Genetic, Structural, and Functional Exploration of the Restrictive Capacity of TRIM Proteins Against Immunodeficiency Viruses

Candidate:

Shmona Simpson

Student ID: 459785
Trinity College
Nuffield Department of Medicine

Supervisors:

Professor Sarah Rowland-Jones
Professor Mary Carrington

Submitted for the award of:

Doctor of Philosophy in Clinical Medicine

Abstract

HIV-2 differs from HIV-1 in that many infected people experience normal survival, whilst only 20% progress rapidly to AIDS. Understanding mechanisms of delayed HIV-2 disease progression could provide new insights into HIV control. The Caio Community Cohort was established in Guinea-Bissau in the setting of high HIV-2 prevalence. This thesis investigates the role of polymorphic host restriction factors of the TRIM family in HIV-2 outcome. TRIM proteins are a family of E3 ubiquitin-ligases, where closely-related TRIM5α and TRIM22 are thought to inhibit HIV-1 transcription, uncoating and budding.

There was an association between TRIM5α amino acid substitution R136Q and reduced HIV-2 viral load/prolonged survival. Conversely, P479L was enriched among HIV-2 infected participants and progressors with CD4+ T cell decline. TRIM22 was highly polymorphic in this cohort, revealing three novel coding variants. Although most substitutions were located in the putative virus-interacting PRYSPRY domain, two in the coiled-coil, D155N and R242T, showed significant and divergent associations with survival. R242T was enriched in HIV-2 infected participants, who progressed to death at twice the rate of wild-type controls. *In silico* studies predicted D282, D360, and R321 of TRIM22 to be highly conserved, exposed residues, for which polymorphisms would be deleterious. When aligned with sequences from the potent HIV-1 restriction factor, rhesus macaque TRIM5α, TRIM22 substitutions R321K, T415I, and D360Y were spatially relevant to residues involved in HIV-1 restriction. The role of TRIM22 in HIV restriction was supported by *in vitro* pilot studies showing that TRIM22 was upregulated by HIV-1 infection in a lymphoid cell line and co-localised with the HIV-1 capsid protein p24. Overexpression of TRIM22 resulted in the restriction of VSV-G pseudotyped HIV-1 and SIVmac. The R242T substitution diminished TRIM22's restriction of HIV-1 and SIVmac: protein analysis suggested that this may be due to the inability of the R242T mutant to fully dimerise.

Statement of Originality

I, Shmona Simpson, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Dedication

This thesis is dedicated with love and gratitude to my father:
Herman Simpson,
who departed this life on 3rd September, 2015.

Every day of my life I have felt deeply loved by you. Every day, I knew exactly where to find you, and knew you would be working hard for our benefit. Every day you were the best father I could hope for. At all times, and in all things, you gave me the financial, spiritual, and emotional tools, and the space, to establish my career. Thank you for inspiring me, believing in me, and teaching me to set my own trajectory, free from the confines of anyone else's opinion or expectation. I am ever imbued by your strength, and ever grateful for the many ways in which you continue to protect me. Being your daughter remains my greatest honour.

Acknowledgements

I sincerely thank my supervisors Professors Sarah Rowland-Jones and Mary Carrington. Sarah, thank you for the opportunity to work in your lab, and for giving me the space to develop my research interests. It has been so inspiring to watch you juggle research, clinical practice, and development work; make an impact at the highest level in all three, and remain so calm and family oriented through it all. My time with your group has sharpened me into a better public speaker, better scientific writer, better teacher, and an independent scientific researcher. Mary, thank you for ensuring this work was well executed, and the data as accurate as it could be. You have this fantastic combination of notable brilliance and humility, which has been very inspiring to witness. I also thank Professor Persephone Borrow; for helping me raise the standard of the functional assays in TRIM22, and for always being so generous with her time and energy. A huge thanks to Professor Robert Gilbert who helped me refine the 3D models presented in this thesis, and for contributing to my life on the DPhil in his role as Director of Graduate Studies. I am indebted to Professors Jan Rehwinkel and Quentin Sattentau for being on my thesis committee for the past two years, and I thank you for lending the wealth of your experience to me.

I am extremely grateful to Dr. Sabelle Jallow who initiated the work on TRIM5α in this cohort. Sabelle, thank you for being so gracious in allowing me to build upon your research, and for your guidance along the way as I attempted to extend beyond host genetics. I am ever so grateful to Dr. Louis-Marie Yindom for teaching me everything about DNA isolation, SNP typing, and sequencing, and for being an ever-present force of encouragement, especially at times when the experiments didn't cooperate. I am indebted to Dr. Thushan de Silva for collecting the 2010 specimens in Caio, and Dr. Carla van Tienen for sharing the cohort's clinical data with me. I am also tremendously grateful to Dr. Simon Brackenridge for helping me establish the molecular assays and for sharing reagents and positivity. Thanks to Dr. Kimiko Fukuhara for teaching me about proteins, and helping to purify TRIM proteins and structurally characterise them. Many thanks to Dr. Emmanouela Reppapi for statistical consultation. I also would like to thank Drs. Stephen Barr at University of Western Ontario for sharing TRIM22 plasmids with me, Eliza Vincenzi and Filippo Turrini of Milan for sharing their permissive and non-permissive U937 clones.

I have had many colleagues who gave me little presents along the way. Many thanks to Kamal Wood for giving of your mind and heart whilst assisting me with the statistical analysis. Thanks to Drs. Yanchun Peng for the C8166 cells, Hongbing yang for the HIV-IIIB isolate, and Eirini Moysi for the HIV-2 CBL-20 isolates. I cannot thank Drs. Shokouh Makvandi-Njead and Geraldine Gillespie enough for their guidance. I also wish to thank Drs. Isabela Pedrosa-Pacheo and Beatrice Ondondo for being very supportive when I was setting up my fist flow panels, and thanks also to my peer Thomas Partridge for his help when I was starting the protein work. I am also grateful to my peers Zoe Wallace for sharing her protocols and reagents to mount the cells for confocal microscopy and Dr. Yada Tansiri for showing me how to keep cells alive in culture.

Above all, I am grateful to my family who have been rooting for me, and holding me in one piece. Mom, I would never have applied to Oxford on either occasion without your suggestion. I thank you for literally jumping into the trenches with me and giving so much of yourself to see me through this program. Thanks from the bottom of my heart to my Grandparents: Vincent and Cynthia, for their incredible foresight, prayer and encouragement. I am grateful to my brothers Munair and Kearton, for encouraging me at every step of this journey. My dearest of friends Ania, Xoliswa, Kelly, Kopano, Makena, Kamz, and Sabi- thank you for sharing the weight of this process. I know that I would not be finishing this programme without each of you. I also wish to thank my mentor and friend, Professor Harold Jaffe, for nurturing my interest in HIV the past 7

years, and for his encouragement and advice along the way. I am also thankful to the Rhodes Trust: the Commonwealth Caribbean selection committee for their tremendous faith in me, and the Rhodes House staff Andrew Graham, Charles Conn, and most especially, Mary Eaton, for their support through the DPhil. I also thank the Nuffield Department of Medicine for awarding me the Prize Studentship to complete the final year of my research.

It has been a privilege to work with the Caio Community Cohort. I am grateful to the team who have continued HIV-2 studies in this community from before I was born, to present. It has been a privilege to be a part of them. The entire basis for the work conducted in this thesis rests on the generosity of participants of the Caio Community Cohort, who have shared pieces of themselves with the international scientific community for over 20 years.

Publications and Presentations

S Simpson, LM Yindom, E Repappi, T Vincent, G Nelson, and S Rowland-Jones. 2017. *"Female sex and TRIM22 Variants Predict Survival in HIV-2 Infected Adults."* Manuscript.

S Simpson, T de Silva, T Vincent, A Jaye, S Rowland-Jones.2015. "*Polymorphisms in TRIM22 are Associated with HIV-2 Acquisition and Disease Progression."* Restricting the Virus Inside and Out. International AIDS Symposium. Vancouver, Canada. July, 2015.

S Simpson, T de Silva, T Vincent, A Leligdowicz, L Yindom, S Rowland-Jones. 2015. *"TRIM22: A Novel Restriction Factor for HIV-2."* Mechanisms of HIV Persistence: Implications for a Cure. Keystone Symposia. Boston, USA. April 2015.

S Simpson, S Jallow. 2015. *"Interactions between HIV-2 and Host Restriction Factors."* Encyclopaedia of AIDS. Editors: Thomas J. Hope, Mario Stevenson, Douglas Richman. ISBN: 978-1-4614-9610-6

S Jallow, **S Simpson**. 2015. *"TRIM5α and HIV-2 Infection."* Encyclopaedia of AIDS. Editors: Thomas J. Hope, Mario Stevenson, Douglas Richman. ISBN: 978-1-4614-9610-6

Table of Contents

List of Abbreviations

AIDS	Acquired Immunodeficiency Syndrome
AP-1	Activator Protein 1
APOBEC	Apolipoprotein B mRNA Editing Enzyme, Catalytic Polypeptide-like
ART	Antiretroviral Therapy
BIV	Bovine Immunodeficiency Virus
BLAT	Basic Local Alignment Search Tool-like Alignment
BMI	Basal Metabolic Rate
BSA	Bovine Serum Albumin
CRF	Circulating Recombinant Form
CRFK	Crandell Feline Kidney Cells
CSWs	Commercial Sex Workers
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleic Triphosphate
dsDNA	Double-stranded Deoxyribonucleic Acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
EIAV	Equine Infectious Anaemia Virus
EIAV	Equine Infectious Anaemia Virus
ELISA	Enzyme-linked Immunosorbent Assay
FCS	Foetal Calf Serum
FIV	Feline Immunodeficiency Virus
GFP	Green Fluorescent Protein
HIV	Human Immunodeficiency Virus
HIV-1	Human Immunodeficiency Virus Type 1
HIV-2	Human Immunodeficiency Virus Type 2
HLA	Human Leuckocyte Antigen
HTLV-1	Human T-Lymphotropic Virus Type 1
IGEPAL	Octylphenoxy poly(ethyleneoxy)ethanol
IDUs	Intravenous Drug Users
IFN-γ	Interferon gamma
INSTIs	Integrase Strand Transfer Inhibitors
Interleukin	IL
KIR	Killer-cell Immunoglobulin-like Receptor
LB	Lysogeny-Broth
LTNPs	Long Term Non Progressors
LTR	Long Terminal Repeat
LTR	Long Terminal Repeat
MMR	Measles Mumps and Rubella
MSM	Men who have Sex with Men
MUAC	Mid Upper Arm Circumference
NFAT	Nuclear Factor of Activated T-cells
NF-kβ	Nuclear Factor kappa beta

NK	Natural Killer
NLS	Nuclear Localisation Signal
N-MLV	N-tropic Murine Leukemia Virus
NNRTIs	Non-nucleoside Reverse Transcriptase Inhibitors
NRTIs	Nucleoside Reverse Transcriptase Inhibitors
nsSNPs	Non-synonymous Single Nucleotide Polymorphisms
PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
PAMP	Pathogen Associated Molecular Pattern
PBMC	Peripheral Blood Mononuclear Cell
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PEI	Polyethyleneimine
PIC	Pre-integration Complex
PMDG	Vesicular Stomatitis Virus Envelope Protein Expression Vector
PolyPhen2	Polymorphism Phenotyping Version 2
PRR	Pattern Recognition Receptor
PRR	Pattern Recognition Receptor
RH	Hydrodynamic Radius
RNA	Ribonucleic Acid
RPM	Revolutions per Minute
RPMI	Roswell Park Memorial Institute Medium
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
SAIDS	Simian Acquired Immunodeficiency Syndrome
SAMHD1	Sterile Alpha Motif (SAM) and Histidine/Aspartic Acid (HD) Domain Protein 1
SDS	Sodium Dodecyl Sulphate
SFFV	Spleen Focus-forming Virus
SIFT	Sorting Intolerant from Tolerant
SIV	Simian Immunodeficiency Virus
SIVagm	Simian Immunodeficiency Virus infecting the African Green Monkey
SIVcpz	Simian Immunodeficiency Virus infecting the Chimpanzee
SIVsm	Simian Immunodeficiency Virus infecting the Sooty Mangabey
SNPs	Single Nucleotide Polymorphisms
SOC	Super Optimal Broth with Catabolite Repression
TAK1	Tat-associated Kinase 1
TNFα	Tumour Necrosis Factor-alpha
TRIM22	Tripartite Motif-containing Protein Number 22
TRIM5α	Tripartite Motif-containing Protein Number 5 Isoform alpha
VL	Viral Load (from plasma; unless otherwise stated)
VL1	Variable Loop 1
VL2	Variable Loop 2
VL3	Variable Loop 3
VL4	Variable Loop 4
VSV-G	Vesicular Stomatitis Virus Envelope Protein
WBC	White Blood Cell

List of Figures

List of Tables

Chapter 1. Introduction

1.1. HIV-2: Origins and Epidemiology

Human Immunodeficiency Virus (HIV) is a lentivirus (slow growing) which infects vital cells of the immune system such as CD4+ T-cells, macrophages, and dendritic cells(1). HIV infection results in Acquired Immunodeficiency Syndrome (AIDS), which is characterised by progressive destruction of the immune system, resulting in an eventually fatal inability to resist opportunistic infections or control malignancies. HIV is transmitted in blood, semen, vaginal fluid, and breast milk, as both free virus particles, and infected cells of the immune system(2). Since its discovery in the 1981, it has resulted in 60 million infections worldwide: however, through public health interventions the incidence of infections has reduced 35% since the year 2000. There are presently 36.7 million people living with HIV (2015 estimate) however only 17 million are receiving treatment, and only 70% of HIV infected pregnant women are receiving prophylaxis to prevent mother to child transmission(3). HIV therefore remains a global public health challenge.

Immunodeficiency viruses exist in multiple animal species, including cats, horses, cows, and non-human primates. Feline immunodeficiency virus (FIV) has morphological, physical and biochemical characteristics that are similar to HIV, and it bears an antigenic relatedness to equine infectious anaemia virus (EIAV)(4). FIV does result in a progressive decline in the immune system leading to clinical presentation bearing much similarity to human AIDS(5). Bovine immunodeficiency virus (BIV) has been recovered from the leucocytes of cattle with persistent clinical signs of immunodeficiency, lesions in the central nervous system, progressive weakness and emaciation(6). EIAV can be transmitted through blood, milk and body secretions, but interestingly can also be transmitted by biting flies, and infects red blood cells(4, 7). Simian immunodeficiency viruses (SIV) are able to infect over 45 species of African non-human primates and, in a large number of cases, appear to be non-pathogenic in their natural hosts(8-13). However, SIV infection of a rhesus macaque of Indian origin results in Simian AIDS (SAIDS)- of similar character to human AIDS(14). HIV exists in two major types, HIV Type 1 (HIV-1)

responsible for the global pandemic, and HIV Type 2 (HIV-2). HIV-1 Group M is believed to have been derived from zoonotic transfer of Simian immunodeficiency virus infecting chimpanzees (SIVcpz) to humans(15). HIV-2 arose during the zoonotic transfer of sooty mangabey Simian Immunodeficiency Virus (SIVsm) to humans on at least eight occasions(16). Of the resulting eight lineages, HIV-2 A-H, only A and B are capable of onward human to human transmission, resulting in an estimated 1-2 million cases mainly across West Africa(17, 18). Five lines of evidence have been used to substantiate the zoonotic origins of HIV-2:

(i) HIV-2 and SIVsm share an identical genome structure, with each virus encoding an accessory protein, Vpx, that has not been observed in any other primate lentivirus(19)

(ii) SIVsm and HIV-2 strains are phylogenetically closely related and cannot be separated into distinct phylogenetic lineages according to their species of origin(20)

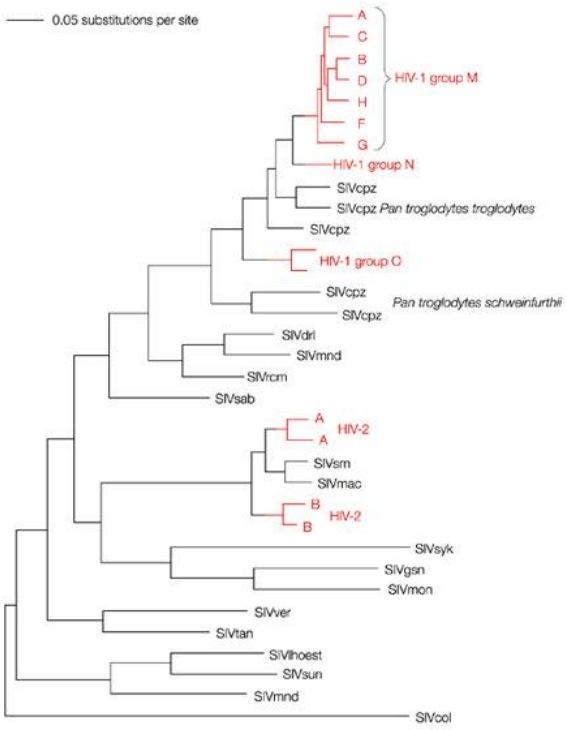


Figure 1.1 Evolutionary Relationships Among Primate Lentiviruses*

*Adapted from Rambaut R., et al, (2004) "The Causes and Consequences of HIV Evolution" Nature Reviews Genetics 5, 52-61 (3) * Sequences based on pol*

(iii) Sooty mangabeys are established in West Africa with an SIV infection frequency as high as 22% in some troops(21)

(iv) Sooty mangabeys are frequently hunted or kept as pets, allowing for frequent encounters of infected animals with humans(20, 22)

(v) There is a geographic coincidence between HIV-2 in endemic areas and the natural habitat of the sooty mangabey. The historical range of the sooty mangabey is coastal west Africa from Senegal to Côte d'Ivoire; in close proximity to the HIV-2 epicenters in Senegal, Guinea-Bissau, Guinea, and Côte d'Ivoire, overlapping Sierra Leone and Liberia, where the most divergent HIV-2 strains have been identified (15)

HIV-2 is largely confined to West Africa and countries with socioeconomic ties to the region. HIV-2 Group A is found throughout West Africa, whereas Group B is mainly prevalent in Cote d'Ivoire(23, 24).

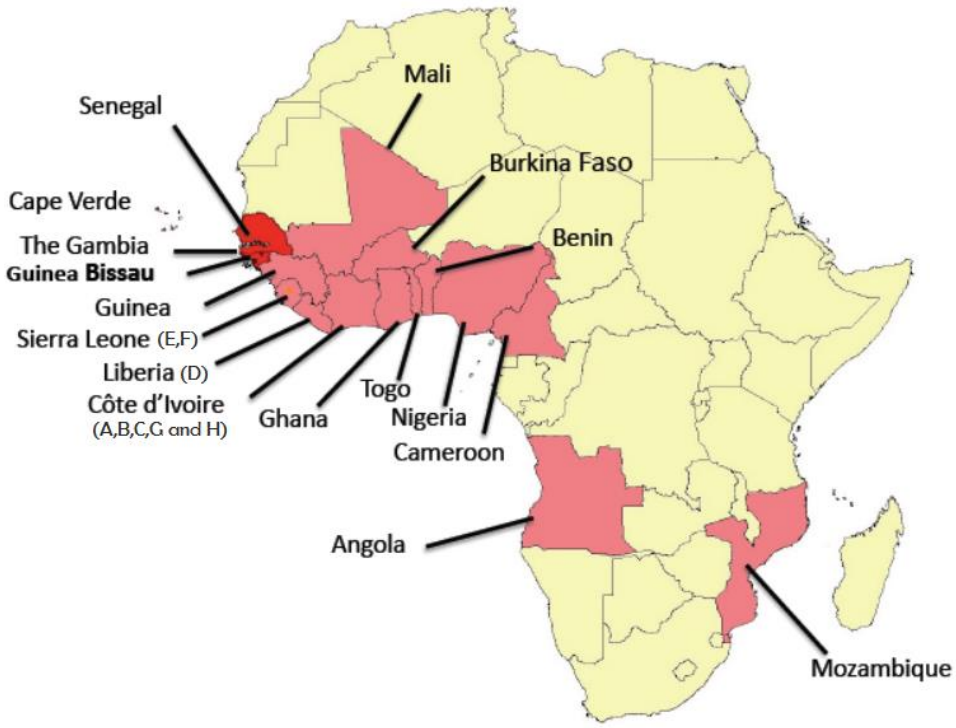


Figure 1.2 Countries with HIV-2 Infection in Africa

The other groups are likely to been single-person transmission events that apparently failed to spread beyond the initial human host. In terms of geography, HIV-2 C, G, and H are related to

SIVsm strains from the Ivory Coast, where Group D is similar to SIVsm from Liberia, and groups E and F are most similar to SIVsm from Sierra Leone(20). There also exists a circulating recombinant form (CRF); CRF01_AB, with a genome of group A and B fragments, isolated in Japan. An isolate with a similar pattern was described in 1994 in Cote d' Ivoire suggesting that this CRF may be more widespread(16, 25). In addition, there are substantive communities of HIV-2 infected individuals in Angola, Mozambique(26), India(27), Brazil(28), Portugal, where it is responsible for 4.5% of AIDS cases(29), France(30) and the United Kingdom, some large enough to have resulted in the establishment of HIV-2 clinical cohorts.

1.2. HIV-2 Virion Structure and Life Cycle

Like all primate lentiviruses, HIV-2 possesses the major structural genes *gag*, *pol*, and *env*, which encode the essential proteins for infection and replication. The genome also contains the regulators of viral expression *tat* and *rev*, along with four accessory genes, *vif*, *vpx, vpr*, *nef.*

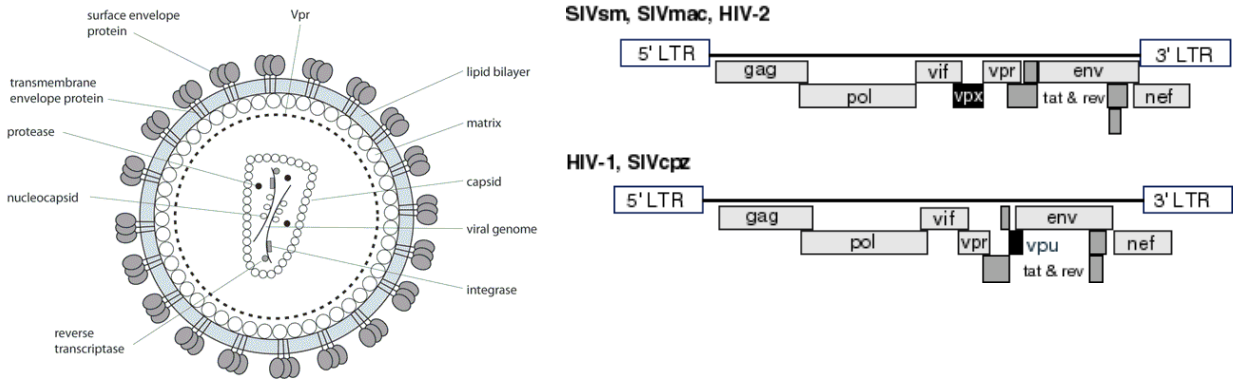


Figure 1.3 Genetic and Cellular Structure of HIV-2.

Reproduced from Beer, B. et al. "Diversity and Evolution of Primate Lentiviruses" Los Alamos National Laboratory. HIV sequence Databank.
http://www.nature.com/nrmicro/journal/v10/n4/full/nrmicro2747.html

At the core of each HIV-2 virion are 2 copies of a ribonucleic acid (RNA) genome and a reverse transcriptase enzyme, encoded by *pol*(31). Both are packaged within a capsid consisting of the protein p26 encoded by *gag*. The capsid is contained within the viral envelope, a lipid bi-layer of host cell origin. Embedded within the envelope are the viral proteins gp125 and gp36, encoded

by *env*, which combine in a trimer that enables the virus to infect the host cell(32). The accessory and regulatory genes *vif, vpx, vpr, rev*, and *tat* encode proteins that work in conjunction with the host cell machinery to facilitate the transcription, maturation, and propagation of the virus(33). It is worthy of note that the corresponding accessory protein to Vpx, is Vpu in HIV-1, which has a role in counteracting the host restriction factor tetherin. The absence of Vpx results in HIV-1's restriction in myeloid and dendritic cells, a feature not observed in HIV-2 infection(34). HIV-2 enters CD4+ T cells and macrophages through plasma membrane fusion of the Env glycoprotein with CD4, facilitated by engagement of viral co-receptors such as CCR5 and CXCR4, which are otherwise involved in immunity and inflammation(35). HIV-2 differs from HIV-1 in that it can utilise a variety of additional coreceptors such as CCR1-5, GPR15, and CXCR6 which HIV-1 cannot(36, 37). The viral core is encased by a shell composed of the viral capsid. After the viral core is released into the cytoplasm, viral uncoating involves cellular factors and the viral proteins matrix (MA), nef, and vif. The viral RNA genome is retrotranscribed into full-length double-stranded deoxyribonucleic acid (dsDNA) by the viral reverse transcriptase(38). Thereafter, directed by Vpr(39), the virus pre-integration complex (PIC) containing viral complementary DNA (cDNA) and viral integrase, is translocated into the nucleus using the cellular proteins lens epithelium-derived growth factor, LEDGF/p75, and Transportin3, which guide the PIC towards active transcription sites and facilitate nuclear translocation respectively(40). In the nucleus, viral-associated integrase inserts viral cDNA into the human chromosome. During the HIV-1 replication cycle within the host cell, the transcription is enhanced by cellular activation. mRNA and full-length viral genome RNA are exported from the nucleus. Viral proteins assemble beneath the plasma membrane and virus particles bud from plasma membrane (41).

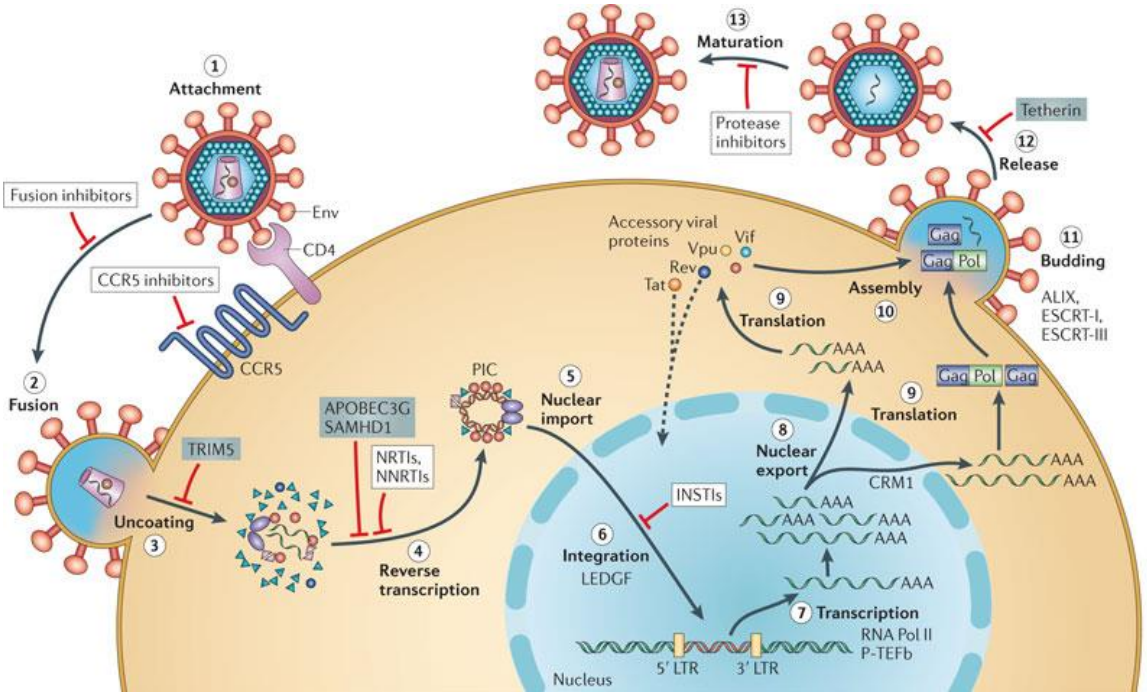


Figure 1.4 The HIV Life Cycle

The HIV Life Cycle: Reproduced from Engelman, A., et al (2012) "The structural biology of HIV-1: mechanistic and therapeutic insights." Nature Reviews Microbiology 10, 279-290(42)

The result of the untreated HIV lifecycle is an initial primary infection, which stabilises into a clinical latency where a complex interplay occurs between the human host and the virus. Over time, which can be 8-12 years in the case of HIV-1, and substantially longer in the vast majority of HIV-2 cases(43), there is progressive destruction of the cells of the immune system, leading to an increase in the burden of cellular virus (proviral load) and free circulating virus (plasma viral load)(44). An AIDS defining illness results when cells of the immune system, symbolised by the CD4+ T cell count, are depleted, and the viral burden begins to overwhelm the host(45). Increased susceptibility to opportunistic infections and cancers results in eventual fatality(46, 47). HIV treatment can drastically change the outcome of HIV infection.

1.3. HIV-2 Treatment

At each key stage of the viral life cycle, a drug has been developed to disrupt HIV replication: the viral targets — envelope, reverse transcriptase, integrase and protease- have drugs on the market which inhibit their action, with the addition of fusion inhibitors which target co-

receptors involved in virus entry. The most common are INSTIs, integrase strand transfer inhibitors; NNRTIs, non-nucleoside reverse transcriptase inhibitors; and NRTIs, nucleoside reverse transcriptase inhibitors. HIV-2 has 90% sequence homology with non-human primate lentiviruses such as SIVsm, but is less homologous to HIV-1 (roughly 60% in gag and 40% in pol)(48). Therefore fundamentally, the use of drugs designed to target HIV-1 will be challenging to use for HIV-2. Many drugs used to treat HIV-1 are not effective for HIV-2. HIV-2 is naturally resistant to first generation NNRTI's(49, 50) and the fusion inhibitor enfuvirtide(51). Due to naturally occurring polymorphisms in reverse transcriptase, the potency of NRTI's is potentially diminished for HIV-2(50). Studies have shown that higher concentrations of zidovudine are required to suppress HIV-2, and that resistance mutations for NRTI's have occurred in treatment naïve individuals(52, 53). Protease inhibitors also show varying degrees of activity due to natural protease polymorphisms, and thus atazavir, fosamprenavir and tipranavir are used with caution in HIV-2 infected patients(54, 55). The promiscuity of HIV-2's Env for a myriad of coreceptors makes the potency of the CCR5 antagonist, maraviroc, uncertain(56). There are other host proteins, termed host restriction factors, that play a role in restricting the replication of HIV. Any of these could become potential antiviral targets if exploited in a way which is not hazardous to the host. Further, the lack of a large observational or randomised treatment studies in HIV-2 infected patients makes treatment and management of HIV-2 infection quite challenging(57). Resource constraints in many low income countries leads to guidelines for the initiation of ART at 350cells/ml where treatment is available(58). However, HIV-2 viral load is often undetectable until CD4+ T-cell count is less than 300cells/ul, and it is the plasma viral load which drives disease progression(59). Furthermore, CD4+ T-cell recovery in treated HIV-2 infected individuals is poorer than expected from HIV-1 data(60). Therefore, initiation of ART is recommended at an HIV-2 plasma viral load about 1000copies/ml(61). At this time, HIV-2 is susceptible to new integrase strand transferase inhibitors like raltegravir. Therefore the preferred treatments for HIV-2 consist of combinations of two NRTI's and an appropriate boosted protease inhibitor.

1.4. The HIV-2 Virus and the Immune Response

Both HIV-1 and HIV-2 have the same degree of infectivity but result in differential disease progression. Proviral DNA loads are similar between the two, but HIV-2 replication is 100x lower than that of HIV-1 resulting in a much lower plasma viral load(62). In terms of innate immunity, natural killer (NK) cells are more cytotoxic in HIV-2 infection and produce more cytokines, but only prior to CD4+ T-cell decline(63). A major hallmark of HIV infection is chronic immune activation that results in CD4+ T-cell recruitment, and thus drives viral replication and eventual CD4+T-cell depletion. Chronic immune activation is characterised by elevated proinflammatory cytokines, soluble CD14 and B2 microglobulin(64) (which has inverse relationship with CD4+ T-cell count), in plasma, and increased T- cell turnover followed by a loss of T-cell regenerative capacity(65). On average, HIV-2 has lower immune activation than HIV-1, however in advanced disease characterised by high viraemia, the levels of immune activation in HIV-1 and HIV-2 are the same. A reason for this early difference may be because HIV-2 intracellular Nef promotes downregulation of the T-cell receptor complex in CD4+ T cells, where HIV-1 nef is unable to do so(66). Structurally, HIV-2 Env has fewer glycoslylation sites in the Env variable loop 3, V3, domain which makes the V3 domain more accessible(67). This provides one explanation for HIV-2's promiscuous co-receptor usage. Another effect of V3 accessibility, is that HIV-2 is more susceptible to antibody neutralisation than HIV-1, producing potent neutralising antibodies specific for epitopes in the gp120 C2, C3, V3, and V4 regions in gp120. Not only do HIV-2 Env trimers expose multiple broadly cross-reactive epitopes susceptible to neutralising antibodies, but these antibodies are deployed soon after infection resulting in notable evolution in the viral envelope; gp120's V2 and V3. Plasma from HIV-2 infected people therefore neutralises a greater proportion of viruses than does plasma from HIV-1 infected people(68). HIV-2 also has a strong cytotoxic T lymphocyte response in asymptomatic patients, where strong cytolytic activity against Gag, Pol, and Nef antigens is inversely correlated with viral load. CD4+ T-cells also show less differentiation, retain better proliferative capacity and elicit more polyfunctional responses

when compared with HIV-1(69). HIV-2 is better at infecting dendritic cells due to the presence of HIV-2's vpx, absent in HIV-1(70), which proteosomally degrades Sterile alpha motif (SAM) and histidine/aspartic acid (HD) domain protein 1 (SAMHD1). SAMHD1 is an intracellular nuclease which hydrolyse dNTPs and inhibits reverse transcription. Rather than vpx mediated SAMHD1 degradation enhancing the pathogenicity of HIV-2, it acts to reduce it(71), because infected dendritic cells trigger a protective type-1 interferon (IFN) response thus reducing virulence(72).

1.5. HIV-2 Pathogenicity: Data from West Africa

Like HIV-1, HIV-2 is primarily transmitted via sexual contact, blood exposure, and from mother to child. Despite similarities in their protein structure and infectivity, the rate of HIV-2 transmission within a population, and the rate at which those infected progress to AIDS, is substantially lower than that of HIV-1. This suggests HIV-2 is a less pathogenic virus, or that it is better restricted by the human host. 35-40% of HIV-2 infected persons exhibit a phenotype similar to that of HIV-1 long-term non progressors. This phenotype is characterised by low to undetectable viral load, and with little to no symptoms of immunodeficiency. The remaining roughly 20% of people living with HIV-2 progress to AIDS in a manner identical to that of HIV-1 infected persons(69). Several HIV cohorts in West Africa have provided evidence for the reduced pathogenicity:

In The Gambia, vertical transmission from HIV-2 infected mothers in the absence of ART is 4% when compared to 24.4% for HIV-1 within the same population, a phenomenon closely related to viral load(73). Observed in Senegal, shedding of HIV-2 in genital secretions is lower for HIV-2 than HIV-1; also closely related to plasma viral load(74). A study of female commercial sex workers (CSWs) in Senegal showed heterosexual spread of HIV-2 being slower than that of HIV-1(75). In Ivory Coast, the rate of perinatal transmission of HIV-2 was 1.2% compared with 24.7% for HIV-1 (76),which was comparable to that observed in the Gambia: the observation a result of the comparably low average viral load in HIV-2 infected women (410copies/mL) when compared

to HIV-1 infected women (15,100 copies/mL). Should viral loads be equal, the odds of perinatal HIV-2 transmission are similar to that of HIV-1. This trend was also observed in the Senegalese sex worker cohort, where almost half had a HIV-2 viral load below the detectable limit, and where viral load was observed, it was 30 fold lower than in HIV-1 at the same point of seroincidence(77). Further, a cohort in Guinea Bissau confirmed this finding, and also indicated that viral load is also able to predict survival of HIV-2(78, 79). Not only is HIV-2 characterised by lower viral loads, but it also has a higher CD4+ T-cell counts at similar stages of infection when compared to HIV-1. In the clinical cohort in The Gambia, HIV-2-infected patients with CD4+ T-cell counts of 500 cells/µl and greater, had a significantly lower mortality rate than HIV-1-infected patients, however those with CD4+ T-cell counts less than 200cells/µl had as high a mortality rate as HIV-1 infected patients(80). Similarly in Guinea Bissau, both plasma viremia and CD4% were independent predictors of survival, with hazard ratios increasing by 1.6 for each log(10) increase of plasma viraemia and 1.7 for each 10% decrease of CD4%(81). Disease progression and survival are therefore closely related to both CD4+ T-cell count and plasma viral load. However it is worthy of note that even at the point of death from HIV-2, the CD4+ T- cell count is much higher than that of HIV-1. Even then, the viral load is comparably low when compared to HIV-1 advanced disease.

The prolonged asymptomatic phase, and slower progression to AIDS has been observed in several cases; a cohort of Senegalese female CSWs showed 100% probability of AIDS-free survival 5 years after HIV-2 seroconversion, compared with only 67% following HIV-1 seroconversion(82). Despite this, HIV-2-infected patients with advanced disease had the same poor prognosis as patients with HIV-1(80). Clinical manifestations such as tuberculosis, multiple cranial neuropathy, NK/T cell lymphoma, demyelinating encephalitis, and wasting syndrome have been observed in AIDS resulting from HIV-2 infection (83). Kaposi's sarcoma- a defining feature of AIDS mediated by HIV-1, is hardly ever seen in the case of HIV-2(84). Understanding the underlying physiological mechanisms which differentiate HIV-2 infected individuals who are

likely to progress to AIDS, from those who remain asymptomatic non-progressors, is vital to the treatment of HIV-2 infected people.

Many studies have been undertaken to determine whether there are characteristics of HIV-2 virology which contribute to long term non progression. As an isolated virus, HIV-2 appears to show reduced replication kinetics when isolated from an aviremic HIV-2 non progressor when compared to a progressor(85). Furthermore, a study of HIV-2 vertical transmission showed that only one in four children of mothers with high HIV-2 viral load showed signs of disease progression(86). At a sequence level, no significant differences have been observed in HIV-2 sequence variation amongst controllers and progressors. A comparison of proviral sequences in *gag, pol, env*, and the LTR in 131 participants in Guinea Bissau, showed no correlation between any particular genotype and CD4+ T-cell count or viral load(87). A study of *gag* p26 from viral RNA derived from the same cohort, showed that prolines at position 120, 159 and 178 were more frequent in participants with lower viraemia when compared to those with higher viraemia(88). Sequence analysis has shown that prolines at the 120th position of the HIV-2 capsid protein confers sensitivity to the host restriction factor Tripartite-motif containing protein, isoform alpha (TRIM5α) in cynomolgus macaques(89). Taken together, the data would suggest that virology is a particularly defining factor of HIV-2 long term non progression, and that the interplay between the virus and the host immune response might provide more insights.

To facilitate the dissection of this phenomenon, several HIV-2 cohorts have been established; the best characterised being the Caio Community Cohort, Guinea Bissau, where the highest recorded community prevalence of HIV-2 has been reported.

1.6. The Caio Community Cohort

Guinea-Bissau is a small country in West Africa with a population of 1.8 million, the majority of which live in the capital Bissau. Great ethnic variety exists, and the most frequent groups observed are the Balante and Fula, who comprise the largest population, the Balanta and Papel

who live in the southern coastal regions, and finally the Mancanha and Manjaku who occupy the northern coastal regions(90). Historically Guinea Bissau was part of the Gabu Empire, a Mandinka kingdom which extended into Senegal. After annexation in 1867, the country suffered from colonialism until its war of independence ending in 1973, and post-colonial political instability, gridlock, and inefficiency in public spending thereafter(91). The economy relies on agriculture, with the main export being cashew nuts, followed by groundnuts and fish, which presently sustains a growth of 5%. Caio, is based on the north western coat of Guinea Bissau. It consists of 10 settlements that are dispersed among cashew and palm tree forests and rice fields, with a population of 10,000 at the last time point in this study which was 2010. 95% of Caio's inhabitants belonged to the Manjaku ethnic group. The inhabitants of Caio are dependent on subsistence farming, mostly of cashew, and have retained their animist belief system. Despite the relative inaccessibility of Caio, the community is migratory, and people, mostly men, travel for both economic and social reasons(92).

As will be discussed later in this introduction, data was collected from this population on 5 occasions throughout a 20 year study period. According to the World Bank, data accessed September 2016, Guinea Bissau has been, and continues to be, one of the world's poorest and most fragile countries with the majority of citizens living on $3.10 per day. Gender parity has improved between 2003 and 2015 with 50% literacy among women aged 15 and older. However, gaps continue in the number of women relative to men enrolled in productive employment. At all points during the study period, women had a higher life expectancy than their male counterparts and were more likely to survive to age 65 than men.

Table 1.1 Key Estimates of Gender Disparity in Guinea Bissau

	1990/91	1996/97	2003	2006	2010	2015
Literacy rate, adult female (% of females ages 15 and above)	-	-	27.49	-	-	46.70
Literacy rate, adult male (% of males ages 15 and above)	-	-	57.61	-	-	71.70
Poverty headcount ratio at $3.10 a day (2011 PPP) (% of population)	61.08	85.01	80.93	-	83.59	-
Employment to population ratio, 15+, female (%)	64.1	65.60	67.1	67.1	68.2	-
Employment to population ratio, 15+, male (%)	73.0	73.40	73.3	73.4	73.4	-
Survival to age 65, female (% of cohort)	49.60	50.30	49.95	52.46	54.26	56.31
Survival to age 65, male (% of cohort)	40.77	44.31	46.50	46.43	47.59	49.14
Life expectancy at birth, female (years)	51.69	52.66	52.74	54.16	55.5	56.99
Life expectancy at birth, male (years)	47.27	49.68	51.17	51.61	52.21	53.42

Data taken from http://www.worldbank.org/en/country/guineabissau

Studies on the Caio Community were initiated in 1989, following an observation made in 1988 by a clinician based in Senegal who noticed a high frequency of HIV-2 infected women originating from Caio(93). Field studies were led by the Medical Research Council in the Gambia, to perform demographic surveillance with an aim of investigating the epidemiological impact of HIV-2. Also studied in this community has been the prevalence of HIV-1 and Human T-cell Leukemia Virus Type 1 (HTLV-1). Unlike HIV-1 and HIV-2, introduced relatively recently, HTLV-1 is an ancient virus with low evolutionary rate (94) and has a high prevalence of 4.9% in the Caio Community(95). In HTLV-1 infection, 95% of the carriers remain asymptomatic, and the remainder will develop adult T-cell leukemia or tropical spastic paraparesis. Like HIV, the routes of transmission are sexual contact, contaminated blood, and breast feeding.

Of the several HIV-2 studies based in Caio, five are amalgamated in this thesis. Of these, three were cross-sectional seroprevalence studies taking place in 1990 (1989–1991)(96) 1997 (1996–1998)(97) and 2007 (2006–2007)(98). In all studies, fieldworkers randomly visited adults in their homes and offered testing for HIV-1 and HIV-2. The three surveys were held in a similar manner which will be described further in the Materials and Methods section of this thesis. As the studies were in the manner of a “sero-census”, HIV-2 positive individuals were identified only during community surveys and not selected from a clinic, therefore avoiding selection bias for infected participants. For each survey, prevalence was measured amongst five different age

categories (15–24, 25–34, 35–44, 45–54, ≥55 years). In 1989, the prevalence of HIV-2 was 9.3% in women, peaking at 17.2% in the 35-44 age group, and 6.6% in men, peaking at 19.1% in the 45-54 age group, resulting in an average HIV-2 prevalence of 8.3%. Overall HIV-2 prevalence remained stable between 1990 and 1997, shifting only slightly from 8.3 to7.9%; in men, peaking at 17.5% in the 35-44 age group, and in women peaking in the oldest age group (>55) at 16.3%.

Table 1.2 HIV-2 Studies in the Caio Community Cohort

Study Year	Study	Study Subjects	Author	Reference
1989-1991	First HIV Serosurvey	Adults 15 years of age and over	A. Wilkins	(96)
1991	First case- control study	HIV+ and HIV- adults	D. Ricard	(99)
1996	Second case-control study	Follw up of HIV+ and controls	K. Ariyoshi	(100)
1997-1998	Second HIV serosurvey	Adults 15 years of age and over	M. Schim van der Loeff	(97, 101)
2003	Natural History Study	Clinical manifestations in HIV+ and HIV- adults	A. Gourlay	(102)
2006-2007	Third HIV case-control study	Adults 15 years of age and over	C. van Tienen/ A. Zaman	(98, 103)
2010	HIV-2 Elite Control Studies	HIV+ and HIV- Subjects	T. de Silva	(104)

By 2007, the overall HIV-2 prevalence reduced to 4.7%. Despite this overall reduction, prevalence remained higher amongst older individuals over the same period peaking at 7.2% and 14.9% among men and women respectively, in the oldest age group (>55 years). By the end of the cross sectional study period in 2007, HIV-2 was largely a disease of older women, compounded by the fact that a large number of the community's male members seek employment in the capital or in neighbouring countries, resulting in the majority of those surveyed being women. HIV-2 infection was more prevalent in women who were widowed or divorced, women whose husbands were living away from the study area, and women who had lived in the capital, Bissau. The relevantly high prevalence amongst older women was also speculated to relate to hormonal changes affecting the vaginal mucosa.

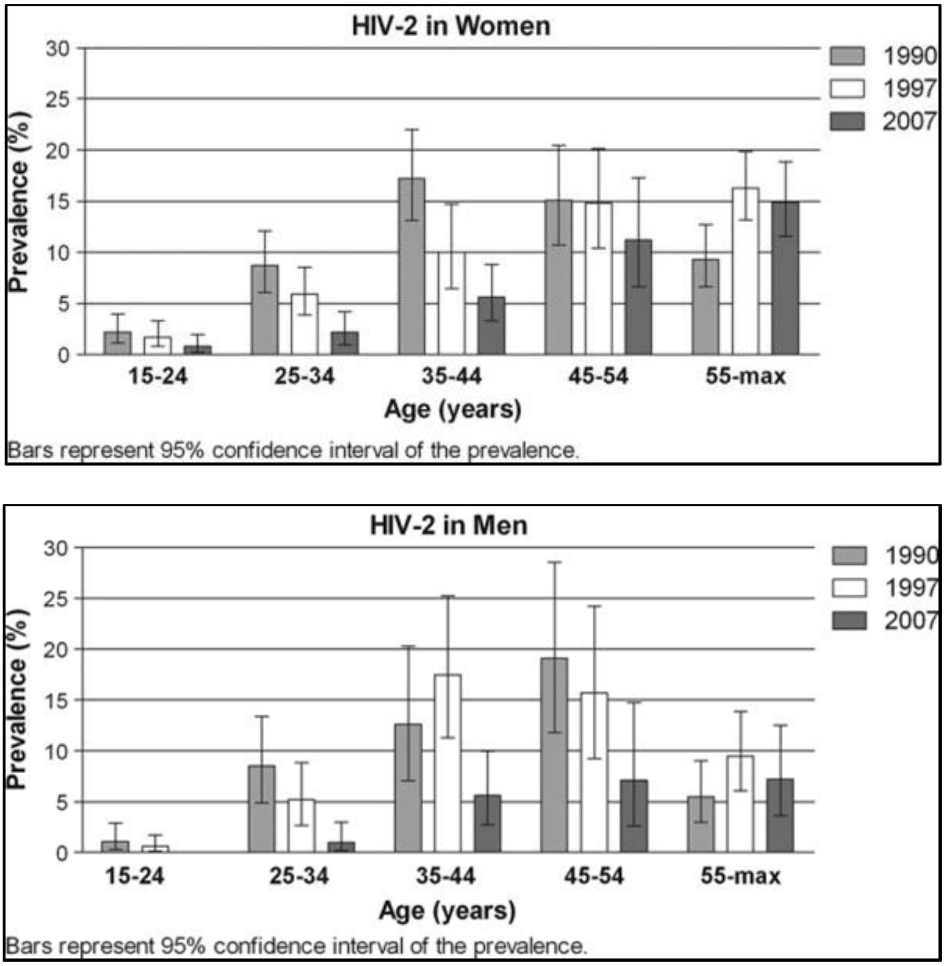


Figure 1.5 HIV-2 prevalence by age in women and men in three surveys (1990, 1997 and 2007) in Caio, Guinea-Bissau

Reproduced from: Tienen, Carla., Rowland-Jones, Sarah ., et al (2010) JAIDS. "Two Distinct Epidemics: The Rise of HIV-1 and Decline of HIV-2 Infection Between 1990 and 2007 in Rural Guinea-Bissau."

Of note, is the fact that HIV-1 prevalence also rose within the same community from 0.5% to 3.6% during the same time period. The same was observed in a Senegalese sex worker cohort, where HIV-2 prevalence fell from 8% in the 1985 to 6% in 2003, while HIV-1 increased from 1% to 13.8% in the same period. HIV-1 is thought to compete with HIV-2 primarily through both strains being more prevalent in high-risk individuals, allowing HIV-1 to preferentially remove HIV-2-infected individuals from the population. More likely factors for the decline of HIV-2 are its low pathogenicity, with sexual and vertical transmission rates much lower than that of HIV-1. Other West African countries also show declining or stable HIV-2 prevalence, compatible with the decreased HIV-2 incidence and mortality caused by old age(105).

Another time point incorporated in thesis was during 2003, where a prospective study designed to detect clinical predictors of HIV-2 mortality was initiated. Here 147 HIV-2 infected and 197 uninfected age and sex matched participants were recruited. Their history was taken, and a clinical examination performed including the collection of laboratory data such as plasma viral load and CD4+ T-cell count. Each participant was scored according to the WHO performance scale ranging from 1 (asymptomatic) to 4 (advanced disease and bedridden 50% of the day), and clinical features like body mass index (BMI) and mid upper arm circumference (MUAC) were noted. Survival was assessed in the 2010 census. Once again the cohort consisted of mostly older females, and the study found that gender was strongly associated with mortality, with females experiencing 0.3 times mortality rate of males. Age also had a strong effect on mortality, the highest mortality rate being in the over 55 age group. The mortality rate of HIV-2 infected participants was found to be 2 times higher than uninfected participants, which contrasts with the over 8 fold increased mortality rate in HIV-1 infected individuals from the same region(106). The findings showed that chronically HIV-2 infected individuals did not show clinical signs of disease progression, as a majority were WHO stage 1 and had normal BMI or MAUC and also a long-term low viral load. However, CD4+ T cell count and viral load were strong predictors of mortality, consistent with previous studies in the cohort and others in the region(79, 81, 107).

The final data point contributing to this thesis resulted from a study carried out 2010. Of the surviving HIV-2 infected individuals at the point of the 2010 census, sixty were recruited to determine the polyfunctionality of T-cell responses associated with control of HIV-2. Roughly half possessed an undetectable viral load, and had done so for at least the preceding 10 years of the study, whilst the other half had a detectable viral load of at least 500IU/ml. It was found that the viraemic controllers possessed a high-magnitude, polyfunctional gag-specific CD8+ T cell response. These T cells exhibited low levels of activation and proliferation(104).

No further samples or clinical data have been collected since 2010. A further study on the influence of compound Human Leukocyte Antigen (HLA) and Killer-cell Immunoglobulin-like receptor (KIR) genotypes on progression related determinants such as viral load and CD4+ T-cell counts relying on data from the 1991, 1996, 2003, and 2006 samples, demonstrated that individuals carrying the common B*1503 allele were more likely to progress to AIDS than those without this allele. There have been no other studies to date exploring the role of host genetics in this unique cohort.

1.7. Host Restriction Factors

Experimental infection of macaques has been the dominant animal model for HIV research, resulting in the discovery of several gene expression signatures that are associated with virus transmission and disease progression. The discovery that HIV-1 could not replicate in rhesus macaques despite its similarity to SIV(108) led to the identification of endogenous factors, termed host restriction factors, that restrict a number of retroviruses(109). Host restriction factors mount a first line response to viral infection, yet, are often hampered by specific viral proteins which directly resist their action(110).

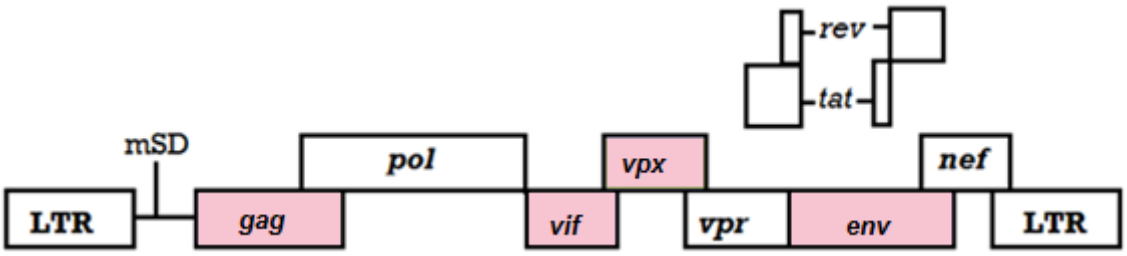


Figure 1.6 HIV-2 Genes which antagonise the function of Host Restriction Factors

Structural proteins gag and env, and accessory proteins vif and vpx, are involved in HIV-2 antagonism of host restriction factors.

The HIV-2 proteins utilise no specific enzymatic function in their roles as antagonists. Instead, they act as “linker molecules”; connecting host restriction factors to cellular proteins involved in sequestration or degradation, and, in so doing-, enabling viral replication and transmission(111). Host restriction factors are encoded by functionally autonomous genes and tend to be unusually variable(112). In fact, their counterparts in non-human primates are also unusually variable,

suggesting that they have evolved under diversifying selection(113, 114). They are believed to have a role in immunity and inflammation because transcription of the genes encoding these proteins is in most cases stimulated by the innate immune system(115). A common feature of host restriction factors, is that several retroviruses have evolved to resist their action by, in most cases, encoding accessory genes that antagonise their function(109, 110, 116). The HIV-2 proteins *gag*, *vif*, *env*, and *vpx* serve to modify the actions of four major host restriction factors: TRIM5α, APOBEC, Tetherin, and SAMHD1 respectively.

1.8. The Interplay between HIV-2's Vif and Host Factor APOBEC

Viral infectivity Factor (Vif) is an accessory protein conserved across all primate lentiviruses. Vif-deficient HIV-1 is unable to spread in physiologically relevant macrophage and T-cell cultured cells(117). Instead, Vif-deficient virions are produced in normal quantities but the ability of these virions to infect the subsequent target cell is compromised(118). This is due to the presence of APOBEC proteins(119, 120). The APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) proteins have several members, A-H, that have been studied in the context of HIV infection. APOBEC3G (A3G), is expressed in the majority of cell types. In the absence of Vif, A3G is packaged into nascent virions, and acts during reverse transcription to hypermutate viral DNA and decrease cellular cDNA levels(121-123). The expression of A3G is sensitive to viral infection and interferon(124-126), and so the amount of A3G incorporated into nascent virions is proportionate to its expression level in the producer cell. The net effect of its increased expression, and packaging into nascent virions, is a restriction of HIV-1 infection(127, 128). APOBEC proteins mediate their effect utilising two zinc-coordinating deaminase domains (Z domains) at each end of the protein(129):

The carboxy-terminal Z domains responsible for deamination(130, 131).

The amino-terminal Z domain mediates A3G's incorporation into viral particles and, is also recognised by HIV-1 Vif(132, 133).

In order to package itself into assembling virions, A3G uses its amino terminal to bind RNA and attach the Z domain to the Nucleocupsid region of Gag(134). Following its incorporation into assembling virions, it is transported from the producer cell to the target cell(135). Once the virion has fused with the target cell, A3G dimerises, and associates with the viral reverse transcriptase complex(136). A3G then deaminates cytidine residues to uridine in nascent, single stranded, negative strand cDNA. These uridine rich transcripts can be degraded. However if salvaged, during the synthesis of the DNA plus strand, adenosines are incorporated instead of the original guanines resulting in G to A mutation(123, 137). Up to 10% of cytidines may be edited resulting in G to A hypermutation of the plus strand sequence, and the loss of genetic integrity(138, 139). However, not all of the hypermutation occurring in Vif-deficient HIV-1 restriction can be attributed to A3G. A3G has a substantial preference for mutating the second C of CC but not TC(140). APOBEC3F (A3F), a sister cytidine deaminase to A3G with high sequence identity, favours TC and also serves to induce G to A hypermutation, resulting in no functional HIV provirus(141). The combined effect of A3G and A3F restriction results in hypermutated proviruses that are largely non-functional, and to decreased levels of cDNA during Vif-deficient HIV infection(142). In the case of wild-type HIV-1, the expression of APOBEC proteins has no effect on viral infectivity. Vif is able to bind A3G and A3F, and recruit them to a cellular ubiquitin ligase complex that comprises cullin5, elongins, Rbx2, and an E2 conjugating enzyme(143-145). The polyubiqutination of APOBEC proteins results in its proteosomal degradation, and thus thwarts their incorporation into nascent viral particles.

Conversely, A3G and A3F do not play a major role in the replication of Vif deficient HIV-2. The Vif-deficient HIV-2 virus is able to replicate in the majority of cell lines, including the lymphocyte and macrophage cell lines that are not permissive to Vif-deficient HIV-1(117). In fact, the Vif of SIVsm is able to rescue the replication of Vif deficient HIV-1 in human H9 cells. Though Vif is conserved across all primate lentiviruses, HIV-2 Vif has only 30% amino acid identity with HIV-1 Vif, and it is evident that APOBEC proteins interact with the two very differently. Regardless of

the amount of A3G expressed in a cell, the infectivity of Vif-deficient HIV-2 remains very high(146, 147). A substantial reduction in infectivity is only observed in the presence of HIV-2 Vif mutants. These Vif mutants confer a reduction in Vif expression, and occur in small motifs essential for Vif-function and stability. It is possible that such mutants make Vif more susceptible to APOBEC proteins, however, the infectivity of these mutants is unaltered by increases in A3G expression. Furthermore, regardless of whether a reduction in infectivity is observed in HIV-2 Vif mutant studies, no alteration is observed in the steady state of deaminase. This indicates that the deamination conferred by A3G is ineffective against HIV-2(117).

Paradoxically, even though A3G is ineffectual for Vif-deficient HIV-2 infectivity, HIV-2 Vif strongly degrades A3G. Furthermore, the effect of HIV-2 Vif is not limited to A3G. HIV-2 Vif can also antagonise A3F, A3H haplotype II, and A3B, with intermediate to high sensitivity, and A3D with low sensitivity. Despite the variation in sensitivity of HIV-2 Vif induced degradation of APOBEC proteins, all known APOBEC proteins except A3A bind to HIV-2 Vif. Now, HIV-2 Vif is better at restricting both A3G and A3F, than HIV-1 Vif. In fact, experiments show that HIV-1 Vif reduces A3G by 90%, and HIV-2 Vif reduces A3G by almost 100%. The effect of HIV-2 Vif antagonism on A3F is even more disparate, showing a 100% reduction compared with only 83% by HIV-1 Vif(141). This is because HIV-2 Vif interacts with separate and distinct domains on APOBEC proteins than those that HIV-1 Vif interacts with. For example, HIV-1 Vif targets a short motif in A3G at position 128-130. This region is essential for HIV-1 Vif's antagonism of A3G. However, HIV-2 Vif targets multiple, non-neighbouring residues found between positions 163 and 321 of A3G(148). This indicates that HIV-2 Vif has a broader range of interaction sites, which may result in a greater propensity to degrade A3G than HIV-1 Vif.

It is evident that APOBEC proteins are less able to resist HIV-2 Vif than HIV-1 Vif, and that HIV-2 Vif is better at antagonising APOBEC proteins than HIV-1 Vif(141). But despite the complex interplay between HIV-2 Vif and APOBEC proteins, to date there is no physiologically relevant evidence for their interaction in the context of HIV-2 infection of the human host. It is possible

that the deamination resulting from APOBEC proteins is countered by another virus mechanism or accessory protein which is presently unknown. Finally the mechanism by which APOBEC proteins are incorporated into nascent HIV-2 virions has yet to be confirmed, and it is very possible that this mechanism may yield important insights into how the effect of APOBEC proteins is managed by the HIV-2 virus.

1.9. The Interplay between HIV-2's Env and Host Factor Tetherin

The Env protein which forms the HIV-2 envelope comprises transmembrane protein gp36 and glycoprotein gp125 and serves to bind and initiate a conformational change in the host cell membrane facilitating virus to cell fusion. HIV-2 Env is not only able to utilise a greater range of coreceptors(36, 37), but its interaction with these coreceptors leading to virus/cell fusion occurs twice as fast(149), when compared to that of HIV-1 Env. In addition to its role in virus entry, HIV-2 Env is responsible for a four to six fold increase in virus export, a function analogous to that of HIV-1 vpu(146). Infection of CD4+ T-cells with HIV-1 mutants lacking the accessory gene *vpu* leads to poor HIV-1 production, and the accumulation of mature virions on cell surfaces and within vacuolar structures. It was observed that the protein tetherin was responsible for this retention of vpu-defective HIV-1 virions. The envelope glycoprotein of some HIV-2 isolates is able to stimulate the release of vpu-defective HIV-1 virions from tetherin positive cells(150).

Tetherin is a type II transmembrane protein which is constitutively expressed on several cell types and is upregulated in response to proinflammatory cytokines(151). Tetherin is found at the cell membrane from which wild type HIV-2 is exported in infected cells, and also in intracellular compartments such as the trans-Golgi network and endosomes which package proteins into membrane-bound vesicles(152, 153). Tetherin is able to facilitate transport between the cell membrane and the golgi apparatus and endosomes using clathrin adaptors(154). Anchored at both the N and C terminals to the plasma membrane, the

extracellular coiled coil domain protrudes(155). Acting as a dimer, tetherin cross links viral membranes, and viral and cellular membranes to each other(156). In this manner, it is able to tether nascent virons preventing their release (Fig2.0). Instead, they remain tethered to the infected cell(157, 158). The virions are then internalised into endosomal compartments where they are degraded. In addition, it is also highy likely that tetherin serves an additional function in cell signalling(159). As antigen presenting cells are targets for HIV-2, the degradation of nascent virions in phagosomes probably leads to the presentation of viral particles to the adaptive immune system. Further, the tetherin promotor contains binding sites for nuclear factor of activated T-cells (NF-AT) and nuclear factor kappa beta NF-kB; IFN and interleukin response elements respectively(160, 161). As tetherin is induced by signals that also trigger type I IFN production, the induction of the type I IFN response and tetherin is linked. Recent data show that tetherin may act as an innate sensor of viral infections that activates NF-kB to induce an inflammatory response(162).

As observed with TRIM and APOBEC proteins, HIV-2 is able to antagonise host restriction factors by the recruitment of ubiquitin ligases, targeting them for degradation. An alternative method employed by HIV-2 to counteract host defences is by altering the trafficking pathways used by the host factors to prevent their expression at the cell surface(163, 164). HIV-2 Env does not reduce total cellular levels of tetherin, as studies have yet to find evidence of tetherin degradation(163). However, HIV-2 Env acts to downregulate the total level of cellular expression of tetherin in two ways: firstly the downregulation of cell-surface tetherin, and secondly through tetherin accumulation in the trans-golgi network, with the net effect of excluding tetherin from virus assembly and export sites. HIV-2 Env attaches its cytoplasmic tail to tetherin, and forms a link between it and the clathrin adaptor complex AP-2(165), which serves to redirect *de novo* or recycling tetherin away from the plasma membrane and to perinuclear compartments via clathrin mediated endocytosis(146, 166). However, while Env promotes the release of virions and facilitates subsequent viral infection,

this may come at a cost to the virus. In the case of HIV-1, the anti-tetherin effect is mediated by an accessory protein: vpu. However, HIV-2 Env is a major structural protein which accompanies tetherin along the clathrin mediated endocytosis pathway, and into its sequestration. This reduction in cellular Env may decrease the levels of Env available during assembly of virus, and thus the amount of virus which can be exported. This is one theory for the reduced virulence of HIV-2.

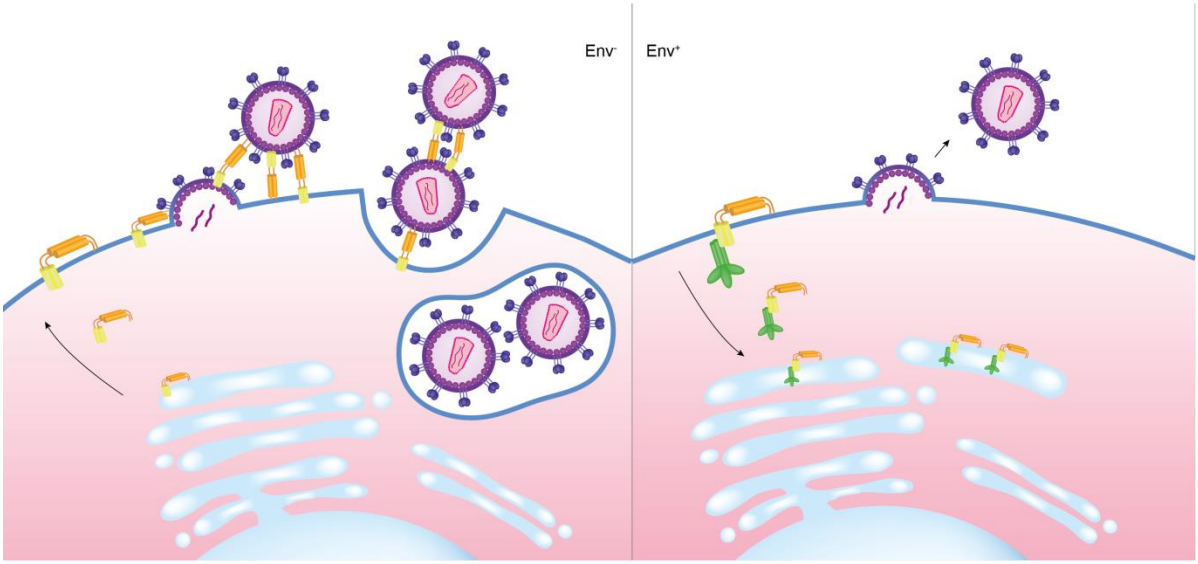


Figure 1.7 Env-mediated Down Regulation of Cell Surface Tetherin

In Env deficient HIV-2 infection, tetherin moves from the ER, through the Golgi network, to the plasma membrane through the anterograde trafficking pathway where it is recycled between the TGN and the cell surface. Tetherin is available to cross link viral membranes, and viral and cellular membranes to each other, resulting in the tethering of nascent virions to the cell. The virions are internalised and degraded. In wild-type HIV-2 infection, HIV-2 Env attaches to tetherin redirecting it away from the plasma membrane and causing its retention in the trans-golgi network. Nascent HIV-2 virions are able to bud from the infected cell.

