Investigating the Association of Host Regulatory Polymorphic Sites with HIV Infection

SABA ALTAF

NUST-00000172110

Supervisor

Dr ADEEB SHEHZAD

Co supervisor

Dr ANEELA JAVED

DEPARTMENT OF BIOMEDICAL SCIENCES

SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD

January, 2019

Investigating the Association of Host Regulatory Polymorphic Sites with HIV Infection

Author

SABA ALTAF

NUST-00000172110

A thesis submitted in partial fulfillment of the requirements for the degree of

MS Biomedical Sciences

Thesis Supervisor:

Dr ADEEB SHEHZAD

Co Supervisor:

Dr ANEELA JAVED

Thesis Supervisor's Signature: _____

Co supervisor Signature: _____

DEPARTMENT OF BIOMEDICAL SCIENCES

SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY, ISLAMABAD

January, 2019

Declaration

I certify that this research work titled "Investigating the Association of Host Regulatory Polymorphic Sites with HIV Infection" is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources it has been properly acknowledged / referred.

SABA ALTAF

2016-NUST-MS-BMES-00000172110

Plagiarism Certificate (Turnitin Report)

This thesis has been checked for Plagiarism. Turnitin report endorsed by Supervisor is attached.

SABA ALTAF

NUST-00000172110

Signature of Supervisor

Copyright Statement

Acknowledgements

ALHUMDULILLAH. all praise to ALLAH, the Lord of Heavens and the Earth… throughout my life, my Lord has given me everything that was better than what I have ever asked for, even when I was not expecting it. A lot of things did not work the way I planned but my Lord make them work out in a better way for me, Almighty did not leave me alone in darkest phase of my life. Indeed "what is meant for you, will reach you even if it is beneath two mountains, and what is not meant for you, won't reach even if it is between your two lips"

I am grateful to my ammi and abbu for being there by my side throughout my life. I would not be here without you people. You have given me wings….i love both of you billion.

I am grateful to my supervisor Dr Adeeb Shehzad and co supervisor Dr Aneela Javed for their support and patience throughout my thesis. It was privilege for me to work with u. I have learned a lot. Thank you for everything.

I would also like to thank my GEC members for being on my thesis guidance and evaluation committee and my senior Dr Uzma Malik and Dr Muneeba, my friends Tehmeena, Shagufta, Ajla ,Aneeqa, my lab fellows Sadia and Maham, Shama etc for their support and cooperation.

Last but not the least, I don't have words to say thank you enough Dr Saadiya Nazeer and you know it how grateful I am. Thank you for everything. You have helped me in time when nobody came forward thou it was not your responsibility even. Thank you very much for guiding me and for your patience.

Dedicated to my Ammi (aaana g) and Abbu

I would not have come out of my cocoon without your support, patience and prayers. You have given me wings. I owe my success (if any) to u.

Abstract

The study investigated the possible association between two polymorphic sites in regulatory region of host genome with HIV-1 infection. Polymorphism was analyzed in Pakistani patent and control samples by allele specific PCR. All HIV positive individuals were either treatment naïve or failed to respond to antiretroviral therapy. The polymorphic sites rs6457282 and rs17064977 being analyzed were proxy SNPs of HIV-1 associated reported SNPs rs7756521 (chromosome 6, HIV-1 control causal variant) and rs8099014 (chromosome 18) respectively. Increased risk of HIV-1 infection was observed with the *rs6457282 C/T* genotype, as compared to healthy controls (OR = 9.5, 95% CI =1.3119-2.6088 and p = 0.0005). Results suggested that C allele confers susceptibility to HIV-1 infection (OR>1) whereas T allele might have a protective effect with OR<1.

When *rs17064977C/T* polymorphic site was genotyped via allele specific PCR, results suggested its association with HIV susceptibility like *rs6457282 T/C* in our case-control study. The patient samples were found to be more heterozygous *(CT)* than control groups. The odd ratios suggested that *rs17064977C/T* heterozygous genotype may be a risk factor in HIV susceptibility (OR=8.1077, 95%CI=3.1125-21.119, P=0.0001) and the T allele may be a protective factor (OR<1) against HIV acquisition.

Key Words: HIV, Allele specific PCR, Regulatory region, Polymorphic sites.

Table of Contents

CHAPTER 1: INTRODUCTION

The research work in this dissertation is a wet lab validation of association of computationally identified regulatory polymorphic sites in host genome with HIV-1 infection. Selected novel proxy SNPs in host genome was analyzed by genotyping Pakistani HIV+ and uninfected representative sample to identify their association with HIV-1 disease.

1.1 Background

Common variant of Human immunodeficiency virus 1 first arise in late 80s as descendent of simian immunodeficiency virus transmitted through chimpanzees (Keele et al., 2006) via chimp/human encounter. Virus isolated in 1983 (Barré-Sinoussi, Ross, & Delfraissy, 2013) has affected more than 30 million people across the globe (Engelman & Cherepanov, 2012).

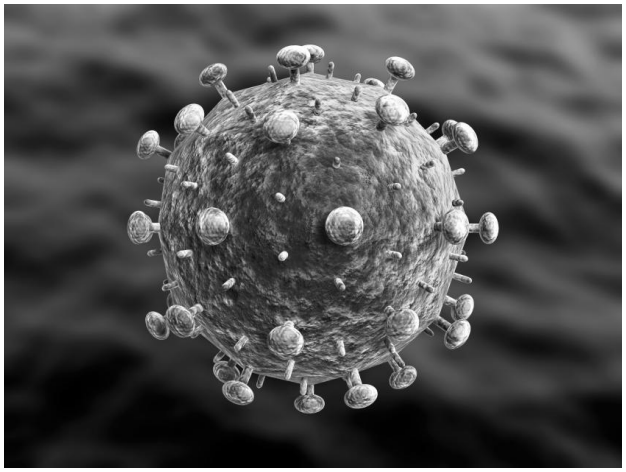


Figure 1.1: *HIV virus Photo, courtesy of Shutterstock.*

1.2 Epidemiology

According to UNAIDS global HIV report July 2018 approximately 36.9 million people had HIV globally with highest number of individuals in Eastern and Southern Africa (about 19.6 million) and 1.8 million new infection in year 2017 (Unaids, 2018).

Although UNAID has reported a decrease in number of AIDS death from 1.9 million back in 2005 to 1 million in 2016 and observed a downward trend of epidemic globally

(The Lancet, 2017) but region based advancement is inconstant, with certain parts showing an increased number of new infections due to lack of awareness and access to antiretroviral treatment (The Lancet, 2017).

1.3 HIV in Pakistan

In Pakistan HIV was first reported by individual and government reports in mid 80s (O. a Khan & Hyder, 2001) and first outbreak occurred in 2004 in district Larkana in community of injection drug users (Shah, Altaf, Mujeeb, & Memon, 2004).

Although The UNAIDS has reported a downward trend of HIV epidemic globally but over the years Pakistan has shown HIV concentrated epidemics in injecting drug users (IDUs) and among men who have sex with men (MSM) (Khanani et al., 2011). The infection is spreading at higher rate from IDUs and heterosexual MSM high risk groups to their spouses and children (non-core group) (Khanani et al., 2011). The common HIV 1 sub type in Pakistan is A but presence of subtype C, D and G confirmed in studies indicates viral evolution (S. Khan, Zehra, Zahid, Iqbal, & Maqsood, 2016).

The UNAIDS 2018 report and National AIDS control Program NACP Pakistan has reported that in Pakistan about 150,000 people are living with HIV including 38000 to 49000 women with HIV prevalence rate of 0.1 and only 15390 people receiving ART (UNAIDS, 2017) (NACP, 2018).

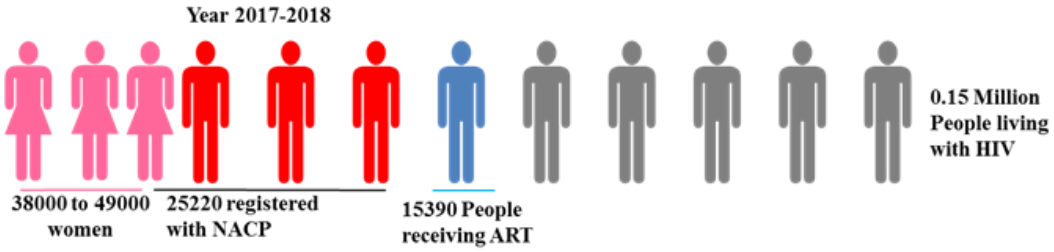


***Figure 1.2**: HIV country based statistics from UNAIDS and NACP Pakistan year 2017-2018*

1.4 Host Genetic Variations

The HIV-1 retrovirus exploits the host innate and adaptive immune system thus making host vulnerable to opportunistic infections. As HIV use certain host factors to enter target cell, the host-virus genetic interplay plays a significant role in viral control, host response and treatment outcome. The genome wide association studies have revealed that HIV host factors are associated with rate of disease progression, varying phenotypic disease markers like viral load and CD4+ T cell count and the response to antiretroviral therapy.

