Gene expression in the female genital tract and mucosal resistance to HIV-1

By Nadine Kaefer

A thesis submitted to the Faculty of Graduate Studies

The University of Manitoba

In partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Medical Microbiology

University of Manitoba

Winnipeg

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

Gene Expression in the Female Genital Tract and Mucosal Resistance to HIV-1

By

Nadine Kaefer

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Master of Science

Abstract

Studies of the immunological environment in the female genital tract are critical for the development of a mucosal human immunodeficiency virus (HIV) vaccine or microbicides. Studies in this area have been confounded by low immune cell recovery; thus, most HIV research has focused on the circulating immune system. Understanding differences between cervical mononuclear cells (CMCs) and peripheral blood mononuclear cells (PBMCs) would aid our understanding of genital tract immune responses. Here, a non-invasive CMC isolation protocol was refined, yielding sufficient cells for gene expression studies. CMCs were compared to matched PBMCs using Illumina gene expression arrays. There were significant differences in gene expression patterns between CMCs and PBMCs, including inflammatory pathways. CMC gene expression from HIV resistant women was then compared to susceptible women to identify markers of HIV resistance. Although differences between these groups of women were minimal, important pathways involved in the immune response were differentially expressed by HIV resistant women.

Acknowledgements

I would first like to acknowledge my supervisor, Dr. Frank Plummer, along with those who also played a supervisory role, Dr. Blake Ball, Dr. Adrienne Meyers and Dr. Martim Songok. I would also to thank my committee members Dr. Keith Fowke and Dr. Teresa Zelinski for their guidance and input.

I would like to gratefully acknowledge the staff at the Health Sciences Centre, Drs. Guijon, Lotoki, Popowich and Dean, and most especially Caroline MacIntosh and Cindy Bousquet, for allowing me to come to their clinic and recruit patients.

From the Kenyan group, I would like to acknowledge Dr. Joshua Kimani, Dr. Charles Wachihi, Jane Njeri, Jane Kamene, Elizabeth Bwibo, Edith Amatiwa, Anne Maingi, Jemima Nyakio and Dr. Makobu Kimani for their warm reception when I was in Kenya and the invaluable contributions they made to the work in this thesis.

I would like to thank Dr. Rafik Sékaly at l'Université de Montréal for allowing us to run our samples on his Illumina system, as well as Nadia Kettaf, Geneviève Boucher, Peter Wilkinson, Ali Filali and Jean-Phillipe Goulet.

I would also like to thank Dr. Rachel Horton, Dr. James Sainsbury, Dr. Adam Burgener, Dr. Ruey Su, John Rutherford and Leslie Slaney for their technical assistance and guidance throughout this research.

Last (but not least) I would like to thank the students from the group, all of whom made my time here a wonderful experience, but especially Paul McLaren, Aida Sivro, Jennifer Juno, Derek Stein and Were Omange for their input into my project.

This research was funded in part by the Bill and Melinda Gates Foundation, Canadian Institutes of Health Research, and the Manitoba Health Research Council.

Dedication

This thesis is dedicated to the women of the Pumwani cohort.

Table of contents

List of tables:

List of figures

List of copyrighted material for which permission was obtained

Figure 1: The HIV life cycle. Reprinted from Nature Reviews: Cancer, Vol. 4, Morini *et al*, "Antitumour Effects of Antiretroviral Therapy", Page 863, Copyright 2004, with permission from Elsevier (obtained May 7th, 2009). Permission from the author received May 11th, 2009.

Figure 3: A global view of HIV-1 prevalence in 2007, from the UNAIDS 2008 Report on the Global AIDS Epidemic. Permission received April 29th, 2009.

Figure 4: Time course of HIV-1 infection and progression to AIDS, from Microbiology and Immunology Online. Permission received March 11th, 2009.

Figure 5: HIV transmission across the endocervix. Reprinted from The Lancet, Vol. 369, J. Balzarini & L. Van Damme, "Microbicide drug candidates to prevent HIV infection", Page 787, Copyright 2007, with permission from Elsevier (obtained March 12th, 2009). Permission received from the author March 11th, 2009.

Figure 6: Kaplan-Meier survival curve. Reprinted from Immunology Letters, Vol. 66, F.A. Plummer *et al*, "Resistance to HIV-1 infection among highly exposed sex workers in Nairobi: what mediates protection and why does it develop?" Page 28, Copyright 1999, with permission from Elsevier (obtained May 7th 2009). Permission from the author received June 25, 2009.

Figure 12: The apoptosis pathway (modified pathway 04210 from KEGG, http://www.genome.jp/dbget-bin/show_pathway?map04210). Figure reproduced with permission from KEGG via Dr. Yamamoto, obtained May 15th, 2009.

Figure 13: The TLR pathway (modified pathway 04620 from KEGG, http://www.genome.jp/dbget-bin/show_pathway?map04620). Figure reproduced with permission from KEGG via Dr. Yamamoto, obtained May 15th, 2009.

Figure 17: Oxidative phosphorylation (modified pathway 00190 from KEGG, http://www.genome.jp/dbget-bin/show_pathway?map00190). Figure reproduced with permission from KEGG via Dr. Yamamoto, obtained May 15th, 2009.

Figure 18: The MAPK signalling pathway (modified pathway 04010 from KEGG, http://www.genome.jp/dbget-bin/show_pathway?map04010). Figure reproduced with permission from KEGG via Dr. Yamamoto, obtained May 15th, 2009.

1. Introduction

1.1. HIV virus

Human Immunodeficiency Virus (HIV) is the virus that causes Acquired Immune Deficiency Syndrome (AIDS) [1]. HIV is a member of the viral family *Retroviridae,* subfamily Orthoretrovirinae and genus *Lentivirus*. Other members of the *Retroviridae* family include Human T-Lymphotropic virus and Feline Leukemia virus [2] This family of viruses is characterized by the presence of two identical positive, single-stranded ribonucleic acid (RNA) molecules as their genomes and by the presence of the viral reverse transcriptase (RT) enzyme. The *Lentivirus* genus causes diseases characterized by long incubation periods and illnesses of long duration. Other viruses in this genus include Simian Immunodeficiency Virus (SIV) and Equine Infectious Anaemia Virus [2].

1.1.1. HIV classifications

There are two distinct types of HIV, Type 1 and Type 2, which arose from separate zoonotic transmissions of two different SIV ancestral strains [3]. The majority of disease found worldwide is caused by HIV-1 [4], and thus will be the focus of this thesis.

There are three groups of HIV-1, designated M (major), N (new) and O (outlier). M is primarily responsible for causing the global HIV epidemic. There are eight clades from the M group of HIV-1: A, B, C, D, F, G, H and J, as well as a diverse array of circulating recombinant forms (CRFs) derived from the clades [2]. Different clades have different global distributions. B is the major clade found in Canada, while clades A and

D predominate in Kenya [2,5]. Clade D and some CRFs have been associated with a decreased time of progression to AIDS when compared with less virulent clades [6].

1.1.2. Origins of HIV-1

In 1999, a new isolate of SIV called SIVcpzUS was isolated from chimpanzees and found to be highly similar to HIV-1 by sequencing [7]. This SIV strain is now believed to have recombined with another SIVcpz isoform to form the ancestral strain of HIV-1 [7]. It is thus hypothesized that the first HIV-1 infection occurred in Africa by contact with the bodily fluids of an infected chimpanzee of the subspecies *Pan troglodytes troglodytes* [7,8]. Although AIDS was first described in the literature in 1981 [9], the first HIV-1 infections may have occurred much earlier, with one study estimating the first zoonotic transmission of HIV-1 as early as 1933 [10].

1.1.3. The HIV-1 life cycle

Figure 1 outlines the HIV-1 life cycle [11]. The first step in the HIV-1 life cycle is to bind to host target cells [12]. Gp120 proteins on the surface of HIV-1 bind to the host cell cluster of differentiation 4 (CD4) receptor and generally one of two co-receptors, chemokine (C-C motif) receptor 5 (CCR5) or chemokine (CXC motif) receptor 4 (CXCR4) [13]. HIV-1 viruses have specific affinities, or tropisms, for these chemokine co-receptors. In the initial stages of HIV-1 infection, the viruses are typically M-tropic and bind CCR5, which is expressed on macrophages and activated T cells [13,14]. At later stages of infection, viruses become predominately T-tropic and bind CXCR4, which is found on activated T cells [13,14].

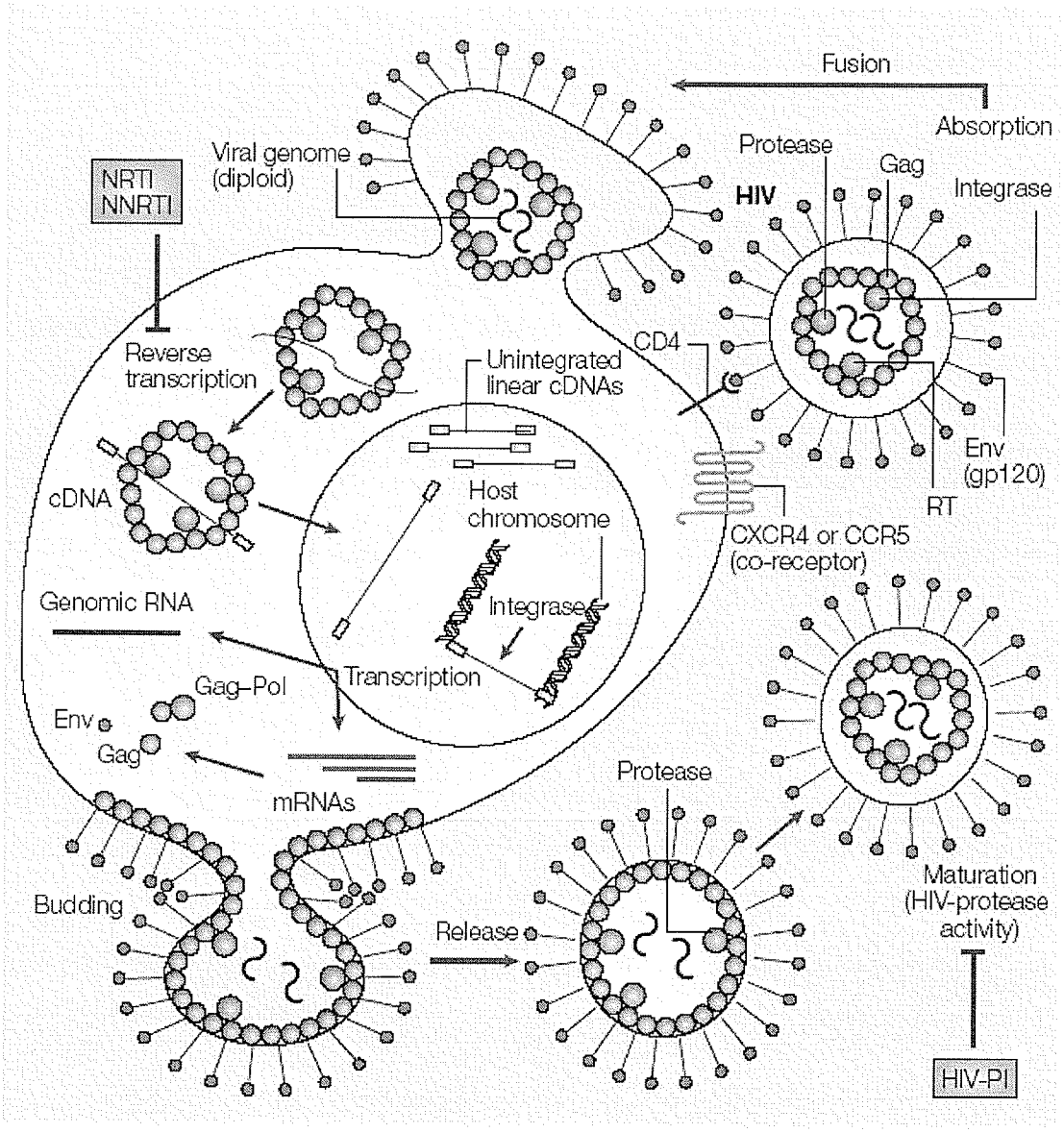


Figure 1: The HIV-1 life cycle [11]. Reprinted from Nature Reviews: Cancer, Vol. 4, Morini *et al*, "Antitumour Effects of Antiretroviral Therapy", Page 863, Copyright 2004, with permission from Elsevier (obtained May 7th, 2009). Permission from the author received May 11th 2009.

Following binding to its receptor and co-receptor, the HIV-1 virion fuses with the host cell and releases its genome and several viral proteins into the cytoplasm [12]. The viral RT converts the single stranded viral RNA genome into double stranded deoxyribonucleic acid (DNA), which is then imported into the nucleus and integrated into the host genome via the viral integrase enzyme, forming the provirus [15,16]. This provirus can remain inactive for several years. Once the cell receives an environmental signal such as tumor necrosis factor alpha (TNFα), nuclear factor kappa-light-chain-enhancer of activated B (NFκB) is transcribed and binds the promoter regions of the HIV-1 genome, inducing HIV-1 genome transcription [17,18]. Host cell-derived RNA polymerase synthesizes RNA copies of the viral genome, which can then be translated and processed into viral proteins. Viral genomes can also be packaged with viral proteins into nascent viral particles. Once new viruses have been made, they then bud out of the host cell, taking part of the host cell membrane with them to make up the viral envelope [11]. Gp41 and gp120 are expressed on the surface of the new viruses and will bind and infect new CD4 cells [2].

HIV-1 is unable to complete its life cycle without the exploitation of a variety of host genes. Three papers published in 2008 identified host factors required by HIV-1 during its life cycle, many of which had not been previously described [19-21]. Although there was little overlap in genes identified by these groups, all three suggest a breadth of host factors that could prove to be potential drug target candidates to limit or prevent the HIV life cycle in infected patients.

1.1.4. The HIV-1 genome

The HIV-1 genome consists of nine genes: group-specific antigen (*gag*),polymerase (*pol*),envelope (*env*), transactivator (*tat*), negative replication factor (*nef*), regulator of virion (*rev*), virion infectivity factor (*vif*), viral protein r (*vpr*) and viral protein u (*vpu*) [2]. Tat–Env–Rev (*Tev*) is another functional HIV-1 gene product, formed by alternative splicing of the HIV-1 genome and the fusion of *tat, rev* and *env* [22]. Long terminal repeat (LTR) regions of the HIV-1 genome flank either side of the genome, and play an important role in the life cycle of HIV-1, although they do not encode proteins. LTRs facilitate integration of the viral DNA into the host genome, and bind host transcription regulatory factors (such as NFκB) and RNA polymerase for transcription of the HIV-1 genome [16,18]. The viral genome is presented in Figure 2.

Gag, pol, and *env* are common to all retroviruses [2]. *Env* is a 2.6 kb (kilobase pair) gene that encodes gp160, which is cleaved with host cell-derived furin to form viral envelope proteins gp120 and gp41 [23]. *Gag* and *pol* are encoded by the 4.3 kb *gag-pol* gene and are synthesized togther as a gag-pol precursor. Gag encodes proteins that make up the viral core, namely capsid (also known as p24), matrix (p17), nucleocapsid (p9) and p6 [2,24]. Pol encodes three proteins important for infecting host cells: RT (with RNase H), integrase and protease [2,24]. RT makes a DNA copy of the RNA genome, forming the provirus, and degrades the RNA template using its RNase H domain [15]. Due to a lack of efficient proof-reading ability, RT makes errors during DNA transcription, leading to high viral mutation rates of the virus and one mechanism HIV-1 uses to escape the host immune response [25]. Integrase is responsible for inserting the DNA into the host genome [26]. Finally, protease processes gag-pol into its substituent proteins [24].

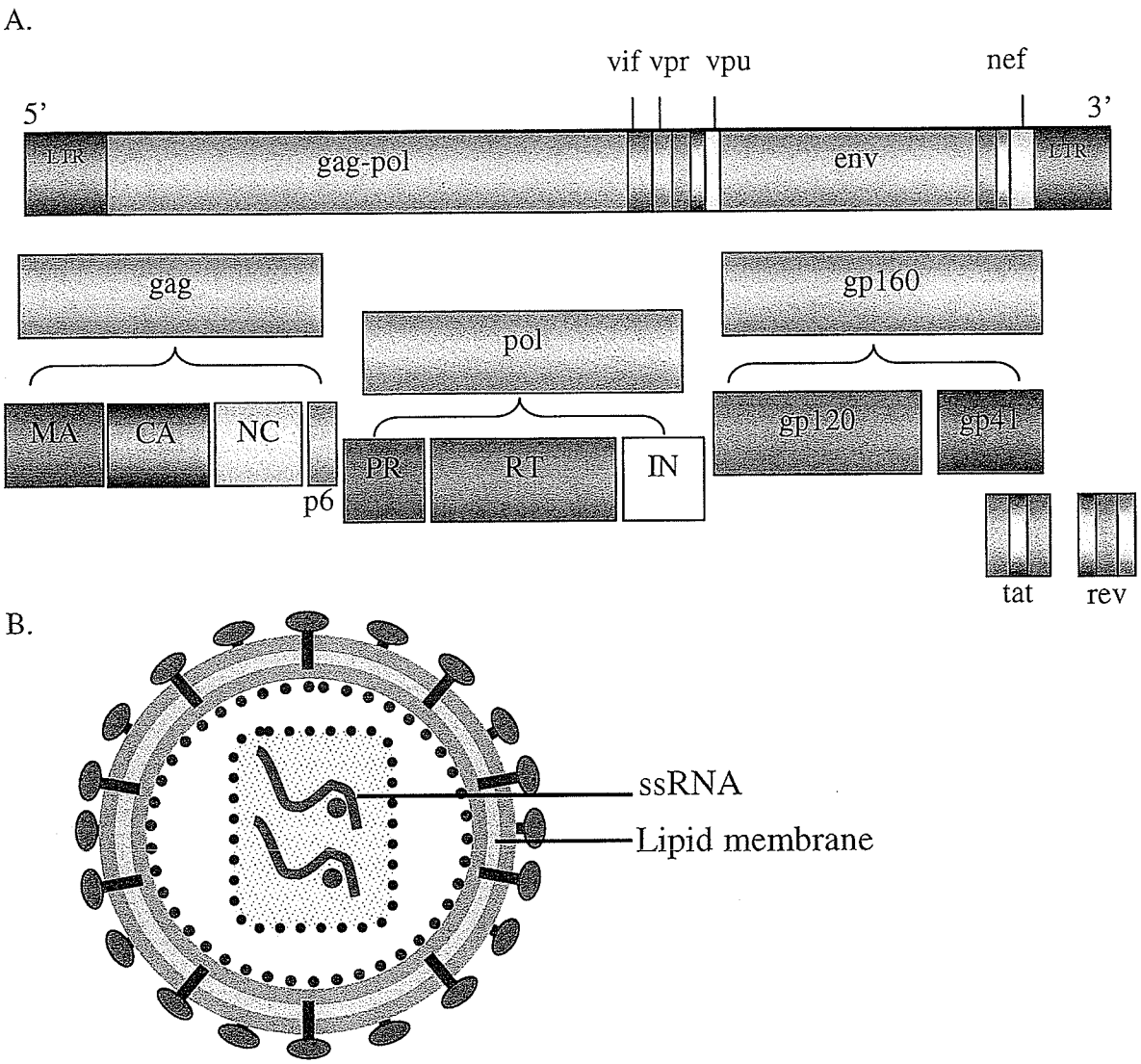


Figure 2: Cartoon schematic of the HIV-1 viral genome and virion. (A) Bars represent HIV-1 proteins and where their corresponding genes are located in the HIV-1 genome. Grey bars represent precursor proteins, and coloured bars represent final protein products matrix (MA), capsid (CA), nucleocapsid (NC), protease (PR), reverse transcriptase (RT), integrase (IN). The locations of the proteins in an assembled HIV-1 virus are shown in (B), along with the HIV-1 lipid membrane and single stranded RNA (ssRNA) genome. (MA = ●, CA = ●, NC = ▭, gp120 = ⬭, gp41 = ▬).

Tat (2.6 kb) and *nef* (0.6 kb) are two products of alternative splicing of the HIV-1 genome (Figure 2). The t*at* gene encodes two proteins that bind the first 59 base pairs (bp) of nascent HIV-1 strands (the transactivator active region), and enhance the transcription of the downstream HIV-1 genome [27]. The n*ef* gene encodes a protein that has been shown to enhance the endocytocis and degradation of CD4 and Major Histocompatibility Complex (MHC) Class I receptors to enhance cell survival [28]. The *rev* gene (2.7 kb) encodes a protein that aids in exporting HIV-1 messenger ribonucleic acid (mRNA) out of the cell nucleus before it undergoes alternative splicing into the mRNAs encoding tat, nef and tev [29]. It acts by binding rev-response elements (RREs) in the HIV-1 genome and then shuttles the complex out of the nucleus. This inhibits the formation of tat and nef proteins, but favours the production of proteins encoded by *gag*, *pol*, and *env* [29]. The protein product of the *vif* gene (0.5 kb) increases the infectivity of HIV-1 by interfering with the activity of host anti-viral protein apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G (APOBEC3G) [30]. The v*pr* gene (291 bp) encodes vpr, which accelerates the production of HIV-1 proteins. It contains nuclear localization signals and mimics importin-β, allowing for the nuclear import of the pre-integration complex (RT, viral RNA and integrase). Vpr has also been shown to stop cells in the G2 phase of the cell cycle [31]. The v*pu* gene (248 bp) encodes a protein that helps in the assembly of new virus particles and aids in their budding out of the cell [32]. *Vpu* has also been implicated in the degradation of CD4 and helps the infected cell survive to produce more viral particles [32]. The role of *tev* in the HIV-1 life cycle has not been well described in the literature, but it has tat and rev functionality *in vitro* and therefore plays a role in the initial regulation of virus expression [22].

1.2. HIV/AIDS epidemic

AIDS is a fatal disease caused by infection with HIV. Since its initial description in 1981 in small groups of men in California and New York, AIDS has reached epidemic proportions worldwide [9,33]. The 2008 Joint United Nations Programme on HIV/AIDS (UNAIDS) report on the global HIV/AIDS epidemic has put the most recent estimates of people living with HIV-1 at 33 to 36 million [4]. Sub-Saharan Africa is the region hardest hit by this epidemic, accounting for 67% of all infections and 75% of AIDS-related deaths (Figure 3) [4].

AIDS is clinically defined as a syndrome comprised of a CD4 count below 200 cells/ml and the presence of opportunistic infections such as Human Herpesvirus 8 (the causative agent of Kaposi's Sarcoma), *Pneumocystis jiroveci* pneumonia and *Candidiasis* (yeast) infections of the respiratory tract, as well as a positive HIV test [2,9]. Figure 4 shows the relationship between HIV-1 virus titres, CD4 counts, and the development of anti-gp120 antibodies from initial HIV-1 infection to the development of AIDS [34]. Although infected patients are capable of mounting an effective cell-mediated and humoral immune response to the HIV-1 infection, eventually the immune system is exhausted and AIDS develops. Compromise of the immune system makes the AIDS patient susceptible to opportunistic infections and certain malignancies that individuals with healthy immune responses are usually able to eliminate.

Several factors influence how long it takes an HIV-1-positive patient to progress to AIDS. The clade of virus the patient is infected with has been shown to play a role, as well as host factors, and will be discussed in a later section of this thesis [6].

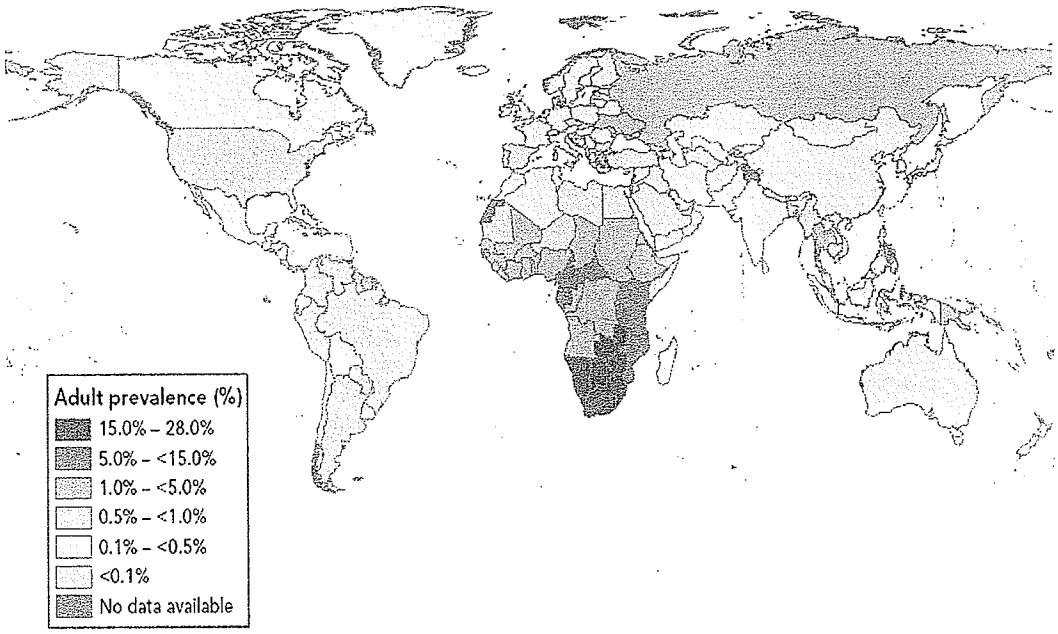


Figure 3: A global view of HIV-1 prevalence in 2007 [4]. Figure reproduced with permission from the 2008 UNAIDS Report on the global AIDS epidemic. Sub-Saharan African countries have some of the highest prevalence rates in the world. Permission obtained April 29th, 2009.