1.10. HIV-2's Vpx Antagonises SAMHD1

HIV-2 encodes two accessory proteins that are homologous to HIV-1 Vpr: Vpr and Vpx. Vpx is unique to HIV-2, SIVsm, and SIV infecting rhesus macaques (SIVmac), and arose from gene duplication of *vpr* during the evolution of primate lentiviruses(167). *vpx* is considered a paralog of HIV-1 *vpr*(168). They are both are packaged into budding viral particles and are able to assemble into an E3 ligase complex with cellular proteins(169). These functions suggest that Vpx

plays an important role in early infection by targeting certain proteins for proteasomal degradation. Early studies showed that Vpx is dispensable for HIV-2 infection of lymphocyte or monocyte derived immortalised cell lines. However, Vpx deficient HIV-2 shows impaired replication in monocyte-derived macrophages(170), PBMCs, and primary T-cells(171-173). This indicates that Vpx is essential for HIV-2 replication only in non-dividing/slow dividing cell lines. Further, though human cells of myeloid lineage and resting CD4 T-cells possess the necessary receptors to allow HIV-1 capture and entry, they are not readily infected by HIV-1. In this case, HIV-1 infection is rescued by the intracellular delivery of Vpx(174). The reason HIV-1 encounters a blockade in non, and slow dividing cells was found to be due to a host restriction factor: Sterile alpha motif (SAM) and histidine/aspartic acid (HD) domain protein 1 (SAMHD1)(175).

SAMHD1 is found in all cell types and localizes to the nucleus. It possesses nuclease and deoxyribonucleoside triphosphate phosphohydrolase (dNTPase) activity. As a dNTPase, SAMHD1 forms a tetramer and acts to sequester dNTPs, thus keeping them at very low levels outside of the S phase of the cell cycle(176). This reduction in cellular dNTP halts reverse transcription. SAMHD1 also possesses exonuclease activity against single-stranded DNAs and RNAs, and preferentially cleaves 3' overhangs in double stranded substrates. This function further contributes to virus restriction by direct cleavage of viral nucleic acids(177). SAMHD1 is regulated by phosphorylation at position T592. In cycling cells, SAMHD1 is phosphorylated, and while its dNTPase activity is unaltered, it is unable to block retroviral restriction(178). Phosphorylation is mediated through SAMHD1's interaction with the cell cycle regulator cyclin dependent kinase (CDK1) (179). CDK1 is inactive in resting cells, which explains why Vpx is only essential for HIV-2 replication in non or slow dividing cells(180). The C-terminal region of SAMHD1 contains a Vpx binding motif, which is important for the ability of Vpx to degrade SAMHD1. Vpx counteracts SAMHD1 restriction by binding to the restriction factor, and coupling with the Cul4/DDB1/DCAF1 ubiquitin ligase complex(167). This results in the proteasomal degradation of SAMHD1 (Fig 1.8). Vpx requires DCAF1 interaction for efficient macrophage

infection by HIV-2(167). Studies on other DCAF-1 interacting proteins revealed that SAMHD1 forms a complex with cyclin L2 which facilitates SAMHD1 interaction with DCAF1. In fact, knockdown of cyclin L2 resulted in a five fold, three fold, and four fold reduction in the replication of HIV-1, HIV-2, and Vpx deficient HIV-2, respectively(181). While in HIV-2 infection Vpx tags SAMHD1 to DCAF, resulting in its proteasomal degradation, it is evident that cyclin L2 assists Vpx in proteasomal degradation of SAMHD1.

Vpx also antagonises APOBEC3A (A3A) and interferon regulatory factor 5 (IRF5)(182), providing insight into why A3A is the only APOBEC protein which does not bind to HIV-2 Vif. Unlike A3G, A3A exerts its effect in myeloid cells, monocytes, dendritic cells and macrophages. Silencing of A3A in primary macrophages, dendritic cells, and monocytes results in an increased susceptibility to HIV-1 and augmented replication of R5 HIV-1 tropic viruses(183). Further, treatment of myeloid cells with IFNα results in increased resistance to HIV-1 infection which coincides with a three fold increase in A3A expression(125). The antiviral effect mediated by A3A results in a decrease in viral DNA accumulation. Like A3G, A3A is believed to act as a cytidine deaminase, destroying viral DNA through extensive deamination of cytosine(184). In addition, A3A exerts a cytotoxic effect by:

inducing strand breaks in genomic DNA which results in cell-cycle arrest(185)

degrading foreign DNA(186), and

blocking restrotransposition of LINE-1, Alu, and TLR(187, 188).

Experimental evidence shows that A3A, but not A3G, co-immunoprecipiates with Vpx proteins derived from HIV-2 and SIVmac. That same study concluded that SIVmac Vpx, antagonises A3A by inducing its degradation in the early stages of infection(189). Biochemical analysis of A3A activity reveals its enclosure in autophagosomal membrane compartments and cell nuclei(190). In the case of IRF5, Vpx inhibits IRF5-mediated transactivation by a mechanism that is not completely understood. IRF5 is involved in the production of pro-inflammatory cytokines and Type 1 interferon. Expression of Vpx can reduce mRNA levels and protein production of Toll-like

receptor dependent IL6, IL12, and TNFα, which would interfere with the upregulation of interferon-sensitive host genes(182).

The interaction of Vpx with SAMHD1, while fascinating, has yet to result in substantial clinically-relevant findings. Vpx contains a highly conserved poly-proline motif (positions 103-109) which is critical for its effective translation(191). Mutation of multiple prolines to alanines in this motif resulted in minimal expression of Vpx(192). However, a study of Vpx sequences from HIV-2 progressors and non-progressors revealed that only one mutation at position 68 reduced SAMHD1 antagonism *in vitro*(193). It is possible that further studies of the interaction between Vpx and SAMHD1 in multiple populations may reveal clinically relevant mutations in host or virus that may have a clinical impact on the course of HIV-2 disease progression.

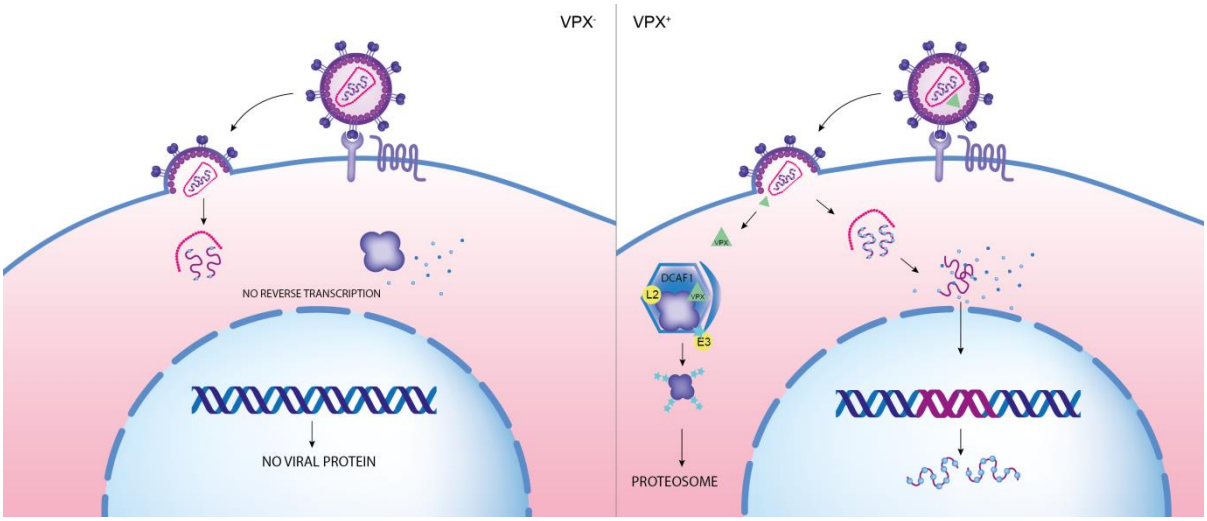


Figure 1.8 SAMHD1 Blocks Reverse Transcription in the Absence of Vpx

In Vpx deficient HIV-2 infection, SAMHD1 reduces the amount of dNTP available for reverse transcription halting the production of viral protein. In wild-type HIV-2 infection, Vpx couples SAMHD1 with the Cul4/DDB1/DCAF1 ubiquitin ligase complex facilitated by cyclin L2, resulting in proteosomal degradation of SAMHD1. As a result, dNTP is available for viral reverse transcription leading to production of viral proteins.

1.11. The Host Restriction Factor TRIM5α and HIV-2

The Tripartite Motif-containing Protein Number 5, Isoform alpha (TRIM5α) is a member of the **Tri**partite **M**otif family of proteins that have the 3 protein domains: RING, B-Box and Coiled Coil. TRIM5α has an additional domain, the PRYSPRY/B30.2 domain that interacts with the retroviral

capsid and is responsible for species specificity (194). The canonical sequence is the first isoform of five, alpha. The species-specific retroviral restriction properties of TRIM5α were discovered because HIV-1 could not grow in certain Old World monkey cells(108). Depending on the mammalian species or cell line, TRIM5α may mediate a post-entry block of lentiviral replication either before or after reverse transcription(195). Each domain plays a different role in retroviral restriction. The RING domain possesses ubiquitin ligase activity. Although the RING domain is required for full antiviral potency, mutants lacking this domain can retain vestigial activity(196). There is no clear function assigned to the B-Box(2) domains, however they appear to be required for antiviral function as restrictive activity is diminished by amino acid substitutions in these domains(197-199). The coiled-coil domain is believed to promote homo-oligomerization, as its deletion prevents TRIM5α protein self-association. The B-Box domains in conjunction with the coiled coil are believed to mediate higher-order self-association of TRIM5α oligomers, which enhances the antiviral activity of TRIM5α *in vitro*(197). TRIM5α proteins assemble into a hexagonal lattice that is complementary to the hexagonal lattice assembled by the HIV-1 capsid protein. The PPYSPRY domain is thought to recognize viral cores, as TRIM5α lacking this domain does not show antiviral activity. Capsid recognition by TRIM5α is thought to be a synergistic combination of direct binding interactions with the PRYSPRY domain, higher-order assembly of TRIM5α, and lattice complementarity(200, 201). The synergistic effect of these domains ensures that TRIM5α is able to multimerise and restrict the productivity of retroviral infection(202-204). Several mechanisms have been proposed with supporting experimental data. Oligomeric TRIMsα recognition of and binding to the incoming intact capsid destabilizes the capsid core, leading to accelerated or premature uncoating, which perturbs reverse transcription(195, 205, 206). The E3 ligase activity of the RING domain can add ubiquitin molecules to itself and to other proteins, leading to proteasome-mediated protein degradation of the capsid-TRIM complex(207). TRIM5α is also a pattern recognition receptor that recognises and binds to the

viral capsid, causing a cascade of innate immune signalling that leads to an antiviral state, restricting retroviral replication(208).

The HIV Capsid encoded by Gag was previously thought to be an inert packaging shell that protects the genomic material. However, it has emerged that it has several domains that allow interaction with host proteins. Recent studies into host restriction factors have shown that the Capsid serves as a pathogen associated molecular pattern (PAMP) for TRIM5α which is a pattern recognition receptor (PRR)(208).

Human TRIM5α has very limited effect on HIV-1, but can restrict N-tropic murine leukemia virus (N-MLV) potently and can limit HIV-2 replication to certain extent. HIV-1 restriction by human TRIM5α maps to the v1 region of the PRYSPRY domain, where a change from arginine or any positively charged residue at position 332 to a proline or any uncharged residue results in potent restriction of both HIV-1 and SIVmac (209, 210). Restriction of N-MLV is different from that of HIV-1, in that a larger area of the PRYSPRY domain is involved; as well as the coiled coil for restriction of N-MLV mutant L117H (209). Human TRIM5α restriction of HIV-2 was also mapped to a single amino acid at position 119 of HIV-2 ROD capsid. HIV-2 viruses with a proline at this position were much more sensitive to human TRIM5α restriction than those without (glutamine or alanine)(89). Further studies showed that presence of hydrophobic amino acids or those with ring structures were associated with sensitivity to TRIM restriction and those with small side chains or amide groups were linked to TRIM5α resistance (211). Three-dimensional modelling of the HIV-1 and HIV-2 capsids showed that the N terminal domains each consists of 7 α-helices from which 3 loops protrude. Position 119 of the HIV-2 capsid is located in the loop between helices 6 and 7. When a proline is present at position 119, the loop between helices 6 and 7 (L6/7) is closer to the loop between helices 4 and 5 (L4/5), a region that directly interacts with cyp A in HIV-1 (211). The presence of hydrophobic or ring-structure residues at position 119 of the capsid maintained a certain conformation at L4/5 that is characteristic of sensitive viruses. Curiously, the equivalent position in N-MLV, at position 110, determined viral susceptibility to

Human TRIM5α: it was suggested that while HIV-1 and HIV-2 are restricted by different mechanisms, human TRIM5α utilises a similar mechanism of recognition for N-MLV and HIV-2 (211). A study from the Caio Community Cohort showed that the presence of three prolines at positions 119, 159 and 178 in the HIV-2 capsid was significantly associated with control of viraemia in the population. The investigators further showed that the presence of these prolines reduced the dimer binding energies of the "PPP" capsid, resulting in a weaker core and more unstable protein (212).

1.12. The Putative Host Restriction Factor TRIM22

Tripartite Motif-containing Protein Number 22, (TRIM22) is a close relative of TRIM5. The two are chromosomally adjacent to each other, with a similar structure and putative function(213). In the case of TRIM22, the RING domain has homology with E3 ligases(214). Unlike TRIM5α, there is only one B-box domain, B-Box-2, present in TRIM22(215). Although the role of the coiled coil region of TRIM22 remains unclear, as in the case of TRIM5α, it has a similar secondary structure and like TRIM5α is also thought to promote self-association and play a role in multimerisation of the protein(216). The role of retroviral restriction by the TRIM22 PRYSPRY domain has yet to be established. That said, the TRIM22 PRYSPRY domain is highly variable which would suggest that this region may confer specificity for viral pathogens(217). TRIM22 has a bipartite nuclear localisation signal (NLS) in the spacer between the coiled coil and the PRYSPRY domain shown to be necessary, but not sufficient, for nuclear localisation. It has been reported that Valine and Cysteine at position 493 and 494 of the PRYSPRY domain are critical for nuclear localisation, while Serine 395, Lysine 396 and Ser 400 in Variable loops 1 ad 3 of the PRYSPRY domain were important for localisation patterns of TRIM22(218). As a result, TRIM22 has been reported to be localised in either the nucleus, cytoplasm, or both, in various cell lines, indicating that cell type specific factors influence its localisation(218, 219).

The antiviral function of TRIM22 is far less well-documented than that of TRIM5α. It has been shown that TRIM22 is expressed in several human tissues and is highly upregulated in response to type 1 and 2 interferons(220-224). The significant upregulation of TRIM22 in response to interferons, together with the finding that TRIM22 has evolved under strong positive selection for millions of years, suggests that TRIM22 is also an antiviral factor. The first experiment to show that TRIM22 could block HIV-1 transcription and replication was carried out in 1995, where exogenous expression of TRIM22 was shown to downregulate transcription from the HIV-1 long terminal repeat (LTR) in Daudi cells(225). Since then, TRIM22 has been found to possess antiviral signatures against HIV-1 both at the level of transcription and virion production in a variety of cells. In U937 cells differentiated by the presence of TRIM22, LTR-mediated transcription was decreased 10 fold in TRIM22 positive cells, an effect which could be diminished by RNAi mediated knockdown in a manner independent of the E3 ligase activity(226). TRIM22 was found to be upregulated in primary monocyte derived macrophages in response to HIV-1 infection, IFNα treatment, or lipopolysaccharide stimulation(227). Cotransfection of TRIM22 with replication defective HIV resulted in diminished production of psuedotyped virus indicating a role in late stage viral production(227). Mechanistically, TRIM22 was observed to interact strongly with HIV-1 gag and was suspected to inhibit HIV-1 replication by two different mechanisms:

i. disruption in the trafficking of gag polyprotein to the plasma membrane mediated by an E3 ligase dependent post translational modification
ii. inhibition of accumulation of intracellular gag protein, either by inhibition of transcription from the LTR or degradation of the gag RNA or polyprotein

in a cell-type dependent manner given the differential localisation patterns(224, 227).

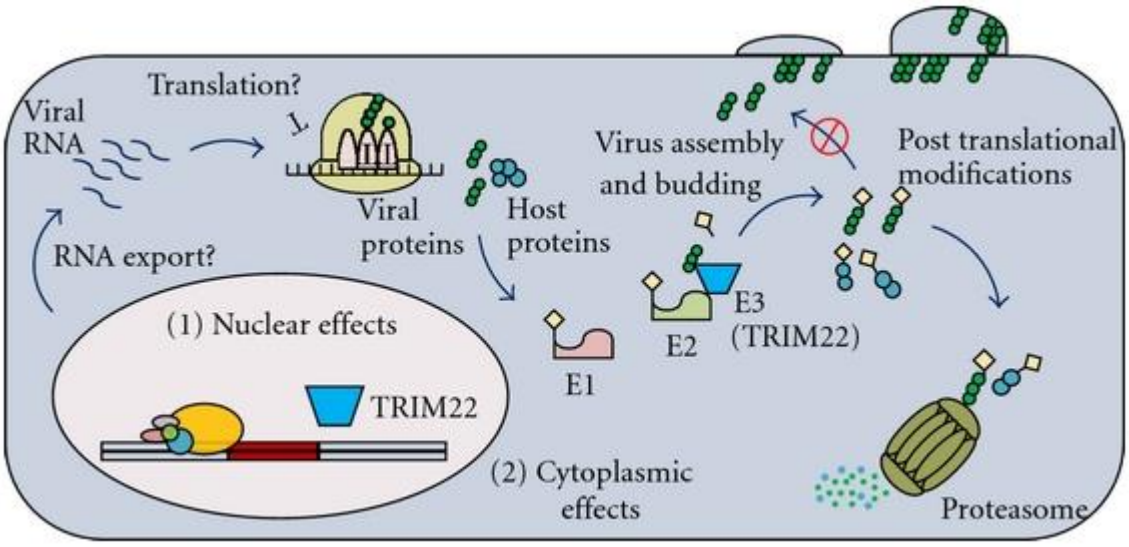


Figure 1.9 Possible Mechanisms of TRIM22 Antiviral Functions

*Reproduced from Hattlmann, CJ., et al. (2012) "*TRIM22: A Diverse and Dynamic Antiviral Protein." Mol. Biology International. 153415. (217) *1. TRIM22 can inhibit viral replication through nuclear-associated effects such as inhibiting viral transcription, RNA export and translation 2. E3 ligase mediated post translational modification of host or viral proteins essential for viral assembly and/or budding, either targeting them for proteasomal degradation, altering their subcellular localization or ability to interact with other entities*

Only in 2011 did Singh *et al*. provide evidence for an anti-HIV role for TRIM22 *in vivo*. In this study, TRIM22 expression was increased in PBMCs of HIV-1 infected individuals, and the level of association was associated with higher CD4+ T-cell counts and lower viraemia in patients from a cohort study in South Africa(223). Data from the Vicenzi group published by Kajaste-Rudnitski *et al*. in 2013 provided the first evaluation of the effect of TRIM22 genetic variation on HIV-1 replication and disease progression. The basis for this work was established in 2011 when the group identified specific TRIM22 variants expressed in HIV-1 non-permissive U937 cell clones(226). Genotyping these two polymorphisms in an HIV-1 infected Caucasian cohort revealed an association between the compound mutations 155N and 242T with advanced disease progression(228).

1.13. Summary

While their suspected founder viruses differ in their respective Simian host, HIV-1 and HIV-2 are highly similar in genetic and virion structure, transmission, infectivity, and cytopathic effect. However, the two are quite different in the proportion of, and rate at which, their human hosts progress to AIDS. Further, a substantial proportion of those infected with HIV-2 remain indistinguishable from uninfected persons in terms of survival, whilst a small proportion of them progress to AIDS in a manner indistinguishable from HIV-1. The genetic, virological, and immunological determinants of these variances require greater exploration in order to find new ways of understanding HIV infection and preventing AIDS deaths from both HIV-2, and HIV-1, where understanding of HIV-2's mechanisms of long-term-non-progression can provide new insights. The longstanding, and well characterised Caio Community Cohort, possessing the highest recorded rates of HIV-2 infection and long-term-non-progression, have not been investigated for the role of host restriction in these traits. Host restriction factors mount a first line response to viral infection, serving to facilitate the sequestration and degradation of key enzymes and components of HIV-2 replication. The restriction factors APOBEC, Tetherin, SAMHD1, and TRIM5α serve to interrupt HIV-2 replication through interaction with vif, env, vpx, and gag respectively. TRIM22 has also demonstrated potent restriction of HIV-1 in early experiments where its close relative, TRIM5α, has failed. Single amino acid changes in TRIM proteins across primate species can confer a change in the type of virus restricted and the level of restriction. Furthermore, the theoretical roles of TRIM proteins in retroviral restriction are diverse; ranging from inhibiting virus transcription, to disrupting virus uncoating and budding, and intracellular signalling. It is therefore possible that polymorphisms in TRIM genes could:

a. modulate how the innate or acquired immune system responds to HIV infection;
b. change how the host cell co-operates with HIV through the restriction factor, therefore modulating the ability of the virus to replicate in human cells

Characterising polymorphisms in TRIM genes, particularly in extreme subsets of HIV disease progression such as those in the Caio Cohort, can assist in re-evaluating the factors influencing susceptibility to HIV infection, and the determinants of the course of disease progression.

1.14. Aims

The thesis aims to investigate whether single nucleotide polymorphisms (SNPs) in TRIM genes are associated with HIV disease progression and to question if this effect is mediated by alteration of resultant protein structure and antiviral function. This can be explored by characterising the variants in TRIM5α and TRIM22 within the same population, to provide insights into which TRIM protein may under selective pressure within the population. The variants observed in TRIM5α and TRIM22 can be examined for associations with HIV-2 acquisition, disease progression as evidenced by CD4+ T-cell count and plasma viral load, and ultimately survival. The thesis also aims to examine the putative effects of these variants on protein structure and function, using *in silico* algorithms based on phylogeny, sequence, and structure information, and also computational methods such as 3D protein models and examination of the impact of polymorphisms to the builds. Given that the restrictive capacity of TRIM22 has yet to be described for any lentivirus other than HIV-1, the work herein aims to pilot test the differential restriction of other retroviruses by TRIM22 and the ability of a deleterious mutation to augment this effect. The deleterious mutation was highlighted from a combination of data from *in vivo* and *in silico* studies.

1.15. Hypotheses

The thesis proposes to test the following hypotheses

1. Description of Polymorphisms in TRIM genes:
    a. PCR based protocols can be optimised to amplify TRIM5α and TRIM22
    b. These amplicons can be sequenced, and sequence data can be examined against published sequences in order to derive amino acid substitutions and infer haplotypes
2. Association Study between TRIM Polymorphisms and HIV-2 Disease Progression:
    a. The above methods can be applied to genomic DNA from HIV-2 infected patients and matched controls.
    b. Using existing data on immunology, clinical phenotype, and viral load, the effect of the variants on disease progression can be quantified using statistical analysis approaches such as regression and mixed effects modelling
3. Computational Analyses of TRIM SNPs:
    a. 3D models of the TRIM proteins and their variants can be generated in the absence of crystal structures
    b. *In silico* prediction algorithms based on phylogeny, sequence, and protein structure can provide inference for effect of non-synonymous polymorphisms on structure and function
4. Functional Analyses of SNPs
    a. TRIM22 and its variants can be overexpressed in TRIM22 deficient cell lines
    b. The ability of TRIM22 and its variants to restrict retroviral infection can be determined by infectivity assays incorporating VSV-G pseudotyped HIV-1, SIVmac239, SIVsm543, and FIV

Chapter 2. General Methods

Each results chapter (Chapters III to VI) of this thesis contains a specific materials and methods section. Therefore, this chapter details only general methods, which are necessary for the execution of the research carried out in subsequent chapters.

2.1. Data Collection from the Caio Community Cohort

This study was conducted among participants aged 15 years and older from the community of Caio, Guinea Bissau. This community has been monitored since 1988 to determine the epidemiological impact of HIV-2 in this community. As described in the introduction, data contributing to this thesis results from three cross-sectional community surveys, which were carried out 1989 to 1991, 1996 to 1998, and 2006 to 2007. In addition, a 2003 study on the clinical manifestations of HIV-2, and a 2010 study on the correlates of elite control, also contribute to data analysed in this thesis. Since 2007, cohort members have had free access to medication and medical care. Antiretroviral treatment became available in 2007, after completion of the 2007 survey, and patients on antiretroviral treatment have been excluded from these analyses.

2.1.1. Sero Survey

Prior to recruitment, the study was discussed with the community. Trained fieldworkers made home visits to study participants. Efforts to contact participants were no different for cohort members with a known HIV status than for other community members. This controlled selection bias. Irrespective of HIV status, participants were only included in the survey if found at home during the survey. Participants were asked to provide consent, complete a demographic and risk factor questionnaire, and provide a 2.5ml venous blood sample. A malaria rapid test and haemoglobin measurement were performed and treatment or referral to the project's physician was provided if necessary. Plasma samples were frozen and transported to Faraja for serology. Patients were tested for HIV, and if positive, invited to visit the project's counsellor to obtain their test result.

2.1.2. HIV-2 Testing

HIV-2 testing was carried out at 5 major time points in the observation period: 1991, 1996, 2003, and 2006, and 2010.

In 1991 and 1996, sera were screened by competitive inhibitory enzyme-linked immunosorbent assay (ELISA) and the Wellcozyme combined test for HIV-1 and HIV-2. Presence of virus was further clarified using a quantitative RNA PCR. RNA was extracted and reverse transcriptase polymerase chain reaction (RT-PCR) was performed in duplicate using oligonucleotides which were specific to the HIV-2 LTR. Products of amplication were quantified by chemiluminescence. The sensitivity of the assay was 100 copies/ml.(81, 96)

In 2003, 2006, and 2010 HIV antibody screening was carried out using the Murex ICE HIV-1-2.0 capture enzyme immunoassay (Murex Diagnostics). Reactive sera were confirmed using a rapid test for viral differentiation (Hexagon HIV, HUMAN GmbH). Indeterminate results were analysed with HIV-2 specific RT-PCR as described above.(229)

2.1.3. CD4+ T-Cell Counts

CD4+ T-cells were counted in participants at 5 major time points in the observation period: 1991, 1996, 2003, and 2006, and 2010.

Prior to 2010, CD4 count was carried out using a manual total white blood cell (WBC) count and differential lymphocyte count based on freshly collected blood. CD4% analysis was done using BD MultiTest reagents and MultiSet fortware (BD Biosciences- Immunocytochemistry Systems) on whole blood stabilised in a 5:1 ratio with TransFix (Cytomark). Absolute CD4+ T-cell counts were calculated using the following formula: (Total WBC) x (% lymphocytes) x (%CD4)(229)

In 2010, CD4+ T-cell subset analysis was performed using fresh whole blood staining with two-colour immunofluorescence reagents (anti-CD45/anti-CD4) and a CyFlow-SL flow cytometer (Partec)(104).

The estimates of absolute CD4+ T-cell count obtained prior to 2010 do introduce considerable variation, especially due to the lack of automated counting machines in the earlier years of the study. However absolute CD4+ T-cell count remains the marker of HIV disease progression, and is the metric on which ART initiation is based. When compared to absolute CD4+ T-cell count, studies have shown that CD4% adds little independent information in relation to progression. (230) Therefore, the remainder of this thesis will report absolute CD4+ T cell count.

2.1.4. Survival and Loss to Follow Up

People who were absent on one occasion were visited a maximum of two additional times by field workers. Whether fieldworkers met the participant or not, if the person was in close contact with those found at the residence and was expected to return, they were considered to be alive. Participants whose whereabouts were unknown were recorded as lost to follow up, and the date of their exit from the trial was recorded as "moved". As a measure of survival, these were not classed as deaths. Only known deaths, reported by the family member or otherwise, were recorded as such in the census even if they occurred abroad.

2.1.5. Study Participants

Altogether, 276 participants were recruited for this study, and divided into a discovery and validation group.

2.1.5.1. Discovery Group

Based on data from the 2010 census, 100 patients were recruited into three groups. These comprised 60 of the surviving HIV-2–monoinfected individuals who were stratified in order to represent "extreme" clinical phenotypes(104). These were examined alongside 40 randomly selected controls.

a. 33 "viraemic controllers" possessing an undetectable viral load (<100 copies per mL)
b. 27 "progressors" possessing a detectable viral load (>500 copies per mL)

c. 40 age and gender matched uninfected controls

2.1.5.2. Validation Group

Based on data the 2003 and 2006 sampling, a further 180 participants were recruited.

a. 70 HIV-2 mono-infected individuals randomly selected

b. 110 age and gender matched uninfected controls

2.2. Methods Involving DNA

The following methods involve the extraction, amplification, and sequencing of DNA.

2.2.1. Salt Precipitation of DNA from Whole Blood

Red blood cells were lysed by mixing whole blood samples with Tris 20mM EDTA ph8.0. The sample stood on ice for 15 minutes, and was then centrifuged at 3500 revolutions per minute (rpm) for 15 minutes allowing white blood cells to be pelleted. The supernatant was aspirated, and the white cell pellet washed again in Tris 20mM EDTA pH 8.0. In order to digest proteins, the sample was agitated with 20µl of a 10mg/ml solution of proteinase K at 42°°C for 12 hours. DNA was precipitated using 2.5ml of 7.5M ammonium acid in the presence of 10ml of chilled 100% ethanol, and then sedimented by centrifuging at 5000rpm for 30 minutes. After discarding the supernatant, the DNA was reprecipitated in 9ml of ice cold ethanol, and sedimented again. DNA was redisolved in 150ul of sterile Tris 20mM EDTA at 58°C for 12 hours. DNA was stored at -20°C until further use.

2.2.2. Amplification of TRIM5α and TRIM22 from DNA

Polymerase chain reaction (PCR) was carried out using the Prime Star GLX kit (Takara Bioscience #R050A) according to manufacturer's instructions. PCR thermocycling conditions are available in Chapter III and Chapter IV for TRIM5α and TRIM22 respectively. The components of each PCR reaction are as follows.

Table 2.1 Master Mix of PCR Reactions using PrimeStar GXL

Reagent	Amount for 1rxn	Final Concentration
Water		
5x PrimeSTAR GXL Buffer	10µl	1x
dNTP Mixture (2.5mM)	1	200µM
Forward Primer	10pmol	0.2µM
Reverse Primer	10pmol	0.2µM
PrimerSTAR GXL Dna Polymerase	1µl	1.25 U/50µl
Template	100ng	100ng

2.2.3. ExoSAP Cleanup of PCR Products

20ul of PCR product was mixed with 0.025ul of Exonuclease I, 0.35ul of SAP, 9.725ul of ultrapure water. The product was then incubated in a thermocycler at 37°C for 30 minutes and then at 85°C for 15 minutes.

2.2.4. Big Dye Terminator Sequencing

Following PCR Amplification, sequencing was performed using the BigDye® Terminator v3.1 Cycle Sequencing Kit (ThermoFisher Scientific #4337455). Briefly, once amplified PCR products were combined with 1x Big Dye Buffer and 1uM of the respective primer. Sequencing was performed using 30 cycles of denaturation at 96°C for 10 seconds, annealing at 50°C for 5 seconds, and elongation at 60 °C for 4 minutes.

2.2.5. Post Sequencing clean-up with Ethanol

To each Big Dye Sequencing Reaction, 4ul of Ethylenediaminetetraacetic acid (EDTA) was added to each sample well, followed by 40ul of 100% ethanol. The plate was incubated at room temperature for 15 minutes. Thereafter sequences were centrifuged at 3400rpm for 15 minutes. Once the supernatant was aspirated, 40ul of 70% ethanol was added to each well, and centrifuged at 1700xg for 15 minutes. DNA crystals were incubated in 10ul of Hi_Di Formamide. Plate was submitted for loading into the 3730XL ABI sequence analyser at the Wetherall Institute of Molecular Medicine, University of Oxford.

2.3. Bacteria and Cell Culture Methods

The following methods involve the growth of plasmids using bacterial culture, and the culture of mammalian cells.

2.3.1. LB Broth and Agar Plates

Lysogeny-Broth (LB) media contained 10g tryptone, 5g yeast extract, 10g NaCl at pH 7.0. It was made up to 1L with Ultrapure water and autoclaved before use. LB Agar contained LB media supplemented with 15g agar. Ampicillin was stored at -20°C, used at 50µg/ml. Agar plates were made with LB agar autoclaved at 15 pounds of pressure per square inch (psi) for 20min and cooled to below 50°C. At this point, ampicillin was added to the plates. Agar was then poured into 9cm Petri dishes under sterile conditions and allowed to cool. Once set, plates were stored at 4°C.

2.3.2. Cell culture

Cells were cultured in incubators set to a temperature of 37°C with 5% CO2. Only two types of media were used, Roswell Park Memorial Institute (RPMI) media (Sigma Aldrich #R8758), or Dulbecco's Modified Eagle's medium (DMEM) (Sigma Aldrich #D6046). Culture conditions are stated for each cell line in the following chapters, however in the majority of cases RPMI was supplemented with 10% foetal calf serum (FCS) and 1% penicillin/streptomycin to form R10 media. DMEM was supplemented with 10% FCS and 1% penicillin/streptavidin to form D10.

2.3.3. Freezing and Thawing of Cells

Freezing media consisted of 10% dimethyl sulfoxide (DMSO) and 90% foetal calf serum (FCS) and was kept at 4°C. Cells to be stored were counted to a final number of 5 million using a haemocytometer. Cells were resuspended in freezing media to a final volume of 1mL in a cryovial. The cryovials were transferred into isopropanol freezing containers, and stored at -80°C in the short term, and at liquid nitrogen thereafter for longer term storage. Frozen cells were

recovered by rapid thawing at 37°C, and resuspension in pre-warmed culture media with 20% FCS. Cells were then washed twice by centrifugation at 1500rpm for 5 minutes. Cells were subsequently resuspended to the appropriate concentration and incubated.

2.3.4. Transformation of Plasmids into Chemically Competent E. coli

DH5alpha competent cells (Life Technologies #18258-012) were thawed on ice. 5ul of vector was added to the cells, and incubated on ice for 30 minutes. The cells were heat shocked at 42°C for 30 seconds in a water bath, and recovered on ice for 2 minutes. 250ul of super optimal broth with catabolite repression (SOC) media was added to the cells, and they were incubated at 37°C for 1 hour at 225 RPM. Varying amounts of each transformation were transformed to pre-warmed LB agar plates containing ampicillin. Plates were inverted and incubated for 12-16 hours at 37°C.

2.3.5. Extraction of DNA from Overnight Culture

A single bacterial colony was selected from the plate of transformants using a sterile pipette tip. The colony was and added to 5ml of LB containing 100ug/ml of ampicillin or 50ug/ml kanamycin depending on resistance of plasmid used, with a sterile pipette tip. The LB broth was then incubated for 16 hours at 37°C in a shaking incubator set to 225RPM. The culture was then centrifuged at 6800xg for 3 minutes at room temperature. The supernatant was discarded, and the plasmids were extracted using the QIAprep Spin Miniprep Kit (QIAGEN #27106). Once extracted, the plasmids were resuspended in 30ul of ultrapure water. The plasmid concentration was measured using the Nanodrop, and stored at -20°C.

2.3.6. Staining of Cells for Flow Cytometry

Cells were recovered from plate or suspension, and washed twice with phosphate buffered saline (PBS) at 1500RPM for 5 mins. Cells were stained with Live/Dead Fixable Aqua Cell Stain (Life Technologies # L34957) and incubated for 20 minutes at room temperature. Thereafter,

cells were washed with PBS twice, and then resuspended in 100ul Beckton Dickenson (BD) Cytofix Cytoperm (BD #554722) and incubated for 20 minutes at 4°C. To remove fixative, cells were washed in 1x BD Perm-Wash Buffer (#554723) and stained with primary antibodies and incubated for 20 minutes. Cells were then washed twice with 1x BD Perm-Wash. If required, secondary antibody staining was conducted and cells were incubated for 15 minutes. Cells were then washed twice with 1xBD Perm Wash, and the cell pellet was resuspended in 400ul of PBS. Cells were enumerated immediately using the Cyan flow cytometer.

Chapter 3. Study of TRIM5α Polymorphisms in the Caio Community Cohort reveal Associations with HIV-2 and Disease Progression

3.1. Introduction

In 2004, Stremlau and colleagues discovered that HIV-1 is able efficiently to enter the cells of Old World monkeys, but encounters a block before reverse transcription. This block was found to be mediated by a cytoplasmic protein: TRIM5α(108). Subsequently, it was observed that rhesus macaques can be infected by SIVmac but are able to restrict HIV-1, N-MLV, and SIV infecting the African Green Monkey (SIVagm) via TRIM5α(231-234). Humans possess an orthologue of TRIM5α which is able to restrict SIV but not HIV-1 or HIV-2.

TRIM5α is a member of the TRIM family of protein isoforms which consist of a RING, 2 B-boxes, and a coiled coil domain. There are over 100 members of the TRIM family, which have been divided into 11 classes(213). The fourth class contains TRIM5, 6, 22, and 34, and they are located in a cluster on chromosome 11. Though the exact mechanism of their action is still being elucidated, they are a family of ubiquitin ligases shown to influence flux through innate signalling pathways by ligating ubiquitin to signal transducers(235). Their expression is sensitive to interferon type 1 and 2, and viral infection(236). They tend to be extremely polymorphic, which has the effect of broadening the antiviral repertoire within a species(113, 237).

Each TRIM5α domain has a role in retroviral restriction. The RING domain possesses ubiquitin ligase activity, and mutants lacking this domain can retain vestigial activity though their cytoplasmic levels are reduced(196). The B-Box[2] domains are required for antiviral activity as restrictive activity is diminished by amino acid substitutions in the B-box domains(198, 199). The B-Box domains in conjunction with the coiled coil are believed to mediate high-order self-association of TRIM5α oligomers(197). Formation of these higher order multimers enhances the antiviral activity of TRIM5α *in vitro*. TRIM5α, unlike other TRIM5 isoforms, possesses an additional carboxy terminal PRYSPRY (also called B30.2) domain which is critical for retroviral capsid restriction(194, 196, 238, 239). The PPYSPRY domain is thought to recognise viral cores, as

TRIM5α lacking this domain does not show antiviral activity. As the interaction between individual capsid monomers and TRIM5α is very weak, capsid recognition by TRIM5α is thought to be a synergistic combination of direct binding interactions with the PRYSPRY domain, higher-order assembly of TRIM5α, template-based assembly, and lattice complementarity(200, 201) TRIM5α proteins assemble into a hexagonal lattice that positions the PRYSPRY domains to match the orientations of their corresponding binding epitopes on retroviral capsids. This assembly serves to significantly amplify otherwise weak intrinsic affinities between the PRYSPRY and capsid subunits, thus being essential to the capsid recognition or even the capsid disruption that accompanies inhibition by TRIM5α(202-204). The synergistic effect of these domains ensures that TRIM5α is able to multimerizse and restrict the productivity of retroviral infection by destabilising the incoming viral capsid, probably through disrupting the timing of retroviral capsid uncoating and subsequent ubiquination of the viral capsid proteins leading to proteasomal degradation. In addition to its role in retroviral restriction, TRIM5α appears to have a second role in innate immune system signalling by activating the tat-associated kinase 1 (TAK1) kinase complex. It was also shown to stimulate NFkB and activator protein 1 (AP-1) signalling and amplifies these activities upon retroviral infection(208). Both signalling mechanisms are dependent upon the ubiquitin ligase activity of TRIM5 and appear to be enhanced following infection with lentiviruses.

Collectively TRIM5α proteins restrict a myriad of retroviruses. As previously mentioned, rhesus macaque TRIM5α can restrict N-MLV, HIV1 and SIVagm. AGM cells have been shown to possess TRIM5α, which restricts HIV-1, HIV-2, N-MLV, Equine Infectious Anaemia Virus (EIAV), and SIVmac infection: yet AGM TRIM5α cannot restrict SIV isolated from AGM or B-MLV(232, 233, 238). Additionally, human TRIM5α is not able to restrict HIV-1 or N-MLV. A pairwise alignment of TRIM5α across primate species (Human, Rhesus macaque, Sooty Mangabey, and African Green Monkey) was conducted and is available in Appendix I. There is very little difference between

the primate TRIM5α proteins: The rhesus macaque TRIM5α, RhTRIM5α, and the human orthologue, HuTRIM5α, share an 89% amino acid identity. AGM TRIM5α and HuTRIM5α share an 84% amino acid identity. Only a difference in three amino acid residues of threonine, phenylalanine, and proline (TFP) at positions 339-341 of RhTRIM5α confer an ability to restrict particular HIV-2 strains(240, 241). Furthermore, a single alteration of arginine 332 (positively charged) to proline (no charge), the residue found in RhTRIM5α, in the HuTRIM5α B30.2 domain, has been shown to create a potent restricting factor for both HIV-1 and SIVmac(242). Capsid specificity of RhTRIM5α has been narrowed down to a 13-amino acid patch of positively selected amino acids in the B30.2 domain(113). A chimeric human/rhesus macaque TRIM5α orthologue was developed incorporating the 13-amino acid patch from RhTRIM5α in place of the corresponding 11-amino acid in HuTRIM5α, and showed a significantly higher level of HIV-1 restriction(194).

It then follows that if subtle amino acid substitutions in TRIM5α can confer a change in the type of virus restricted and the level of restriction, it is very possible that SNPs in TRIM5α could:

a. modulate the innate or acquired immune responses to HIV-1 infection;
b. result in a change in how the host cell co-operates with HIV-1 therefore modulating the ability of the virus to replicate in human cells

Characterising TRIM5α SNPs, particularly in unique subsets of people with different rates of HIV disease progression and in ethnic groups for whom little genetic information is available, can assist in re-evaluating the factors influencing susceptibility to HIV infection, and the determinants of the course of disease progression.

The two seminal papers characterising TRIM5α polymorphisms within human populations hypothesised that protective variants in the TRIM5α coding sequence may be enriched among persons who remain seronegative despite frequent high-risk sexual exposures to HIV-1, or decreased in seropositive individuals. Indicators such as set-point viral load, CD4+ T-cell decline

and altered clinical outcomes were recorded, and association studies were carried out in well over 1500 persons from European(243) and African heritage(244). To determine unreported TRIM5α variants, PCR primers were designed to cover the upstream of the putative promoter, 5'UTF, all exons and introns, and 3'UTR. Several synonymous and non-synonymous SNPs were indentified within the exons, however none were found to exert a significant effect on the likelihood of HIV-1 disease or HIV disease progression, nor were the intronic regions found to have any bearing. In one study, a haplotype containing the substitution R136Q was enriched amongst HIV-1 infected Americans of European heritage(243). In the other study, though R136Q was relatively enriched amongst uninfected Americans of African heritage, the haplotype specifying no variant alleles showed partial restriction of HIV-1 *in vitro,* suggesting that a TRIM5α polymorphism could influence susceptibility to HIV-1(244). Though inconclusive when examined side by side, what these initial studies did reveal were the location of these SNPs on the functional domain regions for the protein. Thereafter, for the sake of efficiency, studies have focussed on SNP discovery using allele-specific PCR or SNP discovery panels in majority European or African-heritage populations. To date, there have been no studies on the effect of TRIM5α polymorphisms in the context of HIV-2 infection.

Since the first report of this virus in 1986, HIV-2 cases remain largely confined to West Africa(26). HIV-1 and HIV-2 share between 30% and 60% nucleotide and amino acid identity and exhibit many similarities in their cytopathic potential and replication kinetics. However, they differ greatly in pathogenicity and transmissibility. Proviral DNA loads are very similar between the two infections but replication rates for HIV-2 are about 100-fold lower than those for HIV-1. Plasma viral loads (VL) for HIV-2 are lower than HIV-1, even though proviral DNA loads are similar at the same stage of disease(75, 77, 79). Clinically, HIV-2 differs from HIV-1 in that a substantial proportion of infected people maintain an undetectable plasma VL for over a decade with no signs of immunodeficiency. Unlike the rare phenomenon of "elite control" in HIV-1

infection, this phenotype is both clinically stable and occurs more commonly than disease progression(80, 245). Generally, this characteristic long-term nonprogression seen in HIV-2 is thought to be a reflection of an effective immune response mounted against the virus, including a vigorous $CD8^+$ T-cell response, maintenance of HIV-specific $CD4^+$ T-cell function, and the presence of a strong neutralising antibody response in many subjects(34, 69, 74). That said, those with HIV-2 who develop disease (15%-20%) do so in a manner and time frame that cannot be distinguished clinically from HIV-1 infection, which can be predicted by detectable vireamia.

In the Caio Community Cohort, 37% of HIV-2 infected participants had stable plasma viral load throughout the 20 years of the study and experienced a normal lifespan. These participants represent a natural model of immunodeficiency virus containment, and a genetic investigation of TRIM5α in such participants in comparison to those who progressed could provide inference regarding beneficial genetic variants involved in long-term non progression. The remainder of this chapter will detail the characterisation of TRIM5α SNPs within a unique cohort of HIV-2 infected, and uninfected at-risk participants.

3.2. Hypotheses and Aims

This chapter aims to investigate whether novel polymorphisms in TRIM5α exist in unique populations not previously characterised such as the Caio Community Cohort. It also aims to investigate whether these polymorphisms in TRIM5α are associated with HIV-2 acquisition or disease progression. This can be explored by characterising the variants, and performing association studies among groups with differing infection statuses and disease progression traits. To investigate the above, this chapter will attempt to address the following hypotheses:

1. Description of SNPs in TRIM5α genes:
    a. PCR based protocols can be optimised to amplify TRIM5α
    b. These amplicons can be Sanger sequenced, and sequence data can be examined against a consensus sequence in order to derive SNPs and infer haplotypes

2. Association Study between TRIM5α SNPs and HIV-2 Disease Progression:
    a. The above methods can be applied to genomic DNA from HIV-2 infected patients and matched controls
    b. Using existing data on immunology, viral load, and survival, the effect of the variants on disease progression can be quantified using statistical analysis approaches such as regression and survival analysis

3.3. Materials and Methods

3.3.1. Sequencing of Full-length TRIM5α from Study Participants

Full-length TRIM5α was sequenced for all participants in both discovery and validation cohorts. Primers were therefore designed based on Genbank Reference Sequence NG_029122 (which is identical to a consensus sequence generated following an aligment of all available TRIM5α transcripts available in NCBI as of 1st March 2013) which comprises 21,148bp, with eight exons, resulting in a 493 amino acid protein. Not all exons coded for protein. While Exons 2-4 code for the RING, BBox-2, and Coiled Coil domains, only Exons 7 and 8 code for the PRYSPRY domain, and Exon 6 for the region which links the two. Primers were therefore designed to flank Exons 2-4, and 6-8. To ensure maximum coverage the primers were crosschecked against polymorphism data from the 1000genomes project and dbSNP(246).

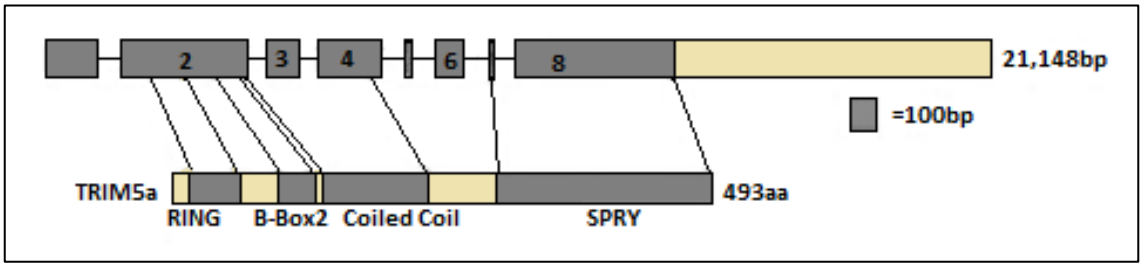


Figure 3.1 Genome and Protein of Human TRIM5α

Schematic illustrating TRIM5 genetic structure (top bar) at 21,148 base pairs (bp), endocing 8 exons. The TRIM5 protein is illustrated underneath (bottom bar); a 493 amino acid protein. Four functional domains are annotated therin, RING, B-Box(2), Coiled coil and PRYSPRY, with the lines connecting the two bars indicating the exons which code for each domain. (not to scale)

Table 3.1 PCR Primers for Amplification of Functional Exons from Human TRIM5α

Exons 2-4	F	TTGGTCCCATTTTAACCTTCC	Tm 53°C
	R	CGAGATCAGTACAAACTATCCAAGG	Tm 55°C
Exons 6-8	F	AGTGGGAAGGCAAACATGAG	Tm 55°C
	R	GCTACGGGAATAAACACTCAA	Tm 53°C

PCR Primers were designed to flank Exons 2-4 and 6-8, in order to capture the regions encoding for functional domains of TRIM5 α. PCR was carried out according to conditions in Table 3.3 and 3.4 for both discovery and validation cohort (n=280)

Table 3.2 Sequencing Primers Used to Characterise Variants in Human TRIM5α

Exon 2, F	TGCAGGGATCTGTGAACAAG
Exon 2, R	CCATCTGGTCCCATTTTCTG
Exon 3, F	TTACCCCTTTCAGCAGGAGA
Exon 3, R	CCATCAGGGAAAGGATGAAG
Exon 4, F	TGTCCCCTCCTTGTGAGAAC
Exon 4, R	TTGATCATCTCCACTACGTCAA
Exon 6 and 7, F	AAGTGCCTCAATGGGTTGAA
Exon 6 and 7, R	GGGTGGCTTATGATAATGTGACT
Exon 8, 1F	CACCCTGCGGCTTATCATTA
Exon 8, 1R	CCAGGATCCAAGCAGTTTTC
Exon 8, 2F	TCCTGGGCTCTCAAAGTATCA
Exon 8, 2R	ACCTGGACAAGAGGTGCTGT

Exons 2-4 and 6-8 were amplified from TRIM5α. Two Sequening Primers were designed to flank each functional exon. Sequencing reaction was carried out as described in Chapter 2, on purified PCR products from both Discovery and Validation Cohorts (n=280)

Table 3.3 PCR Protocol For Amplification of Exons 2-4 of Human TRIM5α

Step	Temperature (°C)	Time (min)	Number of Cycles
Initial Denaturation	95	1	1
Denaturation	95	0.5	40
Annealing	66	0.5	
Elongation	68	3.5	
Final Elongation	68	3.5	1

Conditions for amplification of Exons 2-4 from human TRIM5α from the DNA of participants from both the discovery and validation cohorts (n=280)

Table 3.4 PCR Protocol For Amplification of Exons 6-8 of Human TRIM5α

Step	Temperature (°C)	Time (min)	Number of Cycles
Initial Denaturation	95	1	1
Denaturation	95	0.5	40
Annealing	67	0.5	
Elongation	68	3.5	
Final Elongation	68	3.5	1

Conditions for amplification of Exons 6-8 from human TRIM5 from the DNA of participants from both the discovery and validation cohorts (n=280)

Following PCR amplification, the products were separated on an agarose gel and read under UV light. Products were then cleaned using QIAmp PCR Purification Kit. Sequencing was carried out using Applied Biosystems BigDye Terminatior v3.1 chemistry and the results read on an ABI-3730 DNA analyser. Resulting sequences were viewed using CLCBio Software. Forward and reverse sequences were aligned and the resulting consensus sequence was cleaned via visual examination of the sequence electropherogram. Sequences for each sample were aligned

against the TRIM5α reference sequence and polymorphisms identified. Sequences of polymorphisms and 20 base pairs either side were verified within the UCSC Genome Browser using the blat function(247).

3.3.2. Statistical Analysis

Descriptive statistics and graphs were carried out using R and Graph Pad Prism. Viral loads were transformed into log scale, and the absolute CD4 counts were used and analysed as square root values in order to normalise the data. Kruskal Wallis and Chi squared tests were employed for descriptive data. For clinical data which were not normally distributed, the Wilcoxon test was employed. For normalised data, ANOVA was carried out. Survival analyses were carried out using logrank testing. Kaplan-Meier curves were used to illustrate survival data.

The TRIM5α SNPs were analysed alongside

a. Presence or absence of HIV-2 infection (Status)
b. the CD4+ T cells absolute count (cells/µl)
c. the plasma viral load (copies/ml)
d. controller/progressor status
e. Survival

in order to determine whether they were associated with HIV-2 infection, control of HIV-2 viraemia, and HIV-2 disease progression. Logistic Regression was used for categorical analysis to obtain odds ratios, standard errors, and p values. Linear Regression was used for analysing continuous variables. Statistical analyses were performed using R Software. Haplotype reconstruction and testing for Hardy-Weinberg equilibrium, and haplotype frequency estimations were performed using Arlequin version 3.11. Haploview was used to illustrate linkage. (https://www.broadinstitute.org/haploview/haploview) (Data shown in Appendix)

Statistical Power was calculated for categorical tests and Regression models post-hoc. Effect size was estimated for continuous data using Cohen's F with a cutoff of 0.01. Effect size was

estimated for discrete data using Cramer's V with a threshold of 0.01. The sample size required to achieve a statistical power of 0.8 with a significance of $p<0.05$ with an effect size of at least 0.01 was calculated in R. The sample size is reported for each test up to a maximum of 40,000 participants.

3.4. Results

3.4.1. Demographics of Study Participants

The exome of TRIM5α was sequenced in all participants. The gender distribution and average period of follow-up was consistent across all Study group subsets. Viral load and CD4+ T-cell count were significantly different between HIV-2 viraemic controllers and progressors.

Table 3.5 Demographics of Caio Community Cohort Study Participants

	HIV-2 Infected			p-value*	Uninfected
	All (n=126)	*Controllers (n=33)*	*Progressors (n=27)*		*All (n=150)*
CD4 absolute count (IQR) cells/µL	600.8	749.5	472.4	0.0015	-
Viral Load (IQR)	692.7	193.7	4766.5	0.0001	-
Gender (proportion of females)	77%	79%	74%	NS	76%
Median Survival^	21 years	>50%	23 years	NA	>50%
Follow Up (years)	14.3	14.7	13.0	NS	14.4

Characteristics of HIV-2–infected and Uninfected study participants from both Discovery and Validation Cohorts. Controllers possess an undetectable viral load (less than 100 copies per mL) while progressors have detectable plasma viral loads above 500 copies per mL. (IQR=median)
**1-way Anova,*
^>50% survival indicates that survival exceeded 50% at the longest time point.
Kruskal Wallis Non Parametric Test. NS= non significant

3.4.2. Description of Clinical Data in Discovery and Validation Cohorts

On examination of the entire dataset, it was observed that normalised Absolute CD4+ T-cell count and Viral load were negatively correlated. (Pearson's Correlation = -0.45)

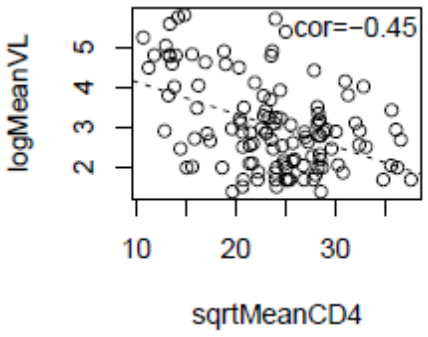


Figure 3.2 Correlation of Normalised Viral Load and Absolute CD4+ T-cell Count

Pearson's Correlation of normalised mean viral load (log10) and normalised mean absolute CD4+ T- cell count (square root) from each HIV-2 infected patient in the discovery and validation cohorts (n=126). Coefficient of -0.45 indicates that CD4 and Viral load were negatively correlated indicating that the high viral load results in CD4+ T-cell depletion as expected in HIV infection.

In the observation periods of 2003 and 2006, Wilcoxon tests showed that the Viral Load was significantly higher in males when compared to females, with p-values of 0.00023 and 0.0032 respectively. ANOVA analysis illustrated that the CD4+ T-cell count was significantly different greater in females when compared to males during the 1996 and 2006 sampling with p values of 0.025 and 0.0242 respectively.

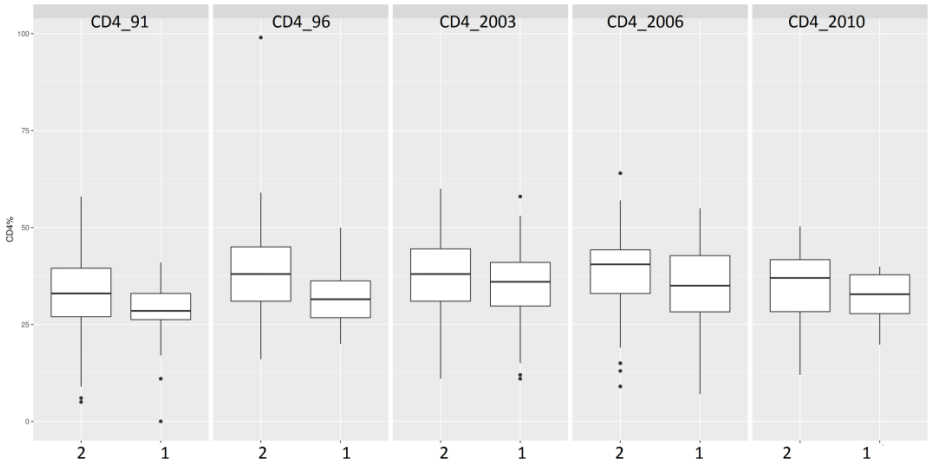


Figure 3.3 Normalised CD4+ T-cell Count Against Gender

Box whisker plots of median normalised CD4+ T-cell counts for Male=1, and Female=2, HIV-2 infected participants in the study. CD4 + T-cell counts were significantly higher females in 1996 (p=0.025) and 2006 (p=0.0242) (ANOVA)

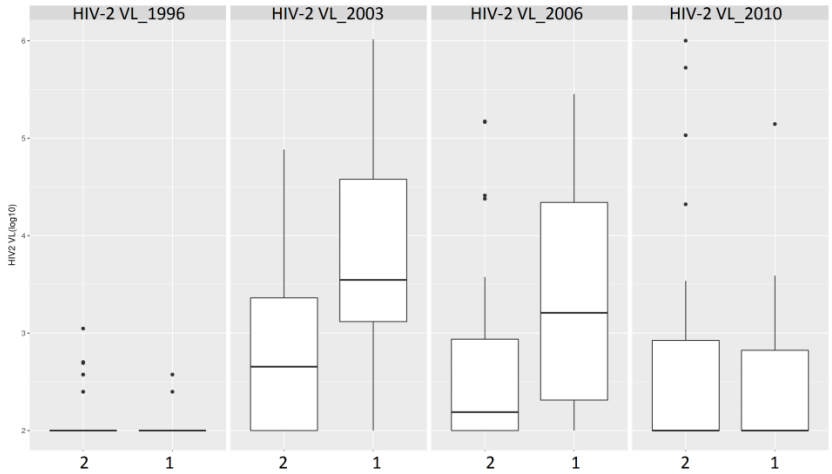


Figure 3.4 Normalised Viral Load Against Gender

Box whisker plots of median normalised HIV-2 viral load for Male=1, and Female=2, HIV-2 infected participants in the study. Viral Load was significantly higher in males when compared to females in 2003 (p=0.00023) and 2006 (p=0.0032) (Wilcoxon)

When taken together, the data would suggest that females showed a reduction in HIV-2 disease progression when compared to their male counterparts.

3.4.3. Examination of Survival among Caio Cohort Subgroups

The association of infection status and gender with survival was next explored. HIV-2 infection resulted in decreased survival (Fig 3.5 A). As expected from the literature, viral load predicted survival among all HIV-2 infected individuals in this cohort, whereby maintenance of an undetectable viral load ensured survival almost equivalent to that of uninfected participants (Fig 3.5 B). Gender was also significantly associated with survival time, with females surviving longer as a whole (Fig 3.5 C), and also once HIV-2 infected (Fig 3.5 D), whether the viral load was detectable or not (Fig 3.5 E) when compared to their male counterparts.

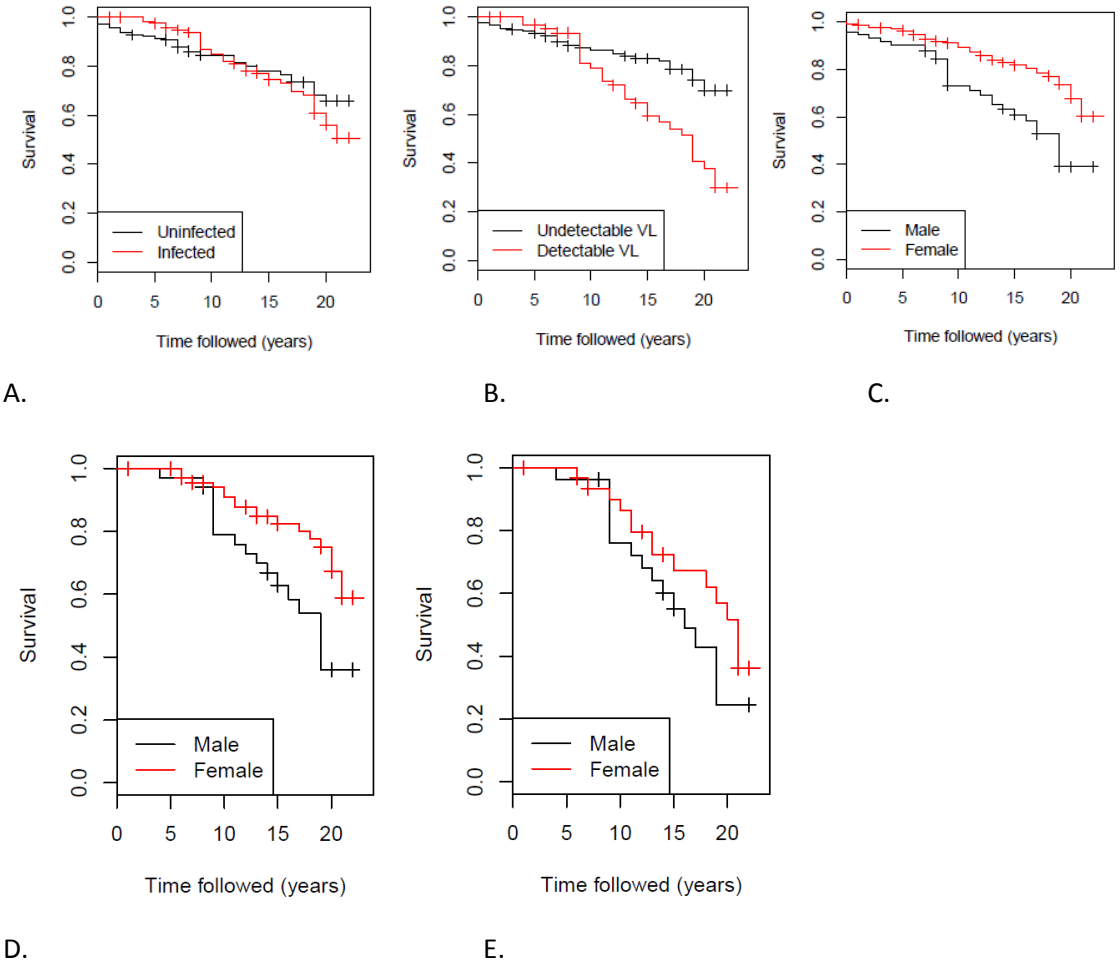


Figure 3.5 Factors which Influence Survival in Caio Cohort

Kaplan-Meier Survival Curves of Study Participants illustrating percentage survival across subgroups. HIV-2 mono-infected vs uninfected participants (A.), HIV-2 mono-infected participants possessing a detectable and undetectable viral load (B.), All male and female participants, infected and uninfected (C.), HIV-2 mono-infected Male and Female Participants(D.), Male and Female Participants with a Detectable Viral Load (E.) Data indicates that in all subgroups, female sex and undetectable viral load predict survival.

3.4.4. Distribution of TRIM5α SNPs Across Study Participant Group Subsets

3.4.4.1. Frequency of TRIM5α Polymorphisms between the Caio Cohort, Yoruba Ethnic Group, and People of African Descent

The majority of polymorphisms were observed in the Coiled Coil Domain, followed by the RING Domain, and finally the PYRPSRY domain. A non-synonymous mutation was observed within the linker region between the coiled coil and the PYRSPRY domain. The allele frequencies within the Caio Community Cohort for both HIV-2 infected participants and controls were calculated, and compared against Yoruba and Africa-at-large allele frequencies available from Hap Map. A significant difference was observed in allele frequencies of L159L, G249D and P479L between the Manjako Community and the Yoruba or Africa-at-large. The minor allele for L159L was enriched in the Caio Community Cohort ($p=0.005$). The minor allele for D249G was enriched in the Caio Community Cohort ($p=0.033$). Finally the minor allele was also enriched in P479L (15.4% compared to 3.35% on average, $p=0.001$). Following correction for multiple comparisons, the enrichment of L159L and P479L remained statistically significant.

3.4.4.2. Frequency of TRIM5α Polymorphisms between Uninfected and Infected Participants

When the frequency of mutations were compared between infected and uninfected participants in both the discovery and validation group (n=100), no alleles were significantly enriched in either subgroup.