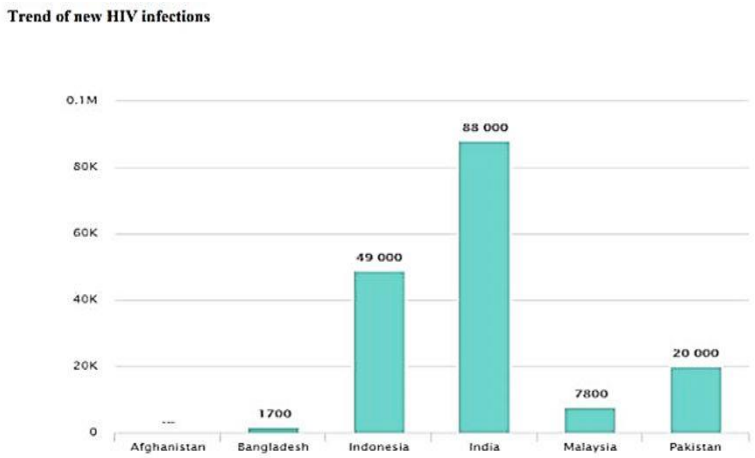


Figure 1.3: *Epidemic transition metrics-Trend of new HIV infection among different Asian and pacific states. Adapted from UNAIDS report*

1.5 Objective

There is a lot of research work done globally to find association of host genetic variants with HIV-1 susceptibility but no such published data is available related to association of common variants with HIV-1 in Pakistani population. The main objective of this study is:

To conduct case control study to evaluate association of novel regulatory host genetic polymorphic sites identified computationally in noncoding regions with HIV-1 infection in representative Pakistani population.

CHAPTER 2: LITERATURE REVIEW

The HIV an enveloped retrovirus awakened through cross species zoonotic transmission (chimpanzee) of simian immunodeficiency virus SIV in late 80s in Gambia and Cameroon (Keele et al., 2006). It was first isolated in France in 1983 coined as lymphadenopathy associated virus LAV (Barré-Sinoussi et al., 2013) and so far it has affected more than 30 million people across the globe (Engelman & Cherepanov, 2012)

T lymphocytes infection by HIV is characterized by expanded activation of cytotoxic CD8+ cells which is followed by decrease in CD4+ T cell count (200 cells then the 500 cells per µL in an uninfected individual) (Volberding & Deeks, 2010) and increase in viral load (associated with host factors) (McLaren et al., 2015)), thus sabotaging both the innate and adaptive immune systems leading to galvanization of opportunistic infections resulting in AIDS.

2.1 Viral Morphology

Extensive research on high resolution 3D HIV structure over the decades helped develop therapeutic interventions. HIV is an enveloped spherical retro virus with a lipid bilayer membrane derived from host cell. This retrovirus encode only sixteen (Engelman & Cherepanov, 2012) structural, catalytic and accessory proteins through reading frame shift.

2.1.1 Envelop Glycoproteins

The envelop surface glycoproteins gP120 and trans membrane gP41 heterodimer linked together via non covalent interaction (Engelman & Cherepanov, 2012) are responsible for viral entry and fusion with host cell. The viral particle incorporate some other proteins in envelop from host plasma membrane like actin and ubiquitin on its way out from host cell. In addition envelop is linked to shell holding matrix protein p17.

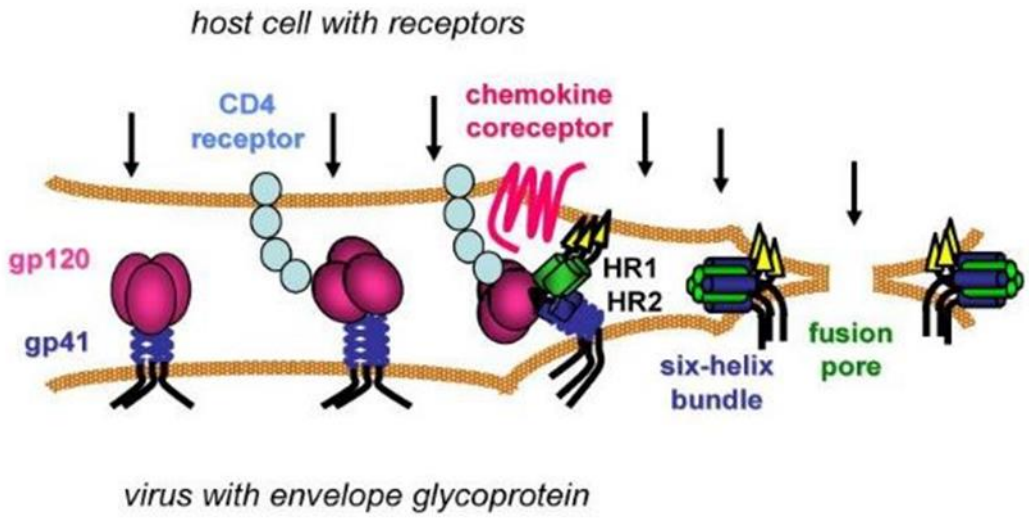


Figure 2.1: *Viral and host fusion mechanism with conformational changes shown* ***(****Weiss, 2013)*

2.1.2 Viral Core

The viral core is composed of capsid protein p24. Encapsulated in capsid there are two unspliced copies of RNA viral genome with a set of 3 viral genes gag, pol, and env (encoding viral proteins), p7 nucleo capsid proteins and three essential viral enzymes integrase, protease and reverse transcriptase encoded by pol gene.

2.1.3 Viral Accessory Proteins

From same set of genes courtesy reading frame shift 6 viral accessory and regulatory proteins located in capsid are produced. Vef, Vpr, Nef, Vpu are accessory proteins while Rev and Tat are regulatory proteins that remains unpackaged while interacting with host (Turner & Summers, 1999). The viral vef protein offset host APOBEC3G a cytidine deaminase thus stopping its ability to suppress HIV 1 infection (Sheehy, Gaddis, Choi, & Malim, 2002). Vpu accessory proteins are testified in literature regarding their tendency to counteract host cytoplasmic HIV limiting component tripartite motif-containing 5α TRIM5α (Stremlau et al., 2004) and antagonize tetherin, HIV release inhibition factor (Neil, Zang, & Bieniasz, 2008) . SAMHD1 is thwarted by Vpx (Laguette et al., 2011). Researchers are finding ways to manipulate these viral antagonists to make host factors able to combat with virus.

2.1.4 Structural proteins

The Gag gene encodes an un cleaved structural poly protein comprising of matrix, capsid and nucleo capsid domains and additional shorter peptides SP1 and SP2 and p6 (Ganser-pornillos, Yeager, & Sundquist, 2008) cleaved later by protease at five specific sites to make functional domains in mature virion. Gag recruits other components like viral replication proteins and genomic RNA thus driving assembly of mature infectious viral particle.

2.2 Viral Replication

The HIV replication initiates with series of interactions that leads to conformational changes in viral structure. Binding of viral particle with CD4+ T lymphocytes or macrophages (or microglia in brain with lower expression of CD4 (Weiss, 2013)), through interaction of gp120 viral envelop surface protein with N terminus of CD4 protein receptor of host cell is the first phase of replication (Turner & Summers, 1999).

Binding is followed by membrane fusion that requires host chemokine co receptors CXCR4 for X4 tropic HIV 1 and CCR5 for R5 tropic HIV 1 for conformational changes facilitating viral entry (Feng, Broder, Kennedy, & Berger, 1996).

Once viral particle enter, reverse transcription of viral genome into DNA take place through reverse transcriptase enzyme in host cytosol facilitated by accessory protein vif and nucleo capsid proteins. The pre integration complex get confined in nucleus by making connection with nuclear import machinery and nucleoporins courtesy vpr assecory proteins (Turner & Summers, 1999). The DNA then covalently binds with host genome catalyzed by integrase enzyme.

In the late viral replication phase, the mRNA is a. highly spliced (encoding Tat, Ref and Nef assecory proteins) b. remains unspliced (encoding Gag and Gag-Pol proteins) or c. singly spliced (enconding Env, Vpu, Vif and Vpr) (Turner & Summers, 1999).