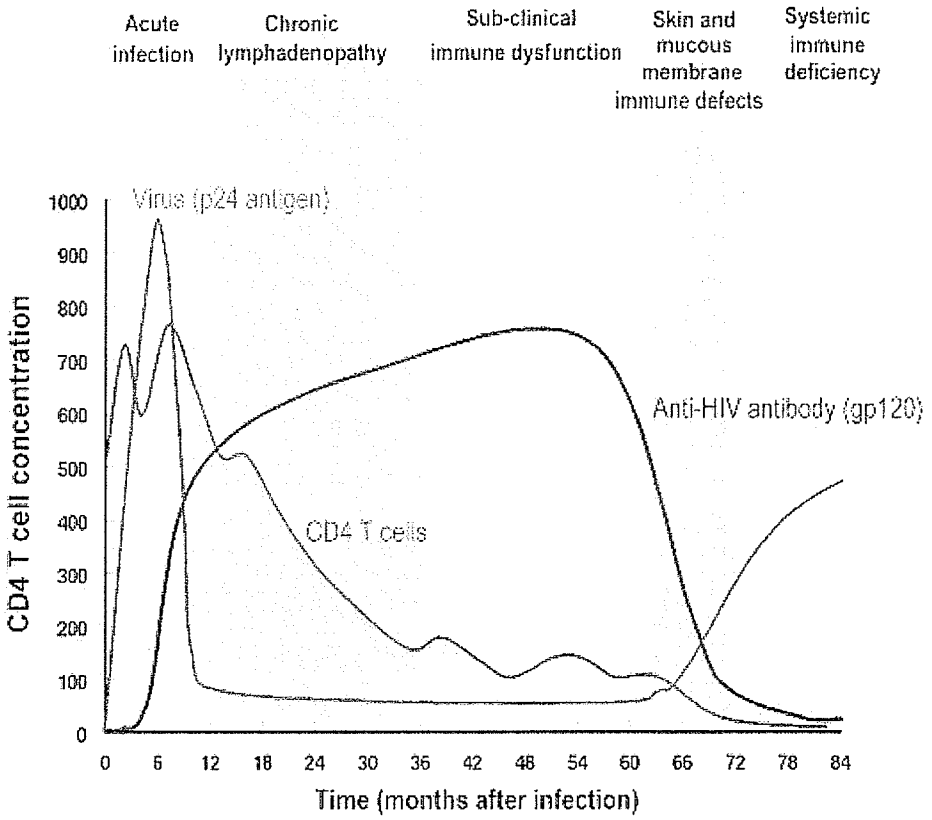


Figure 4: Time course of HIV-1 infection and progression to AIDS. Reproduced with permission from Microbiology and Immunology Online (March 11th, 2009) [34].

1.2.1. Routes of HIV-1 Transmission

HIV-1 is transmitted through direct contact with infected bodily fluids such as blood, semen and vaginal fluids. This is achieved through unprotected sexual contact, by direct inoculation with contaminated needles, by blood transfusions, and from mother to child during delivery or while breastfeeding. In Sub-Saharan Africa, the primary route of HIV-1 infection is through heterosexual intercourse [4].

1.2.2. Treatment

Although there is currently no cure or vaccine against HIV-1, there are many treatment regimes that significantly increase the lifespan of HIV-1 infected individuals [35]. Antiretroviral therapy (ART) is an umbrella term describing several classes of drugs that act against HIV-1 at various stages of the viral life cycle (shown in Figure 1) [36]. Suppressing the viral levels in an infected person allows the immune system to recover and prevents infection by opportunistic pathogens [35]. Different classes of antiretroviral drugs are often taken together to target multiple phases of the virus life cycle. Treatment with combinations of anti-HIV-1 drugs is called Highly Active Antiretroviral Therapy (HAART) [36,37].

The first class of drugs to be discovered and approved for use in patients were nucleoside reverse transcriptase inhibitors (NRTIs) [38]. NRTIs interfere with the activities of RT by mimicking nucleotides, blocking bond formation and ultimately lead to the formation of defective HIV-1 genomes [38]. A second class of drug also interfere with the activity of RT, called non-nucleoside reverse transcriptase inhibitors (NNRTIs).

NNRTIs bind RT and prevent it from moving along the nascent HIV-1 genome, blocking subsequent transcription and effectively freezing it in place [39]. A third class of drug is the protease inhibitors (PIs). These drugs prevent the processing of HIV-1 proteins into their active forms by HIV-1 protease and therefore block the activities and assembly of the HIV-1 virus [40]. Two relatively new classes of drugs are fusion inhibitors and integrase inhibitors. Fusion inhibitors block the initial association of the virus with its target cell, by either blocking one of its receptors on the host cell, such as CD4, CCR5 or CXCR4 (e.g. the CCR5-binding inhibitor Maraviroc) or by binding one of the proteins on the viral surface involved in gaining cell entry, such as gp120 or gp41 (e.g. Enfuvirtide, which binds gp41) [41,42]. Integrase inhibitors (e.g. raltegravir) work by interfering with the activity of integrase, preventing the insertion of the viral genome into host DNA. For example, raltegravir inhibits the strand transfer activity of integrase [43].

Despite the gains made by these drugs in the fight against the HIV/AIDS epidemic, the HIV-1 virus mutates rapidly, reducing the efficacy of the virus-drug interaction and results in drug resistant strains. Furthermore, the aforementioned treatments can often be toxic and have unpleasant side effects [36], and not all HIV-1 infected patients have access to antiretrovirals (ARVs), especially in developing countries [4]. Consequently, an effective vaccine and/or microbicide are still needed to stop the spread of this disease.

1.2.3. Women and HIV-1/AIDS

Interestingly, the rate of transmission of HIV-1 from men to women is actually quite low (approximately 1 in 200 to 1 in 2000 per coital act); indicating a natural and robust innate immunity to HIV-1 in the female genital tract (FGT) [44]. The fact remains,

however, that in some parts of the world, women represent a disproportionately high percentage of people infected with HIV-1 [45]. In 2007, 60% of new infections in Sub-Saharan Africa occurred in women, the vast majority of which were contracted by heterosexual intercourse [4]. Thus, natural immunity to HIV-1 in the FGT is not sufficient to protect women from this virus.

An important factor leading to this increased incidence in women is the lack of effective female-controlled sexually transmitted infection (STI) prevention methods (with the exception of the uncommonly used "female condom") [45]. Not all women are capable of negotiating condom usage, due to cultural mores or financial restraints in the sex trade, and thus are at a high risk of exposure to HIV-1 during heterosexual intercourse [4]. Young women appear to be at increased risk of contracting HIV-1, which may be the result of the innate physiological properties of an immature genital tract or epidemiological factors [45]. Furthermore, some studies have linked the use of hormonal contraception and pregnancy with an increased risk with contracting HIV-1; however, there is currently much debate about these studies in the literature [46-48].

1.2.3.1. HIV-1 transmission across the FGT

The primary deterrent to HIV-1 infection in the FGT is the epithelial layer. In the vagina and ectocervix, the epithelial wall consists of a thick layer of stratified squamous epithelial cells. In the endocervix, however, the epithelial layer is much thinner, consisting of one cell layer of columnar epithelial cells [44]. The endocervix and transformation zone (TZ) between the ecto- and endocervix are believed to be the primary sites of HIV-1 transmission in the FGT [49]; however, Thomas Hope *et al.* has shown HIV-1 can cross an intact stratified vaginal epithelial layer [50].

Once the virus has crossed the FGT epithelial layer, it reaches the lamina propria, which is rich in HIV-1 target cells such as CD4+ T cells. There are five mechanisms described in the literature that are used by HIV-1 to traverse the endocervical layer to establish an infection (Figure 5) [51]. The easiest way for HIV-1 to enter the host is through breaches in the epithelial layer. Breaches can be caused by the presence of other STIs that cause ulcers, such as gonorrhoea, or through trauma to the FGT [44]. HIV-1 has also been shown to cross intact cervical epithelia through infection of the epithelial cells themselves or by transcytocis of the virus by epithelial cells [52]. HIV-1 can also be transmitted via transmigration of HIV-1-infected donor cells across the epithelial layer, or by uptake by sentinel dendritic cells present in the epithelial layer [51]. The role of dendritic cells, especially the subset expressing the cell surface protein Langerin ("Langerhans cells") in permitting HIV-1 infection in the FGT is unclear, as one study has shown the uptake of HIV-1 by Langerhan's cells may provide protection against infection by sequestering HIV-1 virions [53], while another study suggests these cells may become infected and transmit the virus [54]. Once HIV-1 has passed the initial epithelial barrier to find and infect its target cells, infected cells can then migrate to draining lymph nodes, and from there to other lymphoid tissue and finally dissemination in the bloodstream [55]. The time HIV-1 takes to migrate to the draining lymph node varies between studies, from 24 hours [56] to one week [57] after vaginal SIV inoculation in primate models, and Dr. Ashley Haase has suggested this time to dissemination may be a function of the size of the founder HIV-1 virus population in the FGT [55].

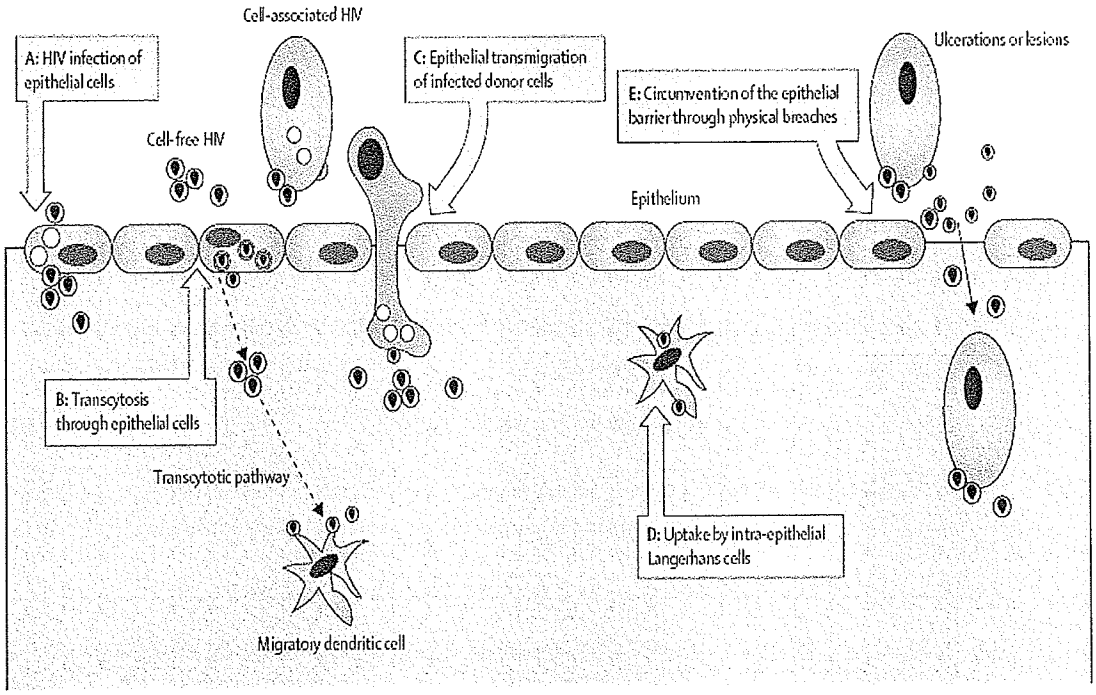


Figure 5: HIV-1 transmission across the endocervix. [51] Reprinted from The Lancet, Vol. 369, J. Balzarini & L. Van Damme, "Microbicide drug candidates to prevent HIV-1 infection", Page 787, Copyright 2007, with permission from Elsevier (obtained March 12th, 2009). Permission received from the author March 11th, 2009.

1.2.3.2. Microbicides

This high rate of infection in women indicates a significant need for the development of a female-controlled method to prevent HIV-1 transmission. One promising method is the development of microbicides and much research has been focused in this area in recent years. There have been many challenges in this field, and the several failed microbicide trials show a lack of basic knowledge of the immunobiology of the FGT [58-62].

Six microbicides have undergone large-scale clinical trials. The first candidate tested was nonoxynol-9 (N9), an anionic surfactant that is currently available as an antimicrobial and in spermicides from several companies. This candidate was tested first as it had been shown to disrupt lipid membranes of enveloped viruses such as HIV-1 [63]. However, trials were stopped when it became apparent the women receiving N9 actually experienced higher rates of infection over women using a placebo when used more than three times per day [60]. Trials testing the polyanionic microbicide candidate cellulose sulphate (Polydex Pharmaceuticals) were also halted in 2007 because women using this microbicide candidate were found to be at a higher risk of contracting HIV-1 over women using a placebo [60].

Carraguard (FMC Biopolymer) is a carrageenan-based microbicide whose trials ended in unsuccessfully 2007 with no safety concerns [60]. Carrageenan is a sulphated polysaccharide that also disrupts the lipid membrane envelopes of HIV-1 viruses [63]. Despite the safety of this microbicide candidate, there was no decrease in HIV-1 transmission in the test group over the placebo group [62]. Microbicide trials testing the antimicrobial agent from Cellegy Pharmaceuticals known as C31G were stopped as this

candidate was also found to have no effect in decreasing susceptibility to HIV-1 infection [60].

PRO2000® from Indevus Pharmaceuticals is a sulfonated polyanionic compound that interferes with the interaction between HIV-1 and dendritic cells present in the FGT, thus blocking infection by HIV-1 [64]. Clinical trials with PRO2000® showed the group of women using this microbicide candidate had a 30% lower risk of HIV-1 infection than women using a placebo.

Finally, BufferGel® (ReProtect, Inc.) has a different mode of action than the previously described microbicides. It is an acidic-buffering system that makes the FGT less hospitable to pathogens [65]. However, in clinical trials it was not shown to reduce the risk of HIV-1 transmission to women during vaginal intercourse.

In addition to the above described microbicide candidates, there many new formulations described in the literature and many more in the pipeline. The efficacy and safety of the new compounds must be determined by clinical trials. One promising candidate, glycerol monolaurate, was recently shown to decrease SIV transmission in rhesus macaque monkeys by decreasing the inflammatory response associated with SIV challenge [66]. Several groups are also working on ARV-based microbicides, such as the inclusion of the HIV-1 entry inhibitor griffithsin in a topical microbicide [67], or by the sustained release of ARVs by vaginal rings [68]. However, as demonstrated by the failure of the N9 and cellulose sulphate trials, more knowledge of the FGT is needed to prevent such failures from happening in the future. Identification of host innate factors that successfully block HIV-1 transmission, such as tetherin, APOBEC3G, and tripartite motif-containing 5 (TRIM5α), is an important and exciting new prospect in the

microbicide field [69-71]. As these compounds would be extensions of the pre-existing immune environment present in the FGT, they could potentially have fewer side effects and provide better protection against HIV-1 transmission.

1.3. The human immune system

The human immune system can be divided into two general types of responses, namely the innate and adaptive immune responses [72]. It can also be divided into systemic and mucosal compartments [73]. The systemic compartment is the circulatory immune system, or the immune response in the blood. Mucosal immune responses occur at surfaces that are points of contact for pathogens, for example the mouth, gut, respiratory tract, or the FGT [72].

1.3.1. Innate immune response

The innate immune response is the first line of defence against microbial infections. It is rapid, less specific than the adaptive immune response, and can recognize general pathogen-associated molecular patterns (PAMPs) [72]. The first and most important components in the innate immune response are barriers. Epithelial cells provide an excellent barrier against microbial infection by physically blocking access to target cells [72]. They also express receptors that recognize PAMPs, such as toll-like receptors (TLRs), and can express innate factors such as defensins and other antimicrobial peptides. Defensins provide another barrier to microbial infection [72]. Defensins are charged particles that cause pores to form in bacterial and fungal cell membranes, causing

cell death. Defensins are also effective in eliminating viruses, and have been shown to inhibit HIV-1 replication [74].

1.3.1.1. Toll-like receptor signalling

Toll was first described in *Drosophila melanogaster*, and its first human homologue, TLR (now known as TLR4), was described by Medzhitov and Janeway in 1997 [75]. TLRs consist of external and internal binding domains. The external binding domain consists of tandem leucine-rich repeats and recognizes PAMPs [75]. The cytoplasmic tail of a TLR has Toll/interleukin-1 (TIR) receptor domain [75]. TIR domains interact with and signal through the TIR-binding domain of either the myeloid differentiation primary response gene (MyD88), which signals through the My-D88-dependent pathway, the TIR domain-containing adaptor protein (TIRAP)/Mal, which signals via the MyD88-dependent and independent pathways, or through the TIR-domain-containing adapter-inducing interferon-β (TRIF) protein in the MyD88-independent pathway [76].

TRIF-mediated TLR signalling leads to the activation of the interferon (IFN)-β promoter and the NFκB-dependent promoter [77]. This results in the production of inflammatory cytokines as well as IFN-β, leading to the formation of an anti-viral state [77]. TLR signalling through MyD88 and TIRAP/Mal leads to the induction of the NFκB-dependent promoter and production of inflammatory cytokines [78-80].

TLR2 is an important molecule in the recognition of peptidoglycan from gram-positive bacterial species [81] and signals through the MyD88-dependent pathway [82]. TLR1 associates with TLR2 and recognizes triacyl lipopeptides, a component of the *Mycobacterial* cell wall [83]. TLR2 also associates with TLR6 to recognize diacyl

lipopeptides, a surface component of *Mycoplasma* spp. [84]. Interestingly, *Lactobacillus* spp. are gram-positive bacteria and are the dominant component of the vaginal microflora and therefore could potentially be recognized by TLR2 [85]. TLR5 recognizes bacterial flagella and also signals through the MyD88-dependent signalling pathway [86].

TLR3 recognises double-stranded RNA [87], and TLRs 7 and 8 recognise short sequences of single-stranded RNA [88]. These three isoforms are therefore important in viral PAMP recognition. TRIF binds the TIR domain of TLR3 and consequently signals through the MyD88-independent pathway [76]. Signalling initiated through TLR3 therefore leads to the production of IFNβ and an anti-viral state [87]. A recent study with Herpes Simplex Virus type-2 (HSV-2) showed triggering TLR3 successfully inhibited HSV-2 replication but also suggested this activation could increase susceptibility to HIV-1 infection due to increased recruitment of activated T cells [89]. TLR7/8 signals through MyD88 and leads to the production of inflammatory cytokines [88]. This pathway also leads to Type I interferon production by IRF7 (interferon regulatory factor 7), and consequently produces an anti-viral state in a MyD88-dependent, TRIF-independent mechanism [76]. Although these molecules lead to the induction of an anti-viral state, activation of TLR7 has been implicated in HIV-1 pathogenesis [90]. HIV-1 transcription is induced by NFκB, and NFκB is induced by the MyD88-dependent TLR signalling pathway. Consequently, increased TLR7 expression in the FGT could increase HIV-1 infection susceptibility and thus would not make a good microbicide target [17,18,90].

TLR4 binds lipopolysaccharide (LPS), a component of the outer membranes of gram-negative bacteria [91]. TLR4 is capable of signalling by MyD88, TIRAP/Mal and

TRIF and consequently leads to the production of inflammatory cytokines and IFN-β [92]. Interestingly, TLR4 is the only bacterial surface PAMP-responsive TLR that induces IFN-β, but may also recognize viral-associated proteins [93]. Several STIs are associated with gram-negative bacteria, including *Neisseria gonnorhea* and *Chlamydia trachomatis*, the causative agents of gonorrhoea and chlamydia, respectively, both of which have been associated with increased susceptibility to HIV-1 infection [94,95].

TLR9 recognizes CpG DNA from both bacterial and viral pathogens [96], and signals via MyD88 [76]. TLR9 produces an anti-viral state using the same mechanism as TLR7/8 [76]. Human Papillomavirus (HPV) has been shown to down regulate TLR9 [97].

1.3.1.2. Cell-mediated innate immunity

There is also a cell-mediated component to the innate immune response. Neutrophils, macrophages, natural killer (NK) cells and dendritic cells (DCs) all play a role in the relatively non-specific killing of foreign organisms [72]. Neutrophils, macrophages and DCs engulf pathogens and destroy them intracellularly. Neutrophils reside in the circulatory immune system and are one of first responders at the site of infection, whereas macrophages tend to arrive later [98]. Dendritic cells are often found localized at mucosal surfaces and act as sentinels [99]. Macrophages and DCs are capable of presenting antigens from bacterial pathogens to the adaptive immune response on Major Histocompatibility Complex (MHC) Class II receptors, triggering the next phase in the immune response [72]. NK cells recognize infected host cells via surface receptors [100]. Activation receptors recognize host stress molecules, and inhibitory receptors recognize Human Leukocyte Antigen (HLA) and MHC markers. The absence

of MHC or HLA tends to be a sign of viral infection, and leads to the disinhibition of NK cells and the thus activation of NK-cell mediated cytotoxicity. NK cells then go on to kill infected host cells, allowing macrophages and DCs to ingest the infecting microbes and link to the adaptive immune response [100].

1.3.1.3. Effector proteins and cytokines

Effector proteins also play an important role in the innate immune response [72]. Complement cascade proteins are found in plasma and cause the lysis of microbes. Side products of the complement cascade can act as signalling molecules and recruit the cell-mediated aspect of the innate immune response to the site of infection [72]. Complement proteins also coat the surface of pathogens, making them easier to be engulfed by neutrophils and macrophages in a process called opsonisation. Other innate proteins involved in opsonisation are the serum proteins C-reactive protein and mannose-binding lectin [72].

Cells of the innate immune system also secrete small proteins called cytokines and chemokines when they are stimulated by the presence of a microbial pathogen. Cytokines and chemokines play several roles in the regulation of the immune response. They act as secondary signals and their presence is critical for the proper activation of the adaptive immune response. They also play roles in augmenting or decreasing the magnitude of the immune response [72]. Cytokines are important coordinators of the immune system and are expressed by nearly all the cells in the body [72]. Upon receiving environmental signals, cytokines are released and bind to receptors on their target cells and trigger intracellular signalling. Cytokines can be loosely grouped into three families based on their tertiary structures: the hematopoietins, the interferons and

the TNF family [72]. Chemokines are a class of cytokines that are primarily involved in chemoattraction [101]. Subclasses of chemokines include the CXC, CC, C and CX_3C families [101]. Cytokine receptors can be divided into four families: Class I, Class II, TNF receptor and the chemokine receptor family [72]. Another useful way to classify cytokines is based on their functionality. Cytokines play a role in the inflammatory response [102] and are also important in T and B cell proliferation and differentiation [103,104]. This thesis focuses on the roles of cytokines in the inflammatory response.

There is no distinct boundary between pro- and anti-inflammatory cytokines, and some cytokines play a role in both processes; however, these are common terms when describing the function of cytokines. Pro-inflammatory cytokines have important roles in activating the inflammatory response [102]. One important trigger of the inflammatory response is the TLR signalling pathway. TLRs are stimulated by PAMPs, leading to intracellular signalling transduction and the eventual transcription of NFκB. NFκB then activates transcription of pro-inflammatory cytokines [80].