Table 3.6 Identification of 8 TRIM5α Polymorphisms From Caio Community Cohort

Exon [a]	Domain [b]	SNP ID [c]	AA change [d]	Allele [e]	AF (n=280) [f]	YRI AF [g]	African AF [h]	p-value [i]	HIV-2 Infected [j] (n=126)	Uninfected [k] (n=150)	p-value [l]	VC [m] (n=33)	RP [n] (n=27)	p-value [o]
2	RING	rs59896509	G31S	C/T	0.9570/ 0.0430	1.0000/ 0.0000	0.9735/ 0.0265	0.149	0.94/0.06	0.96/0.04	0.516	0.94/0.04	0.94/0.04	1.00
2	RING	rs3740996	H43Y	G/A	0.8874/ 0.1126	0.9469/ 0.0531	0.9607/ 0.0393	0.100	0.88/0.12	0.88/0.12	1.00	0.94/0.04	0.88/0.12	**0.041**
2	Coiled coil	rs10838525	R136Q	C/T	0.8744/ 0.1256	0.8805/ 0.1195	0.8971/ 0.1029	0.797	0.84/.015 /0.01	0.87/0.13/0. 00	0.077	0.74/0.26 /0.00	0.72/0.24 /0.04	0.968
3	Coiled coil	rs3740995	L159L	T/C	0.5543/ 0.4457	0.7522/ 0.2478	0.7330/ 0.2670	**0.005 ***	0.49/0.46 /0.05	0.47/0.45/0. 08	0.689	0.50/0.46 /0.04	0.35/0.65 /0.00	**0.007**
4	Coiled coil	rs147678479	L214F	G/A	0.9853/ 0.0147	-	0.9977/ 0.0023	0.992	0.98/0.02	1.00/0.00	0.155	1.00/0.00	0.96/0.04	**0.043**
4	Coiled coil	rs140131201	R238W	G/A	0.9873/ 0.0127	-	0.9924/ 0.0076	0.299	0.99/0.01	0.98/0.02	0.561	0.97/0.03	0.96/0.04	0.700
5	Linker 2	rs11038628	D249G	C/T	0.5084/ 0.4916	0.6327/ 0.3673	0.6876/ 0.3124	**0.033**	0.51/0.39 /0.10	0.53/0.37/0. 10	0.955	0.58/0.33 /0.09	0.52/0.36 /0.12	0.642
8	B30.2/S PRY	rs7104422	P479L	G/A	0.8486/ 0.1514	0.9730/ 0.0270	0.9599/ 0.0401	**0.001 ***	0.81/0.19	0.86/0.14	0.341	0.93/0.07	0.81/0.19	**0.012**

[a] *Exon on which SNP is located* [b] *Domain where SNP is located* [c] *NCBI dbSNP reference* [d] *Predicted amino acid substitution conferred*
[e] *Nucleotide change accompanied by polymorphism based on the TRIM5 consensus sequence*
[f] *Major/minor allele frequency (AF) observed throughout cohort (n=280)*
[g] *Major/minor allele frequency observed in Yoruba (YRI) populations*
[h] *Major/minor allele frequency observed Sub-saharan African and Afro-descendant populations*
*i Chi-squared test, comparing allele proportions across Caio, YRI, and African populations, unadjusted p-value. * remains significant following Bonferroni Correction*
[j] *Major/minor allele frequency observed throughout HIV-2 infected participants (n=126)*
[k] *Major/minor allele frequency observed in uninfected controls (n=150)*
[l] *Chi-squared test, comparing allele proportions in HIV-2 positive and uninfected participants, unadjusted p value*
[m] *Major/minor allele frequency observed in VCs, or Viraemic Controllers, n=33*
[n] *Major/minor allele frequency observed RPs, or Rapid Progressors, n=27*
[o] *Chi-squared test, comparing allele proportions amongst VCs and RPs, unadjusted p value*

3.4.4.3. Frequency of TRIM5α Polymorphisms between Viraemic Controllers and Progressors

The frequency of the non-synonymous mutations between viral controllers and progressors (n=60) was investigated. Heterozygosity (CT) for the polymorphism H43Y was more frequently observed in progressors when compared to controllers (p=0.041). Heterozygosity (AG) (for the synonymous L159L mutation and nonsynonymous mutation L214F was more frequently observed in progressors when compared to controllers (p=0.007 and p=0.043 respectively). The minor allele P479L (CT) was enriched in progressors when compared to controllers. On the other hand, the major allele (CC) seemed enriched in controllers. Following correction for multiple comparisons, no polymorphism was significantly enriched in either viraemic controllers or progressors.

3.4.4.4. TRIM5α Polymorphisms are not associated with Likelihood of HIV-2 Infection in the Discovery Cohort

The likelihood of being HIV-2 infected or uninfected was examined in a logistic regression model (HIV-2 infected=1, Uninfected = 0) against all eight polymorphisms. The frequency of R238W was too low to be examined. None of the SNPs were associated with a likelihood of HIV-2 infection.

Table 3.7 TRIM5α Polymorphisms are not Associated with Likelihood of HIV-2 Infection in the Discovery Cohort

SNP	Allele Change	Allele	Coefficient	P Value	Sample Size Required
G31S	C to T	CT	15.8658	0.9952	7139
H43Y	C to T	CT	-1.0858	0.2534	3023
R136Q	C to T	CT	1.2066	0.0913	256
R136Q		TT	19.1729	0.9977	
L159L	G to A	AG	0.1652	0.9104	18545
L159L		AA	-0.6068	0.6876	
L214F	C to T	CT	18.7467	0.9977	1138
G249D	C to T	CT	0.2808	0.5867	906
G249D		TT	1.6475	0.1307	
P479L	C to T	CT	-0.5098	0.5825	2289

Odds Ratios derived using Logistic Regression model of TRIM5α polymorphisms against HIV-2 infection status (HIV-2 infected =1, and Uninfected = 0) in the Discovery Cohort (n=100) Unadjusted p-values. Sample size required shown to achieve for power of 0.8, effect size cutoff of 0.01, and p value<0.05.

3.4.4.5. TRIM5α Polymorphisms are not associated with Control of HIV-2 in the Discovery Cohort

Control of HIV-2 was investigated using the absolute CD4+ T-cell count and viral load as indicators. ANOVA for normalised CD4+ T-cell data shows there was no association between TRIM5α polymorphisms and Absolute CD4+ T-cell counts.

Table 3.8 TRIM5α Polymorphisms are not associated with Control of HIV-2 in the Discovery Cohort

SNP	Allele Change	Allele	Coefficient	P value	Sample Size Required
G31S	C to T	CT	0.68879	0.867418	6093
H43Y	C to T	CT	-2.19505	0.366184	442
R136Q	C to T	CT	-0.01432	0.993603	1370
R136Q		TT	1.46581	0.801021	
L159L	G to A	AG	1.57068	0.704069	1840
L159L		GG	1.73785	0.680889	
L214F	C to T	CT	6.19876	0.327186	1414
R238W	G to A	GG	3.28129	0.463874	382
G249D	C to T	CT	0.05370	0.968340	119
G249D		TT	-4.74673	0.068386	
P479L	C to T	CT	2.66481	0.256059	11330

ANOVA of TRIM5α SNPs versus Normalised CD4+ T-Cell count in the Validation Cohort. Sample size required shown to achieve for power of 0.8, effect size cutoff of 0.01, and p value<0.05. N/A where sample size sufficient to meet power of 0.8.

Further, when examining normalised viral load in the sixty HIV-2 infected individuals, no significant differences were observed in the case of TRIM5α variants either.

Table 3.9 TRIM5α Polymorphisms are not Associated with HIV-2 Viral Load in the Discovery Cohort

SNP	Allele Change	Allele	Estimate	P value	Required Sample Size
G31S	C to T	CT	-1.3380456	0.211	93
H43Y	C to T	CT	-0.4520796	0.586	277
R136Q	C to T	CT	0.6865163	0.161	131
R136Q		TT	0.5904197	0.634	
L159L	G to A	AG	0.3649868	0.761	1254
L159L		GG	-0.1030846	0.934	
L214F	C to T	CT	-0.7093358	0.614	232
R238W	G to A	GG	0.1918378	0.848	8163
G249D	C to T	CT	-0.0548290	0.885	106
G249D		TT	0.9565547	0.144	
P479L	C to T	CT	-0.2793047	0.648	2498

ANOVA of TRIM5α SNPs versus Normalised Viral Load in the Discovery Cohort Cohort. Sample size required shown to achieve for power of 0.8, effect size cutoff of 0.01, and p value<0.05. N/A where sample size sufficient to meet power of 0.8.

3.4.4.6. TRIM5αPolymorphisms Show No Association with CD4 Count or Viral Load in Discovery Cohort

Graphical analysis of the 60 HIV-2 mono-infected patients in the discovery cohort shows associations consistent with data obtained from the ANOVA and regression models. Regardless of their enrichment in Controllers or Progressors participants homozygous for H43Y, L159L, or P479L wildtype show no significant difference in median Absolute CD4+ T-cell count or Viral Load when compared to their heterozygote counterparts.

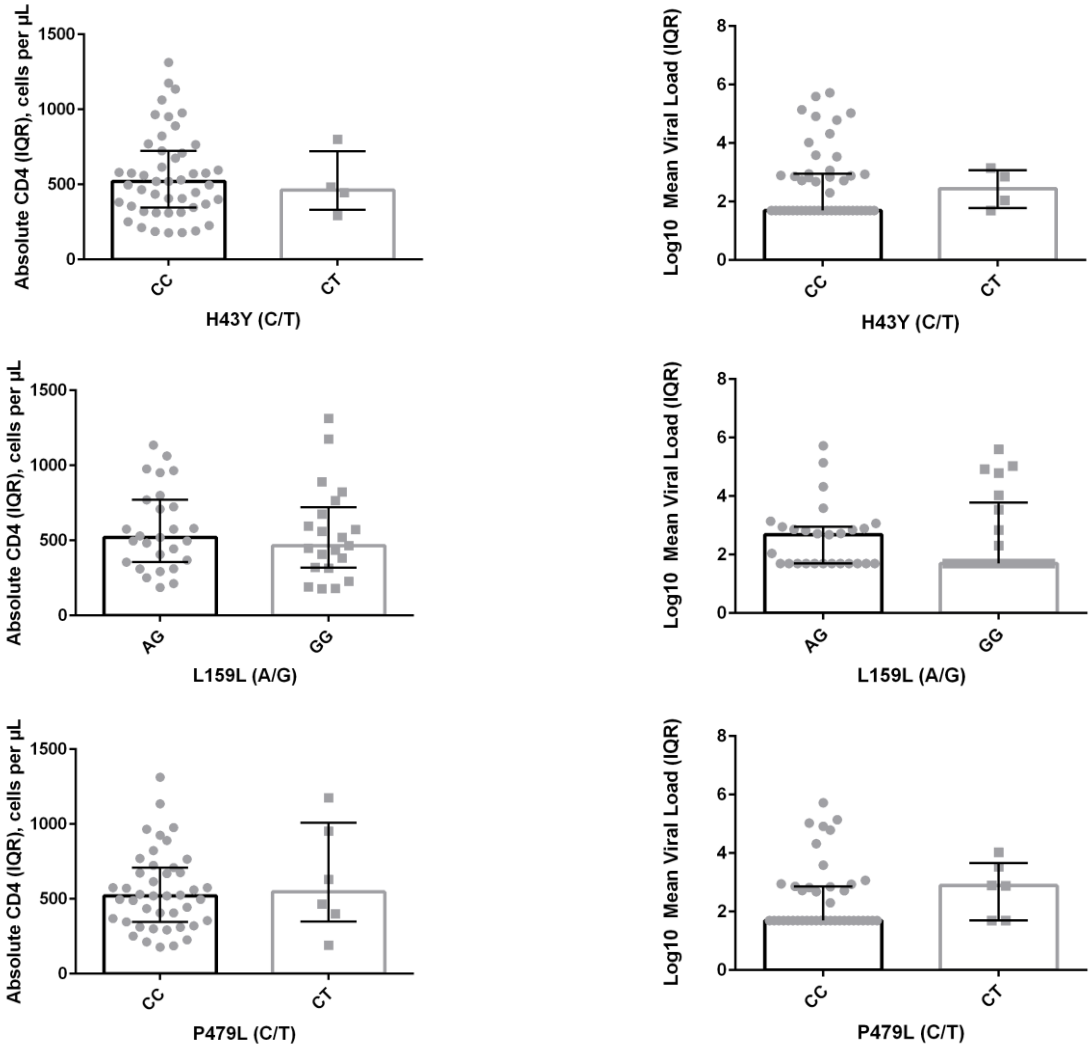


Figure 3.6 TRIM5α Polymorphisms are not associated with CD4+ T-cell Count or Viral Load in Discovery Cohort

Unpaired t-test with error bars showing standard error of CD4+ T-cell Count or Viral Load against H43Y, L159L, and P479L in the Discovery Cohort (n=100). Data shows that no significant difference in CD4+ T-cell count or viral load accompanies the minor allele of H43Y, L159L, and P479L.

3.4.5. Validation Cohort

3.4.5.1. TRIM5α Polymorphisms are not associated with Likelihood of HIV-2 infection amongst Controllers and Progressors in the Validation Cohort

The likelihood of being HIV-2 infected or uninfected was examined in a logistic regression model (HIV-2 infected=1, Uninfected = 0) against all eight polymorphisms. The frequency of R238W was too low for sample size analysis. None of the SNPs were associated with a likelihood of HIV-2 infection.

Table 3.10 TRIM5α Polymorphisms are not Associated with Likelihood of HIV-2 infection in the Validation Cohort

SNP	Allele Change	Allele	Coefficient	P value	Required Sample Size
G31S	C to T	CT	0.43557	0.4005	3627
H43Y	C to T	CT	0.18734	0.6081	4507
H43Y		TT	-13.85555	0.9875	
R136Q	C to T	CT	0.08710	0.8294	2432
R136Q		TT	14.91730	0.9865	
L159L	G to A	AG	0.21637	0.6052	2474
L159L		GG	0.29770	0.5163	
L214F	C to T	CT	1.07951	0.2294	2123
R238W	G to A	GG	0.83800	0.4909	
G249D	C to T	CT	-0.01438	0.9532	79015
G249D		TT	-0.06159	0.8830	
P479L	C to T	CT	0.36455	0.2713	1399
P479L		TT	0.41281	0.7773	

Odds Ratios derived using Logistic Regression model of TRIM5α polymorphisms against HIV-2 infection status (HIV-2 infected =1, and Uninfected = 0) in the Validation Cohort (n=176) Unadjusted p-values. Sample size required shown to achieve for power of 0.8, effect size cutoff of 0.01, and p value<0.05.

3.4.5.2. TRIM5α Polymorphisms Are Associated with Control of HIV-2 in the Validation Cohort

Control of HIV-2 was investigated using the absolute CD4+ T-cell and viral load as indicators. ANOVA for normalised CD4 data showed an association between TRIM5α polymorphism P479L (CT) and CD4 decline (p=0.015). However, this was no longer statistically significant following adjustment for multiple comparisons.

Table 3.11 TRIM5α Polymorphism P479L Associated with Control of HIV-2 in the Validation Cohort

SNP	Allele Change	Allele	P value	Effect Size	Sample Size Required
G31S	C to T	CT	0.979	0.0016	39245
H43Y	C to T	CT	0.502	0.0409	2350
R136Q	C to T	CT	0.952	0.02	8029
L159L	G to A	GA	0.123	0.1269	201
L214F	C to T	CT	0.904	0.0071	39245
R238W	G to A	GG	0.521	0.038	2714
G249D	C to T	CT	0.825	0.0368	2377
P479L	C to T	CT	0.015*	0.1556	N/A

ANOVA of TRIM5α SNPs versus Normalised CD4+ T-Cell count in the Validation Cohort. P479L is associated with CD4 count p=0.015 unadjusted. After adjustment for multiple comparisons, significance at the 5% level is lost. Effect size calculated using Cohen's F. Sample size required shown to achieve for power of 0.8, effect size cutoff 0.01, and p value<0.05. N/A where sample size sufficient to meet power of 0.8.

Graphical analysis of the HIV-2 mono infected patients in the validation cohort show results consistent with data generated from the ANOVA. An unpaired T-test for normalised absolute CD4 in P479L CC and CT participants show a significant difference in absolute CD4+ T-cell count (p=0.0286) where the average CD4 count for P479L CC was 545.8 ± 28.05, n=91, when compared with 412.0 ± 53.30, n=25 for P479L CT.

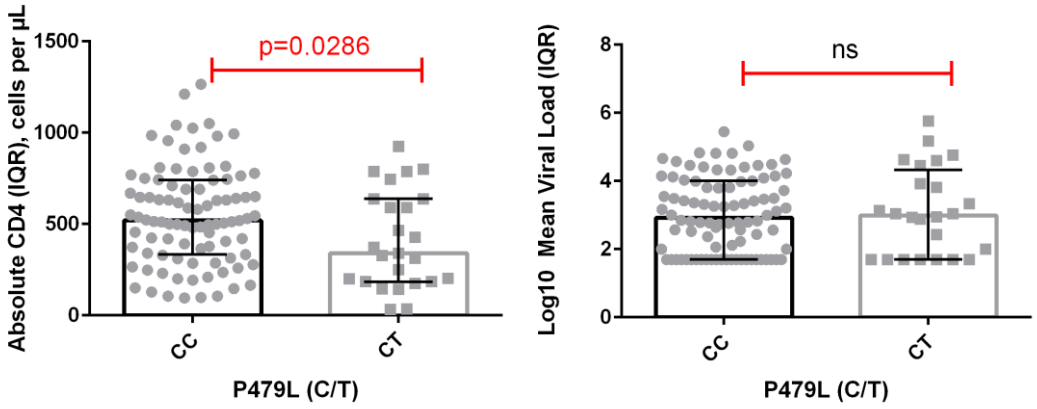


Figure 3.7 P479L associated with CD4+ T-cell Count Decline in Validation Cohort

Unpaired T-test of Absolute CD4+ T-cell Count and Viral Load for P479L. A decrease in CD4 count accompanies the minor allele in P479L, p=0.0286

Further, when examining normalised viral load in the HIV-2 mono infected individuals in the validation cohort, an association was observed between the heterozygote R136Q –CT (p=0.05), however this was no longer statistically significant following adjustment.

Table 3.12 TRIM5α Polymorphism R136Q is Associated with HIV-2 Viral Load

SNP	Allele Change	Allele	Effect Size	P value	Sample Size Required
G31S	C to T	CT	None	0.993	39245
H43Y	C to T	CT	0.0825	0.362	578
R136Q	C to T	CT	0.2303	0.054*	N/A
L159L	G to A	GA	0.0906	0.620	392
L214F	C to T	CT	0.0812	0.356	596
G249D	C to T	TT	0.0399	0.902	2016
P479L	C to T	CT	0.0344	0.717	3326

ANOVA of TRIM5α SNPs versus Normalised Viral Load in the Validation Cohort. Effect size calculated using Cohen's F. R136Q is associated with viral load (unadjusted p=0.054). Upon correction for multiple comparisons, significance at the 5% level is lost. Sample size required shown to achieve for power of 0.8, effect size>0.001, and p value<0.05. N/A where sample size sufficient to meet power of 0.8.

Graphical analysis of the HIV-2 mono infected patients in the validation cohort is consistent with data obtained from the ANOVA. An unpaired T-test for normalised Viral load in R136 CC and CT participants show a significant difference in mean Viral load (p=0.0201) where the average viral load count for R136Q CC was 3.115 ± 0.1103, n=102, when compared with 2.364 ± 0.2105, n=13 for R136Q CT.

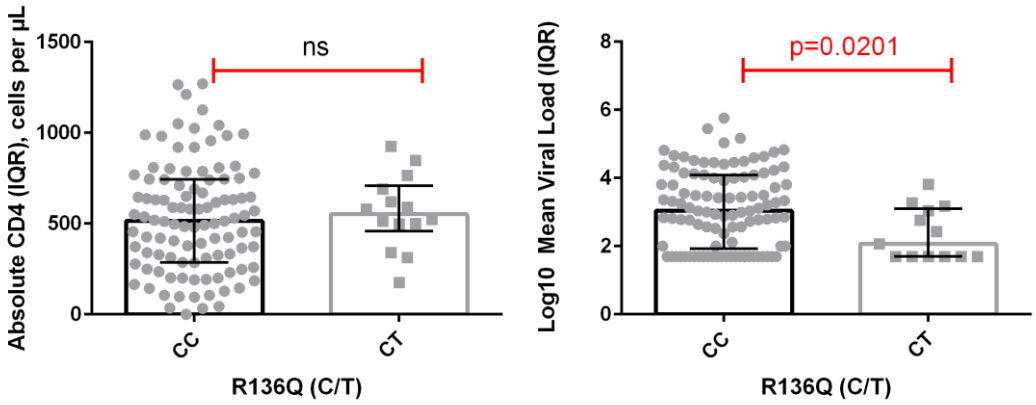


Figure 3.8 R136Q is associated with Decreased Viral Load in Validation Cohort

Unpaired T-test, Error bars show Standard Error. R136Q shows that a significant reduction in viral load accompanies the minor allele CT. p=0.0201)

3.4.5.3. TRIM5α Polymorphisms Do Not Show Differential Survival Amongst HIV-2 Infected Participants

Kaplan-Meier survival, Cox proportional hazard, and logrank tests were performed on HIV-2 mono-infected individuals of the Validation Cohort throughout the 23 year observation period to examine the effect of the seven (R238W excluded) polymorphisms on survival. Participants who were homozygous for SNP R136Q progressed to death 60% more slowly than those who possessed the homozygote wildtype. Those with the SNP R136Q did not achieve fractional survival of 50% whereas those who were wildtype had a median survival time of 16 years. None of the remaining polymorphisms showed any significant association with survival at the 10% level ($p<0.1$).

Table 3.13 TRIM5α Polymorphisms Do Not Show Differential Survival Amongst HIV-2 Infected Participants

SNP	Allele Group	Median Survival	ChiSq	DF	p-value
G31S	CC	17 years	1.913	1	0.1666
	CT/TT	>50% survival			
H43Y	CC	17 years			0.9323
	CT/TT	19 years	0.007	1	
R136Q	CC	16 years	2.81	1	0.0943
	CT	>50% survival			
L159L	GG	18 years	1.139	2	0.5657
	GA	19 years			
	AA	19 years			
L214F	CC	19 years	0.010	1	0.9221
	CT	17.5 years			
G249D	CC	18.5 years	0.883	2	0.6430
	CT	19 years			
	TT	13 years			
P479L	CC	17 years	0.002	1	0.999
	CT	19 years			

Survival data for all HIV-2 infected study participants (n=126) versus TRIM5α genotype. ChiSq shows logrank test for trend, which reports chi-square value associated with appropriate degree of freedom. Table indicates median survival time- defined as the time taken in years at which only 50% of the participants in each genotype group remain alive. Median survival >50% indicates that 50% mortality was not reached within the genotype group during the 23 year study period.

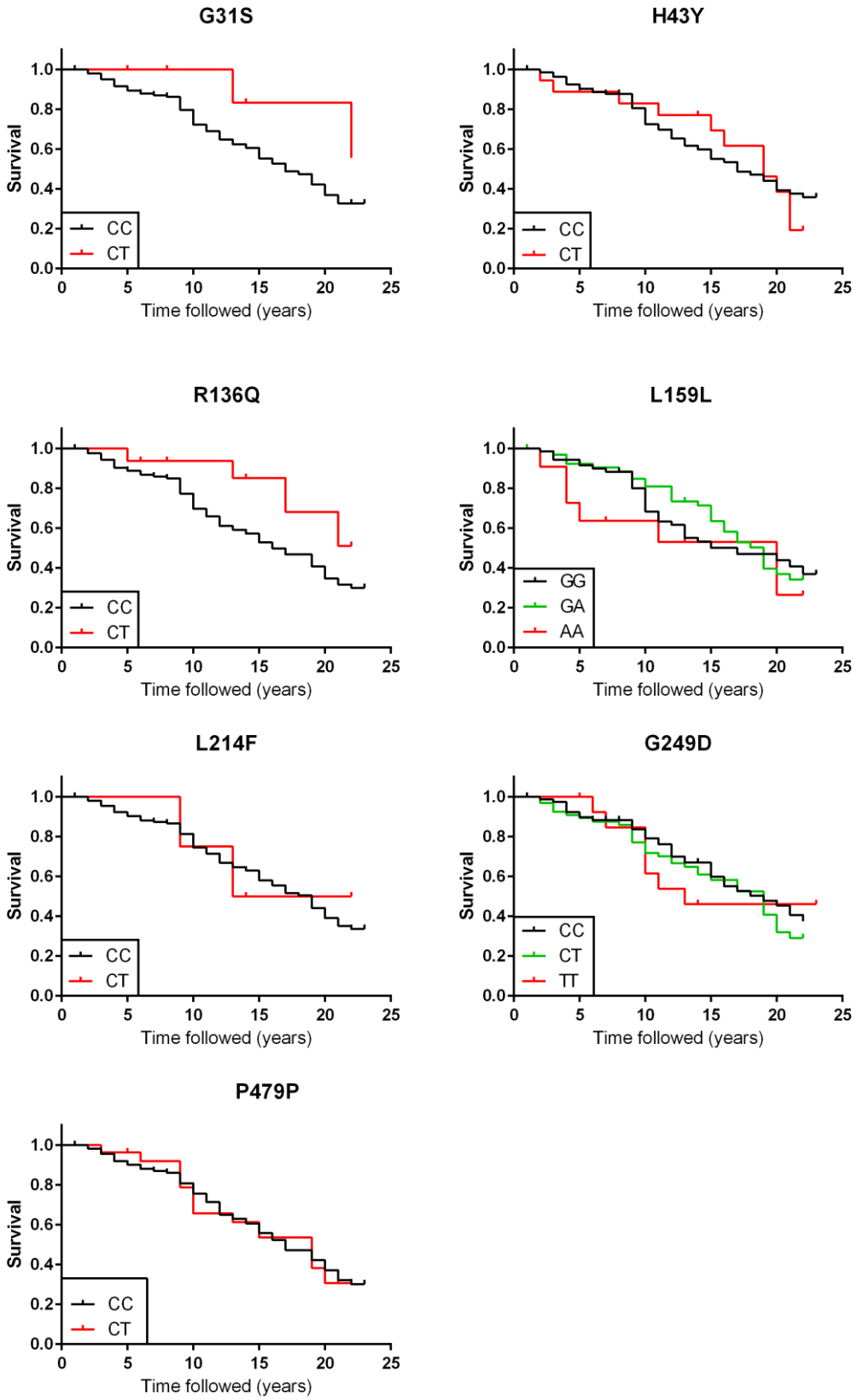


Figure 3.9 TRIM5α Polymorphisms Predict Survival

Kaplan-Meier survival curves illustrating effect of wildtype, heterozygosity, and homozygosity for TRIM5α SNPs on proportional survival amongst all HIV-2 infected study participants (n=126). Adjusted for the effect of gender.

3.5. Discussion

This chapter constitutes the first study of TRIM5α polymorphisms within a distinct and homogenous West African community of this nature, and the first to examine TRIM5α polymorphisms in the context of HIV-2 infected individuals.

Within this subset of the Caio Community Cohort, eight polymorphisms were observed in 276 individuals, encoding 7 non-synonymous missense mutations and 1 synonymous mutation. The majority of non-synonymous single nucleotide polymorphisms (nsSNPs) were found in the coiled coil region, followed by the RING domain, with only one nsSNP observed within the putative capsid-interacting domain; PRYSPRY. When compared to allele frequencies from Africa-at-large, and the Yoruba ethnic group, L159L, D249G, and P479L were enriched in the Caio Community Cohort. With the addition of H43Y which was enriched in HIV-2 rapid progressors, the 4 SNPs appeared relevant as they were differentially enriched among viraemic controllers and progressors within the discovery cohort (n=100). However within ANOVA and regression models, none of the SNPs were significantly associated with likelihood of HIV-2 acquisition or control of HIV-2 as evidenced by CD4+ T-cell count or viral load. Within the validation cohort (n=200), the minor alleles at R136Q and P479L were associated with a reduction in viral load and higher CD4+ T-cell count respectively, however these associations were lost at the 5% level following adjustment for multiple comparisons. Survival analysis showed a significantly positive impact on survival for the R136Q polymorphism which held following adjustment for gender and multiple comparisons at the 10% level ($p<0.1$). Post-hoc power analyses of the other six SNPs show that the effect size required a several fold increase in HIV-2 infected participants to achieve a power of 0.8 and p-values less than 0.05. The study was underpowered to detect the effect of these SNPs on HIV-2 acquisition and disease progression.

Analysis of the entire dataset indicates that two factors influence survival of HIV-2 infected individuals in this cohort: an undetectable viral load, and female sex. In previous studies of this cohort, having a HIV-2 viral load of less than 100 copies/mL predicts normal survival and acts as

a surrogate marker for long-term non progression(79). Undetectable viral load remained an indicator of survival in this subset of individuals in the Caio Community cohort. In light of this, the positive impact on survival in the R136Q polymorphism CT supports initial associations before adjustment suggesting an association between this minor allele and control of viral load.

Sex and gender played significant roles in the analyses. Sex refers to the biological and physiological characteristics that define men and women, whereas Gender refers to the socially constructed roles, behaviours, activities, and attributes that a given society considers appropriate for men and women(248). In this study, females were more likely to have a higher CD4+ T-cell count, and diminished viral load at several time points through the study period. The data suggests that female sex predicts survival within the Caio community cohort in uninfected, HIV-2 infected, and amongst HIV-2 infected individuals who are not controllers. Overall the cohort possesses more female participants of middle age and higher, which may be a result of from a gender trait such as likelihood of participation, a sex trait such as increased risk of HIV-2 acquisition in older women due to mucosal changes, or higher non–HIV-related mortality in men in sub-Saharan Africa (104, 249). Anatomically women are different from men, and possess a different organisation of chromosomes when compared to men, which may result in physiological differences that influence exposure, recognition, and clearance of HIV-2. For example, in the absence of infection, women have higher CD4+ T cell counts and higher CD4/CD8 ratios than age-matched males, whereas males have higher CD8+ T cell frequencies(250-252). Native CD4+ T cells from human females preferentially produce IFN-γ upon stimulation, whereas the same subset from males will produce more IL17(253). Untreated HIV-1 infected women have greater CD8+ T cell activation than men when adjusted for viral load, and over 40% less circulating HIV RNA than men- despite- women having a 1.6 fold higher risk of developing AIDS(254). Further, it also possible that gender traits could also play a role in the increased survival of women with HIV-2 within the Caio cohort. As outlined in the thesis

introduction, women in Guinea-Bissau survive longer when compared to men, but are also less likely to be employed or literate. Differential exposure, health-seeking, and decreased risk-taking behaviours amongst women could result in lower all-cause mortality which may have affected this result.

Several studies have investigated the role of similar TRIM5α polymorphisms on HIV-1 acquisition and disease progression.

The three original studies were carried out in 2006:

1. Wildtype and TRIM5α variants H43Y, V112F, R136Q, R238W, G249D, and H419Y were overexpressed in in Crandel Feline Kidney (CRFK) cells (due to the absence of endogenous TRIM5α). V112F, R136Q, R238W, G249D, and H419Y showed no difference in HIV-1 restriction when compared to TRIM5α wildtype, however it was observed that a relatively highly-expressed H43Y rendered HIV-1 infection identical to that of CRFK cells with no TRIM5α expression(255).
2. In a study of over 3000 participants, roughly half of which were African American and the remainder European American, the frequency of the variant allele encoding an amino acid substitution from R to Q at position 136 TRIM5α was relatively elevated in uninfected individuals, suggesting a possible protective effect. The TRIM5α 136Q protein exhibited slightly better anti-HIV-1 activity in tissue culture than the TRIM5α R136 protein(244).
3. The study most resembling the data presented in this chapter involved investigation of TRIM5α polymorphisms in 110 HIV-1 infected individuals and 96 exposed seronegative individuals with targeted gene sequencing in a further 30 HIV-1 infected individuals who were predominantly men who have sex with men (MSM) of European-American ethnicity. They described 2 non synonymous (L159L, R229R) and eight non synonymous polymorphisms, (H43Y, G110E, V112F, R119W, R119Q, R136Q, V140L, H419Y) a similar

number to those observed within the Caio Community Cohort, the difference being the observation of different polymorphisms altogether, such G110E, V112F, R119W, V140L and H419Y. They did not observe polymorphisms observed in this population such as G31S, R238W, or P479L. No individual polymorphism showed an association with set point viral load after acute infection, CD4+ T cell loss, or frequency among HIV-1 infected subjects relative to exposed seronegative persons. One Haplotype, the 9th most frequently observed, which included the R136Q and no other polymorphism was enriched in HIV-1 infected subjects (243).

Another study of note took place in 2008 with the Amsterdam Cohort which enrolled 364 European MSM, 131 of whom seroconverted during the study. Accelerated disease progression as measured by rapid decline in CD4+ T cell count was observed by individuals with the H43Y minor allele YY when compared to individuals who were wildtype or heterozygous at this position. A protective effect of the 136Q minor allele was observed, not in X5, but only following the emergence of CXCR4 variants(256). A study of 1200 Chinese intravenous drug users (IDUs) observed that the frequency of 43Y homozygote in seronegative IDUs was significantly higher than that in the HIV-1-infected IDUs, suggesting a protective effect among the homozygote subjects(257).

The first investigation of TRIM5α polymorphisms in an African Cohort took place in 2010. The Pumwani sex worker cohort consists of 1032 individuals who are exposed to a 73% risk of seroconversion. 88 individuals are considered resistant, having remained HIV-1 negative for a period of 3 years whilst continuing active sex work. R136Q was enriched in HIV-1 resistant individuals and women with R136Q were less likely to seroconvert(258). Several other studies, included in the summary below, reported no evidence of associations between TRIM5α polymorphisms and HIV-1 acquisition or disease progression, as detailed in the table overleaf.

The literature suggests that variants H43Y and R136Q are typically found at a high allele frequency across multiple ethnic groups. However, the studies reporting the association between H43Y and HIV-1 infection show inconsistent results(243, 244, 256, 257, 259, 260). This could be due to the various routes of transmission, cell lines, and strain of HIV-1 investigated. For example, Feng-Liang et al. studied TRIM5α in over 1,000 IDUs(257); whereas van Mannen et al. studied the effect of the TRIM5a H43Y and R136Q polymorphisms on the clinical course of HIV-1 infection in participants of the Amsterdam Cohort studies, in which all participants are homosexual men, and observed an accelerated disease progression in the group of 43Y homozygotes(256). Additionally, as detailed in Table 3.15, TRIM5α variants were overexpressed in CRFK and HeLa cells, and the effect of HIV-1 replication was studied *in vitro*(255, 261).

Table 3.14 Key Findings from Association Studies between TRIM5α Variants and HIV-1 Disease Progression

Polymorphism	Role in HIV-1	Study
H43Y	Detrimental role when expressed in CRFK cells and challenged in HIV-1 functional studies	(255)
H43Y	Accelerated HIV-1 disease progression in European MSM	(256)
V112F, R136Q, R238W, G249D, and H419Y	No significant difference in HIV-1 restriction when compared to TRIM5α wildtype when overexpressed in CRFK cells and challenged with HIV-1	(255)
H43Y	H34Y enriched in uninfected individuals, yet associated with increased HIV-1 susceptibility in Indian and Japanese participants	(260)
H43Y	1,294 IDU Han and Dai Chinese, H43Y homozygotes for the minor allele were, more frequent amongst HIV-1 infected IDUs	(257)
H43Y, R136Q, G249D, H419Y	Caucasian population shows no evidence of H43Y, G249D, or H419Y association with rapid progression to AIDS. No significant difference in studies of HeLa cells transduced with TRIM5α variants H43Y, R136Q, G249D, H419Y on VSV-G psuedotyped HIV-2 infection.	(261)
H43Y	Population study in two French cohorts, alongside functional study in human Sarcoma (HOS) cells C143, indicating no effect of H43Y in Europeans, Euopean Americans, or Japanese haemophiliacs who are infected with HIV-1 subtype B.	(259)
R136Q	Enriched in HIV-1 infected subjects relative to exposed seronegative Euro-American MSM	(243)
R136Q	Enrichment in HIV-1 resistant but exposed African sex workers, with decreased likelihood of seroconversion	(258)
R136Q	Protective effect in CXC4R mediated HIV-1 infection	(256)
R136Q	Elevated in uninfected African and European American participants, with anti-HIV-1 activity in functional studies	(244)

An *in vitro* transfection study showed that 43Y demonstrated reduced viral restriction when 43Y was expressed in a stably transfected CRFK line. However, such an effect was not found in endogenously expressed 43Y in human B cells, or HeLa cells(255). The data surrounding R136Q is more consistent, indicating a protective effect in X4 mediated HIV-1 infection in two functional studies, an enrichment in uninfected study participants and in at risk, but resistant, sex workers. However one study contradicts these findings showing R136Q enrichment amongst HIV-1 infected individuals in a MSM at risk population.

Residue 43 is located in loop 2 of the RING domain, which confers the E3 ubiquitin ligase activity. Whilst the RING domain is not essential for retrovirus restriction, as previously mentioned, deletion of the RING domain can reduce activity against retroviruses. Loop 2 is believed to be in the junction between E2 and E3 enzymes, and it is hypothesised that altering the positively charged residue (His) to the uncharged residue (Tyr) is likely to enhance the interaction between E2 and E3 enzymes, which may enhance proteasomal degradation of HIV-1. Additionally, the activity of TRIM5α is dependent upon, and enhanced by, its ability to multimerise. R136Q is found in the coiled coil region of TRIM5α, responsible for multimerisation. The arginine residue is positively charged due to the presence of an NH2 group and H2N, whereas the Glutamine residue Q retains the NH2 but loses the positive charge with the addition of O at the H2N position. If R136Q were to act to increase resistance to HIV-1 as suggested in 3 out of four epidemiologised TRIM5α studies, together with the data observed in this chapter regarding HIV-2, it is possible that the change in charge might act to stabilise the hexagonal lattice essential for amplification of the affinities between the PRYSPRY and HIV capsid subunits.

Taken together, the data observed in this chapter suggest that polymorphisms R136Q and P479L in TRIM5α are associated with HIV-2 disease progression. P479L was associated with reduced absolute CD4+ T-cell counts in the validation cohort (unadjusted ANOVA p=0.015, t-test p=0.0286). R136Q was associated with a reduction in viral load (unadjusted ANOVA p=0.054, t-

test p=0.020) and 60% decreased progression to death (logrank p=0.0943). R136Q shows a protective effect against HIV-1 infection and disease progression in several HIV-1 cohorts, and the data shown in this chapter for HIV-2, is consistent with those findings. The small sample size used in this study may conceal stronger associations between P479L and R136Q and progression and control of HIV-2, respectively.

Chapter 4. Genetic Study of TRIM22 in the Caio Community Cohort Reveals Polymorphisms Associated with HIV-2 Acquisition and Disease Progression

4.1. Introduction

The TRIM gene family encodes a diverse group of around 100 proteins that consist of a RING, two B-box, a coiled coil and a C-terminal SPRY domain. Divided into 11 classes, the fourth class of the TRIM family contains TRIM5, TRIM22, TRIM6, and TRIM34, which are located in a cluster on chromosome 11(113). Of the TRIM genes within Class IV, TRIM22 and TRIM5 have a curious relationship: In bovine species, there is an expanded cluster of TRIM5 genes and no TRIM22 gene, whereas in canine species, the TRIM22 gene is present and the TRIM5 gene absent(213, 237). TRIM22 is TRIM5's closest paralogue, and the two possess greater than 60% amino acid identity(113). Studies on TRIM22 sequences from primate genomes show strong evidence of positive selection, typical of that observed for TRIM5(262). However, the positive selection in TRIM5 and TRIM22 were mutually exclusive- with only one gene being positively selected in any given primate lineage(237). The discordant evolution of TRIM5 and TRIM22 suggests that they may both play a similar role in antiviral restriction.

Figure 4.1 TheTRIM6/34/5/22 Cluster on Chromosome 11

Adapted from Sawyer SL, et al., 2007, Discordant evolution of the adjacent antiretroviral genes TRIM22 and TRIM5 in mammals. PLoS Pathog 3(12): e197

In contrast to their evolutionary history, TRIM5 and TRIM22 have many structural similarities which will be investigated further in the following chapter. Like TRIM5α, the TRIM22 RING domain has homology with E3 ligases(214, 263). Unlike TRIM5α, there is only one B-box domain, B-Box-2, present in TRIM22, which has been shown to play a role in the nuclear localization of TRIM22(215, 219). Although the role of the coiled coil region of TRIM22 remains unclear, as in the case of TRIM5α, it has a similar secondary structure and like TRIM5α is also thought to promote self-association and play a role in multimerisation of the protein(216). The role of retroviral restriction by the TRIM22 PRYSPRY domain has yet to be established. That said, the TRIM22 PRYSPRY domain is highly variable which would suggest that this region may confer specificity for viral pathogens(217).

Figure 4.2 Putative Functional Domains in the TRIM22 Protein

The mechanism by which human TRIM5α prevents HIV-1 infection in non-human primate cells has been well established. Rhesus TRIM5α is able to restrict HIV entry by binding to the incoming retroviral capsid using its PRYSPRY domain to cause premature disassembly of the capsid. It has also been postulated to promote the degradation of nascent Gag when expressed at high levels. Human TRIM5α only weakly interacts with the HIV-1 capsid, but has been shown to interact directly with murine leukaemia viruses. Naturally occurring forms of rhesus TRIM5α lacking the PRYSPRY domain fail to restrict HIV-1, showing the pivotal role for the PRYSPRY domain in retroviral restriction. Furthermore, only a difference in three amino acid residues of threonine, phenylalanine, and proline (TFP) at positions 339-341 of Rhesus macaque TRIM5α confers the ability to restrict particular HIV-2 strains. Conversely, HIV-2 restriction by human TRIM5α is weak but specific(264).

TRIM22 is expressed in several human tissues and is highly upregulated in response to type 1 and 2 interferons(222-224). The significant upregulation of TRIM22 in response to interferons, together with the finding that TRIM22 has evolved under strong positive selection for millions of years, suggests that TRIM22 is also an antiviral factor. Since the first experiment in 1995 showing that that TRIM22 could downregulate transcription from the HIV-1 LTR in Daudi cells, (225) TRIM22 has been found to exert antiviral effects against HIV-1 both at the level of transcription and virion production in a variety of cells(227). *In vivo,* TRIM22 expression is increased in PBMCs of HIV-1 infected individuals, and the level of expression was associated with higher CD4+ T-cell counts and lower viraemia in patients followed from acute HIV-1 infection(223). Following the identification of a TRIM22 variant expressed in HIV-1 non-permissive U937 cell clones(226), an association was observed between a haplotype involving the polymorphisms rs795564 and

rs1063303, (corresponding to amino acid substitutions D155N and R242T within the coiled coil region) and advanced disease progression(228). No study to date has investigated polymorphisms in full length TRIM22 within a high risk HIV setting. Furthermore, there is no information in the literature regarding polymorphisms in TRIM22 and their relationship to HIV-2 infection and disease progression.

The previous chapter demonstrated that TRIM5α polymorphisms P479L, and R136Q were respectively associated with CD4+ T-cell decline, and control of HIV-2 leading to increased survival time. R136Q was observed to control viral load and prolong survival; an effect supported by the literature, whereby four previous studies have reported a protective effect of this SNP in HIV-1 infection. However, the literature surrounding polymorphisms in TRIM5α and HIV-1 restriction are often contradictory. Furthermore, human TRIM5α has limited restrictive capacity for HIV-1 and HIV-2 due to poor interaction between the PRYSPRY domain and the viral capsid, whereas studies to date show potent restriction of HIV-1 by TRIM22(224-226). In light of this, together with the clustering of TRIM genes on Chromosome 11, and the discordant evolution between TRIM22 and TRIM5α, is it possible that TRIM22 could have evolved as compensatory antiviral mechanism to TRIM5α, mediating an antiviral effect where TRIM5α cannot? Since TRIM22 appears more adept at HIV restriction than TRIM5α, an investigation into the restrictive capacity of TRIM22 in the context of HIV-2 is warranted. In addition, as amino acid changes in TRIM proteins across primate species can confer a change in the type of virus restricted and the level of restriction, an investigation of TRIM22 polymorphisms and their effect on HIV acquisition and disease progression would also be warranted.

4.2. Hypotheses and Aims

This chapter aims to investigate whether novel polymorphisms in TRIM22 exist in unique populations not previously characterised. It also aims to investigate whether these polymorphisms in TRIM22 are associated with HIV-2 acquisition or disease progression. This can be explored by characterising the variants, and performing association studies among groups with differing infection statuses and disease progression traits. To investigate the above, this chapter will attempt to address the following hypotheses.

1. Description of SNPs in TRIM22 genes:
    a. PCR based protocols can be optimised to amplify TRIM22
    b. These amplicons can be sequenced, and sequence data can be examined against a consensus sequence in order to derive SNPs and infer haplotypes

2. Association Study between TRIM22 SNPs and HIV-2 Disease Progression:
    a. The above methods can be applied to genomic DNA from HIV-2 infected patients and matched controls.
    b. Using existing data on immunology, viral load, and survival, the association between the variants and HIV-2 acquisition or disease progression can be quantified using statistical analysis approaches

4.3. Materials and Methods

4.3.1.1. Genotyping

To genetically characterise TRIM22 within the Discovery Cohort, PCR and sequencing primers were designed first. This was done by downloading all available TRIM22 nucleotide sequences from NCBI and performing an alignment. A consensus sequence was generated, and because the first exon in TRIM5a and TRIM22 is non-coding, PCR primers were designed to flank Exons 2-4, and 5-8, and sequencing primers designed to overlap all functional exons with trifold coverage. Given that a large number of the sequences which formed the consesnsus were based largely on European and African American populations populations, to ensure maximum coverage the primers were cross checked against polymorphism data from the 1000genomes project and dbSNP and dbNSNP(265).

Table 4.1 PCR Primers for Amplification of Functional Exons from Human TRIM22

		Sequence	Tm	GC	Product Size
Exons 2-4	Forward	AATTGCCTTGATGGGGTGAA	57.7	45.0	2790
	Reverse	TGACCCCTGAAATGATGGGC	60.0	55.0	
Exons 5-8	Forward	AATATTAGAGCTCCTCCCTCTTTTT	60.0	55.0	3660
	Reverse	TGCTTTGCTATGGAAGCCCT	59.7	50.0	

PCR Primers were designed to flank Exons 2-4 and 5-8, in order to amplify the regions encoding for functional domains of TRIM22. Full-length amplification was carried out according to conditions in Table 4.3 for the discovery cohort (n=100) and Exons 2-4 only were amplified using DNA from the validation cohort (n=180)

Table 4.2 Sequencing Primers Used to Characterise Variants in Human TRIM22

		Sequence	Tm	GC	Product Size
Exon 2	Exon 2F	TCTCAGCACTGCAGGAGTTT	59.2	50.0	610
	Exon 2R	AGGGCTCTCCGTTCTATGGT	60.0	55.0	
Exon 3	Exon 3F	GGAATGTCAGGTAGGCTCCAA	59.0	52.0	786
	Exon 3R	GCAAAGAGGAAGGGGTCTCC	60.0	60.0	
Exon 4	Exon 4F	GCCAGACTGACTGCCTCTTT	60.0	55.0	427
	Exon 4R	ATAAAAGGGGCATGCCGGG	60.5	57.9	
Exon 5	Exon 5F	GTCCTCTGCAGTGGTCAGTC	60.0	60.0	594
	Exon 5R	TCCGTCGACCCTAACTCTGT	59.0	55.0	
Exon 6	Exon 6F	GGACTCCAGTGTCAGGCATC	60.0	60.0	424
	Exon 6R	CAGACACCTGGTCATTCCCC	60.0	60.0	
Exon 7	Exon 7F	TCCCAGTATTCCCCCTTTCCT	59.0	55.0	277
	Exon 7R	GGAAGAGCTCTGGGTTTGCT	60.0	50.0	
Exon 8	Exon 8F	TGTTCCTCCCCAAACATGCT	59.5	50.0	1017
	Exon 8R	GTCTCAGCACAAGGGCTACT	59.4	55.0	

Exons 2-4 and 6-8 were amplified from TRIM22. Two Sequening Primers were designed to flank each functional exon. Full length sequencing was carried out as detailed in Chapter 2, on purified PCR products from all participants in the discovery cohort (n=100). Subseuqnetly, Exons 3 and 4 were sequenced from all participants in the Validation Cohort (n=180)

Table 4.3 PCR Protocol For Amplification of Exons 2-4 and 5-8 of Human TRIM22

Phase	Step	Temperature (°C)	Time (min)	Number of Cycles
Amplification Phase	Denaturation	95	0.5	20
	Annealing	56	0.5	
	Elongation	68	3.5	
Extension	Final Elongation	68	3.5	1
	Hold	4	Infinity	-

Conditions for amplification of Exons 2-4 from human TRIM22 from the DNA of participants from both the discovery and validation cohorts (n=280), and 5-8 from the discovery cohort (n=100).

Following PCR amplification, the products were visualised on an agarose gel under UV light. Products were then cleaned using QIAmp PCR Purification Kit or ExoSap-IT. Sequencing was carried out using Applied Biosystems BigDye Terminatior v3.1 chemistry and the results read on an ABI-3730 DNA analyser. Resulting sequences were analysed using CLCBio Software. Forward and reverse sequences were aligned against the consensus sequence and SNPs confirmed with visual inspection of the eletropherogram. Polymorphisms found were queried in the UCSC Genome Browser using the blat function, analysed using the Variant Effect Predictor, and checked against dbSNP.

4.3.1.2. Statistical Analyses

Descriptive statistics and graphs were carried out using R and Graph Pad Prism. Viral loads were transformed into log scale, and the absolute CD4+ T-cell counts were used and analysed as square root values in order to normalise the data. For normalised data, ANOVA was carried out. Survival analyses were carried out using Cox Proportional Hazards. Kaplan-Meier curves were used to illustrate survival data.

The TRIM22 SNPs were analysed alongside

a. Presence or absence of HIV-2 infection (Status)
b. the CD4+ T cells absolute count (cells/µl)
c. the plasma viral load (copies/ml)
d. controller/progressor status
e. Survival

in order to determine whether they were associated with HIV-2 infection, viraemic control of HIV-2, and HIV-2 disease progression. For continuous variables, Akaike information criteria was used (AIC) to highlight the best model. Logistic Regression was used for categorical analysis to obtain odds ratios, standard errors, and p values. Linear Regression was used for analysing continuous variables. Statistical analyses were performed using R Software. Haplotype reconstruction and testing for Hardy-Weinberg equilibrium, and haplotype frequency estimations were performed using Arlequin version 3.11. The Broad Institute's Haploview (https://www.broadinstitute.org/haploview/haploview) was used to illustrate linkage. (Appendix III)

4.4. Results

4.4.1. Genetic Diversity of Caio Cohort Participants

The exome of TRIM22 was sequenced in all 100 discovery cohort patients, providing 200 allele positions: n=33 HIV-2 mono-infected controllers (viral load <100 copies per mL) n=27 HIV-2 mono-infected "progressors" (viral load >500 copies per mL) and 40 age and sex matched uninfected controls. Exons 2-4, comprising the Coiled Coil region was successfully sequenced in a further 176 participants: 66 HIV-2 infected, and 110 uninfected age and sex matched controls. The predominantly female and older age of the subjects may be due to an increased risk of HIV-2 acquisition in older women and higher non–HIV-related mortality in men in sub-Saharan Africa(104, 249).

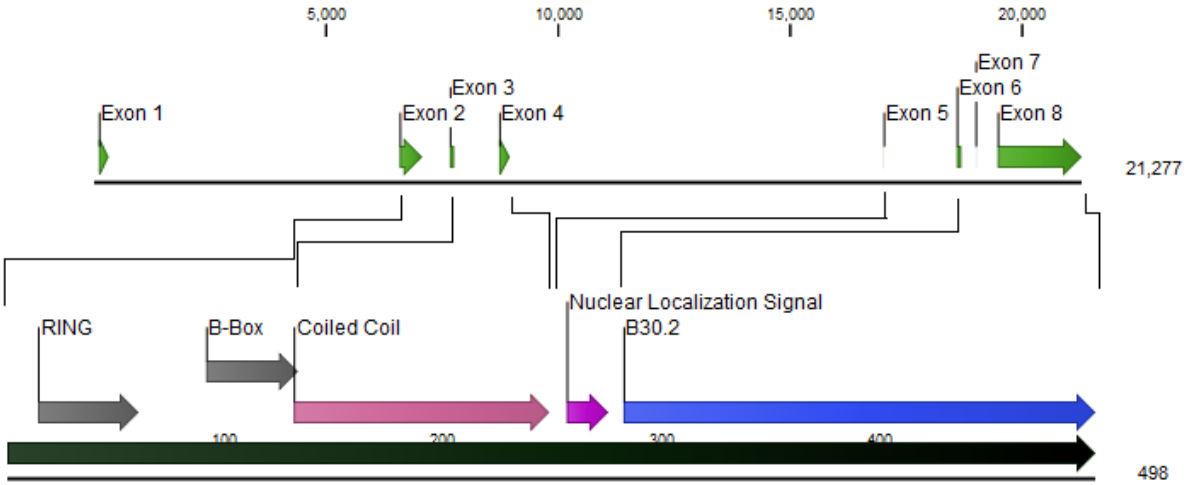


Figure 4.3 Genome and Proteome of Human TRIM22

Schematic illustrating TRIM22 genetic structure (top bar) at 21,277 base pairs (bp), endocing 8 exons. The TRIM22 protein is illustrated underneath (bottom bar); a 498 amino acid protein. Four functional domains are annotated therin, RING, B-Box, Coiled coil and PRYSPRY, with the lines connecting the two bars indicating the exons which code for each domain.

The 200 allele positions revealed 52 single nucleotide polymorphisms: (n=33 intronic, n=21 exonic). Of the 21 exonic SNPs, 10 were synonymous mutations, and 11 were non-synonymous missense mutations. Twenty-one coding polymorphisms were observed, 11 synonymous and 10 non-synonymous. Four polymorphisms were observed in this population; not previously reported. Two resulted in synonymous mutations at residue N188 and V359, and non-synonymous D209K. Though the D360Y mutation has been previously reported, the mutation

frequently observed is a codon change GA**T**/GA**C.** In the Caio Community, the change observed was **G**AT/**T**AT resulting in the same amino acid change due to the redundancy of the amino acid code. Rs7935564 and Rs1063303 encoding for substitutions D155N and R242T respectively, were significantly out of linkage disequilibrium. (Appendix III)

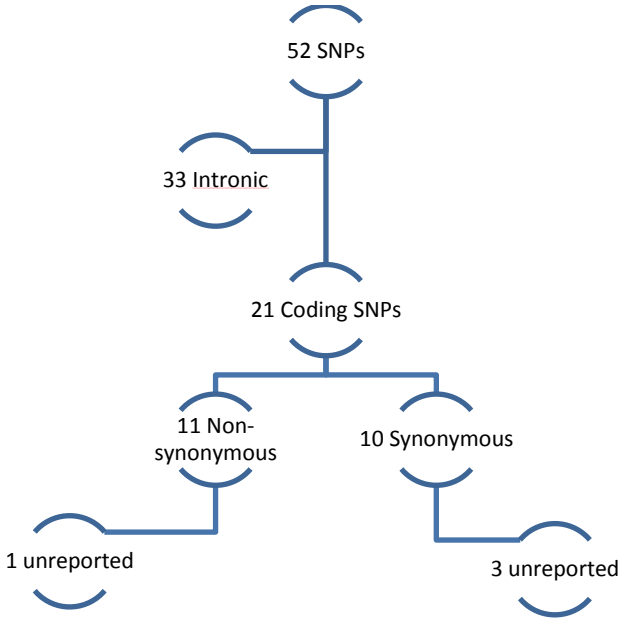


Figure 4.4 Illustration of Polymorphisms Observed in Discovery Cohort

Full length sequencing of TRIM22 in participants from the discovery cohort revealed 52 SNPs, 21 of which were coding. Of these, 11 were non-synonymous and 10 were synonymous. Four novel polymorphisms were discovered.

4.4.2. Distribution of TRIM5α SNPs Across Study Participant Group Subsets

4.4.2.1. Frequency of TRIM5α Polymorphisms between the Caio Cohort, Yoruba Ethnic Group, and People of African Descent

The majority of polymorphisms were observed in the PRYSPRY domain, followed by the coiled coil domain. The RING and B-Box domains were the least polymorphic regions in the protein. No polymorphisms were observed in Exon 5. Data on allele frequencies were obtained for polymorphisms where available. Yoruba allele frequencies were used for comparison, and Africa-at-large (comprising Sub-saharan African and African descendant populations) were also incorporated. No significant differences in allele frequencies following adjustment for multiple

comparisons were observed between those of the Caio Community Cohort, Yoruba, or Africa-at-large. Taken together, where TRIM22 polymorphisms have been previously observed, allele frequencies were typical of those identified in the region.

No polymorphisms observed were out of Hardy-Weinberg equilibrium with the exception of E209K. There were no heterozygotes observed at this position, only two copies of the minor allele (AA) observed at frequency of 2% (Full Data in Appendix III). Haplotype reconstruction and frequency estimations generated 15 possible haplotypes observed at greater than 1%. The three most observed haploytpes are reported below, and the remainder are available in the Appendices.

Table 4.4 TRIM22 Haplotypes Observed in Caio Community Cohort

Frequency	rs10838543(T/C)	rs617352271(A/G)	rs200668710(G/A)	rs7935564(G/A)	rs200924168(G/C)	*5718516(C/T)	rs78484876(T/C)	rs2291842(T/C)	rs1063303(G/C)	rs61735273(C/T)	*5719589(T/C)	rs2291843(A/G)	COSM1509111(G/A)	rs73404240(C/A)	*5730459(G/T)	rs150095329(C/T)	rs61735276(C/T)	*5730458(A/T)
0.261	T	A	G	G	G	C	T	T	G	C	T	A	G	C	G	C	C	A
0.118	C	A	G	A	G	C	T	T	G	C	T	A	G	C	G	C	C	A
0.088	C	A	G	A	G	C	T	T	C	C	T	A	G	C	G	C	C	A

Haplotypes were constructed using Arlequin 3.0, representing the three most frequently occurring alleles at each non-synonymous TRIM22 SNP position.

Table 4.5 Identification of 21 nsSNPs Observed from The Discovery Cohort

Exon[a]	Domain[b]	A.A.[c]	SNP ID/ Position[d]	Allele[e]	Total AF[f]	YRI AF[g]	African AF[h]	p-value[i]	Uninfected[j] (n=40)	HIV-2 +[k] (n=60)	p-value[l]	VC[m] (n=33)	RP[n] (n=27)	p-value[o]
2	RING/ BBox	E83K	rs200668710	G/A	0.989/0.011	-	0.999/0.001	0.316	0.98/0.02	0.95/0.05	0.248	0.97/0.03	0.91/0.09	0.074
		H100H	rs10838543	T/C	0.652/0.348	0.679/0.371	0.679/0.371	0.999	0.42/0.43/0.15	0.40/0.52/0.08	0.220	0.42/0.52/0.06	0.41/0.52/0.09	0.744
		I106I	rs61735271	A/G	0.989/0.011	-	0.977/0.023	0.561	1.00/0.00	0.98/0.02	0.652	1.00/0.00	0.95/0.05	**0.024**
3	Coiled Coil	R150T	rs200924168	G/C	0.907/0.093	-	0.986/0.014	**0.009**	0.92/0.08	0.89/0.11	0.334	0.87/0.13	0.91/0.09	0.366
		D155N	rs7935564	G/A	0.571/0.429	0.473/0.527	0.527/0.473	0.362	0.42/0.37/0.21	0.28/0.50/0.21	0.094	0.26/0.55/0.19	0.33/0.43/0.24	0.237
		A171A	rs779550205	C/T	0.978/0.021	-	-	-	1.00/0.00	0.96/0.04	**0.043**	0.94/0.06	1.00/0.00	**0.013**
4	Coiled Coil	N188N	Chr11:*5719589 (AAT/ AAC)	T/C	0.995/0.005	-	-	-	1.00/0.00	0.96/0.04	**0.043**	0.94/0.06	1.00/0.00	**0.013**
		G208G	rs78484876	T/C	0.874/0.126	-	0.917/0.083	0.249	0.83/0.15/0.02	0.62/0.38/0.00	**0.001***	0.71/0.29	0.67/0.33	0.541
		E209K	Chr11:*5719650 (GAG/ AAG)	G/A	0.978/0.022	-	-	-	1.00/0.00	0.97/0.03	0.079	0.97/0.03	0.95/0.05	0.470
		D214D	rs2291842	T/C	0.772/0.228	0.681/0.319	0.789/0.211	0.163	0.62/0.32/0.03	0.52/0.44/0.04	0.187	0.61/0.36/0.03	0.38/0.57/0.05	**0.005**
		T232A	rs2291843	A/G	0.995/0.005	0.996/0.004	0.999/0.001	0.370	1.00/0.00	0.97/0.02	0.115	0.97/0.03	1.00/0.00	0.081
		R242T	rs1063303	G/C	0.636/0.364	0.628/0.372	0.667/0.333	0.827	0.57/0.35/0.07	0.26/0.55/0.17	**0.001***	0.26/0.58/0.16	0.29/0.52/0.19	0.688
		S244L	rs61735273	C/T	0.962/0.038	-	0.844/0.156	**0.005**	0.92/0.08	0.97/0.03	0.183	0.94/0.06	1.00/0.00	**0.013**
6	PRYSPRY	D282N	COSM1509111	G/A	0.995/0.005	-	-	-	1.00/0.00	0.97/0.03	0.115	0.97/0.03	1.00/0.00	0.081
7	PRYSPRY	T294K	rs73404240	C/A	0.965/0.035	0.500/0.500	0.911/0.089	**0.0152**	0.98/0.02	0.90/0.10	**0.018**	0.85/0.15	0.96/0.04	**0.008**
8	PRYSPRY	L305L	rs61735276	C/T	0.995/0.005	-	0.974/0.026	0.308	0.97/0.03	0.97/0.03	1.000	0.96/0.04	0.95/0.05	0.733
		R321K	rs12364019	G/A	1.000/0.000	0.999/0.001	1.000/0.000	0.999	1.00/0.00	0.99/0.01	0.316	1.00/0.00	1.00/0.00	1.00
		V359V	Chr11:*5730458 (GTA/GTT)	A/T	0.995/0.005	-	-	-	1.00/0.00	0.98/0.02	0.115	0.97/0.03	1.00/0.00	0.081
		D360Y	Chr11:*5730459 (GAT/TAT)	G/T	0.990/0.010	-	-	-	1.00/0.00	0.97/0.03	0.079	0.97/0.03	0.94/0.06	0.306
		L378L	rs61735325	G/T	0.995/0.005	-	-	-	1.00/0.00	0.98/0.02	0.115	0.97/0.03	1.00/0.00	0.081
		T415I	rs150095329	C/T	0.990/0.010	-	0.985/0.015	1.000	100/0	0.97/0.03	0.079	1.00/0.00	0.95/0.05	**0.024**

Full length TRIM22 was sequenced from participants in the discovery cohort (n=100), and 21 coding polymorphisms were identified. The minor allele of rs78484876 and rs1063303 are significantly enriched in HIV2+ individuals(p=0.001) following adjustment for multiple comparisons/

a Exon on which SNP is located b Domain where SNP is located c Predicted amino acid substitution and position

d SNP identified using NCBI dbSNP reference, novel polymorphisms annotated according to the GrCh37 assembly; chromosome number, position, and codon change

e Nucleotide change accompanied by polymorphism based on the TRIM22 consensus sequence f Major/minor allele frequency (AF) observed throughout cohort (n=100)

g Major/minor allele frequency observed in Yoruba (YRI) populations h Major/minor allele frequency observed Sub-saharan African and Afro-descendant populations

*i Chi-squared test, comparing allele proportions across Caio, YRI, and African populations, unadjusted p-value. * remains significant following Bonferroni Correction*

[j] Major/minor allele frequency observed throughout HIV-2 infected participants (n=60) [k] Major/minor allele frequency observed in uninfected controls (n=40)

[l] Chi-squared test, comparing allele proportions in HIV-2 positive and uninfected participants, unadjusted p value

[m] Major/minor allele frequency observed in VCs, or Viraemic Controllers, n=33) [n] Major/minor allele frequency observed RPs, or Rapid Progressors, (n=27)

[0] Chi-squared test, comparing allele proportions amongst VCs and RPs, unadjusted p value

4.4.2.2. Frequency of TRIM22 Polymorphisms between Uninfected and Infected Participants in the Discovery Cohort

When the frequency of non-synonymous mutations were compared between infected and uninfected participants in the discovery group (n=100), heterozygous (GC) and homozygous (CC) rs1063303 were enriched HIV-2 infected individuals when compared to uninfected individuals following adjustment for multiple comparisons (p=0.001), where the homozygous wildtype (GG) was more frequently observed. Heterozygotes (TC) for rs78484876 was also enriched amongst HIV-2 infected participants when compared to individuals homozygous for the wildtype following correction for multiple comparisons (p=0.001).

4.4.2.3. Frequency of TRIM22 Polymorphisms between Controllers and Progressors in the Discovery Cohort

The frequency of the non-synonymous mutations between viral controllers (n=33) and progressors (n=27) was investigated. There were no statistically significant difference in allele frequency between controllers and progressors following adjustment for multiple comparisons.

4.4.3. Discovery Cohort

4.4.3.1. TRIM22 Polymorphisms Associated with Likelihood of HIV-2 infection amongst Viral Controllers and Progressors in the Discovery Cohort

The likelihood of being HIV-2 infected or uninfected was examined in a logistic regression model (HIV-2 infected=1, Uninfected = 0) against the non-synonymous polymorphisms. Of all the SNPs, only rs1063303 and was found to be associated with HIV-2 infection risk with $p<0.05$. Persons possessing the wild-type of rs1063303 (GG) were found to be 2.5x more likely to be uninfected than those who possessed a mutation at this position (OR 2.47, p=0.005). When adjusted for gender and multiple comparisons, the data was no longer statistically significant at the 5% level. However, the data would suggest that the rs1063303 mutation makes a person more likely to be HIV-2 positive, thus exerting a detrimental effect.

Table 4.6 TRIM22 Polymorphisms Associated with Likelihood of HIV-2 Infection

Polymorphism	Base Change	Genotype	OR	*p-value*	*p-value corrected*
rs1063303	G to C	GG	-2.4674	0.0045	0.0896

The likelihood of being HIV-2 infected or uninfected was examined in a logistic regression model (HIV-2 infected=1, Uninfected = 0) against the non-synonymous polymorphisms. Participants with rs1063303 (GG) found to be 2.5x more likely to be uninfected than those who possessed a mutation at this position (OR 2.47, p=0.005). p-value corrected for multiple comparisons and effect of gender.