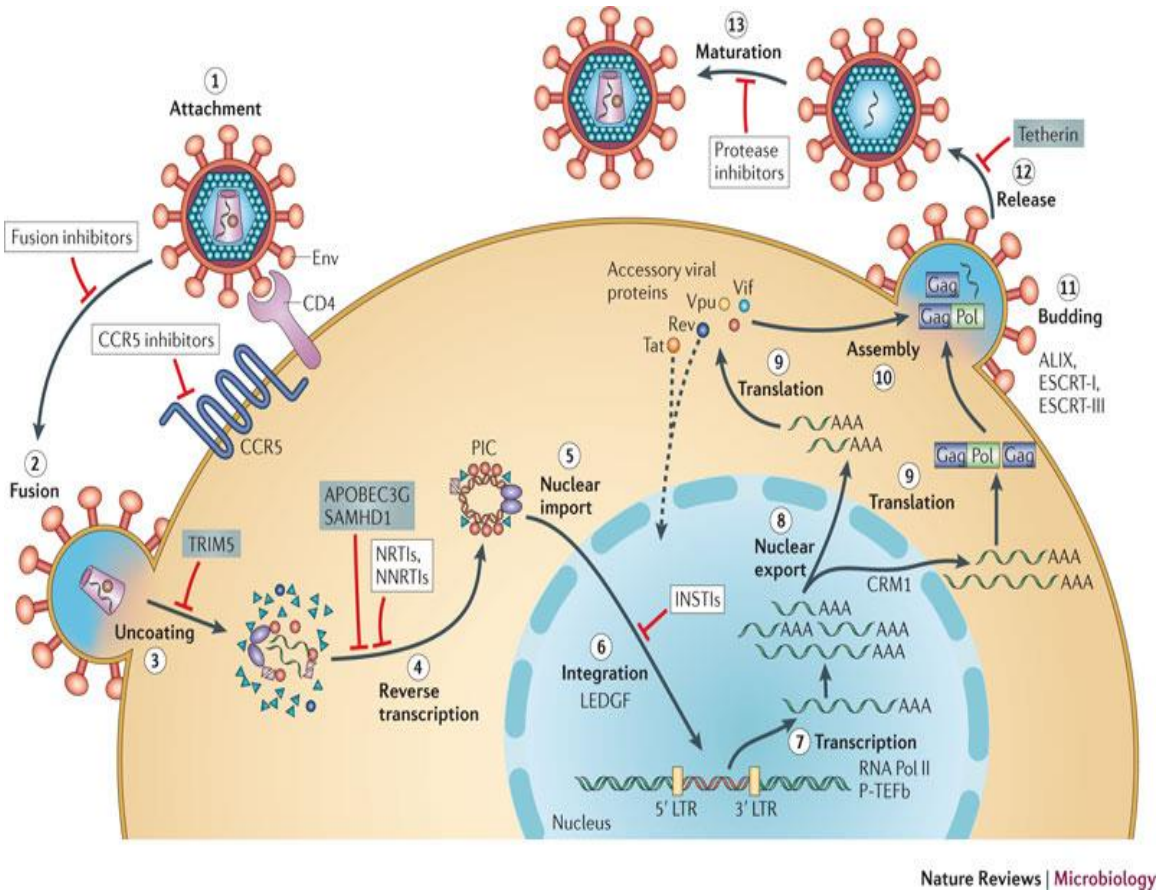


Figure 2.2: *HIV Replication Cycle. Figure taken from Natures Review (Engelman & Cherepanov, 2012)*

2.3 Antiretroviral Therapy

The viral pathogenicity and genome wide associations studies (GWAS) have been supportive in developing and improving the design of antiretroviral vaccines (McLaren & Carrington, 2015). The antiretroviral drugs are classified based on steps they inhibit in viral replication, the host factors or the structure of drug (Volberding & Deeks, 2010).

The Pharmaceutical industries have developed many drugs 23 of which got approved and are still in use. So far there are 17 drugs (Volberding & Deeks, 2010). The treatment response varies from patient to patient due to host genetic factors but drugs are well tolerated with short and long term toxic effects with patients having short life expectancy. If treatment is started before the severe immunological suppression eg CD4+ cells count > 200 per µL then the treatment is effective to an extent that it can decrease the plasma RNA concentration to level (50 to 200 RNA copies per ml) (Volberding & Deeks, 2010).

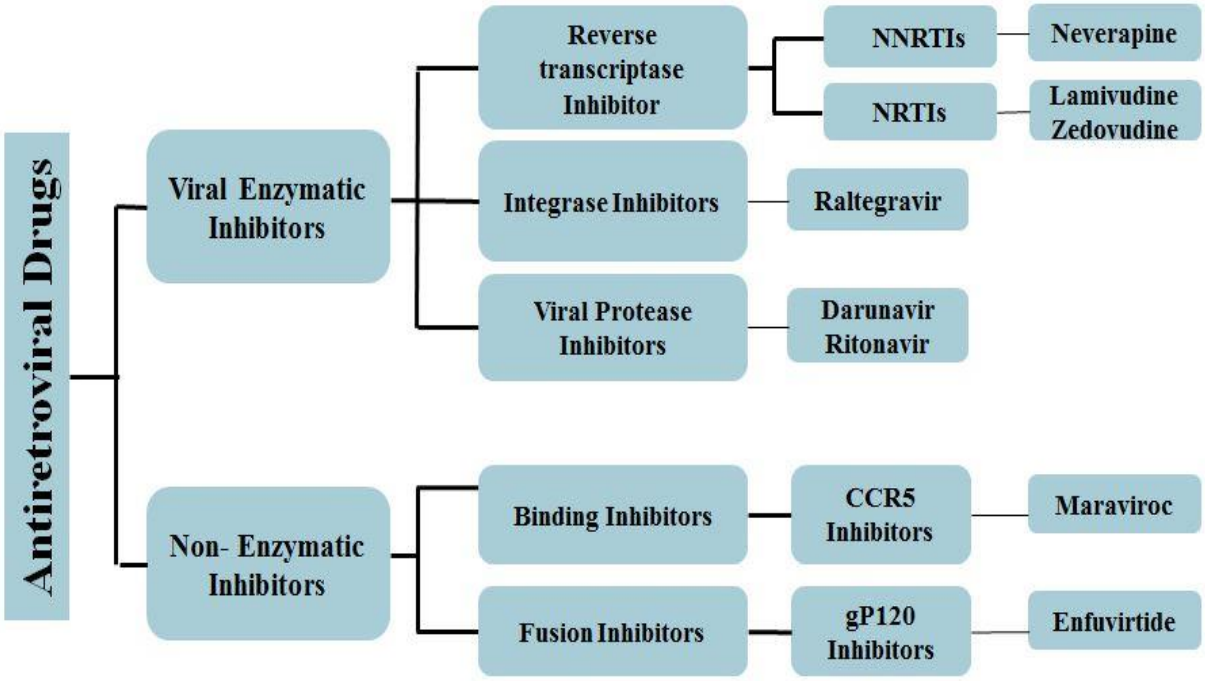


Figure 2.3: *Antiretroviral drugs, their target sites and examples*

2.4 Host Genetic Variations

An intricate collaboration between host and viral factors affects the host susceptibility to infection, disease progression and treatment response by controlling adaptive and innate immune response. The variation in host genome has been a topic of great concern for researchers since its impact on HIV pathogenesis is established. The viral factors are not always ideal target for drug development because of viral ability to show escape mutations but host genetics factors are relatively stable(Ping An & Winkler, 2010). The variants identified over years are related to immune system (adaptive and innate immunity), host antiviral restriction factors or those that are needed by virus for its replication(Ping An & Winkler, 2010). Some variations in host genome are found to have a low frequency protective affect through their additive or synergistic interactions with other variants while some others contribute to disease progression.

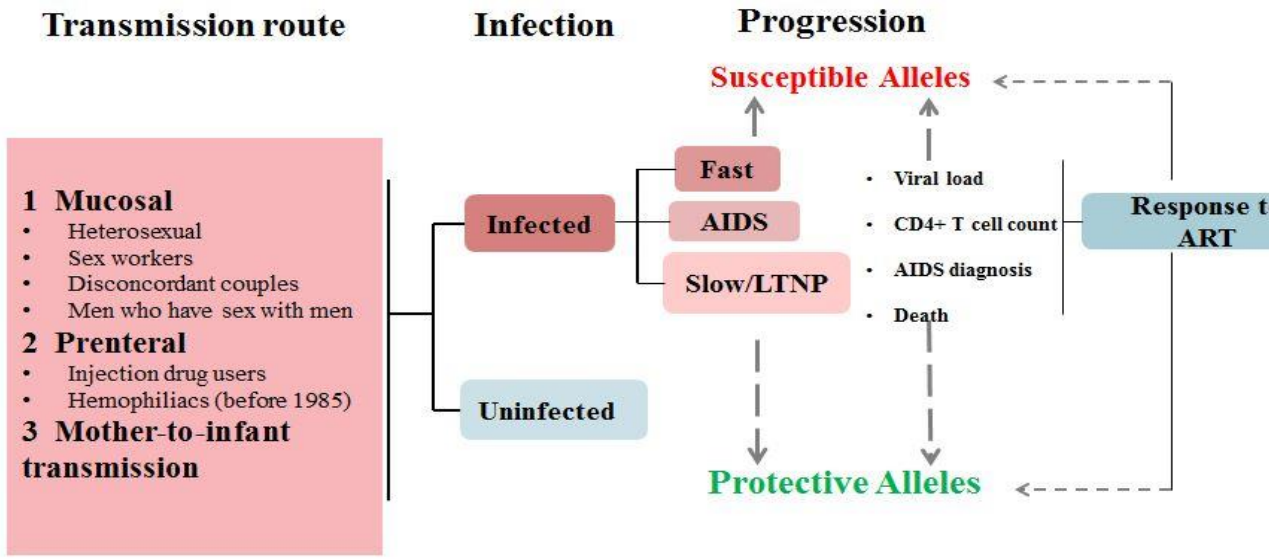


Figure 2.4: *Taken from (Ping An & Winkler, 2010) with slight modifications depicting how host variations are contributing to HIV-1.*

2.4.1 HIV Control and Human Leucocytes Antigen Class 1 Variants

HLA class 1 genes on p arm of chromosome 6 encode cell surface proteins that present the cellular and viral peptides to natural killer cells and influence the response of immune system by making it differentiate between self and foreign antigen(Ping An & Winkler, 2010). To escape the immune response, viral particles scheme is to down regulate HLA cellular expressions(Ping An & Winkler, 2010).