The inflammatory response plays an important role in the immune system. Inflammation is characterized by heat, pain, redness and swelling [105]. At the onset of an infection, it ensures the cells of the immune system reach the site of infection. To cause this, pro-inflammatory cytokines such as IL1 and IFNγ increase vascular permeability [106]. IL1 and TNF function to increase the adhesive properties of the endothelial wall by triggering them to increase expression of E- and P-selectins. These selectins interact with surface markers of neutrophils, allowing them to bind the endothelial wall and exit the vascular system [107]. Chemokines, such as MIP-1α/β, IL8 and CCL1, then act as chemoattractants, causing the neutrophils and the later-arriving

monocytes to migrate to the site of infection. Once they have arrived, neutrophils and macrophages (mature monocytes) engulf pathogens and further secrete inflammatory cytokines to attract other cells to the site. Resident dendritic cells also engulf pathogens, at which point they mature and travel to a nearby lymph node to present antigens to lymphocytes [72].

Although the inflammatory process is important in mounting an effective response to pathogens, over-activation of this process leads to cell and tissue damage. High levels of inflammatory cytokines and tissue damage have been suggested to increase susceptibility to HIV infection [108]. Anti-inflammatory cytokines are important to mediate the effects of the inflammatory response. These cytokines generally operate by suppressing the transcription of cytokines or by antagonizing their activity [109].

1.3.2. Adaptive immune response

The adaptive immune response is larger in magnitude and more specific than the innate immune response [72]. This arm of the immune response also has a memory component, whereby if the same pathogen infects again, the second response is greater and faster than the initial response. The innate immune response provides two signals for the activation of the adaptive immune response [110]. Antigen presentation by macrophages and DCs are important to direct specificity to a specific pathogen, but a second signal is also required to ensure only the proper activation of the adaptive immune response [110]. One example of a secondary signal is CD80, which is up-regulated by monocytes upon activation of TLRs [72]. The adaptive immune response is divided into

two components: cell-mediated and humoral immunity. Important cell types in the cell-mediated arm of the adaptive immune response are T and B cells

MHC (called HLA in humans) receptors are important molecules in the immune response. MHC Class I receptors are found on all nucleated cells and present normal "self" peptides (or "antigens"). When a cell is infected with an intracellular pathogen, MHC Class I receptors display non-self antigens and thus provide evidence to the immune system that the cell must be eliminated. MHC Class II receptors are only found on professional antigen presenting cells (APCs) and activated T cells. APCs link the innate and adaptive response, ensuring a highly specific and effective immune response to a pathogen [72].

T cells are primarily responsible for the cell-mediated adaptive immune response. There are two kinds of T cells: CD8 positive (CD8+) and CD4 positive (CD4+), so named due to markers on their cell surfaces. CD4+ cells bind MHC Class I and II receptors on professional APCs through their T cell receptors and become activated if their antigens are recognized. CD8+ cells are also referred to as cytotoxic T lymphocytes (CTLs), as they recognize and kill infected cells presenting foreign antigen on their MHC Class I receptors. CD4+ cells are also referred to as T helper cells (T_H) because they coordinate the adaptive immune response. There are two T_H subsets: T_H1 and T_H2. In general, T_H1 cells assist CD8+ cells in eliminating infected host cells, and T_H2 cells assist B cells, which produce antibodies [72]. CD4+ cells are also the cells that are targeted by HIV-1 [2].

The humoral immune response is provided by the actions of antibodies. Antibodies are proteins that bind specific antigens and are secreted by B cells. Antibodies play

several important roles in the elimination of microbes. First, they can neutralize pathogens by binding them, preventing them from entering a cell. Second, they can protect the host by binding toxins produced by infecting organisms. Antibodies can also lead to the elimination of infected host cells by binding them and attracting NK cells. Finally, antibodies can also coat the surface of bacteria, for opsonisation or to activate the complement cascade [72].

1.3.3. Vaccines

The principle behind vaccination is to provide safe exposure to a pathogen or part of a pathogen and prime the immune response so that at a subsequent exposure the adaptive immune response will be quicker and of larger magnitude. There have been several HIV-1 vaccine trials; however, none have provided protection, and the Merck STEP trial may have increased susceptibility to HIV-1 infection in people receiving the vaccine over control groups [111]. HIV-1 vaccine research is confounded by many issues, such as the breadth of HIV-1 sequence diversity across clades, which makes epitope selection challenging, and the adaptability of the virus that allows it to evade the immune response [112]. Furthermore, there is no clear understanding of which epitopes would provide the best protection. Finally, as HIV-1 infects activated immune cells, vaccination could put people at risk for infection [112]. An effective vaccine is critical to stop the spread of HIV-1, but it is also important to ensure the vaccine provides protection at the site of infection. For Sub-Saharan African women, this means a vaccine capable of producing a safe and robust immune response in the FGT.

1.3.4. Mucosal immunity

The FGT is a mucosal immune site. Generally, mucosal sites are primarily protected by innate immune mechanisms, but adaptive immune responses are also present [72]. Like most mucosal sites, there must be a balance between the induction of a full immune response in the presence of a pathogen, but not in the presence of benign or essential external factors such as sperm or natural flora [99].

Cell-mediated aspects of the innate and adaptive immune responses play important roles in immune protection in the FGT. High levels of macrophages and DCs are found in the FGT, and are at particularly high concentrations in the TZ between the ecto- and endocervix [113]. Intraepithelial lymphocytes (IELs) are a heterogeneous group of lymphocytes present at mucosal surfaces [114]. The distribution of cells varies widely throughout the FGT, but appear to be at their highest concentration in the TZ [113]. One study on the histopathology of cells in the TZ showed an abundance of CD8+ and CD4+ IELs, indicating this site as a potential target for HIV-1 infection *in vivo* [113].

There is also an antibody-mediated aspect to mucosal immunity. In general, the most abundant type of antibody present at mucosal surfaces is IgA; however, secretory IgG is the dominant isotype in the FGT [115].

The natural barriers present at the FGT alone are clearly insufficient to protect against all FGT infections. However, the FGT is necessarily a site of immunological balance. Not only must the host be protected from pathogens, but this site must also be permissible to sperm and natural flora. Consequently, care must be taken to ensure that any alterations to immunological activity at this site (for example, the application of a microbicide) do not upset this delicate balance [116].

1.3.4.1. Tolerance

Mucosal sites represent points of contact between the host and the environment. Consequently, there is near constant antigenic challenge at these sites, most of which is caused by harmless environmental factors. Mucosal surfaces thus tend to take an "immunotolerant profile", or a state where the presence of antigens does not always trigger an immune response [99]. The mechanisms causing mucosal sites to adopt an immunotolerant profile have not been fully described in the literature. One suggested mechanism for the induction of tolerance at gut mucosal surfaces is tolerogenic DCs. DCs are abundant at mucosal surfaces, and are constantly processing antigens from these sites [99]. Semi-activated DCs express low levels of co-stimulatory molecules (CD80, CD86 and CD40) and sample the environment continuously. When the incompletely activated DCs travel to their draining lymph node, they present antigens to T cells and cause them to undergo apoptosis or differentiate into regulatory T cells [117]. Although this phenomenon has not been described in the FGT, it is reasonable to anticipate a similar mechanism to work at this site.

1.3.4.2. Mucosal immune responses to HIV-1

Several studies have documented host factors present in the FGT that play a role in both protection from and susceptibility to infection, including the role of natural microbiota [116,118]. A low potential of hydrogen (pH) and a diverse array of anti-microbial peptides provide natural protection against some pathogens [118,119], and it is reasonable to assume that similar factors may protect against HIV-1. Mucosal responses against HIV-1 are best characterized in individuals who remain relatively resistant to

HIV-1 infection despite repeated exposure. Mechanisms of HIV-1 resistance will be discussed in a later section.

1.4. HIV-1 resistance

HIV-1 resistance is defined as remaining HIV-1-uninfected despite documented exposure to HIV-1 [95]. This phenomena is documented in many groups: commercial sex workers [95], discordant couples (where one partner is HIV-1 positive and the other is not, and the couple continues to have unprotected sex [120]), newborns from HIV-1-infected mothers [121], injection drug users [122], exposed health care workers [123], and people who engage in other forms of higher risk sexual activities (e.g. men who have unprotected sex with men [124]).

Long term non-progressors (LTNPs) are a group of individuals who are HIV-1 infected and progress more slowly to AIDS than other HIV-1 infected patients. LTNPs can act as another model of HIV-1 resistance because, although they are HIV-1 infected, they are resistant to disease progression. These individuals sometimes share immunological features with HIV-1 resistant women, and represent another important study group to understand the resistance and susceptibility to HIV-1 infection and pathogenesis [125].

1.4.1. Correlates of protection

There is no one "resistance phenotype" across all highly exposed, persistently seronegative (HEPS) cohorts [125]. However, different HEPS groups do tend to show

immune responses to HIV-1, including cell-mediated, humoral and innate immune responses. Interestingly, the maintenance of HIV-1 resistance seems to require continued exposure, and several studies have documented the loss of HIV-1 resistance after a period of decreased HIV-1 exposure [125,126]. Some studies suggest increased immune activation as a potential mechanism for HIV-1 resistance [127,128] while other studies suggest lower basal immune activation is important for the maintenance of the HEPS phenotype [125,129].

1.4.1.1. Innate immune and genetic markers of HIV-1 resistance

Several alleles and genetic polymorphisms for genes encoding parts of the innate immune system have been described as reducing susceptibility to HIV-1 infection as well as slowed disease progression. Interestingly, some polymorphisms have different impacts, depending on the ethnic background of the HEPS cohorts studied [130]. Some alleles encoding receptors for NK cells [131] and polymorphisms in the promoter region of the Regulated on Activation Normal T Cell Expressed and Secreted (RANTES) gene, a chemokine whose receptor is also CCR5 [130], have been correlated to decreased rate of disease progression. Increased RANTES levels have also been correlated with HIV-1 resistance [132]. CCL3L1 is another ligand for CCR5, and increased copy numbers of this gene are correlated with reduced susceptibility to HIV-1 infection, and lead to a decrease in CCR5 expression [133]. Polymorphisms in other chemokine genes have also been correlated with altered susceptibility to HIV-1 infection in a variety of cohorts [130]. Increased expression of APOBEC3G [134], acute-phase amyloid A protein [135], IL-22 [135], and polymorphisms in the TRIM5α gene [136] have also been correlated with decreased susceptibility to HIV-1 infection and slower disease progression.

Genetic polymorphisms of genes involved in the adaptive immune response play a role in susceptibility to HIV-1 infection. The widely-described CCR5Δ32 mutation confers resistance to infection as it prevents HIV-1 from binding the CCR5 co-receptor [130]. Different HLA alleles have been associated with protection against HIV-1 progression in different HEPS groups [130]. HLA-B*57 and B*27 are associated with decreased disease progression, and HLA supertype A2/6802 has been associated with HIV-1 resistance in the Pumwani cohort [130,137]. It should be noted that these alleles do not provide protection against HIV-1 infection in all populations [130].

1.4.1.2. Humoral markers of resistance

Some (but not all) HEPS cohorts have antibodies directed against HIV-1 receptors or mucosal antibodies directed against HIV-1 [125]. In the systemic compartment, anti-CD4 antibodies bind CD4 antigens that are exposed only after the binding of gp120 were described in a cohort of HEPS individuals in 1996 [125]. There have also been anti-CCR5 antibodies found in HEPS cohorts described in the literature, and these antibodies could possibly lead to the down regulation of CCR5 expression and a consequent reduction of HIV-1 infection [138]. Protective IgA antibodies in the FGT have also been described in HEPS cohorts and are hypothesized to bind CCR5 or HIV-1. Anti-CCR5 IgA has been shown to block transcytocis of HIV-1 in one cohort [139], but in others the protective capacity of antibodies is less clear [140]. Interestingly, HIV-1-neutralizing IgA has been isolated from HEPS individuals without the presence of circulatory IgM or IgG, suggesting protective mechanisms present in the genital tract are not necessarily reflected in the systemic compartment [125]. Secreted IgG may also be important for mucosal HIV-1 resistance [115]. However, some studies have shown that the presence of

anti-HIV-1 IgA is more of a correlate of exposure than of resistance to HIV-1 infection [141].

1.4.1.3. Cell-mediated immune markers of resistance

Several HEPS cohorts have described the presence of anti-HIV-1 CD8+ T cells in the systemic and mucosal compartments as a correlates of protection against infection [142-146]. These cells recognize different epitopes than the CD8+ cells found in HIV-1 infected individuals [147] and respond with a lower magnitude as measured by IFNγ [148].

HIV-1 specific CD4+ cells have also been described in HEPS cohorts [146,149,150]. These cells recognize different HIV-1 epitopes than CD4+ cells isolated from HIV-1-infected patients and expressed lower levels of CXCR4 [125,150]. HEPS individuals were also characterized as having more systemic regulatory T cells (CD4+CD25+), suggesting a role for decreased immune activation in resistance to HIV-1 infection [151,152].

1.4.2. The Pumwani sex worker cohort

The Pumwani sex worker cohort in Nairobi, Kenya was first established in 1985 and studies of the immunobiology of HIV-1 infection in these women have been ongoing for over twenty years. Over 3000 women have been enrolled in this cohort to date. Fowke *et al.* observed in 1996 that there was a decreased risk of the women acquiring an HIV-1 infection the longer they were enrolled in the cohort (Figure 6) [95]. Since that time, many studies have come out of work with this cohort in an attempt to characterize

the mechanism(s) of HIV-1 infection and protection from infection. One of the most common correlates of HIV-1 resistance, the CCR5Δ32 mutation, does not occur in African populations and thus does not appear to play a role in HIV-1 resistance in this cohort [130].

There are many correlates of protection against HIV-1 infection that have been described for this cohort, both systemically and in the genital tract. In terms of cell-mediated immune responses, HEPS women have been shown to have broad HIV-1-specific CD8+ responses, both systemically and in the FGT [142,143,153]. Interestingly, the magnitude of the CD8+ responses were greater in the FGT than in the peripheral circulation, suggesting higher protection against HIV-1 infection in the genital tract of HEPS women, which is important as the primary route of transmission in this cohort is heterosexual intercourse [154]. There is an increased level of circulating regulatory T cells in HIV-1 resistant women, again suggesting a role for reduced immune activation in HIV-1 resistance [129].

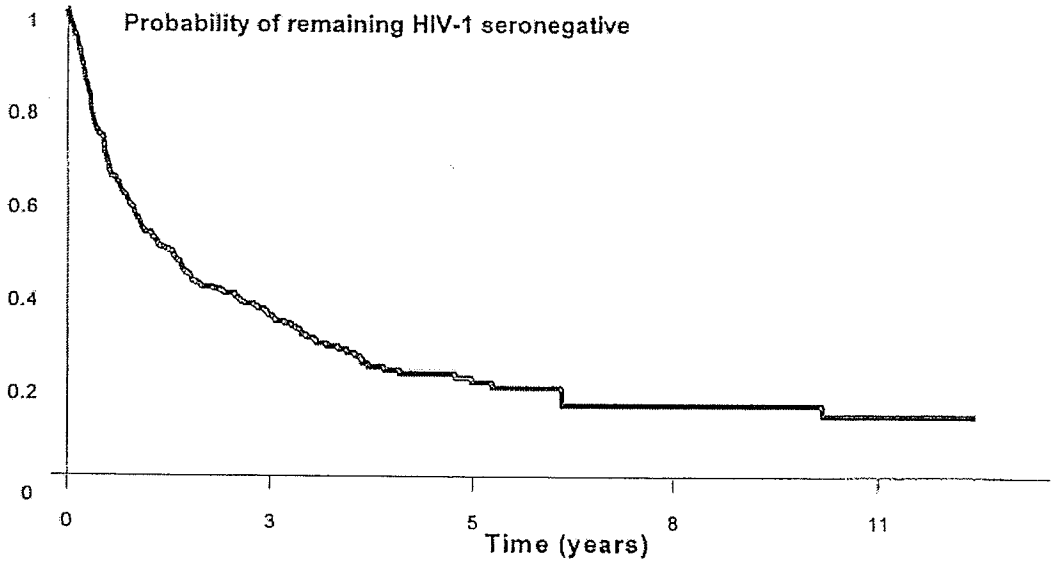


Figure 6: Kaplan-Meier survival curve [155]. This data shows a decrease in HIV-1 seroconversion as a function of increased time of follow up in the Pumwani cohort. Reprinted from Immunology Letters, Vol. 66, F.A. Plummer *et al*, "Resistance to HIV-1 infection among highly exposed sex workers in Nairobi: what mediates protection and why does it develop?" Page 28, Copyright 1999, with permission from Elsevier (obtained May 7th 2009). Permission from the author received June 25, 2009.

There have also been humoral correlates of protection described for this cohort, primarily in the FGT. The presence of mucosal anti-HIV-1 IgA has been described in these women [156,157], although it is unclear if antibodies are a marker of protection or an indication of exposure [141].

There are several mucosal correlates of protection against HIV-1 infection described in the Pumwani HEPS women. Mucosal HIV-1-specific CD8+ cells and increased T cell recruitment have been shown to be correlated with HIV-1 resistance [132,143]. Increased mucosal levels of the chemokine RANTES, as well as anti-proteases from the serpin B family, alpha-2 macroglobulin-like 1, and cystatin A have all been correlated with resistance in this cohort [132,158]. RANTES binds CCR5, and elevated levels could potentially inhibit the binding of HIV-1 to its co-receptor, although it may also recruit CD4 cells to the site of infection [132]. Serpins and alpha-2 macroglobulin-like 1 inhibit the inflammatory proteases cathepsins and elastases [159-161]. These proteins could potentially play an important role in decreasing inflammation and therefore decreasing the availability of activated T cells for HIV-1 infection. Cystatins interfere with HIV-1 processing, which is again of obvious benefit to prevent HIV-1 infection [158,162].

There have been several studies on determining genetic correlates of resistance in the Pumwani cohort. HLA alleles at the A*0202, A*6802, DPA1, DQB1, DRB, and other loci have been associated with HIV-1 resistance, which could play a role in HIV-1 antigen presentation and the development of an effective immune response [137,163-165]. Other genetic correlates of protection identified in this cohort are polymorphisms in interferon regulatory factor 1 (IRF-1) [166]. IRF-1 regulates the transcription of

interferons, which play a key role in antiviral responses and thus the role of polymorphisms in providing a protective immune response to HIV-1 is critical to understand [166]. To better understand cellular and immunological processes occurring in HIV-1 resistant women, and to discover new genetic correlates of resistance, whole genome expression analyses have been performed on CD4+ cells and whole blood of HEPS women. These expression arrays have both shown a distinct genetic profile in HIV-1 resistant women, further suggesting a genetic basis for HIV-1 resistance [129,167]. However, to date no whole genome expression analyses have utilized leukocytes from the genital tract of HIV-1 resistant women.

Although there have been many correlates of resistance described in this cohort, it can be difficult to understand how all these different factors are related. One model proposed by Fowke *et al* to explain the HIV-1 resistant phenotype of this cohort is that of immune quiescence [129,152]. Lower basal immune gene expression as shown by gene expression profiling of CD4+ cells, combined with increased levels of circulating regulatory T cells of HIV-1 resistant women suggest lower basal immune activation over HIV-1 susceptible women [129,152]. This decreased immune activation and a well-described HIV-1-specific CD8+ response suggest HEPS women may be able to successfully avoid HIV-1 infection despite exposure. This model would be strengthened by more in-depth gene expression analyses at the genital tract level of these same women to confirm this proposed decreased basal immune activation at the site of HIV-1 exposure.

1.5. HIV-1 research and genome-wide studies

Genome-wide expression studies allow for the discovery of completely novel factors to be identified in characterizing biological phenomena. These studies also allow for the study of entire cell signalling pathways that would not be elucidated by cell phenotyping or more targeted biochemical studies. To date, there have been no published genome-wide expression studies on cells from the female genital tract. There are, however, disadvantages to genome-wide association studies. First, it is difficult to get an idea of the differences in cell population composition between two different test groups when examining RNA isolated from mixed cell populations (e.g. lymphocytes). Also, small and potentially important differences in gene expression in one specific cell subset can be lost [168,169]. Despite the drawbacks, genome-wide studies are currently one of the best mechanisms to discover new biological processes between different phenotypes.

Several interesting genome-wide and proteome-wide studies have already emerged from the Pumwani cohort [129,158,167]. In addition, Brass *et al.* Zhou *et al.* and Konig *et al* used genome-wide siRNA analysis to identify novel host factors exploited by HIV-1 during its lifecycle [19-21]. This thesis aims to add to the current body of knowledge of HIV-1 resistance by characterizing the gene expression profiles cells from the FGT of indivduals who are resistant to HIV-1 infection.

2. Hypothesis, rationale, and specific objectives

The FGT is the primary site of HIV-1 transmission in the women from Pumwani cohort. Identification of host innate factors that inhibit HIV-1 transmission is an exciting new prospect in the microbicide field, as these compounds would be extensions of the pre-existing immune environment present in the FGT and could potentially have fewer side effects and provide better protection against HIV-1 transmission. Previous work has shown there is differential protein expression detected in the cervicovaginal lavages of HEPS women as compared to susceptible women [132,158]. There is a distinct gene expression profile in CD4+ cells and the whole blood of HIV-1 resistant women [129,167] Furthermore, CD8+ responses to HIV-1 are detected in the genital tract of HIV-1 resistant women and are of greater magnitude in the FGT when compared with systemic responses, suggesting a difference in the magnitude of HIV-1 responses between the genital tract and the periphery [143]. Although there have been many immune factors described in the genital tract of HIV-1 resistant women in the Pumwani cohort, there is little information regarding factors regulating these responses, especially gene expression patterns in these women. A better idea of cellular processes occurring in the genital tract would be of great benefit to understanding mechanisms of HIV-1 resistance in the FGT. Taking this previous work into consideration, the rationale for the following hypothesis was formulated:

There is differential gene expression in the genital tract of HIV-1 resistant women when compared to susceptible women that mediates biochemical pathways that reduce their susceptibility to HIV-1 infection.

To test this hypothesis, this thesis attempts to fulfill three specific objectives:

Objective 1: Examine gene expression patterns in cervical mononuclear cells (CMCs) and peripheral blood mononuclear cells (PBMCs) to identify important differences between the systemic and mucosal immune compartments.

Objective 2: Compare gene expression patterns of CMCs isolated from HIV-1 resistant women and HIV-susceptible women to identify markers of HIV-1 resistance in the FGT.

Objective 3: Examine changes in gene expression patterns of CMCs from HIV-1 resistant and susceptible women after HIV-1 antigenic challenge.