4.4.3.2. TRIM22 Polymorphisms are Associated with control of HIV-2 in Discovery Cohort

Control of HIV-2 was investigated using the absolute CD4+ T-cell count and viral load as indicators. When examining viral load, no significant differences were observed where wild type and TRIM22 variants were concerned. In a linear regression model for normalised CD4+ T-cell data, nsSNPs rs1063303 and rs7935564 were associated with lower and higher CD4+ T-cell counts respectively. Participants with rs106330 wildtype had a significant increase in average CD4+ T-cell count (Est 6.04, p=0.007), all factors held constant. Conversely, participants with the wildtype for rs7935564 had a significant decrease in CD4 on average when compared with persons positive for the mutation (Est -3.7188, p=0.04). Following adjustment for multiple comparisons and gender, none of the associations retained statistical significance at the 5% level. Overall the data would suggest a protective effect for the rs7935564 mutation in HIV-2 disease progression, a detrimental effect for the rs1063303 mutation in the case of HIV-2 infection and disease progression, and a protective effect for the rs10838543 mutation in terms of HIV-2 infection.

Table 4.7 TRIM22 Polymorphisms Associated with Absolute CD4+ T-cell Counts

Polymorphism	Base Change	Genotype	Estimate	*p-value*	*p-value corrected*
rs1063303	G to C	GC	4.0971	0.06210	ns
rs1063303	G to C	GG	6.0421	0.00653	0.1306
rs7935564	G to A	GG	-3.7188	0.04181	0.8362

Linear regression model for normalised CD4+ T-cell data, nsSNPs rs1063303 and rs7935564 were associated with lower (p=0.007) and higher CD4+ T-cell counts (p=0.04) respectively.

4.4.3.3. TRIM22 Polymorphisms Are Associated with CD4 Count and Infection Status in Discovery Cohort

Graphical analysis of the 60 HIV-2 mono-infected patients show is consistent with data obtained from the linear and logistic regression models. Participants with rs7935564 AA show a higher median absolute CD4+ T-cell count when compared to participants with rs7935564 GG (wildtype). Participants with any allelic distribution of rs7935564 were not enriched among uninfected or infected groups. Conversely, rs1063303 showed a stepwise decrease in median absolute CD4+ T-cell count through participants who possessed the wildtype, heterozygote, or homozygote. When examining the proportion possessing each 1063303 genotype, 1063303-CC had a greater proportion of HIV-2 infected participants than 1063303-GG, the wildtype. The trend observed was once again stepwise in manner.

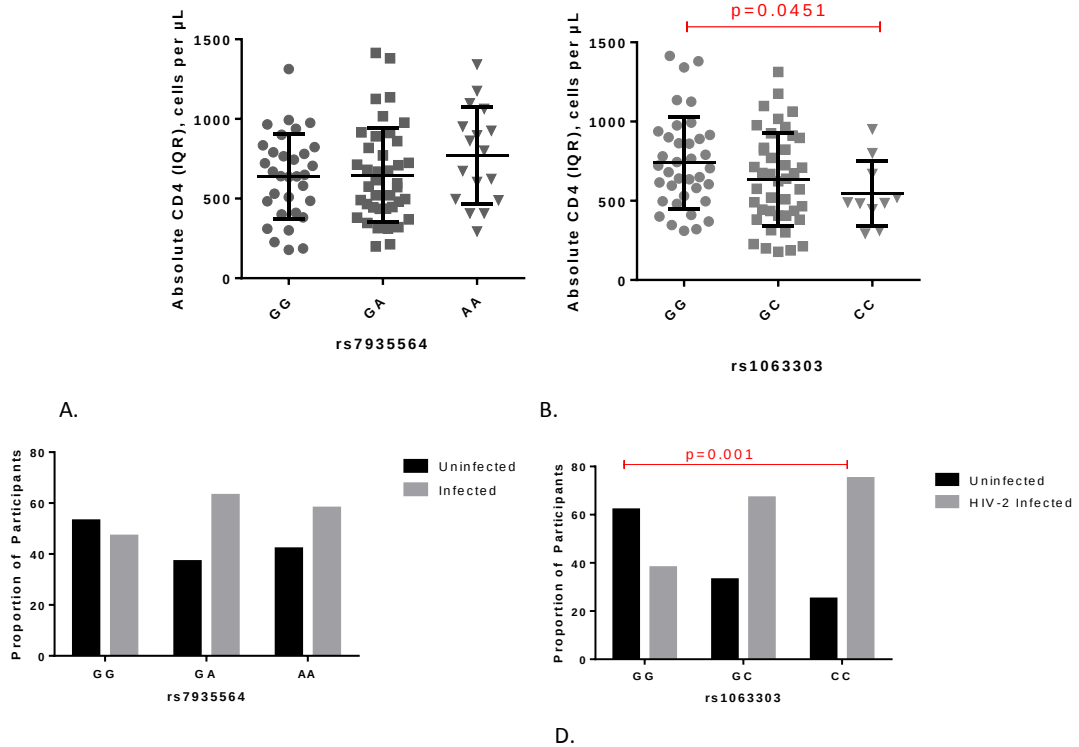


Figure 4.5 TRIM22 SNPs are Associated with Absolute CD4+ T-cell Count and Infection Status in Discovery Cohort

T-tests of Absolute CD4+ T-cell Count in HIV-2 monoinfected individuals with rs7935564 genotypes (A), and rs1063303 genotypes (B.) Chi Squared Test of Proportion in Participants who are either Uninfected or HIV-2 Infected according to rs7935564 genotype (C), and rs1063303 genotype (D).

4.4.4. Validation Cohort

4.4.4.1. TRIM22 Polymorphisms Are associated with Infection Status in Validation Cohort

When examining all participants in the validation cohort (n=176) there were no significant differences between the viral load or CD4+ T-cell count of patients possessing any of the rs7935564 or 1063303 alleles (Fig 4 A-D). However, twice as many participants with the 1063303 CC genotype were HIV-2 infected, when compared to a greater proportion of Uninfected participants who possessed the wildtype. (p=0.0536)

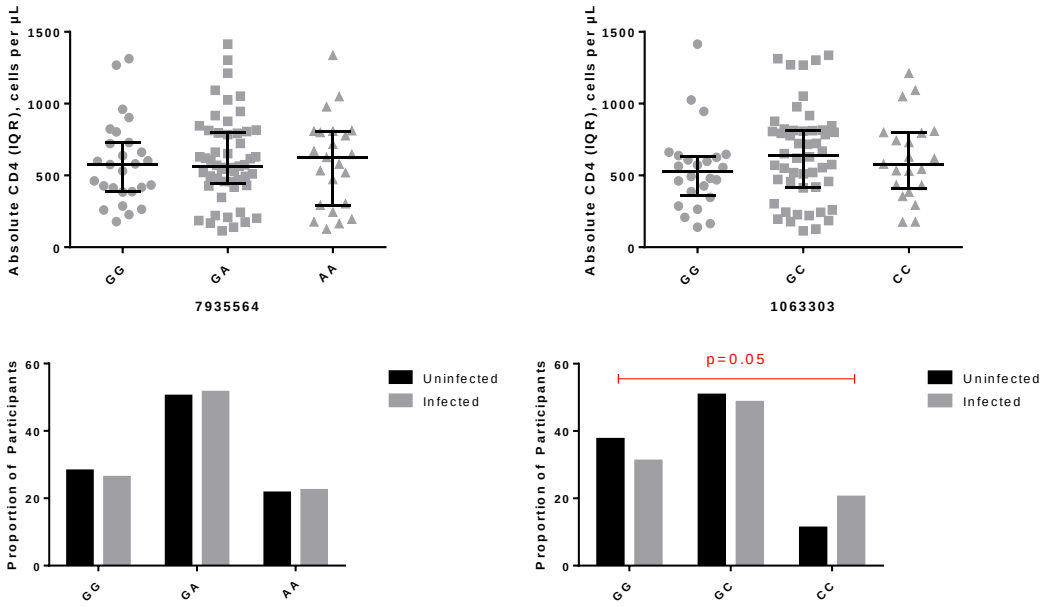


Figure 4.6 TRIM22 SNPs Show Association with Infection Status in Discovery Cohort

T-tests of Absolute CD4 Count in HIV-2 monoinfected individuals with rs7935564 genotypes (A), and rs1063303 genotypes (B.) 2-way ANOVA in participants who are either Uninfected or HIV-2 Infected according to rs7935564 genotype (C.), and rs1063303 genotype (D.)

4.4.4.2. Examination of the Effect of Polymorphisms rs1063303 and 7935564 on Infection Status and Clinical Progression throughout Data Collection Points

As there was no statistical significance observed between either polymorphism and status or disease progression in the validation cohort, in the Validation cohort, associations with each SNP were further validated against CD4% for each year that a substantial number of representative samples were available. This was an extra analysis to confirm absence of the associations. Neither SNP was significantly associated with %CD4 in 1991, 1996, 2003, or 2006.

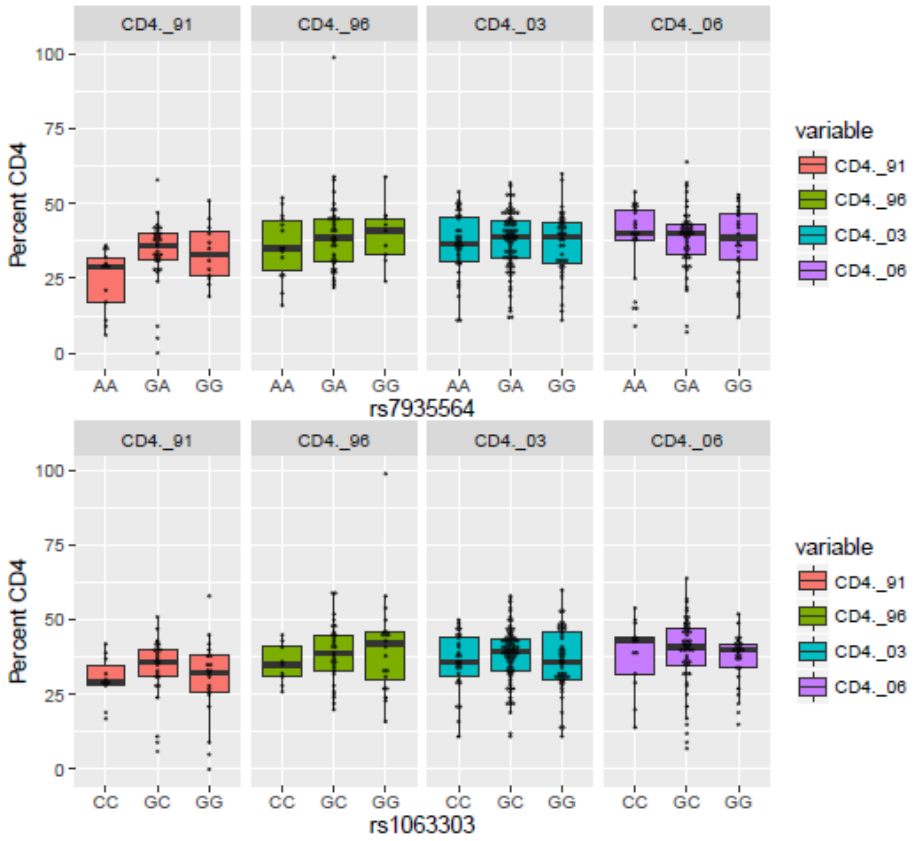


Figure 4.7 TRIM22 Polymorphisms Show no Association with Percentage CD4+T-Cell Count throughout Data Collection Points in the Validation Cohort

Distribution of Percentage CD4 Count in HIV-2 monoinfected individuals with rs7935564 and rs1063303 genotypes across study time points. Data shows no statistically significant variation in percentage CD4 data for the various alleles across the time points.

In the Validation cohort, associations with each SNP were further validated against Viral Load for each year that substantial representative data was available. Neither SNP was significantly associated with Viral Load in 1996, 2003, or 2006 following adjustment for multiple comparisons.

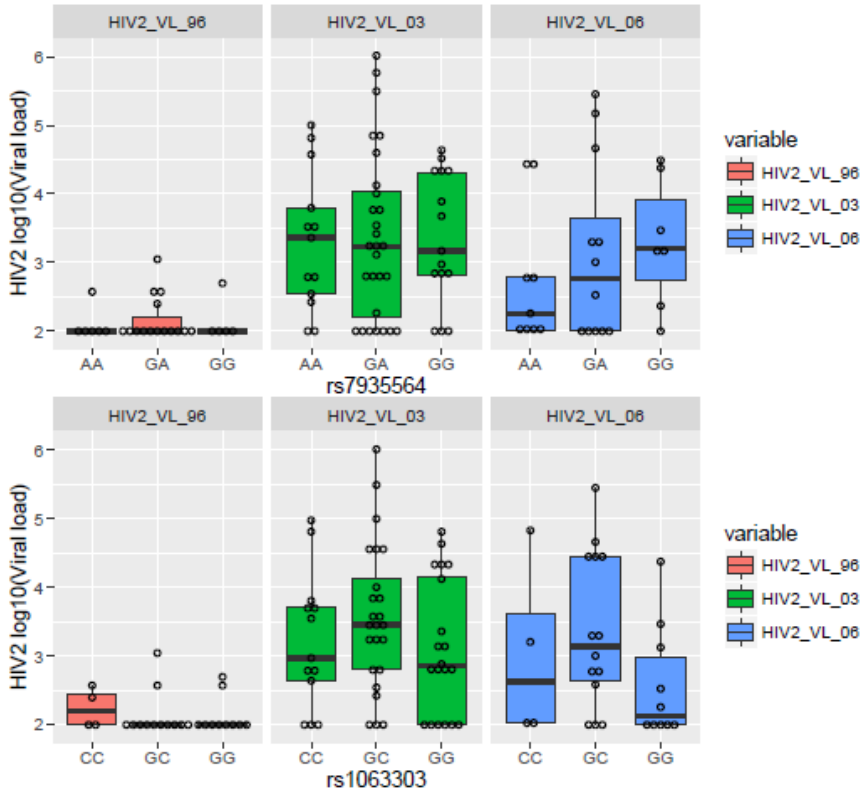


Figure 4.8 TRIM22 Polymorphisms Show no Association with Viral Load throughout Data Collection Points in the Validation Cohort

Distribution of Normalised Viral Load in HIV-2 monoinfected individuals with rs7935564 and rs1063303 genotypes. Data shows no statistically significant variation in normalised viral load for the various alleles across the time points.

4.4.4.3. TRIM22 Polymorphisms Show Differential Survival Among HIV-2 Infected Participants

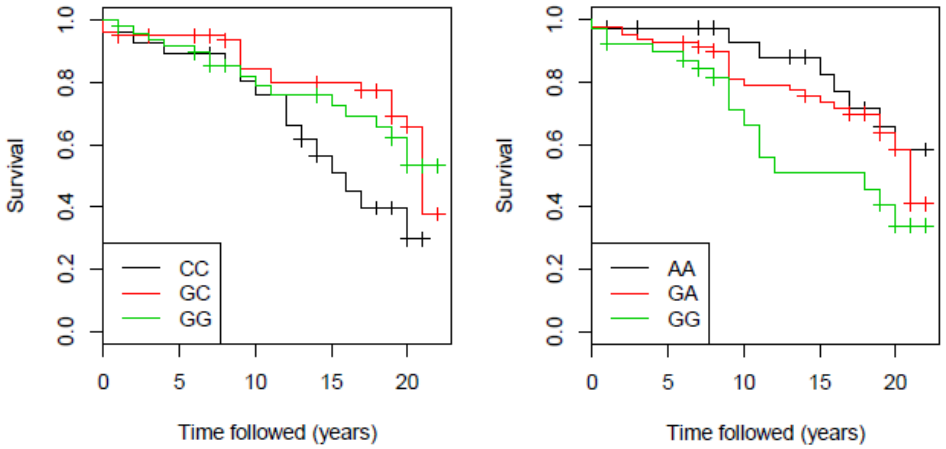
Kaplan-Meier survival, proportional hazard, and logrank tests were performed on HIV-2 mono-infected individuals throughout the 23 year observation period to examine the effect of each polymorphism on survival. Those who were homozygous for SNP 1063303 (CC) progressed to death at over twice the rate of those who were homozygous for the wildtype(GG). Those with wildtype (GG) did not achieve fractional survival of less than 50%, where as those with 1063303-CC had a median survival time of 16 years. Conversely, patients who were homozygous for SNP

7935564 (AA) progressed to death at just under half the rate of those homozygous for the wildtype(GG). Though 50% of the patients died within the 23 year follow up period in both groups, those possessing 7935564 SNP AA the median survival time was 23 years, when compared to 19 years for those with the 7935564 wildtype.

Table 4.8 Median Survival and Cox Proportional Hazard Data Examining the Effect of rs1063303 and rs7935564 on Survival Following HIV-2 infection

Figure	Group [a]	Median Survival [b]	ChiSq[c]	DF[d]	p-value[e]	P-value adjusted[f]
A	1063303 GG	>50% survival	8.994	2	0.0488	0.01111
	1063303 CC	16 years				
B	7935564 GG	19 years			0.045	0.003996
	7935564 AA	23 years	11.45	2		

Survival data for all HIV-2 infected study participants (n=126) versus TRIM22 genotypes
[a] TRIM22 SNPs rs1063303 and rs7935564 are grouped by allele
[b] Median survival time- defined as the time taken in years at which only 50% of the participants in each genotype group remain alive. Median survival >50% indicates that 50% mortality was not reached within the genotype group during the 23 year study period.
[c] ChiSq shows logrank test for trend, which reports chi-square value associated with
[d] 2 degrees of freedom
[e] Unadjusted p-value
[f] p-value adjusted for the effect of gender



A. B.

Figure 4.9 TRIM22 Polymorphisms Predict Survival

Kaplan-Meier survival curves illustrating effect of wildtype, heterozygosity, and homozygosity for rs1063303 (A) and rs7935564 (B) on proportional survival. Adjusted for the effect of gender. Data indicates that two copies of the minor allele at rs1063303 is significantly associated with diminished survival (p=0.011) and that the presence of the minor allele at rs7935564 is significantly associated with improved survival (p=0.004)

4.5. Discussion

This chapter describes the genetic diversity of TRIM22 in a unique West African cohort with unprecedented levels of HIV-2 infection, revealing several novel polymorphisms. Out of the 52 TRIM22 polymorphisms observed, only 2 were significantly associated with HIV-2 disease acquisition and/or disease progression:

Rs1063303 was associated with HIV-2 infection risk. Regression models revealed that participants possessing the wild-type of rs1063303 were found to be 2.5x more likely to be uninfected than those who possessed a mutation at this position. When compared with the rs1063303 homozygote, participants with the rs1063303 wildtype had a significant increase on average, in CD4+ T-cell count, all factors held constant. The effect is stepwise: with heterozygous participants also showing a significant, to a lesser extent than the wildtype, increase in CD4+ T-cell with all other factors held constant.

The two mutations were not found to be in linkage disequilibrium, however they are both frequently occurring. Wild type at both 7935564 and 1063303 was within the most frequently observed haplotype, followed by A/G, and A/C in 7935564 and 1063303 respectively out of 20 possible recombinations occurring at greater than 1% within the population. The two mutations were characterised in a further 176 HIV-2 infected, and uninfected participants from the same community. Among this group, neither polymorphism was associated with CD4+ T cell count. Only homozygotes for rs1063303 remained enriched in the HIV-2 infected patients.

The data indicate that TRIM22 polymorphisms can predict median survival time within a group of HIV-2 infected individuals. Over 50% of patients that maintained an undetectable viral load survived at the end of 23 year follow with HIV-2 infection, confirming that undetectable viral load indeed predicts normal survival as previously reported. Like patients with undetectable viral load, 50% of patients with absent detrimental polymorphism 1063303 also survived

following 23 years with HIV-2 infection. This supports data observed in the association study, insofar as HIV-2 infected participants with the 1063303 wildtype were more likely to maintain an undetectable viral load, and thus show a trend towards normal survival, when compared to those possessing the 1063303 SNP.

Taken together, the association study data suggest a protective effect for the rs7935564 mutation in HIV-2 disease progression and a detrimental effect for the rs1063303 mutation in the case of HIV-2 susceptibility and disease progression.

The differential effect of TRIM22 polymorphisms has been reported in 4 studies to date: in a study of genetic factors associated with granulomatous colitis and severe perianal disease, rs187416296, located in the V3 loop of the PRYSPRY domain, was found in a patient with severe disease. Functional studies demonstrated that the variant disrupted the ability of TRIM22 to regulate NOD2dependent activation of IFN-β and Nf-kβ(266). The study also reported three other polymorphisms observed in patients with severe disease; rs61735273, rs200924168 and rs12364019, for which *in silico* studies predicted a damaging effect and potential disruption of TRIM22 higher order assembly due to disruption of the coiled coil domain.(267) In a study of SNPs associated with multiple rubella-specific immune response outcomes post-measles, mumps, and rubella (MMR) immunization, TRIM22 rs2291842 was significantly associated with rubella-specific immune response outcomes. Increasing numbers of the minor allele were associated with an increase in NA titers (1.1-fold) and IFN-γ secretion (2.0-fold); however, the major allele seemed to correlate with enhanced secretion of IL-6 (1.1-fold)(268). Of the 5 TRIM22 SNPs investigated in other contexts, rs61735273, rs200924168, and rs2291842 were also observed in this study but showed no correlated with HIV-2 acquisition or disease progression.

The only other genetic study conducted to date on TRIM22 polymorphisms and their impact on HIV disease progression was by Ghezzi and colleagues in 2013. The study cohort consisted predominantly of European patients with varying progression levels for HIV-1 and a number of uninfected controls. The infected patients were stratified from progressors to long-term-non-progressors (LTNPs). All patients were genotyped for rs1063303 and rs7935564, following discovery of these polymorphisms in HIV-1 permissive and non-permissive cells. The study showed that that HIV-1 replication was more efficient in PBMCs from donors with wildtypes in both 1063303 and 7935563.(228) In addition, the study showed that wildtype and heterozygote forms of rs7935564 were more frequent in advanced progressors than in LTNP or normal progressors. In this study, rs1063303 showed no association with HIV-1 disease progression. However, in 2014, using *in silico* methods, Barr and colleages showed that rs1063303 was under positive selection, and that the frequency of the rs1063303 polymorphism varied 10-fold between ethnicities. 31% of European individuals were homozygous for the nsSNP (C/C) genotype when compared to 3% of Asian individuals, 15% of those with African ancestry, and 19% of Americans. They also functionally demonstrated that the rs1063303CC increased TRIM22 expression and decreased the antiviral activity of TRIM22. (269) The data from this chapter corresponds with two accounts observed in the literature which suggest polymorphisms rs7935564 and rs1063303 in TRIM22 play a role in differential restriction of HIV-1. Like in the 2013 Ghezzi paper, the absence of homozygosity for the rs7935564-AA is associated with disease progression. Though this group discovered no detrimental role for rs1063303, independent studies show that rs1063303 diminishes the restrictive capacity of TRIM22 *in vitro*.(224)

This chapter contains the first study which characterises full length TRIM22, and the impact of TRIM22 polymorphisms in a retrovirus other than HIV-1. This is also the first study to examine the effect of TRIM22 polymorphisms on survival. Though no effect of TRIM22 on other

lentiviruses has been functionally investigated, the data presented herein provides a basis for further investigation into the restrictive capacity of TRIM22 and its variants in HIV-2 infection, and that of other lentiviruses. The previous chapter illustrated that polymorphisms in TRIM5α were not significantly associated with HIV-2 acquisition, disease progression, or survival. Given the discordant evolution of TRIM5α and TRIM22, an investigation into the structural changes conferred by TRIM5α polymorphisms when compared to those of TRIM22 would be worthy of investigation. These will be explored in the next chapter.

Chapter 5. Chapter V: Interpreting Mutations in TRIM5α and TRIM22: an Examination of Phylogenetic, Sequence, Structural, and Heuristic Features

5.1. Introduction

Host restriction factors are encoded by functionally autonomous genes and tend to be unusually variable. Their counterparts in non-human primates are also unusually variable, suggesting that they have evolved under diversifying selection. Single nucleotide polymorphisms (SNPs) occur within the human genome at a frequency of approximately every 100-300 base pairs, however the vast majority of these SNPs are neutral allelic variants. (270) Amino acid substitutions, resulting from non-synonymous polymorphisms, are represented in the final protein structure, can affect protein stability, structure, and function, and therefore provide crucial evidence for the molecular basis of host restriction factors and their interaction with retroviruses. It is therefore necessary to extend population genetics work, such as that observed in Chapter III and Chapter IV of this thesis, to analyses of the effects of genetic variation on protein evolution, structure and function.

Protein evolution is computed by the rate of amino-acid altering DNA changes (**N**on-synonymous) relative to silent DNA changes (**S**ynonymous mutations,) or d**N**/d**S** ratio, which provides information about the degree of selection operating on a system. TRIM5α genes from multiple species have shown evidence of a high dN/dS ratio.(237) Such evidence suggests that TRIM5α is continually under pressure to maintain fitness, perhaps mediated by their restrictive capacity for retroviruses. A consequence of this pressure is that advantageous mutations can rise in frequency, disadvantageous mutations can also arise given the interplay between retrovirus and host, and that these mutations will be accompanied by proximal mutations in paralogs. The latter may be driven by the process of neofunctionalisation: a defence mechanism against the accumulation of mutations that affect fitness where a duplicate of a gene may gather a beneficial function not present in the ancestral gene and may also be positively selected for, while the other copy retains its original function.

The TRIM5/6/22/34 gene cluster is likely to have arisen through tandem gene duplication as the four genes are paralogs.(237) The relationship between TRIM5 and TRIM22 is an interesting one because the cow genome encodes multiple TRIM5 genes, while the dog genome encodes no TRIM5 at all. Such evidence suggests that multiple retroviral pressures in the cow may have resulted in an increase in the number of TRIM5 genes, or that another gene has acquired the restrictive function of TRIM5 in dogs leading to loss of the gene. Both hypotheses are more difficult to unravel in a situation where the presence of either protein is mutually exclusive, and where there will be species divergent retroviral selection pressures. A genetically homogenous group which possesses both TRIM5α and TRIM22, and a common set of retroviral selection pressures, makes a more conducive environment to investigate selective pressures on both proteins as evidenced by the mutations which exist in both, and an assessment of fitness and restrictive capacity therein. The Caio Community Cohort has genetically homogenous participants which, in theory, are maintained under a common and longstanding set of selective pressures due to their relative geographical isolation and animist culture where marriages occur only within the community. Chapter III and IV characterised polymorphisms in TRIM5α and TRIM22 respectively within this cohort observing different frequencies and distributions of nsSNPs, and also the unadjusted significance of TRIM5α, and high statistical relevance of TRIM22 in HIV-2 disease progression within the same population. Is it possible that TRIM22 might have evolved as compensatory antiviral mechanism to TRIM5α, mediating an antiviral effect where TRIM5α cannot?

In addition to evaluating the selective pressures on both TRIM5α and TRIM22 as illustrated by their differential genetic variability, it is important to determine if any of the nsSNPs have an impact on protein structure and function. In the case of structure, TRIM proteins are notoriously difficult to crystallise due to their highly flexible structure and penchant for oligomerisation. To date, only crystal structures of TRIM protein fragments exist, as summarised below. The lack of

solved structures for TRIM proteins makes structural examinations TRIM proteins, and analyses of the effect of amino acid substitutions on protein structure, difficult.

Table 5.1 List of Successfully Crystallised TRIM Fragments

Domain	***Species***	***TRIM Protein***	***Protein Data Bank ID***
PRYSPRY	*Rhesus*	*TRIM5α*	*4B3N*
	Mouse	*TRIM21*	*2VOK*
	Human	*TRIM20*	*2WL1*
	Human	*TRIM25*	*4B8E*
	Human	*TRIM72*	*3KB5*
	Human	*TRIM21*	*2IWG*
Coiled Coil	*Human*	*TRIM25*	*4LTB*
	Human	*TRIM69*	*4NQJ*
B-Box & Coiled Coil	*Rhesus*	*TRIM5α*	*4TN3/ 5IEA/ 5F7T*
B-Box	*Rhesus*	*TRIM5*	*5K3Q/ 5EIU*
RING	*Human*	*TRIM25*	*5EYA*

Further, although experimental evidence provides the strongest evidence for the functional role of a genetic variant, it is often not feasible to perform laboratory experiments for all nsSNPs in a single gene, especially those highly polymorphic in nature like host restriction factors. Further, investigating the synergistic or antagonistic effects of different combinations of the nsSNPs is also quite experimentally challenging. A way of addressing these shortcomings involves the use of *in silico* analyses to predict the effect of a variant on protein structure and function.

Computational methods may be useful for identifying SNPs with functional significance from the pool of alleles. Computational methods provide rapid predictions of functionally significant SNPs. On their own, they are insufficient to predict the significance of nsSNPs, however when combined with sequence and structural information, computational methods can make reliable predictions of the functional significance of nsSNPs. Predictions of the functional significance of nsSNPs may utilise:

a. Phylogenetic information
b. Structural analysis and sequence properties(271)
c. Heuristic features

Phylogenetic approaches rely on two assumptions. The first is that variants that affect a protein's biochemical function or cause a detrimental phenotype also cause a loss of evolutionary fitness, making them deleterious in the evolutionary sense, and thus subject to purifying selection. The second is that a deleterious variant in the human population may also be deleterious in homologous genes in other living species. These allow the functional significance of a SNP to be predicted from the pattern of amino acids observed in that corresponding position of a multiple alignment of related sequences. The major phylogenetic tools are Sorting Intolerant from Tolerant (SIFT) and Polymorphism Phenotyping version 2 (PolyPhen2), which compute the deleterious nature of a SNP using sequence weights or structural comparisons respectively.(272, 273) A major limitation of studies relying on phylogeny like SIFT is that they are unable to distinguish between strongly deleterious and moderately deleterious SNPs. This is because a variant associated with a small loss of fitness has little chance of being represented in a population at a frequency greater than 5%. As a result, the amino acid position is likely to be conserved despite being predicted damaging by phylogenetic methods.

The second approach to predicting the effect of nsSNPs relies on structural information. These methods rely on the assumption that disadvantageous mutations are structurally destabilising. The change in stability conferred by a SNP can be measured by the protein's folding free energy, or ΔΔG. The measurement of ΔΔG is quite intensive, and tools such as I-Mutant3.0 resolve this by using a combination of sequence and structural methods predicting a measure of ΔΔG that correlates with empirically measured ΔΔG at a correlation coefficient of 0.8.

Both methods; phylogenetic and sequence/structure are imperfect. However these inaccuracies can be somewhat mitigated using predictions of protein stability which examine heuristic features such as the character of the mutation site, secondary structure, comparisons of

mutants within a domain in light of putative domain functions, models of tertiary structure, and examining the properties of wildtype and mutant amino acids for areas of solvent accessibility, hydrophobicity, charge and volume.

The population genetics experiments in Chapter III and IV uncovered seven nsSNPs in TRIM5α, and twelve in TRIM22 among participants from the Caio Community Cohort, 3 of which, TRIM5α's R136Q, and TRIM22's D155N and R242T suggested a role in HIV-2 disease progression. This chapter aims to use a combination of phylogenetic, sequence and structural algorithms to infer the effect of nsSNPs observed in TRIM5α and TRIM22 from the Caio Community cohort. The computational estimates will be bolstered using heuristic evaluations of primary and secondary structure, alterations in protein folding and confirmation conferred by nsSNPs, and structural and biochemical properties of wildtype and mutant amino acids.

5.2. Aims

This chapter aims to use computational methods to predict the alterations to structure and function conferred by the non-synonymous polymorphisms in TRIM5α and TRIM22 which were observed in the Caio Community Cohort. These predictions will be largely based on phylogeny, sequence, structure, and heuristic features. It also aims to investigate patterns in the type and distribution of non-synonymous mutations between TRIM5α and TRIM22 within this genetically homogenous population for clues regarding positive selection. The chapter seeks to address the following:

1. To determine whether TRIM5α and TRIM22 nsSNPs observed in the Caio Community Cohort are deleterious as determined by computational approximations based on phylogeny
2. Determine whether TRIM5α and TRIM22 nsSNPs observed in the Caio community cohort are deleterious as determined by computational approximations based on sequence and structure
3. In the absence of solved crystal structures for human TRIM5α or TRIM22, computationally generate 3D models of the wildtype and mutant versions of each protein
4. Calculate the free energy score and relative solvent accessible area of TRIM5α and TRIM22 wildtype and mutants, based on both sequence and structure
5. Use secondary structure alignments to describe the position of TRIM5α and TRIM22 nsSNPs observed in the Caio cohort relative to each other in order to determine areas of selective pressure
6. Use the 3D models of TRIM5α and TRIM22 mutants to describe the structural and biochemical properties of the amino acid changes, describe mutation sites relative to tertiary structure and determine their ability to change protein confirmation, and observe the distribution of polymorphisms at a tertiary level

5.3. Materials and Methods

To predict the functional effect of TRIM5α and TRIM22 nsSNPs observed in the Caio Community Cohort, three methods were employed:

- Phylogenetic analysis using the SIFT algorithm (http://sift.jcvi.org/) and the Consurf server (http://consurf.tau.ac.il/2016/)
- Sequence and Structural analysis using

  Polyphen2 (http://genetics.bwh.harvard.edu/pp2)

  Protein Variation Effect Analyser (Provean) (http://provean.jcvi.org/index.php)

  I-Mutant3.0 http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi
- Heuristic examination involving the generation of secondary structure alignments of TRIM5α and TRIM22 and annotation of variant positions, generation of 3D protein models of TRIM5α and TRIM22 mutants, and examination of the properties of the amino acid substitutions

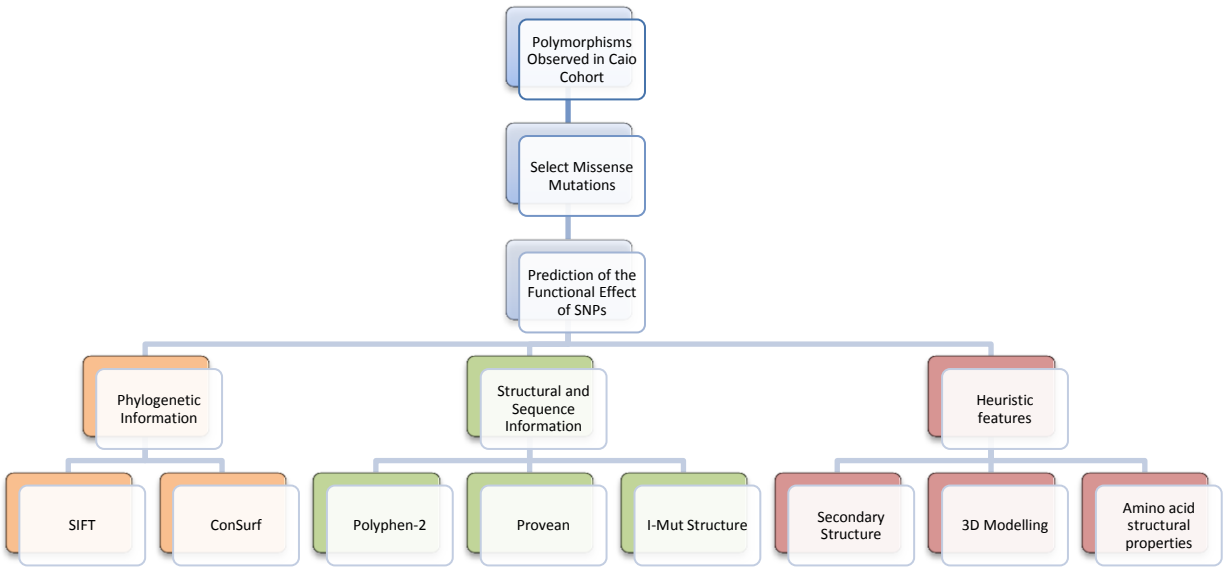


Figure 5.1 Computational SNP Analysis- Methods Workflow

Of the TRIM5α and TRIM22 polymorphisms observed in the Caio Cmmunity Cohort, the non-synonymous mutations will be selected for prediction of their effect. Predictions will be carried out using phylogenetic methods (SIFT, Consurf) Sequence and Structure Data (Polyphen-2, Provean, I-Mutant) and an examination of Heuristic features using secondary structure alignments, generation of 3D structures and examination of amino acid structural properties.

5.3.1. Predicting the Effect of Polymorphisms on Protein Structure and Function using SIFT, PolyPhen2, and Provean

The chromosome number, position, and base pair mutations subjected to the SIFT and Polyphen2 algorithms. The SIFT algorithm predicts whether an amino-acid substitution affects protein function based on the degree of conservation of amino acid residues throughout evolution. Based on SIFT scores, substitutions were classified as damaging (0.00-0.05), potentially damaging (0.051-0.10), borderline (0.101-0.20) or tolerant (0.201-1.00).(273)

SIFT Score	Effect Class
0.00-0.05	Damaging
0.051-0.10	Potentially damaging
0.101-0.20	Borderline
0.201-1.00	Tolerant

The Polyphen-2 tool predicts the possible impact of an amino acid substitution on the structure and function of a human protein using physical and comparative considerations of the available structure. The PolyPhen-2 scores were used to classify the substitutions as probably damaging (1.00-0.80), possibly damaging (0.80-0.40) or benign (0.40-0.0).(272)

Polyphen2 Score	Effect Class
0.80-1.00	Probably Damaging
0.20-0.80	Possibly Damaging
0.00-0.40	Benign

Provean extends both methods by predicting the damaging effect of not only amino acid substitutions but also in-frame insertions or deletions. Provean generates an alignment-based score which measures the change in sequence similarity of a query sequence to a protein sequence homolog before and after the introduction of an amino acid variation to the query sequence. When the Proven score is equal to or below a predefined threshold -2.5, the amino

acid substitution is predicted to have a "deleterious" effect. If the provean score is above the threshold, the variant is predicted to have a "neutral" effect.(274)

Provean Score	Effect Class
<-2.5	Deleterious
> -2.5	Neutral

5.3.2. Calculation of Gibbs Free Energy Score and Solvent Accessible Area for TRIM5α and TRIM22 Mutants using I-Mutant3

I-Mutant3.0 was used to evaluate nsSNP-induced changes in protein stability. Sequences of TRIM22 and TRIM5α, along with amino acid substitutions for each variant, were submitted to I-Mutant3. I-Mutant3 estimates the free energy change value (DDG) by calculating the unfolding Gibbs free energy value (ΔG) for the wild type protein and subtracting it from that of the mutant protein (DDG or ΔΔG = ΔG mutant – ΔG wild type). I-Mutant3 predicts the sign (increase or decrease) of the free energy change value (DDG), along with a reliability index for the results (RI: 0–10, where 0 is the lowest reliability and 10 is the highest reliability). It also determines the Relative Solvent Accessible Area (RSA) which is an intrinsic measure of protein flexibility.(275) The pH was set to 7 and the temperature was set to 25°C for all submissions.

ΔΔG	Effect Class
< -0.5	Large Decrease in Protein Stability
-0.5<= x <=0.5	Neutral Stability
> 0.5	Large Increase in Protein Stability

5.3.3. Secondary Structure Alignments of TRIM5α and TRIM22

Uniprot was used to retrieve sequences for Rhesus macaque TRIM5α (Identifier: Q0PF16), *Homo Sapiens* TRIM5α (Identifier: Q9C035-1), *Homo Sapiens* TRIM22 (Identifier: Q8IYM9). Sequence alignments were derived using Clustal Omega- Multiple Sequence Alignment.(276) Sequences were mapped onto secondary structures using Espript 3.0 Advanced, Global Score 0.7, and

Difference Score 0.5. (http://espript.ibcp.fr)(277) In order to derive the secondary structure of TRIM domains; sequence alignments from Clustal were mapped onto 3D protein models generated in Phyre2 (succeeding section) or from Protein Data Bank (PDB). To derive secondary structure of the PRYSPRY and coiled coil domain of the rhesus macaque, the crystallised domains were used as a templates PRYSPRY RCSB Identifier: 4B3N (278), Coiled coil RCSB Identifier 4TN3 (279)} To derive secondary structure of the Human B-box and coiled coil domain, and PRYSPRY domain, for human TRIM5α and TRIM22, 3D protein models were derived generated as described below, and used as scaffolds from which predict secondary structure. Secondary structure alignments are annotated as follows:

α-helices: medium squiggle

3_{10}-helices: small squiggle

π-helices: large squiggle

β-strands: →

Strict β-turns: **TT**

Strict α-turns: **TTT**

3_{10}-helix alone: η

Regions of identity were highlighted in red, and gaps shown as dots.

5.3.4. Generation of 3D Protein Models

In order to build models of Human TRIM5α and Human TRIM22: amino acid sequences were submitted to Phyre 2.0(280) for intensive modelling.

(http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) A minimum of 120 structures crystal structures from proteins with at least 98% confidence in the homology, were used generate 3D models of full length Rhesus TRIM5α, along with Human TRIM5α, Human TRIM22 and their mutants.(280) Models were visualised using an educational version of PyMol.(281) To

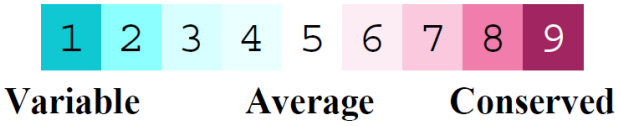
describe the position of mutations on individual domains, non-synonymous polymorphisms derived from population studies were mutated onto 3D structures using PyMol Mutagenesis. Distances between amino acid residues were also calculated in PyMol and represented in Angstroms.

5.3.5. Amino Acid Properties

The type of charge, molecular weight, and hydrophobicity are reported for each wildtype residue and nsSNP. The information for each residue is derived from Lide's Handbook of Chemistry and Physics.(282) The hydrophobicity values as reported are normalised so that the most hydrophobic residue is given a value of 100 relative to glycine, which is considered neutral (0 value). The scales were extrapolated to residues which are more hydrophilic than glycine.

5.3.6. Evolutionary Conservation

ConSurf (http://consurftest.tau.ac.il) estimates the evolutionary conservation of amino acid positions in a protein based on the phylogenetic relationships between homologous sequences [25]. Amino acid sequences of proteins were entered to the server using default settings (Homolog search algorithm: CS- BLAST, Maximal %ID Between Sequences 95 and Minimal % ID For Homologs: 35, for a maximum of 150 homologues). The conservation scores were calculated based on the evolutionary relationships among the protein and its homologs where a conservation score between 1 and 4 is considered variable, whereas a score of 5–6 is intermediate, and a score in the range of 7 to 9 indicates conserved.



5.4. Results

5.4.1. *In silico* Analyses TRIM5α

As shown in Chapter III, there were seven nsSNPs in TRIM5α observed in the Caio cohort. These seven amino acid substitutions (G31S, H43Y, R136Q, L214F, R238W, D249G, P479L) were subjected to analysis by four programmes which aim to predict the effect of an amino acid substitution on protein structure and function. Though not observed in the Caio Cohort, a known TRIM5α polymorphism, E204K, was added to these analyses due to its overlap with TRIM22's E209K, for comparison. The following table presents the results obtained through the SIFT, Polyphen, Provean, and I-Mutant Suite. Because of the difficulty of phylogenetic analyses such as SIFT to distinguish possibly deleterious from deleterious, all potentially deleterious SNPs were classified as benign for the purposes of analysis. As previously mentioned: a variant associated with a small loss of fitness has little chance of being represented in a population at a frequency greater than 5%. As a result, the amino acid position is likely to be conserved despite being predicted damaging by phylogenetic methods.(271)

Table 5.2 Phylogenetic, Sequence, and Structural Computational Analyses of TRIM5α Polymorphisms

Exon	domain	SNP ID	Allele	Allele Frequency	Residue Change	Polyphen	Provean	SIFT	ΔΔG (Kcal/mol)	RSA
2	RING	rs59896509	C/T	0.9570/0.0430	G31S *	**1**	**-5.12**	**0.002**	-1.25	38.1
2	RING	rs3740996	G/A	0.8874/0.1126	H43Y	**0.036**	**-1.63**	**1.000**	-0.33	27.5
2	Coiled coil	rs10838525	C/T	0.8744/0.1256	R136Q	**0**	**1.84**	**1.000**	-0.20	47.6
4	Coiled coil	rs374402267	C/T	1.000/0.0000	E204K	**0.016**	**-1.28**	**0.280**	-0.42	63.5
4	Coiled coil	rs147678479	G/A	0.9853/0.0147	L214F *	**0.801**	**-2.80**	**0.039**	-0.56	39.9
4	Linker 2	rs140131201	G/A	0.9873/0.0127	R238W	**0.077**	**-6.30**	**0.003**	-0.70	22.7
5	Linker 2	rs11038628	C/T	0.5084/0.4916	D249G	**0.001**	**4.00**	**1**	-1.29	36.8
8	B30.2/SPRY	rs7104422	G/A	0.8486/0.1514	P479L *	**0.582**	**-8.30**	**0.012**	-1.37	0.0

**Agreement of deleterious effect in all four algorithms*

Phylogenetic methods in SIFT predated 50% of the nsSNP as damaging (G31S, L214F R238W, P479L). The SNPs were further analysed using Polyphen2 based on structural information and multiple sequence alignments, indicating 1 SNP, G31S, as damaging, and 2 SNPs L214F, and P479L as possibly damaging. Polyphen2 characterised 57.5% of SNPs damaging. Provean, which also relies on multiple sequence alignments, had identical predictions to the SIFT algorithm. To

improve the prediction accuracy of structure-based tools, the I-Mutant suite was then used. It was observed that the same 4 SNPs detected by SIFT as damaging in addition to D249G, had ΔΔG values of less than =-0.5kcal/mol, indicating a significant loss of stability in the resulting protein. The predictive power of determining the functional impact of a given nsSNP can be significantly increased by combining information from a variety of computational tools. (283) Taken together, all four indicators predicted G31S, L214F, and P479L as nsSNPs likely to have a functional impact. In addition, the relative solvent accessible area of P479L dropped to 0 suggesting a significant loss in flexibility in the resulting protein structure.

5.4.2. *In silico* Analyses TRIM22

As show in Chapter IV, TRIM22 has twelve non-synonymous SNPs represented in the Caio Cohort, two of which were not previously observed; E209K and D360Y. The proceeding table presents the results obtained through the SIFT, Polyphen, Provean, and I-Mutant Suite for TRIM22's nsSNPs.

Table 5.3 Phylogenetic, Sequence, and Structural Computational Analyses of TRIM22 Polymorphisms

Exon	Domain	SNP ID/Position	Allele	Allele Frequency	Residue Change	Polyphen	Provean	SIFT	ΔΔG (Kcal/mol)	RSA
2	RING/BBox	rs200668710	G/A	0.989/0.011	E83K	0.134	-2.66	0.042	-0.78	83.6
3	Coiled Coil	rs200924168	G/C	0.907/0.093	R150T	0.012	-3.1	0.135	-0.68	59.0
		rs7935564	G/A	0.571/0.429	D155N	0.039	-0.31	0.009	-0.34	57.6
		Chr11:*5719650	G/A	0.978/0.022	E209K	0.997	-2.11	0.331	-0.42	81.9
		rs2291843	A/G	0.995/0.005	T232A	0.004	0.46	0.015	-0.88	51.9
		rs1063303	G/C	0.636/0.364	R242T	0	-1.46	0.036	-0.23	97.8
		rs61735273	C/T	0.962/0.038	S244L	0.803	-4.55	0.001	-0.20	41.0
6	Linker-2	COSM1509111	G/A	0.995/0.005	D282N *	1	-4.00	0.028	-1.20	50.0
7	PRYSPRY	rs73404240	C/A	0.965/0.035	T294K	0.007	-2.58	0.146	-0.67	64.9
		rs12364019	G/A	1.000/0.000	R321K *	0.831	-2.59	0.004	-1.06	25.3
		Chr11:*5730459	G/T	0.990/0.010	D360Y *	1	-8.13	0.000	-0.74	22.9
		rs150095329	C/T	0.990/0.010	T415I	0.356	-2.21	0.080	-0.16	67.1

**Agreement of deleterious effect in all four algorithms*

Phylogenetic methods in SIFT predicted a high proportion of the SNPs, 66.7% as damaging (E83K, D155N, T232A, R242T, S244L, D282N, R321K, D360Y). The SNPs were further analysed using Polyphen2 based on structural information and multiple sequence alignments, which

indicated 3 nsSNPs, E209K, D282N, and D360Y as damaging, and 2 nsSNPs, S244L, and R321K as possibly damaging. Polyphen2 characterised 41.6% of SNPs damaging. Provean, which also relies on multiple sequence alignments, did not have identical predictions to either of the algorithms used previously. It predicted R150T and T294K as damaging where neither SIFT or Polyphen2 had. To improve the prediction accuracy of structure-based tools, the I-Mutant suite (58.3% prediction as damaging) showed similar estimates to that of Provean (58.3% prediction as deleterious) with the omission of S244L which had a free energy score which suggested it as tolerated. Given that the predictive power of determining the functional impact of a given substitution can be significantly increased by combining information from a variety of computational tools, taken together, the data suggests the substitutions D282N, R321K, and D360Y likely to have a functional impact on TRIM22. Further, two of these substitutions, R321K and D360Y had the lowest relative solvent accessibility suggesting a significant loss in flexibility in the resulting protein structure.

5.4.3. Primary Structure Spatial Distribution of nsSNPs in TRIM5α and TRIM22

Aligment of TRIM22 and TRIM5α was performed using Clustal omega and annotated in ESPRIPT. The domain regions were drawn on the alignment, and the nsSNPs observed in the Caio Chort were highlighted on each respective sequence.

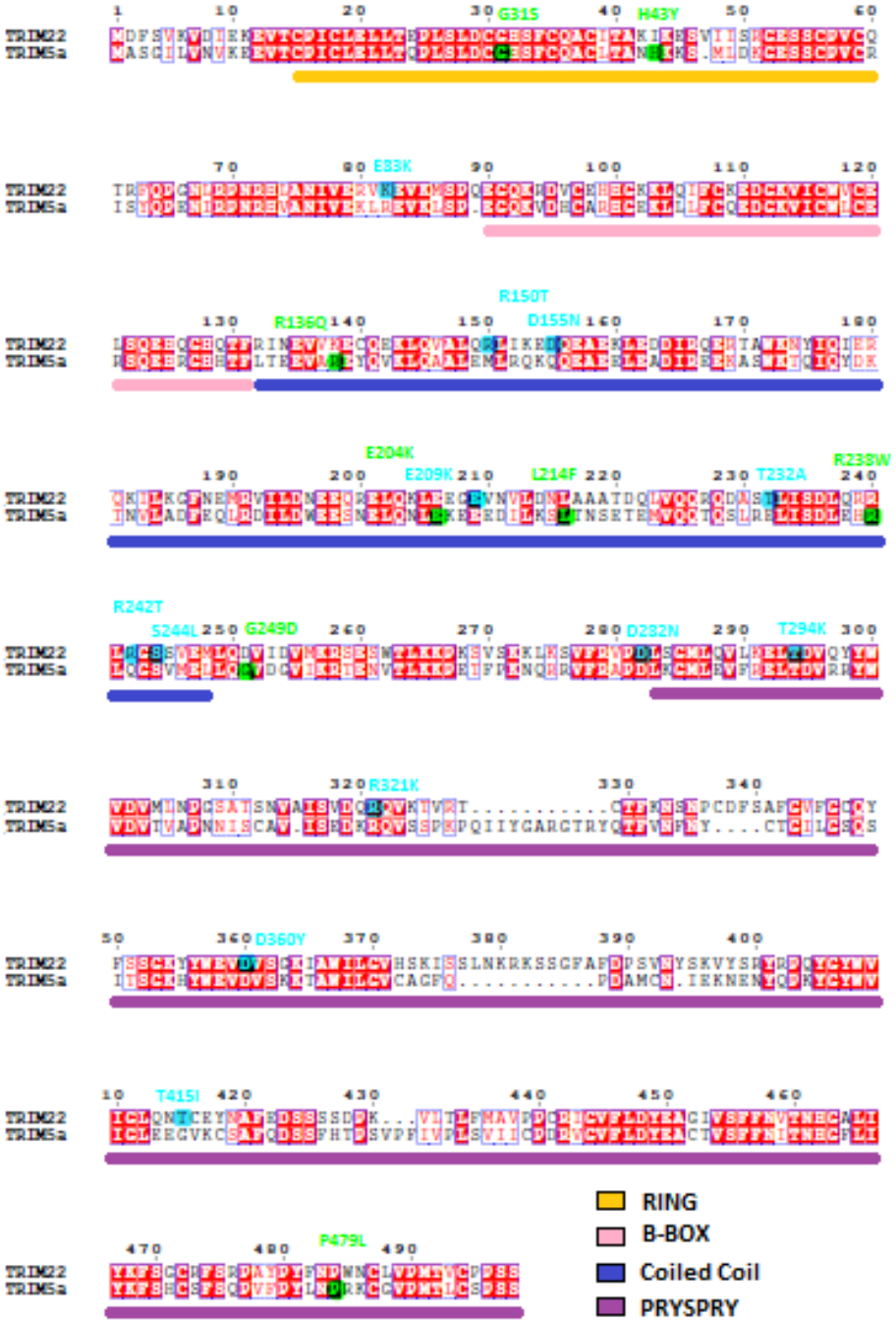


Figure 5.2 Distribution of Polymorphisms in TRIM22 (blue) and TRIM5α (green) Observed in the Caio Cohort.

Distribution of polymorphisms in TRIM22 (blue) and TRIM5α (green) observed in the Caio Cohort. Sequences: 55.95% amino acid identity. Functional domains highlighted as shown.

12 non-synonymous mutations were observed in TRIM22. The majority of the non-synonymous polymorphisms were observed in the coiled coil domain (n=6) and PRYSPRY domains (n=4). Two were located in the link regions between the RING and B-Box domain, and the Coiled Coil and PRYSPRY domain. None were observed in the RING domain. Seven nsSNPs were obsereved in TRIM5α. Like TRIM22 the majority of ns polymorphisms were observed in the coiled coil domain (n=3) Unlike TRIM22, two SNPs were observed in the RING domain where TRIM22 had no mutations at all. There was only 1 nsSNP in the PRYPSRY domain of TRIM5α. Given that this region is virus interacting, this finding suggests low selective pressure on TRIM5α in comparison to TRIM22. Though resulting from gene duplication, structurally, TRIM22 and TRIM5α only possess 55.95% Identity with most of the divergence being in the PRYSPRY domain. Also, TRIM22 possesses only 1 B-Box domain, whereas TRIM5α has two. TRIM22 however possesses a longer coiled coil (116amino acids vs 111amino acids). When compared to TRIM5α within the same population, TRIM22 was more polymorphic. Further, the distribution of polymorphisms was dissimilar between the two. No substitutions were shared between TRIM22 and TRIM5α. E204K of TRIM5α which has been observed in other populations, shared proximity with E209K of TRIM22 which was observed in the Caio Community.

5.4.4. Secondary Structure Examination of TRIM5α Polymorphisms Observed in the Caio Cohort Secondary Structure- TRIM5α- Coiled Coil

Primary and secondary and structure is largely conserved between rhesus (RhTRIM5α) and human (HuTRIM5α). HuTRIM5α and RhTRIM5α have 89% amino acid homology. Predictions of TRIM5α secondary and tertiary structure are presumed more accurate due to the availability of solved crystal structures of rhesus coiled coil and PRYSPRY domains. The TRIM5α coiled coil was modelled in Phyre 2.0 and served as a scaffold for a secondary structure alignment of TRIM5α wildtype sequence and mutated sequence, in order to decipher the arrangement of polymorphisms at the secondary structure level. The coiled coil of TRIM5α is 112 amino acids in length, and is comprised of 1 medium(α1), and 1 large alpha helix(α2). The SNP R136Q is located on α1, whereas E204K (not observed in the Caio Cohort) and L214F are located in α2.

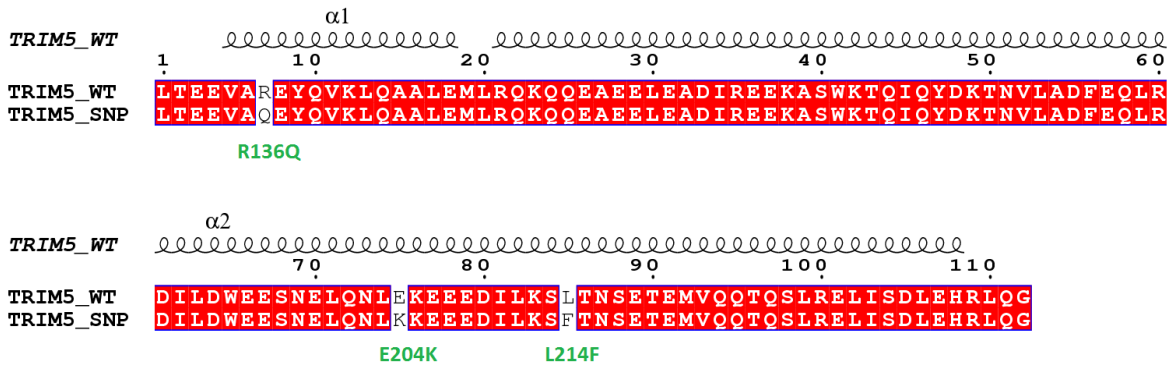


Figure 5.3 Secondary Structure Alignment of the Coiled Coil of TRIM5α:

The TRIM5α coiled coil is comprised of one medium, and one long alpha helix; α1 and α2 respectively. Polymorphisms observed in the Caio Cohort are annotated in Green.

5.4.5. Secondary Structure- TRIM5α- PRYSPRY

The secondary structure alignment of the TRIM5α PRYSPRY domain was derived in the same manner as that of the coiled coil. The structure of both human and rhesus PRYSPRY shows a highly conserved inner core comprised of 14 β sheets where the majority of conservation occurs. The remainder of the PRYSPRY domain is comprised of 4 Variable loops (V1-V4). In the case of Rhesus macaque TRIM5α, V1, V2, and V3 have been implicated in HIV-1 viral capsid interaction. The variable loops are the sites of least conservation between the PRYSPRY Domains in Rhesus TRIM5α and Human TRIM5α (Appendix). Data from the Caio Community Cohort revealed only one polymorphism within the human TRIM5α PRYSPRY domain. P479L, predicted to be deleterious by all four algorithms in a previous section, was located on the cusp of the V4 loop, which has yet to be implicated in virus interaction.

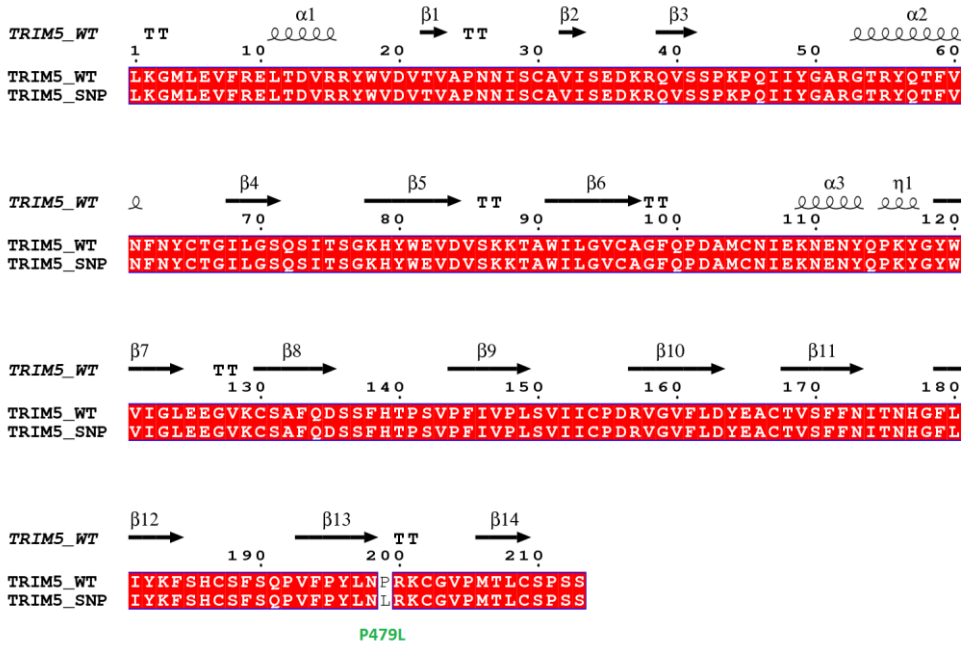


Figure 5.4 Secondary Structure Alignment of the PRYSPRY of TRIM5α:

The TRIM5α PRYSPRY is comprised 14 β sheets and 4 Variable loops. *Polymorphisms observed in the Caio Cohort are annotated in Green.*

5.4.6. Secondary Structural Examination of TRIM22 Polymorphisms Observed in the Caio Cohort Secondary Structure-TRIM22- Coiled Coil

The impact of TRIM22 polymorphisms on secondary structure may be less definitive due to the absence of a solved structure for TRIM22 or a close relative. The TRIM22 coiled coil was modelled in Phyre 2.0 and served as a scaffold for a secondary structure alignment of a TRIM22 wildtype sequence and TRIM22 mutant sequence in order to decipher the arrangement of polymorphisms at the secondary structure level. The coiled coil of TRIM22 is 116 amino acids in length, and like TRIM5α, is comprised of 1 medium (α1), and 1 large alpha helix (α2) and possesses a strict alpha turn where it is elongated in comparison to TRIM5α. The substitutions R150T and D155N are located in α1, whereas E209K and T232A are located in α2. R242T and S244L are located in the hinge between the second helix and the strict alpha turn which allow them to potentially alter the flexibility at this region.

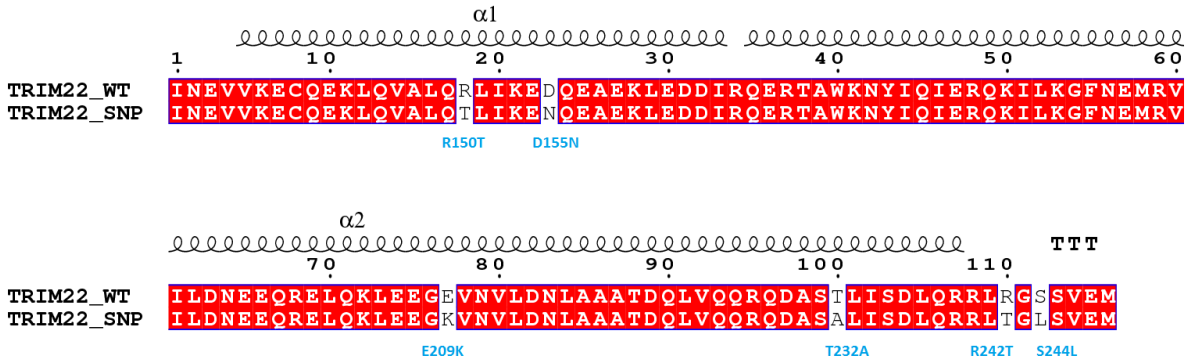


Figure 5.5 Secondary Structure Alignment of the Coiled Coil of TRIM22:

The TRIM22 coiled coil is comprised of one medium, and one long alpha helix; α1 and α2 respectively, and a strict alpha turn. Polymorphisms observed in the Caio Cohort are annotated in Blue.

5.4.7. Secondary Structure-TRIM22- PRYSPRY

The secondary structure alignment of the TRIM22 PRYSPRY domain was derived in the same manner as that of the TRIM22 coiled coil. The TRIM22 PRYSPRY domain is 215 amino acids in length, and is comprised of 14 β- strands, 4 β-turns, 2 3_{10}-helices and an alpha turn. R321K, D360Y, and T425I are located within the folds of β -strands, whereas T294K is located within the alpha helix at the start of the domain.

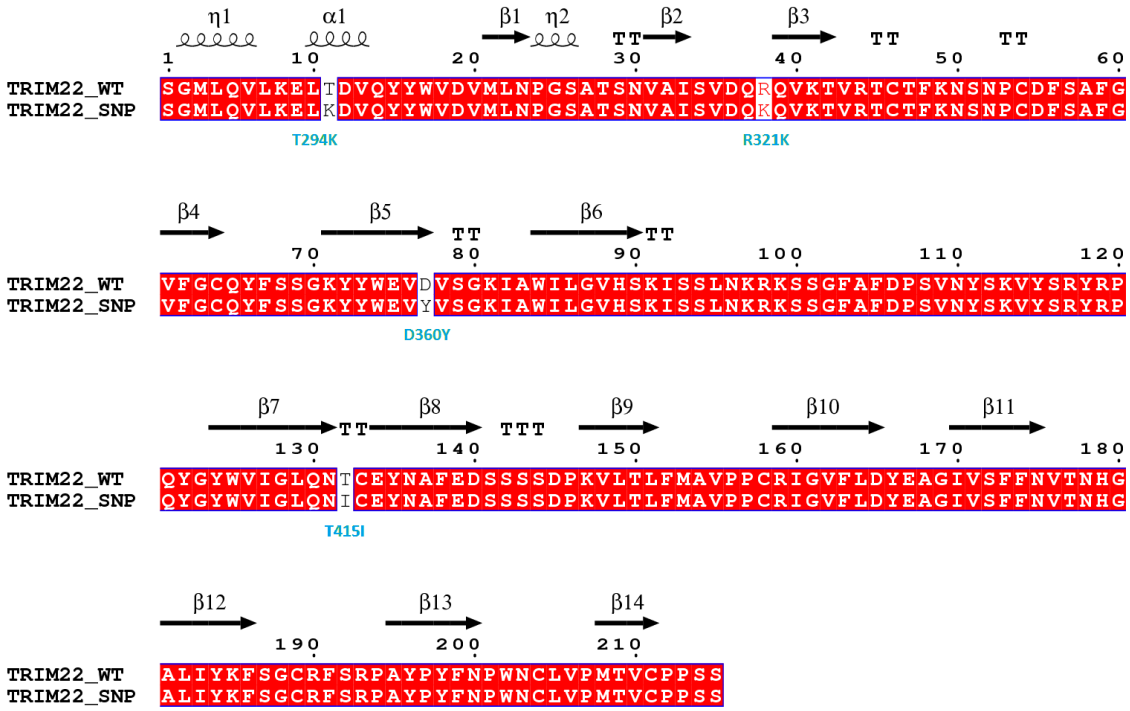


Figure 5.6 Secondary Structure Alignment of the PRYSPRY Domain of TRIM22

The TRIM22 PRYSPRY is comprised 14 β sheets and 4 Variable loops. *Polymorphisms observed in the Caio Cohort are annotated in Blue.*

5.4.8. Secondary Structure-TRIM22- PRYSPRY: Spatial Distribution of nsSNPs Relative to Restrictive Function

Many polymorphisms in the Rhesus TRIM5α PRYSPRY domain have been characterised for their role in altering the restrictive capacity of rhesus TRIM5α, specificity of TRIM5α for lentiviruses, or the strength of capsid binding. The PRYSPRY polymorphisms in Human TRIM22 were aligned with notable mutation hotspots in the rhesus TRIM5α PRYSPRY domain. A secondary structure alignment of a TRIM22 PRYSPRY mutant sequence and Rhesus PRYSPRY sequence, using the Rhesus macaque crystal structure 43BN as a scaffold to decipher the arrangement of polymorphisms at the secondary structure level.