GWAS and HIV genetic studies have proposed that HLA-B*5701 polymorphism (through its combined or additive affect with other variants) is associated with HIV control by keeping CD4+ T cells decline slow and lowering plasma RNA concentration per milliliter (Fellay et al., 2007) (Migueles et al., 2000) The disease control is associated with amino acids in peptide binding groove of HLA-B57 (Pereyra et al., 2010). The protective role of HLA-B hetero zygosity was confirmed in study conducted in European population too (McLaren et al., 2015). Studies conducted on individuals with HLA-B*35 Px variant (Gao et al., 2001) and HLA-B*27variant (Carrington & O'Brien, 2003) have reported rapid and delayed disease progression respectively for the infected individuals.

The profusion of HLA-C on cell surface was found to be associated with HIV control in a study (Apps et al., 2013) but the HLA-C cell surface expression is influenced by variation in its

3’ UTR inhibitory polymorphic micro RNA 148a binding site (only intact site) (Smita Kulkarni et al., 2011). The variation in HLA-C 3’UTR binding site make HLA-C escape the inhibitory effect of miRna-148a thus more expression of HLA-C on cell surface presenting HIV peptide leads to disease control (S. Kulkarni et al., 2013). The SNP rs9264942 upstream the HLA-C was reported to be in linkage disequilibrium with HLA-C miR-148a binding site (Smita Kulkarni et al., 2011).

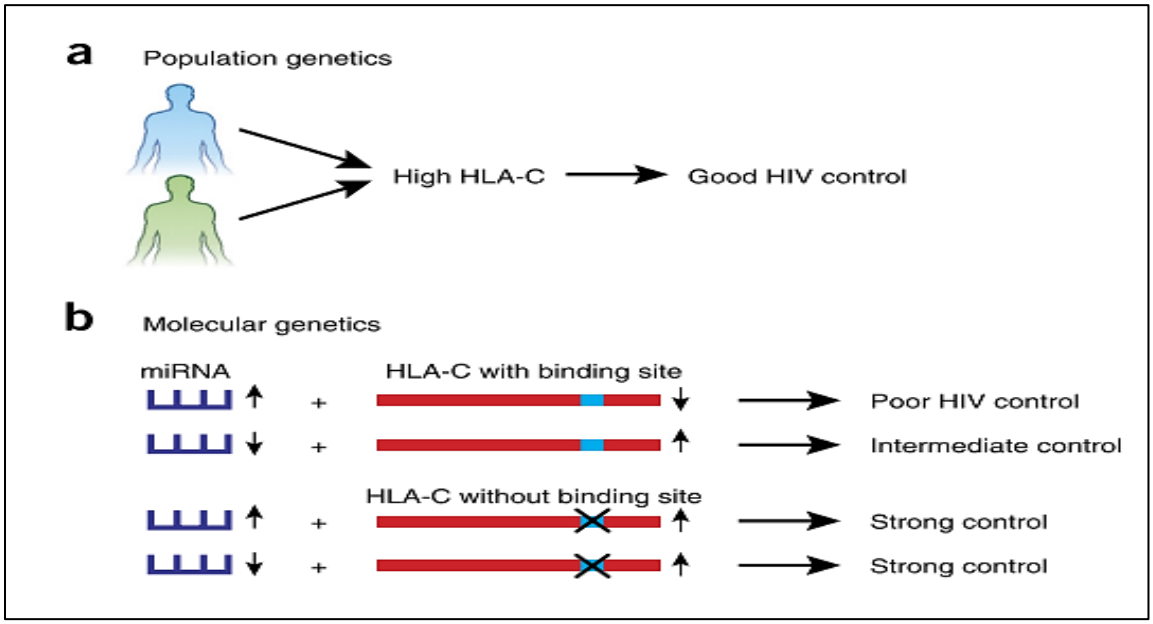


Figure 2.5: *HLA-C expression and role of miRNA-148a (McLaren & Carrington, 2015)*

2.4.1.1 HLA Variants associated with Set Point Viral Load

Some SNPs on chromosome 3 and 6 are testified to have a link with disease set point viral load in infected European population study, like rs59440261 in MHC region is in linkage disequilibrium with SNP rs2395029 of chromosome 3 (McLaren et al., 2015).

Amino acids variation in peptide binding groove of HLA-A at position 77, 95 and HLA-B at position 45, 67, 97 were reported to be associated with set point viral load with underlying alleles at the pointed out positions (combined effect) making 12.3 % of the variation in set point viral load (McLaren et al., 2015).

2.4.2 Role of Killer Immuno globin Like Receptors

The viral epitopes are when presented by class 1 molecule to natural killer cells, NK cells make sure that the target cell has displayed a correct number and type of class 1 molecules on its surface. KIR3DS1 receptor allele in a study has revealed its association with disease progression in case when required HLA-B ligands are absent but show protective effect in presence of its HLA-B alleles **7,8 of**(Ping An & Winkler, 2010)**.**

2.4.3 Role of CCR5 and Other Associated Variants

A number of studies have confirmed complete protective role of secondary chemokine receptor homozygous deletion mutation CCR5Δ32 against HIV host susceptibility, reduced set point viral load and rate at which disease advances as this mutation results in coding of nonfunctional truncated proteins (Dean et al., 1996)(Smith et al., 1997)(Ping An & Winkler, 2010)(McLaren et al., 2015). The heterozygous mutation in CCR5Δ32 is associated with slow progression of HIV (Dean et al., 1996). CCR5Δ32 stem cells transplantation in HIV positive patient has shown the significant results in favor of its protective role too.

Similarly infected individuals with co receptor CCR5 promoter variant Hap P1 were reported to undergo disease progression at a greater rate than infected individuals with other variants at this promoter site (Martin et al., 1998)(McDermott et al., 1998) with increased set point viral load(McLaren et al., 2015).

A chromosome 3 SNP rs1015164 is weakly related with CCR5Δ32 polymorphism thus linked to rate at which disease progresses (McLaren et al., 2015). Several SNPs in CCR5 region (liable for regulating its expression) were identified in study associated autonomously with viral load set point(McLaren et al., 2015).

The studies over the years have shown influence of genetic variations in chemokine receptor CCR5 and human leucocyte antigen HLA on host susceptibility and severity of disease. Variation in these two regions is responsible for 25% of phenotypic variability in viral load that ultimately leads to variation in HIV progression rate (McLaren et al., 2015).

2.4.4 Role of CCR5 Ligands Variation in HIV

Studies have shown that HIV strain R5 has to compete with various natural ligands like RANTES (CCL5) and CCRL3L1 to bind with its co receptor CCR5. The up or down regulated expression of these ligands has an effect on replication of R5 HIV strain. Three SNPs in intron

region, 3'UTR and promoter region are found to be associated with an increased HIV infection and progression due to decreased CCL5 gene expression (P. An et al., 2002).A polymorphic region upstream the RANTES gene (-28G) is found to be associated with increased ligand expression in Japanese individuals, thus competing with HIV R5 strain to bind with co receptor CCR5 delaying disease progression(Liu et al., 1999). Similarly lower levels of CCRL3L1 is associated with an increased HIV risk (Gonzalez et al., 2005).

2.4.5 Immune Regulatory Genes Polymorphism and HIV

HIV disease advancement and acquisition is also associated with polymorphism in regulatory and coding regions of several host cytokine genes like IL10 and interferon-γ (IFNG) (O'Brien & Nelson, 2004). The genetic variants in noncoding regions are reported to play significant role in altering gene expression thus affecting rate of disease progression.

2.4.6 Role of SNPs in HIV

Table1: *Reported SNPs and their role in HIV*

Phenotype	SNP ID		Reference study
	rs4118325	rs1360517	
	rs10800098	rs1556032	Genome wide association study of a rapid progression cohort identifies new susceptibility alleles for AIDS (ANRS Genome wide Association Study
	rs1020064	rs1360517	
	rs3108919		
AIDS progression	rs8321		Genome wide association study of an AIDS-non progression cohort emphasizes the role played by HLA genes (ANRS Genome wide Association Study 02
	rs10484554		A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1
	rs2395029		The major genetic determinants of HIV-1 control affect HLA class I peptide presentation

AIDS	rs11239930 rs6441975 rs477687 rs7217319	rs11884476 rs1015164 rs6467710	Genome-wide association study implicates PARD3B-based AIDS restriction
HIV-1 Replication	rs12483205		Genome-wide association study identifies single nucleotide polymorphism in DYRK1A associated with replication of HIV-1 in monocyte-derived macrophages
HIV-1 Susceptibility	rs11924930 rs4751185 rs285757	rs842304 rs1240773 rs12604483	Common human genetic variants and HIV-1 susceptibility: a genome-wide survey in a homogeneous African population
HIV control**	rs10159302 rs12122100 rs6751715	rs2069084 rs1405262 rs9367630	Common genetic variation and the control of HIV-1 in humans
	rs3131018 rs2523590	rs9264942 rs2523608	The major genetic determinants of HIV-1 control affect HLA class I peptide presentation
	rs1413191		Host determinants of HIV-1 control in African Americans

2.5 Proteins Associated with Novel SNPs and their role in HIV 1

2.5.1 MYC

SNP rs6457282 was found to be associated with MYC protein that being a transcriptional factor up or down regulates the expression of certain genes. The role of c-myc in AIDS pathogenesis through regulation of cytokines like TNF-alpha and IL-6 (via activation of MAPK, ERK1/2 PKR) was also reported in a study as expression of c-myc was up regulated by HIV-1 viral TAT protein (J.C.H., J.C.B., & H.C.H., 2012). The role of c-myc in nuclear import of HIV-

1 viral DNA was reported in another study thus depicting its importance in viral life cycle (Sun & Clark, 1999).