3. Materials and Methods

3.1. General laboratory materials:

3.1.1. Solutions

PBS + 2% FCS

48.5 g PBS (137.93 mM NaCl, 2.67 mM KCl, 8.1 mM Na_2HPO_4, 1.47 mM KH_2PO_4)

20 ml Fetal Calf Serum (Invitrogen: Burlington, Ontario, Canada)

to 1L with sterile-filtered double distilled H_2O (ddH_2O)

75% Ethanol

750 ul 100% ethanol

250 ul RNase-free water (from RNeasy MinElute Kit)

TE buffer pH 8.0

100 ul 1 M Tris-HCl pH 8.0, Cat. No. VW1500-01 (VWR International: Mississauga, Ontario, Canada)

20 ul 0.5 M EDTA pH 8.0, Cat. No. 15576-028 (Invitrogen: Burlington, Ontario, Canada)

99.8 ml ddH_2O

3% sodium hypochlorite solution

500 ml Lavo-6 Bleach (6% v/v sodium hypochlorite) (Lavo: Montreal, Quebec, Canada)

500 ml water

0.1 M glycine pH 2.5

7.5 g glycine, Cat. No. 161-0718 (Bio-Rad: Mississauga, Ontario, Canada)

1 L water

Adjust pH with to 2.5 with HCl

10X TBS

87.7 g NaCl, 12.1 g Tris, 4 ml HCl

ddH_2O to 1 L

1X TBS + 0.1% Tween-20

100 ml 10X TBS

1 ml Tween-20, Cat. No. H5151 (Promega: Madison, Wisconsin, U.S.A.)

ddH_2O to 1 L

SNAP i.d. Blocking Solution

100 ml 1X TBS + 0.1% Tween-20

0.5g skim milk powder

Cell lysis buffer for protein isolation

6.7g NaCl (115 mM final concentration)

6.05 g Tris (50 mM final concentration)

50 ul NP-40, Cat. No. 11 754 599 001 (Roche Applied Science: Laval, Quebec, Canada)

ddH_2O to 1 L

1 protease inhibitor cocktail tablet per 10 ml lysis buffer, Cat. No. S8820 (Sigma-Aldrich: Oakville, Ontario, Canada)

3.1.2. Reagents

Cell culture media

AIM-V media with BSA, Cat. No. 31035-025. (Invitrogen: Burlington, Ontario, Canada)

RPMI media with L-Glutamine + 10% Fetal Calf Serum + 2% penicillin, streptomycin and fungazone (HyClone: Logan, Utah, USA)

Cell separation

Lymphoprep, Product No. 11 14547 (Axis-Shield: Oslo, Norway)

Trypan Blue, Cat. No T8154 (Sigma-Aldrich: Oakville, Ontario, Canada)

40 um nylon cell strainer, Cat. No. 352340 (BD Falcon: Mississauga, Ontario, Canada)

RNA Extraction

TRIzol Reagent, Cat. No. 15596-018 (Invitrogen: Burlington, Ontario, Canada)

Chloroform

Isopropanol

RNase-free glycogen, Cat. No 10814-010 (Invitrogen: Burlington, Ontario, Canada)

Stimulations

DMSO (Dimethylsulfoxide), Cat. No. D8418 (Sigma-Aldrich: Oakville, Ontario, Canada)

CEF (Cytomegalovirus, Epstein Barr and Flu) positive control peptide pool, Cat. No. 408-452-5055 (Anaspec: San Jose, California, USA)

HIV-1 gag custom peptide library, Order No. 71615 (Sigma Genosys: St. Louis, Missouri, USA)

PHA (Phytohemagglutinin), Cat. No. L1668 (Sigma: St. Louis, Missouri, USA)

Antibodies

OSM (N-1) Cat. No. sc-129 (SantaCruz Biotechnology: Santa Cruz, California, USA)

p52 (447) Cat. No. Sc-848 (SantaCruz Biotechnology: Santa Cruz, California, USA)

$FLIP_{S/L}$(H-202) Cat. No. Sc-8347 (SantaCruz Biotechnology: Santa Cruz, California)

Actin (I-19) Cat. No. Sc-1616 (SantaCruz Biotechnology: Santa Cruz, California, USA)

Goat anti-rabbit IgG-HRP Cat. No. 111-035-003 (Jackson ImmunoResearch, Westgrove, Pennsylvania, USA)

Rabbit anti-goat IgG-HRP Cat. No. 305-035-003 (Jackson ImmunoResearch, Westgrove, Pennsylvania, USA)

3.1.3. Commercial kits

RNA extraction and purification

RNeasy Plus Mini Kit, Cat. No. 74134 . (Qiagen: Mississauga, Ontario, Canada)

RNase-free DNase Set, Cat. No. 79254. (Qiagen: Mississauga, Ontario, Canada)

RNeasy MinElute Kit, Cat. No. 74204. (Qiagen: Mississauga, Ontario, Canada)

RNA quantification

RNA Nano 6000 Series II Kit: Chips, Cat. No. 5067-1511; Reagents, 5067-1512; Ladder, 5067-1529. (Agilent: Santa Clara, California, USA)

Illumina Gene Expression BeadArrays®

Ambion Illumina TotalPrep RNA Amplification Kit, Cat. No. IL1791 (Ambion: Foster City, California, USA)

Illumina HumanRef-8 v2 Expression BeadChip Kit, Cat No. BD-25-113 – includes 10x High-Temp Wash Buffer, HYB, HCB, Wash E1BC Buffer and Block E1 Buffer (recipes proprietary) (Illumina: San Diego, California, USA)

Streptavidin-Cy3, Cat. No. PA43001 (Amersham Biosciences: Piscataway, New Jersey, USA)

100% Ethanol

RT-PCR

QuantiTect Reverse Transcription Kit, Cat. No. 205311 (Qiagen: Mississauga, Ontario)

QuantiTect SYBR Green PCR Kit, Cat. No. 204143. (Qiagen: Mississauga, Ontario)

IFNG QuantiTect Primer, Cat. No. QT00000525 (Qiagen: Mississauga, Ontario)

MyD88 QuantiTect Primer, Cat. No. QT00203490 (Qiagen: Mississauga, Ontario)

CFLAR QuantiTect Primer, Cat. No. QT00064554 (Qiagen: Mississauga, Ontario)

OSM QuantiTect Primer, Cat. No. QT00209286 (Qiagen: Mississauga, Ontario)

RRN18S QuantiTect Primer, Cat. No. QT00199367 (Qiagen: Mississauga, Ontario)

FirstChoice® Human Brain Reference RNA (1 mg/ml) Cat. No. AM6050 (Ambion: Foster City, California, USA)

TransPlex® Complete Whole Transcriptome Amplification Kit (10 reactions). Cat. No. WTA2 (Sigma-Aldrich: Oakville, Ontario, Canada)

Protein quantification and preparation

BCA protein assay kit, Cat. No. 71285-3 (Novagen: Gibbstown, New Jersey, U.S.A.)

Gel electrophoresis

4x NuPAGE® LDS Sample Buffer, Cat. No. NP0008 (Invitrogen: Burlington, Ontario)

10x NuPAGE® Reducing Agent (500 mM dithiothreitol), Cat. No. NP0004 (Invitrogen: Burlington, Ontario, Canada)

NuPAGE® Novex 12% Bis-Tris Gel 1.0 mm, 12 well, Cat. No. NP0342BOX (Invitrogen: Burlington, Ontario, Canada)

NuPAGE® MES SDS Running Buffer (for Bis-Tris Gels only) (20X) Cat. No. NP0002 (Invitrogen: Burlington, Ontario, Canada)

NuPAGE® Antioxidant, Cat. No. NP0005 (Invitrogen: Burlington, Ontario, Canada)

Novex® Sharp Pre-stained Protein Standard, Cat. No. LC5800 (Invitrogen: Burlington, Ontario, Canada)

Western blotting

iBlot Transfer Stack (Nitrocellulose) Cat. No. IB3010-01 (Invitrogen: Burlington, Ontario, Canada)

SNAP i.d. Single Well Blot Holder Cat. No. WBAVDBH01 (Millipore: Billerica, Massachusetts, U.S.A.)

Novex® ECL Chemiluminescent Substrate Reagent Kit, Cat. No. WP20005 (Invitrogen: Burlington, Ontario, Canada)

3.1.4. Specialized equipment

Illumina Gene Expression BeadArrays®

BeadChip Hyb Chamber, Part No. 210948 (Illumina: San Diego, California, USA)

BeadChip Hyb Chamber inserts, Part No. 222682 (Illumina: San Diego, California, USA)

BeadChip coverseals, Part No. 195750 (Illumina: San Diego, California, USA)

Illumina Hybridization Oven, Part No. 198361 (Illumina: San Diego, California, USA)

Rocker Attachment for Hybridization Oven, Part No. 11188299 (Illumina: San Diego, California, USA)

BeadChip wash rack, Part No. 200037 (Illumina: San Diego, California, USA)

Glass wash dishes, Part No. 198205 (Illumina: San Diego, California, USA)

BeadChip deep well wash tray, Part No. 11174188 (Illumina: San Diego, California)

BeadChip deep well wash tray cover, Part No. 195717 (Illumina: San Diego, California)

Hybex Microarray Incubation System Heating Base, Cat No. 1057-30-0 (SciGene: Sunnyvale, California, USA)

BeadStation 500X, Part No. SC-16-103 (Illumina: San Diego, California, USA)

RNA Quantification

NanoDrop 1000 Spectrophotometer (ThermoScientific: Wilmington, Delaware U.S.A.)

Agilent BioAnalyzer 2100, Part No. G2938C (Agilent: Santa Clara ,California, USA)

Quantitative Real-Time RT-PCR

LightCycler® System 1.5, Cat. No. 04484495001 (Roche Applied Science: Laval, Quebec, Canada)

LightCycler® Capillaries, 20ul, Cat. No. 04929292001 (Roche Applied Science: Laval, Quebec, Canada)

Western Blotting

iBlot® Gel Transfer Device, Cat. No. IB1001 (Invitrogen: Burlington, Ontario, Canada)

SNAP i.d. Protein Detection System, Cat. No. WBAVDBASE (Millipore: Billerica, Massachusetts, U.S.A.)

Molecular Imager ChemiDoc XRS+ System , Cat. No. 170-8252 (Bio-Rad: Mississauga, Ontario, Canada)

3.2. Protocols

3.2.1. Specimen sampling – Winnipeg

Women donated cervical scrapings during routine pelvic exams at the Health Sciences Centre Department of Obstetrics & Gynaecology Colposcopy Clinic in Winnipeg, Manitoba. Each participant had the procedure explained to her by the gynaecologist or a nurse and consent was obtained before the exam was performed. Only women who were healthy and not scheduled for biopsies or actively menstruating were asked to participate in the study. The gynaecologist inserted a cytobrush into the endocervix and rotated 360°. The brush was then placed in 10 ml AIM-V media in a 50 ml Falcon tube. The gynaecologist then used a wooden scraper to scrape the available immune cells from the cervical opening (os). The scraper was also placed in the 10 ml AIM-V media and kept at 4°C or on ice until ready for processing.

3.2.2. Specimen sampling – Nairobi

Commercial sex workers from the ML (Malaya) clinic in the Pumwani slum of Nairobi, Kenya, donated cervical scrapings during a voluntary pelvic exam. Each participant had the procedure explained to her by the physician or nurse and consent was obtained when she signed up as a cohort participant. Women menstruating at the scheduled time of exam were asked to return at a later date for the examination and sample collection. The remaining procedure follows as in 3.2.1 but the scraper used in the exam was plastic and the scrapings were stored in a cooler in 10 ml PBS (without FCS) instead of AIM-V media.

3.2.3. Cytobrush processing

Samples were transported either on ice or in a cooler from the clinic to the lab and processed within 2 to 3 hours after sampling. The tubes were gently mixed for 25 seconds and the mucus was scraped off the cytobrush with the scraper. The brush and scraper were then discarded in a 3% sodium hypochlorite solution. The samples were mechanically sheared by repeated pipetting and vortexed again for a few seconds until the cervical mucus was disrupted. When no more pieces of mucus were visible, the samples were layered onto 3 ml Lymphoprep in a 15 ml Falcon tube and centrifuged for 25 minutes at 357 relative centrifugal force (r.c.f.). The lymphocyte layer was removed to a new 15 ml tube and the volume was adjusted to 12 ml with AIM-V media. Cells were pelleted by centrifugation at 514 r.c.f. for 10 minutes, the supernatant was decanted and pellets resuspended in 1.5 ml AIM-V media. Cell viability was assessed by staining with trypan blue and live cells were counted using a hemocytometer.

Cells were then transferred to a 1.5 ml tube and pelleted by centrifuging at 585 r.c.f. for 2 minutes. The supernatant was removed, 800 uL of TRIzol® reagent was added and the cell pellets were homogenized by repeated pipetting until there were no visible solids in the solution. Samples were then frozen at -70°C until RNA was extracted.

Variations to the above protocol included incubating the brushes in 10 ml AIM-V media or with 2 ml TrypLE (a commercial trypsin solution) for one hour at 37°C, separating the CMCs from epithelial cells using nylon wool in a 10mL syringe or a 40 um cell strainer in lieu of Lymphoprep.

3.2.4. Blood processing

Approximately 8 ml of peripheral blood was collected into heparinised tubes from matched CMC donors at the time of cervical exam. PBMCs were isolated following described Ficol-Hypaque protocols [153]. Briefly, tubes were centrifuged for 7 minutes at 514 r.c.f. and the plasma was removed and discarded. Blood was diluted with PBS + 2% FCS to approximately twice the starting volume, layered onto 3 ml Lymphoprep in a 15 ml Falcon tube and centrifuged at 448 r.c.f. for 25 minutes. The mononuclear cell layer was removed to a new 15 ml Falcon tube and diluted to 12 ml with PBS + 2% FCS. PBMCs were then centrifuged for another 10 minutes at 585 r.c.f. and resuspended in 10 ml AIM-V media. Cell viability was assessed by staining with trypan blue and live cells were counted using a hemocytometer.

Cells were transferred to a 1.5 ml tube and pelleted by centrifuging at 1600 r.c.f. for 2 minutes. The supernatant was removed, 800 uL of TRIzol® reagent was added and the cell pellets were homogenized by repeated pipetting until there were no visible solids in the solution. Samples were then frozen at -70°C until RNA was extracted.

3.2.5. Peptide pools

Cells underwent three treatment conditions: HIV-1 peptide pool or CEF peptide pool, PHA as a positive control and DMSO as a negative control. Concentrations of each stimulation condition are described in sections 3.2.6 and 3.6.7.

CEF is a pool of 32 peptides from Cytomegalovirus, Epstein Barr and Influenza viruses and was used to set up stimulation experiments with the rationale that everyone at one point in their lives should have been exposed to these viruses and would have an

antigen-specific immune memory recall response. The peptides were reconstituted to a final concentration of 4 mg/ml of total peptide in DMSO.

The HIV-1 gag peptide pool was made from a peptide library of Clade A and ancient consensus gag peptide sequences. Stock peptides were resuspended in DMSO to a concentration of 20 mg/ml. The peptides range from 13 to 15 amino acids in length and each successive peptide overlaps with its amino(N)-terminal predecessor by 10 amino acids (Appendix B). The HIV-1 gag protein consisted of a pool of 56 peptides starting from the N-terminus and continues until halfway through the gag protein at NKIVRMYSPVSILDI. The pool covered only the N-terminal half of the HIV-1 gag sequence as the protein is too large to be completely represented in a single pool without having cytotoxic levels of DMSO. This peptide pool was optimised on HIV-1 positive bloods in Winnipeg and on CMCs isolated in Nairobi. The final concentration of the pool was 330 ug/ml of each peptide.

PHA (phytohemagglutinin) is a mitogen that agglutinates leukocytes together, producing robust T and B cell responses [170], and acts as a positive control for all stimulation experiments in this thesis.

3.2.6. Stimulation conditions – Winnipeg

CMCs or PBMCs were re-suspended to a concentration of 600,000 cells/ml in AIM-V media, and a 0.5 ml aliquot was made for each stimulation condition. HIV-1 uninfected CMCs or PBMCs were treated with either 1.51 ul DMSO (final concentration 0.3% v/v), 5 ul 1:10 dilution 4 ug/ul stock CEF peptides (final concentration 2 ug/ml total peptide) or 2.5 ul 5 mg/ml PHA (final concentration 5 ug/ml) as a positive control. HIV-

1-positive donor bloods were stimulated with either 1.51 ul DMSO, 1.51 ul 330 ng/ul HIV-1 gag peptide pool (final concentration 1ug/ml of each peptide, DMSO 0.3% v/v), or 2.5 ul 5 mg/ml PHA (final concentration 5 ug/ml) as a positive control. Cells were either stimulated for 3, 6 or 16 hours at 37°C, and the stimulation was ended by centrifuging the cells for 5 min at 2,000 r.c.f., removing and saving the supernatants, and resuspending the cell pellets in 800 uL TRIzol. Samples were frozen at -80°C until further processing could be conducted.

3.2.7. Stimulation Conditions – Nairobi

CMCs from each patient were equally divided into 4 x 500 ul aliquots: untreated (baseline), DMSO-, HIV-1 gag- and PHA-treated. Scrapings yielding less than 1 million CMCs were either saved for exclusively baseline studies or separated into 2 of the stimulation conditions. CMCs were stimulated overnight at 37°C using stimulation concentrations previously described in section 3.2.6. The following morning, cells were transferred into 1.5 ml tubes and centrifuged for 5 min at 5,000 r.c.f. Supernatants were removed and saved for later analysis, and cell pellets were homogenized in 800 ul TRIzol®. Samples were stored at -80°C until shipment back to Winnipeg in liquid nitrogen. In Winnipeg, samples were stored at -135°C until processed as described in 3.2.8, 3.2.9 and 3.2.10.

3.2.8. RNA extraction

Samples frozen in TRIzol® were thawed and incubated for 5 min at room temperature (r.t) to dissociate nucleoprotein complexes. 0.16 ml of chloroform was added, and samples were capped securely and shaken vigorously for 15 sec. Next, they were incubated at r.t. for 3 minutes and then centrifuged at 12,000 r.c.f. for 10 min at 2°C. The upper aqueous phase was removed to a new 1.5 ml tube containing 450 ul isopropanol and 10 ug RNase-free glycogen. RNA was precipitated for 10 min at r.t. and then centrifuged for 10 min at 12,000 r.c.f. at 2°C. The RNA pellet was washed with 1 ml RNase-free 75% EtOH, mixed for a few seconds and then centrifuged at 7,500 r.c.f. for 5 min at 2°C. Supernatants were decanted and the RNA pellet was air dried for no more than 10 min. The pellet was resuspended in 17 ul RNase-free water and incubated at 55°C for 10 min to completely resuspend the pellet. Samples were then DNase digested and purified for downstream analysis.

3.2.9. DNase digestion and RNA cleanup

Carryover genomic DNA (gDNA) was removed with the Qiagen RNase-free DNase set as per the manufacturer's instructions. Briefly, the RNA was diluted to 85 ul final volume in RNase-free water. Ten ul buffer RDD and 5 ul DNase I were added and samples were incubated at r.t. for 10 minutes. Samples were then immediately concentrated and purified according to the Qiagen RNeasy MinElute Kit protocol [171]. Samples were eluted from the column in 14 ul RNase-free water. A 2 ul aliquot was reserved to determine RNA concentration and purity, and the remaining RNA was stored at -70°C.

3.2.10. Alternate RNA isolation protocol

Issues arising from low yields of RNA from samples from Nairobi led to the development of an alternative RNA isolation protocol for several samples. This procedure follows the method described in 3.2.8 until the removal of the upper aqueous phase after the 15 min spin at 12,000 r.c.f. This protocol has the advantage of skipping the RNA precipitation step and therefore should increase RNA yields. An equal volume (500 ul) of 70% ethanol was added to the isolated upper aqueous phase and thoroughly mixed. This was then loaded onto the RNeasy MinElute column, centrifuged for 15 sec at 10,000 r.c.f. at r.t. The remainder of the protocol is followed according to pages 26 and 27 of the Qiagen RNeasy MinElute Kit protocol book. Samples were eluted from the column in 14 ul RNase-free water. A 2 ul aliquot was reserved to determine RNA concentration and purity, and the remaining RNA was stored at -70°C. The samples extracted using this protocol include CMCs from the Spring 2007 Resurvey and samples from the Spring 2008 Resurvey shown in Table 1. All other samples were treated as described in sections 3.2.8 and 3.2.9.

3.2.11. RNA Quantification

RNA for qRT-PCR was quantified using the Nanodrop ND-1000 Spectrophotometer following the manufacturer's instructions [172]. RNA for the Illumina® BeadChip were quantified and their degradation and purity were assessed using the Agilent BioAnalyzer 2100 system according to the RNA Nano 6000 Series II Kit protocol [173].

Table 1: Samples from May 2008 resurvey treated as described in 3.2.10

Unstimulated	DMSO-treated	Gag-treated	PHA-treated
ML1025	ML2345	ML2345	ML2345
ML1938	ML2367	ML2367	ML2367
ML1952	ML2811	ML2661	ML2661
ML2006		ML2811	ML2811
ML2141			
ML2138			
ML2403			
ML2404			
ML2669			
ML2701			
ML2764			
ML2787			
ML2811			
ML2914			

3.2.12. cDNA preparation

First strand complimentary DNA (cDNA) synthesis for real-time, quantitative reverse transcription PCR (qRT-PCR) was performed using the QuantiTect Reverse Transcription Kit according to the protocol included with the kit. QuantiTect Primer Assays were resuspended in 1.1 ml TE buffer pH 8.0 and frozen in 100 ul aliquots. Amplification and product detection was performed using the Roche LightCycler® System 1.5 and QuantiTect SYBR Green PCR Kit according to the kit-enclosed protocol [174]. Thermocycler conditions are outlined in Table 2.

3.2.13. Whole transcriptome amplification

The Sigma TransPlex® Complete Whole Transcriptome Amplification Kit was assessed to improve gene signals and to increase sample sizes for qRT-PCR microarray follow up studies. Five, 25 and 125 ng of standard brain RNA from Ambion® underwent amplification according to the enclosed kit instructions [175]. Amplified cDNA was quantified on Nanodrop and the expression levels of 18S and CFLAR were evaluated using the Roche LightCycler System 1.5 and QuantiTect SYBR Green PCR Kit according to the kit-enclosed protocol [174].

Table 2: Thermocycler conditions for RT-PCR amplification.

Step	Time	Temperature	Ramp	Cycles
PCR initial activation step	15 min	95°C	20°C/s	1
Denaturation	15 s	94°C	20°C/s	40
Annealing	20 s	55°C	20°C/s	
Extension	20 s	72°C	20°C/s	
Data acquisition	5 s	--	20°C/s	
Melting curve analysis	--	--	--	1

3.2.14. qRT-PCR

To quantify the PCR product, a standard curve was generated for the CMC and PBMC stimulation study as follows: 6 million PBMCs were incubated for 24 hours with 5 mg/ml PHA. RNA was extracted using the Qiagen RNeasy Plus Mini Kit according to the manufacturer's instructions [176] and diluted 1:10, 1:100 and 1:1000 for the IFNγ assay. Three ul of each dilution was reverse-transcribed into cDNA and then included in the LightCycler run. For the CMC vs. PBMC gene expression confirmation, 1 ug brain standard RNA from Ambion® was reverse transcribed into cDNA, and this standard was 5-fold serially diluted to a final dilution of 1:125. SYBR Green concentrations and gene product crossing points were calculated using the second derivative maximum method, or where the acceleration of the amplification of the PCR product is at a maximum [177] using LightCycler® software v. 3.5.3.

3.2.15. Calculation of fold-change expression values for qRT-PCR

Gene expression levels were determined using the mathematical model previously described [178]. This model incorporates the PCR efficiency (as determined by standard curve) as well as normalizes the results to the expression of a reference gene to calculate gene expression. The equation is as follows:

$R = (E_{target})^{\Delta Cp_{target}(control-sample)}/(E_{reference})^{\Delta Cp_{reference}(control-sample)}$, where R is the ratio of expression of the gene of interest between media and stimulated cells (or between PBMCs and CMCs), which is normalized by the expression of the housekeeping gene 18S; E_{target} is the measured amplification efficiency of the gene of interest; $E_{reference}$ is the

measured amplification efficiency of the housekeeping gene, ΔCP_{target}(control-sample) is a measure of the difference in crossing points of the gene of interest between the control condition (i.e. unstimulated cells or PBMCs) and sample condition (i.e. stimulated cells or CMCs); $\Delta CP_{reference}$(control-sample) is a measure of the difference in crossing points of the reference gene between the control and sample conditions.

3.2.16. RNA preparation for Illumina BeadChip® arrays

150 ng purified CMC and PBMC RNA were taken for gene expression analysis on the Illumina HumanRef-8 v2 Expression BeadChip arrays. RNA was reverse-transcribed and amplified before running on arrays according to the Illumina TotalPrep RNA Amplification Kit protocol. First, the RNA is reverse-transcribed into double-stranded cDNA and then amplified by *in-vitro* transcription (IVT) into cRNA.

3.2.16.1. First strand cDNA synthesis

150 ng purified RNA was brought to a final volume of 11 ul. To each sample, 1 ul T7 Oligo(dT) Primer (100 ng), 2 ul 10X First strand buffer, 4 ul dNTP mix, 1 ul RNase inhibitor and 1 ul of ArrayScript reverse transcriptase were added. Samples were then incubated at 42°C for 2 hours. The resulting cDNA product had a T7 promoter that allowed this cDNA to act as a template for IVT [179].

3.2.16.2. Second strand cDNA synthesis

To each First Strand cDNA synthesis sample, 63 ul nuclease-free water, 10 ul 19X Second Strand Buffer, 4 ul dNTP mix, 2 ul DNA Polymerase and 1 ul RNase H were added. Samples were then incubated for 2 hours at 16°C to allow the

complimentary cDNA strand to be synthesized, as well as to degrade the starting RNA material. This double-stranded, T7-promoter-containing cDNA was then used as template for IVT and amplification of cRNA [179].

3.2.16.3. cDNA purification

Before the cDNA was used for IVT, it was first purified on a filter cartridge to remove RNA, salts, enzymes and primers from the previous 2 steps. To each cDNA sample, 250 ul cDNA Binding Buffer was added and loaded onto the cDNA Filter Cartridge. The samples were centrifuged at 10,000 r.c.f. for one minute. Next, 500 ul Wash Buffer was added and samples were centrifuged at 10,000 r.c.f. for one minute. Finally, the cDNA was eluted with 19 ul 55°C nuclease-free water. This purified cDNA product was then used as template for IVT [179].