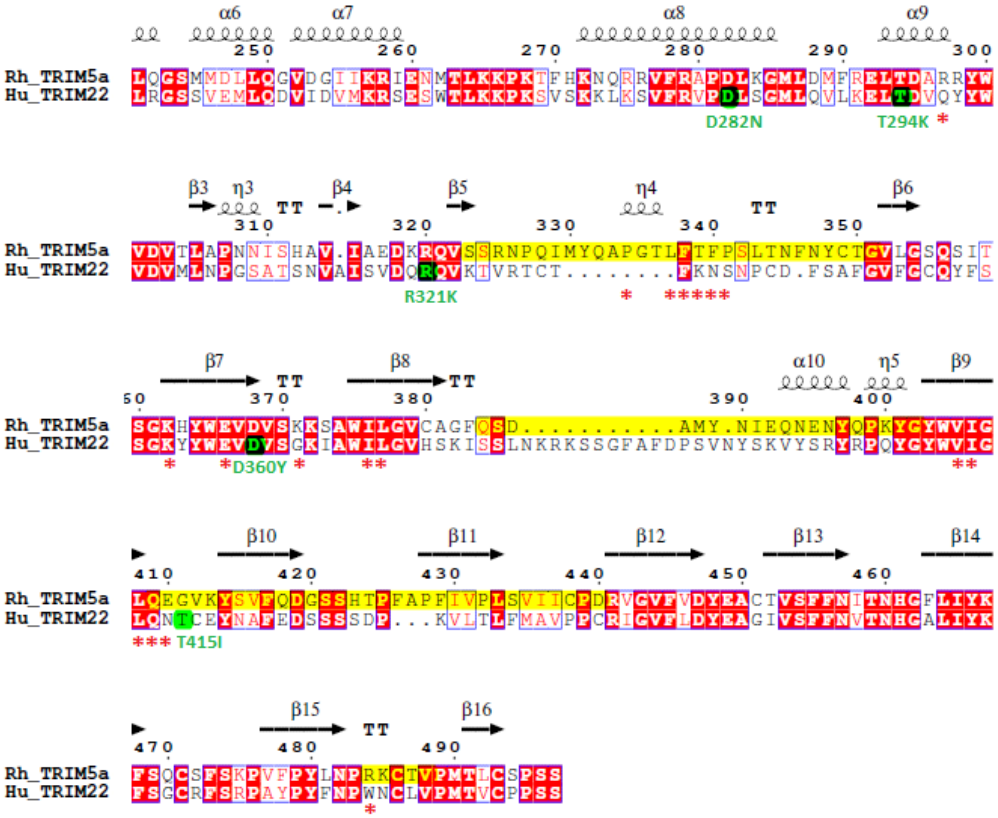


Figure 5.7 Secondary Structure Alignment of Rhesus TRIMalpha and Human TRIM22 PRYSPRY Domain

*Variable Loops=yellow, *Virus Interacting Residues, TRIM22 nsSNPs=green*

RhesusTRIM5α and Human TRIM22 PRYSPRY domains share only 50% amino acid identity, and yet both bind to and potently restrict HIV-1. Both domains possess a highly conserved inner core comprised of 14 β sheets where the majority of conservation occurs, and 4 Variable loops,

where V1 V2 V3 have been implicated in HIV-1 viral capsid interaction in the Rhesus macaque. The variable loops are the sites of least conservation between the PRYSPRY Domains in Rhesus TRIM5α and Human TRIM22, especially Variable loop 1 which is truncated in TRIM22 when compared to Rhesus TRIM5α. Data from the Caio Community Cohort revealed four polymorphisms within the human TRIM2 PRYSPRY domain, and one adjacent in the Linker 2 region.

The spatial distribution of human TRIM22 SNPs are relevant to areas of retroviral restriction in Rhesus TRIM5α. When aligned, TRIM22's R321K is located on the cusp of the Variable loop 1 region responsible for the restriction of HIV-1 and maintaining function of the PRYSPRY domain in RhTRIM5α. Further, T415I is located within the Variable 3 loop which is virus-interacting in Rhesus TRIM5α, and adjacent to three loss-of-function residues. D360Y is sandwiched between E366 and K371 which are associated with N-MLV restriction, and T294K is within three residues of R297 which is also associated with N-MLV restriction.

Table 5.4 Mutagenesis Hotspots in Rhesus Macaque Trim5α

Mutant Role	**Mutant in Rhesus Trim5α**	**Ref**
Gain the ability to inhibit HIV-1, SIV, or B/NB-MLV	P334	(209, 210)
Gain the ability to inhibit HIV-1 and BMLV and NMLV slightly	L337	(284)
Gain the ability to inhibit B-MLV	F338	(285)
Gain the ability to inhibit HIV2	T339 F340 P341	(286)
Associated with N-MLV restriction but not necessarily CA binding	R297, E366, K371, K362, R484	(287)
Required for N-MLV capsid binding	R295A (17), E410 and Q409	(239)
Loss of Function	L377	(239)
	L408	(194, 288)
	I406	(194, 288)
	E410	(239)
	Q409	(239)
	V405	(194, 288)
	I376	(194, 288)
	V1 loop	(113)

5.4.9. Tertiary Structural Examination of TRIM5α Polymorphisms Observed in the Caio Cohort TRIM5α Polymorphisms in the RING, B-Box and Coiled Coil

The majority of TRIM5α SNPs were located in the coiled coil with a further two in the RING domain. All were solvent exposed residues with the exception of L214F which was buried.

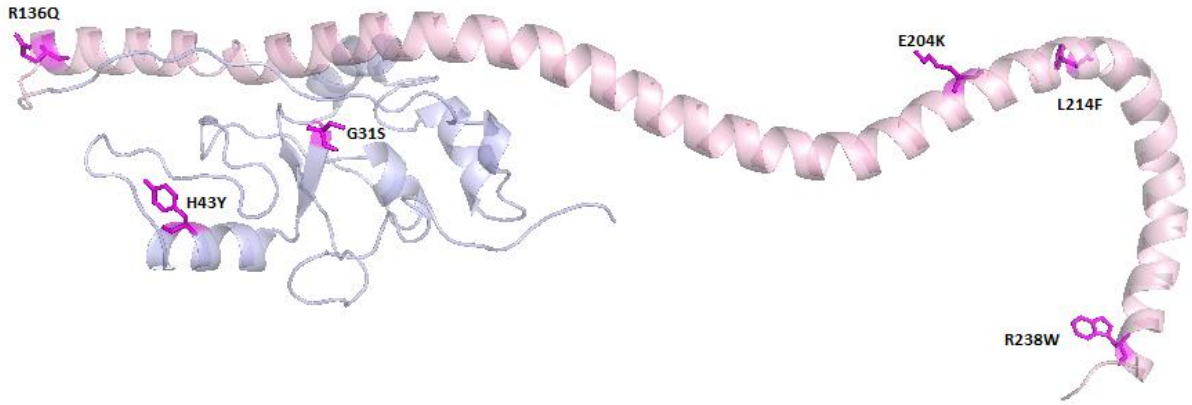


Figure 5.8 The TRIM5α RING, B-Box(2), and Coiled Coil Domain

RING/B-Box- lightblue, Coiled Coil- lightpink, Mutants- Magenta

5.4.10. TRIM5α Polymorphisms in the Linker2 and PRYSPRY Domain

Only one polymorphism was found in the TRIM5α PRYSPRY domain, P479L, which is an exposed residue. D249G was observed in the linker region, however it was not exposed.

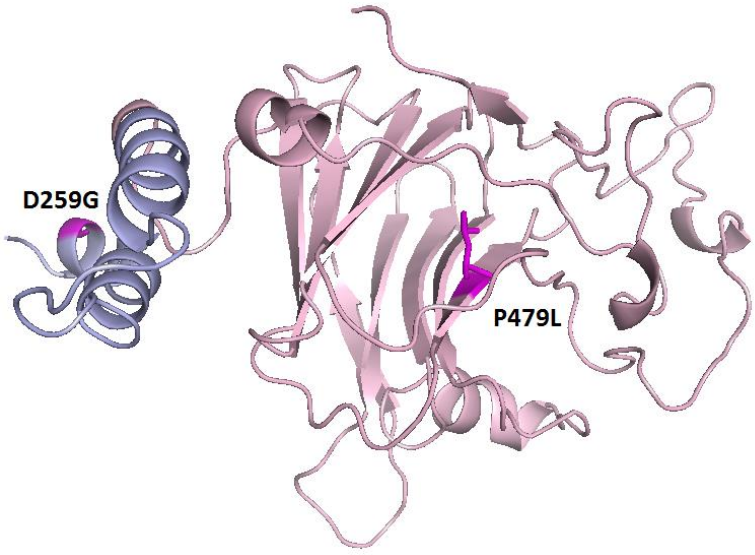


Figure 5.9 The TRIM5α Linker2 and PRYSPRY Domain

Linker2- lightblue, PRYSPRY- lightpink, Mutants- Magenta

5.4.11. 3D Structure of TRIM5α Mutants

Using Phyre 2.0, a 3D structure was generated for each TRIM5α containing each nsSNP observed in the Caio Cohort. The 3D model of wildtype TRIM5α illustrates a 74 angstrom cleft between the active cysteines of the RING domain and a key interaction point, position 332 in the PRYSPRY domain.

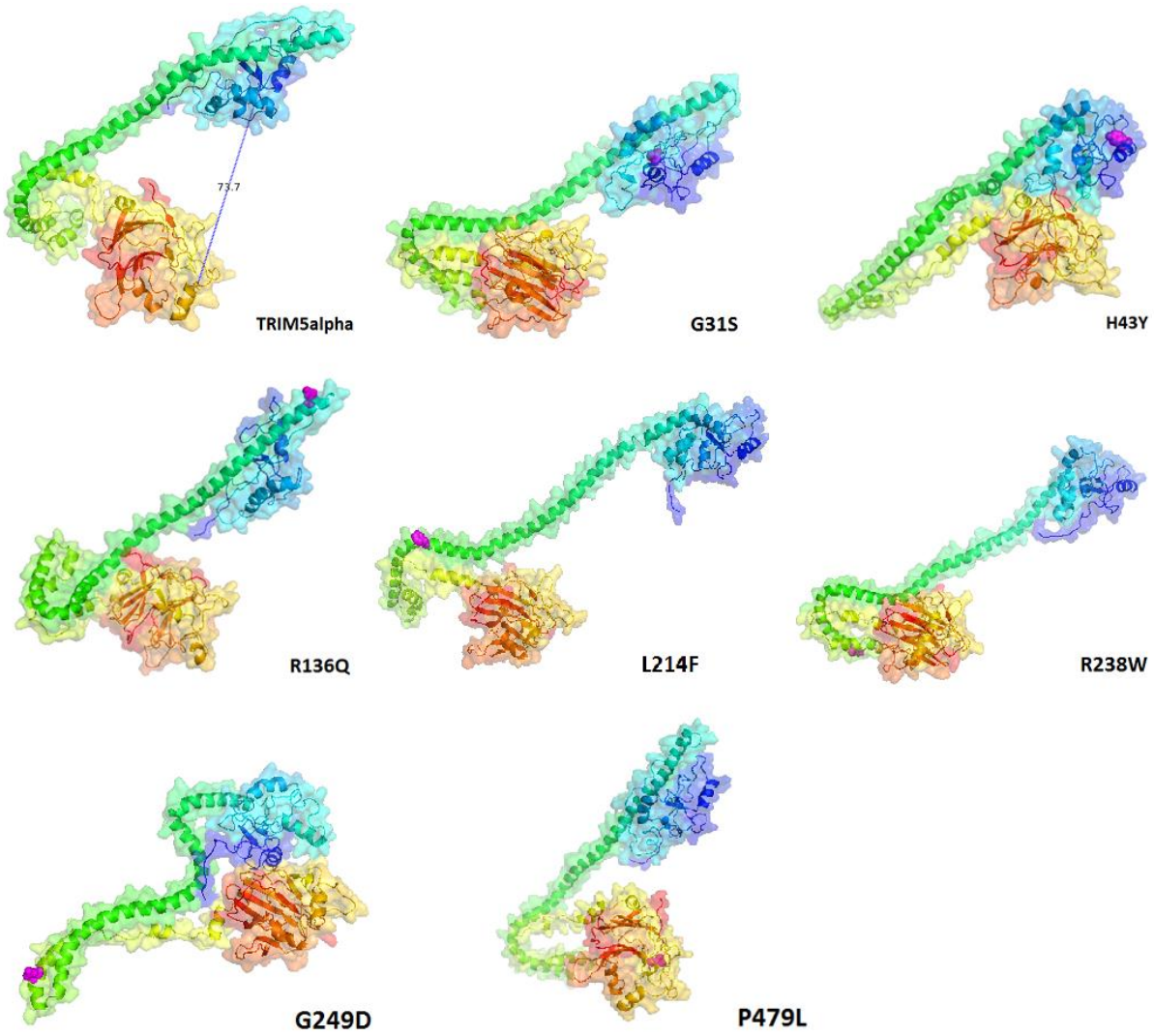


Figure 5.10 3D Models of TRIM5α and Mutants

3D Models of TRIM5alplha were generated in Phyre2.0. The domains were annotated as follows: RING- navy, BBox- aqua, Coiled- Coil- green, PRYSPRY- yellow/orange, and the amino acid substitution shown in magenta

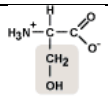
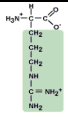
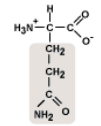
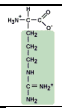
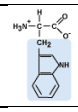
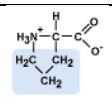
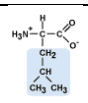
The addition of substitutions (illustrated in magenta) in modelling TRIM5α resulted in notable changes in the tertiary structure of the protein; shifting the orientation of the RING and B-Box domains, adjusting the flex in the coiled coil. In the case of H43Y, R136Q, and G249D, the cleft

between the virus-interacting PRYSPRY domain, and the functional E3-ubiquitin ligase in the RING domain, was narrowed. Of further interest is G249D, which resulted in a change in charge from neutral to acidic, and also resulted in a reduced hydrophobicity which may have the effect of destabilising a hydrophobic core. The resultant change in structural prediction shows a kink in the Linker 2 where the peptide interacted with a helix of the coiled coil domain. The same can be said for R136Q where substitution resulted in a kink in the coiled coil domain, changing the orientation of the RING and B-Box domains and drawing them closer to the PRYSPRY domain.

5.4.12. Amino Acid Substitutions in TRIM5α

Examination of the structural properties of amino acids showed profound changes in the PRYSPRY domain when compared to the other domains. There were substantive changes in hydrophobicity for P479L and R238W which both moved from hydrophilic to very hydrophobic amino acids. Given that hydrophobic amino acids are likely to be found in the interior, whereas hydrophilic amino acids are likely to be in contact with the aqueous environment, changes in hydrophobicity of this magnitude in exposed residues or hinge regions can affect structure and function a great deal. L214F and G249D had a change in size of 34 and 58 Daltons respectively,(282) though the change in charge was inconsequential. In the case of G249D, the substantial increase in volume within the turn of the in the linker region between the coiled coil and PRYSPRY domain could give rise to the significant change in conformation observed in the G249D mutant 3D model.

Table 5.5 Structural Properties of Amino Acid Substitutions in TRIM5α

nsSNP	Wildtype Residue	SNP Residue	Wildtype Charge	SNP Charge	Wildtype MW (Da)	SNP MW(Da)	Wildtype Hydrophobicity	SNP Hydrophobicity
G31S			Non Polar	Polar	75.07	105.09	Neutral (0)	Neutral (-7)
H43Y			Basic	Non-polar	155.16	181.19	Hydrophillic (-42)	Hydrophobic (49)
R136Q			Basic	Polar	174.20	146.15	Hydrophillic (-26)	Neutral (-18)
E204K			Acidic	Basic	147.13	146.19	Neutral (8)	Hydrophillic (-37)
L214F*			Non polar	Non Polar	131.18	165.19	Very hydrophobic (100)	Very hydrophobic (92)
R238W *			Basic	Non Polar	174.20	204.23	Hydrophillic (-26)	Very hydrophobic (84)
G249D*			Non Polar	Acid	75.07	133.11	Neutral (0)	Neutral -18
P479L *			Non polar	Non polar	115.13	131.18	Hydrophillic (-46)	Very hydrophobic (100)

5.4.13. Tertiary Structural Examination of TRIM22 Polymorphisms Observed in the Caio Cohort TRIM22 Polymorphisms in the RING, B-Box and Coiled Coil

The majority of TRIM22 polymorphisms were located in the coiled coil domain, and all were highly solvent exposed residues with the exception of T232A which was buried.

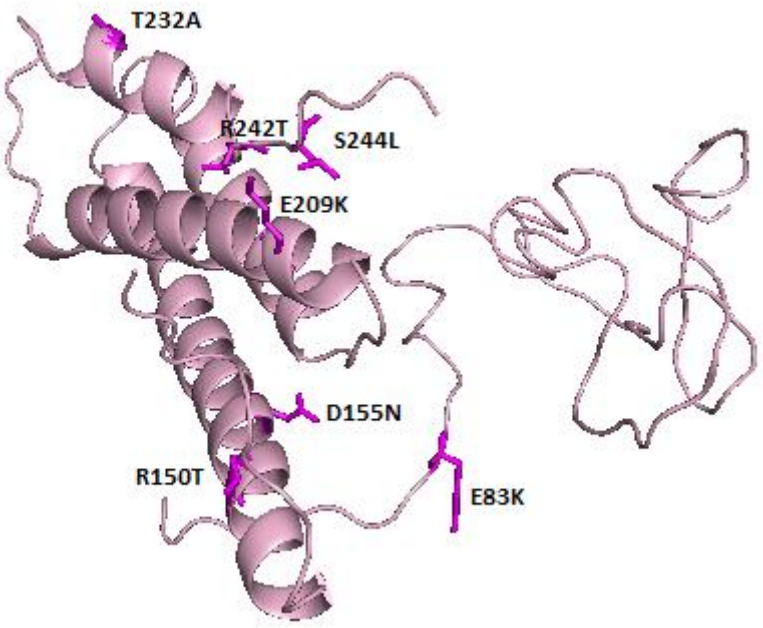
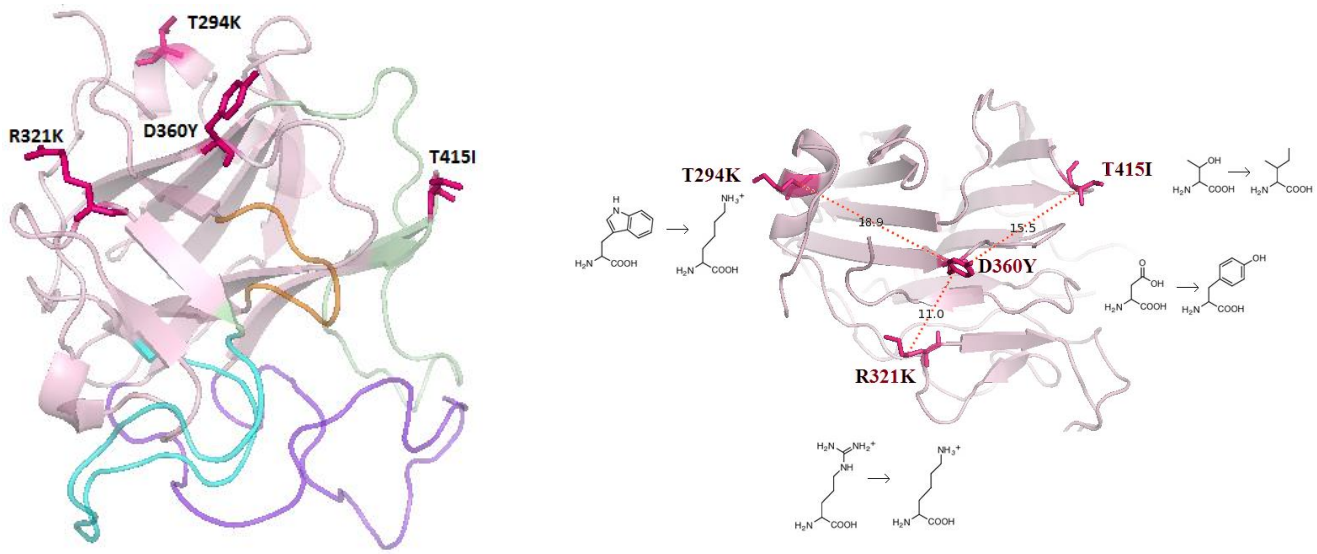


Figure 5.11 Diagram of TRIM22's RING, B-Box(1), and Coiled Coil Domain

3D model of TRIM22 was generated using Phyre2.0 and amino acid substitutions made using the mutagenesis feature in PyMol. Structure is shown in lightpink and the mutants in magenta

5.4.14. TRIM22 Polymorphisms in the PRYSPRY Domain

All of the TRIM22 polymorphisms in the PRYSPRY domain were located in the conserved β sheet with the exception of T415I which was located on the interior side of the V3 loop. Yet the polymorphisms within the PRYSPRY domain clustered 11-18 angstroms between them.



A. *B.*

Figure 5.12 Diagram of TRIM22's PRYSPRY Domain

A. The TRIM22 PRYSPRY Domain shows a highly conserved inner core comprised 14 B sheets with 4 variable loops protruding on the surface of the structure. Polymorphisms- hotpink V1- cyan, V2-purple, V3- lime green, V4- orange. T415I is located in the putative virus-interacting V3 loop, whereas T294K, R321K, and D360 are located among the conserved β-sheets on the posterior side of the PRYSPRY domain. B. All SNPs of TRIM22's PRYSPRY domain cluster together, with 11-18 Angstroms between them.

5.4.15. 3D Structure of TRIM22 Mutants

Using Phyre 2.0, a 3D structure was generated for TRIM22 containing each nsSNP observed in the Caio Cohort. The 3D model of TRIM22 illustrates an 81 angstrom cleft between the active cysteines of the RING domain and a key virus interaction point, position 332, in the PRYSPRY domain. This represents a larger distance between the same two points than in TRIM5α.

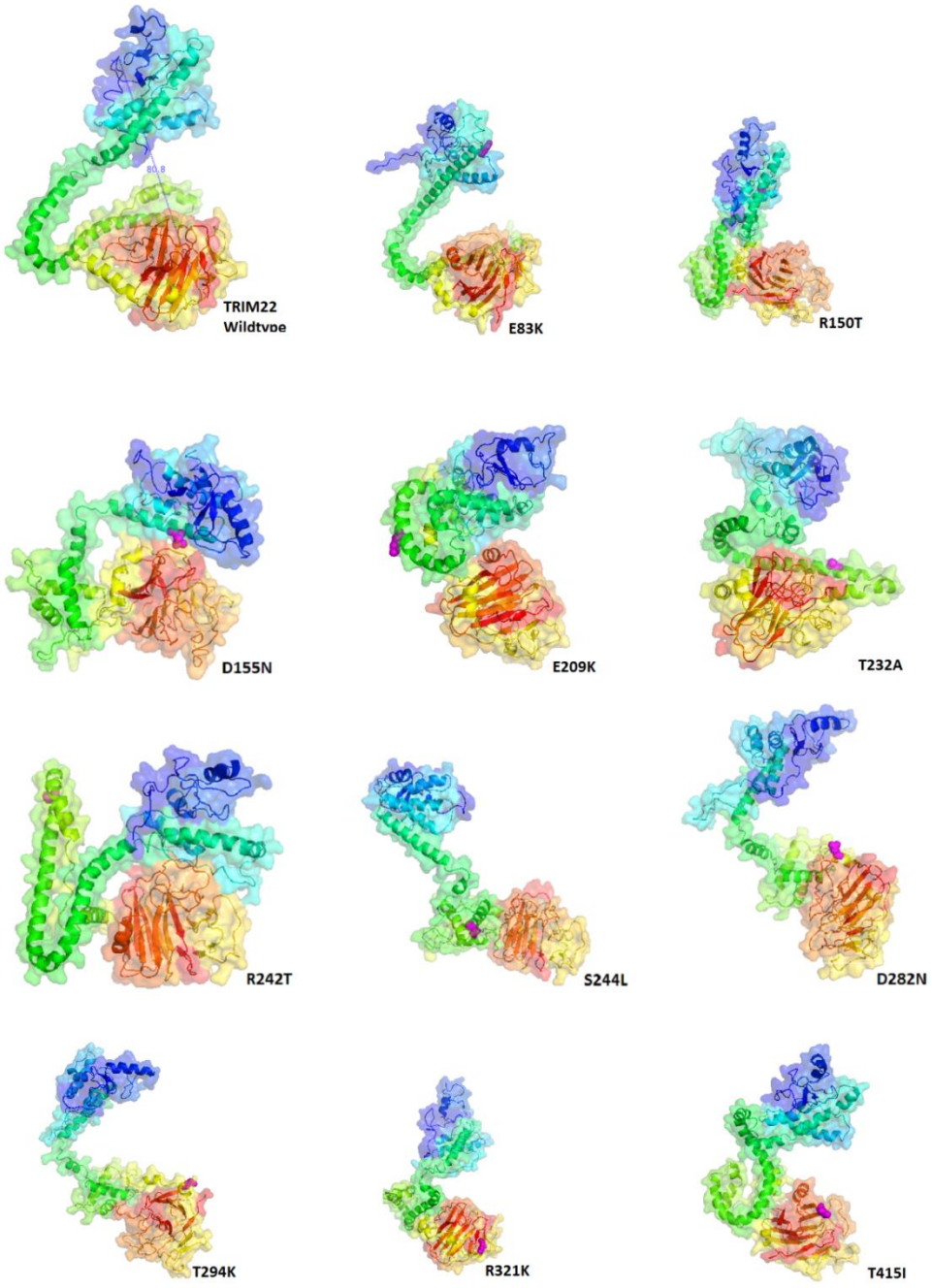


Figure 5.13 3D Models of TRIM22 and Mutants

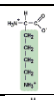
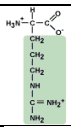
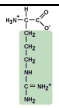
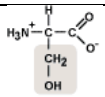
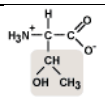
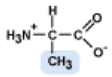
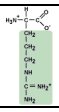
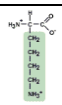
3D Models of TRIM22 were generated in Phyre2.0. The domains were annotated as follows: RING-navy, BBox- aqua, Coiled- Coil- green, PRYSPRY- yellow/orange, and the amino acid substitution shown in magenta

The addition of nsSNPs (illustrated in magenta) in modelling TRIM22 resulted in notable changes in the tertiary structure of the TRIM22 protein models. In some cases, the cleft between the virus-interacting PRYSPRY domain, and the functional E3-ubiquitin ligase in the RING domain, was narrowed or widened. D155N showed notable changes in structure, narrowing the cleft of TRIM22 due to a change in charge which may have drawn the PRYSPRY closer. The coiled coil in T232A folded upon itself, suggesting the hydrophobic nature of A may have altered the conformation allowing the residue to be buried. Also the R242T model, showed folding of the coiled coil acting to close the cleft between the RING and PRYSPRY domains.

5.4.16. Amino Acid Substitutions in TRIM22

On examination of the amino acid properties of TRIM22 polymoprhisms, R150T and R242T showed substantive decreases in volume in the magnitude of 58 Daltons. Both 3D structures exhibited large conformational changes disjointing the coiled coil region of the protein, and closing the cleft between the RING and Pryspry domains. S244L and T415I resulted in changes to very hydrophobic residues but would appear to be in regions inconsequential for the overall structure for the protein. In 3D models a profound change was exhibited by T232A and D155N, both of which resulted in folding of the coiled coil upon itself either obstructing the protein cleft in the case of 232A or closing it in the case of D155N.

Table 5.6 Structural Properties of Amino Acid Substitutions in TRIM22

nsSNP	Wildtype Residue	SNP Residue	Wildtype Charge	SNP Charge	Wildtype MW (Da)	SNP MW(Da)	Wildtype Hydrophobicity	SNP Hydrophobicity
E82K			Acidic	Basic	147.13	146.19	Neutral (8)	Hydrophillic (-37)
R130T			Basic	Neutral	174.20	119.12	Hydrophillic (-26)	Neutral (13)
D155N			Acid	Polar	133.11	132.12	Neutral (-18)	Hydrophillic (-41)
R150T*			Basic	Polar	174.20	119.12	Hydrophillic (-26)	Neutral (13)
R242T*			Basic	Polar	174.20	119.12	Hydrophillic (-26)	Neutral (13)
S244L*			Polar	Non polar	105.09	131.18	Neutral (-7)	Very hydrophobic (100)
T232A			Polar	Non polar	119.12	89.10	Neutral (13)	Hydrophobic (47)
D282N			Acid	Polar	133.11	132.12	Neutral -18	Hydrophillic(-41)
T294K			Polar	Basic	119.12	146.19	Neutral (13)	Hydrophillic (-37)
T415I*			Polar	Non Polar	119.12	131.18	Neutral (13)	Very hydrophobic (100)
R321K			Basic	Basic	174.20	146.19	Hydrophillic (-26)	Hydrophillic (-37)

Consurf Results: TRIM22

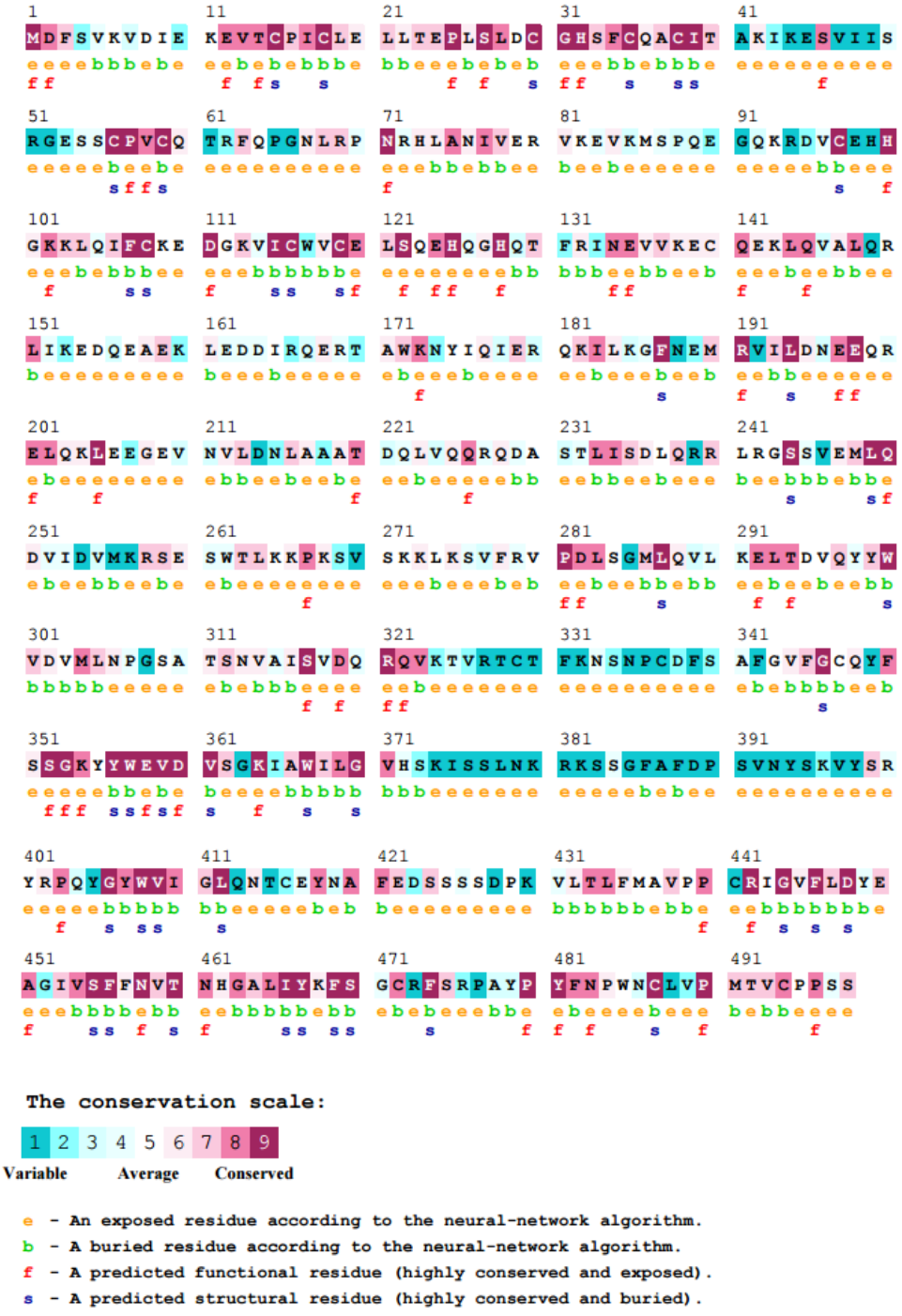


Figure 5.14 Consurf Results for TRIM22

Colours of the ConSurf output indicate the level of sequence conservation. Purple indicates conservation and blue indicates variability. Residues are predicted to be exposed (e), buried (b), functional (i.e., highly conserved and exposed; f), or structural (i.e., highly conserved and buried, s).

Consurf Results TRIM5α

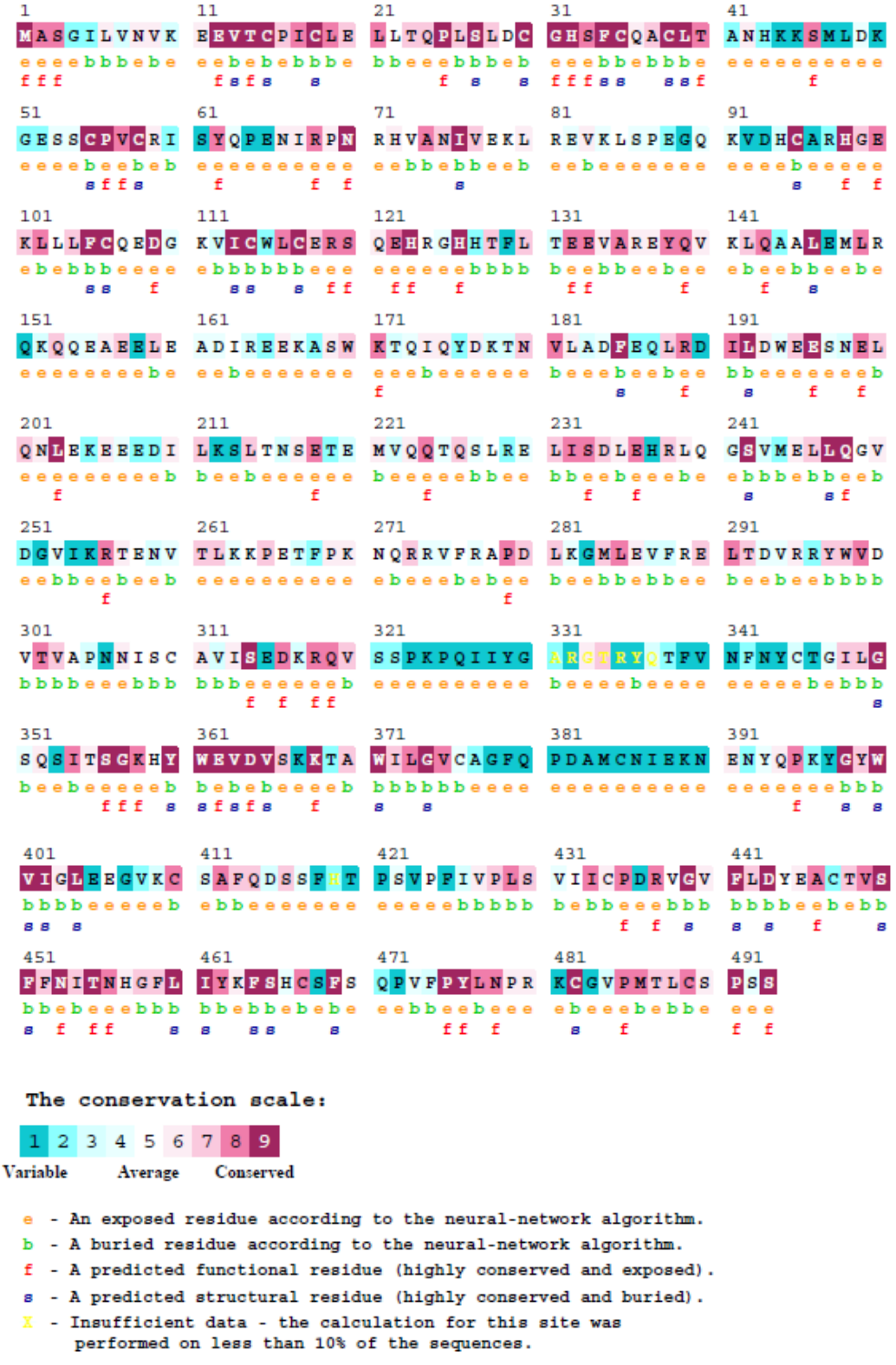


Figure 5.15 Consurf Results for TRIM5α

Colours of the ConSurf output indicate the level of sequence conservation. Purple indicates conservation and blue indicates variability. Residues are predicted to be exposed (e), buried (b), functional (i.e., highly conserved and exposed; f), or structural (i.e., highly conserved and buried, s).

5.5. Discussion

Taken together, phylogenetic, sequence, and structure based algorithms predicted G31S, L214F, and P479L as TRIM5α nsSNPs likely to have a functional impact. However on examination of the 3D models produced of mutant proteins, the H43Y, R136Q, and G249D substitutions appeared to have a profound effect in structural predictions; providing substantive changes in amino acid weight and hydrophobicity to alter the cleft of the TRIM5 active site. Amino acid changes in hydrophobicity were profound in the case R238W which both moved from hydrophilic to a very hydrophobic amino acid. Further, L214F and G249D resulted in substantial changes in weight, which in the case of G249D may have given rise to the demonstrable change in confirmation when compared to the wildtype in predicted 3D models. Consurf predicted none of these SNPs to be evolutionarily conserved, however all were predicted to be exposed residues. When combining the phylogenetic, sequence, structural and heuristic exploration of TRIM5α polymorphisms, P479L and G249D appear to be the most potentially relevant in TRIM5α structure and function.

In the case of TRIM22, phylogenetic, sequence, and structure based algorithms predicted only SNPs in and around the PRYSPRY domain as likely to have a functional impact: D282N, D360Y, and R321K. However the 3D models told a different story, showing SNPs within the coiled coil domain like R150T, D155N, R242T, and T232A, capable of exerting profound changes on protein confirmation and the distance between the PRYSPRY and RING domain. Again, R150T and R242T resulted in substantive decreases in amino acid weight whereas S244L and T415I resulted in substantial changes to very hydrophobic residues. An examination of secondary structure showed both R150 and 242 were located around the strict alpha turn of the coiled coil domain which acts as the hinge controlling the distance between the RING and the PRYSPRY domain, consistent with 3D models of the two mutants. An examination of the tertiary structure of the 4

TRIM22 PRYSPRY domain substitutions, despite their linear separation in the primary structure, all clustered together on the exterior of the PRYSPRY domain. T415I protruded from the virus-interacting Variable 3 loop, while 360Y, 321K and 294K were distributed in and around the β strands. Consurf predicted all of these residues as exposed, and in addition D360, R321, and T415 were predicted to be functional residues on account of their being highly conserved and exposed. Consurf also indicated S244 as a highly conserved residue though buried, and D155 and R242 as solvent exposed without conservation, and finally R150 as conserved and exposed. The linker position D282 was predicted to be highly conserved and exposed. Taken together, the data suggests high structure function involvement for the substitutions D282, R321K, and D360Y, however the only way to define the changes conferred by these amino acid substitutions is to crystallise the TRIM proteins and their mutants under similar conditions.

There were 50% more polymorphisms observed in TRIM22 when compared to TRIM5α in the Caio Cohort. Furthermore, there is no spatial similarity between polymorphisms observed in TRIM5α and TRIM22. However, there is more spatial similarity between polymorphisms observed in TRIM22's PRYSPRY domain, and the antiviral patches of Rhesus TRIM5α which are major determinants of retroviral recognition and viral specificity. In both TRIM22 and RhesusTRIM5α, the PRYSPRY domain consists of β strands in tandem that form a core of mostly conserved residues, with loops of different lengths interspersed between them. In Rhesus TRIM5α, these loops, especially Variable Loop 1 around the area of β strand 2 and 3, contain virus interacting regions. In TRIM22, R321K was located on the cusp of the Variable loop 1 region where T415I was located within the Variable 3 loop and aligned adjacent to three loss-of-function residues in Rhesus TRIM5α. Further D360Y and T294K were adjacent to residues associated with N-MLV restriction. Proportionally, TRIM5α possessed a greater ratio of nsSNPs to synonymous SNPs. Regardless, it would appear that the nsSNPs in TRIM22 were more

spatially relevant in the virus-interacting PRYSPRY domain and in areas which have the potential to alter protein conformation and oligomerisation.

Should retroviral selection pressure from HIV-2 be exerted on either of the TRIM proteins within the Caio Community Cohort, it would appear that it is more likely to be TRIM22. Further, association study data from Chapter III and IV of this thesis suggest that substitutions appear more relevant in the context of TRIM22 when compared to TRIM5α as HIV-2 outcome was significantly affected by TRIM22 SNPs at the 5% level following adjustment; an effect not observed in the case of TRIM5α SNPs. The distribution of substitutions in TRIM22 when compared to human TRIM5α would suggest that the positive selection of TRIM22 has been concentrated in the same regions as positively selected residues, which also contribute to retroviral specificity in rhesus TRIM5α. However, though the substitutions in TRIM22 were associated with HIV-2 progression more so than in TRIM5α, it is difficult to determine if virus restriction plays a role in driving the evolution of these two proteins within this community, and if so which virus. The pressure exerted on these proteins is not likely to be due to HIV-2 given that it was introduced relatively recently into the population. HTLV-1 has been endemic in West Africa for tens of thousands of years and has a high prevalence of 4.9% in the Caio Community(95) but has a low evolutionary rate suggesting there may be little selective pressure exerted. (94) The relationship between TRIM proteins and HTLV has not been investigated to date.

In light of these findings, from an evolutionary perspective, the fact that the cow encodes multiple TRIM5 genes whereas the dog encodes no TRIM5, suggests that another gene has absorbed TRIM5's antiretroviral function, which could be TRIM22. Within the Caio Community Ccohort, it is evident that the frequency of polymorphisms is unusually high in both TRIM5 and TRIM22, but that the substitutions in TRIM22 are more frequent, and more functionally relevant, when compared to TRIM5α.

Chapter 6. Pilot Examination of the Restrictive Capacity of TRIM22 Against Retroviruses and the Ability of the rs1063303 Mutation to Diminish This Effect

6.1. Introduction

In the context of retroviral infection, restriction factors are characterised by a major criterion:

A significant decrease in retroviral activity during their overexpression, or attenuation of retroviral activity when they are deficient(109, 289).

Known restriction factors, such as TRIM5α, APOBEC, Tetherin, and SAMHD1 have additional features in common such as:

• Derivation from functionally autonomous genes that are unusually variable(113, 290-292)

• Sensitivity to immune activation and viral infection resulting in notable changes in their expression(208, 293, 294)

• The identification of specific viral proteins which directly resist their action(70, 166, 295-297)

Chapters III and IV of this thesis have illustrated a high degree of genetic variation in TRIM5α and TRIM22 in the Caio Community Cohort. Chapter IV highlighted two polymorphisms which appear relevant in HIV-2 acquisition and disease progression: It demonstrated that while TRIM22 mutation rs7936654 has a positive impact on survival, rs1063303 results in increased susceptibility to HIV-2 *in vivo*, and faster disease progression culminating in decreased survival time. Furthermore, Chapter V of this thesis illustrated the signatures of positive selection with *in silico* algorithms and models, illustrating the multiple TRIM22 residues in the PRYSPRY domain which cluster in putative regions of interaction with HIV-1, and the potential impact on protein structure derived from polymorphisms in the coiled coil domain of TRIM22. Given the evidence that TRIM22 has undergone positive selection in primates(237), the frequency of non-synonymous mutations in TRIM22 observed this study, and strong association of two SNPs with HIV-2 related mortality: TRIM22 shows evidence of an evolutionary conflict with an agent which it may act to restrict.

There have been 3 major studies carried out to investigate the restrictive capacity of TRIM22:

1. The first in 2008, identified TRIM22 as one of the most strongly upregulated genes in response to treatment with IFN-β(224). The authors utilised HOS cells generated to express CD4/CXCR4 and stably expressed TRIM22 in a portion of these cells prior to HIV-1 infection. They observed that cells with TRIM22 released less virus than control cells. The paper went on to examine the point in the HIV-1 lifecycle at which TRIM22 acted: in the presence of TRIM22, substantial amounts of intracellular gag were retained in cells, however gag failed to accumulate in the extracellular supernatant in various cell types. To determine how TRIM22 expression at physiological levels and conditions would affect HIV-1 replication, they compared the effects of interferon β on cells depleted for TRIM22 using RNAi and found that HIV particle release was significantly increased in the knockdown cell types. Confocal microscopy of cells transfected with Gag-GFP and TRIM22 in the presence of cycloheximide showed that only 12% of cells had fluorescence at or near the plasma membrane, when compared to 65% of cells without TRIM22 expression. The research also confirmed an association between TRIM22 and Pr55Gag in immunoprecipitation assays.
2. The second study, in 2011(226), involved the use of permissive and nonpermissive promonocytic cells to HIV-1 IIIB infection. The U937 cell line, characterised by the constitutive expression of CD4 and CXCR4 (but not CCR5)(298), has been broadly utilised to dissect viral and/or host determinants modulating CXCR4-dependent human immunodeficiency virus type 1 (HIV-1) replication(299). Through limited dilution cloning, the Vicenzi group (226)identified clones of the U937 cell line that were either permissive or nonpermissive for CXCR4 dependent HIV-1 replication. There were no differences between these permissive and nonpermissive cells, other than a constitutive release of tumor necrosis factor alpha (TNF-α), which was shown to boost virus replication and CXCR4 expression. These clones were investigated further in the search for interferon-

induced host restriction factors that could explain their differential restriction during the early stages of HIV-1 life cycle. TRIM22 was the only other known restriction factor differentially expressed at the level of transcription in nonpermissive, but not permissive cells, confirming earlier accounts that TRIM22 could inhibit HIV replication at transcriptional, post-transcriptional, or post-translational levels(224, 225). Further, TRIM22 was observed in the nucleus of nonpermissive and absent in permissive U937 cells. This study showed that non permissive cells impaired the activity of VSV-G lentiviral vectors with a reporter gene under the control of the LTR of HIV-1 (clade B), that this effect could be rescued with shRNA mediated knockdown of TRIM22, and that the inhibition occurred independently of its E3 ubiquitin-ligase activity and of the tandem NF-κB binding site within the viral promoter(226).

3. The third study, in 2014, went further, clarifying not only the impact of TRIM22 on restriction, but also assessing the functional impact of the rs1063303 mutation on HIV replication. The study demonstrated that TRIM22 carrying the rs1063303 mutation showed 40 fold higher mRNA expression than the TRIM22 wildtype in qPCR experiments, and 10.3 fold higher protein expression in densiometric analysis of western blots. Further, in HOS or HeLa cells constransfected with proviral HIV-1 (pR9) and either TRIM22 wildtype or TRIM22-SNP, cells expressing increasing concentrations of wildtype TRIM22 exhibited decreasing levels of intracellular HIV-1 Gag protein expression. In contrast, cells expressing increasing concentrations of 1063303-TRIM22 did not exhibit a substantial reduction in intracellular Gag protein production. Taken together, these data showed that the SNP rs1063303 variant had an inverse functional impact where it increased TRIM22 expression and decreased the antiviral activity of TRIM22(269).

The mechanism of TRIM22 restriction has yet to be established, and two of the three functional studies of TRIM22 investigate the role of the E3 ligase activity mediated by the RING domain, with each study generating conflicting results. Barr *et al*., in 2008(224), described late stage effects of TRIM22 where E3 ligase activity was essential for restriction, whereas Vincenzi in 2011(226) described a post entry effect of TRIM22, and its inhibition of LTR driven HIV-1 transcription. TRIM5α, the close relative of TRIM22, is a classed as a restriction factor. It is unknown whether TRIM22 restriction is mediated by the disruption of virus entry and uncoating, in a similar mechanism to that of TRIM5α(195). Rhesus TRIM5α has been shown to interact with the HIV capsid protein, Gag(113, 194). One of the mechanisms that TRIM5α is thought to employ is the premature uncoating of HIV-1 Gag. This effect is largely dependent on coiled coil-mediated multimerisation of TRIM5α around the incoming HIV capsid. Though the mechanism of TRIM22 restriction has yet to be established, it is clear that there is an interaction with HIV-1 Gag(224). Therefore an important key to the puzzle would be to determine whether TRIM22 restriction is dependent upon the multimerisation and eventual hexameric structure around the HIV-1 capsid in a manner essential for TRIM5α-mediated restriction(199). The relevance of TRIM22 multimerisation to restriction has yet to be described. Further, Human TRIM5α is largely incapable of restricting HIV-1, but it possesses activity against SIVmac, N-MLV, and through three amino acid modifications, HIV-2. From an evolutionary perspective, the cow has multiple copies of TRIM5α, with an absence of TRIM22, where dogs and cats have complexly lost the TRIM5 gene, retaining its paralogue TRIM22(237). Given the vast similarities between, but the discordant evolution which occurs between TRIM22 and TRIM5α, it is possible that TRIM22 could have a compensatory restriction mechanism for HIV-1. Human TRIM22 may therefore lend itself to the restriction of lentiviruses where human TRIM5α may have failed.

No study to date has investigated the restrictive capacity of TRIM22 on any other lentivirus than HIV-1, or determined whether the detrimental rs1063303 mutation observed *in vivo* and in functional studies of HIV-1 has a detrimental impact on any other lentivirus. In the light of data

presented in Chapter IV of this thesis suggesting an association between TRIM22 polymorphisms and HIV-2 disease progression and survival, it would be advantageous to determine whether this effect is shared by TRIM22 *in vitro*, and if the rs1063303 mutation, resulting in amino acid substitution R242T, is detrimental to this restriction. In light of *in silico* data presented in Chapter V of the this thesis, indicating that R242T also would be predicted to result in changes in protein structure and function, it would be advantageous to interrogate whether the R242T substitution results in profound changes in protein structure and function *in vitro*. The binding of TRIM22 to HIV-1 gag has been well characterised by immunoprecipitation studies(224), but it would now be expedient to determine whether TRIM22 also interacts with the HIV-2 capsid.

6.2. Hypothesis and Aims

In light of the deficiencies in information surrounding the restrictive capacity of TRIM22 and its variants, this chapter aims to provide initial evidence to warrant an extensive investigation into the restrictive capacity of TRIM22 against a broad spectrum of lentiviruses. This chapter proposes to investigate whether a naturally TRIM22-deficient system would result in increased HIV-1 and HIV-2 activity in comparison to an identical system with endogenous TRIM22. Further it intends to investigate the restrictive capacity of TRIM22 and TRIM22-1063303 mutant (TRIM22-106) in the context of other retroviruses such as SIVmac239, SIVsm543, and FIV. Further, as felines do not possess TRIM5α, and FIV can be mildly restricted by human TRIM5α in a feline cell line, the study also aimed to investigate whether TRIM22 had the potential to provide a compensatory post entry block to FIV infection. Finally, the relevance of the multimeric states of the highly similar TRIM5α protein to HIV-1 restriction, in conjunction with the protein models generated in Chapter V of this thesis, gives rise to the need for examination of the oligomeric states of TRIM22 and TRIM22-106 in order to determine whether the variation in restriction lies in the ability of the respective proteins to oligomerise. This chapter therefore

aims to also provide evidence for future proteomic studies of TRIM22 and TRIM22 with the variant R242T.

This chapter therefore aims to:

a. Determine whether TRIM22 possesses antiviral capacity against HIV-2 in the same manner in as it does HIV-1 within systems that naturally express and lack endogenous TRIM22

b. Examine the restrictive capability of exogenous TRIM22 against lentiviruses other than HIV-1, and determine whether the TRIM22-106 mutation diminishes this effect as seen *in vivo* for HIV-2

c. Given that TRIM5α restriction is dependent on its multimeric states, generate materials, methods, and preliminary data to facilitate an investigation of the oligomeric states of TRIM22 and TRIM22- 106

6.3. Materials

6.3.1. Cell Lines

6.3.1.1. C8166

C8166 cells were used as host to HIV-1-IIIB and HIV-2-CBL20 infection. C8166 cells are highly sensitive to HIV and show extensive syncytium formation within a short time post infection(300). They cluster in suspension culture, express CD64, CD32, and CD16 (markers of activated T-cells) on their surface and are able to transmit non-neutralised HIV-1 into CD4 negative cells(301). C8166 cells were obtained from the NIH-AIDS Programme. They were cultured at 37°C with 5% CO_2, and maintained in cell culture media RPMI-1640+10% FBS +2mML-Glutamine +100 U/ml Penicillin and 0.1mg/ml Streptomycin at a concentration of 0.5 million cells/ml.

6.3.1.2. HEK

Human Embryonic Kidney (HEK293T) cells were used for:

a. Plasmid transfection in order to created pseudotyped viruses
b. Infection with VSV-G pseudotyped viruses (due to demonstrating that they are TRIM22 deficient)
c. Overexpression of TRIM22 and variants

This cell line is a highly transfectable derivative of HEK 293 cells, and contains the SV40 T-antigen(302). This leads to high levels of replication when transfected with vectors carrying the SV40 region of replication, and therefore gives high titres when used to produce pseudotyped retroviruses(303). They were cultured at 37°C with 5% CO_2, and maintained in cell culture media DMEM + 10% FBS+2mM L-Glutamine+100 U/ml Penicillin and 0.1mg/ml Streptomycin at a concentration of 0.5 million cells/ml and split with Trypsin EDTA 0.25%.

6.3.1.3. U937

The U937 cell line is characterised by the constitutive expression of CD4 and CXCR4 (but not CCR5)(299). Clones of the U937 promonocytic cell line that were either permissive or nonpermissive for HIV-1 replication were identified(304). These clones were investigated further in the search for host restriction factors that could explain their differential capacity to support HIV-1 replication. TRIM22 was the only factor constitutively expressed in the nucleus of nonpermissive and absent in permissive U937 cells(226). At the genetic level, the TRIM22 in nonpermissive U937s is heterozygous for both rs7935564 and rs1063303 which code for the polymorphisms R242T and D155N. Permissive and nonpermissive U937s were infected with VSV-G psueodtyped HIV-1 and HIV-2. They were cultured at 37°C with 5% C02. They were maintained in cell culture media RPMI-1640 + 10% FBS+2mM L-Glutamine+100 U/ml Penicillin and 0.1mg/ml Streptomycin at a concentration of 0.5 million cells/ml.

6.3.2. Viruses

6.3.2.1. HIV-1 IIIB

The HIV-1 IIIB (clade B, CXCR4-tropic) isolate has a high capacity to replicate in human T-cell lines(305). It was obtained from the Program EVA Centre for AIDS Reagents, National Institute for Biological Standards and Control (NIBSC) and expanded. HIV-1 IIIB viral stocks were prepared by propagation in primary CD4+ cells and virus-containing supernatant was harvested at day 6 post-infection, aliquoted and frozen at −80 C. 50% tissue culture infectious dose (TCID50) was calculated as described previously(306).

6.3.2.2. HIV-2 CBL-20

The HIV-2 CBL-20 isolate utilises multiple co-receptors including CXCR5 and CCR5, and grows rapidly in T-cell lines(307). It was originally isolated in 1988 from a Gambian patient with AIDS(308).

6.3.2.3. VSV-G Pseudotyped Viruses

Pseudotyped viruses were prepared by the transient transfection of multiple plasmids into HEK293T cells using GeneJuice Transfection Reagent. For HIV-1, confluent HEK293T cells were transfected with PMDG (Vesicular stomatitis virus envelope protein (VSV-G) expression vector], HIV-1gag-pol expression vector p8.9, and retroviral expression vector encoding eGFP SIN CGSW in a 1:1:2 ratio. [CSGW is an HIV-1 vector expressing eGFP under the control of the spleen focus-forming virus (SFFV). P8.9 expresses HIV-1 gag and pol from the CMVie promoter and bears inactivating mutations in env.] Similar experiments were performed using VSV-G pseudotyped GFP encoding retroviral vectors derived from HIV-2, FIV, and SIVmac in the vector combinations described below(309). 48 hours after each transfection, the supernatant was harvested, filtered and stored at -80.

Table 6.1 VSV-G Pseudotyped Lentiviral Vectors

eGFP Expressing Virus	Virus Clone	Gag Pol	Env	Reference
HIV-1	pCSGW	P8.9	VSV-G	(310, 311)
HIV-2	pSVRΔNBDM	pSVRΔNBeGFPΔH	VSV-G	(312)
SIVmac	pSIVmac	SIV3+	VSV-G	(310, 313)
SIVsm	SIVsm	-	VSV-G	*Gift from Prof. Greg Towers, UCL*
FIV	FIV	FIVgp	VSV-G	(309, 314)
EIAV	pONY8.0	pONY3.1	VSV-G	(315)

6.3.3. Antibodies

Antibodies were used for the applications listed below:

Table 6.2 Antibodies used in Functional Assays

Target	Conjugate	Clone	Clonality	Source	Application
HIV-1 p24, HIV-2 p26	FITC	KC57*	Polyclonal	Beckman Coulter	Flow/Confocal
TRIM22	Unconjugated	-	Monoclonal	Abcam	Flow/Confocal
TRIM5α	Unconjugated	-	Monoclonal	Abcam	Flow/Confocal
Goat Anti-mouse	AF-568	-		Abcam	Flow/Confocal
Goat Anti-mouse	APC	Poly4053		Biolegend	Flow
Anti-Flag	PE	-		Cambridge Biosciences	Flow
Actin (rabbit)	Unconjugated		Polyclonal	Abcam	Western
Flag (mouse)	Unconjugated	M2	Monoclonal	Sigma	Western
TRIM22 (rabbit)	Unconjugated	-	Polyclonal	Sigma	Western
HIV-1/2 p24 (mouse)	Unconjugated	B1216M	Monoclonal	Abcam	Western
IR Dye Goat Anti-Mouse IgG	Unconjugated	-	-	Licor	Western
IR Dye Donkey Anti-Rabbit IgG	Unconjugated	-		Licor	Western
Heatshock Protein 70	Unconjugated	C92F3A5	Monoclonal	Stressmarg Bio	Western
Histone H3	Unconjugated	-	Polyclonal	Abcam	Western

**Cross reacts with HIV-2 p26 (34)*

6.4. Methods

6.4.1. Generation of Flag-tagged TRIM22 wildtype and TRIM22-106

The coding region of wild-type TRIM22 (GenBank Accession NM_006074.4) was subcloned into p3xFLAG-CMV-10 (Sigma, St. Louis, MO) using HindIII and XbaI restriction sites to generate a plasmid encoding flag-tagged wild-type TRIM22 (pTRIM22). The flag-tagged nsSNP rs1063303:G>C plasmid (TRIM22-106) was created using site-directed mutagenesis. TRIM22-106 contains C in place of the ancestral allele G at nucleotide position 725 of the TRIM22 coding region (GenBank Accession NM_006074). The following primers were used to amplify pTRIM22 and generate PCR fragment 1: Forward WT (5' ACG TAA GCT TAT GGA TTT CTC AGT AAA GG 3') and Reverse SNP (5' GAC GAT CCC GTC AAC CTC CGC TGG AGA 3'). Similarly, PCR fragment 2 was generated using following primers: Forward SNP (5' TCT CCA GCG GAG GTT GAC GGG ATC GTC 3') and Reverse WT (5' ACG TTC TAG ATC AGG AGC TCG GTG GGC ACA CAG 3'). A 1:25 dilution of PCR fragment 1 and PCR fragment 2 was added with "Forward WT" and "Reverse WT" primers together in a PCR reaction. The amplified TRIM22 coding region containing the nsSNP was

cloned into p3xFLAG-CMV-10 (Sigma) using HindIII and XbaI to generate pTRIM22-106. The entire coding region of TRIM22-106 was sequenced and no other PCR-introduced variations were detected.

6.4.2. Transfection of TRIM22 Flag Tagged Plasmids into HEK293Ts

Cells were seeded into 150cm^3 flasks and plated to a final count of 15million cells. 10ug of TRIM22 and TRIM22-106 plasmid were transfected into the HEK293Ts using GeneJuice transfection reagent. As a control for transfection efficiency, 1ug of a plasmid encoding enhanced green fluorescent protein (GFP), [Clontech, Mountain View, CA] was added.

6.4.3. Cell Lysis and Purification of Flag Tagged Proteins

48 hours later, HEK293T Cells were dissociated using 0.25% Trypsin EDTA, and washed in DMEM and Phosphate Buffered Saline (PBS). Immunoprecipitation of Flag-tagged proteins was carried out using Sigma FlagM Purification kit. Briefly, 15 million HEK293T cells expressing exogenous flag-tagged TRIM22, and TRIM22-106 were suspended in 250ul of Sigma Cell Lytic-M ® with protease inhibitors (Sigma Complete Protease Inhibitors). Lysates were purified using affinity Sigma M2 FLAG antibody covalently attached to agarose resin. The TRIM22 was eluted from the resin using competition with 200ng/ul of Flag Peptide (Sigma F4799) to a final volume of 1ml in Tris ph7.4.

6.4.4. Dynamic Light Scattering

Dynamic light scattering (DLS) was used to determine the oligmeric state of TRIM22 and TRIM22-106. DLS determines the hydrodynamic radius (RH) distribution of samples through measurement of the decay rates of scattered light and calculation of translational diffusion coefficients. The hydrodynamic radii were obtained from the diffusion coefficients (D) via the Stokes–Einstein equation: RH = kBT 6D where is the viscosity of the solvent, T is the absolute temperature, and kB the Boltzmann constant. Dynamic light scattering experiments were

performed at 25 ◦C, using a Viscotek 802 DLS instrument (Viscotek Europe Ltd., Basingstoke, UK) using a 50 mW diode laser with wavelength 827.9 nm. 30ul of sample was used for each experiment. Samples were allowed to settle for 10 minutes due to absence of column purification. 10 scans of 10 s were obtained for each sample. Data were analysed using Omnisize 3.0 software (Viscotek Europe Ltd).

6.4.5. SDS PAGE

Cells were lysed using IGEPAL + Protease Inhibitors (Sigma Complete Protease Inhibitors) and nuclei separated. Lysate was added 1 in 2 in NuPage Sample buffer, with 10ug/ml of dithiothreitol (DTT), and boiled at 70°C for 10 minutes. Proteins were resolved using sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) on 4-12% NuPAGE® Bis-Tris Precast Gels at 120V for 60 minutes using Kaleidescope Protein ladder. Gels were then blotted onto a nitrocellulose membrane (Hybond C by Amersham) at 15V for 30 minutes. Blots were blocked in 50% Odyssey blocking buffer in PBS for 18 hours. Primary antibody staining was done using TRIM22 antibody (Sigma HPA00175 at a final concentration of 1:250 in 50% Odyssey blocking buffer) or actin primary (ab8227 at a final concentration of 1:5000) at room temperature for 1 hour and washed. Secondary antibody staining was carried out using LICOR IRD dye Anti-rabbit antibody at a final concentration of 1:15,000. Blots were read on the LICOR system. Washing steps were performed in triplicate with PBS 0.05% tween.

6.4.6. Native PAGE

HEK293T cells were lysed in IGEPAL Protease Inhibitors (Sigma). Sample was resuspended 1 in 2 in Novex Native PAGE Sample Buffer. Proteins were resolved on 3-12% acrylamide Novex Native precast gels in the absence of SDS or DTT, at 150V for 90 minutes alongside the native mark Unstained protein ladder. Gels were then blotted onto ImmunBlot PVDF membrane using the TransBlot Semi Dry System at 25V for 50 minutes. Blots were blocked in 50% Odyssey blocking buffer in PBS for 12 hours. Blots were stained using Flag antibody (Sigma M2 Flag at a final

concentration of 1:250 in 50% Odysstey blocking buffer) or actin primary (ab8227 at a final concentration of 1:5000.) The primary antibody staining was carried out at room temperature for 1 hour. Secondary antibody staining was carried out using LICOR IRD dye Anti-rabbit and Antimouse antibody at a final concentration of 1:15,000. Blots were read on the LICOR system. Washing steps were performed in triplicate in PBS 0.05% tween.

6.4.7. Preparation and Transduction Titre Determination of VSV-G Pseudotyped Viruses

400,000 HEK293T were plated and allowed to rest. 1ug of total plasmid encoding HIV-1, gag/pol, and VSV-G env, was transfected into cell using GeneJuice Transfection Reagent. 48 hours later, 2ml of supernatant was harvested, spun down at 2000rpm, and the supernatant filtered through a 0.45micron filter. The resulting pseudovirus was diluted in media with 2ug/ml of Polybrene reagent as shown.

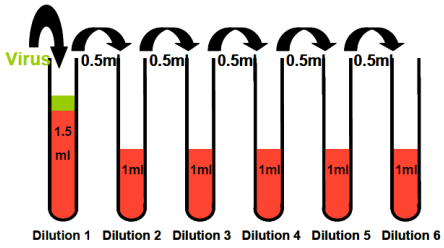


Figure 6.1 Dilution Series of Lentiviral Transductions

500,000 HEK293T cells were transduced with 0.5ml or diluent of resultant pseudovirus and incubated for 72 hours. This was repeated for each combination of viruses to encode HIV-2, SIVmac, SIVsm, and FIV, VSV-G pseudotyped virus. GFP positive cells were enumerated by FACS analysis and the titre of transducing units/ml was calculated: ((%GFP positive cells/100) x # cells transduced)/Volume of Virus(ml). Each reporter plasmid combination produced a varying number of transducing units, with HIV-1 capable of much higher transduction events than the other pseudoviruses. Experiments using these cell lines in a myriad of cell lines from human, feline, bovine, and non-human primates illustrate that these differences are intrinsic to the viruses rather than HEK293T cells differentially restricting the viruses(316).

