. "Role of HIV-1 Tat in AIDS pathogenesis: its effects on cytokine dysregulation and contributions to the pathogenesis of opportunistic infection :", AIDS, 07/2010	<1%

Publication

22	Li, Luna, Satinder Dahiya, Sandhya Kortagere, Benjamas Aiamkitsumrit, David Cunningham, Vanessa Pirrone, Michael R. Nonnemacher, and Brian Wigdahl. "Impact of Tat Genetic Variation on HIV-1 Disease", Advances in Virology, 2012. Publication	<1%
23	www.retrovirology.com Internet Source	<1%
24	Vinod B. Patel, Grant Theron, Laura Lenders, Brian Matinyena et al. "Diagnostic Accuracy of Quantitative PCR (Xpert MTB/RIF) for Tuberculous Meningitis in a High Burden Setting: A Prospective Study", PLoS Medicine, 2013 Publication	<1%
25	www.scribd.com Internet Source	<1%
26	pure.uva.nl Internet Source	<1%
27	circres.ahajournals.org Internet Source	<1%
28	Marcella Sarzotti-Kelsoe, Robert T. Bailer, Ellen Turk, Chen-li Lin et al. "Optimization and validation of the TZM-bl assay for standardized	<1%

assessments of neutralizing antibodies against HIV-1", Journal of Immunological Methods, 2014
Publication

29	Karn, J.. "Tackling tat", Journal of Molecular Biology, 19991022 Publication	<1%
30	ulspace.ul.ac.za Internet Source	<1%
31	Verhoef, K.. "Determination of the Minimal Amount of Tat Activity Required for Human Immunodeficiency Virus Type 1 Replication", Virology, 19971027 Publication	<1%
32	www.hiv.lanl.gov Internet Source	<1%
33	Thomas, Mary C., and Cheng-Ming Chiang. "The General Transcription Machinery and General Cofactors", Critical Reviews in Biochemistry and Molecular Biology, 2006. Publication	<1%
34	www.ncbi.nlm.nih.gov Internet Source	<1%
35	www.nature.com Internet Source	<1%

www.mediander.com

36	Internet Source	<1%
37	www.csus.edu Internet Source	<1%
38	qmro.qmul.ac.uk Internet Source	<1%
39	M. E. Garber. "CDK9 Autophosphorylation Regulates High-Affinity Binding of the Human Immunodeficiency Virus Type 1 Tat-P-TEFb Complex to TAR RNA", Molecular and Cellular Biology, 09/15/2000 Publication	<1%
40	www.faqs.org Internet Source	<1%
41	discovery.ucl.ac.uk Internet Source	<1%
42	Smyth, A., H. M. Reid, A. H. Baker, and H. McGlynn. "Modifications of the radiosensitivity of a renal cancer cell line as a consequence of stable TIMP-1 overexpression", International Journal of Radiation Biology, 2007. Publication	<1%
43	arizona.openrepository.com Internet Source	<1%
44	www.taqpolymerase.org	

	Internet Source	<1%
45	Vassil St. Georgiev. "Introduction", National Institute of Allergy and Infectious Diseases NIH, 2009 Publication	<1%
46	serval.unil.ch Internet Source	<1%
47	www.freshpatents.com Internet Source	<1%
48	V. B. Patel. "Cerebrospinal T-Cell Responses Aid in the Diagnosis of Tuberculous Meningitis in a Human Immunodeficiency Virus- and Tuberculosis-Endemic Population", American Journal of Respiratory and Critical Care Medicine, 08/15/2010 Publication	<1%
49	scholarbank.nus.edu.sg Internet Source	<1%
50	Sylvain Egloff, Cécilia Studniarek, Tamás Kiss. "7SK small nuclear RNA, a multifunctional transcriptional regulatory RNA with gene-specific features", Transcription, 2017 Publication	<1%
51	www.paradeclassroom.com Internet Source	<1%

52	oro.open.ac.uk Internet Source	<1%
53	peds.oxfordjournals.org Internet Source	<1%
54	www.lume.ufrgs.br Internet Source	<1%
55	Mahkonen, Anna, Heli Putaala, Harri Mustonen, Nina Rautonen, and Pauli Puolakkainen. "Lactobacillus acidophilus 74-2 and Butyrate Induce Cyclooxygenase (COX)-1 Expression in Gastric Cancer Cells", Immunopharmacology and Immunotoxicology, 2008. Publication	<1%
56	www.tandfonline.com Internet Source	<1%
57	open.uct.ac.za Internet Source	<1%
58	docs.di.fc.ul.pt Internet Source	<1%
59	Melanie Denizot, Vincent G., Shiril Kumar, Jim W., Philippe Gasque. "Chapter 3 Virus-Induced Encephalitis and Innate Immune Responses – A Focus on Emerging or Re-Emerging Viruses", InTech, 2011 Publication	<1%