3.2.16.4. *In vitro* transcription and biotin labelling of cDNA

This step amplifies starting mRNA, allowing it to be detected on the gene expression arrays. Ambion® uses its patented MEGAScript® IVT technology to maximize the amplification of the starting material. To each purified cDNA sample, 2.5 ul 10X T7 Reaction Buffer, 2.5 ul T7 Enzyme Mix, and 2.5 ul biotin-NTP Mix were added. Samples were incubated at 37°C for 14 hours, at which time 75 ul nuclease-free water was added to each sample to stop the reaction. This amplified cRNA is next purified before hybridization onto the Illumina HumanRef-8 v2 Expression BeadChip arrays [179].

3.2.16.5. cRNA purification

This purification step is important to remove enzymes, salts and unincorporated nucleotides, which can inhibit the hybridization of the biotinylated complimentary RNA (cRNA) onto the arrays. 350 ul cRNA Binding Buffer was added to each cRNA sample and loaded onto a cRNA Filter Cartridge. Samples were centrifuged for one minute at 10,000 r.c.f., washed with 650 ul Wash buffer and eluted with 100 ul 55°C nuclease-free water. This purified biotinylated cRNA was then ready to be hybridized onto the Illumina HumanRef-8 v2 Expression BeadChip arrays [179].

3.2.17. Hybridization

750 ng of purified cRNA from section 3.2.16.5 was taken in 5 ul water for hybridization onto Illumina HumanRef-8 v2 Expression BeadChip arrays. The BeadChip is a novel method for gene expression microarray technology. Each bead has thousands of unique, gene-specific 50-mers covalently bound to the surface of the bead. 24,500 gene transcripts were represented an average of 30 times on each BeadChip, and 8 samples can be analysed on one chip. Illumina uses a patented sequence tag ("address tag") for each bead to identify the location and gene identity of each chip [180]. The samples were hybridized onto the BeadChips as follows:

Each sample had 10 ul of 58°C HYB (Hybridization Buffer) added, followed by incubation at 65°C for 5 min and then cooled to r.t. The Hyb chamber was assembled according to manufacturer's instructions and 200 ul HCB (Humidity Control Buffer) was added to the humidifying buffer reservoirs. Beadchips with coverseals were loaded into Hyb chamber inserts, and 15 ul of the HYB-sample mix were added to the BeadChips by

the inlet port. Hyb chamber inserts were loaded into the Hyb chamber and the chamber lid was locked on. The chamber was loaded in the Hybridization oven and set to incubate at 58°C for 16 hours with rocking [181].

3.2.18. First wash and blocking

After the 16-hour hybridization, the Hyb chamber was disassembled and BeadChips® were submerged in a beaker containing 1L Wash E1BC buffer. While submerged, the BeadChip coverseals were removed. BeadChips were moved to a slide rack in a staining dish with 250 ml Wash E1BC buffer. Chips were transferred via slide rack into 500 ml 55°C 1X High-Temp Wash Buffer in the Hybex Waterbath and incubated for 10 min. The slide rack was then returned to the staining dish containing 250 ml Wash E1BC buffer and plunged into the solution 10 times. The dish was then shaken for 5 min at a medium low speed at r.t. BeadChips were then washed in 250 ml 100% Ethanol by plunging the rack into the solution 10 times and rotation for 10 minutes at medium-low speed at r.t. Sample rack was transferred to 250 ml fresh of Wash E1BC buffer, plunged into the solution 10 times, and rotated for an additional 2 minutes at r.t. BeadChips were transferred to a BeadChip® Wash Tray containing 4 ml Block E1 buffer and rocked for 10 min at r.t. [181]

3.2.19. Stain, second wash, and drying

BeadChips were transferred to a new BeadChip Wash Tray with 2 ml Block E1 buffer with streptavadin-Cy3 (1:1000 dilution) and rocked for 10 minutes with wash tray covers. BeadChips were then transferred back to the slide rack in a staining dish

containing 250 ml Wash E1BC. BeadChips were washed by plunging the slide rack into the buffer 5 times and shaken for 5 minutes at r.t. BeadChips were then transferred to a centrifuge with microtiter plate holders, spun at 275 r.c.f. for 4 min at r.t and stored in the dark until scanning [181].

3.2.20. Scanning and image processing

Dried BeadChips were loaded into the adaptor tray to have spot intensities read by the Illumina BeadStation 500X. The BeadStation 500X is a scanning laser confocal microscope system with red and green lasers for BeadChip scanning. The adaptor tray was loaded into the BeadStation, and the default settings for DirectHyb Gene Expression were used in the scanning process. Spot signal intensities were measured by the green laser. The scanned images were saved and converted into text files. Text files were exported from BeadScan for statistical analysis elsewhere [182].

3.2.21. Gene array expression analysis

Raw chip intensity text files exported from BeadScan were imported into Bioconductor (www.bioconductor.org) an online program for bioinformatics analysis, for further processing [183]. The R software package was used for sample pre-processing. Background normalization was performed by surrogate normal minimum replacement. Briefly, background hybridization signal was calculated by the average signal intensity from designated background and "address only" beads. The data was then filtered to remove genes falling below background levels in all samples. Values were converted log(2) and this background integer was added to any remaining sample probes falling

below average background intensity. Consequently, any probes below the average background intensity were given a positive integer.

Averaged raw CMC intensities were divided by PBMC raw intensities to calculate fold change values in Linear Models for Microarray Data (LIMMA). P-values were also calculated in LIMMA using a paired T-test, and the false discovery rate was adjusted using the Benjamini-Hochberg method [184]. An absolute fold change > 1.3 with an adjusted p-values < 0.05 was considered significant.

Alternatively, averaged raw HEPS CMC intensities were divided by NN CMC intensities to calculate fold changes between the two test groups. P-values were calculated as described above and p-values < 0.05 were considered significant.

3.2.22. Hierarchical clustering

Hierarchical clustering of the gene expression data was performed using Cluster® program [185]. This algorithm sorts arrays and genes based on similarity of expression patterns in a step-wise manner. A detailed explanation of the methods is available in the Cluster and TreeView manual [185]. The arrays with pre-processed signal values were inputted into the Cluster program, median centred, and then hierarchically clustered by genes and by arrays. For hierarchical clustering, the "distance" between genes (in terms of expression level) is first calculated by determining the uncentred correlation coefficient I, for each gene, using the formula $r = N^{-1}\Sigma^{N}_{i=1}[X_i/(N^{-1}\Sigma^{N}_{i=1}(X_i)^2)^{1/2}][Y_i/(N^{-1}\Sigma^{N}_{i=1}(Y_i)^2)^{1/2}]$ where N = all genes on all arrays, and X_i and Y_i are the spots whose distances are being calculated. Distance was calculated between all possible pairs of gene spots from the data. The distance was found from the correlation coefficient using

the formula $d = 1 - r$. Once distance had been calculated, the data were clustered in a pair wise manner. Data points with the smallest distance between them are clustered more closely, and thus can be described as being more similar. Genes are then grouped based on similarity in a hierarchical manner. The two most similar genes are paired into a pseudo-item. Successive pseudo-items are joined based on the average distance of the genes between the pseudo-items, until all data were part of the same group (average-linked agglomerative clustering). Clustered data were then visualized using the TreeView® program [185]. Red was assigned to spots whose intensities fell below the median of the gene, and green was assigned to spots whose intensities were above the median expression level of the gene.

3.2.23. DAVID Functional Annotation Database

Entrez gene identification numbers of significant genes found from calculations described in 3.2.21 were inputted into the online Database for Annotation, Visualization and Integrated Discovery (DAVID) (http://david.abcc.ncifcrf.gov/). This program tests for biological theme and pathway enrichment in the inputted data. First, the gene IDs are functionally annotated for biological themes using over 20 gene and protein identifiers and 40 functional annotation categories from several public annotation databases [186]. Next, genes from pathways over-represented in the inputted data list are found using the Expression Analysis Systematic Explorer (EASE) test. This calculation is a modified one-tailed Fisher's Exact test [187]. In Fisher's Exact test, the number of genes from a given pathway (a) is compared to the number of genes from this pathway annotated in public databases (b), and the number of genes in the list not in the pathway (c) are

compared to the number of genes in the human genome that are not in the pathway (d) [187]. EASE modifies this calculation by comparing a – 1 genes found from the inputted gene list to b [187]. This decreases the probability of a pathway with very few annotated genes being significantly enriched when one gene appears in the inputted gene list by random chance. Over represented pathways from the KEGG (Kyoto Encyclopaedia of Genes and Genomes) database were selected for follow up studies [188-190].

3.2.24. Gene Set Enrichment Analysis

Gene Set Enrichment Analysis (GSEA) was used as an additional analysis method to identify biological processes differentially expressed in CMCs from HIV-1 resistant women as compared to susceptible women [191]. A list of genes (gene list L) sorted by fold changes between two test groups is inputted into the program. GSEA then tests if genes involved in a specific process or pathway (gene list S) are enriched at the top or bottom of the gene list L (where significantly differentially expressed genes generally fall). The program runs sequentially through the genes in gene list L, increasing a running-sum enrichment score (ES) statistic whenever a gene from gene list S is encountered (P_{hit}) and decreasing it when a gene not from gene list S is encountered using the following formulas (P_{miss}) [191]. ES is a weighted Kolmogorov-Smirnov-like statistic, and is equal to the maximum derivation of zero of $P_{hit} - P_{miss}$, where:

$P_{hit}(S,i) = \Sigma_{g_j=S, j\leq i} |r_j|/(\Sigma_{g_j=S}|r_j|)$, and $P_{miss}(S,i) = \Sigma_{g_j\neq S, j\leq i}(1/(N - N_H))$

S represents pathway genes, i is the gene position on the list L, g_j is the gene at position i, r_j represents correlation of the gene to the phenotype of interest (in this case, resistant

women), N represents all genes in gene list L, and N_H represents the number of genes in gene list S.

The statistical significance of ES is determined by an empirical phenotype-based permutation test procedure. Here, the phenotype labels are randomly assigned and the gene list L is permutated 1000 times and ES is re-calculated ("null distribution"). The observed ES is then compared to the null distribution ES score and an empirical, nominal p-value is estimated. To correct for gene sets of different size, the ES is normalized to the gene list input size, generating the Normalized Expression Score (NES). The false discovery rate is then calculated for each NES by comparing observed and null distributions of NES [184].

GSEA is linked to the Molecular Signatures Database (MsigDB), which is a depository for previously generated gene lists and gene expression pathways. Two of the databases were accessed to look for pathway enrichment in HIV-1 resistant CMCs. The curated gene lists (C2) is a depository for pathway annotations and other published gene lists, and gene ontology (GO) database (C5) has genes grouped based on common biological themes.

3.2.25. Protein isolation and western blotting

To further confirm genes identified as significantly differentially expressed between CMCs and PBMCs, 8 unmatched CMC and 5 PBMC donors were recruited for protein quantification studies. CMCs and PBMCs were isolated as described in sections 3.2.3 and 3.2.4, respectively and pelleted by centrifugation at 1600 r.c.f. for 5 minutes. CMCs and PBMCs were resuspended in 50 ul and 200 ul cell lysis buffer, respectively.

Samples were centrifuged for 10 minutes at 10,000 r.c.f. at 4°C, and stored at -40°C. Protein concentrations were determined using the Novagen BCA Protein Assay Kit following manufacturer's micro-scale assay instructions.

Ten ug of each protein sample were loaded onto a 12% Nupage® mini gel and electrophoresed at 200V for approximately 35 minutes. The gel was then washed in ddH_2O for 2 minutes, and proteins were transferred onto a nitrocellulose membrane using the Invitrogen iBlot® gel transfer device according to the manufacturer's instructions. The nitrocellulose membranes were then washed with ddH_2O for 2 minutes and loaded into a pre-wet SNAP i.d. Single Well Blot Holder. Membranes were blocked with a 0.5% milk blocking solution for 10 minutes, and followed by 20 seconds of vacuum to remove the solution. Three ml of a 1:100 dilution of primary antibody was applied to each blot, incubated at room temperature for 30 minutes, and antibody solution was removed by vacuum. Membranes were washed 4 times with 30 ml 1X TBS + 0.1% Tween-20 under continuous vacuum. Membranes were then incubated with 3 ml of a 1:3200 dilution of goat anti-rabbit secondary antibody conjugated to horse-radish peroxidise for 15 minutes and washed 4 times with 30 ml 1X TBS + 0.1% Tween-20 under continuous vacuum. Membranes were quickly rinsed with 10 ml ddH_2O and incubated with Novex® ECL Chemiluminescent Substrate Reagent Kit for one minute. Chemiluminescence was detected with the Bio-Rad ChemiDoc XRS system and band intensities were determined using the Bio-Rad Quantity One Software. Membranes were then stripped with a 0.1 M glycine stripping solution, washed for one hour in 10 ml 1X TBS + 0.1% Tween-20 at r.t., and re-probed with actin and imaged as described above. OSM, $cFLIP_{S/L}$ and NFκB p52 protein levels were normalized to total loaded protein by

dividing the target protein band intensities by the actin band intensity for each sample. Target protein fold changes between CMCs and PBMCs were calculated by dividing this normalized protein value in CMCs by the one calculated for PBMCs. P-values were calculated using GraphPad Prism 4.

3.2.26. Alternative Western Blot protocol

In an attempt to increase the protein signal intensity across samples, an alternative western blot protocol was tested across PBMC and CMC protein isolates. Samples were run on gel and transferred as above. Membranes were then blocked for one hour with a 1X TBS + 0.1% Tween-20 + 5% skim milk blocking solution at room temperature and then incubated overnight with 15 mL of a 1:500 dilution of the same primary antibodies described in section 3.2.25. Membranes were washed twice, each time for 5 minutes followed by three times for 15 minutes each wash with 30 ml of a 1X TBS + 0.1% Tween-20 solution. Membranes were then incubated with a 1:10,000 dilution of secondary antibody for one hour at room temperature, and washed as above. Proteins were detected as described in section 3.2.25.

3.2.27. Statistics

P-values for the of qRT-PCR data, Western blots and CMC counts were calculated in GraphPad Prism 4 using a Mann Whitney U test, and p-values less than 0.05 were considered to be not significant. Statistical tests for other experiments were performed as described in their individual sections.

4. Results

4.1. Objective 1: Gene expression analysis in CMCs versus PBMCs

4.1.1. Patients

Healthy women were recruited from the Health Sciences Centre Department of Obstetrics & Gynaecology Colposcopy Clinic in Winnipeg, Manitoba. The women were recruited to optimize protocols and to compare gene expression patterns between donor-matched CMCs and PBMCs. To minimize difference in gene expression caused by the menstrual cycle or the presence of an infection, women scheduled for cervical biopsies, menstruating, or pregnant were excluded from the study.

4.1.2. Cell counts

The first 35 women recruited for this study were asked to donate cervical scrapings to optimize CMC yields from the scrapings. Several modifications of the technique described by Iqbal *et al.* [132] were tested, and the highest CMC yields were achieved via treatment with TrypLE for one hour at 37°C (section 3.2.3) (Figure 7). The filter and nylon wool separation techniques yielded no CMCs, and heating the samples without TrypLE did not show any appreciable increase in CMC yields. It was not possible to determine what effect TrypLE treatment would have on gene expression in cells, and in the end rigorous shearing of the samples to break up mucus followed by separation on Ficol Hypaque was determined to be the best protocol for cell yields and minimal gene expression changes.

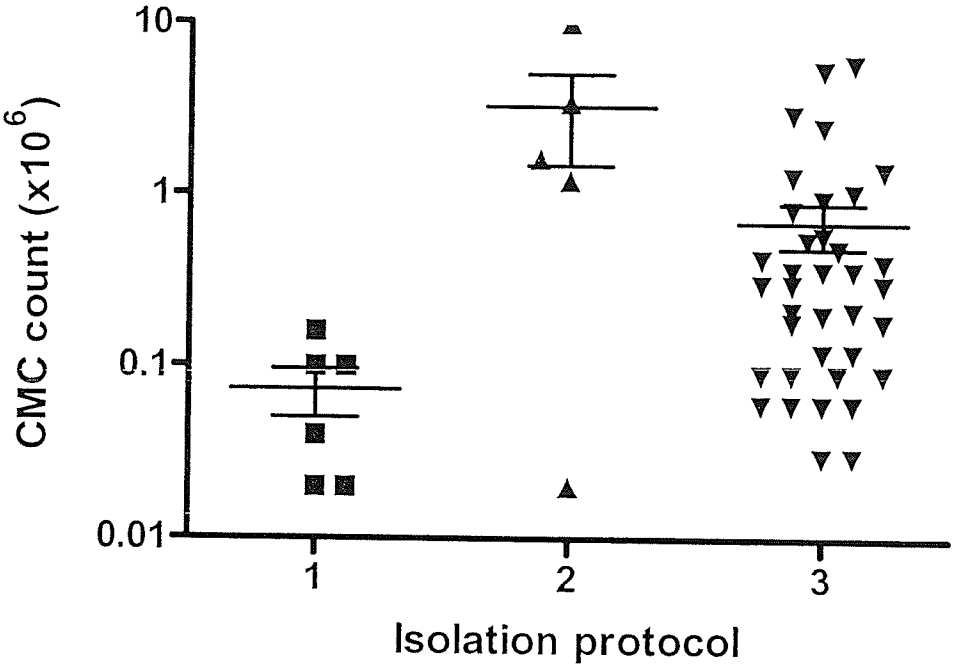


Figure 7: Winnipeg CMC yields. Several protocol variations were attempted to maximize CMC yields from cervical scrapings. Protocol 1 was adapted from Iqbal *et. al.* [132]. Protocol 2 included a one-hour incubation of the cytobrushes with TrypLE, and protocol 3 was similar to protocol 1 but included vigorous pipetting of the samples to disintegrate cervical mucus. Although protocol 2 gave the highest cell yields, the effect on gene expression by this cell treatment could not be determined. Therefore, protocol 3 was determined to be the best isolation protocol for future CMC genetic analysis. The higher cell yields for protocol 3 over protocol 1 is significant at $p < 0.05$. Error bars represent one standard deviation from the average CMC count obtained from each isolation protocol.

4.1.3. RNA quality

Two RNA isolation techniques were tested and evaluated in an attempt to maximize RNA yields from the isolated CMCs. The Qiagen RNeasy Plus Kit is well described in the literature to have high RNA yields, but did not to isolate any significant amount of quality RNA from the 5 CMC samples tested (Figure 8a). TRIzol® is a much more caustic RNA isolation reagent, and was highly effective in isolating quality RNA from CMCs (Figure 8b). However, this technique is more sensitive to phenol contamination, which can have an effect on downstream applications [192]. Consequently, RNA had to undergo an additional cleanup and DNase-digestion before being run on arrays. RNA extraction trials from CMCs donated in Winnipeg indicated that 300,000 CMCs was the minimum number of cells needed to reliably have sufficient RNA to run on the Illumina HumanRef-8 v2 Expression BeadChips.

4.1.4. Differential gene expression

Twenty-five women were recruited from the Winnipeg Health Sciences Centre to donate matched CMCs and PBMCs for this study, from which 10 CMC samples had RNA of high enough quality to be run on gene expression arrays. Matched CMC and PBMC samples were taken for this study to increase the statistical power of the results. Of the 24,500 genes represented on the Illumina HumanRef-8 v.2 Gene Expression BeadChip array, 11,586 genes were expressed in at least one CMC or PBMC sample.

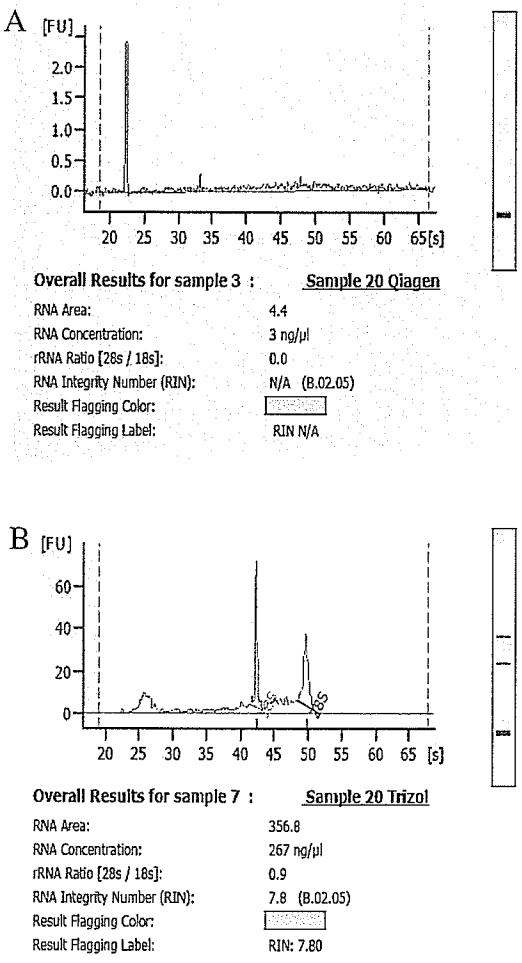


Figure 8: Representative RNA extractions from CMCs. CMCs from donor 20 was divided equally and extracted with (A) Qiagen RNeasy Plus Mini Kit and (B) TRIzol®. [FU] represents the fluorescence units of RNA-intercalating dye. The peak at 20[s] in (A) represents the marker peak of the Agilent BioAnalyzer. The peaks at 43 and 50 [s] in (B) represent 18S and 28S ribosomal RNA bands, respectively. RNA integrity number (RIN) is a measure of degradation of the RNA sample [193]

Of these, 5345 genes were significantly differentially regulated between CMCs and PBMCs with an absolute fold change of 1.3 or greater and an adjusted p-value of <0.05 (gene list: all_diff_exp). Fold changes of 1.3 or greater have been previously shown to be reproducible from Illumina gene expression data [194] and therefore was the fold-change cut-off for this experiment. This list was divided into over expressed (CMC_up) (2342) or under expressed genes (3003) (CMC_down) in CMCs.

4.1.5. Clustering

Normalized gene expression intensities of all genes expressed by matched CMCs and/or PBMCs were inputted into Cluster® for unsupervised hierarchical clustering. CMCs and PBMCs from these matched donors clustered independently, as shown by Figure 9. Heatmaps were also generated for all significantly differentially expressed and the 100 most significantly differentially expressed genes (Figure 9). All three heatmaps show PBMCs and CMCs clustering independently, suggesting CMCs are genetically distinct from PBMCs. The data indicated CMCs and PBMCs do have unique gene expression patterns and represent distinct functional phenotypes based upon gene expression.

The majority of this altered gene expression between matched CMCs and PBMCs is most likely explained by the different cell subset makeup at the two sites, as has been previously published [195]. Differences in cell subsets between CMCs and PBMCs were characterized by 10 colour flow cytometry in parallel with these gene expression studies (Kaefer and Horton, submitted [196]).

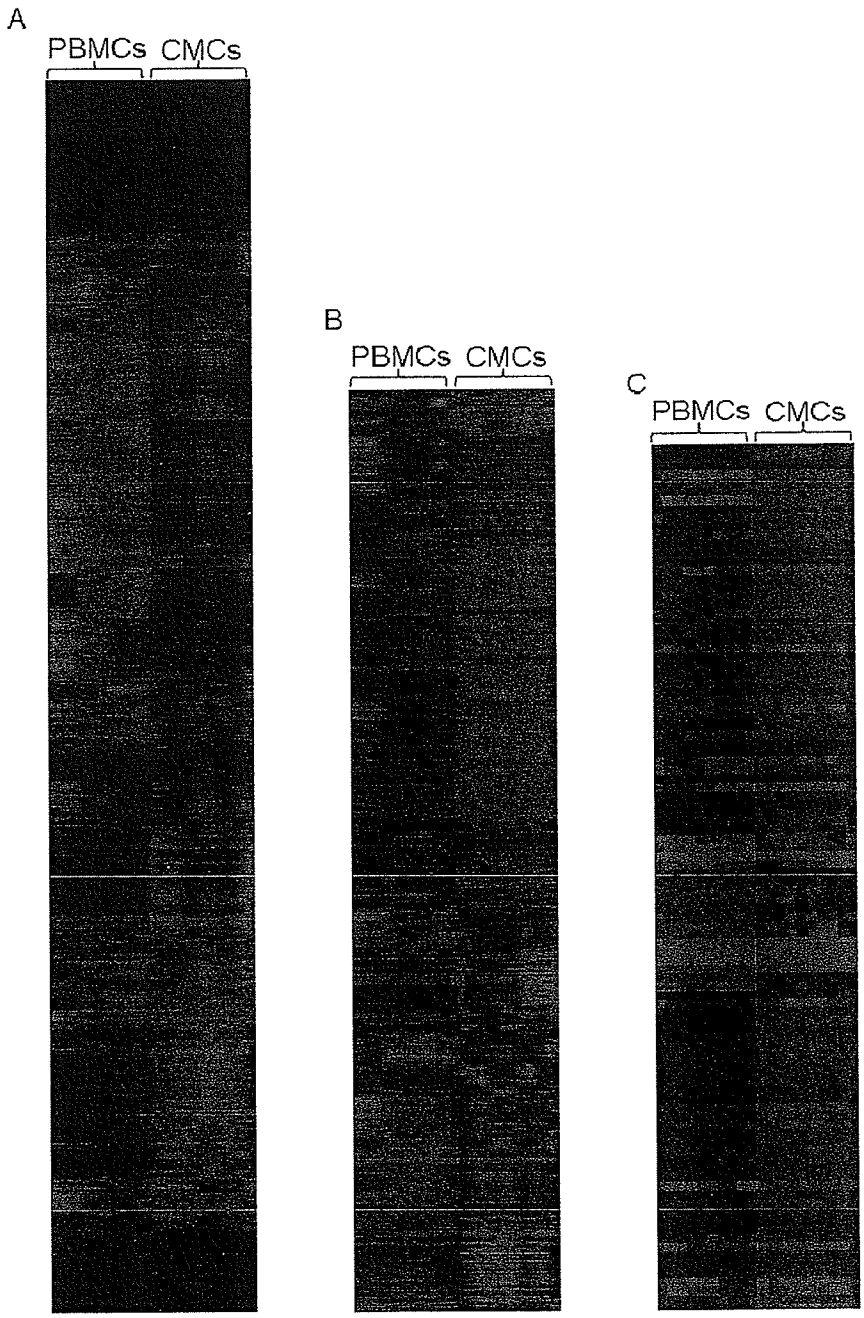


Figure 9: Overall gene expression in matched CMCs and PBMCs. CMCs and PBMCs cluster independently using unsupervised hierarchical clustering. Red indicates a relative increase and green represents a relative decrease in gene expression intensity. (A) All genes expressed by CMCs and PBMCs. (B) Significantly differentially expressed genes only. (C) The 100 most significantly differentially expressed genes.