Viruses were stored at -80 for downstream experiments at the following transducing units:

Table 6.3 Infectious Titres of VSV-G Psuedotyped Viruses

eGFP Expressing Virus	Infectious Units/ml
HIV-1	$2.1x10^6$
HIV-2	$1.5x10^5$
SIVmac239	$7.7x10^4$
SIVsm543 *	$8.1x10^3$
FIV *	$4.6x10^4$

*Transducing units of VSV-G pseudotyped viruses. *consistently low infectious titres*

6.4.8. VSVG-pseudotyped infection of U937 Permissive and Non-permissive cells

80,000 permissive or non-permissive U937 cells were seeded in a 48 well plate in triplicate and allowed to rest. 24 hours later, VSV-G psuedotyped HIV-1 and HIV-2 were serially diluted 10-fold, five times, and spinoculated into the cells in a flat bottom tube at 2000rpm at 32°C for 2 hours.

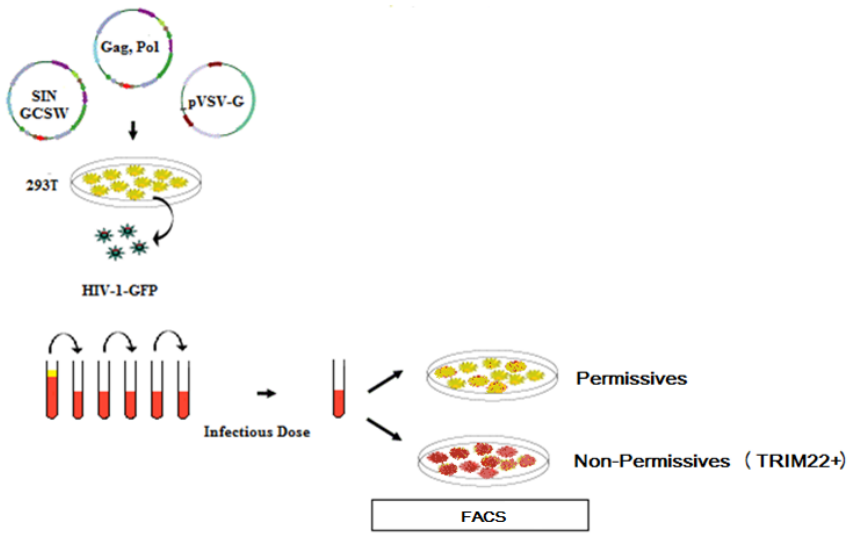


Figure 6.2 Flow diagram of Methods for VSV-G Psuedotped HIV Infection of U937 Clones

VSV-G-pseudotyped retrovirus was produced by 3 plasmid transfection of 293T cell and titrated to determine infectious dose. Permissive and Non-Permissve U937 cells, differentiated by the absence and presence of TRIM22 respectively, were spinoculated with VSV-G pseudotyped HIV-1 and HIV-2 and GFP positive cells enumerated on the flow cytometer.

6.4.9. Overexpression of TRIM22 and Subsequent Challenge with VSV-G pseudotyped Virus

400,000 HEK293T cells were plated and allowed to rest. 1ug of total plasmid DNA encoding TRIM22 or TRIM22 106 was transfected into cells using GeneJuice Transfection Reagent in

triplicate. In addition empty vector pgl3 and eGFP were used as controls for infection and transfection efficiency respectively. 24 hours later, cells were removed from plate, placed into a flat bottom tube, and transduced with $1x10^6$, $3.9x10^4$, $4.1x10^3$, and $2.4x10^4$ infectious units of VSV-G psuedotyped HIV-1, SIVmac, SIVsm, and FIV, respectively(317). Half the supernatant was removed, and fresh DMEM added, and cells incubated. 72 hours later cells were stained with live/dead Aqua fixable stain, fixed, and enumerated for presence of GFP. At least 1 million events were recorded.

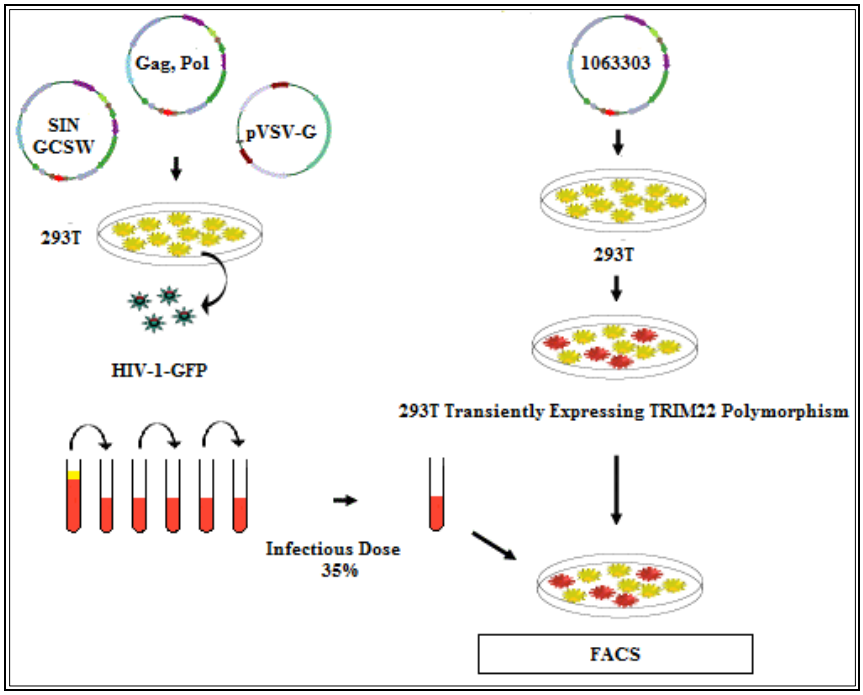


Figure 6.3 Flow Diagram of Methods for VSV-G Pseudotyped Infection of HEK293T cells

VSV-G-pseudotyped retrovirus was produced by 3 plasmid transfection of HEK293T cells. HEK293T cells are transduced with either TRIM22 wildtype or TRIM22-106, and challenged with VSV-G pseudotyped virus 24 hours later. 72 hours later, GFP expression is measured using flow cytometry.

6.4.10. Propagation of HIV1 and HIV2 on C8166 cells

2 million C8166 cells were added to a culture tube, and centrifuged at 350g. HIV-1-IIIB or HIV-2-CBL20 stock was titrated in media to a range a virus concentrations from 0.1 MOI to 0.001 MOI. The percentage infection was enumerated 3, 5, and 7 days later in the case of HIV-IIIB, and 14, 16, and 18 days later in the case of HIV-2 CBL-20, by fixation and permeabilisation of the cells, staining with Gag-specific antibody KC57 and Aquafluorescin L/D fixable stain, and acquisition on

the Cyan ADP Analyser. The HIV-IIIB titre selected for subsequent experiments infected 45% of unmodified cells. The HIV-2 CBL-20 titre selected for subsequent experiments infected 3% of unmodified cells. The appropriate dilution of virus was then used for spinoculation in 2 million C8166 pelleted cells by centrifugation at 2000rpm at 27°C for 2 hours. Following spinoculation, half of the supernatant was removed, and cells were resuspended in media to a final count of 1 million cells per ml.

6.4.11. Immunofluorescence

C8166 cells were plated on glass coverslips, washed with PBS and fixed with 4% paraformaldehyde in PBS before being permeabilised with 0.5% Triton in PBS for 10 min at room temperature, washed with PBS and then blocked with 5% Bovine Serum Albumin (BSA) in PBS containing 0.1% Tween for 1 h at room temperature. Cells were incubated for 1 hour with the primary TRIM22 monoclonal antibody (Abcam) at 1:250 dilution and HIV p24 FITC, washed three times with PBS and then incubated for 1h with a secondary antibody (Alexa-568 conjugated anti-mouse IgG (Invitrogen)) at 1:400 dilution. Coverslips were mounted onto glass slides using Vectashield mounting medium without DAPI (Vector Labs) and imaged using an Inverted Olympus FV1200 System. ImageJ was used to correct background fluorescence and analyse co-localisation using Pearson's correlation coefficient and fluorescence intensity under identical conditions.

6.4.12. Flow Cytometry and Intracellular Cytokine Staining

Flow cytometry was performed on a CyAN ADP Analyser (Beckman Coulter) FlowJo software (Treestar, Ashland, OR, USA) was used for data analysis. Compensation was performed using the BD anti-mouse compensation beads and the Arc reactive amine beads where appropriate. Cell viability was obtained using Aqua Live/Dead Fixable stain (Invitrogen.)

An example gating strategy for HEK293T cells expressing GFP is shown below.

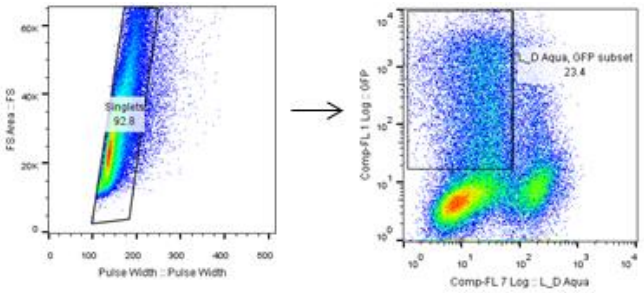


Figure 6.4 Gating Strategy for GFP expressing HEK293Tcells

(Left) HEK293T Doublets were excluded using Pulse Width versus forward scatter (FSC-A) (Right) Live and GFP positive cells were enumerated

An example gating strategy for C8166 cells infected with HIV-1-IIIB or HIV-2-CBL20 and subsequently stained for TRIM22-PE is shown below.

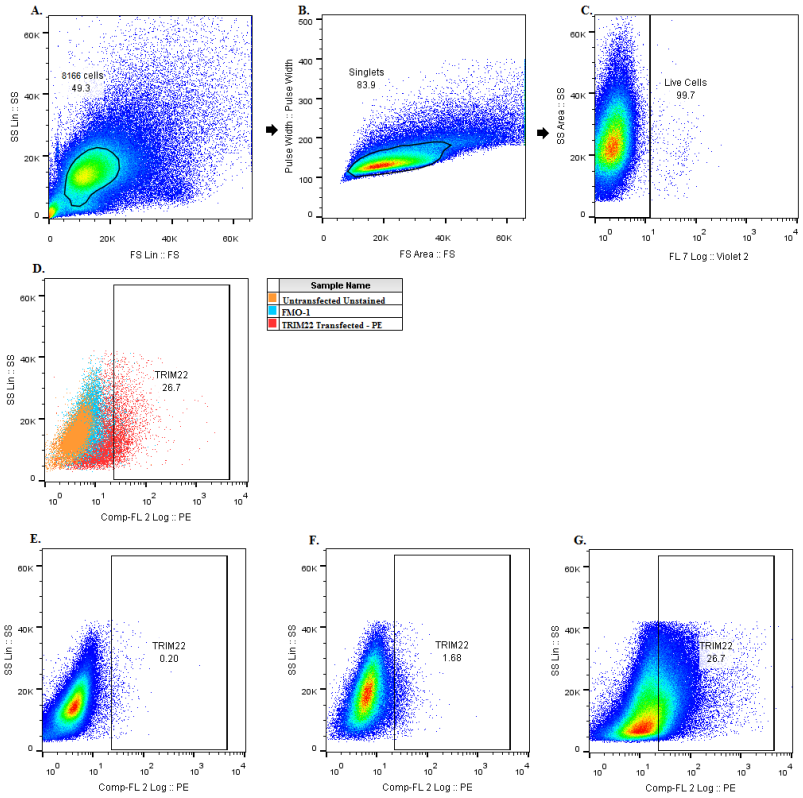


Figure 6.5 Gating Strategy for stained C8166 cells

A. *293Tcells were gated based on side scatter (SS-Lin) and forward scatter (FS-Lin).* ***B.*** *Doublets were excluded using Pulse Width versus forward scatter (FSC-A)* ***C.*** *The live cell population was identified using side scatter (SS-Lin) versus FL 7 Log Violet 2 which were the Live/Dead Fixable Stain Negative population.* ***D.*** *Fluorescence-minus-one (FMO) is shown illustrated to show the true PE positive population.* ***E.*** *Untransfected unstained cells.* ***F****. FMO-1 to identify nonspecific fluorescence in PE.* ***G.****Cells stained for transiently transfected TRIM22 conjugated to PE, showing the true positive population for PE.*

6.5. Results

6.5.1. Preliminary Experiments

6.5.1.1. Determining the Endogenous Expression of TRIM22 in HEK293T, U937 Permissive and Non- Permissive Clones to Facilitate Downstream Experiments

In order to perform studies on the structure or restrictive capacity of TRIM22, it was first necessary to characterise cell lines devoid of endogenous TRIM22. This allows the cell lines to be used as systems from which to isolate exogenously overexpressed protein, and measure the effect of exogenous overexpression on retroviral restriction. It was also necessary to characterise comparable cell lines with endogenous TRIM22 expression for comparison.

Firstly, determining whether they possess TRIM22 mutations rs1063303 and rs7935564 (associated with outcome to HIV-2 infection in Chapter IV) was essential information, as these may serve to affect the restrictive capacity of the cells to HIV. DNA was isolated from permissive and non-permissive U937s, and HEK293Ts using alcohol precipitation. Exons 2-4 of TRIM22 were amplified from the isolated DNA using methods previously described in Chapter IV, and Sanger sequenced.

Table 6.4 TRIM22 Genotypes in Cell Lines

Cell Line	TRIM22 Gene Present/Absent	rs1063303 (G to C)	rs7935564 (G to A)
U937 Permissive	Present	G/C (het)	G/A (het)
U973 Non-permissive	Present	G/C (het)	G/A (het)
HEK293T	Present	G/G (wildtype)	G/A (het)

DNA was precipitated from HEK293T cells, Permissive and Non-permissive U937 cells, and Exons 2-4 of TRIM22 were successfully amplified from all three lines. Exons 3-4 of each cell line were then sequenced, and typed for the presence of the rs1063303 and rs7935564 mutants. Het=heterozygous

The TRIM22 gene was present in all three cell lines. The 1063303 (R242T) mutation was present as a heterozygote in both U937 Permissive and non-permissive cells, and absent in HEK293T cells. The 7935564 mutation was present as a heterozygote in U937s permissive, non-permissive, and HEK293T cells.

It was then essential to determine whether TRIM22 was present at the protein level in each cell line. Lysate was obtained from each cell line and the nuclear compartment separated. The cytoplasmic compartment was subjected to SDS-reducing PAGE and Western Blotting. Blots were stained for TRIM22 and β-Actin. It was observed that TRIM22 could not be detected in the cytoplasmic compartment of the non-permissive cells and was absent from the permissive and HEK293T cells. In order to clarify the results of downstream experiments where these cells would be infected, it was essential to determine whether the expression of TRIM22 would be altered upon infection or similar stimulation. Further, it was then important to clarify whether this effect was compartment dependent. 1 million cells from each of the three lines were stimulated with 10,000U of β-IFN in triplicate. The cells were then lysed at 24, 36, or 48 hours post stimulation, and the lysates subjected to SDS-reducing PAGE and Western Blotting. Blots were stained for TRIM22 along with Heatshock protein 70 (HSP70) and Histone H3 to determine cytoplasmic and nuclear localisation respectively.

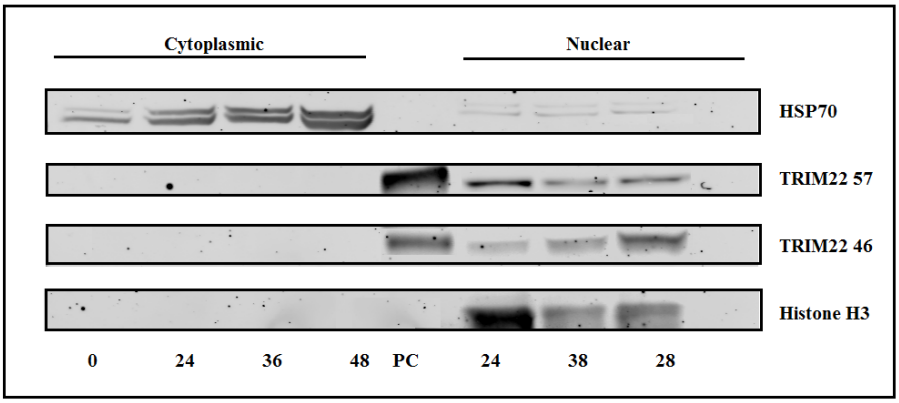


Figure 6.6 Lysate of Cytoplasmic and Nuclear Compartments from HEK293T Cells

HEK293T cells were stimulated with 10,000U of β-IFN and harvested at 24, 36, and 48 hours. The expression of TRIM22 was sensitive to β-IFN. HSP70= Heatschock Protein 70 (cytoplasmic marker), TRIM22 57=full length TRIM22, TRIM22 46= N-terminal cleaved TRIM22.

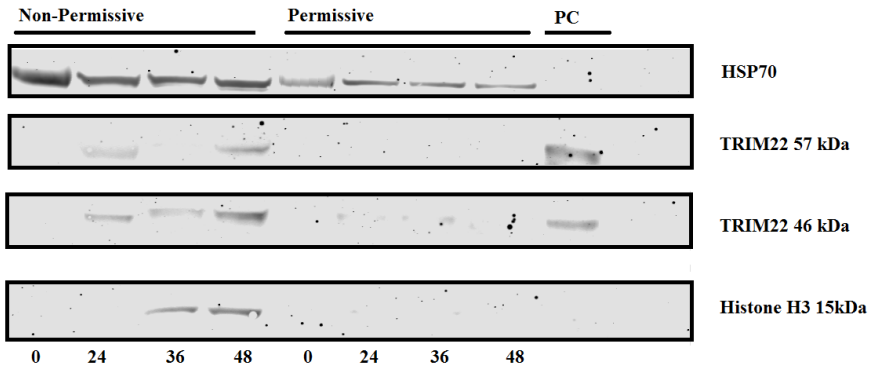


Figure 6.7 Lysate of Cytoplasmic Compartments from Permissive and Non-Permissive U937 cells

Permissive and Non-permissive U937 Cells were stimulated with 10,000U of β-IFN and harvested at 24, 36, and 48 hours. The expression of TRIM22 was sensitive to β-IFN in non-permissive cell cytoplasm. HSP70= Heatschock Protein 70 (cytoplasmic marker), TRIM22 57=full length TRIM22, TRIM22 46= N-terminal cleaved TRIM22.

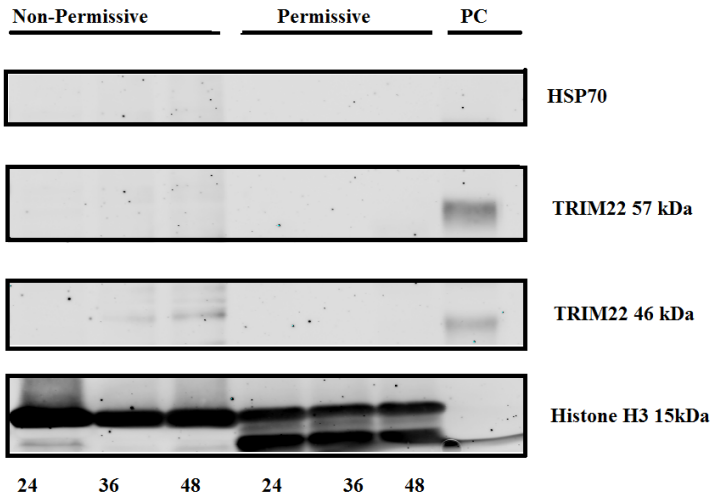


Figure 6.8 Lysate of Nuclear Compartments from Cell Lines (Stimulated with IFNβ)

Permissive and Non-permissive U937 Cells were stimulated with 10,000U of β-IFN and harvested at 24, 36, and 48 hours. The expression of TRIM22 was sensitive to β-IFN in the nucleus non-permissive cells. HSP70= Heatschock Protein 70 (cytoplasmic marker), TRIM22 57=full length TRIM22, TRIM22 46= N-terminal cleaved TRIM22.

Permissive cells possessed nominal TRIM22 following stimulation in any time point in either nuclear or cytoplasmic compartments. Non-permissive cells possessed TRIM22 in the cytoplasmic compartment which was sensitive to stimulation. A 46kDa version of TRIM22 was also observed in the nuclear compartment of the non-permissive cells. The 46kDa isoform is believed to have been cleaved at the N-terminus resulting in a truncated protein. HEK293Ts

possessed no cytoplasmic expression of TRIM22, however the protein was observed in the nuclear fraction at all IFN-β stimulated time points.

In light of the data, HEK293Ts were chosen for downstream experiments to overexpress and isolate exogenous TRIM22 and TRIM22-106. HEK293Ts were also used to determine the effect of exogenous TRIM22 expression and the dissection of the effect of polymorphism 1063303 given that the endogenous gene expressed was wildtype at this position. Non-permissive cells were chosen to determine the ability of endogenous TRIM22 to restrict VSV-G psuedotyped virus infection, and permissive cells selected as controls for this experiment.

6.5.1.2. Purification of Flag-Tagged TRIM22 and TRIM22-106

In order to produce purified TRIM22 and TRIM22-106 for downstream structural analysis, 15million HEK293T cells overexpressing either Flag-TRIM22 or Flag-TRIM22-106 were lysed, and the lysate was applied to the Sigma Flag Purification Column. The column was washed twice, and bound proteins eluted via competition binding. The first fraction, wash, and final eluate were collected and subjected to SDS- PAGE in the presence of 10mM DTT, and the gel was western blotted. The blot was stained using anti-TRIM22 polyclonal antibody.

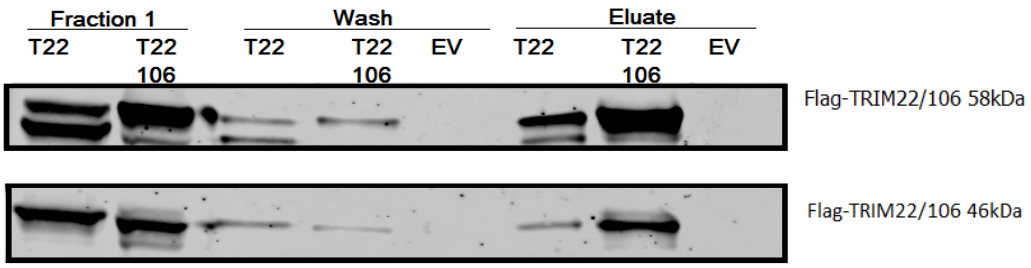


Figure 6.9 Flag Purification of Exogenous TRIM22 and TRIM22-106 From HEK293T Cells

Exogenous Flag-TRIM22/TRIM22-106 was purified from HEK293T cells in a through a column. EV= Empty Vector.

Blots indicated that both Flag-TRIM22 and Flag-TRIM22-106 could be purified through processing through the column. A 58 kDa monomer was observed for both TRIM22 and TRIM22-106. A band of 46kDa was also observed in the fraction: N-terminal degraded Flag-TRIM22/TRIM22 106. Both showed a 56kDa Monomer, which appears to be showing cleavage of

the Flag tag during electrophoresis. A Western Blot (Appendix VII) of the same components stained using B-Actin and Flag Antibody shows only the 58kDa and 46kDa Monomer). The 58kDa monomer of TRIM22 106 demonstrated greater expression than in TRIM22.

6.5.2. Structural Studies of TRIM22 and TRIM22-106

6.5.2.1. Dynamic Light Scattering Reveals Differential RH between Flag-TRIM22-106 and Flag-TRIM22

The flag-purified eluate of HEK293T cells transfected with Flag-TRIM22 or Flag-TRIM22-106 were subjected to 10 scans of 10s Dynamic Light Scattering. The hydrodynamic radius (RH) distribution of Flag-TRIM22 and Flag-TRIM22-106 were overlaid.

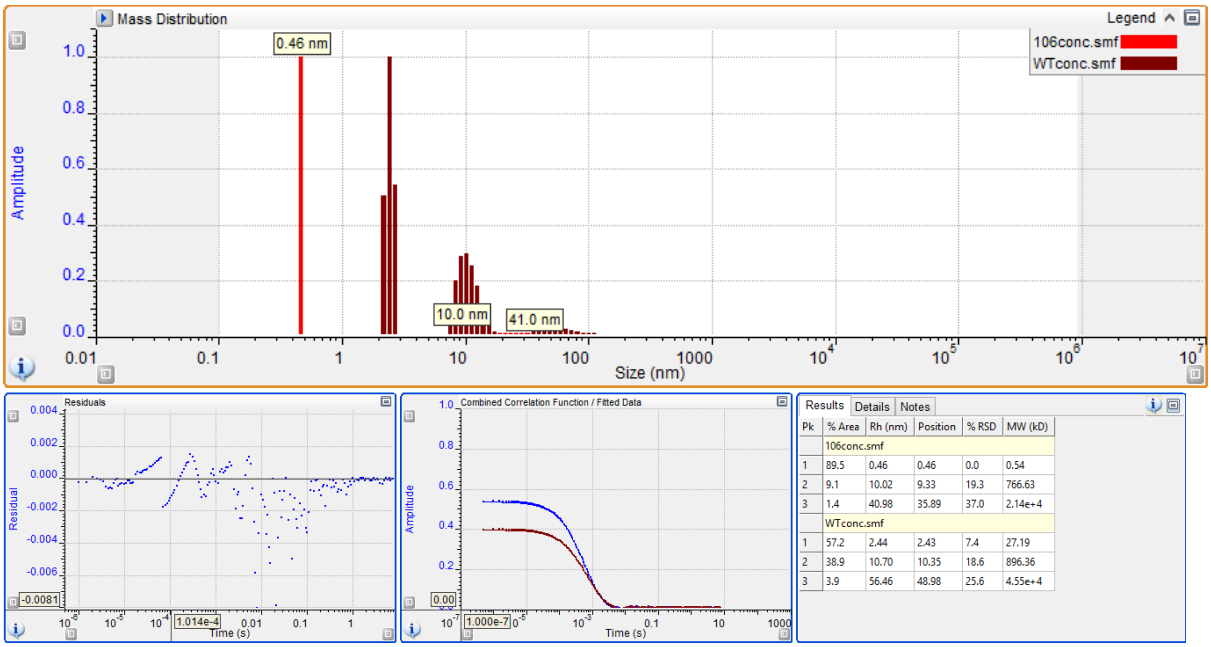


Figure 6.10 Dynamic Light Scattering of TRIM22 and TRIM22-106
(Both proteins were stored under identical conditions)

TRIM22 showed a major peak (57.% by weight) at RH of 2.42nm that probably corresponded to dimeric TRIM22 and the absence of monomeric TRIM22. Aggregates were also visible at an RH of 10.70 and 56.46 corresponding to 38.9% and 3.9% by weight. The profile of TRIM22 106 was similar in this regard, showing aggregates at RH of 10.02 and 40.98 at a weight of 9.1% and 1.4% respectively. However, the major peak at RH of 2.42nm seen in the TRIM22 wildtype was deficient, and the major peak was instead visible at an RH 0.46nm with a weight of 89.5%. This

could be due to the high precipitation of the TRIM22-106 monomer in the solution which is absent in the TRIM22 wildtype.

6.5.2.2. Confocal Microscopy Shows TRIM22 Sensitivity to, and Co-localisation with, HIV-1

It was previously shown that TRIM22 expression is sensitive to HIV-1 infection in monocyte-derived macrophages(227), and that TRIM22 interacts with the HIV-1 capsid in imaging and immunoprecipitation assays(224). To determine the extent to which the HIV-1 capsid protein interacts with TRIM22 *in vitro*, C8166 cells (a CD4 T-cell line expressing CXCR4 and CCR5) were infected with 0.1 MOI of HIV-IIIB. After 72 hours cells were blocked, stained for HIV-p24 (KC57) conjugated to FITC (green) and TRIM22 with a secondary antibody conjugated to AF568 (red). The cells were mounted on slides, imaged using confocal microscopy, and analysed for pixel intensity and colocalisation in the ImageJ software platform. When compared with uninfected cells TRIM22 was significantly upregulated in the presence of HIV-IIIB infection, and strongly colocalised with HIV-1 p24 (Colocalisation Coefficient: 0.881; Pearson's Coefficient 0.6931; Uninfected T22 Average Pixel Intensity- 0.661, Infected TRIM22 Average Pixel Intensity- 3.160**). This corresponded with upregulation flow cytometry data measuring a 4 fold change in TRIM22 expression at peak HIV-1 infection.

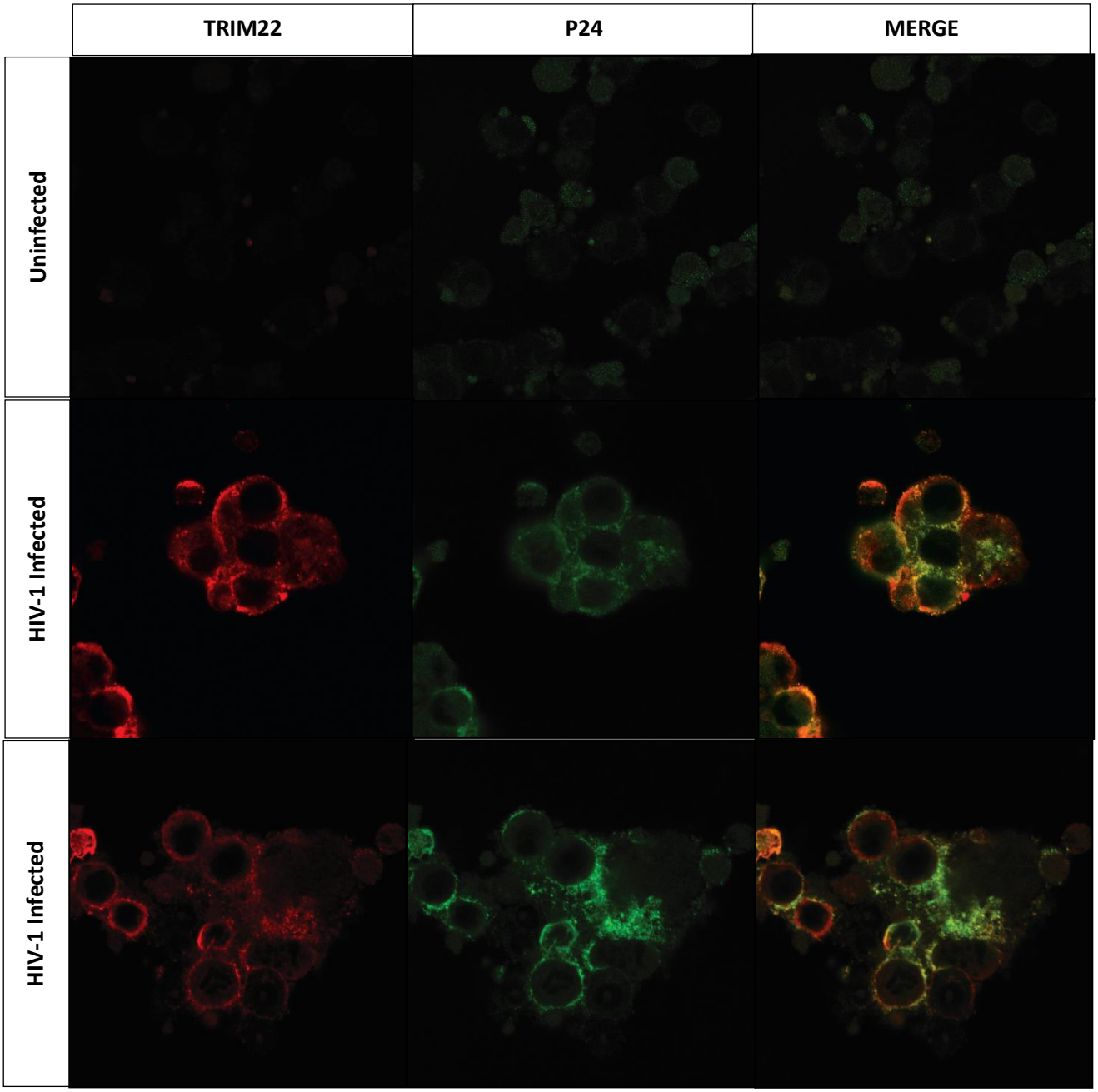


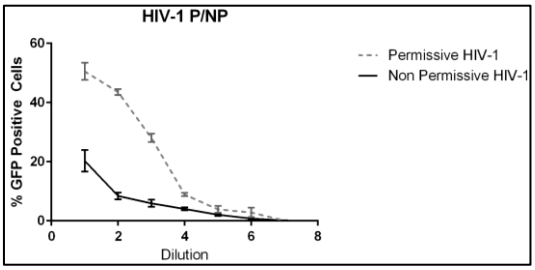
Figure 6.11 TRIM22 Co-localises with HIV-1 Capsid p24

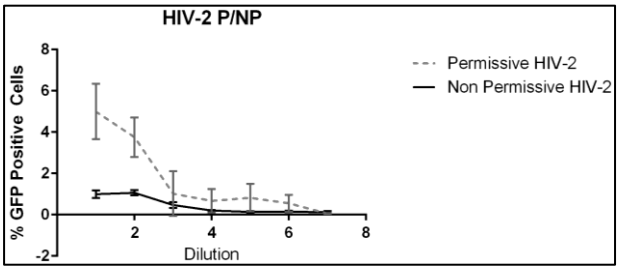
TRIM22 Confocal immunofluorescence microscopy revealed that TRIM22 (red) is substantially upregulated in C8166 cells following 72 hours infection with HIV-IIIB. The Average Pixel Intensity of the TRIM22 staining is 0.661 in uninfected cells, whereas the average pixel intensity of TRIM22 is 3.160 in HIV-IIIB infected cells. TRIM22 co-localises with the HIV-1 capsid protein p24 (green) at a co-localisation coefficient of 0.881. Pearson's Correlation 0.6931*

6.5.3. Functional Studies of TRIM22

6.5.3.1. Determining the Restrictive Capacity of Endogenous TRIM22

This experiment aimed to determine whether the endogenous TRIM22 expressed in non-permissive U937 cells could restrict VSVG-psuedotyped HIV-1 and HIV-2, in comparison to permissive U973 cells which were devoid of endogenous TRIM22, but otherwise transcriptionally identical to their non-permissive counterparts. To do this, non-permissive (TRIM22+) and permissive (TRIM22-) were spincoluated with 1050000 and 75000 infectious units of VSV-G psuedotyped HIV-1 and HIV-2 respectively, in triplicate. Infections with 10 fold dilutions of each virus were also performed. After 72 hours, the GFP positive cells were enumerated on FACS. Permissive cells, devoid of endogenous TRIM22, showed greater replication of both VSV-G psuedotyped HIV-1 and HIV-2 than their non-permissive counterparts.





Dilution	Permissive HIV-1			Non Permissive HIV-1		
	Trial #1	Trial #2	Trial #3	Trial #1	Trial #2	Trial #3
100000	47.300	51.400	52.900	16.400	20.700	23.700
10000	44.600	43.400	42.600	7.270	8.490	9.550
100	29.200	26.500	28.400	4.690	5.990	7.160
10	9.560	8.660	8.460	3.500	4.060	4.510
1	3.420	2.970	5.170	1.600	2.190	2.460
0.1	1.910	1.910	4.680	0.680	0.840	0.810
0	0.042	0.046	0.081	0.076	0.120	0.170

Dilution	Permissive HIV-2			Non Permissive HIV-2		
	Trial #1	Trial #2	Trial #3	Trial #1	Trial #2	Trial #3
100000	5.030	3.630	6.320	0.830	0.950	1.200
10000	4.020	2.690	4.550	0.960	1.030	1.200
100	0.570	0.230	2.260	0.430	0.350	0.630
10	0.420	0.270	1.320	0.230	0.180	0.200
1	0.560	0.320	1.580	0.120	0.120	0.180
0.1	0.360	0.280	1.030	0.160	0.120	0.190
0	0.042	0.046	0.081	0.076	0.120	0.170

Figure 6.12 Permissive and Non-Permissive Cells Infected with HIV-1 and HIV-2

Permissive U937s cells lacking functional TRIM22 show several fold higher HIV-1 (left) and HIV-2(right) replication than their non-permissive (TRIM22+) counterparts, as evidenced by %cells infected with VSV-G pseudotyped virus. Lines shows mean and standard error of biological triplicates. HIV-1/2 permissive and non-permissive slope fitted with linear regression, and tested for covariance (ANCOVA). HIV-1P/NP: p=0.0001. HIV-2 P/NP: p=0.0001

6.5.3.2. Determining the Restrictive Capacity of Exogenous TRIM22 and TRIM22-106

In order to determine the restrictive capacity of exogenous TRIM22 and the mutant TRIM22-106, 1ug of either Flag-TRIM22, Flag-TRIM22-106, an empty vector, or eGFP, was transfected

into 400,000 HEK293T cells in triplicate. 24 hours later, the triplicates were transduced with $1x10^6$, $3.9x10^4$, $4.1x10^3$, and $2.4x10^4$ infectious units of VSV-G psuedtyped HIV-1, SIVmac, SIVsm, and FIV, respectively, and incubated at 37°C. 72 hours later, cells were stained with Aquafluorescin Live/Dead Stain, fixed, and GFP expression was enumerated.

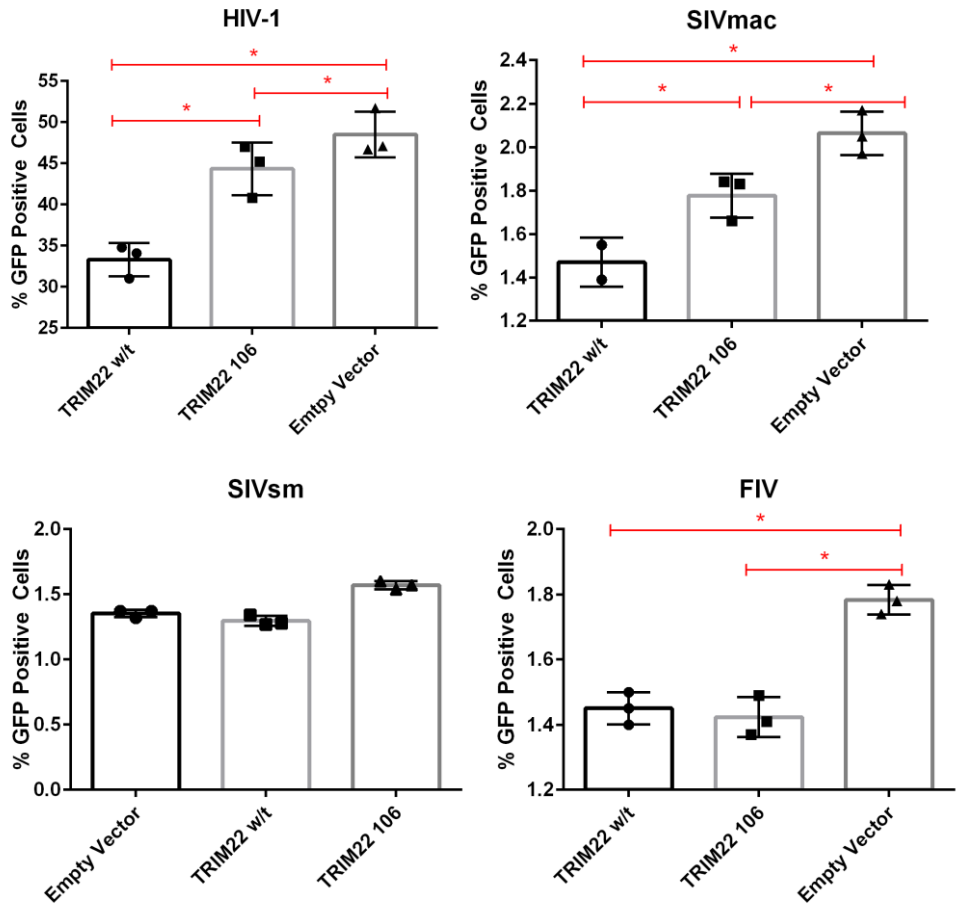


Figure 6.13 Examination of the Restrictive Capacity of TRIM22 Against VSV-G Pseudotyped Lentiviruses

*Exogenous expression of TRIM22 in HEK293T cells leads to significant restriction of HIV-1, SIVmac, and FIV. Though FIV showed significant in the t-tests, the infectious titres are very low in cultures with TRIM22 expressed, and this may be a spurious finding. The mutant TRIM22-106 is less efficient at restriction in the case of HIV-1 and SIVmac, but not FIV. (Biological Repeats) Unpaired t-tests. *p<0.05*

The overexpression of TRIM22 in HEK293T cells led to potent restriction of VSV-G psuedotyped HIV-1, SIVmac, and FIV (though the infectious titres were very low), and not SIVsm infection, when compared to cells which expressed the empty vector only. When compared to TRIM22, the overexpression of TRIM22-106 lead to significantly less restriction of VSV-G pseudotyped

virus infection in the case of HIV-1 and SIVmac. Restrictive capacity of FIV was unaltered by the presence of the TRIM22-106.

6.6. Discussion

Chapters III, IV, and V provided evidence that polymorphisms in TRIM5α and TRIM22 were associated with differential outcome to HIV-2 infection, and that there was phylogenetic and structural evidence to support the findings. This chapter pilots an extension to this work, by attempting to test the findings from population based genetics, structurally, and functionally, *in vitro.* This chapter aimed to test whether TRIM22 could restrict HIV-1 and other lentiviruses, to determine whether a deleterious mutation, rs1063303 (R242T) could diminish this effect as observed in population studies, and to facilitate downstream studies to determine how.

The centrality of the E3 ligase activity, and PRYSPRY domain, to the restrictive capacity of TRIM22 has already been documented(224, 225). The centrality of multimerisation has not been investigated. The coiled coil region of TRIM proteins are essential for mediating higher order assembly, a mechanism which has been reported to be essential in virus restriction for TRIM5α. In Chapter IV, a polymorphism of the TRIM22 coiled coil domain, rs1063303 or R242T, was found to predict decreased CD4+ T-cell count, and survival in HIV-2 infected individuals within the Caio Community Cohort. The first aim of the chapter and future work is to clarify differences in structure between TRIM22 and the TRIM22-106 to determine if there was a structural reason, related to oligomerisation, why the two would behave differently.

To facilitate the pilot, it was necessary to find systems devoid of endogenous TRIM22, and systems with endogenous TRIM22 that performed as suitable comparisons. Firstly, cell lines were investigated for the presence and absence of endogenous TRIM22. HEK293Ts were found to be deficient of TRIM22 in the cytoplasm but not the nucleus upon stimulation with IFN-β. In addition the permissive U937 cells were also devoid of TRIM22 protein expression in both

cytoplasm and nucleus, whereas the non-permissive U937 clones had TRIM22 in both compartments.

TRIM22 has been shown to be highly upregulated in primary monocyte-derived macrophages in response to HIV-1 infection, IFNα treatment, or stimulation with lipopolysaccharide (LPS)(227). To determine this in a lymphoid cell whine which best mimics the natural course of HIV infection, a highly infectable CD4+T cell line, C8166, was infected with a potent strains of HIV-1 IIIB. A optimisation time course was carried out to measure the infectious dose required for maximum infection and cell viability, and also the peak HIV infection as measured by p24 staining by flow cytometry. It was determined that peak HIV-1 infection occurred 72 hours post infection. Confocal microscopy showed that TRIM22 was upregulated around 5 fold at peak HIV-1 infection from two independent repeats in C8166 cells:

a. 4.78 fold change in average pixel intensity observed by confocal microscopy
b. 5.42 fold change in median fluorescence intensity observed by FACS (Appendix VI)

and that TRIM22 and HIV-1 p24 colocalised at a coefficient of 0.881, (Pearson's Correlation 0.6931).

HEK293Ts were used as a system for overexpression of exogenous TRIM22. Flag-TRIM22 and Flag-TRIM22-106 were purified from this system, and subjected to Dynamic Light Scattering and Size Exclusion Chromatography (Appendix VII). Dynamic Light Scattering aims to separate proteins on the basis of their hydrodynamic radius, and is typically speculative and requires highly purified samples to produce somewhat accurate findings. As these samples were not highly purified, and instead allowed to settle, both fractions showed minor populations of 10.02 and 40.98 which are likely to be cellular debris due to the impurities within the sample. Flag-TRIM22-106 showed a 5.3 fold reduction in radius of the aggregate observed at the major population, consistent with findings of an additional monomeric fraction present in the Flag-TRIM22-106 which is not present in the Flag-TRIM22.

The second aim of the chapter was functionally to further the results of the *in vivo* studies, by firstly investigating whether TRIM22 on its own could restrict HIV-2, and secondly, mediate restriction of other lentiviruses, and finally, if so, whether the 1063303 mutation diminished TRIM22's restrictive capacity. This was investigated by examining the ability of U937s which had lost the expression of TRIM22 over time (whilst maintaining steady levels of APOBEC3-A, G and F, tetherin TRIM5α, TRIM19, TRIM37), to restrict VSV-G pseudotyped HIV-1 and HIV-2 virus infection. This was also investigated by overexpressing TRIM22 and TRIM22-106 in endogenously deficient HEK293T cells, and subsequently challenging them with VSV-G psuedotyped HIV-1 and other lentiviruses of interest such as SIVmac and SIVsm (for their phylogenetic links to HIV-1 and HIV-2), and FIV given that felines discordantly express genes for TRIM22 and not TRIM5α.

The functional studies show that endogenous TRIM22 is able to restrict VSV-G pseudotyped lentivirus infection both in the monocytic U937 cell line, and when expressed in endogenously deficient HEK293T cells. TRIM22 showed potent restriction of HIV-1 and HIV-2 in U937 non-permissive clones in comparison to clones devoid of TRIM22. Overexpression of Flag-TRIM22 in HEK293T cells lead the restriction of HIV-1, SIVmac, and FIV whereas no significant restriction was observed for SIVsm which may have been a facet of low infectious titres. Flag-TRIM22-106 mediated restriction of VSV-G HIV-1, SIVmac, was significantly less than that of the Flag-TRIM22 wildtype. The Flag-TRIM22-106 restriction of FIV was comparable to that of Flag-TRIM22. However, the infectious titres of VSV-G FIV were reproducibly low. This resulted in less than 2% of cells being infected in the control experiment. Though the restriction of FIV was reproduced in three separate experiments, the low infectious titres make this result potentially spurious. The results clearly demonstrate that TRIM22 wildtype is capable of restricting VSV-G HIV-1 and SIVmac, and that the R242T mutation makes TRIM22 less efficient at restriction. Further, when compared to the actin probe, the Flag-TRIM22-106 was better expressed than Flag-TRIM22 which indicates that the detrimental effects of this allele may even be slightly underestimated. It

would be beneficial to investigate the restrictive capacity of TRIM22-106 for HIV-2 in another cell line, or using virus isolate, to determine whether this increased replication is reproducible. A better way to clarify this would be to investigate restriction in a dose dependent manner, for example, transfecting increasing amounts of TRIM22 into a deficient system with subsequent FIV challenge, or knocking down TRIM22 in CRFK cells.

What does the data tell us about how TRIM22's restriction is mediated?

It has been demonstrated that the restriction of HIV-1 by RhTRIM5α is dependent on its self-association, and multimiersation to form a hexamer of dimers(202, 203). In the absence of such oligomerisation, HIV-1 restriction by TRIM5α is rendered ineffective. In light of DLS data on the oligomeric state of Flag-TRIM22 and Flag-TRIM22-106, it would be expected that the R242T mutation within the coiled coil renders TRIM22 less able to dimerise. If the data from studies on TRIM5α demonstrating the centrality of lattice formation to HIV restriction are directly applicable to TRIM22, it is possible that the 1063303 mutation disrupts the formation of the disulphide bonds required between the coiled coil domains of two monomers and therefore prevents oligomer mediated restriction. This complements the functional data, where TRIM22 restricted (HIV-1, HIV-2, SIVmac,) but the TRIM22-106 mutation rendered it less effective at doing so in every case except FIV. A possible mechanism for deficient restriction of the TRIM22-106 could therefore be the disruption of higher order assembly of TRIM22.

This chapter provides preliminary evidence supporting an investigation into TRIM22's restriction of various lentiviruses in varied cell lines, which include myeloid and lymphoid lineages. It also prompts further investigation of the structure of TRIM22 and TRIM22 with mutations such as R242T.

Chapter 7. Discussion and Conclusion

7.1. Discussion

Despite over thirty years of research into HIV; the mechanisms surrounding disease progression remain poorly understood. HIV-1 and HIV-2 infected people are eight, and two times, respectively more likely to die when compared to their uninfected counterparts. This likelihood is greatly reduced by the provision of ART, however only half of HIV infected individuals receive treatment do date, and a number of these drugs remain ineffective as HIV-2 treatment. Further the cumulative toxicity of long-term ART poses significant challenges, where tradeoffs are being made between ART-induced morbidity and HIV mortality. New insights into HIV control remain essential in the future's fight against HIV. This thesis is concerned with disease progression in HIV-2 infection; where the majority of individuals exhibit a long term non progressor phenotype, and roughly 20% progress to AIDS in a manner identical to that of HIV-1 infected people. Examining the mechanisms behind the halted progression in the majority of HIV-2 infected people may provide inference for slowing the progression of HIV-1. Further clarifying the mechanisms that distinguish HIV-2 progressors from controllers can provide insight into novel approaches to control of HIV. It is likely that the control of HIV-2 within subsets of the infected population is due to a combination of host and viral factors. The fact that HIV-2 and HIV-1 exhibit differing ranges of viral loads, CD4+ T cell counts, and markers of immune activation even at the point of death- provides credence for a virological component to disease control.

The Caio Cohort represents a rare, ancestrally, and virologically, homogenous population of HIV-2 infected people, and age and sex matched controls, with the highest recorded rates of HIV-2 infection in any community. Examining the mechanisms that distinguish progressors from controllers in this population is key. The HIV-2 virus has been well characterised in this cohort; where studies show the presence of a recent common ancestor in both controllers and progressors in both *gag* and *env* phylogenies(43). This suggests that pathogenicity has a strong host component within this cohort, and yet, only one genetic study has been performed to date

in this community which investigated compound HLA/KIR genoptyes. Host restriction factors have been well researched in the context of HIV-1. The closely related counterparts of host restriction factors from other species show marked restriction of HIV infection, showing that small changes in the genetic structure of these factors can modulate their viral specificity and the magnitude of restriction.

Chapter III illustrates an association study between TRIM5α polymorphisms and HIV-2 disease progression. Preliminary studies showed that female sex and undetectable viral load predicted survival in this subset of the Caio Community Cohort. Eight polymorphisms, synonymous and non-synonymous were observed in this subset. When compared to allele frequencies from neighbouring ethnic groups, namely Yoruba, and Africa-At-Large (which comprises Africa and African descendants worldwide) L159L, D249G, P479L were significantly enriched in the Caio Cohort. However this enrichment counted for little in terms of disease progression. P479L was enriched in HIV-2 infected patients and progressors, and correlated with significantly lower CD4+ T-cell count in t-tests and regression models, but significance was not retained following adjustment for multiple comparisons. Further R136Q, was associated with depressed viral load, but this too was not retained following adjustment. However this SNP had a statistically significantly positive impact on survival at the 10% level following adjustment. P479L has not been reported as affecting HIV acquisition, disease progression, or survival in any previous study. However, three previous studies have indicated a protective effect of the R136Q mutation. It was enriched in HIV-1 resistant but exposed sex workers and associated with decreased likelihood of seroconversion, showed a protective effect against CXCR4 mediated HIV-1 infection, and was elevated in uninfected participants with anti-HIV-1 activity in functional studies. This piece of research constitutes the fourth study suggesting a protective effect of the R136Q mutation. As power calculations for TRIM5α indicated the need for a larger cohort to generate more robust associations, given the small sample size, the association data involving

R136Q, viral load and survival give rise to a genuine signal which warrants reporting. It would be interesting to clarify the functional effect of P479L and R136Q in studies involving HIV-2 challenge *in vitro*.

Chapter IV contained the first exploration of TRIM22 polymorphisms in relation to HIV-2, and is also the first study to examine the effect of TRIM22 polymorphisms on survival. In comparison to the 8 SNPs found in TRIM5α, 21 polymorphisms in TRIM22 were observed in this cohort; 11 of which were non-synonymous and 4 of which had not been previously reported. Allele frequencies of previously reported SNPs were normal for the geographic region. Though highly polymorphic, only two polymorphisms were significantly associated with any differential outcome to infection. Both located in the coiled coil, rs7935564 (D155N) and rs1063303 (R242T) were associated with disease control and disease progression respectively. D155N was associated with a 50% reduction in progression to death. Rs1063303 was enriched in HIV-2 infected participants, associated with a 2.5 fold increased likelihood of being HIV-2 infected, a stepwise reduction in absolute CD4+ T-cell count with each substitution to the minor allele, and twice the rate of progression to death following adjustment for gender. The presence of both D155N and R242T has been shown to exert a negative impact on HIV-1 progression in studies conducted in Italy. This data suggests a dominant effect of the R242T mutation. Functional studies have shown the R242T mutation to be highly expressed and detrimental to TRIM22's restriction. These data provided a basis for a functional examination of the effect of TRIM22 on HIV-2 and other lentiviruses, and an assessment of whether the R242T mutation would diminish the effect of restriction, as was suggested by the population genetics studies. Further the imbalance of polymorphisms in TRIM22 when compared to TRIM5α, two closely related proteins, prompted further characterisation of these polymorphisms in relation to each other, and also clarification as to whether any of the SNPs had a deleterious effect on protein structure.

Chapter V served as an extension to the population genetics work described in the preceding two chapters, to analyse the effect of TRIM5α and TRIM22 genetic variation on protein evolution, structure, and function. Because of the discordant evolution of TRIM5α and TRIM22, exploring polymorphisms within the two proteins, in an ancestrally homogenous group exposed to a common set of retroviral selection pressures, had the potential to provide important insights on the protein fitness and restrictive capacity of the two proteins. Such an analysis has not been performed for TRIM proteins to date. *In silico* algorithms based on phylogeny, sequence, and structure predicted G31S, L214F, and P479L as variants likely to have a functional impact on TRIM5α. They also predicted D282N, R321K, and D360Y (all in and around the PRYSPRY domain) as likely to have a functional impact on TRIM22. Despite only 56% amino acid identity, the two possessed the same predicted secondary and tertiary structure, with the exception of a truncation in the coiled coil of TRIM22 and the presence of only 1 B-Box domain where TRIM5α has two. There were 50% more nsSNPs in TRIM22 than in TRIM5α from the same population, and while there was no spatial overlap between them, nsSNPs in TRIM22 appeared spatially relevant to regions in the rhesus TRIM5α PRYSPRY domain that are known to be responsible for protein fitness, strength, and specificity of restriction. Further, variants in the coiled coil mapped onto hinge regions that had the potential to alter protein conformation through the coiled coil- narrowing the cleft between the RING and PRYSPRY domains, and potentially affecting oligomerisation. In relation to the TRIM5α and TRIM22 variants associated with disease progression in Chapter III and IV: only P479L was predicted to have a functional impact, but was not evolutionarily conserved. R136Q showed a notable change in protein conformation in 3D models, but appeared irrelevant in the algorithms on a whole. 3D models of TRIM22 showed mutations D155N and R242T to be capable of exerting profound changes on protein conformation, narrowing the distance between the PRYSPRY and RING domains, and Consurf analysis showed that both were highly conserved and exposed residues. Taken together the data suggested that when compared to TRIM5α, variants in TRIM22 were more frequent,

and more relevant to retroviral restriction. Should a virus be exerting selective pressure on TRIM22 it would likely be HTLV-1, which studies assert has existed in West Africa for thousands of years. In light of the clinical data on TRIM22 illustrating that R242T was associated with poor outcome following HIV-2 exposure and infection, and the profound changes exerted by R242T to the 3D model of TRIM22, the protrusion of the residue, and its evolutionary conservation; the restrictive capacity of TRIM22 and the ability of the R242T mutation to augment this effect was explored in the next chapter.

Chapter VI provided the first evidence of the restrictive capacity of TRIM22 for lentiviruses other than HIV-1. Data showed that TRIM22 was sensitive to HIV-1 infection in C8166 cells (expressing CXCR4 and CCR5) and that TRIM22 co-localised with the capsid protein p24. Permissive and non-permissive U937 cells, which are differentiated by the absence of, and possession of, endogenous TRIM22 respectively, were challenged with VSV-G pseudotyped HIV-1 and HIV-2. There was singificantly greater HIV-1 and HIV-2 replication in permissive when compared to non-permissive clones. Following the expression of exogenous TRIM22 in HEK293T cells (devoid of endogenous TRIM22 expression in the cytoplasm) it was observed that TRIM22 restricted VSV-G pseudotyped HIV-1, SIVmac, and potentially FIV infection. When TRIM22-106 was expressed in HEK293Ts, the restrictive capacity was diminished in HIV-1 and SIVmac. As demonstrated in this study, this effect was independent of gene expression, as TRIM22-106 is more strongly expressed than TRIM22 wildtype. DLS data showed that in purified solution, the TRIM22 wildtype exists as a disulphide mediated dimer, whereas the TRIM22-106 has a high amount of a protein with a much smaller radius which may be a TRIM22 monomer. This is unlikely to be a feature of the plasmid as both were identical, and able to produce flag-tagged protein which could be purified from the column in both forms. These data correspond with recent developments in Rhesus TRIM5α that demonstrate the centrality of TRIM oligomerisation to HIV restriction.

Table 7.1 Key Findings Arising From This Thesis by Chapter

III	▪ The majority of TRIM5α Polymorphisms are not within the putative virus-interacting PRYSPRY Domain ▪ Allele frequencies corresponding to substitutions L159L, D249G, and P479L are significantly higher in the Caio Cohort when compared to Yoruba populations or and Africa-at-large ▪ Undetectable viral load and female sex predicts survival in the Caio Community Cohort ▪ P479L was enriched in HIV-2 infected patients and progressors, and was associated with CD4+ T-cell decline in t-tests and ANOVA. Significance was not retained following adjustment. ▪ R136Q was associated with depressed viral load which was not retained following adjustment, and showed significantly positive impact on survival at the 10% level following adjustment
IV	▪ The majority TRIM22 polymorphisms were located within the putative virus-interacting PRYSPRY domain, followed by the coiled coil region ▪ TRIM22 SNP allele frequencies were similar in the Caio Cohort when compared to those observed in Yoruba and other African populations ▪ 3 new synonymous polymorphisms were observed in the population; 2 in the coiled coil and 1 in the PRYSPRY domain. 1 new codon sequence was observed with non synonymous: D360Y ▪ Rs7935564 (D155N) was associated with a 50% reduction in progression to death ▪ Rs1063303 was enriched in HIV-2 infected participants, significantly associated with 2.5x increased likelihood of being HIV-2 at the 10% level, a stepwise reduction in absolute CD4+ T-cell count with the addition of the minor allele, and twice the rate of progression to death following adjustment for gender
V	▪ *In silico* algorithms predicted: o G31S, L214F, and P479L as nsSNPs likely to have a functional impact on TRIM5α o D282N, R321K, and D360Y (all in and around the PRYSPRY domain) as likely to have a functional impact on TRIM22 ▪ There were 50% more nsSNPs in TRIM22 when compared TRIM5α from the same population, and there was no spatial similarity between them ▪ TRIM5α P479L was located on the cusp of the V4 loop- not implicated in virus interaction. TRIM5α's R136Q in the coiled coil consisted of a change from hydrophilic to neutral amino acid drawing the RING and PRYSPRY domains closer in 3D models ▪ There was more similarity between TRIM22 PRYSPRY polymorphisms and major residues associated with retroviral recognition and specificity in the Rhesus TRIM5α (321K-VL1, T415I-VL3, D360Y/T294K-residues involved in NMLV restriction) ▪ 294K, 321K,360Y, and 415I clustered together on the exterior of the PRYSPRY domain in models ▪ Consurf showed no evolutionary conservation of TRIM5α residues, but D282, D360, R321, and T415 in TRIM22 were predicted to be functional residues on account of their being highly conserved and exposed ▪ In comparison to TRIM5α, TRIM22 would appear under more selection pressure in the Caio Community Cohort, and the pressure is likely to be exerted HTLV-1 which has been endemic in West Africa historically
VI	▪ TRIM22 was sensitive to interferon stimulation in a lymphoid cell line, HIV-1 infection, and co-localised strongly with p24 in imaging studies ▪ DLS shows the presence of a protein with a small radius in a solution of purified TRIM22-106, that is not present in a corresponding solution of TRIM22 wildtype, suggesting high concentrations of a monomer ▪ Non-permissive U937 cells containing endogenous TRIM22 restrict VSV-G psudotyped HIV-1 and HIV-2, 2.5 and 10 fold better, respectively, when compared to their permissive counterparts which lack endogenous TRIM22 ▪ Overexpression of TRIM22 results in the restriction of VSV-G psuedotyped HIV-1, SIVmac, and FIV but not SIVsm, at maximum infectious titres ▪ TRIM22-106 results in diminished restriction of HIV-1 and SIVmac

The work herin has shown the ability of TRIM polymorphisms, especially those in TRIM22, to predict the outcome of HIV-2 infection. It is evident that considerably more selective pressure acts on TRIM22 within this population, when compared to TRIM5α.

Based on phylogeny and structural predictions, polymorphisms in TRIM5α and TRIM22 possess the ability to alter protein structure radically. This work was extended beyond the traditional remit of population genetics studies, by presenting preliminary data on the restrictive capacity of TRIM22 for a broad range of lentiviruses and the ability of a deleterious mutation to disrupt its effect in a manner independent of the strength of TRIM expression. Though demonstrating the centrality of host restriction factors, like TRIM proteins, to the story of HIV-2 control *in vivo*, the story of TRIM and HIV-2's relationship requires a sequel in order to answer the following:

What are the functional implications of TRIM5α's R136Q- a polymorphism associated with HIV-2 viral load and survival?

If overexpressed in human cells that express endogenous TRIM22, would the TRIM22 with the D155N mutation (associated with improved survival) lead to increased HIV restriction, possibly providing a therapeutic role for TRIM22?

Though it has been demonstrated that TRIM22 can restrict VSV-G pseudotyped lentiviruses is it able to restrict isolated virus in the same manner- and is this dose dependent?

How do the oligomeric states of TRIM22 and TRIM22 with the R242T substitution differ in proteomic studies?

7.2. Concluding Remarks and Future Work

While the population genetics studies have provided substantive evidence for the role of TRIM proteins in HIV-2 outcome showing evidence of selective pressure; all components of this thesis could be extended. A major limitation of Chapter III and IV (particularly Chapter III where the effect size of the variants was smaller), is that there was a small number of participants for the study. Many other studies of host genetics related to HIV have used multiple centres in order to generate sample sizes in the order of several hundreds. However, this has the potential to introduce bias from different HIV-2 strains and ethnicities. A benefit of this style of study, is that results have been generated from a small number of patients who are genetically homogenous and mono-infected with HIV-2 from the same group. In addition, participants have been matched for typical confounders such as age, gender, and period of follow-up. A benefit of doing a propspective study with such a long follow up (1988-2010) such as that of Caio, is the provision of key determinants of long-term-non-progression and survival data. A caveat is that the time points which have been described in this thesis are roughly five years apart. Therefore, it is possible that those newly infected with HIV-2 who progressed rapidly to AIDS may have been omitted, and that the dataset may be inherently enriched with long-term-non-progressors. Though rs1063303 in TRIM22 was strongly associated with rapid progression (a phenomenon that could be reproduced in functional assays); this finding was best observed *in vivo* when contrasting progressors and controllers in the discovery cohort. When the analyses were extended to the validation cohort, though rs1063303's enrichment amongst HIV-2 infected participants was retained, it was no longer associated with CD4+ T-cell count or viral load. This would suggest that an enrichment of LTNPs in this cohort may mask the impact of polymorphisms associated with disease progression.

A disadvantage of the Caio cohort, which has accompanied genetic homogeneity, is the fact that it is genetically distinct. Participants from the Caio Community have not mixed significantly with other ethnic groups. The only other study of the role of genetics in HIV-2 disease progression

investigated the role of HLA and KIR, both of which showed a significantly different distribution from those reported in Gambian or Senegalese populations. This same study, showed that HLA-B*1503 was linked to low CD4+ T-cell count and increased viral load, and that HLA-B*08 was associated with HIV-2 susceptibility(103). This is in stark contrast to previous studies of HLA, where HLA-B35 suggested an increased risk of HIV-2 disease progression in Senegal(318), or where HLA-B27, 57, and 51 have been implicated in HIV-1 disease progression(319). Therefore our ability to extrapolate these findings observed in the Caio Community to HIV-2 in West Africa on a whole, are quite limited due to the fact that people from other countries in the region may have different genetic backgrounds and carry potentially different HIV-2 strains. That said, as described previously; data from this thesis involving R136Q in TRIM5α and D155N, and R242T in TRIM22, are in line with previous findings related to HIV-1 disease progression.

A major limitation of performing gene-disease association studies is that they can only measure statistical associations between polymorphisms and clinical factors, but cannot ascribe causality. This thesis attempted to address this limitation in Chapters V and VI. *In silico* algorithims are extremely useful in quickly assessing the probable effects of variants on protein structure and function. This is especially helpful in studies of host restriction factors where a large number of polymorphisms occur. However, they are no substitute for crystallographic studies that derive the true structure, and oligomeric state, of proteins. Future work would greatly benefit from the cyrstallisation of the coiled coil domain of TRIM22, and that of TRIM22-106 to concretely determine the true impact of the R242T mutation on TRIM22 multimerisation.