60	archiv.ub.uni-heidelberg.de Internet Source	<1%
61	T. Rinsoz. "Exposure to Bioaerosols in Poultry Houses at Different Stages of Fattening; Use of Real-time PCR for Airborne Bacterial Quantification", Annals of Occupational Hygiene, 06/11/2008 Publication	<1%
62	Daniel Garin, Christophe Peyrefitte, Jean-Marc Crance, Alain Le Faou, Alain Jouan, Michèle Bouloy. "Highly sensitive Taqman® PCR detection of Puumala hantavirus", Microbes and Infection, 2001 Publication	<1%
63	www.jourlib.org Internet Source	<1%
64	Patrycja Lech, Nikunj V. Somia. "Retrovirus Vectors", S. Karger AG, 2008 Publication	<1%
65	Edna Maria Vissoci Reiche, Andrea Name Colado Simao. "Chapter 2 The Role of Human Immunodeficiency Virus Type 1 (HIV-1) Proteins and Antiretroviral Drug Therapy in HIV-1-Induced Oxidative Stress", InTech, 2011 Publication	<1%
	Jean K Carr. "Viral diversity as a challenge to	

66	HIV-1 vaccine development", Current Opinion in HIV and AIDS, 07/2006 Publication	<1%
67	hs.hu.wikimiki.org Internet Source	<1%
68	bloodjournal.hematologylibrary.org Internet Source	<1%
69	hms.harvard.edu Internet Source	<1%
70	Orville Heslop. "Chapter 14 Genetic Diversity and the Human Immunodeficiency Virus Type-1: Implications and Impact", InTech, 2012 Publication	<1%
71	janelia.org Internet Source	<1%
72	mdpi.com Internet Source	<1%
73	genome.cshlp.org Internet Source	<1%
74	Lucile Espert. "Autophagy in HIV-Induced T Cell Death", Current Topics in Microbiology and Immunology, 2009 Publication	<1%
75	www.einj.org Internet Source	<1%

76	shodhganga.inflibnet.ac.in Internet Source	<1%
77	Won Tae Kim, Ye Hwan Kim, Pildu Jeong, Sung-Pil Seo et al. "Urinary cell-free nucleic acid IQGAP3: a new non-invasive diagnostic marker for bladder cancer", Oncotarget, 2018 Publication	<1%
78	www.jstage.jst.go.jp Internet Source	<1%
79	Jennifer D Watkins. "Homonuclear 1H NMR and circular dichroism study of the HIV-1 Tat Eli variant", Retrovirology, 2008 Publication	<1%
80	Alon Herschhorn. "Retroviral reverse transcriptases", Cellular and Molecular Life Sciences, 04/01/2010 Publication	<1%
81	S. Misumi. "Three Isoforms of Cyclophilin A Associated with Human Immunodeficiency Virus Type 1 Were Found by Proteomics by Using Two-Dimensional Gel Electrophoresis and Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry", Journal of Virology, 09/03/2002 Publication	<1%
82	"Human Immunodeficiency Virus (HIV)",	<1%

	Transfusion Medicine and Hemotherapy, 2016 Publication	
83	spandidos-publications.com Internet Source	<1%
84	docplayer.net Internet Source	<1%
85	www.spandidos-publications.com Internet Source	<1%
86	Markus Schneider, Aylin Yigitliler, Frank Stubenrauch, Thomas Iftner. "Cottontail Rabbit Papillomavirus E1 and E2 Proteins Mutually Influence Their Subcellular Localizations", Journal of Virology, 2018 Publication	<1%
87	academic.oup.com Internet Source	<1%
88	Enan, Khalid A, Takeshi Nabeshima, Toru Kubo, Corazon C Buerano, Abdel Rahim El Hussein, Isam M Elkhidir, Eltahir AG Khalil, and Kouichi Morita. "Survey of causative agents for acute respiratory infections among patients in Khartoum- State, Sudan, 2010--2011", Virology Journal, 2013. Publication	<1%
89	thesis.library.caltech.edu Internet Source	<1%

90	Hult, Britta, Gursharan Chana, Eliezer Masliah, and Ian Everall. "Neurobiology of HIV", International Review of Psychiatry, 2008. Publication	<1%
91	Bonini, Carlo, Lucia Chiummiento, Margherita De Bonis, Nadia Di Blasio, Maria Funicello, Paolo Lupattelli, Rocco Pandolfo, Francesco Tramutola, and Federico Berti. "Synthesis of New Thienyl Ring Containing HIV-1 Protease Inhibitors: Promising Preliminary Pharmacological Evaluation against Recombinant HIV-1 Proteases§", Journal of Medicinal Chemistry, 2010. Publication	<1%
92	doc.rero.ch Internet Source	<1%
93	"Researchers from University of Cape Town Detail Findings in HIV/AIDS (Isolation and characterization", AIDS Weekly, Feb 23 2015 Issue Publication	<1%
94	megasoftware.net Internet Source	<1%
95	www.oncotarget.com Internet Source	<1%
96	www.plosone.org	