.

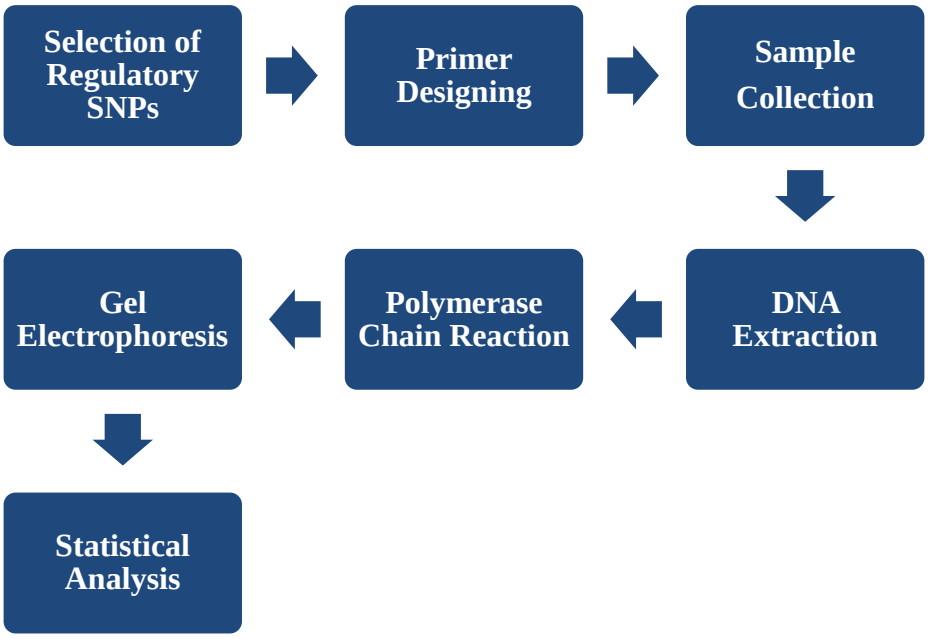
2.5.2 CTCF

CTCF a transcriptional factor is associated with negative regulation of c-myc gene. The expression of CTCF was reported to be up regulated by 18 folds in HIV-1 resistant human T-lymphocyte cell line secreting HIV-1 resistance factor (HRF) (Kartvelishvili, Lesner, Szponar, & Simm, 2004).

CHAPTER 3: MATERIAL AND METHODS

3.1 Study Design

Following steps in sequence were followed to validate the effect of our novel regulatory SNPs in noncoding regions of HIV host genome during the course of research project.



3.2 Regulatory SNPs Selection

The regulatory SNPs were identified and selected by using GWAS catalog and regulome DB online search engines by another research group.

3.3 Primers Designing

Four sets, containing 4 primers each targeting the four specific selected regulatory single nucleotide polymorphic region of HIV host genome were designed manually or by using primer3 software (Koressaar & Remm, 2007) (Untergasser et al., 2012).An online computational engine Oligo Calc: oligonucleotide properties calculator (Kibbe, 2007) and Multiple Primer Analyzer (Breslauer, Frank, Blocker, & Marky, 1986) were used to check Primer specificity, GC content, dimer formation, hair pin loop.

Each SNP region was targeted by four primers (one common reverse, one common forward and 2 allelic primers one for each of the specific allele in separate reaction).100uL

volume 1:10 dilutions of primers were prepared from stock made according to manufacturer guidelines.

Table 1: *Selected SNPs, corresponding primer sequence and temperature specifications*

No	SNP ID	Primer ID	Primer Sequence	Product Type	Product Size	Tm
1	**rs6457282**	rs6457282 **F**	CCCTTTTCTGCCTGGAGTC	Control	317bp	62.3
		rs6457282 **R**	GAGAGCCTGTGGTTATTTAGG			60.6
		rs6457282F**(C)**	GCCAGAACAAAAGCTGAGC	Allele Specific	201bp	60.2
		rs6457282F**(T)**	GCCAGAACAAAAGCTGAGT			58.0
2	**rs9261105**	rs9261105**F**	CCGACGATGGGGTCATG	Control	516bp	62
		rs9261105**R**	GGAGTAAGTGATCTGCGC			59.9
		rs9261105F**(A)**	GAGGAGCGAAAGGTCCA	Allele Specific	401bp	59.6
		rs9261105F**(G)**	GAGGAGCGAAAGGTCCG			62
3	**rs17064977**	rs17064977**F**	GCAAAAATGCTCAAGGTGATAAG	Control	290bp	59.2
		rs17064977**R**	AAAATAAAGAAACAGGTCAAAGGC			57.7
		rs17064977F**(C)**	TGGAAGAGGAGCGACTCC	Allele Specific	213bp	62.2
		rs17064977F**(T)**	TGGAAGAGGAGCGACTCT			59.9
4	**rs3130350**	rs3130350**F**	GAGGCAGATGAAGCAGATG	Control	298bp	60.2
		rs3130350**R**	TCTGGAAGAGCAAAGCTG			57.6
		rs3130350F(**G)**	AGTGTCACAATAGGAGGGTG	Allele Specific	199bp	60.4
		rs3130350F**(T)**	AGTGTCACAATAGGAGGGTT			58.4

3.4 Ethical Statement

Patients and controls were first made comfortable by informing them about research project prior the sample collection. Patients were asked to sign a written consent form with patient and sample particulars.

3.4.1 Blood Sampling

Blood samples collected in 3ml sterile EDTA K3 tubes were stored at -4 degree Celsius (control sample) or processed same day (patient samples) for downstream applications like DNA extraction courtesy NIH.

HIV positive blood samples were collected in Pakistan Institute of Medical Sciences Referral Lab. All the standard precautionary measures were taken to avoid infection.

***Table 3**: Inclusion/exclusion criteria of samples*

Group	Mean Age	Inclusion Criteria
Control	30±8	HIV negative individuals with no other disease history.
Case	38±8	HIV positive individuals with high viral load but either • Treatment Naïve • Failed to respond to ART

3.5 Host DNA Extraction, Quantification and Amplification

DNA was extracted from host blood using phenol chloroform method and trizol LS method followed by its quantification and PCR amplification using housekeeping β-actin gene.

3.5.1 DNA Extraction by Phenol Chloroform Method from Control Samples

To extract genomic DNA from whole blood control sample, nonorganic phenol chloroform based method (Grimberg et al., 1989) was used with slight modifications. The method involves SDS and proteinase K mediated digestion and protein removal by phenol, chloroform and isoamyl alcohol. Purification was carried out with absolute alcohol. Two cell lysis buffers were prepared beforehand for extraction. The cell lysis buffer 1 (CLB1) was prepared from 0.32M sucrose, 10mM Tris (pH 7.5) and 5Mm MgCl2. pH of freshly prepared

solution was maintained followed by autoclave. 1% Triton X-100 was added in solution before use. This buffer ruptures plasma membrane and releases the cell components. Lysis buffer 2 (CLB2) was prepared from 10Mm Tris (pH7.5), 400Mm NaCl and 2Mm EDTA (pH=8.0). This buffer separate proteins and used for DNA precipitation. All centrifugation steps were done at 13000 rpm.

First equal volume (750uL) of blood and ice cold CLB1 was taken in 2 mL eppendorf tubes. Buffer was gently mixed with blood by inversion and left at room temperature for 10 minutes followed by standard centrifugation for 60 seconds. Pellet was washed with 400µL of CLB1 buffer again by centrifugation at same conditions. Next pellet was dissolved in 400 µL of CLB2, 14 µL of 20% SDS and 5 µL of Proteinase K followed by overnight incubation at room temperature to denature proteins and for inactivation of nucleases.

400 µL of phenol and a solution X (freshly prepared Chloroform and Isoamyl alcohol solution in 24:1 ratio) was added to tube and centrifuged for 10 minutes. Upper aqueous layer was collected in new 1.5 mL tube and 500 µL of Solution X was added again followed by centrifugation for 10 minutes. Aqueous layer was collected in new tube for the second time and DNA was precipitated by adding 55 µL of 3M Sodium Acetate (pH 6.0) and 500 uL of chilled Iso-propanol. Tubes were inverted several times and centrifuged for another 10 minutes. DNA was washed with 200 µL of 100% chilled ethanol initially followed by washing with 70% ethanol with centrifugation for 7 minutes. DNA pellet was dried by incubation at 37°C for 20-30 minutes followed by addition of 60 to 75µL of TE buffer to solubilize the DNA pellet.

3.5.3 Quality and Quantity of Extracted DNA

To check the integrity of extracted DNA, 5 µL of DNA was run on 0.6% agarose gel (stained with Ethidium Bromide) at 110V for 30 minutes and visualized under UV using DigiGenius gel documentation system (Syngene, United Kingdom).

Optical density (OD) of DNA was measured at 260nm on Nano Drop 2000 UV-Vis Spectro photometer (Thermo Scientific). A260/A280 absorbance ratio of 1.7 to 2.0 indicated purity of DNA. DNA dilutions were prepared and stored at -4◦C for downstream applications.

3.5.4 Amplification with House Keeping Gene (β-Actin)

DNA quality was further evaluated by performing amplification reaction with housekeeping gene β actin.