In addition, the effect of host restriction factors is usually saturable. Typically restriction is only observed when a small amount of virus is used for challenge. In this study, the use of exogenous expression in HEK293Ts allowed for challenge with high infectious titres of VSV-G psuedotyped virus and presented no need for virus titration. It would be beneficial to examine the effect of

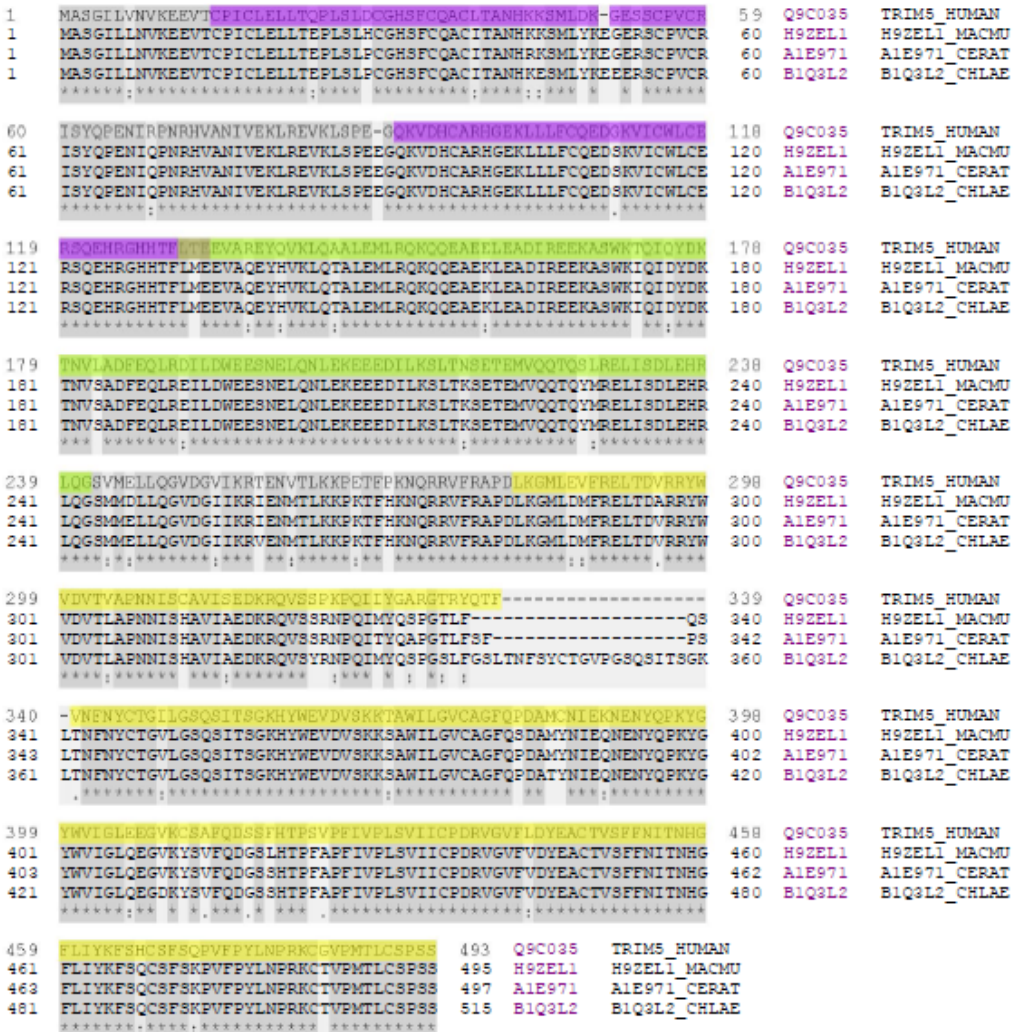
TRIM22 on virus isolates, and determine whether the effect is dose dependent on TRIM22 or infectious virus.

Finally, the D155N mutation, which was associated with survival in the Caio Cohort presents a unique opportunity to determine the therapeutic benefit of TRIM22. An essential next step would be the administration of whole TRIM22-D155N to cell cultures or animal models to determine its contribution to HIV-1 and HIV-2 restriction.

Appendix

Appendix I. Sequence Alignment and Phylogeny of TRIM5α from Non-Human Primates

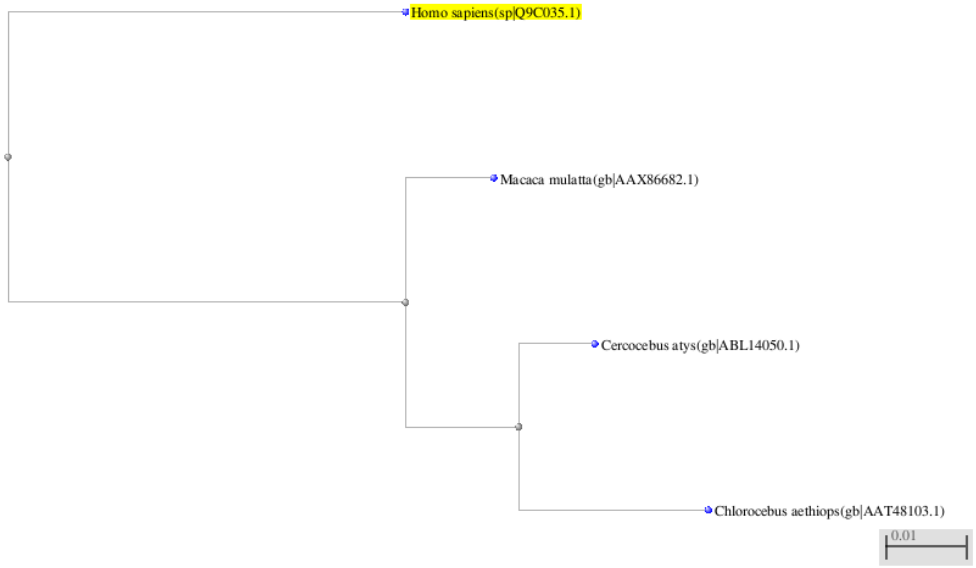
An alignment of TRIM5α from Humans (UniprotID: Q9C035-1), Rhesus Macaque (*Macaca mulatta) (Uniprot ID:* H9ZEL1), Sooty Mangabey (*Cercocebus atys*) (UniprotID: A1E971), and Green Monkey (*Chlorocebus atheiops*) (UniprotID: B1Q3L2) was performed to determine degree of identity between human TRIM5α from non-human primates capable of restricting immunodeficiency viruses. TRIM5α shared the most identity between human and rhesus macaque. TRIM5α is highly conserved amongst primates with the exception of the PRYSPRY domain, the domain.



Appendix Figure I- 1 Alignment of TRIM5α from Non-human Primates

TRIM5α Amino Acid Alignment from Human, Rhesus Macaque (Macaca mulatta), Sooty Mangabey (Cercocebus atys), and Green Monkey (Chlorocebus atheiops)

*RING and B-Box in purple, coiled coil in green, and PRYSPRY domain in yellow. *Pairwise alignment conducted using BLAST Alignment Tool*



Appendix Figure I- 2 Phylogenetic tree of TRIM5α from Non-human Primates

(human, rhesus macaque, sooty mangabey and green monkey)

Tree produced using Fast Minimum Evolution Tree Method

Appendix II. HWE and Haplotype Reconstruction for SNPs in TRIM5α

Appendix Table II- 1 Hardy Weinberg Equilibrium Calculations for TRIM5α SNPs

Locus	#Genot value	Obs.Het.	Exp.Het.	P-value	s.d.	Steps done
1	92	0.03261	0.03225	1.00000	0.00000	1001000
2	92	0.08696	0.08363	1.00000	0.00000	1001000
4	91	0.19780	0.19671	1.00000	0.00000	1001000
5	90	0.54444	0.41782	0.00429	0.00007	1001000
6	99	0.02020	0.02010	1.00000	0.00000	1001000
7	100	0.32000	0.36663	0.26917	0.00045	1001000
8	88	0.10227	0.09760	1.00000	0.00000	1001000

Appendix Table II- 2 Haplotypes in TRIM5α

Hap	Freq	sd	G31S	H43Y	R136Q	L159L	L214F	R238W	G249D	P479L
15	0.292	0	C	C	C	C	G	G	C	C
10	0.191	0	C	C	C	C	A	G	C	C
17	0.132	0	C	C	C	C	G	G	T	C
22	0.066	0	C	C	C	T	G	G	C	C
26	0.025	0	C	T	C	C	A	G	C	C
18	0.023	0	C	C	C	C	G	G	T	T
13	0.010	0	C	C	C	C	A	G	T	C
11	0.008	0	C	C	C	C	A	G	C	T
20	0.005	0	C	C	C	T	G	A	T	C
27	0.005	0	C	T	C	T	A	A	C	C
23	0.003	0	C	C	C	T	G	G	C	T
19	0.001	0	C	C	C	T	A	G	C	T

Haplotype reconstruction and frequency estimations generated 7 possible haplotypes observed at greater than 1% frequency.

Appendix III. HWE and Haplotype Reconstruction for SNPs in TRIM22

Appendix Table III- 1 Hardy Weinberg Equilibrium Calculations for TRIM22 SNPs

#	Name	Position	ObsHET	PredHET	HWpval	%Geno	FamTrio	MendErr	Alleles
1	rs200668710	5696479	0.043	0.043	1	100	0	0	G:A
2	rs10838543	5696532	0.478	0.454	0.8174	100	0	0	T:C
3	rs61735271	5696550	0.011	0.011	1	100	0	0	A:G
4	rs200924168	5697273	0.098	0.093	1	100	0	0	G:C
5	rs7935564	5697287	0.446	0.49	0.4774	100	0	0	G:A
6	rs779550205	5697337	0.022	0.022	1	100	0	0	C:T
7	snp5719589	5698358	0.011	0.011	1	100	0	0	T:C
8	rs78484876	5698419	0.239	0.227	1	100	0	0	T:C
9	snp5719650	5698420	0	0.043	2.00E-04	100	0	0	G:A
10	rs2291842	5698437	0.391	0.352	0.4833	100	0	0	T:C
11	rs2291843	5698489	0.011	0.011	1	100	0	0	A:G
12	rs1063303	5698520	0.467	0.463	1	100	0	0	G:C
13	rs61735273	5698526	0.054	0.073	0.2232	100	0	0	C:T
14	COSM1509111	5708243	0.011	0.011	1	98.9	0	0	G:A
15	rs73404240	5708583	0.077	0.074	1	98.9	0	0	C:A
16	rs61735276	5709064	0.011	0.011	1	98.9	0	0	C:T
17	rs12364019	5709113	0	0	1	98.9	0	0	G:G
18	snp5730458	5709228	0.011	0.011	1	98.9	0	0	A:T
19	snp5730459	5709229	0.022	0.022	1	98.9	0	0	G:T
20	rs61735325	5709285	0.011	0.011	1	98.9	0	0	G:T
21	rs150095329	5709395	0.011	0.011	1	98.9	0	0	C:T

Appendix Table III-2 Haplotypes of TRIM22 nsSNPS Observed at >1% Frequency

No	Frequency	rs10838543(T/C)	rs61735271(A/G)	rs200668710(G/A)	rs7935564(G/A)	rs200924168(G/C)	*5718516(C/T)	rs78484876(T/C)	rs2291842(T/C)	rs1063303(G/C)	rs61735273(C/T)	*5719589(T/C)	rs2291843(A/G)	COSM1509111(G/A)	rs73404240(C/A)	*5730459(G/T)	rs150095329(C/T)	rs61735276(C/T)	*5730458(A/T)
43	0.261	T	A	G	G	G	C	T	T	G	C	T	A	G	C	G	C	C	A
13	0.118	C	A	G	A	G	C	T	T	G	C	T	A	G	C	G	C	C	A
12	0.088	C	A	G	A	G	C	T	T	C	C	T	A	G	C	G	C	C	A
34	0.062	T	A	G	G	G	C	C	T	G	C	T	A	G	C	G	C	C	A
36	0.056	T	A	G	G	G	C	T	C	C	C	T	A	G	C	G	C	C	A
23	0.029	T	A	G	A	G	C	T	C	C	C	T	A	G	C	G	C	C	A
10	0.028	C	A	G	A	G	C	T	C	C	C	T	A	G	C	G	C	C	A
38	0.026	T	A	G	G	G	C	T	C	G	C	T	A	G	C	G	C	C	A
6	0.026	C	A	G	A	C	C	T	T	C	C	T	A	G	C	G	C	C	A
27	0.026	T	A	G	A	G	C	T	T	G	C	T	A	G	C	G	C	C	A
31	0.021	T	A	G	G	G	C	C	C	G	C	T	A	G	C	G	C	C	A
25	0.013	T	A	G	A	G	C	T	T	C	C	T	A	G	C	G	C	C	A
42	0.013	T	A	G	G	G	C	T	T	C	C	T	A	G	C	G	C	C	A
33	0.010	T	A	G	G	G	C	C	T	C	C	T	A	G	C	G	C	C	A
9	0.010	C	A	G	A	G	C	C	C	C	C	T	A	G	C	G	C	C	A

Linkage Disequilibrium Plots for TRIM5α and TRIM22 SNPs

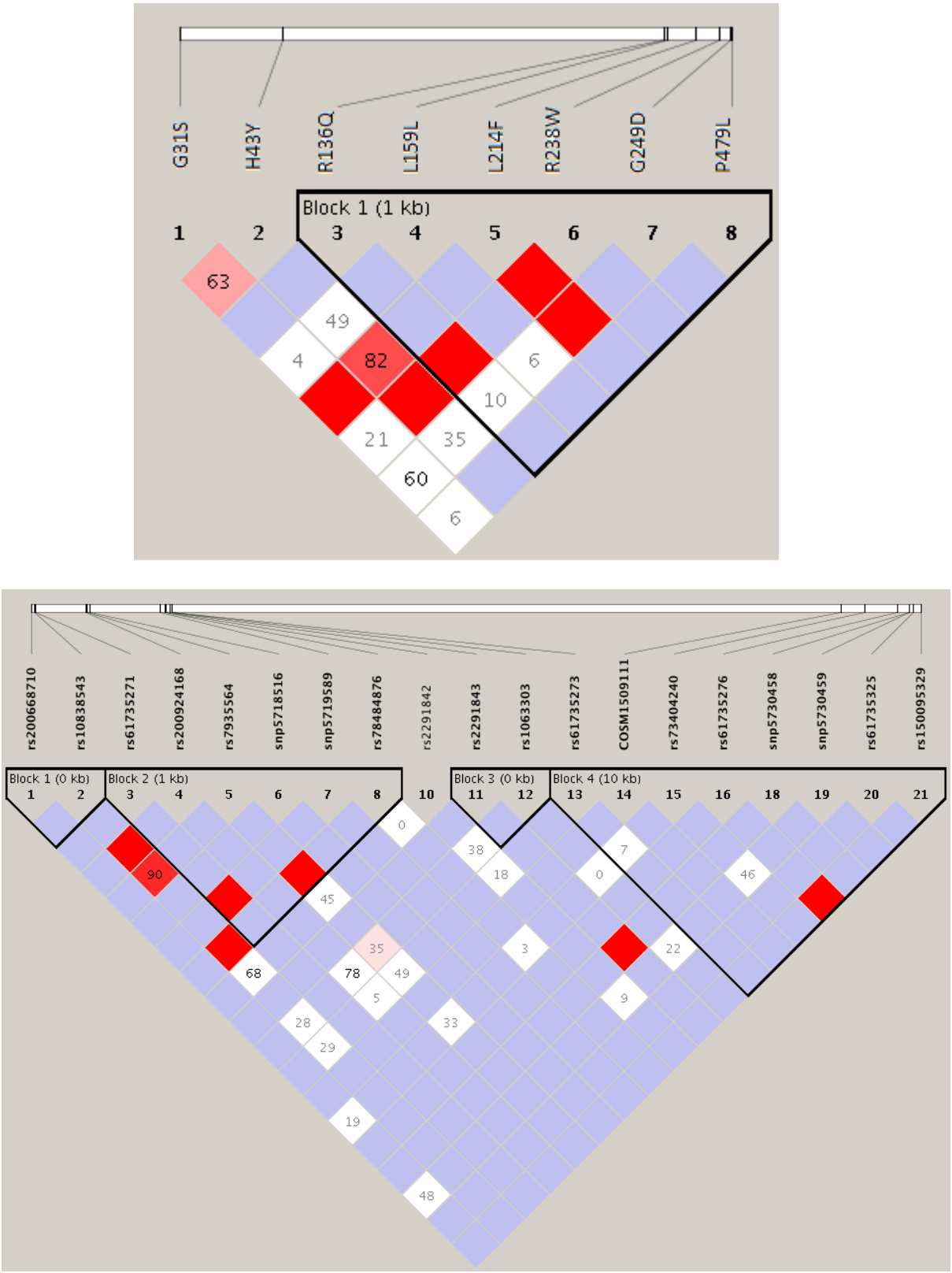
Haploview (Broad Institute) analyses of linkage disequilibrium, and illustrates two key outputs:

D' = value of D prime between the two loci, where D'=1 shows complete LD

LOD= log of the likelihood odds ratio, a measure of confidence in the value of D'

The data is illustrated in blocks, and coded as follows:

	D' < 1	D' = 1
LOD < 2	white	blue
LOD ≥ 2	shades of pink/red	bright red



Appendix Figure III- 1 Linkage Disequilibrium Plots for TRIM5α and TRIM22 SNPs

TRIM5α (Top) TRIM22 (Bottom)

Appendix IV. PDB Files Used to Generate 3D Protein Models in Phyre

The amino acid sequence of human TRIM5α (Uniprot ID Q9C035-1) and human TRIM22 (Uniprot ID: Q8IYM9) were submitted to Phyre 2.0 in order to generate 3D models based on crystal structures from proteins that share amino acid homology. The models generated are shown in Chapter V of this thesis, however this appendix contains the templates used to build the models for wildtype TRIM5α and TRIM22, their confidence score, percentage identity and additional information about the template which are described as follows:

Template: The templates are ranked by a raw alignment score (not shown) that is based on the number of aligned residues and the quality of alignment. This in turn is based on the similarity of residue probability distributions for each position, secondary structure similarity and the presence or absence of insertions and deletions.

Confidence: represents the probability (from 0 to 100) that the match between the selected TRIM sequence and this template is a true homology. A match with confidence of 90% or higher indicates great certainty that the protein adopts the overall fold, where the core of the protein is modelled at high accuracy (2-4Å rmsd from native,true structure). However, surface loops will probably deviate from the native.

% Identity: This indicates the percentage identity between the TRIM sequence and the template. Ideally a high accuracy model requires greater than 30% identity. However, even with sequence identities lower than this, the models provide useful information as long as the confidence is high.

Appendix Table IV- 1 Templates used to Generate 3D Model of Human TRIM5α

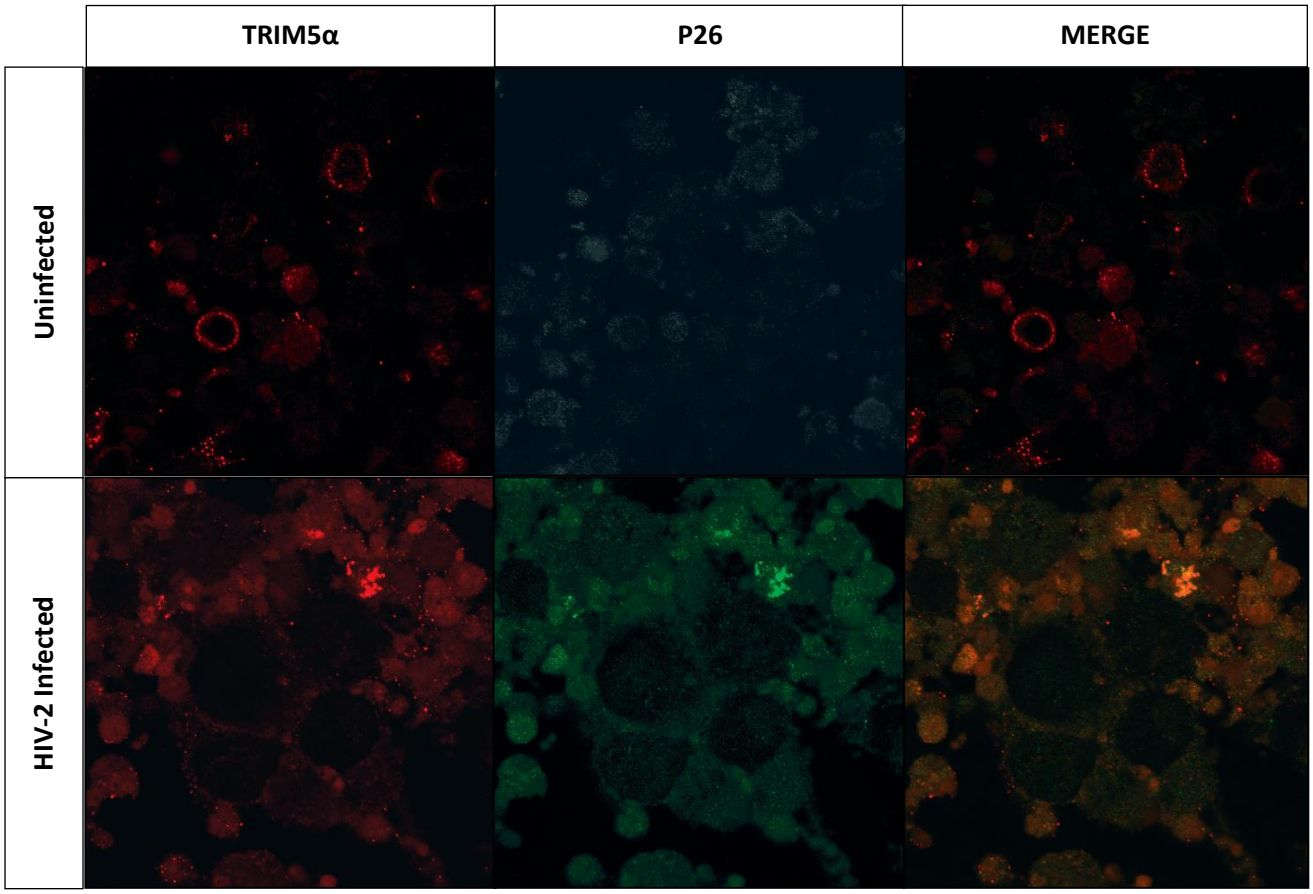
Template (PDB ID)	Confidence	% Identity	Template Information
4CG4	100	31	Coiled coil and B30.2 domains of TRIM20
2LM3	100	83	Rhesus TRIM5α PRYSPRY domain
4WVM	100	16	Stonustoxin subunit alpha
4B3N	100	74	Structure of the rhesus TRIM5α PRYSPRY domain
2FBE	100	30	Concanavalin A-like lectins/glucanases SPRY domain
4WVM	100	17	Stonustoxin subunit beta
2WL1	100	35	Pyrin pryspry domain
4N7U	100	32	Intracellular b30.2 domain of butyrophilin subfamily 3 member a1
2IWG	100	42	Concanavalin A-like lectins/ glucanases PRYSPRY domain
3KB5	100	26	Tripartite motif-containing protein 72 PRYSPRY domain
2FNJ	99.8	18	Concanavalin A-like lectins/ glucanases PRYSPRY domain
2AFJ	99.7	17	Concanavalin A-like lectins/ glucanases PRYSPRY domain
4GT6	99.7	21	PRYSPRY domain of human herc1
4XW3	99.6	18	PRYSPRY domain of the human dead-box protein2 (ddx1)
2ECW	99.5	59	ring2 finger domain of tripartite motif protein 30
2ECV	99.5	100	ring2 finger domain of TRIM5
4TXA	99.5	19	N-terminus of roquin (mouse)
5FER	99.5	33	e3 ubiquitin/isg15 ligase trim25; PDBTitle: complex of trim25 ring with ubch5-ub
2F42	99.4	15	dimerization and u-box domains of zebrafish c-terminal of hsp702 interacting protein

Appendix Table IV- 2 Templates used to Generate 3D Model of Human TRIM22

Template (PDB ID)	Confidence	% Identity	Template Information
4CG4	100	29	Coiled coil and B30.2 domains of TRIM20
4WVM	100	17	Stonustoxin subunit alpha
2LM3	100	50	Rhesus TRIM5α PRYSPRY domain
4WVM	100	17	Stonustoxin subunit beta
2FBE	100	24	Concanavalin A-like lectins/glucanases SPRY domain
2WL1	100	36	Pyrin pryspry domain
4B8E	100	24	PRYSPRY domain of TRIM25
4B3N	100	42	Rhesus TRIM5α PRYSPRY domain
4N7U	100	34	Intracellular b30.2 domain of butyrophilin subfamily 3 member a1
2IWG	100	43	Concanavalin A-like lectins/ glucanases PRYSPRY domain
3KB5	100	26	Tripartite motif-containing protein 72 PRYSPRY domain
2FNJ	99.9	21	Concanavalin A-like lectins/ glucanases PRYSPRY domain
2AFJ	99.8	20	Concanavalin A-like lectins/ glucanases PRYSPRY domain
4GT6	99.7	19	PRYSPRY domain of human herc1
4XW3	99.7	17	PRYSPRY domain of the human dead-box protein2 (ddx1)
4TXA	99.6	20	N-terminus of roquin (mouse)
3TOJ	99.6	15	Set1/ash2 histone methyltransferase complex subunit (ash2l) PRYSPRY domain
2ECW	99.5	56	TRIM30 RING domain
5FER	99.5	32	TRIM25 RING domain
2F42	99.5	19	BBox domains of zebrafish c-terminal of hsp702 interacting protein
2ECV	99.5	100	RING domain of TRIM5
2EGP	99.5	60	TRIM34 RING domain

Appendix V. Pilot Colocalisation of HIV-2 Capsid and TRIM5α /TRIM22

Data is widely available on the antiviral potential of TRIM5α especially in relevance to HIV-1. However there are less data available of relationship between TRIM5α and HIV-2. To complete the data gap, the interaction between p26 and TRIM5α was also investigated. C8166 cells were infected with 0.1 MOI of HIV-2 CBL20. After 18 days, cells were blocked, stained for HIV2-p26 (KC57) conjugated to FITC (green) and TRIM5α with a secondary antibody conjugated to AF568 (red). The cells were mounted on slides, imaged using confocal microscopy, and analysed for pixel intensity and colocalisation in the ImageJ software platform.

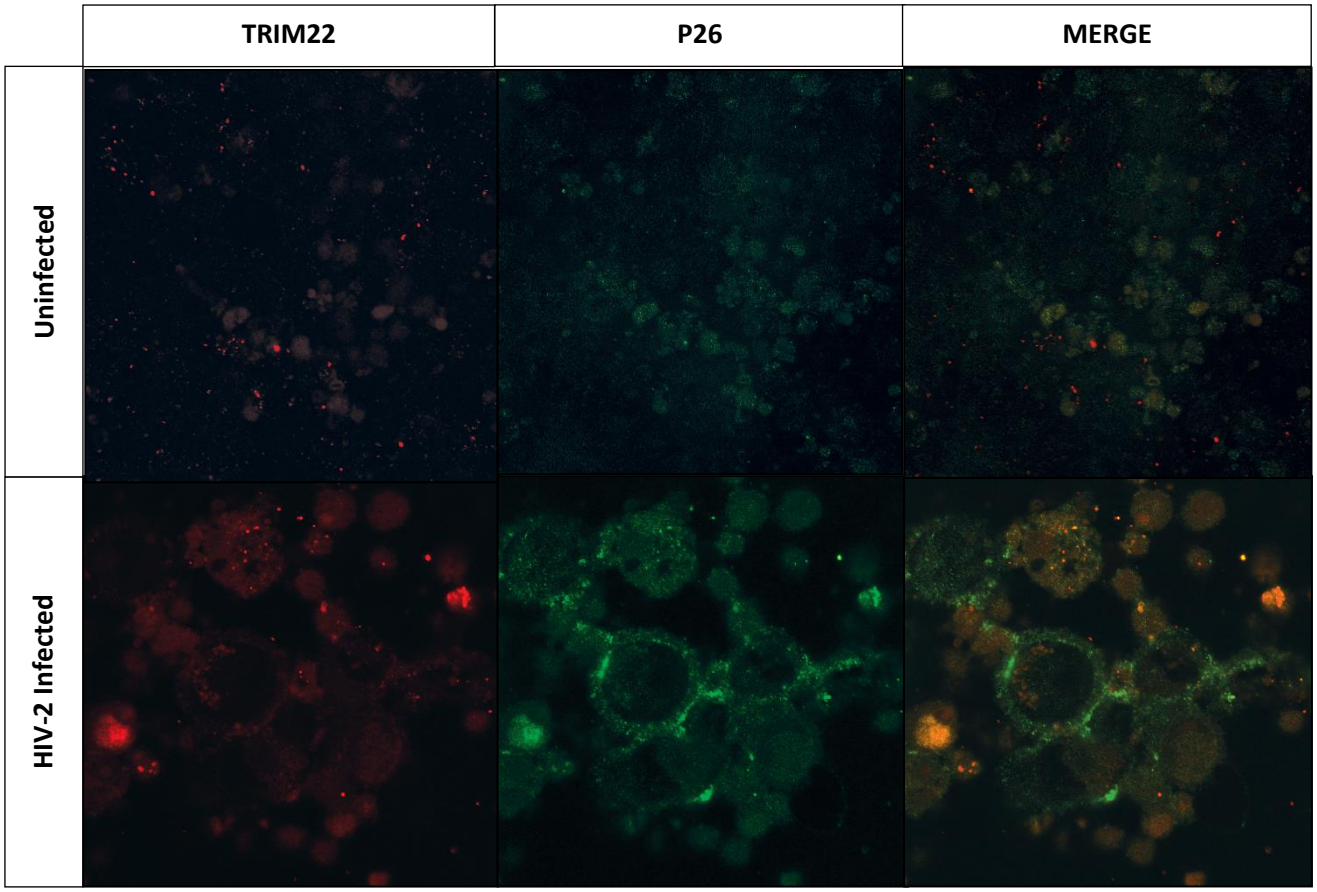


Appendix Figure V- 1 TRIM5α Co-localises with HIV-2 Capsid p26

TRIM22 Confocal immunofluorescence microscopy revealed that TRIM5α (red) is not substantially upregulated in C8166 cells following 18 days infection with HIV-2 CBL20. The Average Pixel Intensity of the TRIM5α staining is 2.568 in uninfected cells, whereas the average pixel intensity of TRIM22 is 2.670 in HIV-2-CBL20 infected cells. TRIM5α co-localises with the HIV-2 capsid protein p26 (green) at a co-localisation coefficient of 0.946.*

When compared with uninfected cell slides (shown in Chapter VI) TRIM5α was not significantly upregulated in the presence of HIV- 2 CBL20 infection, however it more strongly colocalised with HIV-2 p26. This is primarily because TRIM5α seems more ubiquitously expressed in this cell line when compared to TRIM22.

The data shown in the association studies would benefit greatly from an understanding of whether TRIM22 and HIV-2 do interact in immunoprecipitation assays. A precursor to these assays would involve imaging studies to determine whether TRIM22 also co-localises with the capsid protein in the case of HIV-2 infection, C8166 cells were infected with 0.1 MOI of HIV-2 CBL20. After 18 days, cells were blocked, stained for HIV2-p26 (KC57) conjugated to FITC (green) and TRIM22 with a secondary antibody conjugated to AF568 (red). The cells were mounted on slides, imaged using confocal microscopy, and analysed for pixel intensity and colocalisation in the ImageJ software platform. When compared with uninfected cell slides TRIM22 was significantly upregulated in the presence of HIV-2-CBL20 infection, and strongly colocalised with HIV-2 p26. (Colocalisation Coefficient: 0.908; Uninfected TRIM22 Average Pixel Intensity- 1.061. Infected TRIM22 Average Pixel Intensity- 3.111**).

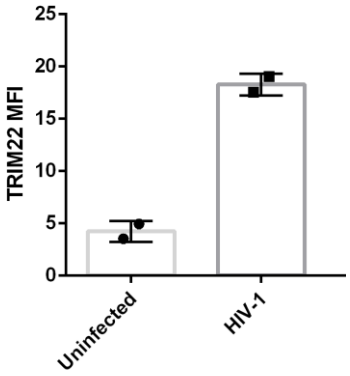


Appendix Figure V-2 TRIM22 Co-localises with HIV-2 Capsid p26

TRIM22 Confocal immunofluorescence microscopy revealed that TRIM22 (red) is substantially upregulated in C8166 cells following 18 days infection with HIV-2 CBL20. The Average Pixel Intensity of the TRIM22 staining is 1.061 in uninfected cells, whereas the average pixel intensity of TRIM22 is 3.111 in HIV-2-CBL20 infected cells. TRIM22 co-localises with the HIV-2 capsid protein p26 (green) at a co-localisation coefficient of 0.908.*

Appendix VI. TRIM22 Upregulation in C8166 cells upon infection with HIV-1 and HIV-2

C8166 cells were infected with HIV-1-IIIB and stained for TRIM22 and capsid at 3 days days post infection, respectively. When compared to uninfected cells, the expression of TRIM22 substantially increased in response to retroviral infection. There was 4.33 times more TRIM22 expressed following infection with HIV-1-IIIB (t-test, p=0.005)



Appendix Figure VI- 1 TRIM22 Upregulation in C8166 Cells Infected with HIV-1 and HIV-2

MFI=Median Fluorescence Intensity of TRIM22 within infected C8166. Duplicate technical repeats

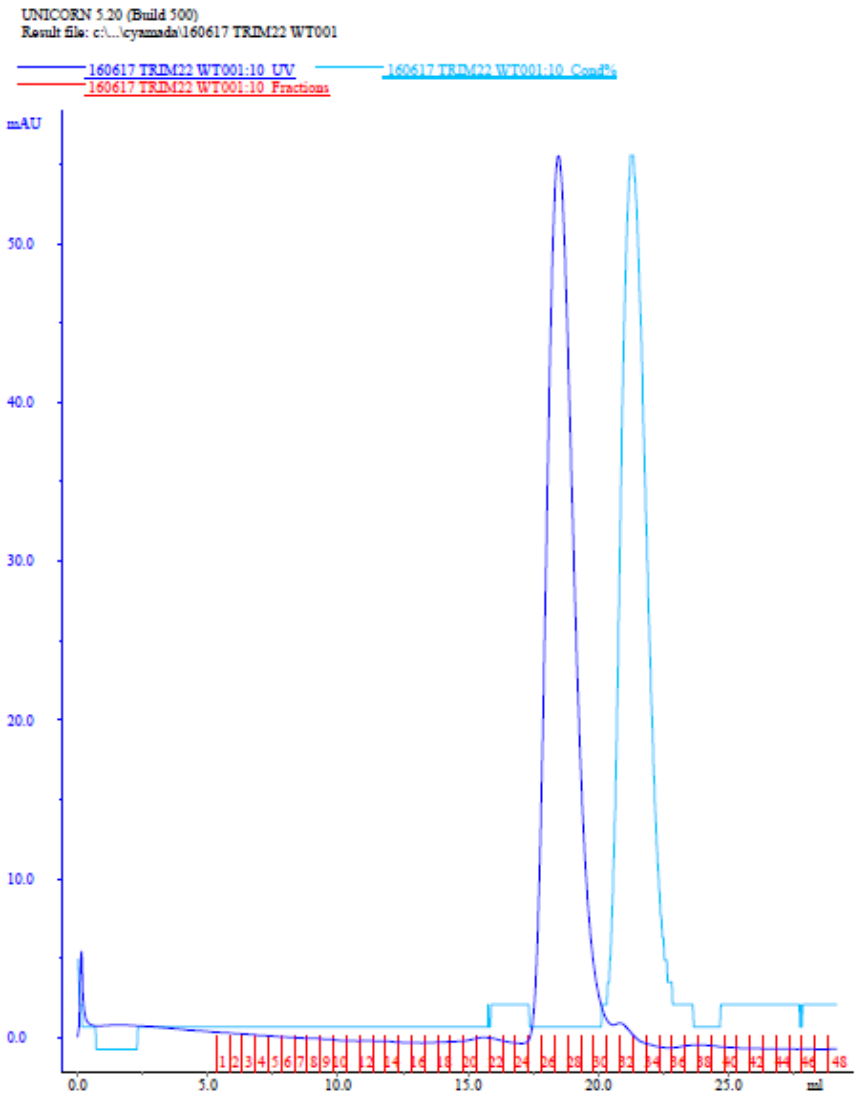
Appendix VII. Size Exclusion Chromatography to determine Oligomeric State of TRIM22 and TRIM22-106

In order to determine the oligomeric state of TRIM22 wildtype and TRIM22-106, purified TRIM22 and TRIM22-106 were subject to Gel Filtration Chromatography. This technique separates proteins on the basis of size, and incorporates passing the proteins through a bed of porous beads. Small molecules diffuse further into the pores of the beads, and move slowly through the bed. Larger molecules diffuse quickly, because often they pass into the beads less, or not at all.

TRIM22 and TRIM22-106 were separated on a Superdex 200 Increase 10/300 GL Column (GE Lifesciences # 17-5175-01) in a buffer containing 10 mM sodium phosphate and 140 mM sodium chloride, at pH of 7.4. The flow rate was 0.4 ml/min at room temperature. As both conformation and size contribute to the degree of retention, it would be expected for dimeric structures to leave the column before monomers. A chart of relative molecular mass compared to flow rate is available at:

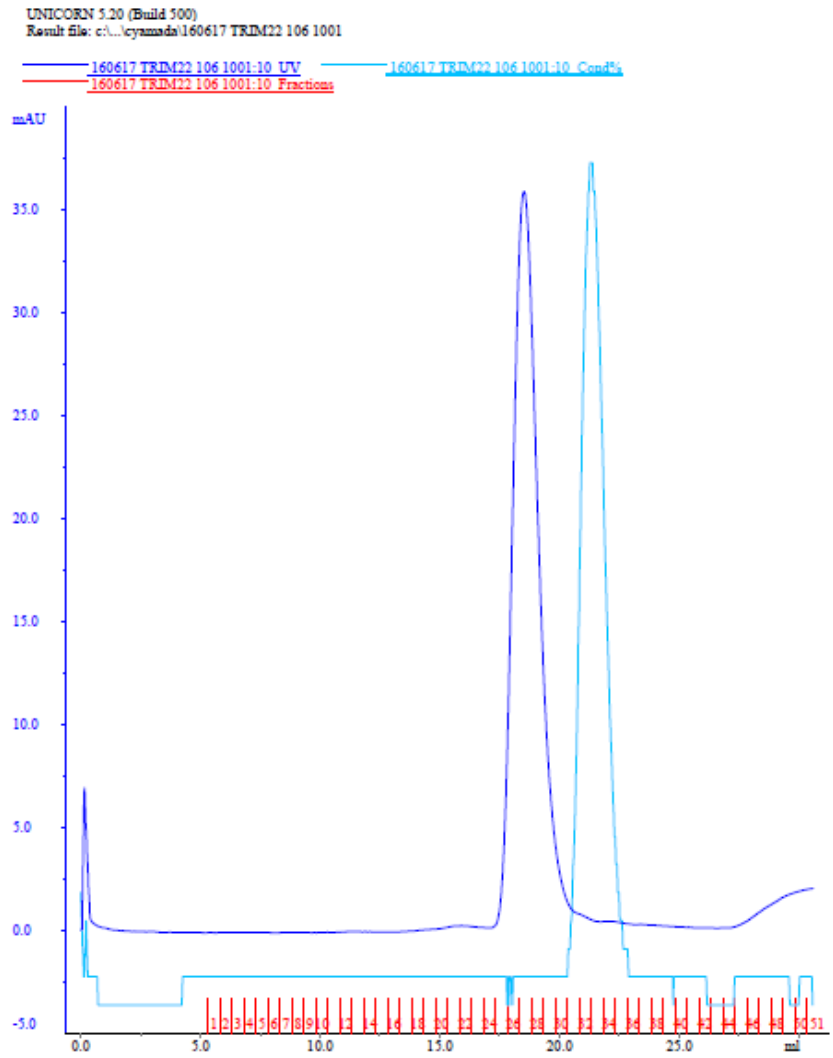
http://research.med.helsinki.fi/corefacilities/akta/manuals/columns/Superdex75and200_10-300GL.pdf

Figures below show the Size Exclusion Chromatography data for TRIM22 and TRIM22-106 where volume is shown on the x axis in mL (representing flow rate) and the mAU is on the y axis representing the UV absorption of the protein (concentration dependent). Both sets of data show two peaks of low concentration at relative mass (Mr) of 13600Da, and 1355 which are too small to represent TRIM22 in its monomer or dimer form. The fractions at these two peaks were then analysed using SDS-PAGE.



Appendix Figure VII- 1 Size Exclusion Chromatography on Purified TRIM22 Wildtype

TRIM22 Wildtype show two peaks at 21 and 27ml, corresponding roughly to a size relative mass (Mr) of 13600Da, and 1355Da, both at mAU of 57. The result indicates that there is no wildtype TRIM22 as a dimer or monomer despite the Western Blot of the same sample showing presence of the dimer

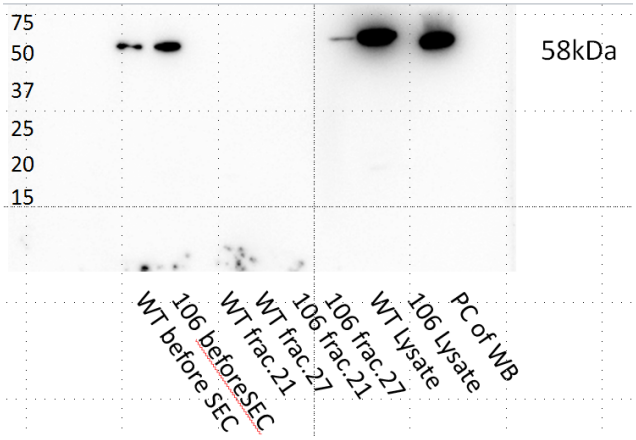


Appendix Figure VII- 2 Size Exclusion Chromatography on Purified TRIM22-106

TRIM22 Wildtype show two peaks at 21 and 27ml, corresponding roughly to a size relative mass (Mr) of 13600Da, and 1355Da, both at mAU of 57. The result indicates that there is no wildtype TRIM22 as a dimer or monomer despite the Western Blot of the same sample showing presence of the dimer

SDS PAGE of Eluents Obtained from Size Exclusion Chromatography

The fractions obtained from the SEC were subjected to SDS-PAGE in the presence of DTT. The gel was then blotted, and stained for actin antibody. The sample of TRIM22 WT (WT before SEC) and TRIM22-106(106 before SEC) had TRIM22 monomer of 58kDa before Size Exclusion Chromatography. The proteins were also present in high concentrations in the original cell lysate. However, the fragments obtained at the two peaks, 21 and 27 in SEC, did not contain TRIM22. It was observed that the concentration of TRIM22 and TRIM22-106 within the purified sample, was too low for detection using SEC. Therefore, Dynamic Light Scattering as detailed in Chapter VI was used to determine oligomeric state of TRIM22.



Appendix Figure VII- 3 SDS-PAGE and Western Blot of Eluents from Size Exclusion Chromatography

References

1. Pantaleo G, Fauci AS. Immunopathogenesis of HIV infection. Annual review of microbiology. 1996;50:825-54. PubMed PMID: 8905100.
2. Cohen MS, Hellmann N, Levy JA, DeCock K, Lange J. The spread, treatment, and prevention of HIV-1: evolution of a global pandemic. The Journal of clinical investigation. 2008 Apr;118(4):1244-54. PubMed PMID: 18382737. Pubmed Central PMCID: 2276790.
3. Joint United Nations Programme on HAU. 2015 Report on the Global AIDS Epidemic. 2015.
4. Egberink H, Horzinek MC. Animal immunodeficiency viruses. Veterinary microbiology. 1992 Nov;33(1-4):311-31. PubMed PMID: 1336243.
5. Pedersen NC, Ho EW, Brown ML, Yamamoto JK. Isolation of a T-lymphotropic virus from domestic cats with an immunodeficiency-like syndrome. Science. 1987 Feb 13;235(4790):790-3. PubMed PMID: 3643650.
6. St-Louis MC, Cojocariu M, Archambault D. The molecular biology of bovine immunodeficiency virus: a comparison with other lentiviruses. Animal health research reviews. 2004 Dec;5(2):125-43. PubMed PMID: 15984320.
7. Leroux C, Cadore JL, Montelaro RC. Equine Infectious Anemia Virus (EIAV): what has HIV's country cousin got to tell us? Veterinary research. 2004 Jul-Aug;35(4):485-512. PubMed PMID: 15236678.
8. Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR. An African primate lentivirus (SIVsm) closely related to HIV-2. Nature. 1989 Jun 01;339(6223):389-92. PubMed PMID: 2786147.
9. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, et al. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. Science. 2006 Jul 28;313(5786):523-6. PubMed PMID: 16728595. Pubmed Central PMCID: 2442710.
10. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, et al. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature. 1999 Feb 04;397(6718):436-41. PubMed PMID: 9989410.
11. Takehisa J, Kraus MH, Ayouba A, Bailes E, Van Heuverswyn F, Decker JM, et al. Origin and biology of simian immunodeficiency virus in wild-living western gorillas. Journal of virology. 2009 Feb;83(4):1635-48. PubMed PMID: 19073717. Pubmed Central PMCID: 2643789.
12. Broussard SR, Staprans SI, White R, Whitehead EM, Feinberg MB, Allan JS. Simian immunodeficiency virus replicates to high levels in naturally infected African green monkeys without inducing immunologic or neurologic disease. Journal of virology. 2001 Mar;75(5):2262-75. PubMed PMID: 11160730. Pubmed Central PMCID: 114810.
13. Rey-Cuille MA, Berthier JL, Bomsel-Demontoy MC, Chaduc Y, Montagnier L, Hovanessian AG, et al. Simian immunodeficiency virus replicates to high levels in sooty mangabeys without inducing disease. Journal of virology. 1998 May;72(5):3872-86. PubMed PMID: 9557672. Pubmed Central PMCID: 109612.
14. Kestler H, Kodama T, Ringler D, Marthas M, Pedersen N, Lackner A, et al. Induction of AIDS in rhesus monkeys by molecularly cloned simian immunodeficiency virus. Science. 1990 Jun 1;248(4959):1109-12. PubMed PMID: 2160735.
15. Hahn BH, Shaw GM, De Cock KM, Sharp PM. AIDS as a zoonosis: scientific and public health implications. Science. 2000 Jan 28;287(5453):607-14. PubMed PMID: 10649986.
16. Gao F, Yue L, Robertson DL, Hill SC, Hui H, Biggar RJ, et al. Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. Journal of virology. 1994 Nov;68(11):7433-47. PubMed PMID: 7933127. Pubmed Central PMCID: 237186.

17. Brigitte E. Beer EB, Paul M. Sharp, and Vanessa M. Hirsch. Diversity and Evolution of Primate Lentiviruses.
18. Damond F, Worobey M, Campa P, Farfara I, Colin G, Matheron S, et al. Identification of a highly divergent HIV type 2 and proposal for a change in HIV type 2 classification. AIDS research and human retroviruses. 2004 Jun;20(6):666-72. PubMed PMID: 15242544.
19. Kappes JC, Morrow CD, Lee SW, Jameson BA, Kent SB, Hood LE, et al. Identification of a novel retroviral gene unique to human immunodeficiency virus type 2 and simian immunodeficiency virus SIVMAC. Journal of virology. 1988 Sep;62(9):3501-5. PubMed PMID: 3136256. Pubmed Central PMCID: 253477.
20. Chen Z, Luckay A, Sodora DL, Telfer P, Reed P, Gettie A, et al. Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. Journal of virology. 1997 May;71(5):3953-60. PubMed PMID: 9094672. Pubmed Central PMCID: 191547.
21. Chen Z, Telfier P, Gettie A, Reed P, Zhang L, Ho DD, et al. Genetic characterization of new West African simian immunodeficiency virus SIVsm: geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey troop. Journal of virology. 1996 Jun;70(6):3617-27. PubMed PMID: 8648696. Pubmed Central PMCID: 190237.
22. Apetrei C, Metzger MJ, Richardson D, Ling B, Telfer PT, Reed P, et al. Detection and partial characterization of simian immunodeficiency virus SIVsm strains from bush meat samples from rural Sierra Leone. Journal of virology. 2005 Feb;79(4):2631-6. PubMed PMID: 15681464. Pubmed Central PMCID: 546599.
23. Santiago ML, Range F, Keele BF, Li Y, Bailes E, Bibollet-Ruche F, et al. Simian immunodeficiency virus infection in free-ranging sooty mangabeys (Cercocebus atys atys) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. Journal of virology. 2005 Oct;79(19):12515-27. PubMed PMID: 16160179. Pubmed Central PMCID: 1211554.
24. Lemey P, Pybus OG, Wang B, Saksena NK, Salemi M, Vandamme AM. Tracing the origin and history of the HIV-2 epidemic. Proceedings of the National Academy of Sciences of the United States of America. 2003 May 27;100(11):6588-92. PubMed PMID: 12743376. Pubmed Central PMCID: 164491.
25. Ibe S, Yokomaku Y, Shiino T, Tanaka R, Hattori J, Fujisaki S, et al. HIV-2 CRF01_AB: first circulating recombinant form of HIV-2. Journal of acquired immune deficiency syndromes (1999). 2010 Jul;54(3):241-7. PubMed PMID: 20502347.
26. Schim van der Loeff MF, Aaby P. Towards a better understanding of the epidemiology of HIV-2. Aids. 1999;13 Suppl A:S69-84. PubMed PMID: 10885765.
27. de Silva T, Weiss RA. HIV-2 goes global: an unaddressed issue in Indian anti-retroviral programmes. The Indian journal of medical research. 2010 Dec;132:660-2. PubMed PMID: 21245611. Pubmed Central PMCID: 3102451.
28. Cortes E, Detels R, Aboulafia D, Li XL, Moudgil T, Alam M, et al. HIV-1, HIV-2, and HTLV-I infection in high-risk groups in Brazil. The New England journal of medicine. 1989 Apr 13;320(15):953-8. PubMed PMID: 2927478.
29. Valadas E, Franca L, Sousa S, Antunes F. 20 years of HIV-2 infection in Portugal: trends and changes in epidemiology. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2009 Apr 15;48(8):1166-7. PubMed PMID: 19292644.
30. Tubiana R, Le Chenadec J, Rouzioux C, Mandelbrot L, Hamrene K, Dollfus C, et al. Factors associated with mother-to-child transmission of HIV-1 despite a maternal viral load <500 copies/ml at delivery: a case-control study nested in the French perinatal cohort (EPF-ANRS CO1). Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2010 Feb 15;50(4):585-96. PubMed PMID: 20070234.

31. L'Hernault A, Greatorex JS, Crowther RA, Lever AM. Dimerisation of HIV-2 genomic RNA is linked to efficient RNA packaging, normal particle maturation and viral infectivity. Retrovirology. 2007 Dec 13;4:90. PubMed PMID: 18078509. Pubmed Central PMCID: 2222663.
32. Taviera N. HIV-2 Envelope: Structure, Diversity, and Evolution. In: Thomas J. Hope MS, Douglas Richman, editor. Encyclopedia of AIDS: Springer Science+Business Media New York 2013; 2014.
33. Lever AM. Molecular Biology of HIV-2. In: Thomas J. Hope MS, Douglas Richman, editor. Encyclopedia of AIDS: Springer Science+Business Media New York 2013; 2014.
34. Duvall MG, Lore K, Blaak H, Ambrozak DA, Adams WC, Santos K, et al. Dendritic cells are less susceptible to human immunodeficiency virus type 2 (HIV-2) infection than to HIV-1 infection. Journal of virology. 2007 Dec;81(24):13486-98. PubMed PMID: 17913821. Pubmed Central PMCID: 2168847.
35. Pierson TC, Doms RW. HIV-1 entry and its inhibition. Current topics in microbiology and immunology. 2003;281:1-27. PubMed PMID: 12932074. Epub 2003/08/23. eng.
36. Morner A, Bjorndal A, Leandersson AC, Albert J, Bjorling E, Jansson M. CCR5 or CXCR4 is required for efficient infection of peripheral blood mononuclear cells by promiscuous human immunodeficiency virus type 2 primary isolates. AIDS research and human retroviruses. 2002 Feb 10;18(3):193-200. PubMed PMID: 11839153.
37. Morner A, Bjorndal A, Albert J, Kewalramani VN, Littman DR, Inoue R, et al. Primary human immunodeficiency virus type 2 (HIV-2) isolates, like HIV-1 isolates, frequently use CCR5 but show promiscuity in coreceptor usage. Journal of virology. 1999 Mar;73(3):2343-9. PubMed PMID: 9971817. Pubmed Central PMCID: 104479.
38. Herschhorn A, Hizi A. Retroviral reverse transcriptases. Cellular and Molecular Life Sciences. 2010;67(16):2717-47.
39. Popov S, Rexach M, Zybarth G, Reiling N, Lee MA, Ratner L, et al. Viral protein R regulates nuclear import of the HIV-1 pre-integration complex. The EMBO journal. 1998 Feb 16;17(4):909-17. PubMed PMID: 9463369. Pubmed Central PMCID: 1170440.
40. Matreyek KA, Engelman A. Viral and cellular requirements for the nuclear entry of retroviral preintegration nucleoprotein complexes. Viruses. 2013 Oct 07;5(10):2483-511. PubMed PMID: 24103892. Pubmed Central PMCID: 3814599.
41. Lama J, Planelles V. Host factors influencing susceptibility to HIV infection and AIDS progression. Retrovirology. 2007;4:52. PubMed PMID: 17651505. Pubmed Central PMCID: 1978541.
42. Engelman A, Cherepanov P. The structural biology of HIV-1: mechanistic and therapeutic insights. Nature reviews Microbiology. 2012 Mar 16;10(4):279-90. PubMed PMID: 22421880. Pubmed Central PMCID: 3588166.
43. de Silva TI, van Tienen C, Onyango C, Jabang A, Vincent T, Loeff MF, et al. Population dynamics of HIV-2 in rural West Africa: comparison with HIV-1 and ongoing transmission at the heart of the epidemic. Aids. 2013 Jan 02;27(1):125-34. PubMed PMID: 23032414.
44. Langford SE, Ananworanich J, Cooper DA. Predictors of disease progression in HIV infection: a review. AIDS research and therapy. 2007 May 14;4:11. PubMed PMID: 17502001. Pubmed Central PMCID: 1887539.
45. Ariyoshi K, Jaffar S, Alabi AS, Berry N, Schim van der Loeff M, Sabally S, et al. Plasma RNA viral load predicts the rate of CD4 T cell decline and death in HIV-2-infected patients in West Africa. Aids. 2000 Mar 10;14(4):339-44. PubMed PMID: 10770535.
46. Roberts RB, Murray HW, Rubin BY, Masur H. Opportunistic infections and impaired cell-mediated immune responses in patients with the acquired immune deficiency syndrome. Transactions of the American Clinical and Climatological Association. 1984;95:40-51. PubMed PMID: 6433529. Pubmed Central PMCID: 2279602.

47. Centers for Disease C. Update on acquired immune deficiency syndrome (AIDS)--United States. MMWR Morbidity and mortality weekly report. 1982 Sep 24;31(37):507-8, 13-4. PubMed PMID: 6815471.
48. Guyader M, Emerman M, Sonigo P, Clavel F, Montagnier L, Alizon M. Genome organization and transactivation of the human immunodeficiency virus type 2. Nature. 1987 Apr 16-22;326(6114):662-9. PubMed PMID: 3031510.
49. Witvrouw M, Pannecouque C, Van Laethem K, Desmyter J, De Clercq E, Vandamme AM. Activity of non-nucleoside reverse transcriptase inhibitors against HIV-2 and SIV. Aids. 1999 Aug 20;13(12):1477-83. PubMed PMID: 10465070.
50. Isaka Y, Miki S, Kawauchi S, Suyama A, Sugimoto H, Adachi A, et al. A single amino acid change at Leu-188 in the reverse transcriptase of HIV-2 and SIV renders them sensitive to non-nucleoside reverse transcriptase inhibitors. Archives of virology. 2001;146(4):743-55. PubMed PMID: 11402860.
51. Witvrouw M, Pannecouque C, Switzer WM, Folks TM, De Clercq E, Heneine W. Susceptibility of HIV-2, SIV and SHIV to various anti-HIV-1 compounds: implications for treatment and postexposure prophylaxis. Antiviral therapy. 2004 Feb;9(1):57-65. PubMed PMID: 15040537.
52. Reid P, MacInnes H, Cong ME, Heneine W, Garcia-Lerma JG. Natural resistance of human immunodeficiency virus type 2 to zidovudine. Virology. 2005 Jun 05;336(2):251-64. PubMed PMID: 15892966.
53. Boyer PL, Sarafianos SG, Clark PK, Arnold E, Hughes SH. Why do HIV-1 and HIV-2 use different pathways to develop AZT resistance? PLoS pathogens. 2006 Feb;2(2):e10. PubMed PMID: 16485036. Pubmed Central PMCID: 1364504.
54. Desbois D, Roquebert B, Peytavin G, Damond F, Collin G, Benard A, et al. In vitro phenotypic susceptibility of human immunodeficiency virus type 2 clinical isolates to protease inhibitors. Antimicrobial agents and chemotherapy. 2008 Apr;52(4):1545-8. PubMed PMID: 18227188. Pubmed Central PMCID: 2292533.
55. Masse S, Lu X, Dekhtyar T, Lu L, Koev G, Gao F, et al. In vitro selection and characterization of human immunodeficiency virus type 2 with decreased susceptibility to lopinavir. Antimicrobial agents and chemotherapy. 2007 Sep;51(9):3075-80. PubMed PMID: 17576848. Pubmed Central PMCID: 2043247.
56. Flexner C. HIV drug development: the next 25 years. Nature reviews Drug discovery. 2007 Dec;6(12):959-66. PubMed PMID: 17932493.
57. Ekouevi DK, Tchounga BK, Coffie PA, Tegbe J, Anderson AM, Gottlieb GS, et al. Antiretroviral therapy response among HIV-2 infected patients: a systematic review. BMC infectious diseases. 2014 Aug 26;14:461. PubMed PMID: 25154616. Pubmed Central PMCID: 4156654.
58. Organization WH. CONSOLIDATED GUIDELINES ON HIV PREVENTION, DIAGNOSIS, TREATMENT AND CARE FOR KEY POPULATIONS. 2014.
59. Matheron S, Damond F, Benard A, Taieb A, Campa P, Peytavin G, et al. CD4 cell recovery in treated HIV-2-infected adults is lower than expected: results from the French ANRS CO5 HIV-2 cohort. Aids. 2006 Feb 14;20(3):459-62. PubMed PMID: 16439883.
60. Soares RS, Tendeiro R, Foxall RB, Baptista AP, Cavaleiro R, Gomes P, et al. Cell-associated viral burden provides evidence of ongoing viral replication in aviremic HIV-2-infected patients. Journal of virology. 2011 Mar;85(5):2429-38. PubMed PMID: 21159859. Pubmed Central PMCID: 3067805.
61. Gilleece Y, Chadwick DR, Breuer J, Hawkins D, Smit E, McCrae LX, et al. British HIV Association guidelines for antiretroviral treatment of HIV-2-positive individuals 2010. HIV medicine. 2010 Nov;11(10):611-9. PubMed PMID: 20961377.
62. Popper SJ, Sarr AD, Gueye-Ndiaye A, Mboup S, Essex ME, Kanki PJ. Low plasma human immunodeficiency virus type 2 viral load is independent of proviral load: low virus production in

vivo. Journal of virology. 2000 Feb;74(3):1554-7. PubMed PMID: 10627569. Pubmed Central PMCID: 111493.
63. Nuvor SV, van der Sande M, Rowland-Jones S, Whittle H, Jaye A. Natural killer cell function is well preserved in asymptomatic human immunodeficiency virus type 2 (HIV-2) infection but similar to that of HIV-1 infection when CD4 T-cell counts fall. Journal of virology. 2006 Mar;80(5):2529-38. PubMed PMID: 16474159. Pubmed Central PMCID: 1395408.
64. Jaffar S, Van der Loeff MS, Eugen-Olsen J, Vincent T, Sarje-Njie R, Ngom P, et al. Immunological predictors of survival in HIV type 2-infected rural villagers in Guinea-Bissau. AIDS research and human retroviruses. 2005 Jun;21(6):560-4. PubMed PMID: 15989461.
65. Campbell-Yesufu OT, Gandhi RT. Update on human immunodeficiency virus (HIV)-2 infection. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2011 Mar 15;52(6):780-7. PubMed PMID: 21367732. Pubmed Central PMCID: 3106263.
66. Schindler M, Munch J, Kutsch O, Li H, Santiago ML, Bibollet-Ruche F, et al. Nef-mediated suppression of T cell activation was lost in a lentiviral lineage that gave rise to HIV-1. Cell. 2006 Jun 16;125(6):1055-67. PubMed PMID: 16777597.
67. Shi Y, Brandin E, Vincic E, Jansson M, Blaxhult A, Gyllensten K, et al. Evolution of human immunodeficiency virus type 2 coreceptor usage, autologous neutralization, envelope sequence and glycosylation. The Journal of general virology. 2005 Dec;86(Pt 12):3385-96. PubMed PMID: 16298986.
68. Ozkaya Sahin G, Holmgren B, da Silva Z, Nielsen J, Nowroozalizadeh S, Esbjornsson J, et al. Potent intratype neutralizing activity distinguishes human immunodeficiency virus type 2 (HIV-2) from HIV-1. Journal of virology. 2012 Jan;86(2):961-71. PubMed PMID: 22072782. Pubmed Central PMCID: 3255861.
69. Nyamweya S, Hegedus A, Jaye A, Rowland-Jones S, Flanagan KL, Macallan DC. Comparing HIV-1 and HIV-2 infection: Lessons for viral immunopathogenesis. Rev Med Virol. 2013 Jul;23(4):221-40. PubMed PMID: 23444290.
70. Hrecka K, Hao C, Gierszewska M, Swanson SK, Kesik-Brodacka M, Srivastava S, et al. Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. Nature. 2011 Jun 29;474(7353):658-61. PubMed PMID: 21720370. Pubmed Central PMCID: 3179858.
71. Schaller T, Goujon C, Malim MH. AIDS/HIV. HIV interplay with SAMHD1. Science. 2012 Mar 16;335(6074):1313-4. PubMed PMID: 22422971.
72. Manel N, Hogstad B, Wang Y, Levy DE, Unutmaz D, Littman DR. A cryptic sensor for HIV-1 activates antiviral innate immunity in dendritic cells. Nature. 2010 Sep 9;467(7312):214-7. PubMed PMID: 20829794. Pubmed Central PMCID: 3051279.
73. O'Donovan D, Ariyoshi K, Milligan P, Ota M, Yamuah L, Sarge-Njie R, et al. Maternal plasma viral RNA levels determine marked differences in mother-to-child transmission rates of HIV-1 and HIV-2 in The Gambia. MRC/Gambia Government/University College London Medical School working group on mother-child transmission of HIV. Aids. 2000 Mar 10;14(4):441-8. PubMed PMID: 10770548.
74. Gottlieb GS, Hawes SE, Agne HD, Stern JE, Critchlow CW, Kiviat NB, et al. Lower levels of HIV RNA in semen in HIV-2 compared with HIV-1 infection: implications for differences in transmission. Aids. 2006 Apr 4;20(6):895-900. PubMed PMID: 16549974. Pubmed Central PMCID: 3726185.
75. Kanki PJ, Travers KU, S MB, Hsieh CC, Marlink RG, Gueye NA, et al. Slower heterosexual spread of HIV-2 than HIV-1. Lancet. 1994 Apr 16;343(8903):943-6. PubMed PMID: 7909009.
76. Adjorlolo-Johnson G, De Cock KM, Ekpini E, Vetter KM, Sibailly T, Brattegaard K, et al. Prospective comparison of mother-to-child transmission of HIV-1 and HIV-2 in Abidjan, Ivory Coast. JAMA : the journal of the American Medical Association. 1994 Aug 10;272(6):462-6. PubMed PMID: 8040982.
77. Popper SJ, Sarr AD, Travers KU, Gueye-Ndiaye A, Mboup S, Essex ME, et al. Lower human immunodeficiency virus (HIV) type 2 viral load reflects the difference in pathogenicity of HIV-1

and HIV-2. The Journal of infectious diseases. 1999 Oct;180(4):1116-21. PubMed PMID: 10479138.
78. Andersson S, Norrgren H, da Silva Z, Biague A, Bamba S, Kwok S, et al. Plasma viral load in HIV-1 and HIV-2 singly and dually infected individuals in Guinea-Bissau, West Africa: significantly lower plasma virus set point in HIV-2 infection than in HIV-1 infection. Archives of internal medicine. 2000 Nov 27;160(21):3286-93. PubMed PMID: 11088091.
79. van der Loeff MF, Larke N, Kaye S, Berry N, Ariyoshi K, Alabi A, et al. Undetectable plasma viral load predicts normal survival in HIV-2-infected people in a West African village. Retrovirology. 2010 May 19;7:46. PubMed PMID: 20482865. Pubmed Central PMCID: 2887382.
80. Schim van der Loeff MF, Jaffar S, Aveika AA, Sabally S, Corrah T, Harding E, et al. Mortality of HIV-1, HIV-2 and HIV-1/HIV-2 dually infected patients in a clinic-based cohort in The Gambia. Aids. 2002 Sep 6;16(13):1775-83. PubMed PMID: 12218389.
81. Berry N, Jaffar S, Schim van der Loeff M, Ariyoshi K, Harding E, N'Gom PT, et al. Low level viremia and high CD4% predict normal survival in a cohort of HIV type-2-infected villagers. AIDS research and human retroviruses. 2002 Nov 1;18(16):1167-73. PubMed PMID: 12487822.
82. Marlink R, Kanki P, Thior I, Travers K, Eisen G, Siby T, et al. Reduced rate of disease development after HIV-2 infection as compared to HIV-1. Science. 1994 Sep 9;265(5178):1587-90. PubMed PMID: 7915856.
83. Martinez-Steele E, Awasana AA, Corrah T, Sabally S, van der Sande M, Jaye A, et al. Is HIV-2- induced AIDS different from HIV-1-associated AIDS? Data from a West African clinic. Aids. 2007 Jan 30;21(3):317-24. PubMed PMID: 17255738.
84. Ariyoshi K, Schim van der Loeff M, Cook P, Whitby D, Corrah T, Jaffar S, et al. Kaposi's sarcoma in the Gambia, West Africa is less frequent in human immunodeficiency virus type 2 than in human immunodeficiency virus type 1 infection despite a high prevalence of human herpesvirus 8. Journal of human virology. 1998 Mar-Apr;1(3):193-9. PubMed PMID: 10195242.
85. Blaak H, van der Ende ME, Boers PH, Schuitemaker H, Osterhaus AD. In vitro replication capacity of HIV-2 variants from long-term aviremic individuals. Virology. 2006 Sep 15;353(1):144-54. PubMed PMID: 16814357.
86. Ota MO, O'Donovan D, Alabi AS, Milligan P, Yamuah LK, N'Gom PT, et al. Maternal HIV-1 and HIV-2 infection and child survival in The Gambia. Aids. 2000 Mar 10;14(4):435-9. PubMed PMID: 10770547.
87. Grassly NC, Xiang Z, Ariyoshi K, Aaby P, Jensen H, Schim van der Loeff M, et al. Mortality among human immunodeficiency virus type 2-positive villagers in rural Guinea-Bissau is correlated with viral genotype. Journal of virology. 1998 Oct;72(10):7895-9. PubMed PMID: 9733826. Pubmed Central PMCID: 110115.
88. Onyango CO, Leligdowicz A, Yokoyama M, Sato H, Song H, Nakayama EE, et al. HIV-2 capsids distinguish high and low virus load patients in a West African community cohort. Vaccine. 2010 May 26;28 Suppl 2:B60-7. PubMed PMID: 20510746.
89. Song H, Nakayama EE, Yokoyama M, Sato H, Levy JA, Shioda T. A single amino acid of the human immunodeficiency virus type 2 capsid affects its replication in the presence of cynomolgus monkey and human TRIM5alphas. Journal of virology. 2007 Jul;81(13):7280-5. PubMed PMID: 17475650. Pubmed Central PMCID: PMC1933286. Epub 2007/05/04. eng.
90. Central Intelligence Agency U. Guinea Bissau 2016 [cited 2016]. Available from: https://www.cia.gov/library/publications/the-world-factbook/geos/pu.html.
91. Company BB. Guinea Bissau Profile-Timeline 2016 [cited 2006]. Available from: http://www.bbc.co.uk/news/world-africa-13579838.
92. Bank TW. Guinea Bissau Country Profile 2016 [cited 2016]. Available from: http://www.worldbank.org/en/country/guineabissau.
93. Naucler A, Andreasson PA, Costa CM, Thorstensson R, Biberfeld G. HIV-2-associated AIDS and HIV-2 seroprevalence in Bissau, Guinea-Bissau. Journal of acquired immune deficiency syndromes (1999). 1989;2(1):88-93. PubMed PMID: 2918464.