	Internet Source	<1%
97	superfund.berkeley.edu Internet Source	<1%
98	Jenkins, T.M.. "Phylogeography illuminates maternal origins of exotic Coptotermes gestroi (Isoptera: Rhinotermitidae)", Molecular Phylogenetics and Evolution, 200703 Publication	<1%
99	dm5migu4zj3pb.cloudfront.net Internet Source	<1%
100	www.allindianpatents.com Internet Source	<1%
101	prr.hec.gov.pk Internet Source	<1%
102	Mitra, P., L. A. Pereira, Y. Drabsch, R. G. Ramsay, and T. J. Gonda. "Estrogen receptor-Â recruits P-TEFb to overcome transcriptional pausing in intron 1 of the MYB gene", Nucleic Acids Research, 2012. Publication	<1%
103	www.selleck.cn Internet Source	<1%
104	Patel, J.. "HIV-1 Tat-coated nanoparticles result in enhanced humoral immune responses and	<1%

neutralizing antibodies compared to alum adjuvant", Vaccine, 20060424
Publication

105	Shojania, Shaheen, and Joe D. O'Neil. "Order from Disorder: Structure, Function, and Dynamics of the HIV-1 Transactivator of Transcription", Flexible Viruses Structural Disorder in Viral Proteins, 2011. Publication	<1%
106	academicjournals.org Internet Source	<1%
107	S. WANARATANA. "The potential of house flies to act as a vector of avian influenza subtype H5N1 under experimental conditions", Medical and Veterinary Entomology, 03/2011 Publication	<1%
108	rnajournal.cshlp.org Internet Source	<1%
109	content.karger.com Internet Source	<1%
110	bv.fapesp.br Internet Source	<1%
111	Sanchez-Duffhues, G.. "Denbinobin, a naturally occurring 1,4-phenanthrenequinone, inhibits HIV-1 replication through an NF-@kB-dependent pathway", Biochemical	<1%

Pharmacology, 20081115
Publication

112	Temin, Howard M.. "Retroviruses and evolution", Cell Biophysics, 1986. Publication	<1%
113	www.biochemj.org Internet Source	<1%
114	Helio A., Ana F., Paula M., Ana S.. "Chapter 12 Lentiviral Gene Therapy Vectors: Challenges and Future Directions", InTech, 2013 Publication	<1%
115	prezi.com Internet Source	<1%
116	Mediouni, Sonia, Maria Cecilia Garibaldi Marcondes, Courtney Miller, Jay P. McLaughlin, and Susana T. Valente. "The cross-talk of HIV-1 Tat and methamphetamine in HIV-associated neurocognitive disorders", Frontiers in Microbiology, 2015. Publication	<1%
117	www.sogc.com Internet Source	<1%
118	Internet Source	<1%
119	Maldarelli, F., M. Kearney, S. Palmer, R. Stephens, J. Mican, M. A. Polis, R. T. Davey, J.	<1%

Kovacs, W. Shao, D. Rock-Kress, J. A. Metcalf, C. Rehm, S. E. Greer, D. L. Lucey, K. Danley, H. Alter, J. W. Mellors, and J. M. Coffin. "HIV Populations are Large and Accumulate High Genetic Diversity in Nonlinear Fashion", Journal of Virology, 2013.
Publication

120	www.jleukbio.org Internet Source	<1%
121	www.jove.com Internet Source	<1%
122	www.freepatentsonline.com Internet Source	<1%
123	Esposito, Francesca Carli, Ilaria Del Ve. "Sennoside A, derived from the traditional Chinese medicine plant Rheum L., is a new dual HIV-1 inhib", Phytomedicine: International Journal of , Nov 15 2016 Issue Publication	<1%
124	Guy C.J. Abell, John P. Bowman. "Ecological and biogeographic relationships of class Flavobacteria in the Southern Ocean", FEMS Microbiology Ecology, 2005 Publication	<1%
125	www.tdx.cat Internet Source	<1%