3.6 Sequence Specific PCR for SNPs Genotyping

Sequence specific polymerase chain reaction was performed for genotyping four selected regulatory SNPs in case and control host genome. For each sample two allele specific reactions were performed. Each reaction mixture contains 3 primers, control forward, common reverse (to amplify control sequence) and one of the two allelic primers (to amplify DNA only in the presence of given allele).

Reaction mixture was prepared with 5µL of template DNA (100-250ng), 1µLof each primer (10mM) and 3µL of 5X Solis Bio Dyne cyber green master mix with addition of nuclease free water to make a final volume of 25µL. The reaction was performed with following thermo cycler conditions. Initial denaturation at 95°C for 5 minutes followed by 25 cycles of denaturation at 95°C for 45 seconds, annealing at 62°C for 45 seconds, extension at 72°C for 45 seconds and final extension at 72°C for 7 minutes followed by infinite hold at 4°C.

3.7 Gel Electrophoretic Analysis of PCR Products

The PCR product of each reaction mixture was run on 2% agarose gel (stained with Ethidium Bromide) at 100V for 35 minutes and visualized under UV on gel doc system to analyze the presence of bands for a specific allele.

3.8 Statistical Analysis

The SNP data was analyzed manually and allelic/genotypic frequencies were calculated for each SNP. To find significance of results Chi square test was performed using Graph Pad Prism 5.01 software (Graph-Pad Software, San Diego, CA USA) by calculating p-values. Odd Ratios (OR) and risk ratios (RR) were calculated using MEDCALC statistical software.

CHAPTER 4: RESULTS

4.1 Gel Electrophoretic Images

The agarose gels of varying concentrations (0.6% and 2%) stained with ethidium bromide were used to visualize the extracted DNA and the PCR products respectively. The gel were visualized under UV using gel doc system and images were saved and labeled.

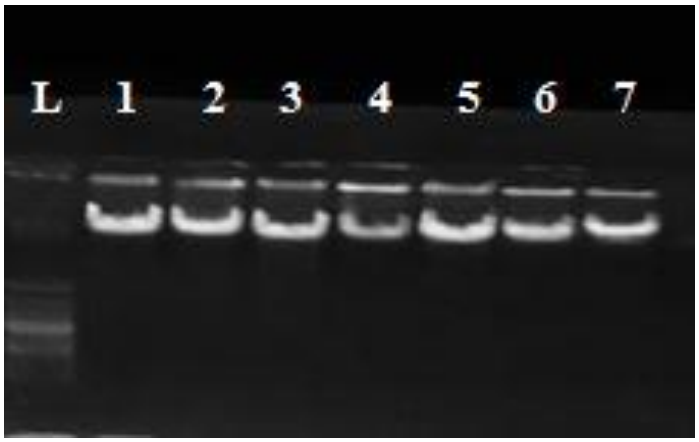


Figure 4.1*: Gel electrophoretic image of extracted genomic DNA on 0.6% Agarose gel stained with Ethidium bromide (L=100bp ladder)*

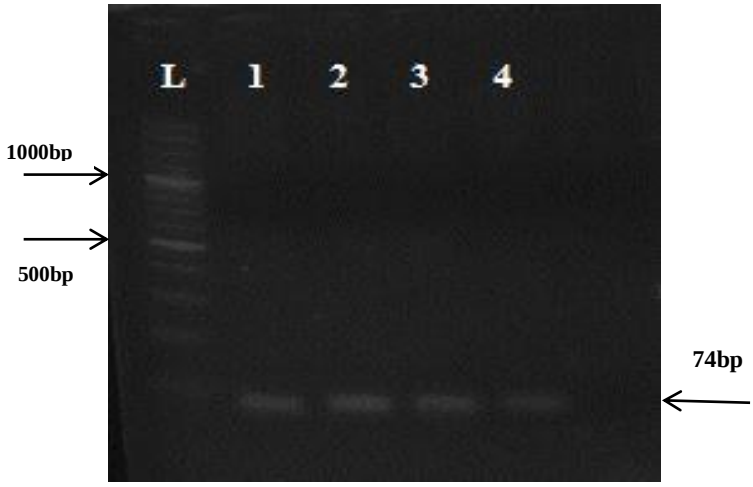


Figure 4.2: *PCR amplification of GAPDH gene on extracted DNA visualized on 2%agarose gel stained with Ethidium Bromide.*

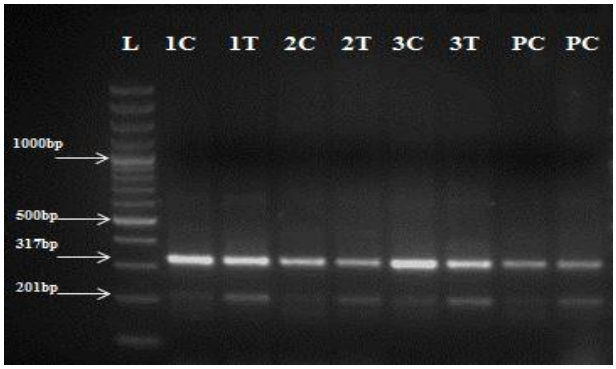


Figure 4.3: *rs6457282 polymorphism screening in three samples. PC represents positive control. C represents C allele and T represents T allele. L=100bp ladder.*

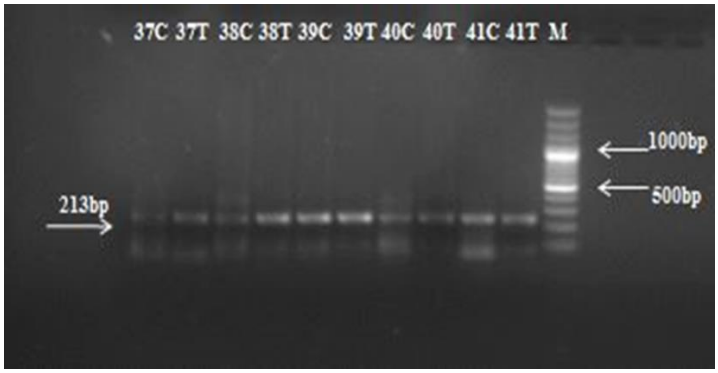


Figure 4.4: *rs17064977polymorphism screening. C represents C allele and T represents T allele. L=100bp ladder.*

4.2 Statistical Analysis

4.2.1 Genotypic/Allelic Frequency Distribution

The frequency was calculated manually by counting number of samples with specific genotype and odd ratios and risk ratios were calculated through online software engines.

Table 4: *Frequency, odd ratio, risk ratio and P-value of* ***rs6457282*** *and* ***rs17064977*** *SNP genotype and alleles observed in Pakistani Population.*

		HIV+(n=50)	***HIV¯(n=37)***	***Odd ratio OR***	***Relative risk RR***	***95% CI***	***P value***	***Significance P<0.05***
rs6457282								
Genotypic Distribution	*CC*	0.04(2/50)	0.08(3/37)	0.4722	0.493	0.0868-2.8051	0.4256	----
	TT	0.06(3/50)	0.43(16/37)	0.0838	0.138	0.0436-0.4415	0.0008	***
	CT	0.9(45/50)	0.49(18/37)	9.50	1.85	1.3119-2.6088	0.0005	***
Allelic Frequency	*C*	0.49(49/100)	0.32(24/74)	2.002	1.51	1.028-2.220	0.0356	*
	T	0.51(51/100)	0.67(50/74)	0.4996	0.75	0.588-0.9679	0.0266	*

		HIV+(n=47)	***HIV¯(n=41)***	***Oddratio OR***	***95%CI***	***P- Value***	***Significance P<0.05***
rs17064977							
Genotypic Distribution	*CC*	0.06(3/47)	0.145(6/41)	0.3977	0.0928-1.7046	0.2143	----
	TT	0.21(10/47)	0.609(25/41)	0.1730	0.0676-0.4424	0.0002	***
	CT	0.72(34/47)	0.24(10/41)	8.1077	3.1125-21.119	0.0001	***
Allelic Frequency	*C*	0.43(40/94)	0.268(22/82)	2.0202	1.0684-3.8200	0.0305	*
	T	0.57(54/94)	0.73(60/82)	0.4950	0.2618-0.9360	0.0305	*