94. Verdonck K, Gonzalez E, Van Dooren S, Vandamme AM, Vanham G, Gotuzzo E. Human T-lymphotropic virus 1: recent knowledge about an ancient infection. The Lancet infectious diseases. 2007 Apr;7(4):266-81. PubMed PMID: 17376384.
95. van Tienen C, McConkey SJ, de Silva TI, Cotten M, Kaye S, Sarge-Njie R, et al. Maternal proviral load and vertical transmission of human T cell lymphotropic virus type 1 in Guinea-Bissau. AIDS research and human retroviruses. 2012 Jun;28(6):584-90. PubMed PMID: 22066980.
96. Wilkins A, Ricard D, Todd J, Whittle H, Dias F, Paulo Da Silva A. The epidemiology of HIV infection in a rural area of Guinea-Bissau. Aids. 1993 Aug;7(8):1119-22. PubMed PMID: 8397950.
97. Schim van der Loeff MF, Aaby P, Aryioshi K, Vincent T, Awasana AA, Da Costa C, et al. HIV-2 does not protect against HIV-1 infection in a rural community in Guinea-Bissau. Aids. 2001 Nov 23;15(17):2303-10. PubMed PMID: 11698704.
98. Tienen C, van der Loeff MS, Zaman SM, Vincent T, Sarge-Njie R, Peterson I, et al. Two distinct epidemics: the rise of HIV-1 and decline of HIV-2 infection between 1990 and 2007 in rural Guinea-Bissau. Journal of acquired immune deficiency syndromes (1999). 2010 Apr;53(5):640-7. PubMed PMID: 19841588.
99. Ricard D, Wilkins A, N'Gum PT, Hayes R, Morgan G, Da Silva AP, et al. The effects of HIV-2 infection in a rural area of Guinea-Bissau. Aids. 1994 Jul;8(7):977-82. PubMed PMID: 7946109.
100. Ariyoshi K, Berry N, Cham F, Jaffar S, Schim van der Loeff M, Jobe O, et al. Quantification of Human T-lymphotropic virus type I (HTLV-I) provirus load in a rural West African population: no enhancement of human immunodeficiency virus type 2 pathogenesis, but HTLV-I provirus load relates to mortality. The Journal of infectious diseases. 2003 Dec 1;188(11):1648-51. PubMed PMID: 14639534.
101. Lagarde E, Schim van der Loeff M, Enel C, Holmgren B, Dray-Spira R, Pison G, et al. Mobility and the spread of human immunodeficiency virus into rural areas of West Africa. International journal of epidemiology. 2003 Oct;32(5):744-52. PubMed PMID: 14559743.
102. Gourlay AJ, van Tienen C, Dave SS, Vincent T, Rowland-Jones SL, Glynn JR, et al. Clinical predictors cannot replace biological predictors in HIV-2 infection in a community setting in West Africa. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 2012 May;16(5):e337-43. PubMed PMID: 22387142. Pubmed Central PMCID: 3324712.
103. Yindom LM, Leligdowicz A, Martin MP, Gao X, Qi Y, Zaman SM, et al. Influence of HLA class I and HLA-KIR compound genotypes on HIV-2 infection and markers of disease progression in a Manjako community in West Africa. Journal of virology. 2010 Aug;84(16):8202-8. PubMed PMID: 20519398. Pubmed Central PMCID: 2916551.
104. de Silva TI, Peng Y, Leligdowicz A, Zaidi I, Li L, Griffin H, et al. Correlates of T-cell-mediated viral control and phenotype of CD8(+) T cells in HIV-2, a naturally contained human retroviral infection. Blood. 2013 May 23;121(21):4330-9. PubMed PMID: 23558015.
105. Fryer HR, Van Tienen C, Van Der Loeff MS, Aaby P, Da Silva ZJ, Whittle H, et al. Predicting the extinction of HIV-2 in rural Guinea-Bissau. Aids. 2015 Nov 28;29(18):2479-86. PubMed PMID: 26485121. Pubmed Central PMCID: 4645958.
106. Prince PD, Matser A, van Tienen C, Whittle HC, Schim van der Loeff MF. Mortality rates in people dually infected with HIV-1/2 and those infected with either HIV-1 or HIV-2: a systematic review and meta-analysis. Aids. 2014 Feb 20;28(4):549-58. PubMed PMID: 23921613.
107. Hansmann A, Schim van der Loeff MF, Kaye S, Awasana AA, Sarge-Njie R, O'Donovan D, et al. Baseline plasma viral load and CD4 cell percentage predict survival in HIV-1- and HIV-2-infected women in a community-based cohort in The Gambia. Journal of acquired immune deficiency syndromes (1999). 2005 Mar 1;38(3):335-41. PubMed PMID: 15735454.
108. Stremlau M, Owens CM, Perron MJ, Kiessling M, Autissier P, Sodroski J. The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. Nature. 2004 Feb 26;427(6977):848-53. PubMed PMID: 14985764.

109. Malim MH, Bieniasz PD. HIV Restriction Factors and Mechanisms of Evasion. Cold Spring Harb Perspect Med. 2012 May;2(5):a006940. PubMed PMID: 22553496. Pubmed Central PMCID: 3331687.
110. Malim MH, Emerman M. HIV-1 accessory proteins--ensuring viral survival in a hostile environment. Cell host & microbe. 2008 Jun 12;3(6):388-98. PubMed PMID: 18541215.
111. Strebel K. HIV accessory proteins versus host restriction factors. Current opinion in virology. 2013 Dec;3(6):692-9. PubMed PMID: 24246762. Pubmed Central PMCID: 3855913.
112. Bushman FD, Malani N, Fernandes J, D'Orso I, Cagney G, Diamond TL, et al. Host cell factors in HIV replication: meta-analysis of genome-wide studies. PLoS pathogens. 2009 May;5(5):e1000437. PubMed PMID: 19478882. Pubmed Central PMCID: 2682202.
113. Sawyer SL, Wu LI, Emerman M, Malik HS. Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. Proceedings of the National Academy of Sciences of the United States of America. 2005 Feb 22;102(8):2832-7. PubMed PMID: 15689398. Pubmed Central PMCID: 549489.
114. Stoye JP. Studies of endogenous retroviruses reveal a continuing evolutionary saga. Nature reviews Microbiology. 2012 May 08;10(6):395-406. PubMed PMID: 22565131.
115. Kirchhoff F. Immune evasion and counteraction of restriction factors by HIV-1 and other primate lentiviruses. Cell host & microbe. 2010 Jul 22;8(1):55-67. PubMed PMID: 20638642.
116. Zheng YH, Jeang KT, Tokunaga K. Host restriction factors in retroviral infection: promises in virus-host interaction. Retrovirology. 2012 Dec 20;9:112. PubMed PMID: 23254112. Pubmed Central PMCID: 3549941.
117. Ribeiro AC, Maia e Silva A, Santa-Marta M, Pombo A, Moniz-Pereira J, Goncalves J, et al. Functional analysis of Vif protein shows less restriction of human immunodeficiency virus type 2 by APOBEC3G. Journal of virology. 2005 Jan;79(2):823-33. PubMed PMID: 15613310. Pubmed Central PMCID: 538526.
118. Gaddis NC, Chertova E, Sheehy AM, Henderson LE, Malim MH. Comprehensive investigation of the molecular defect in vif-deficient human immunodeficiency virus type 1 virions. Journal of virology. 2003 May;77(10):5810-20. PubMed PMID: 12719574. Pubmed Central PMCID: 154025.
119. Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature. 2002 Aug 08;418(6898):646-50. PubMed PMID: 12167863.
120. Sheehy AM, Gaddis NC, Malim MH. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. Nature medicine. 2003 Nov;9(11):1404-7. PubMed PMID: 14528300.
121. Kao S, Khan MA, Miyagi E, Plishka R, Buckler-White A, Strebel K. The human immunodeficiency virus type 1 Vif protein reduces intracellular expression and inhibits packaging of APOBEC3G (CEM15), a cellular inhibitor of virus infectivity. Journal of virology. 2003 Nov;77(21):11398-407. PubMed PMID: 14557625. Pubmed Central PMCID: 229358.
122. Shindo K, Takaori-Kondo A, Kobayashi M, Abudu A, Fukunaga K, Uchiyama T. The enzymatic activity of CEM15/Apobec-3G is essential for the regulation of the infectivity of HIV-1 virion but not a sole determinant of its antiviral activity. The Journal of biological chemistry. 2003 Nov 07;278(45):44412-6. PubMed PMID: 12970355.
123. Zhang H, Yang B, Pomerantz RJ, Zhang C, Arunachalam SC, Gao L. The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. Nature. 2003 Jul 03;424(6944):94-8. PubMed PMID: 12808465. Pubmed Central PMCID: 1350966.
124. Refsland EW, Stenglein MD, Shindo K, Albin JS, Brown WL, Harris RS. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. Nucleic acids research. 2010 Jul;38(13):4274-84. PubMed PMID: 20308164. Pubmed Central PMCID: 2910054.

125. Koning FA, Newman EN, Kim EY, Kunstman KJ, Wolinsky SM, Malim MH. Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. Journal of virology. 2009 Sep;83(18):9474-85. PubMed PMID: 19587057. Pubmed Central PMCID: 2738220.
126. Mehta HV, Jones PH, Weiss JP, Okeoma CM. IFN-alpha and lipopolysaccharide upregulate APOBEC3 mRNA through different signaling pathways. Journal of immunology. 2012 Oct 15;189(8):4088-103. PubMed PMID: 22972924. Pubmed Central PMCID: 3531240.
127. Miyagi E, Opi S, Takeuchi H, Khan M, Goila-Gaur R, Kao S, et al. Enzymatically active APOBEC3G is required for efficient inhibition of human immunodeficiency virus type 1. Journal of virology. 2007 Dec;81(24):13346-53. PubMed PMID: 17928335. Pubmed Central PMCID: 2168852.
128. Jin X, Brooks A, Chen H, Bennett R, Reichman R, Smith H. APOBEC3G/CEM15 (hA3G) mRNA levels associate inversely with human immunodeficiency virus viremia. Journal of virology. 2005 Sep;79(17):11513-6. PubMed PMID: 16103203. Pubmed Central PMCID: 1193574.
129. Navarro F, Bollman B, Chen H, Konig R, Yu Q, Chiles K, et al. Complementary function of the two catalytic domains of APOBEC3G. Virology. 2005 Mar 15;333(2):374-86. PubMed PMID: 15721369.
130. Schumacher AJ, Hache G, Macduff DA, Brown WL, Harris RS. The DNA deaminase activity of human APOBEC3G is required for Ty1, MusD, and human immunodeficiency virus type 1 restriction. Journal of virology. 2008 Mar;82(6):2652-60. PubMed PMID: 18184715. Pubmed Central PMCID: 2259018.
131. Chen KM, Harjes E, Gross PJ, Fahmy A, Lu Y, Shindo K, et al. Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. Nature. 2008 Mar 06;452(7183):116-9. PubMed PMID: 18288108.
132. Luo K, Liu B, Xiao Z, Yu Y, Yu X, Gorelick R, et al. Amino-terminal region of the human immunodeficiency virus type 1 nucleocapsid is required for human APOBEC3G packaging. Journal of virology. 2004 Nov;78(21):11841-52. PubMed PMID: 15479826. Pubmed Central PMCID: 523292.
133. Harris RS, Liddament MT. Retroviral restriction by APOBEC proteins. Nature reviews Immunology. 2004 Nov;4(11):868-77. PubMed PMID: 15516966.
134. Schafer A, Bogerd HP, Cullen BR. Specific packaging of APOBEC3G into HIV-1 virions is mediated by the nucleocapsid domain of the gag polyprotein precursor. Virology. 2004 Oct 25;328(2):163-8. PubMed PMID: 15464836.
135. Zennou V, Perez-Caballero D, Gottlinger H, Bieniasz PD. APOBEC3G incorporation into human immunodeficiency virus type 1 particles. Journal of virology. 2004 Nov;78(21):12058-61. PubMed PMID: 15479846. Pubmed Central PMCID: 523273.
136. Autore F, Bergeron JR, Malim MH, Fraternali F, Huthoff H. Rationalisation of the differences between APOBEC3G structures from crystallography and NMR studies by molecular dynamics simulations. PloS one. 2010 Jul 12;5(7):e11515. PubMed PMID: 20635000. Pubmed Central PMCID: 2902501.
137. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature. 2003 Jul 03;424(6944):99-103. PubMed PMID: 12808466.
138. Sato K, Izumi T, Misawa N, Kobayashi T, Yamashita Y, Ohmichi M, et al. Remarkable lethal G-to-A mutations in vif-proficient HIV-1 provirus by individual APOBEC3 proteins in humanized mice. Journal of virology. 2010 Sep;84(18):9546-56. PubMed PMID: 20610708. Pubmed Central PMCID: 2937654.
139. Yu Q, Konig R, Pillai S, Chiles K, Kearney M, Palmer S, et al. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. Nature structural & molecular biology. 2004 May;11(5):435-42. PubMed PMID: 15098018.
140. Beale RC, Petersen-Mahrt SK, Watt IN, Harris RS, Rada C, Neuberger MS. Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with

mutation spectra in vivo. Journal of molecular biology. 2004 Mar 26;337(3):585-96. PubMed PMID: 15019779.
141. Wiegand HL, Doehle BP, Bogerd HP, Cullen BR. A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. The EMBO journal. 2004 Jun 16;23(12):2451-8. PubMed PMID: 15152192. Pubmed Central PMCID: 423288.
142. Holmes RK, Koning FA, Bishop KN, Malim MH. APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. The Journal of biological chemistry. 2007 Jan 26;282(4):2587-95. PubMed PMID: 17121840.
143. Shirakawa K, Takaori-Kondo A, Kobayashi M, Tomonaga M, Izumi T, Fukunaga K, et al. Ubiquitination of APOBEC3 proteins by the Vif-Cullin5-ElonginB-ElonginC complex. Virology. 2006 Jan 20;344(2):263-6. PubMed PMID: 16303161.
144. Yu X, Yu Y, Liu B, Luo K, Kong W, Mao P, et al. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. Science. 2003 Nov 07;302(5647):1056-60. PubMed PMID: 14564014.
145. Zheng YH, Irwin D, Kurosu T, Tokunaga K, Sata T, Peterlin BM. Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. Journal of virology. 2004 Jun;78(11):6073-6. PubMed PMID: 15141007. Pubmed Central PMCID: 415831.
146. Abada P, Noble B, Cannon PM. Functional domains within the human immunodeficiency virus type 2 envelope protein required to enhance virus production. Journal of virology. 2005 Mar;79(6):3627-38. PubMed PMID: 15731257. Pubmed Central PMCID: 1075700.
147. Bishop KN, Verma M, Kim EY, Wolinsky SM, Malim MH. APOBEC3G inhibits elongation of HIV-1 reverse transcripts. PLoS pathogens. 2008 Dec;4(12):e1000231. PubMed PMID: 19057663. Pubmed Central PMCID: 2584787.
148. Smith JL, Izumi T, Borbet TC, Hagedorn AN, Pathak VK. HIV-1 and HIV-2 Vif Interact with Human APOBEC3 Proteins Using Completely Different Determinants. Journal of virology. 2014 Sep;88(17):9893-908. PubMed PMID: WOS:000341232300035.
149. Gallo SA, Reeves JD, Garg H, Foley B, Doms RW, Blumenthal R. Kinetic studies of HIV-1 and HIV-2 envelope glycoprotein-mediated fusion. Retrovirology. 2006 Dec 04;3:90. PubMed PMID: 17144914. Pubmed Central PMCID: 1693918.
150. Exline CM, Yang SJ, Haworth KG, Rengarajan S, Lopez LA, Droniou ME, et al. Determinants in HIV-2 Env and tetherin required for functional interaction. Retrovirology. 2015 Aug 07;12:67. PubMed PMID: 26248668. Pubmed Central PMCID: 4528709.
151. Neil SJ, Sandrin V, Sundquist WI, Bieniasz PD. An interferon-alpha-induced tethering mechanism inhibits HIV-1 and Ebola virus particle release but is counteracted by the HIV-1 Vpu protein. Cell host & microbe. 2007 Sep 13;2(3):193-203. PubMed PMID: 18005734. Pubmed Central PMCID: 3793644.
152. Neil SJ, Zang T, Bieniasz PD. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature. 2008 Jan 24;451(7177):425-30. PubMed PMID: 18200009.
153. Masuyama N, Kuronita T, Tanaka R, Muto T, Hirota Y, Takigawa A, et al. HM1.24 is internalized from lipid rafts by clathrin-mediated endocytosis through interaction with alpha-adaptin. The Journal of biological chemistry. 2009 Jun 05;284(23):15927-41. PubMed PMID: 19359243. Pubmed Central PMCID: 2708888.
154. Dietrich I, Hosie MJ, Willett BJ. The role of BST2/tetherin in feline retrovirus infection. Veterinary Immunology and Immunopathology. 2011 10/15/;143(3–4):255-64.
155. Andrew AJ, Miyagi E, Kao S, Strebel K. The formation of cysteine-linked dimers of BST-2/tetherin is important for inhibition of HIV-1 virus release but not for sensitivity to Vpu. Retrovirology. 2009 Sep 08;6:80. PubMed PMID: 19737401. Pubmed Central PMCID: 2754425.
156. Kupzig S, Korolchuk V, Rollason R, Sugden A, Wilde A, Banting G. Bst-2/HM1.24 is a raft-associated apical membrane protein with an unusual topology. Traffic. 2003 Oct;4(10):694-709. PubMed PMID: 12956872.

157. Hinz A, Miguet N, Natrajan G, Usami Y, Yamanaka H, Renesto P, et al. Structural basis of HIV-1 tethering to membranes by the BST-2/tetherin ectodomain. Cell host & microbe. 2010 Apr 22;7(4):314-23. PubMed PMID: 20399176. Pubmed Central PMCID: 2859121.
158. Dube M, Roy BB, Guiot-Guillain P, Binette J, Mercier J, Chiasson A, et al. Antagonism of tetherin restriction of HIV-1 release by Vpu involves binding and sequestration of the restriction factor in a perinuclear compartment. PLoS pathogens. 2010 Apr 08;6(4):e1000856. PubMed PMID: 20386718. Pubmed Central PMCID: 2851737.
159. Perez-Caballero D, Zang T, Ebrahimi A, McNatt MW, Gregory DA, Johnson MC, et al. Tetherin inhibits HIV-1 release by directly tethering virions to cells. Cell. 2009 Oct 30;139(3):499-511. PubMed PMID: 19879838. Pubmed Central PMCID: 2844890.
160. Yin X, Hu Z, Gu Q, Wu X, Zheng YH, Wei P, et al. Equine tetherin blocks retrovirus release and its activity is antagonized by equine infectious anemia virus envelope protein. Journal of virology. 2014 Jan;88(2):1259-70. PubMed PMID: 24227834. Pubmed Central PMCID: 3911658.
161. Matsuda A, Suzuki Y, Honda G, Muramatsu S, Matsuzaki O, Nagano Y, et al. Large-scale identification and characterization of human genes that activate NF-kappaB and MAPK signaling pathways. Oncogene. 2003 May 22;22(21):3307-18. PubMed PMID: 12761501.
162. Hotter D, Sauter D, Kirchhoff F. Emerging Role of the Host Restriction Factor Tetherin in Viral Immune Sensing. Journal of molecular biology. 2013 12/13/;425(24):4956-64.
163. Le Tortorec A, Neil SJ. Antagonism to and intracellular sequestration of human tetherin by the human immunodeficiency virus type 2 envelope glycoprotein. Journal of virology. 2009 Nov;83(22):11966-78. PubMed PMID: 19740980. Pubmed Central PMCID: 2772693.
164. Hauser H, Lopez LA, Yang SJ, Oldenburg JE, Exline CM, Guatelli JC, et al. HIV-1 Vpu and HIV-2 Env counteract BST-2/tetherin by sequestration in a perinuclear compartment. Retrovirology. 2010 Jun 07;7:51. PubMed PMID: 20529266. Pubmed Central PMCID: 2890665.
165. Noble B, Abada P, Nunez-Iglesias J, Cannon PM. Recruitment of the adaptor protein 2 complex by the human immunodeficiency virus type 2 envelope protein is necessary for high levels of virus release. Journal of virology. 2006 Mar;80(6):2924-32. PubMed PMID: 16501101. Pubmed Central PMCID: 1395462.
166. Le Tortorec A, Willey S, Neil SJ. Antiviral inhibition of enveloped virus release by tetherin/BST-2: action and counteraction. Viruses. 2011 May;3(5):520-40. PubMed PMID: 21994744. Pubmed Central PMCID: 3185764.
167. Planelles V, Barker E. Roles of Vpr and Vpx in modulating the virus-host cell relationship. Mol Aspects Med. 2010 Oct;31(5):398-406. PubMed PMID: 20558198. Pubmed Central PMCID: 2967654.
168. Tristem M, Marshall C, Karpas A, Petrik J, Hill F. Origin of vpx in lentiviruses. Nature. 1990 Sep 27;347(6291):341-2. PubMed PMID: 2145513.
169. Fujita M, Otsuka M, Nomaguchi M, Adachi A. Multifaceted activity of HIV Vpr/Vpx proteins: the current view of their virological functions. Rev Med Virol. 2010 Mar;20(2):68-76. PubMed PMID: 20069611.
170. Goujon C, Riviere L, Jarrosson-Wuilleme L, Bernaud J, Rigal D, Darlix JL, et al. SIVSM/HIV-2 Vpx proteins promote retroviral escape from a proteasome-dependent restriction pathway present in human dendritic cells. Retrovirology. 2007 Jan 09;4:2. PubMed PMID: 17212817. Pubmed Central PMCID: 1779362.
171. Baldauf HM, Pan X, Erikson E, Schmidt S, Daddacha W, Burggraf M, et al. SAMHD1 restricts HIV-1 infection in resting CD4 + T cells. Nature medicine. 2012;18(11):1682-7.
172. Bonifati S, Daly MB, St Gelais C, Kim SH, Hollenbaugh JA, Shepard C, et al. SAMHD1 controls cell cycle status, apoptosis and HIV-1 infection in monocytic THP-1 cells. Virology. 2016 Aug;495:92-100. PubMed PMID: 27183329. Pubmed Central PMCID: 4912869.
173. St Gelais C, de Silva S, Amie SM, Coleman CM, Hoy H, Hollenbaugh JA, et al. SAMHD1 restricts HIV-1 infection in dendritic cells (DCs) by dNTP depletion, but its expression in DCs and

primary CD4+ T-lymphocytes cannot be upregulated by interferons. Retrovirology. 2012 Dec 11;9:105. PubMed PMID: 23231760. Pubmed Central PMCID: 3527137.
174. Sharova N, Wu Y, Zhu X, Stranska R, Kaushik R, Sharkey M, et al. Primate lentiviral Vpx commandeers DDB1 to counteract a macrophage restriction. PLoS pathogens. 2008 May;4(5):e1000057. PubMed PMID: 18451984. Pubmed Central PMCID: 2323106.
175. Hrecka K, Hao C, Gierszewska M, Swanson SK, Kesik-Brodacka M, Srivastava S, et al. Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. Nature. 2011;474(7353):658-61.
176. Yan J, Kaur S, DeLucia M, Hao C, Mehrens J, Wang C, et al. Tetramerization of SAMHD1 is required for biological activity and inhibition of HIV infection. The Journal of biological chemistry. 2013 Apr 12;288(15):10406-17. PubMed PMID: 23426366. Pubmed Central PMCID: 3624423.
177. Beloglazova N, Flick R, Tchigvintsev A, Brown G, Popovic A, Nocek B, et al. Nuclease activity of the human SAMHD1 protein implicated in the Aicardi-Goutieres syndrome and HIV-1 restriction. The Journal of biological chemistry. 2013 Mar 22;288(12):8101-10. PubMed PMID: 23364794. Pubmed Central PMCID: 3605629.
178. White TE, Brandariz-Nunez A, Valle-Casuso JC, Amie S, Nguyen LA, Kim B, et al. The retroviral restriction ability of SAMHD1, but not its deoxynucleotide triphosphohydrolase activity, is regulated by phosphorylation. Cell host & microbe. 2013 Apr 17;13(4):441-51. PubMed PMID: 23601106. Pubmed Central PMCID: 3864637.
179. Endicott JA, Noble ME, Tucker JA. Cyclin-dependent kinases: inhibition and substrate recognition. Current opinion in structural biology. 1999 Dec;9(6):738-44. PubMed PMID: 10607671.
180. Nigg EA. Cellular substrates of p34(cdc2) and its companion cyclin-dependent kinases. Trends in cell biology. 1993 Sep;3(9):296-301. PubMed PMID: 14731846.
181. Kyei GB, Cheng X, Ramani R, Ratner L. Cyclin L2 Is a Critical HIV Dependency Factor in Macrophages that Controls SAMHD1 Abundance. Cell host & microbe. 2015 Jan 14;17(1):98-106. PubMed PMID: 25532805. Pubmed Central PMCID: 4297224.
182. Cheng X, Ratner L. HIV-2 Vpx protein interacts with interferon regulatory factor 5 (IRF5) and inhibits its function. The Journal of biological chemistry. 2014 Mar 28;289(13):9146-57. PubMed PMID: 24532789. Pubmed Central PMCID: 3979364.
183. Berger A, Munk C, Schweizer M, Cichutek K, Schule S, Flory E. Interaction of Vpx and apolipoprotein B mRNA-editing catalytic polypeptide 3 family member A (APOBEC3A) correlates with efficient lentivirus infection of monocytes. The Journal of biological chemistry. 2010 Apr 16;285(16):12248-54. PubMed PMID: 20178977. Pubmed Central PMCID: 2852964.
184. Love RP, Xu H, Chelico L. Biochemical analysis of hypermutation by the deoxycytidine deaminase APOBEC3A. The Journal of biological chemistry. 2012 Aug 31;287(36):30812-22. PubMed PMID: 22822074. Pubmed Central PMCID: 3436324.
185. Landry S, Narvaiza I, Linfesty DC, Weitzman MD. APOBEC3A can activate the DNA damage response and cause cell-cycle arrest. EMBO reports. 2011 May;12(5):444-50. PubMed PMID: 21460793. Pubmed Central PMCID: 3090015.
186. Mitra M, Hercik K, Byeon IJ, Ahn J, Hill S, Hinchee-Rodriguez K, et al. Structural determinants of human APOBEC3A enzymatic and nucleic acid binding properties. Nucleic acids research. 2014 Jan;42(2):1095-110. PubMed PMID: 24163103. Pubmed Central PMCID: 3902935.
187. Richardson SR, Narvaiza I, Planegger RA, Weitzman MD, Moran JV. APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. eLife. 2014 Apr 24;3:e02008. PubMed PMID: 24843014. Pubmed Central PMCID: 4003774.
188. Bogerd HP, Wiegand HL, Doehle BP, Lueders KK, Cullen BR. APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. Nucleic acids research. 2006;34(1):89-95. PubMed PMID: 16407327. Pubmed Central PMCID: 1326241.

189. Berger G, Durand S, Fargier G, Nguyen XN, Cordeil S, Bouaziz S, et al. APOBEC3A is a specific inhibitor of the early phases of HIV-1 infection in myeloid cells. PLoS pathogens. 2011 Sep;7(9):e1002221. PubMed PMID: 21966267. Pubmed Central PMCID: 3178557.
190. Pham P, Landolph A, Mendez C, Li N, Goodman MF. A biochemical analysis linking APOBEC3A to disparate HIV-1 restriction and skin cancer. The Journal of biological chemistry. 2013 Oct 11;288(41):29294-304. PubMed PMID: 23979356. Pubmed Central PMCID: 3795231.
191. Miyake A, Fujita M, Fujino H, Koga R, Kawamura S, Otsuka M, et al. Poly-proline motif in HIV-2 Vpx is critical for its efficient translation. The Journal of general virology. 2014 Jan;95(Pt 1):179-89. PubMed PMID: 24114794.
192. Ciftci HI, Fujino H, Koga R, Yamamoto M, Kawamura S, Tateishi H, et al. Mutational analysis of HIV-2 Vpx shows that proline residue 109 in the poly-proline motif regulates degradation of SAMHD1. FEBS letters. 2015 Jun 04;589(13):1505-14. PubMed PMID: 25936766.
193. Yu H, Usmani SM, Borch A, Kramer J, Sturzel CM, Khalid M, et al. The efficiency of Vpx-mediated SAMHD1 antagonism does not correlate with the potency of viral control in HIV-2-infected individuals. Retrovirology. 2013 Mar 05;10:27. PubMed PMID: 23497283. Pubmed Central PMCID: 3599662.
194. Stremlau M, Perron M, Welikala S, Sodroski J. Species-specific variation in the B30.2(SPRY) domain of TRIM5alpha determines the potency of human immunodeficiency virus restriction. Journal of virology. 2005 Mar;79(5):3139-45. PubMed PMID: 15709033. Pubmed Central PMCID: 548447.
195. Stremlau M, Perron M, Lee M, Li Y, Song B, Javanbakht H, et al. Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. Proceedings of the National Academy of Sciences of the United States of America. 2006 Apr 4;103(14):5514-9. PubMed PMID: 16540544. Pubmed Central PMCID: 1459386.
196. Javanbakht H, Diaz-Griffero F, Stremlau M, Si Z, Sodroski J. The contribution of RING and B-box 2 domains to retroviral restriction mediated by monkey TRIM5alpha. The Journal of biological chemistry. 2005 Jul 22;280(29):26933-40. PubMed PMID: 15897199.
197. Li X, Yeung DF, Fiegen AM, Sodroski J. Determinants of the higher order association of the restriction factor TRIM5alpha and other tripartite motif (TRIM) proteins. The Journal of biological chemistry. 2011 Aug 12;286(32):27959-70. PubMed PMID: 21680743. Pubmed Central PMCID: 3151041.
198. Li X, Sodroski J. The TRIM5alpha B-box 2 domain promotes cooperative binding to the retroviral capsid by mediating higher-order self-association. Journal of virology. 2008 Dec;82(23):11495-502. PubMed PMID: 18799578. Pubmed Central PMCID: 2583650.
199. Diaz-Griffero F, Qin XR, Hayashi F, Kigawa T, Finzi A, Sarnak Z, et al. A B-box 2 surface patch important for TRIM5alpha self-association, capsid binding avidity, and retrovirus restriction. Journal of virology. 2009 Oct;83(20):10737-51. PubMed PMID: 19656869. Pubmed Central PMCID: 2753111.
200. Diaz-Griffero F. Caging the beast: TRIM5alpha binding to the HIV-1 core. Viruses. 2011 May;3(5):423-8. PubMed PMID: 21994740. Pubmed Central PMCID: 3186010.
201. Li X, Song B, Xiang SH, Sodroski J. Functional interplay between the B-box 2 and the B30.2(SPRY) domains of TRIM5alpha. Virology. 2007 Sep 30;366(2):234-44. PubMed PMID: 17543365. Pubmed Central PMCID: 2040257.
202. Wagner JM, Roganowicz MD, Skorupka K, Alam SL, Christensen D, Doss G, et al. Mechanism of B-box 2 domain-mediated higher-order assembly of the retroviral restriction factor TRIM5alpha. eLife. 2016;5. PubMed PMID: 27253059. Pubmed Central PMCID: 4936894.
203. Li YL, Chandrasekaran V, Carter SD, Woodward CL, Christensen DE, Dryden KA, et al. Primate TRIM5 proteins form hexagonal nets on HIV-1 capsids. eLife. 2016;5. PubMed PMID: 27253068. Pubmed Central PMCID: 4936896.
204. Ganser-Pornillos BK, Chandrasekaran V, Pornillos O, Sodroski JG, Sundquist WI, Yeager M. Hexagonal assembly of a restricting TRIM5alpha protein. Proceedings of the National

Academy of Sciences of the United States of America. 2011 Jan 11;108(2):534-9. PubMed PMID: 21187419. Pubmed Central PMCID: 3021009.
205. Zhao G, Ke D, Vu T, Ahn J, Shah VB, Yang R, et al. Rhesus TRIM5alpha disrupts the HIV-1 capsid at the inter-hexamer interfaces. PLoS pathogens. 2011 Mar;7(3):e1002009. PubMed PMID: 21455494. Pubmed Central PMCID: 3063768.
206. Wu X, Anderson JL, Campbell EM, Joseph AM, Hope TJ. Proteasome inhibitors uncouple rhesus TRIM5alpha restriction of HIV-1 reverse transcription and infection. Proceedings of the National Academy of Sciences of the United States of America. 2006 May 9;103(19):7465-70. PubMed PMID: 16648264. Pubmed Central PMCID: 1464362.
207. Anderson JL, Campbell EM, Wu X, Vandegraaff N, Engelman A, Hope TJ. Proteasome inhibition reveals that a functional preintegration complex intermediate can be generated during restriction by diverse TRIM5 proteins. Journal of virology. 2006 Oct;80(19):9754-60. PubMed PMID: 16973579. Pubmed Central PMCID: 1617233.
208. Pertel T, Hausmann S, Morger D, Zuger S, Guerra J, Lascano J, et al. TRIM5 is an innate immune sensor for the retrovirus capsid lattice. Nature. 2011 Apr 21;472(7343):361-5. PubMed PMID: 21512573. Pubmed Central PMCID: 3081621.
209. Yap MW, Nisole S, Stoye JP. A single amino acid change in the SPRY domain of human Trim5alpha leads to HIV-1 restriction. Current biology : CB. 2005 Jan 11;15(1):73-8. PubMed PMID: 15649369.
210. Li Y, Li X, Stremlau M, Lee M, Sodroski J. Removal of arginine 332 allows human TRIM5alpha to bind human immunodeficiency virus capsids and to restrict infection. Journal of virology. 2006 Jul;80(14):6738-44. PubMed PMID: 16809279. Pubmed Central PMCID: 1489046.
211. Miyamoto T, Yokoyama M, Kono K, Shioda T, Sato H, Nakayama EE. A single amino acid of human immunodeficiency virus type 2 capsid protein affects conformation of two external loops and viral sensitivity to TRIM5alpha. PloS one. 2011;6(7):e22779. PubMed PMID: 21829511. Pubmed Central PMCID: PMC3145752. Epub 2011/08/11. eng.
212. Onyango CO, Leligdowicz A, Yokoyama M, Sato H, Song H, Nakayama EE, et al. HIV-2 capsids distinguish high and low virus load patients in a West African community cohort. Vaccine. 2010;28(SUPPL. 2):B60-B7.
213. Sardiello M, Cairo S, Fontanella B, Ballabio A, Meroni G. Genomic analysis of the TRIM family reveals two groups of genes with distinct evolutionary properties. BMC evolutionary biology. 2008 Aug 01;8:225. PubMed PMID: 18673550. Pubmed Central PMCID: 2533329.
214. Duan Z, Gao B, Xu W, Xiong S. Identification of TRIM22 as a RING finger E3 ubiquitin ligase. Biochemical and biophysical research communications. 2008 Sep 26;374(3):502-6. PubMed PMID: 18656448.
215. Torok M, Etkin LD. Two B or not two B? Overview of the rapidly expanding B-box family of proteins. Differentiation; research in biological diversity. 2001 Mar;67(3):63-71. PubMed PMID: 11428128.
216. Parry DA, Fraser RD, Squire JM. Fifty years of coiled-coils and alpha-helical bundles: a close relationship between sequence and structure. Journal of structural biology. 2008 Sep;163(3):258-69. PubMed PMID: 18342539.
217. Hattlmann CJ, Kelly JN, Barr SD. TRIM22: A Diverse and Dynamic Antiviral Protein. Molecular biology international. 2012;2012:153415. PubMed PMID: 22649727. Pubmed Central PMCID: 3356915.
218. Herr AM, Dressel R, Walter L. Different subcellular localisations of TRIM22 suggest species-specific function. Immunogenetics. 2009 Apr;61(4):271-80. PubMed PMID: 19212762. Pubmed Central PMCID: 3085756.
219. Sivaramakrishnan G, Sun Y, Rajmohan R, Lin VC. B30.2/SPRY domain in tripartite motif-containing 22 is essential for the formation of distinct nuclear bodies. FEBS letters. 2009 Jun 18;583(12):2093-9. PubMed PMID: 19481078.

220. Rajsbaum R, Stoye JP, O'Garra A. Type I interferon-dependent and -independent expression of tripartite motif proteins in immune cells. European journal of immunology. 2008 Mar;38(3):619-30. PubMed PMID: 18286572.
221. Gongora C, Tissot C, Cerdan C, Mechti N. The interferon-inducible Staf50 gene is downregulated during T cell costimulation by CD2 and CD28. Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research. 2000 Nov;20(11):955-61. PubMed PMID: 11096452.
222. Wang Y, Gao B, Xu W, Xiong S. BRG1 is indispensable for IFN-gamma-induced TRIM22 expression, which is dependent on the recruitment of IRF-1. Biochemical and biophysical research communications. 2011 Jul 08;410(3):549-54. PubMed PMID: 21683060.
223. Singh R, Gaiha G, Werner L, McKim K, Mlisana K, Luban J, et al. Association of TRIM22 with the type 1 interferon response and viral control during primary HIV-1 infection. Journal of virology. 2011 Jan;85(1):208-16. PubMed PMID: 20980524. Pubmed Central PMCID: 3014203.
224. Barr SD, Smiley JR, Bushman FD. The interferon response inhibits HIV particle production by induction of TRIM22. PLoS pathogens. 2008 Feb;4(2):e1000007. PubMed PMID: 18389079. Pubmed Central PMCID: 2279259.
225. Tissot C, Mechti N. Molecular cloning of a new interferon-induced factor that represses human immunodeficiency virus type 1 long terminal repeat expression. The Journal of biological chemistry. 1995 Jun 23;270(25):14891-8. PubMed PMID: 7797467.
226. Kajaste-Rudnitski A, Marelli SS, Pultrone C, Pertel T, Uchil PD, Mechti N, et al. TRIM22 inhibits HIV-1 transcription independently of its E3 ubiquitin ligase activity, Tat, and NF-kappaB-responsive long terminal repeat elements. Journal of virology. 2011 May;85(10):5183-96. PubMed PMID: 21345949. Pubmed Central PMCID: 3126207.
227. Bouazzaoui A, Kreutz M, Eisert V, Dinauer N, Heinzelmann A, Hallenberger S, et al. Stimulated trans-acting factor of 50 kDa (Staf50) inhibits HIV-1 replication in human monocyte-derived macrophages. Virology. 2006 Dec 5-20;356(1-2):79-94. PubMed PMID: 16926043.
228. Ghezzi S, Galli L, Kajaste-Rudnitski A, Turrini F, Marelli S, Toniolo D, et al. Identification of TRIM22 single nucleotide polymorphisms associated with loss of inhibition of HIV-1 transcription and advanced HIV-1 disease. Aids. 2013 Sep 24;27(15):2335-44. PubMed PMID: 23921607.
229. Leligdowicz A, Yindom LM, Onyango C, Sarge-Njie R, Alabi A, Cotten M, et al. Robust Gag-specific T cell responses characterize viremia control in HIV-2 infection. The Journal of clinical investigation. 2007 Oct;117(10):3067-74. PubMed PMID: 17823657. Pubmed Central PMCID: 1964515.
230. Gebo KA, Gallant JE, Keruly JC, Moore RD. Absolute CD4 vs. CD4 percentage for predicting the risk of opportunistic illness in HIV infection. Journal of acquired immune deficiency syndromes (1999). 2004 Aug 15;36(5):1028-33. PubMed PMID: 15247555.
231. Perron MJ, Stremlau M, Song B, Ulm W, Mulligan RC, Sodroski J. TRIM5alpha mediates the postentry block to N-tropic murine leukemia viruses in human cells. Proceedings of the National Academy of Sciences of the United States of America. 2004 Aug 10;101(32):11827-32. PubMed PMID: 15280539. Pubmed Central PMCID: 511059.
232. Yap MW, Nisole S, Lynch C, Stoye JP. Trim5alpha protein restricts both HIV-1 and murine leukemia virus. Proceedings of the National Academy of Sciences of the United States of America. 2004 Jul 20;101(29):10786-91. PubMed PMID: 15249690. Pubmed Central PMCID: 490012.
233. Hatziioannou T, Perez-Caballero D, Yang A, Cowan S, Bieniasz PD. Retrovirus resistance factors Ref1 and Lv1 are species-specific variants of TRIM5alpha. Proceedings of the National Academy of Sciences of the United States of America. 2004 Jul 20;101(29):10774-9. PubMed PMID: 15249685. Pubmed Central PMCID: 490010.
234. Ylinen LM, Keckesova Z, Wilson SJ, Ranasinghe S, Towers GJ. Differential restriction of human immunodeficiency virus type 2 and simian immunodeficiency virus SIVmac by

TRIM5alpha alleles. Journal of virology. 2005 Sep;79(18):11580-7. PubMed PMID: 16140735. Pubmed Central PMCID: 1212619.
235. Ozato K, Shin DM, Chang TH, Morse HC, 3rd. TRIM family proteins and their emerging roles in innate immunity. Nature reviews Immunology. 2008 Nov;8(11):849-60. PubMed PMID: 18836477. Pubmed Central PMCID: 3433745.
236. Carthagena L, Bergamaschi A, Luna JM, David A, Uchil PD, Margottin-Goguet F, et al. Human TRIM gene expression in response to interferons. PloS one. 2009;4(3):e4894. PubMed PMID: 19290053. Pubmed Central PMCID: 2654144.
237. Sawyer SL, Emerman M, Malik HS. Discordant evolution of the adjacent antiretroviral genes TRIM22 and TRIM5 in mammals. PLoS pathogens. 2007 Dec;3(12):e197. PubMed PMID: 18159944. Pubmed Central PMCID: 2151084.
238. Nakayama EE, Miyoshi H, Nagai Y, Shioda T. A specific region of 37 amino acid residues in the SPRY (B30.2) domain of African green monkey TRIM5alpha determines species-specific restriction of simian immunodeficiency virus SIVmac infection. Journal of virology. 2005 Jul;79(14):8870-7. PubMed PMID: 15994780. Pubmed Central PMCID: 1168783.
239. Perron MJ, Stremlau M, Sodroski J. Two surface-exposed elements of the B30.2/SPRY domain as potency determinants of N-tropic murine leukemia virus restriction by human TRIM5alpha. Journal of virology. 2006 Jun;80(11):5631-6. PubMed PMID: 16699044. Pubmed Central PMCID: 1472168.
240. Newman RM, Hall L, Connole M, Chen GL, Sato S, Yuste E, et al. Balancing selection and the evolution of functional polymorphism in Old World monkey TRIM5alpha. Proceedings of the National Academy of Sciences of the United States of America. 2006 Dec 12;103(50):19134-9. PubMed PMID: 17142324. Pubmed Central PMCID: 1679755.
241. Kono K, Bozek K, Domingues FS, Shioda T, Nakayama EE. Impact of a single amino acid in the variable region 2 of the Old World monkey TRIM5alpha SPRY (B30.2) domain on anti-human immunodeficiency virus type 2 activity. Virology. 2009 May 25;388(1):160-8. PubMed PMID: 19342071.
242. Nakayama EE, Shioda T. TRIM5alpha and Species Tropism of HIV/SIV. Frontiers in microbiology. 2012;3:13. PubMed PMID: 22291694. Pubmed Central PMCID: 3264904.
243. Speelmon EC, Livingston-Rosanoff D, Li SS, Vu Q, Bui J, Geraghty DE, et al. Genetic association of the antiviral restriction factor TRIM5alpha with human immunodeficiency virus type 1 infection. Journal of virology. 2006 Mar;80(5):2463-71. PubMed PMID: 16474153. Pubmed Central PMCID: 1395369.
244. Javanbakht H, An P, Gold B, Petersen DC, O'Huigin C, Nelson GW, et al. Effects of human TRIM5alpha polymorphisms on antiretroviral function and susceptibility to human immunodeficiency virus infection. Virology. 2006 Oct 10;354(1):15-27. PubMed PMID: 16887163.
245. Schim van der Loeff MF, Hansmann A, Awasana AA, Ota MO, O'Donovan D, Sarge-Njie R, et al. Survival of HIV-1 and HIV-2 perinatally infected children in The Gambia. Aids. 2003 Nov 7;17(16):2389-94. PubMed PMID: 14571192.
246. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012 11/01/print;491(7422):56-65.
247. Kent WJ. BLAT--the BLAST-like alignment tool. Genome research. 2002 Apr;12(4):656-64. PubMed PMID: 11932250. Pubmed Central PMCID: 187518.
248. Organization WH. What do we mean by "sex" and "gender"? 2016. Available from: http://apps.who.int/gender/whatisgender/en/index.html.
249. Brinkhof MW, Boulle A, Weigel R, Messou E, Mathers C, Orrell C, et al. Mortality of HIV-infected patients starting antiretroviral therapy in sub-Saharan Africa: comparison with HIV-unrelated mortality. PLoS medicine. 2009 Apr 28;6(4):e1000066. PubMed PMID: 19399157. Pubmed Central PMCID: 2667633.

250. Lee BW, Yap HK, Chew FT, Quah TC, Prabhakaran K, Chan GS, et al. Age- and sex-related changes in lymphocyte subpopulations of healthy Asian subjects: from birth to adulthood. Cytometry. 1996 Mar 15;26(1):8-15. PubMed PMID: 8809475.
251. Lisse IM, Aaby P, Whittle H, Jensen H, Engelmann M, Christensen LB. T-lymphocyte subsets in West African children: impact of age, sex, and season. The Journal of pediatrics. 1997 Jan;130(1):77-85. PubMed PMID: 9003854.
252. Uppal SS, Verma S, Dhot PS. Normal values of CD4 and CD8 lymphocyte subsets in healthy indian adults and the effects of sex, age, ethnicity, and smoking. Cytometry Part B, Clinical cytometry. 2003 Mar;52(1):32-6. PubMed PMID: 12599179.
253. Klein SL, Flanagan KL. Sex differences in immune responses. Nature reviews Immunology. 2016 Oct;16(10):626-38. PubMed PMID: 27546235.
254. Griesbeck M, Scully E, Altfeld M. Sex and gender differences in HIV-1 infection. Clinical science. 2016 Aug 1;130(16):1435-51. PubMed PMID: 27389589.
255. Sawyer SL, Wu LI, Akey JM, Emerman M, Malik HS. High-frequency persistence of an impaired allele of the retroviral defense gene TRIM5alpha in humans. Current biology : CB. 2006 Jan 10;16(1):95-100. PubMed PMID: 16401428. Epub 2006/01/13. eng.
256. van Manen D, Rits MA, Beugeling C, van Dort K, Schuitemaker H, Kootstra NA. The effect of Trim5 polymorphisms on the clinical course of HIV-1 infection. PLoS pathogens. 2008 Feb 8;4(2):e18. PubMed PMID: 18248091. Pubmed Central PMCID: 2222955.
257. Liu FL, Qiu YQ, Li H, Kuang YQ, Tang X, Cao G, et al. An HIV-1 resistance polymorphism in TRIM5alpha gene among Chinese intravenous drug users. Journal of acquired immune deficiency syndromes (1999). 2011 Apr;56(4):306-11. PubMed PMID: 21107267.
258. Price H, Lacap P, Tuff J, Wachihi C, Kimani J, Ball TB, et al. A TRIM5alpha exon 2 polymorphism is associated with protection from HIV-1 infection in the Pumwani sex worker cohort. Aids. 2010 Jul 31;24(12):1813-21. PubMed PMID: 20588169. Pubmed Central PMCID: 2921035.
259. Nakayama EE, Carpentier W, Costagliola D, Shioda T, Iwamoto A, Debre P, et al. Wild type and H43Y variant of human TRIM5alpha show similar anti-human immunodeficiency virus type 1 activity both in vivo and in vitro. Immunogenetics. 2007 Jun;59(6):511-5. PubMed PMID: 17406861. Epub 2007/04/05. eng.
260. Nakajima T, Nakayama EE, Kaur G, Terunuma H, Mimaya JI, Ohtani H, et al. Impact of novel TRIM5alpha variants, Gly110Arg and G176del, on the anti-HIV-1 activity and the susceptibility to HIV-1 infection. Aids. 2009 Oct 23;23(16):2091-100. PubMed PMID: 19710594. Epub 2009/08/28. eng.
261. Goldschmidt V, Bleiber G, May M, Martinez R, Ortiz M, Telenti A, et al. Role of common human TRIM5alpha variants in HIV-1 disease progression. Retrovirology. 2006;3:54. PubMed PMID: 16925802. Pubmed Central PMCID: 1560158.
262. Han K, Lou DI, Sawyer SL. Identification of a genomic reservoir for new TRIM genes in primate genomes. PLoS genetics. 2011 Dec;7(12):e1002388. PubMed PMID: 22144910. Pubmed Central PMCID: 3228819.
263. Meroni G, Diez-Roux G. TRIM/RBCC, a novel class of 'single protein RING finger' E3 ubiquitin ligases. BioEssays : news and reviews in molecular, cellular and developmental biology. 2005 Nov;27(11):1147-57. PubMed PMID: 16237670.
264. Zhang F, Hatziioannou T, Perez-Caballero D, Derse D, Bieniasz PD. Antiretroviral potential of human tripartite motif-5 and related proteins. Virology. 2006 Sep 30;353(2):396-409. PubMed PMID: 16828831.
265. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012 Nov 1;491(7422):56-65. PubMed PMID: 23128226. Pubmed Central PMCID: 3498066.

266. Yu S, Gao B, Duan Z, Xu W, Xiong S. Identification of tripartite motif-containing 22 (TRIM22) as a novel NF-kappaB activator. Biochemical and biophysical research communications. 2011 Jul 01;410(2):247-51. PubMed PMID: 21651891.
267. Li Q, Lee CH, Peters LA, Mastropaolo LA, Thoeni C, Elkadri A, et al. Variants in TRIM22 That Affect NOD2 Signaling Are Associated With Very-Early-Onset Inflammatory Bowel Disease. Gastroenterology. 2016 May;150(5):1196-207. PubMed PMID: 26836588. Pubmed Central PMCID: 4842103.
268. Ovsyannikova IG, Salk HM, Larrabee BR, Pankratz VS, Poland GA. Single nucleotide polymorphisms/haplotypes associated with multiple rubella-specific immune response outcomes post-MMR immunization in healthy children. Immunogenetics. 2015 Oct;67(10):547-61. PubMed PMID: 26329766. Pubmed Central PMCID: 4587390.
269. Kelly JN, Woods MW, Xhiku S, Barr SD. Ancient and recent adaptive evolution in the antiviral TRIM22 gene: identification of a single-nucleotide polymorphism that impacts TRIM22 function. Human mutation. 2014 Sep;35(9):1072-81. PubMed PMID: 24863734.
270. Ke X, Taylor MS, Cardon LR. Singleton SNPs in the human genome and implications for genome-wide association studies. European journal of human genetics : EJHG. 2008 Apr;16(4):506-15. PubMed PMID: 18197193.
271. Jordan DM, Ramensky VE, Sunyaev SR. Human allelic variation: perspective from protein function, structure, and evolution. Current opinion in structural biology. 2010 Jun;20(3):342-50. PubMed PMID: 20399638. Pubmed Central PMCID: 2921592.
272. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nature methods. 2010 Apr;7(4):248-9. PubMed PMID: 20354512. Pubmed Central PMCID: 2855889.
273. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature protocols. 2009;4(7):1073-81. PubMed PMID: 19561590.
274. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. PloS one. 2012;7(10):e46688. PubMed PMID: 23056405. Pubmed Central PMCID: 3466303.
275. Marsh JA, Teichmann SA. Relative solvent accessible surface area predicts protein conformational changes upon binding. Structure. 2011 Jun 8;19(6):859-67. PubMed PMID: 21645856. Pubmed Central PMCID: 3145976.
276. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology. 2011 Oct 11;7:539. PubMed PMID: 21988835. Pubmed Central PMCID: 3261699.
277. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic acids research. 2014 Jul;42(Web Server issue):W320-4. PubMed PMID: 24753421. Pubmed Central PMCID: 4086106.
278. Yang H, Ji X, Zhao G, Ning J, Zhao Q, Aiken C, et al. Structural insight into HIV-1 capsid recognition by rhesus TRIM5alpha. Proceedings of the National Academy of Sciences of the United States of America. 2012 Nov 6;109(45):18372-7. PubMed PMID: 23091002. Pubmed Central PMCID: 3494900.
279. Goldstone DC, Walker PA, Calder LJ, Coombs PJ, Kirkpatrick J, Ball NJ, et al. Structural studies of postentry restriction factors reveal antiparallel dimers that enable avid binding to the HIV-1 capsid lattice. Proceedings of the National Academy of Sciences of the United States of America. 2014 Jul 1;111(26):9609-14. PubMed PMID: 24979782. Pubmed Central PMCID: 4084454.
280. Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. Nature protocols. 2009;4(3):363-71. PubMed PMID: 19247286.
281. The PyMOL Molecular Graphics System VS, LLC.
282. Haynes WM. CRC Handbook of Chemistry and Physics. 97th Edition.2016-2017.

283. B R, C GP. Path to facilitate the prediction of functional amino acid substitutions in red blood cell disorders--a computational approach. PloS one. 2011;6(9):e24607. PubMed PMID: 21931771. Pubmed Central PMCID: 3172254.
284. Pham QT, Bouchard A, Grutter MG, Berthoux L. Generation of human TRIM5alpha mutants with high HIV-1 restriction activity. Gene Ther. 2010 Jul;17(7):859-71. PubMed PMID: 20357830.
285. Maillard PV, Reynard S, Serhan F, Turelli P, Trono D. Interfering residues narrow the spectrum of MLV restriction by human TRIM5alpha. PLoS pathogens. 2007 Dec 28;3(12):e200. PubMed PMID: 18166079. Pubmed Central PMCID: 2156100.
286. Kono K, Song H, Shingai Y, Shioda T, Nakayama EE. Comparison of anti-viral activity of rhesus monkey and cynomolgus monkey TRIM5alphas against human immunodeficiency virus type 2 infection. Virology. 2008 Apr 10;373(2):447-56. PubMed PMID: 18201746.
287. Sebastian S, Grutter C, Strambio de Castillia C, Pertel T, Olivari S, Grutter MG, et al. An invariant surface patch on the TRIM5alpha PRYSPRY domain is required for retroviral restriction but dispensable for capsid binding. Journal of virology. 2009 Apr;83(7):3365-73. PubMed PMID: 19153241. Pubmed Central PMCID: 2655600.
288. Arriagada G, Muntean LN, Goff SP. SUMO-interacting motifs of human TRIM5alpha are important for antiviral activity. PLoS pathogens. 2011 Apr;7(4):e1002019. PubMed PMID: 21490953. Pubmed Central PMCID: 3072370.
289. Laguette N, Sobhian B, Casartelli N, Ringeard M, Chable-Bessia C, Segeral E, et al. SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. Nature. 2011 May 25;474(7353):654-7. PubMed PMID: 21613998. Pubmed Central PMCID: 3595993.
290. Emerman M, Malik HS. Paleovirology--modern consequences of ancient viruses. PLoS biology. 2010 Feb 09;8(2):e1000301. PubMed PMID: 20161719. Pubmed Central PMCID: 2817711.
291. Lim ES, Malik HS, Emerman M. Ancient adaptive evolution of tetherin shaped the functions of Vpu and Nef in human immunodeficiency virus and primate lentiviruses. Journal of virology. 2010 Jul;84(14):7124-34. PubMed PMID: 20444900. Pubmed Central PMCID: 2898239.
292. Laguette N, Rahm N, Sobhian B, Chable-Bessia C, Munch J, Snoeck J, et al. Evolutionary and functional analyses of the interaction between the myeloid restriction factor SAMHD1 and the lentiviral Vpx protein. Cell host & microbe. 2012 Feb 16;11(2):205-17. PubMed PMID: 22305291. Pubmed Central PMCID: 3595996.
293. Liberatore RA, Bieniasz PD. Tetherin is a key effector of the antiretroviral activity of type I interferon in vitro and in vivo. Proceedings of the National Academy of Sciences of the United States of America. 2011 Nov 01;108(44):18097-101. PubMed PMID: 22025715. Pubmed Central PMCID: 3207693.
294. Goujon C, Schaller T, Pedro Galao R, Amie SM, Kim B, Olivieri K, et al. Evidence for IFN alpha-induced, SAMHD1-independent inhibitors of early HIV-1 infection. Retrovirology. 2013 Feb 25;10. PubMed PMID: WOS:000315892700001.
295. Hauser H, Lopez LA, Yang SJ, Oldenburg JE, Exline CM, Guatelli JC, et al. HIV-1 Vpu and HIV-2 Env counteract BST-2/tetherin by sequestration in a perinuclear compartment. Retrovirology. 2010 Jun 7;7. PubMed PMID: WOS:000279544600001.
296. Hrecka K, Hao C, Gierszewska M, Swanson SK, Kesik-Brodacka M, Srivastava S, et al. Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. Nature. 2011 Jun 30;474(7353):658-U137. PubMed PMID: WOS:000292204300043.
297. Le Tortorec A, Neil SJD. Antagonism to and Intracellular Sequestration of Human Tetherin by the Human Immunodeficiency Virus Type 2 Envelope Glycoprotein. Journal of virology. 2009 Nov 15;83(22):11966-78. PubMed PMID: WOS:000271084100054.

298. Guignard F, Combadiere C, Tiffany HL, Murphy PM. Gene organization and promoter function for CC chemokine receptor 5 (CCR5). Journal of immunology. 1998 Jan 15;160(2):985-92. PubMed PMID: 9551938.
299. Moriuchi H, Moriuchi M, Arthos J, Hoxie J, Fauci AS. Promonocytic U937 subclones expressing CD4 and CXCR4 are resistant to infection with and cell-to-cell fusion by T-cell-tropic human immunodeficiency virus type 1. Journal of virology. 1997 Dec;71(12):9664-71. PubMed PMID: 9371631. Pubmed Central PMCID: 230275.
300. Salahuddin SZ, Markham PD, Wong-Staal F, Franchini G, Kalyanaraman VS, Gallo RC. Restricted expression of human T-cell leukemia--lymphoma virus (HTLV) in transformed human umbilical cord blood lymphocytes. Virology. 1983 Aug;129(1):51-64. PubMed PMID: 6412453.
301. McLain L, Dimmock NJ. A human CD4+ T-cell line expresses functional CD64 (Fc gamma RI), CD32 (Fc gamma RII), and CD16 (Fc gamma RIII) receptors but these do not enhance the infectivity of HIV-1-IgG complexes. Immunology. 1997 Jan;90(1):109-14. PubMed PMID: 9038720. Pubmed Central PMCID: 1456708.
302. Gama-Norton L, Botezatu L, Herrmann S, Schweizer M, Alves PM, Hauser H, et al. Lentivirus production is influenced by SV40 large T-antigen and chromosomal integration of the vector in HEK293 cells. Human gene therapy. 2011 Oct;22(10):1269-79. PubMed PMID: 21554103.
303. Merten OW, Hebben M, Bovolenta C. Production of lentiviral vectors. Molecular therapy Methods & clinical development. 2016;3:16017. PubMed PMID: 27110581. Pubmed Central PMCID: 4830361.
304. Franzoso G, Biswas P, Poli G, Carlson LM, Brown KD, Tomita-Yamaguchi M, et al. A family of serine proteases expressed exclusively in myelo-monocytic cells specifically processes the nuclear factor-kappa B subunit p65 in vitro and may impair human immunodeficiency virus replication in these cells. The Journal of experimental medicine. 1994 Oct 01;180(4):1445-56. PubMed PMID: 7931077. Pubmed Central PMCID: 2191703.
305. Popovic M, Sarngadharan MG, Read E, Gallo RC. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. Science. 1984 May 04;224(4648):497-500. PubMed PMID: 6200935.
306. Yang H, Yorke E, Hancock G, Clutton G, Sande N, Angus B, et al. Improved quantification of HIV-1-infected CD4+ T cells using an optimised method of intracellular HIV-1 gag p24 antigen detection. Journal of immunological methods. 2013 May 31;391(1-2):174-8. PubMed PMID: 23500782.
307. McKnight A, Dittmar MT, Moniz-Periera J, Ariyoshi K, Reeves JD, Hibbitts S, et al. A broad range of chemokine receptors are used by primary isolates of human immunodeficiency virus type 2 as coreceptors with CD4. Journal of virology. 1998 May;72(5):4065-71. PubMed PMID: 9557695. Pubmed Central PMCID: 109635.
308. Schulz TF, Whitby D, Hoad JG, Corrah T, Whittle H, Weiss RA. Biological and molecular variability of human immunodeficiency virus type 2 isolates from The Gambia. Journal of virology. 1990 Oct;64(10):5177-82. PubMed PMID: 1975844. Pubmed Central PMCID: 248013 immunodeficiency virus (HIV) 1 and 2 are indistinguishable from each other, there is some evidence that HIV-2 isolates may have a longer incubation period. Thus, an investigation was conducted of the biological properties and molecular variability of the spectrum of HIV-2 isolates existing in The Gambia. Serum samples were obtained from 20 HIV-2-positive individuals attending a sexually transmitted diseases clinic in Fajara. Seven new HIV-2 isolates (CBL-20 to CBL-26) were identified. Each differed in terms of its growth rate, cytopathogenicity in vitro, and sensitivity to neutralizing antibodies in patient sera. In addition, there was a close association between the isolates' in vitro cytopathogenicity and the clinical cytopathogenic strains, were obtained from the two patients with acquired immunodeficiency syndrome (AIDS). In contrast, patients from whom CBL-25 and 26 were isolated have been asymptomatic for at least 3 years. CBL-22 was highly sensitive to neutralization by HIV-2 sera and was cross-neutralized by some

HIV-1 sera. It is speculated that the observed differences in sensitivity to neutralization reflect differences in antigenic epitope expression or the number and presentation of envelope glycoprotein molecules on an infectious virion. Analysis of the molecular weight of the envelope precursor and the outer envelope protein of various HIV-2 isolates revealed that all the CBL isolated and SBL-6669 have smaller envelope proteins than the prototype strain, LAV-2 ROD. Also observed were differences in the amino acid sequences of these HIV-2 isolates.
309. Wilson SJ, Webb BL, Maplanka C, Newman RM, Verschoor EJ, Heeney JL, et al. Rhesus macaque TRIM5 alleles have divergent antiretroviral specificities. Journal of virology. 2008 Jul;82(14):7243-7. PubMed PMID: 18480454. Pubmed Central PMCID: 2446970.
310. Cowan S, Hatziioannou T, Cunningham T, Muesing MA, Gottlinger HG, Bieniasz PD. Cellular inhibitors with Fv1-like activity restrict human and simian immunodeficiency virus tropism. Proceedings of the National Academy of Sciences of the United States of America. 2002 Sep 3;99(18):11914-9. PubMed PMID: 12154227. Pubmed Central PMCID: 129368.
311. Besnier C, Takeuchi Y, Towers G. Restriction of lentivirus in monkeys. Proceedings of the National Academy of Sciences of the United States of America. 2002 Sep 3;99(18):11920-5. PubMed PMID: 12154231. Pubmed Central PMCID: 129369.
312. Griffin SD, Allen JF, Lever AM. The major human immunodeficiency virus type 2 (HIV-2) packaging signal is present on all HIV-2 RNA species: cotranslational RNA encapsidation and limitation of Gag protein confer specificity. Journal of virology. 2001 Dec;75(24):12058-69. PubMed PMID: 11711596. Pubmed Central PMCID: 116101.
313. Negre D, Duisit G, Mangeot PE, Moullier P, Darlix JL, Cosset FL. Lentiviral vectors derived from simian immunodeficiency virus. Current topics in microbiology and immunology. 2002;261:53-74. PubMed PMID: 11892253.
314. Poeschla EM, Wong-Staal F, Looney DJ. Efficient transduction of nondividing human cells by feline immunodeficiency virus lentiviral vectors. Nature medicine. 1998 Mar;4(3):354-7. PubMed PMID: 9500613.
315. Mitrophanous K, Yoon S, Rohll J, Patil D, Wilkes F, Kim V, et al. Stable gene transfer to the nervous system using a non-primate lentiviral vector. Gene Ther. 1999 Nov;6(11):1808-18. PubMed PMID: 10602376.
316. Saenz DT, Teo W, Olsen JC, Poeschla EM. Restriction of feline immunodeficiency virus by Ref1, Lv1, and primate TRIM5alpha proteins. Journal of virology. 2005 Dec;79(24):15175-88. PubMed PMID: 16306589. Pubmed Central PMCID: 1316021.
317. Hatziioannou T, Cowan S, Goff SP, Bieniasz PD, Towers GJ. Restriction of multiple divergent retroviruses by Lv1 and Ref1. The EMBO journal. 2003 Feb 3;22(3):385-94. PubMed PMID: 12554640. Pubmed Central PMCID: 140727.
318. Diouf K, Sarr AD, Eisen G, Popper S, Mboup S, Kanki P. Associations between MHC class I and susceptibility to HIV-2 disease progression. Journal of human virology. 2002 Jan-Feb;5(1):1-7. PubMed PMID: 12352262.
319. International HIVCS, Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, et al. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science. 2010 Dec 10;330(6010):1551-7. PubMed PMID: 21051598. Pubmed Central PMCID: 3235490.