When compared with age-matched PBMCs, CMCs showed increased proportions of CD14+ monocytes (5.5 ± 1.9% vs. 28.2 ± 5.6%; p = 0.005) and lineage (CD3, CD14, CD19, CD56) negative, HLA-DR+ DCs (2.2 ± 0.4% vs. 3.5 ± 0.3%; p = 0.025). CMCs showed relatively lower levels of CD3+ T cells (61.0 ± 4.0% vs. 32.3 ± 10.0%; p = 0.028), while levels of CD56+ NK cells were the same (7.3 ± 2.3% vs. 8.3 ± 2.3%; p = ns) and CD19+ B cells (9.0 ± 1.4% vs. 4.4 ± 1.5%; p = 0.059) trended towards a lower level in CMCs (Kaefer and Horton, submitted [196]). There was no significant difference in CD4 expression when measured by flow cytometry or by microarray analysis. Figure 10a shows the relative differences in cell phenotype between CMCs and PBMCs and Figure 10b shows relative differences in gene expression of the cell markers.

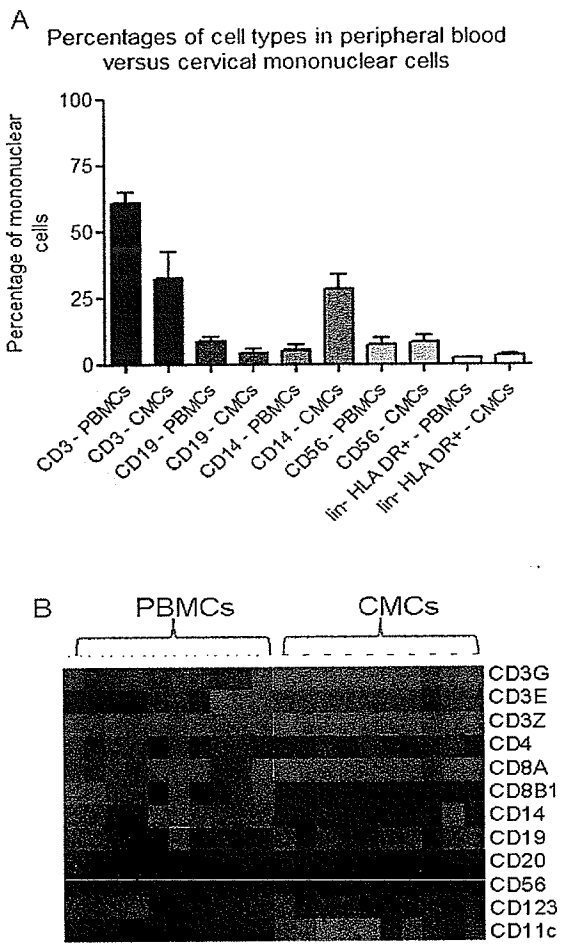


Figure 10: Differences in cell phenotype between CMCs and PBMCs. (A) Cell phenotype comparison between the two cell populations by flow cytometry. There was a significant increase in CD14+ and lineage negative HLA-DR cells and a significant decrease CD3+ cells in CMCs as compared to PBMCs. There was no significant difference between CD56+ cells, and there was a trend toward significantly fewer CD19+ cells in CMCs. (B) Gene expression of the phenotyping markers described in (A), where red indicates an increase in gene expression and green indicates a decrease in gene expression. The gene expression results correlate with the cell phenotyping data, although the expression of CD19 was significantly lower in CMCs using this method ($p = 0.0014$) (from Kaefer and Horton, submitted [196]).

4.1.6. Pathway analysis

4.1.6.1. Pathways differentially regulated by CMCs

The all differentially expressed gene list was uploaded into DAVID for pathway analysis [197]. 5011 GeneIDs were recognized, 1301 (26% of inputted gene list) of which were significantly enriched (EASE score < 0.05) in 25 KEGG pathways. All pathways identified by DAVID are listed in Table 3. Several pathways involved in regulating the immune response were identified as being differentially expressed by CMCs. Such pathways included the T cell receptor signalling pathway (48 genes), B cell receptor signalling (30 genes), natural killer cell-mediated cytotoxicity (53 genes), and the hematopoietic cell lineage (37 genes). Several cellular process and metabolic pathways were also identified, including cell cycle (53 genes), apoptosis (35 genes), oxidative phosphorylation (57 genes), and ubiquitin-mediated proteolysis (51 genes) (Table 3).

4.1.6.2. Pathways up-regulated in CMCs

Of the 2342 genes that were up-regulated in CMCs, 2180 GeneIDs were recognized by the program. 28 KEGG pathways had 653 genes (30%) significantly enriched. Pathways from the complete significant gene list overlapped with many of the up-regulated pathways but several novel pathways were found by dividing the list. Such pathways included the toll-like receptor signalling pathway (29 genes), complement and coagulation cascades (22 genes), and cytokine-cytokine receptor interactions (51 genes) (Table 3).

Table 3: Pathways differentially regulated by CMCs identified by DAVID

Pathway	List[a]	Pop[b]	p	Gene list
T cell receptor signalling pathway	48	94	6.85E-05	all_diff_exp[c]
Ribosome	50	100	9.32E-05	all_diff_exp
Epithelial cell signalling in *H. pylori* infection	37	68	1.09E-04	all_diff_exp
Cholera – Infection	25	41	2.08E-04	all_diff_exp
Alzheimer's disease	19	28	2.59E-04	all_diff_exp
Cell cycle	53	114	5.54E-04	all_diff_exp
Oxidative phosphorylation	57	128	0.00121	all_diff_exp
Galactose metabolism	18	32	0.00696	all_diff_exp
B cell receptor signalling pathway	30	65	0.0128	all_diff_exp
NK cell mediated cytotoxicity	53	129	0.0130	all_diff_exp
Hematopoietic cell lineage	37	85	0.0155	all_diff_exp
Pyrimidine metabolism	37	88	0.0277	all_diff_exp
N-Glycan biosynthesis	20	42	0.0342	all_diff_exp
Apoptosis	35	84	0.0373	all_diff_exp
Ubiquitin mediated proteolysis	51	131	0.0422	all_diff_exp
Insulin signalling pathway	52	134	0.0426	all_diff_exp
Aminoacyl-tRNA biosynthesis	18	38	0.0490	all_diff_exp
Complement/coagulation cascades	22	69	0.00119	CMC_up[d]
TLR signalling pathway	29	102	0.00122	CMC_up
Galactose metabolism	12	32	0.00642	CMC_up
Aminosugars metabolism	11	28	0.00682	CMC_up
Adipocytokine signalling pathway	19	73	0.0263	CMC_up
Huntington's disease	10	29	0.0264	CMC_up
Cytokines and cytokine receptors	51	256	0.0418	CMC_up
Fructose and mannose metabolism	12	42	0.0498	CMC_up
Purine metabolism	38	145	9.24E-04	CMC_down[e]
RNA polymerase	11	23	0.00121	CMC_down
Folate biosynthesis	14	41	0.00667	CMC_down
DNA polymerase	10	24	0.00692	CMC_down
Pyruvate metabolism	14	42	0.00834	CMC_down
Parkinson's disease	9	21	0.00967	CMC_down
Citrate cycle	10	30	0.0322	CMC_down
Aminophosphonate metabolism	7	17	0.0357	CMC_down
Phe, Tyr and Trp biosynthesis	5	9	0.0371	CMC_down
Histidine metabolism	12	41	0.0417	CMC_down
Proteasome	8	22	0.0417	CMC_down

[a]Number of genes from inputed list involved in the pathway

[b]Number of genes in the population involved in the pathway. Total gene population: 4214

[c]Total number of genes from gene list recognized by DAVID: 1301

[d]Total number of genes from gene list recognized by DAVID: 653. CMC_up and CMC_down to do not include pathways previously discovered by all_diff_exp

[e]Total number of genes from gene list recognized by DAVID: 651

4.1.6.3. Pathways down-regulated in CMCs

Of the 3003 genes down-regulated in CMCs, 2863 GeneIDs were recognized by the program. 19 KEGG pathways had 651 (23%) genes significantly enriched. Again, there were many overlaps with the pathways identified by DAVID analysis of the complete significant gene list. Several new pathways were identified, however, including the citric acid cycle (10 genes) and the proteasome (8 genes) (Table 3).

4.1.7. Pathways for follow-up studies

Three pathways were selected for follow up studies based on their relevance to the immunobiology of the female genital tract. The pathways selected for follow up studies were TLR signalling, apoptosis and cytokine-cytokine receptor interaction. Normalized fluorescence intensities of the identified pathways (including genes not significantly differentially expressed by CMCs) also underwent unsupervised hierarchical clustering. Again, CMCs and PBMCs cluster independently, consistent with the pathways being significantly differentially expressed between the two cell types (Figure 11).

4.1.7.1. Cytokine-cytokine receptor interaction

A total of 78 cytokines or their receptors were differentially expressed by CMCs. There were a significantly higher number of cytokines up-regulated by CMCs than down-regulated (56 up-regulated vs. 22 down-regulated, $\chi^2 = 13.14$, df = 1, P = 0.0003).

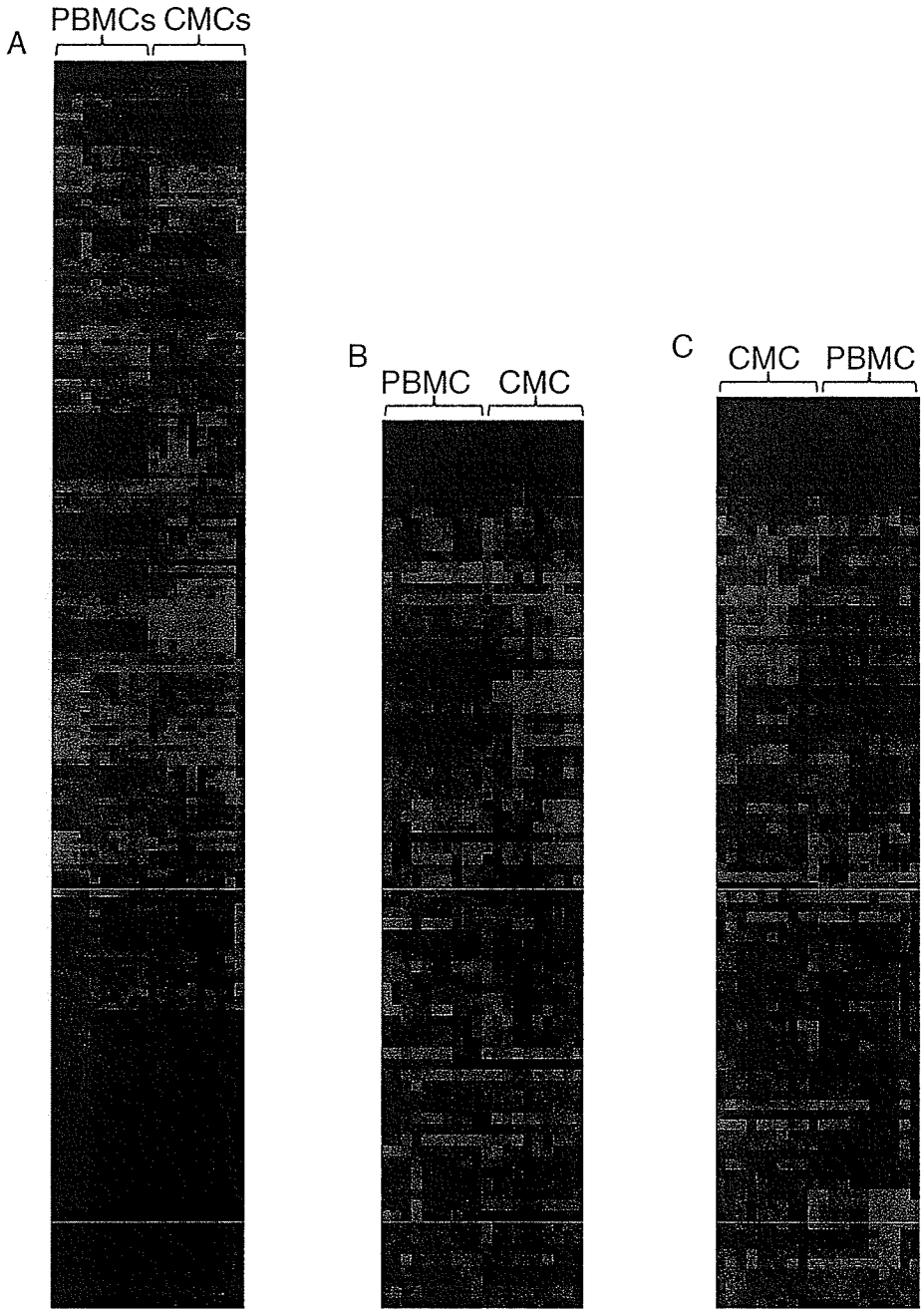
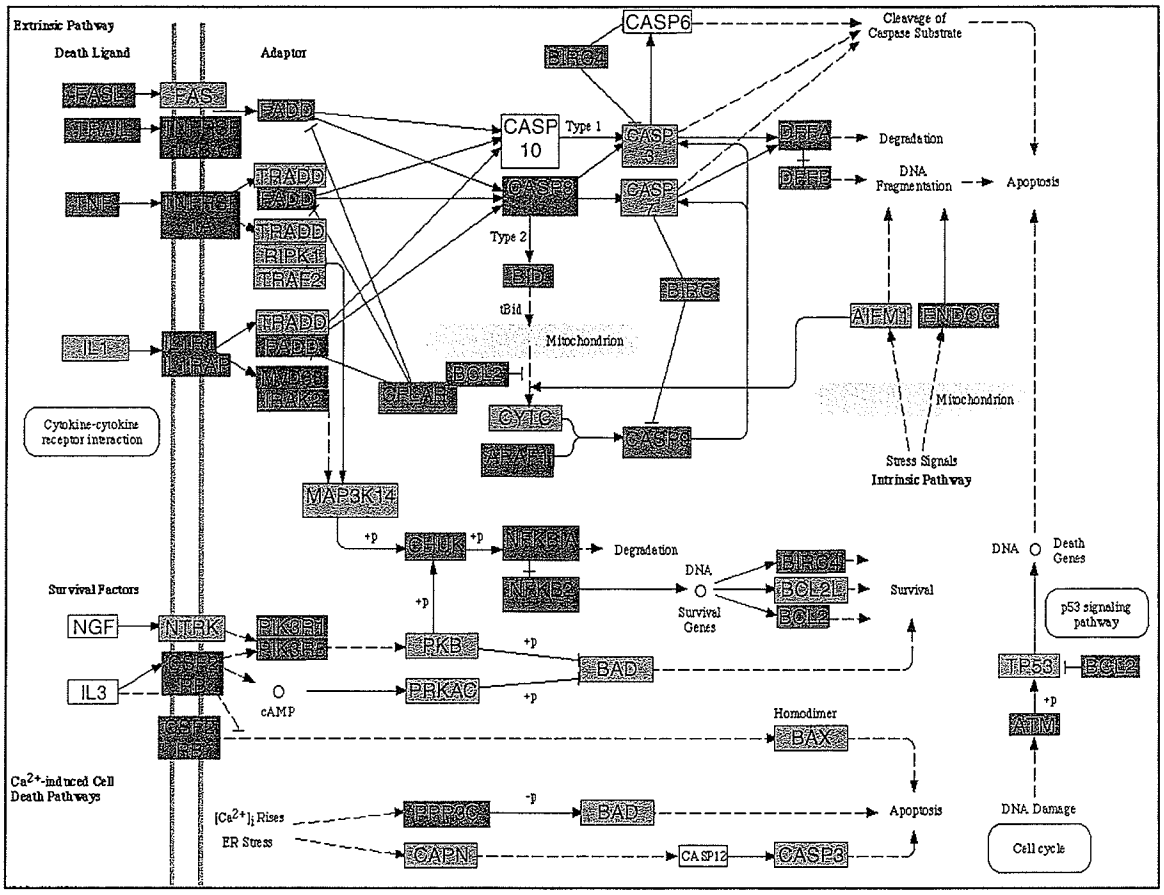


Figure 11: Heatmaps of differentially expressed pathways identified by DAVID. Red indicates a relative increase and green represents a relative decrease in gene expression intensity. Gene symbols are listed on the right of each heatmap. (A) All expressed genes involved in the cytokine-cytokine receptor interaction pathway. (B) Genes expressed in the apoptosis pathway. (C) All expressed genes involved in TLR signalling.

These data did not show a significant enrichment of pro- or anti-inflammatory cytokines in the FGT over blood, but. suggest an overall increase in cytokine-mediated immune signalling. Figure 11a shows a heatmap of all cytokine genes expressed by CMCs and PBMCs.

4.1.7.2. Apoptosis

Apoptosis pathways were significantly enriched in two gene lists – all_diff_exp and CMC_up. Despite the apparent increase in the expression of apoptosis-triggering cytokines such as TNF (3.02-fold up-regulated, $p = 5.8 \times 10^{-4}$), there was a concomitant increase in downstream anti-apoptotic factors, including caspase 8 and Fas-associated death domain-like apoptosis regulator (CFLAR), which was 2.04-fold up-regulated ($p = 0.0028$) and a 1.73-fold up-regulation of X-linked inhibitor of apoptosis (XIAP) ($p = 1.9 \times 10^{-4}$). There was also a decrease in the expression of cytochrome c (1.49-fold, $p = 0.026$), an important inducer of the intrinsic apoptosis pathway. Caspase 9, one of the first proteins in the caspase cascade, was significantly over expressed by CMCs (2.11-fold, $p = 1.6 \times 10^{-4}$); however, genes responsible for apoptotic end functions such as DNA degradation (DNA fragmentation factor B [DFFB]: 1.87-fold down-regulated, $p = 7.63 \times 10^{-6}$) and caspase cascade initiation were either significantly down-regulated (Fas-associated death domain [FADD]: 1.54-fold down-regulated, $p = 3.9 \times 10^{-5}$), not significantly differentially expressed (caspase 3: 1.17-fold down-regulated, $p = 0.45$), or not expressed by either cell type (Figure 12). Overall, the data for this pathway suggests an anti-apoptotic environment in the FGT due to downstream suppression of apoptotic pathways in the presence of upstream apoptosis-inducing signals.



Increased in CMCs Decreased in CMCs Not significantly differentially expressed

Not expressed in CMCs or PBMCs → Activation --→ Indirect interaction

—| Inhibition —+— Dissociation ___ Binding p Phosphorylation o→ Gene expression

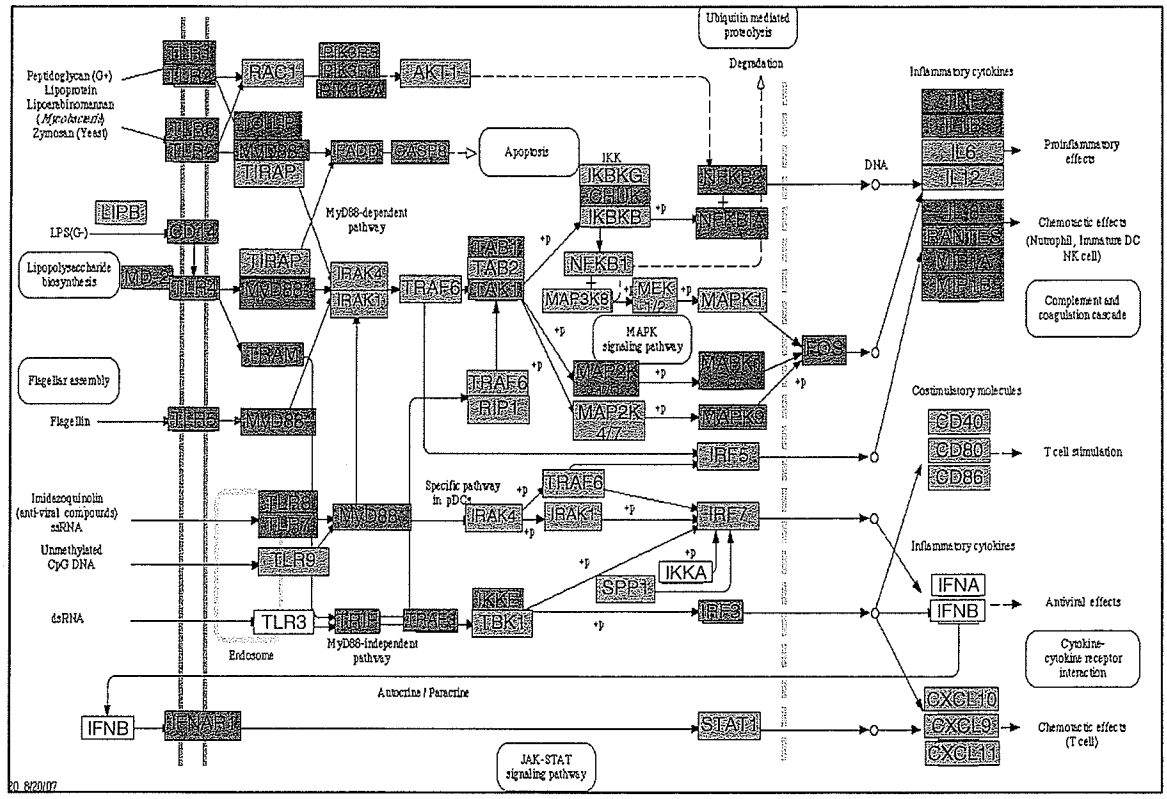
Figure 12: The apoptosis pathway (modified pathway 04210 from KEGG, http://www.genome.jp/dbget-bin/show_pathway?map04210) [188-190]. CFLAR, denoted by *, is an important inhibitor of apoptosis and is significantly up-regulated by CMCs by a factor of 2. Figure reproduced with permission from KEGG via Dr. Yamamoto, obtained May 15th, 2009.

4.1.7.3. Toll-like receptor signalling

Several TLR pathway genes were significantly enriched in the CMC_up gene list. Nearly all the TLR receptors were differentially expressed, with TLRs 1, 2, 4, 5, 6 and 8 significantly up-regulated and TLR7 significantly down-regulated. Neither cell population expressed TLR3 nor was there a significant difference in expression of TLR9 between CMCs and PBMCs. MyD88, an important regulator of the TLR signalling pathway for all TLRs except TLR3, was significantly up-regulated in CMCs 2.39-fold ($p = 2.6x10^{-4}$). There was also an increase in some cytokine end products of TLR signalling, including a 5.24-fold increase in macrophage inflammatory protein 1α (MIP-α, $p = 0.0016$) and a 5.52-fold increase in MIP-1β expression ($p = 0.0035$), but not DC activation markers CD86 or CD40, nor antiviral factors IFNα, IFNβ or IFNγ (Figure 13). Differences in TLR gene expression could be due to differences in cell populations between CMCs and PBMCs (Kaefer and Horton, submitted [196]); however, the data also show increased downstream TLR signalling complete with increased cytokine expression, which is important in the context of HIV-1 transmission.