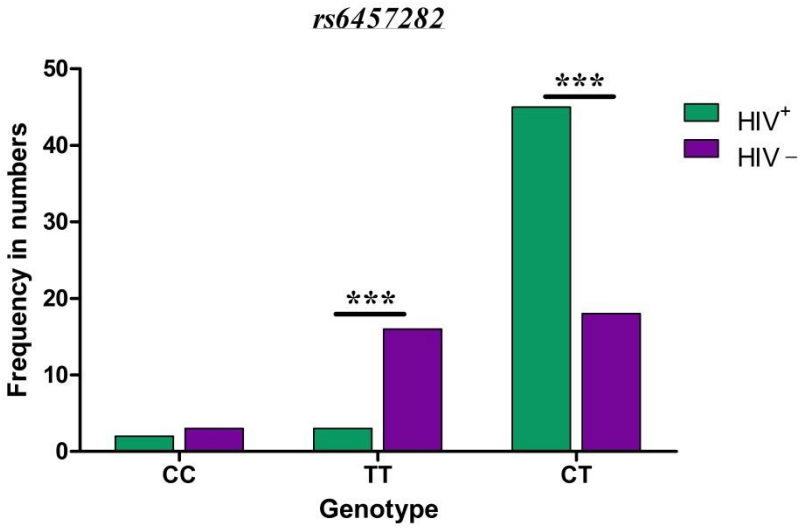


Figure 4.5 : *Genotypic Distribution of rs6457282*

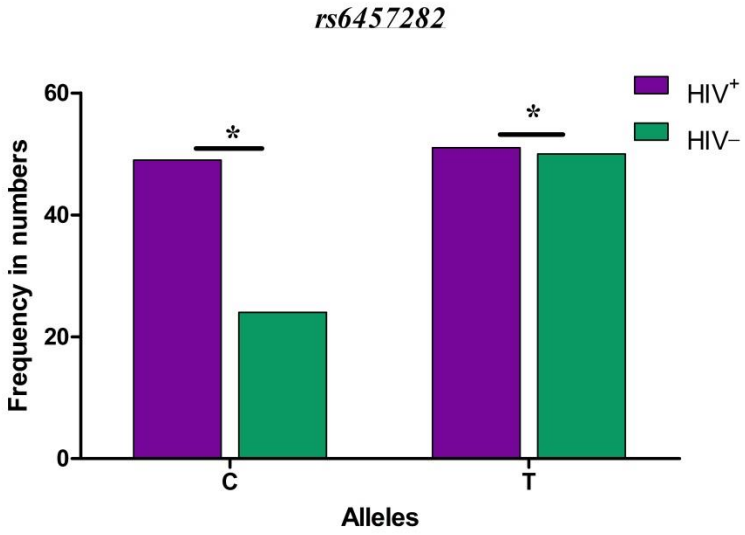


Figure4.6: *Allelic frequency of rs6457282*

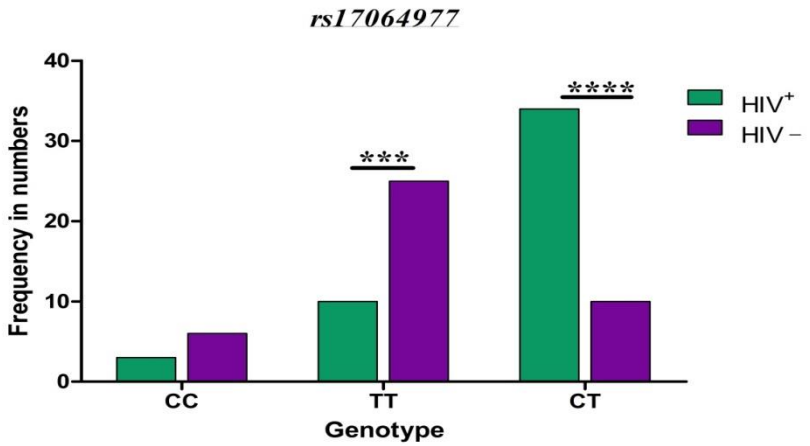


Figure4.7: *Genotypic distribution of rs17064977*

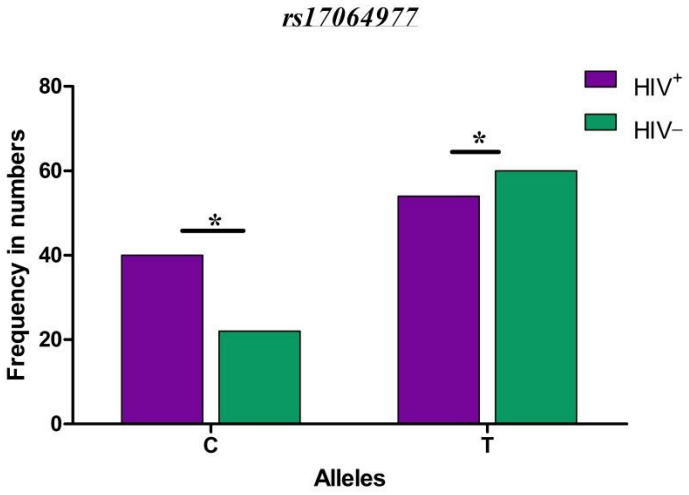


Figure 4.8: *Allelic frequency of SNP rs117064977*

CHAPTER 5: DISCUSSION

Over the years, the association of human genomic sequence variation with various diseases has been reported by researchers. These sequence variations can be single nucleotide polymorphisms. The significance of polymorphic sites in terms of their influence on disease outcome and effect on gene expression cannot be denied. Many single nucleotide polymorphic sites in coding or intronic region of host genome, are reported in various ethnic groups to influence HIV in terms of rate of disease progression, variation in host phenotypic outcomes like CD4+cell count and viral RNA concentrations and response to antiretroviral therapy.

As in Pakistan, HIV has shown a progress in terms of new HIV infections, the host population must be analyzed genetically to find the variations in genome that can help the virus to take over or making host more vulnerable to disease, influence the rate of disease progression or alter the response to ART.

In this study two novel polymorphic sites identified computationally, were analyzed in HIV^{+} and HIV^{-} samples collected at PIMS referral lab representing Pakistani population. All HIV+ samples were either treatment naïve or did not show any expected response to ART. The two computationally identified novel SNPs were linked with SNPs on chromosome 6 and 18 that were reported to be associated with HIV-1.

Our results demonstrated the possible association of *rs6457282 T/C* polymorphism with risk of HIV-1 infection in Pakistani population. Increased risk of HIV-1 infection was observed with the *rs6457282 C/T* genotype, as compared to healthy controls (OR = 9.5, 95% CI =1.3119-2.6088 and p = 0.0005). Results suggested that C allele confers susceptibility to HIV-1 infection (OR>1) whereas T allele might have a protective effect with OR<1.

Our first proxy SNP *rs6457282* is positioned 9735 bases from *rs7756521* (a C/T nucleotide variation) on chromosome 6 affects the discoidin domain receptor tyrosine kinase 1 (2KB Upstream Variant) DDR1 gene expression and the MYC protein binding. *rs7756521 is* reported as causal variant of HIV-1 control (Jin et al., 2018). It is reported that pro inflammatory cytokines like TNF-α play its part in inducing HIV, and c-myc regulate TNF-α expression thus contributing to disease progression.

Similarly the second proxy SNP *rs17064977* positioned 2162 bases from *rs8099014* on chromosome 18 can alter the Contig35673_RC (cis), LMAN1 (cis) and TLR4 (cis) expression and CTCF protein binding. CTCF is reported to play role in HIV disease control.

When *rs17064977C/T* polymorphic site was genotyped via allele specific PCR, results suggested its association with HIV susceptibility like *rs6457282 T/C* in our case-control study. The patient samples were found to be more heterozygous *(CT)* than control groups. The odd ratios suggested that *rs17064977C/T* heterozygous genotype may be a risk factor in HIV susceptibility (OR=8.1077, 95%CI=3.1125-21.119, P=0.0001) and the T allele may be a protective factor (OR<1) against HIV acquisition.

CONCLUSION:

The computationally identified SNPs might be a real risk factor associated with HIV-1 disease susceptibility but the limited number of samples was analyzed in this study. The greater number of samples might change the outcome. The SNP association with phenotypic outcomes needs to be evaluated too in population of varying ethnicity. The associated gene and protein expression of studies can help identify the new aspects of HIV-1 disease.

CHAPTER 6: REFERENCES

An, P., Nelson, G. W., Wang, L., Donfield, S., Goedert, J. J., Phair, J., … Winkler, C. A. (2002). Modulating influence on HIV/AIDS by interacting RANTES gene variants. *Proceedings of the National Academy of Sciences*, *99*(15), 10002–10007. https://doi.org/10.1073/pnas.142313799

An, P., & Winkler, C. A. (2010). Host genes associated with HIV/AIDS: advances in gene discovery. *Trends in Genetics*, *26*(3), 119–131. https://doi.org/10.1016/j.tig.2010.01.002

Apps, R., Qi, Y., Carlson, J. M., Chen, H., Gao, X., Thomas, R., … Carrington, M. (2013). Influence of HLA-C expression level on HIV control. *Science*, *340*(6128), 87–91. https://doi.org/10.1126/science.1232685

Barré-Sinoussi, F., Ross, A. L., & Delfraissy, J. F. (2013). Past, present and future: 30 years of HIV research. *Nature Reviews Microbiology*, *11*(12), 877–883. https://doi.org/10.1038/nrmicro3132

Breslauer, K. J., Frank, R., Blocker, H., & Marky, L. A. (1986). Predicting DNA duplex stability from the base sequence. *Proceedings of the National Academy of Sciences*, *83*(11), 3746–3750. https://doi.org/10.1073/pnas.83.11.3746

Carrington, M., & O'Brien, S. J. (2003). The Influence of *HLA* Genotype on AIDS. *Annual Review of Medicine*, *54*(1), 535–551. https://doi.org/10.1146/annurev.med.54.101601.152346

Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Smith, M. W., Allikmets, R., … O'Brien, S. J. (1996). Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE. *Science*, *273*(5283), 1856–1862. https://doi.org/8791590

Engelman, A., & Cherepanov, P. (2012). The structural biology of HIV-1: Mechanistic and

therapeutic insights. *Nature Reviews Microbiology*, *10*(4), 279–290. https://doi.org/10.1038/nrmicro2747

Fellay, J., Shianna, K. V., Ge, D., Colombo, S., Ledergerber, B., Weale, M., … Goldstein, D. B. (2007). A whole-genome association study of major determinants for host control of HIV-1. *Science*, *317*(5840), 944–947. https://doi.org/10.1126/science.1143767