126	da Cunha, Inailson Marcio Costa de Souza. "Genetic multipartitions based on D-loop sequences and chromosomal patterns in Brown chromis, Chromis", BioMed Research International, Annual 2014 Issue Publication	<1%
127	ro.ecu.edu.au Internet Source	<1%
128	Chang, H.-W., L. J. Tartaglia, J. B. Whitney, S.-Y. Lim, S. Sanisetty, C. L. Lavine, M. S. Seaman, C. Rademeyer, C. Williamson, K. Ellingson-Strouss, L. Stamatatos, J. Kublin, and D. H. Barouch. "Generation and Evaluation of Clade C Simian-Human Immunodeficiency Virus Challenge Stocks", Journal of Virology, 2014. Publication	<1%
129	Hideto Teranishi, Nanae Ohzono, Ippei Miyata, Shoko Wakabayashi et al. "Incidence of Viremia With DNA Viruses in Oncology Patients With Febrile Neutropenia", Journal of Pediatric Hematology/Oncology, 2018 Publication	<1%
130	iovs.arvojournals.org Internet Source	<1%
131	Y Pilehvar, SM Hosseini, N Sheikhi, A	<1%

Ghalyanchi, M Shahidi, MT Kheiri. "Avian Influenza Survey in North-West of Iran from January to March 2008, Using RT-PCR", Iranian Journal of Virology, 2009
Publication

132	www.intechopen.com Internet Source	<1%
133	Wang, Y.. "Interleukin-10 inhibits HIV-1 LTR-directed gene expression in human macrophages through the induction of cyclin T1 proteolysis", Virology, 20060901 Publication	<1%
134	paperity.org Internet Source	<1%
135	Dörr de Quadros, Patrícia, Vanessa Sacramento Cerqueira, Juciana Clarice Cazarolli, Maria do Carmo R. Peralba, Flávio A.O. Camargo, Adriana Giongo, and Fátima Menezes Bento. "Oily sludge stimulates microbial activity and changes microbial structure in a landfarming soil", International Biodeterioration & Biodegradation, 2016. Publication	<1%
136	"A Historical Overview of the Epidemiology of HIV/AIDS in the United States", Social Disparities in Health and Health Care, 2016. Publication	<1%

137 Burcu S. Aytar, Mark R. Prausnitz, David M. Lynn. "Rapid Release of Plasmid DNA from Surfaces Coated with Polyelectrolyte Multilayers Promoted by the Application of Electrochemical Potentials", ACS Applied Materials & Interfaces, 2012 <1%
Publication

138 Kashyap, R. Ramteke, S. Gaherwar, H. Des. "Evaluation of BioFM liquid medium for culture of cerebrospinal fluid in tuberculous meningitis to id", Indian Journal of Medical Microbiology, Oct-Dec 2010 Issue <1%
Publication

139 Petr Nemec. "Association of Polymorphisms in Interleukin-10 Gene Promoter with Autoantibody Production in Patients with Rheumatoid Arthritis", Annals of the New York Academy of Sciences, 09/2009 <1%
Publication

140 Grant R Campbell. "What does the structure-function relationship of the HIV-1 Tat protein teach us about developing an AIDS vaccine?", Retrovirology, 2009 <1%
Publication

141 Q. Zhou. "The Yin and Yang of P-TEFb Regulation: Implications for Human Immunodeficiency Virus Gene Expression and <1%

Global Control of Cell Growth and Differentiation", Microbiology and Molecular Biology Reviews, 09/01/2006
Publication

---

142 Elijah M. Songok. "Identification of *env* CRF-10 among HIV Variants Circulating in Rural Western Kenya", AIDS Research and Human Retroviruses, 02/2003 <1%
Publication

---

143 "Mosaic RAS/MAPK variants cause sporadic vascular malformations which respond to targeted therapy.(CLINICAL MEDICINE) (Report)", Journal of Clinical Investigation <1%
Publication

---

144 Liddell, Lucette Holcomb, Maura. "Nodules on the ear.(CME CE: DERMATOLOGY COURSES: Dermatologic Look-Alikes)", Clinical Advisor, April 2015 Issue <1%
Publication

---

145 Vetter, Beatrice N., Vanessa Orlowski, Katrien Fransen, Christoph Niederhauser, Vincent Aubert, Marcel Brandenberger, Diana Ciardo, Günter Dollenmaier, Thomas Klimkait, Stephan Regenass, Patrick Schmid, Volkmar Schottstedt, Franziska Suter-Riniker, Sabine Yerly, Cyril Shah, Jürg Böni, and Jörg Schüpbach. "Generation of a Recombinant Gag <1%

Virus-Like-Particle Panel for the Evaluation of p24 Antigen Detection by Diagnostic HIV Tests", PLoS ONE, 2014.
Publication

146 Hermancia S., Ted M.. "Chapter 9 Immunotherapies and Vaccines", InTech, 2011 <1%
Publication

Exclude quotes	On	Exclude matches	< 3 words
Exclude bibliography	On		