4.1.8. Quantitative Real-Time RT-PCR (qRT-PCR)

To confirm the differences in gene expression between CMCs and PBMCs elucidated by the gene expression arrays, one gene was selected for quantification by qRT-PCR from each pathway and 18S was used as a housekeeping gene to normalize gene expression levels. Five matched CMC and PBMC samples from the original gene expression arrays had sufficient RNA (100 ng) for qRT-PCR confirmation studies.



Increased in CMCs Decreased in CMCs Not significantly differentially expressed

Not expressed in CMCs or PBMCs → Activation --→ Indirect interaction

—| Inhibition —+— Dissociation ___ Binding p Phosphorylation o→ Gene expression

Figure 13: The TLR pathway (modified pathway 04620 from KEGG, http://www.genome.jp/dbget-bin/show_pathway?map04620) [188-190]. MyD88, denoted by *, is an important inhibitor of apoptosis and is significantly up-regulated by CMCs by a factor of 2.39. Figure reproduced with permission from KEGG via Dr. Yamamoto, obtained May 15th, 2009.

Oncostatin-M (OSM) from the cytokine-cytokine receptor interaction pathway, CFLAR from the extrinsic apoptosis pathway and MyD88 from the TLR signalling pathway were selected as representative genes to confirm the gene expression arrays.

The OSM gene was found to be 18-fold up-regulated in CMCs with an adjusted p-value of $4x10^{-4}$ as determined by the Illumina gene expression BeadArrays. OSM is a potent mediator of inflammation [198]. In general, it is not a highly expressed gene, and it was initially difficult to confirm the differential expression of this gene. One ug of RNA from 19 RNA tissue standards (purchased from Ambion) as well as RNA from PHA-stimulated PBMCs were screened to identify a suitable sample with detectable OSM signals. This sample was then used to generate a standard curve to quantify the differences in OSM expression between CMCs and PBMCs by qRT-PCR. A pool of RNA from CMCs and PBMCs would have made the ideal standard curve, but this was not possible due to limited CMC RNA availability at the time of this study. Of the 20 tissues screened, only the brain standard RNA had OSM levels detected after 35 PCR cycles, which was the limit of detection of the LightCycler [199] and therefore 5-fold serial dilutions of cDNA made from 1 ug of brain RNA were used as a standard curve. Despite efforts to improve quantification of this gene by qRT-PCR, only 2 CMC and 2 PBMC samples tested had detectable levels of OSM. Furthermore, the qRT-PCR results indicate there was a 1.7-fold down-regulation of OSM in CMCs with an insignificant p-value of 0.45, which was inconsistent with the results from the gene expression microarrays.

CFLAR from the apoptosis pathway was chosen as a second candidate gene for qRT-PCR confirmation due to its importance as an inhibitor of apoptosis (Figure 12).

The brain RNA standard curve was also used to quantify this qRT-PCR data. From the BeadArrays, there was a 2.04-fold up-regulation of this gene in CMCs, with an adjusted p-value of 0.0028. This gene was more highly expressed than OSM and was quantifiable in all 5 CMC and PBMC samples taken for confirmation studies. The qRT-PCR data showed an opposite trend to the gene expression array results, with CFLAR down-regulated 1.9-fold with a p-value of 0.13.

MyD88 is an important intermediary signalling molecule in the TLR signalling pathway, and was found to be 2.39-fold over expressed by CMCs with an adjusted p-value of 2.6×10^{-4} from the gene expression arrays. MyD88 was also detected in all 5 matched CMC and PBMC samples tested. qRT-PCR was once again not able to confirm the array results, showing a 5.2-fold down-regulation of MyD88 in CMCs ($p = 0.005$).

4.1.8.1. Troubleshooting the qRT-PCR protocol

One significant difference between the gene expression array data and the qRT-PCR data was that samples undergo *in vitro* transcriptional amplification before being hybridized to the gene expression arrays. Therefore, to troubleshoot the poor correlation between the qRT-PCR data and the microarray data, the Sigma TransPlex® Complete Whole Transcriptome Amplification Kit was tested to determine if it would be effective in increasing the gene intensities to levels seen on the gene expression arrays. Furthermore, increasing the amount of available RNA from the samples would allow for the generation of an ideal standard curve from pooled samples, and the inclusion of more samples in the qRT-PCR study would increase the statistical power of the experiment. To that end, 5, 25 and 125 ng of the same standard brain RNA used to make the standard curve in the original qRT-PCR experiment was linearly amplified and expression levels

of 18S and CFLAR were measured. 18S amplified as expected, where the amount of gene detected increased approximately 5-fold for each 5-fold increase in input RNA. However, for CFLAR, the expression levels decreased as the amount of input RNA increased. CFLAR levels in the 25 ng preparation were 3-fold lower than in the 5 ng sample, and CFLAR levels in the 125 ng amplified sample were below the limit of detection of the machine. By comparison, 125 ng of unamplified brain standard RNA had detectable CFLAR levels, suggesting this amplification protocol inhibited the detection of this gene. Therefore, this confirmation method was abandoned.

4.1.9. Western blot confirmation of gene expression microarrays

Eleven CMC donors from the Winnipeg Colposcopy clinic and seven age-matched female local PBMC donors (CMC-matched PBMCs were not available from these donors) were recruited for protein confirmation studies. Five of the 8 CMC donors had sufficient protein to be tested for expression levels of 3 proteins. The sample from donor 133 was later excluded from all analyses as excessive mucus in this sample inhibited protein migration. The proteins selected for confirmation of gene expression pathway analysis results were analyzed by Western blot and included OSM, protein 52 (p52), which is the protein product of NFKB2, and cellular FLICE-(caspase 8)-like inhibitory protein (cFLIP), which is the protein product of CFLAR. MyD88 was replaced by p52 to confirm TLR signalling as p52 is a subunit of NFκB, which is one of the end products of this pathway and actually induces inflammatory cytokine transcription. However, p52 is involved in pathways other than TLR signalling and,

therefore, MyD88 expression levels should also be done at a later date to demonstrate a specific increase in TLR signalling.

Although the OSM antibody hybridized many proteins in CMC and PBMC samples using both the SnapID and alternative Western blot protocols, none of the bands detected were at or near the expected size of 28 kDa. The antibody product information sheet did not indicate it would pick up different band sizes, nor was there data in the literature to suggest there were alternative isoforms of this protein. Thus it was impossible to determine if this antibody was binding to OSM and therefore was not a suitable means to confirm the gene expression arrays.

Western blots with the p52 antibody were unsuccessful when tested with both protocols. Only one CMC sample and no PBMC sample had p52 levels within the detection limit of this protocol. Blots were re-probed with actin and showed good protein levels, indicating p52 was possibly too lowly expressed by the cells to be detected by Western blotting.

The cFLIP antibody was more specific, and detected a band at the predicted size of 28 kDa ($cFLIP_S$) in 7 of 11 CMC samples tested but in none of the 7 PBMC samples tested. A band at 55 kDa ($cFLIP_L$) was also detected in the same 7 of 11 CMC samples and in 4 of 7 PBMC samples. Band intensities were measured, normalized to actin levels, and compared between CMC and PBMCs. There was statistically significant 13.5-fold increase in $cFLIP_S$ expression in CMCs as compared to PBMCs ($p = 0.038$). There was no significant difference in expression of $cFLIP_L$ between CMCs and PBMCs. Figure 14 shows differential protein expression of cFLIP in CMCs and PBMCs.

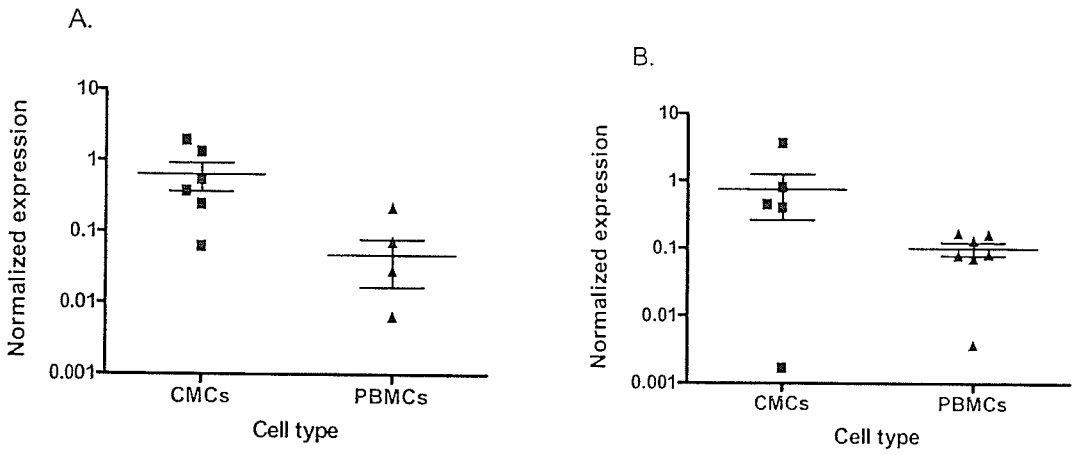


Figure 14: Western blot confirmation of the protein product of CFLAR. (A) Protein expression of cFLIPs isoform. There was a statistically significant 13.5-fold increase in cFLIPs expression in CMCs over PBMCs ($p = 0.038$). (B) Protein expression of the cFLIPL isoform. There was no significant difference in expression of this isoform between CMCs and PBMCs.

When the blots were re-probed with actin, the CMC samples show actin levels well below those of PBMCs, despite loading equal amounts of total protein for each sample, suggesting CMC protein concentrations were skewed by mucus proteins.

4.2. Objective 2: HIV-1-resistant CMC gene expression

CMCs were isolated in Kenya in May and June 2008 and were successfully shipped to Winnipeg in TRIzol®. In order to perform baseline and stimulations (described in section 4.3) on the same patient, a minimum of 1.2 million CMCs were required from each CMC donor. HIV-1 susceptible women were represented by HIV-1 uninfected women newly enrolled in the Pumwani cohort. These women were referred to as "new negatives" (NN) and were included in this study as a control group because they are from the same socio-economic group and have similar lifestyles as the HIV-1-resistant sex workers.

4.2.1. Cell yields

A total of 122 women donated cervical scrapings during the Spring 2008 resurvey for this project. The CMC yields in Nairobi were noticeably lower than those in Winnipeg. Cell yields were highly variable, but there was an apparent pattern towards an increase in CMC counts in HIV-1 infected individuals, specifically those who are currently on antiretroviral (ARV) therapy (ARV+). HIV+ donors not on ARVs had the lowest cell yields (317,000 CMCs), followed by NN and HEPS women (483,077 and 835,833 CMCs, respectively) (Figure 15).

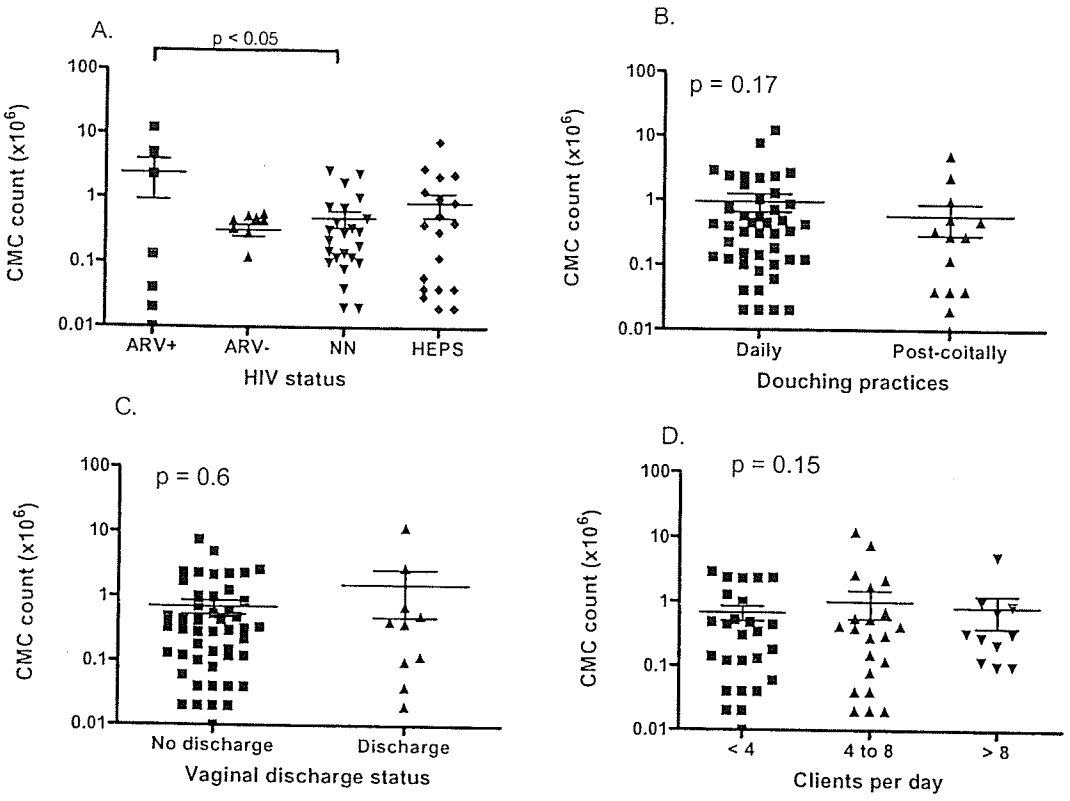


Figure 15: CMC yields from the Pumwani cohort in Kenya. (A) Comparison of cell counts between HEPS, HIV+ and NN patients. There was high variability in CMC counts, and the only significantly different mean CMC counts were between HIV+/ARV+ CMC donors and NN. Analysis was Kruskal-Wallis with Tukey's *post hoc* tests between all test groups. (B) Comparison of CMC yields between women who douche post-coitally or daily using a Mann-Whitney U test. (C) CMC yields of women reporting discharge at time of exam as compared to women without discharge symptoms, same analysis as (B). (D) Examination of the effects of clients per day on CMC yields. P-value was calculated using a Kruskal-Wallis test. None of the average CMC counts in B – D were statistically significant.

However, the only groups who had significantly different cell counts as determined by ANOVA and a Tukey's *post hoc* test were the HIV+/ARV+ donors and NN (mean CMC counts 2,432,083 vs. 483,077 respectively, $p < 0.05$). There was no significant difference between HIV-1 resistant and susceptible women ($p = 0.34$), and thus decreased susceptibility to HIV-1 infection could not be explained by a decrease in total cell number in the endocervix. Furthermore, there were no differences in cell yields when comparing women who were experiencing vaginal discharge at the time of examination to women not reporting discharge, between women who douche daily or post-coitally, or between women reporting different numbers of clients per day (group average was 4 clients/day) (Figure 15). Therefore, behavioural and biological confounders, with the exception of being HIV-1 infected and taking ARVs, do not seem to play a major role in influencing CMC yield from donors.

4.2.2. RNA quality

Despite the safe and timely shipment of the samples from Nairobi to Winnipeg, RNA yields from samples isolated in the resurvey were much lower than those from CMCs in Winnipeg. Although the average CMC yields were lower in Kenya than in Winnipeg, the CMC samples that were included in this study did have counts comparable to those obtained from donors in Winnipeg. Therefore, the low quality and quantity of RNA cannot be completely explained by lower CMC counts. None of the samples reached the minimum RNA concentration (150 ng in 11 ul, or 14 ng/ul) required for the Illumina HumanRef-8 v2 Expression BeadChips. Consequently, previously frozen samples from the Spring 2007 resurvey were used to determine baseline CMC gene

expression patterns in HEPS and NN. RNA was extracted from 55 HIV-1 uninfected donors (37 HEPS, 18 NN). Of the 55 isolations, 22 HEPS and 13 NN donors were run on the arrays. All women reported douching on the day of examination, and all but two women douched with soap and water (ML1998 used an antiseptic solution and salt water, and ML2671 douched with water alone). There was no significant difference in douching frequency (daily vs. post-coitally) between HEPS and NN.

4.2.3. Differential gene expression in HIV-1-resistant women

Of the 24,500 genes represented on the Illumina HumanRef-8 v.2 Gene Expression BeadChip array, 11,857 genes were expressed by one or more donor. Of these, 1268 were significantly differentially expressed with an absolute fold change greater than 1.3 and a p-value < 0.05. Again, a fold-change of 1.3 was chosen as the cut-off as it is the lowest fold change that can be realiably reproduced with Illumina gene expression arrays [194]. No genes were significantly differentially expressed when the p-values were adjusted for the false discovery rate (FDR). As differences between gene expression can be subtle between two phenotypes, to reduce the risk of having a high false negative rate, genes were still considered to be significantly differentially expressed without passing FDR [191]. However, any gene identified as being significantly differentially expressed by HEPS women must be confirmed by another method before publication. Of the 1268 significantly differentially expressed genes, 665 genes were under expressed and 603 were over expressed by HEPS women.

4.2.4. HEPS Clustering

Normalized gene expression intensities of all genes expressed by HEPS women and/or NN were inputted into Cluster® for unsupervised hierarchical clustering. HEPS and NN did not cluster independently, as shown by Figure 16.

To determine if different subsets of HEPS women cluster independently, whole genome expression analysis of the 10 HEPS women who have been active in sex work the longest were compared to 10 NN who have been active in sex work for the shortest amount of time. However, HEPS and NN again did not cluster independently (Figure 16). Gene intensities for only the significantly differentially expressed genes were also inputted into Cluster® for clustering analysis and HEPS and NN did not cluster independently. Heatmaps were generated for the 100 most significantly differentially expressed genes but again HEPS and NN did not cluster independently (Figure 16).

4.2.5. HEPS Pathway analysis

To detect more subtle phenotypic gene expression differences between HEPS and NN women, significantly differentially expressed gene lists were inputted into DAVID and GSEA pathway and biological theme databases.

4.2.5.1. Differentially expressed pathways by HEPS women from DAVID

All DAVID pathways identified as differentially expressed by HEPS CMCs as compared to NN are listed in Table 4. The gene list consisting of the 1268 significantly differentially expressed genes (all_diff_exp2) was inputted into the DAVID database, 146 (11.5%) of which were significantly enriched (EASE score < 0.05) in 8 KEGG pathways. The identified pathways included several cancer and metabolic processes (Table 4).

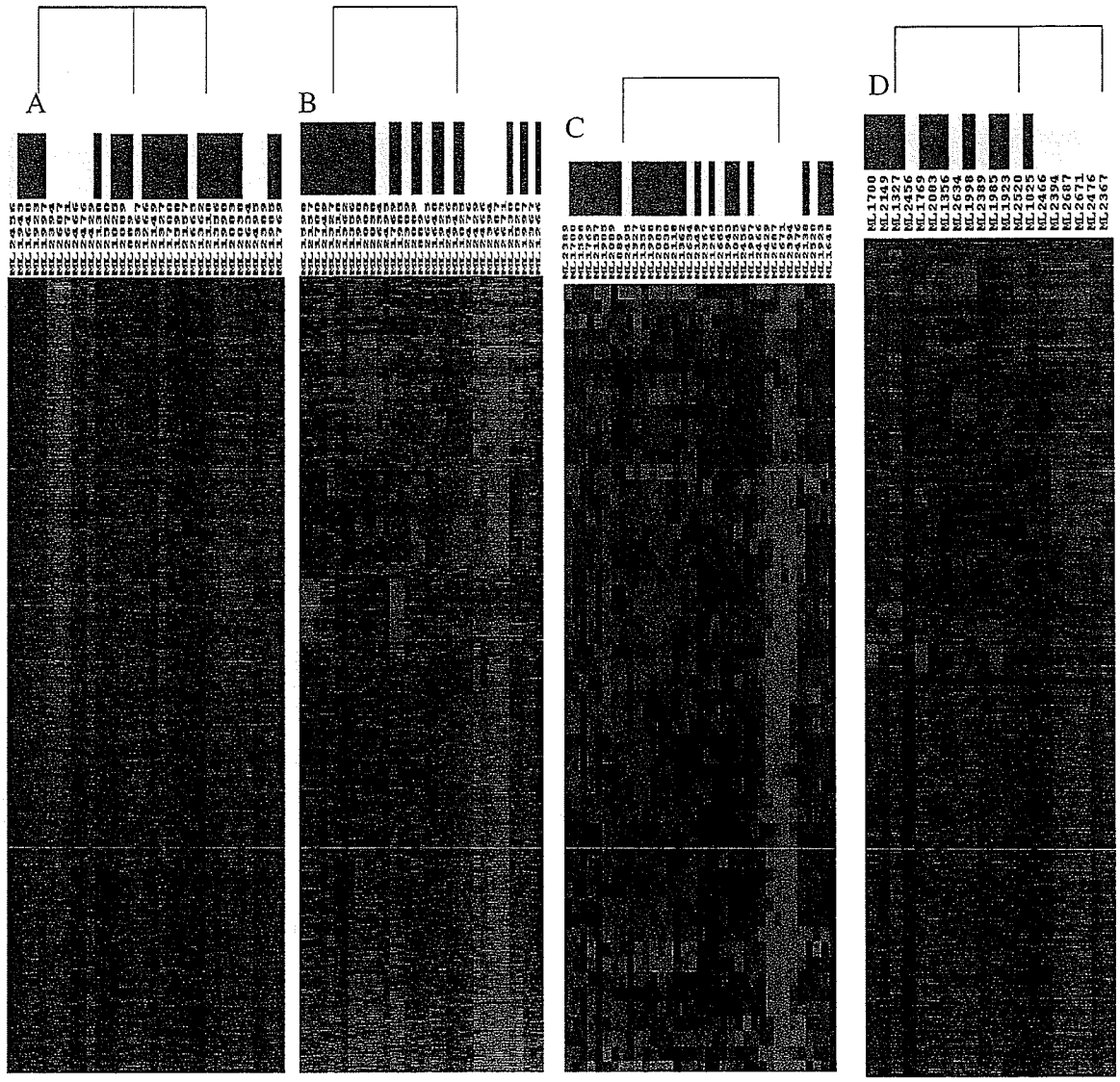


Figure 16: Overall gene expression in CMCs from HEPS and NN. Yellow boxes represent NN donors, blue boxes represent HEPS donors. HEPS and NN do not cluster independently with unsupervised hierarchical clustering. Red indicates a relative increase and green represents a relative decrease in gene expression intensity. (A) All genes expressed by HEPS and NN CMCs. (B) Genes significantly differentially expressed by HEPS CMCs as compared to NN. (C) 100 most significant differentially expressed genes (D) All genes expressed by the 10 most resistant sex workers and 10 newest sex workers.

Table 4: Pathways differentially regulated by HEPS women from DAVID

Pathway	List[a]	Pop[b]	p	Gene list
Chronic myeloid leukemia	14	74	0.003982	all_diff_exp2[c]
Bladder cancer	9	40	0.010301	all_diff_exp2
Pancreatic cancer	12	73	0.023715	all_diff_exp2
Acute myeloid leukemia	10	55	0.024062	all_diff_exp2
Long-term potentiation	11	65	0.026618	all_diff_exp2
Glutathione metabolism	8	39	0.028155	all_diff_exp2
Pyruvate metabolism	8	42	0.040505	all_diff_exp2
Axon guidance	17	132	0.047574	all_diff_exp2
Pyruvate metabolism	8	42	6.71E-04	HEPS_up[d]
Oxidative phosphorylation	12	128	0.006292	HEPS_up
Propanoate metabolism	6	34	0.007015	HEPS_up
Limonene and pinene degradation	5	29	0.019561	HEPS_up
Tryptophan metabolism	7	60	0.02065	HEPS_up
The citrate cycle	5	30	0.02195	HEPS_up
Fatty acid metabolism	6	45	0.02240	HEPS_up
Butanoate metabolism	6	45	0.02240	HEPS_up
Amyotrophic lateral sclerosis	4	19	0.029671	HEPS_up
Ascorbate and aldarate metabolism	3	9	0.039871	HEPS_up
MAPK signalling pathway	22	259	0.001843	HEPS_down[e]
Renal cell carcinoma	9	67	0.005642	HEPS_down
Acute myeloid leukemia	8	55	0.006677	HEPS_down
Pancreatic cancer	9	73	0.009441	HEPS_down
Long-term potentiation	8	65	0.016264	HEPS_down
Bladder cancer	6	40	0.022606	HEPS_down
Prostate cancer	9	87	0.025302	HEPS_down
Chronic myeloid leukemia	8	74	0.030951	HEPS_down
Type II diabetes mellitus	6	45	0.035653	HEPS_down
NK cell mediated cytotoxicity	11	129	0.037917	HEPS_down

[a]Number of genes from inputed list involved in the pathway
[b]Number of genes in the population involved in the pathway. Total gene population: 4214
[c]Total number of genes from gene list recognized by DAVID: 146
[d]Total number of genes from gene list recognized by DAVID: 67
[e]Total number of genes from gene list recognized by DAVID: 98

Lists of genes exclusively over expressed (HEPS_up) by HEPS CMCs or under expressed (HEPS_down) by HEPS CMCs were also inputted into DAVID. Of the 603 genes over expressed by HEPS CMCs, 67 genes (11.1%) were significantly enriched in 10 KEGG pathways (Table 4). Nearly all of the identified pathways were metabolic processes, including pyruvate metabolism (8 genes, $p = 0.00067$), oxidative phosphorylation (12 genes, $p = 0.006$) and the citrate cycle (5 genes, $p = 0.022$). Of the 665 under expressed genes, 98 genes (14.7%) were significantly enriched in 10 KEGG pathways. These pathways were primarily cancer and cancer-related, such as the MAPK (mitogen-activated protein kinase) signalling pathway (22 genes, $p = 0.00184$) and acute and chronic myeloid leukemia (8 genes, $p = 0.0067$ and 8 genes, $p = 0.031$ respectively) (Table 4).