Feng, Y., Broder, C. C., Kennedy, P. E., & Berger, E. A. (1996). HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*, *272*(5263), 872–877. https://doi.org/10.1126/science.272.5263.872

Ganser-pornillos, B. K., Yeager, M., & Sundquist, W. I. (2008). The structural biology of HIV assembly. https://doi.org/10.1016/j.sbi.2008.02.001

Gao, X., Nelson, G. W., Karacki, P., Martin, M. P., Phair, J., Kaslow, R., … Carrington, M. (2001). Effect of a Single Amino Acid Change in MHC Class I Molecules on the Rate of Progression to AIDS. *New England Journal of Medicine*, *344*(22), 1668–1675. https://doi.org/10.1056/NEJM200105313442203

Gonzalez, 1* Hemant Kulkarni, 1* Hector Bolivar, 1*. Andrea Mangano, 2*. (2005). The Influence of CCL3L1 Gene– Containing Segmental Duplications The Influence of CCL3L1 Gene– Containing Segmental Duplications on HIV-1/AIDS Susceptibility Containing Segmental Duplications on HIV-1/AIDS Susceptibility. *Science*, *37*, 296–318. https://doi.org/10.1126/science.1101160

Grimberg, J., Nawoschik, S., Belluscio, L., McKee, R., Turck, A., & Eisenberg, A. (1989). A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. *Nucleic Acids Research*, *17*(20), 8390.

J.C.H., P., J.C.B., L., & H.C.H., Y. (2012). The role of c-Myc in regulating HIV-1 Tat protein induced cytokine expression and consequent effects on opportunistic infection. *Cytokine*, *59*(3), 522. https://doi.org/10.1016/j.cyto.2012.06.096

Jin, Y., Gittelman, R. M., Lu, Y., Liu, X., Li, M. D., Ling, F., & Akey, J. M. (2018). Evolution of dnaase i hypersensitive sites in mhc regulatory regions of primates. *Genetics*, *209*(2), 579–

589. https://doi.org/10.1534/genetics.118.301028

Kartvelishvili, A., Lesner, A., Szponar, M., & Simm, M. (2004). Microarray analysis of differentially expressed genes in cells resistant to HIV-1. *Immunology Letters*, *93*(1), 79–86. https://doi.org/10.1016/j.imlet.2004.02.009

Keele, B. F., Van Heuverswyn, F., Li, Y., Bailes, E., Takehisa, J., Santiago, M. L., … Hahn, B. H. (2006). Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science*, *313*(5786), 523–526. https://doi.org/10.1126/science.1126531

Khan, O. a, & Hyder, A. A. (2001). Responses to an emerging threat: HIV/AIDS policy in Pakistan. *Health Policy and Planning*, *16*(2), 214–218. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11358924

Khan, S., Zehra, F., Zahid, M., Iqbal, A., & Maqsood, N. (2016). Pattern of concentrated HIV epidemic in Pakistan. *Pathology*, *48*, S132. https://doi.org/10.1016/j.pathol.2015.12.356

Khanani, M. R., Somani, M., Rehmani, S. S., Veras, N. M. C., Salemi, M., & Ali, S. H. (2011). The spread of HIV in Pakistan: Bridging of the epidemic between populations. *PLoS ONE*, *6*(7), 1–6. https://doi.org/10.1371/journal.pone.0022449

Kibbe, W. A. (2007). OligoCalc: An online oligonucleotide properties calculator. *Nucleic Acids Research*, *35*(SUPPL.2), 43–46. https://doi.org/10.1093/nar/gkm234

Koressaar, T., & Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*, *23*(10), 1289–1291. https://doi.org/10.1093/bioinformatics/btm091

Kulkarni, S., Qi, Y., O'hUigin, C., Pereyra, F., Ramsuran, V., McLaren, P., … Carrington, M. (2013). Genetic interplay between HLA-C and MIR148A in HIV control and Crohn disease. *Proceedings of the National Academy of Sciences*, *110*(51), 20705–20710. https://doi.org/10.1073/pnas.1312237110

Kulkarni, S., Savan, R., Qi, Y., Gao, X., Yuki, Y., Bass, S. E., … Carrington, M. (2011). Differential microRNA regulation of HLA-C expression and its association with HIV control. *Nature*, *472*(7344), 495–498. https://doi.org/10.1038/nature09914

Laguette, N., Sobhian, B., Casartelli, N., Ringeard, M., Chable-Bessia, C., Ségéral, E., … Benkirane, M. (2011). SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature*, *474*(7353), 654–657. https://doi.org/10.1038/nature10117

Liu, H., Chao, D., Nakayama, E. E., Taguchi, H., Goto, M., Xin, X., … Shioda, T. (1999). Polymorphism in RANTES chemokine promoter affects HIV-1 disease progression. *Proceedings of the National Academy of Sciences*, *96*(8), 4581–4585. https://doi.org/10.1073/pnas.96.8.4581

Martin, M. P., Dean, M., Smith, M. W., Winkler, C., Cerrard, B., Michael, N. L., … Carrington, I. M. (1998). Genetic Acceleration of AlDS, *282*(DECEMBER). https://doi.org/10.1002/acp.1528

McDermott, D. H., Zimmerman, P. A., Guignard, F., Kleeberger, C. A., Leitman, S. F., & Murphy, P. M. (1998). CCR5 promoter polymorphism and HIV-1 disease progression. *Lancet*, *352*(9131), 866–870. https://doi.org/10.1016/S0140-6736(98)04158-0

McLaren, P. J., & Carrington, M. (2015). The impact of host genetic variation on infection with HIV-1. *Nature Immunology*, *16*(6), 577–583. https://doi.org/10.1038/ni.3147

McLaren, P. J., Coulonges, C., Bartha, I., Lenz, T. L., Deutsch, A. J., Bashirova, A., … Fellay, J. (2015). Polymorphisms of large effect explain the majority of the host genetic contribution to variation of HIV-1 virus load. *Proceedings of the National Academy of Sciences*, *112*(47), 14658–14663. https://doi.org/10.1073/pnas.1514867112

Migueles, S. A., Sabbaghian, M. S., Shupert, W. L., Bettinotti, M. P., Marincola, F. M., Martino, L., … Connors, M. (2000). HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. *Proceedings of the National Academy of Sciences*, *97*(6), 2709–2714. https://doi.org/10.1073/pnas.050567397

Neil, S. J. D., Zang, T., & Bieniasz, P. D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature*, *451*(7177), 425–430. https://doi.org/10.1038/nature06553

O'Brien, S. J., & Nelson, G. W. (2004). Human genes that limit AIDS. *Nature Genetics*, *36*(6), 565–574. https://doi.org/10.1038/ng1369

Pereyra, F., Jia, X., McLaren, P. J., Telenti, A., De Bakker, P. I. W., & Walker, B. D. (2010). The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science*, *330*(6010), 1551–1557. https://doi.org/10.1126/science.1195271

Shah, S. A., Altaf, A., Mujeeb, S. A., & Memon, A. (2004). An outbreak of HIV infection among injection drug users in a small town in Pakistan: Potential for national implications [2]. *International Journal of STD and AIDS*, *15*(3), 209. https://doi.org/10.1258/095646204322916713

Sheehy, A. M., Gaddis, N. C., Choi, J. D., & Malim, M. H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature*, *418*(6898), 646–650. https://doi.org/10.1038/nature00969.Published

Smith, M. W., Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Lomb, D. A., … O'Brien, S. J. (1997). Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort. *Science (New York, N.Y.)*, *277*(5328), 959–965. https://doi.org/10.1126/science.277.5328.959

Stremlau, M., Owens, C. M., Perron, M. J., Kiessling, M., Autissier, P., & Sodroski, J. (2004). The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature*, *427*(6977), 848–853. https://doi.org/10.1038/nature02343

Sun, Y., & Clark, E. A. (1999). Expression of the c-myc proto-oncogene is essential for HIV-1 infection in activated T cells. *J Exp Med*, *189*(9), 1391–1398. https://doi.org/10.1084/jem.189.9.1391

The Lancet. (2017). The global HIV/AIDS epidemic—progress and challenges. *The Lancet*, *390*(10092), 333. https://doi.org/10.1016/S0140-6736(17)31920-7

Turner, B. G., & Summers, M. F. (1999). Structural biology of HIV [Review]. *Journal of*

Molecular Biology, *285*(1), 1–32.

Unaids. (2018). *F a C T S H E E T – J U L Y 201 8 Global Hiv Statistics(2017)*. Retrieved from http://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Research*, *40*(15), 1–12. https://doi.org/10.1093/nar/gks596

Volberding, P. A., & Deeks, S. G. (2010). Antiretroviral therapy and management of HIV infection. *The Lancet*, *376*(9734), 49–62. https://doi.org/10.1016/S0140-6736(10)60676-9

Weiss, R. A. (2013). Thirty years on: HIV receptor gymnastics and the prevention of infection. *BMC Biology*, *11*. https://doi.org/10.1186/1741-7007-11-57

Arthur, L. O., Bess, J. W., Sowder, R. C., Benveniste, R. E., Mann, D. L., Chermann, J. C., & Henderson, L. E. (1992). Cellular proteins bound to immunodeficiency viruses: implications for pathogenesis and vaccines. *Science, 258*(5090), 1935-1938.