4.2.5.2. Oxidative phosphorylation

Several genes involved in the protein complexes making up the electron transport chain (ETC) were found to be significantly increased in HEPS CMCs. Seven of 45 genes involved in the first complex, nicotinamide adenine dinucleotide (NADH) dehydrogenase were found to be significantly increased in HEPS CMCs. Subunit A of the succinate-Q oxidoreductase complex, an alternative ETC entry point [200], was significantly over expressed 1.37-fold ($p = 0.02$). The next protein complex in the ETC is the cytochrome b_1-c complex with 11 subunits [201], one of which, ubiquinol-cytochrome c reductase iron-sulphur subunit (UQCRFS1: 1.3-fold increased, $p = 0.03$) was significantly increased in HEPS CMCs. Genes differentially expressed by HEPS CMCs involved in oxidative phosphorylation are shown by red stars in Figure 17.

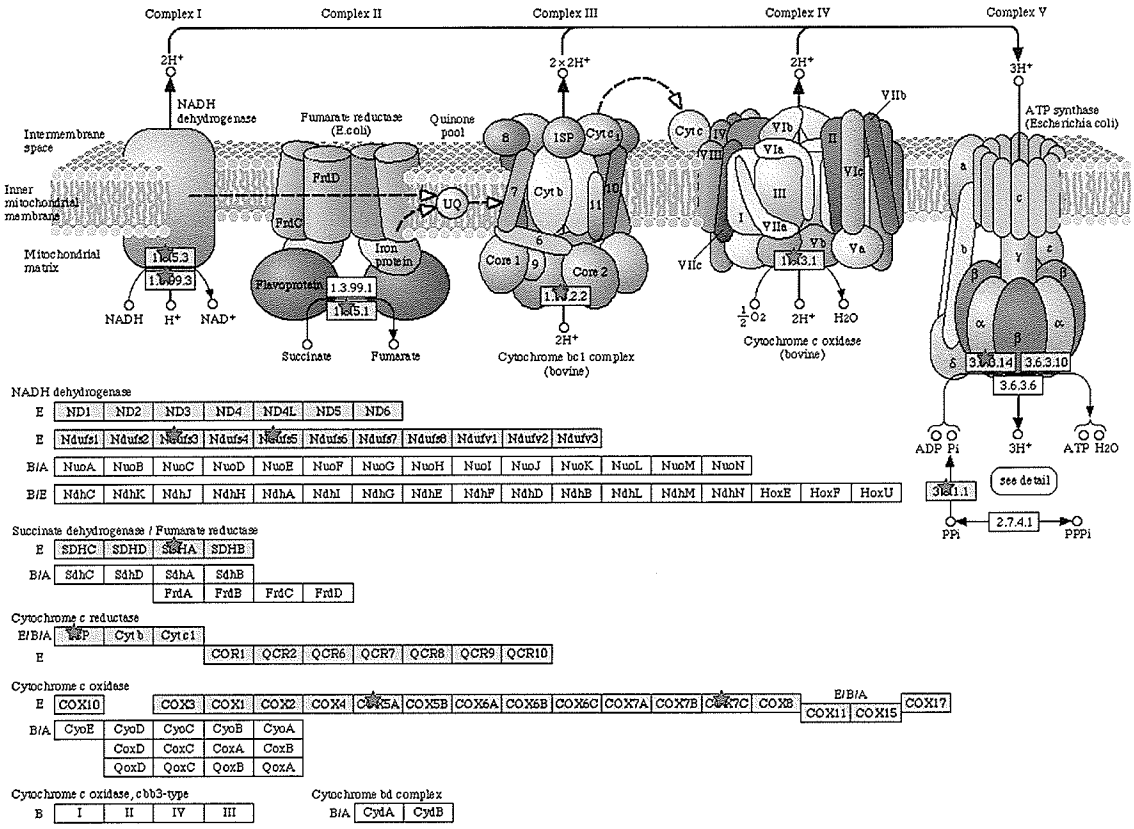


Figure 17: Oxidative phosphorylation (modified pathway 00190 from KEGG, http://www.genome.jp/dbget-bin/show_pathway?map00190) [188-190]. The red stars indicate genes significantly differentially expressed by HEPS CMCs as compared to NN. Figure reproduced with permission from KEGG via Dr. Yamamoto, obtained May 15th, 2009.

4.2.5.3. MAPK Pathway

The mitogen-activated protein kinase (MAPK) pathway is a signalling cascade that allows cells to respond to external stimuli and triggers a variety of cellular processes, including proliferation and apoptosis, and was found to be significantly under expressed by HEPS CMCs. Genes significantly under expressed by HEPS CMCs involved in the MAPK pathway are identified with red stars in Figure 18. Important MAPK cascade genes that are significantly differentially expressed by HEPS CMCs include a 1.45-fold decrease in MAPK1 expression ($p = 0.03$), a 1.47-fold decrease in v-raf murine sarcoma viral oncogene homolog B1 (BRAF) expression ($p = 0.01$) and a 1.31-fold decrease in v-raf murine sarcoma 3611 viral oncogene homolog (ARAF) expression ($p = 0.03$). MAPK frequently ends in the activation of several transcription factors involved in the cell cycle, such as activating transcription factor 4 (ATF4), which had 1.77-fold lower expression in HEPS CMCs as compared to NN ($p = 0.009$) [202]. The MAPK pathway is also inhibited by a variety of gene products, including dual specificity phosphatase 5 (DUSP5), which was 1.42-fold increased in HEPS CMCs ($p = 0.04$) [203].

4.2.5.4. Differentially expressed pathways and gene sets from GSEA

As relatively few pathways were identified as being significantly differentially expressed, GSEA analysis was also done to identify biological themes differentially expressed by CMCs from HEPS women as compared to NN. The curated gene set database (C2) and GO annotation database (C5) were accessed to identify biological themes (Table 5). The C2 database includes unconfirmed gene lists from microarray experiments; consequently, hits that were only published in one paper were excluded.

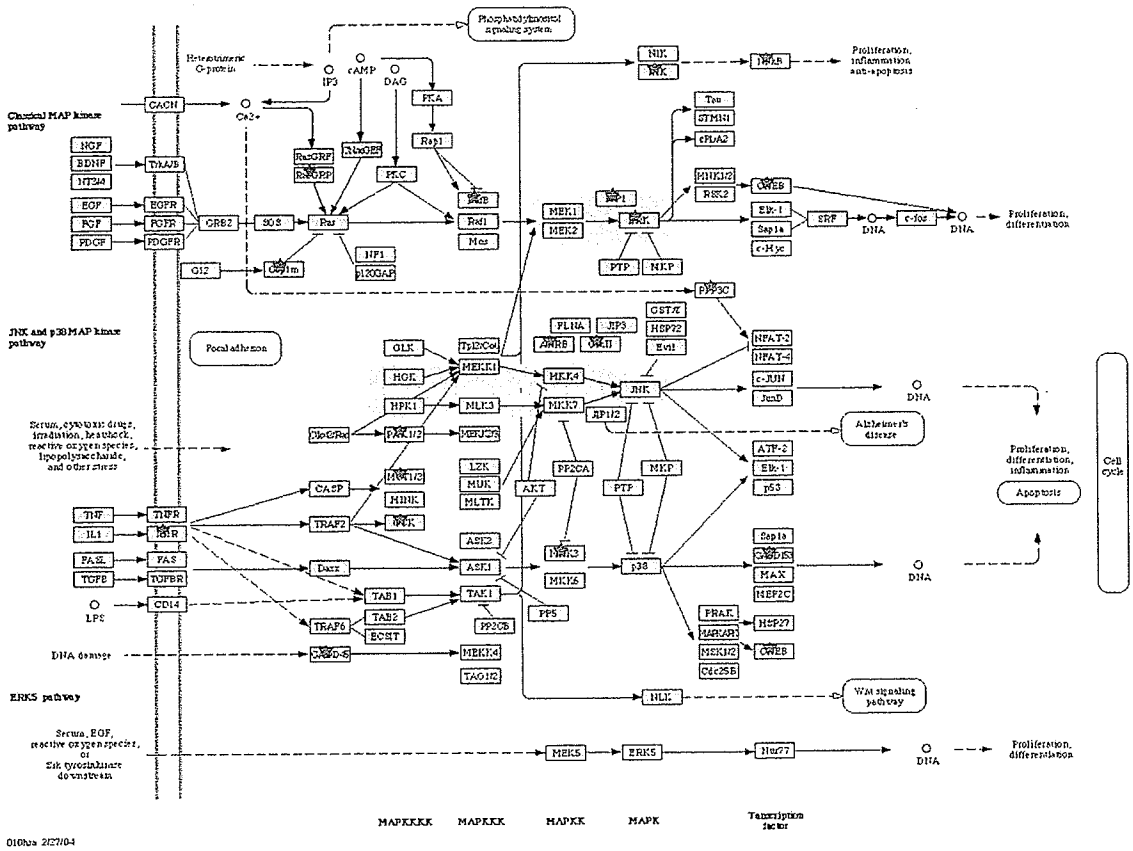


Figure 18: The MAPK signalling pathway (modified pathway 04010 from KEGG, http://www.genome.jp/dbget-bin/show_pathway?map04010) [188-190]. The red stars indicate genes significantly under expressed by HEPS CMCs as compared to NN. Figure reproduced with permission from KEGG via Dr. Yamamoto, obtained May 15th, 2009.

Table 5: Differentially expressed biological themes from GSEA

Name	ES[a]	NES[b]	FDR q-val	Gene list[c]
Electron Transport Chain	0.5289	2.4383	0	C2_Up
Mitochondria	0.3997	2.3247	7.78E-04	C2_Up
Spliceosome, methylation, Ubiquination	0.4165	2.2903	9.42E-04	C2_Up
Ribosomal proteins	0.4929	2.2693	0.001034	C2_Up
Oxidative Phosphorylation	0.4626	2.2651	0.001112	C2_Up
TGFβ, TNF, Inflammatory, Apoptotic, IFN [204]	-0.5780	-3.4082	0	C2_Down
ISRE, Influenza, Antiviral, IFN-γ/α	-0.7056	-3.3886	0	C2_Down
Toll-like receptor signalling pathway	-0.5396	-2.6399	0	C2_Down
Granulocytes, monocytes, Myeloid, ERK, Necrosis	-0.4848	-2.5497	0	C2_Down
Myeloid lineage	-0.4739	-2.5135	0	C2_Down
Structural Constituent of Ribosome	0.5642	2.61775	0	C5_Up
Mitochondrial Matrix	0.5777	2.38762	3.91E-04	C5_Up
Mitochondrial Part	0.4767	2.3764	2.61E-04	C5_Up
Mitochondrial Lumen	0.5777	2.3256	4.10E-04	C5_Up
Organellar Ribosome	0.6629	2.3068	3.28E-04	C5_Up
Defence response	-0.4498	-2.6034	0	C5_Down
Cellular Defence response	-0.5215	-2.2272	0.007386	C5_Down
Hematopoetin Interferon Class D200 Domain Cytokine Receptor Activity	-0.6154	-2.1911	0.009138	C5_Down
Immune response	-0.3846	-2.1803	0.008479	C5_Down
Inflammatory response	-0.4211	-2.0983	0.018092	C5_Down

[a]Expression score

[b]Normalized expression score

[c]Gene list refers to the database and input gene list. C2_up represents pathways over expressed by HEPS CMCs from the C2 database, C2_down represents pathways under expressed by HEPS CMCs from the C2 database, C5_up represents pathways over expressed by HEPS CMCs from the C5 database, and C5_down represents pathways under expressed by HEPS CMCs from the C5 database.

Twenty-four gene sets from the MsigDB C2 (curated gene set database) were found to be significantly over expressed (FDR p-value < 0.05) in CMCs from HEPS women (C2_up). Of the identified gene sets, 10 gene sets have been verified by other databases and the 5 most significantly enriched are present in Table 5. Ninety-nine gene sets from this database were found to be significantly under expressed by CMCs from HEPS women (C2_down), and 25 of the gene sets have been verified by other databases. The 5 most significantly enriched gene sets are also presented in Table 5.

Nineteen biological themes from the MsigDB C5 (GO annotation database) were found to be significantly over expressed in CMCs from HEPS women (C5_up). The 5 most significantly enriched genes are presented in Table 5. 28 biological themes were identified as being significantly under expressed in CMCs from HEPS women (C5_down), the top 5 of which are also presented in Table 5.

4.2.5.5. Pathways related to the immune response

Several pathways related to the immune response were identified by GSEA as significantly under expressed in CMCs from HEPS donors, including inflammation, TLRs, and Hematopoetin Interferon Class D200 Domain Cytokine Receptor Activity. Due to the important role TLRs play in inflammation, this pathway will be the focus of follow up analysis and discussion.

CMCs from HEPS donors show a decrease in the expression of TLR6 (which recognizes cell wall components from *Mycoplasma* spp. [84]) by 1.55-fold ($p = 0.02$) and TLR10 (1.33-fold decreased in HEPS CMCs, $p = 0.04$). MyD88, which is the next step in the TLR signalling pathway for these particular TLRs, trended toward being significantly decreased 1.32-fold with a p-value of 0.056. Most importantly, however,

was that NFKB2, a subunit of NFκB, was significantly decreased 1.48-fold ($p = 0.04$). In addition to TLR pathway genes, other important immune response genes were significantly differentially expressed by HEPS CMCs. Interleukin 1 receptor type I (IL1R1), the receptor for a potent inflammatory cytokine and also plays a role in inducing NFκB transcription [205], was identified as the most significantly differentially expressed gene in HEPS CMCs(1.86-fold decreased in HEPS CMCs, $P = 2.2 \times 10^{-4}$). CD86, an important secondary signal for T cell activation [110], was also found to be significantly decreased by HEPS CMCs (1.54-fold decreased, $p = 0.02$).

4.2.6. Pathway clustering

To determine if the differentially expressed pathways were able to distinguish HEPS and NN CMCs, genes involved in oxidative phosphorylation, the immune response and the MAPK pathway were subjected to unsupervised hierarchical clustering (Figure 19). HEPS and NN samples did not cluster independently when exclusively clustering enriched pathway gene sets.

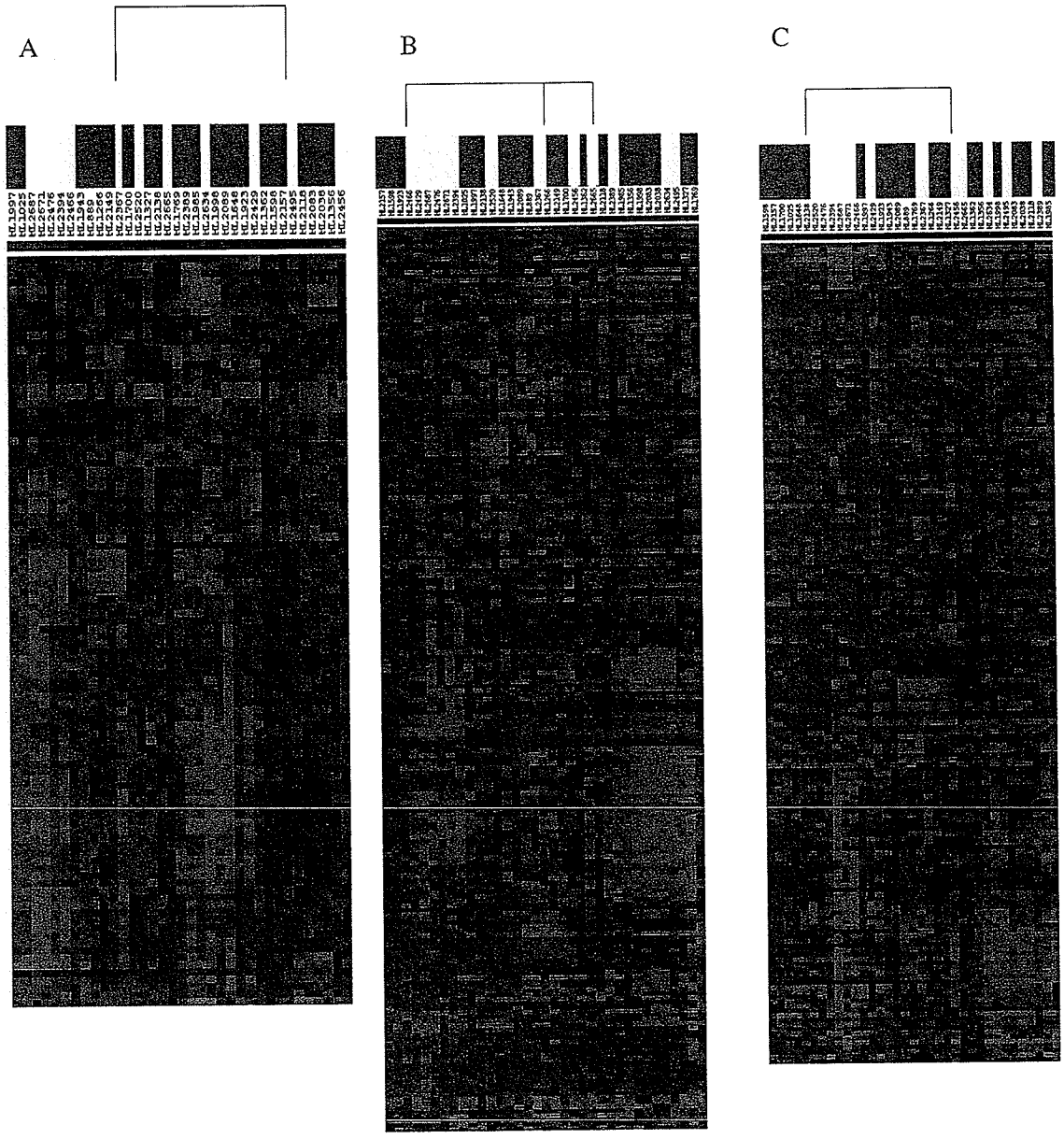


Figure 19: Unsupervised hierarchical clustering of genes in selected pathways. (A) Oxidative phosphorylation (B) the immune response and (C) the MAPK signalling pathway. Yellow boxes represent NN donors, blue boxes represent HEPS. HEPS and NN do not cluster independently with unsupervised hierarchical clustering. Red indicates a relative increase and green represents a relative decrease in gene expression intensity.

4.3. Objective 3: Stimulations

To determine if there were unique cell signalling pathways present upon HIV-1 antigenic challenge in HIV-1 resistant women, CMCs were to be stimulated with a pool of clade A and ancient consensus gag peptide sequences in Kenya. In order to establish ideal stimulation conditions for CMCs, an additional 25 Winnipeg CMC donors were recruited for this study. PHA was included as a positive control, and a CEF peptide pool was used in place of the HIV-1 peptides, as there were no CMCs from HIV-1 infected donors available in Winnipeg. However, the gag peptide pool was tested on PBMCs donated from HIV-1 infected donors from the Cadham Provincial Laboratory in Winnipeg. qRT-PCR was used to measure IFNγ expression as a marker of immune activation.

4.3.1. CEF and PHA stimulations on CMCs and PBMCs

PBMCs from HIV-1 uninfected donors showed robust responses to PHA and CEF by IFNγ levels measured by qRT-PCR. PBMC stimulation trials were performed on 3 donors, and a representative sample is shown in Figure 20a. PBMCs were stimulated at 37°C for 3, 6, and 16 hours and IFNγ measured in equal amounts of RNA. PBMCs showed a peak of IFNγ output at 6h for CEF and overnight for PHA. DMSO was used as a negative control.

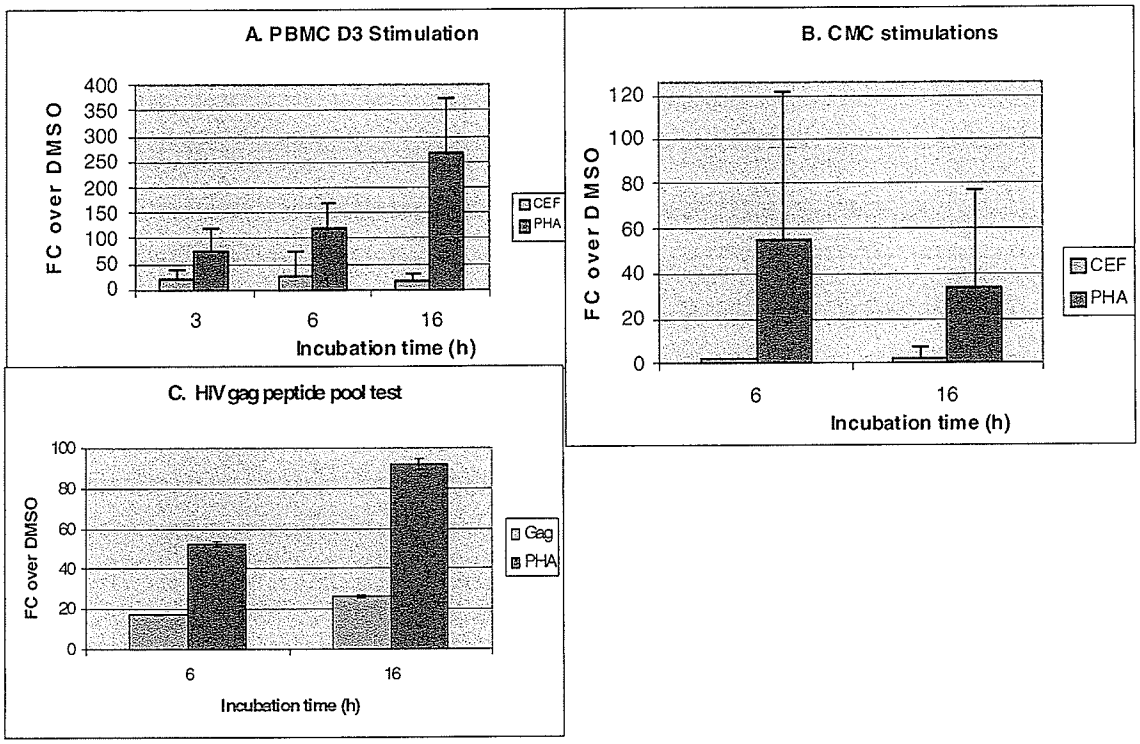


Figure 20: CMC stimulation experiments. IFNγ outputs expressed as fold-change (FC) over media treatment alone. (A) A representative trial of CEF and PHA stimulations of HIV-1-negative PBMCs. qRT-PCR samples were run in duplicate and the error bars represent the standard error of the IFNγ readings. (B) Averaged IFNγ FC of CMC stimulation trials. Error bars represent the standard deviation of the 6 trials. (C) Representative trial of stimulations of PBMCs from HIV-1 infected with a gag peptide pool. qRT-PCR samples were run in duplicate and error bars represent the standard error of the IFNγ readings.

Performing stimulations on CMCs was challenging, and there was high variability in IFNγ outputs between subjects. As such there was no representative CMC stimulation trial and Figure 20b represents an average of IFNγ fold changes for 6 subjects. Due to limited cell availability, only the 6 and 16-hour time points were tested. CMCs tended to have the opposite trend to PBMCs, showing the highest IFNγ output for CEF after an overnight incubation while PHA peaked after 6 hours. The high variability between experiments suggests there was no statistical difference between the incubation times (Figure 20b). These data suggested that the best incubation time for antigen recall responses in CMCs, as tested by stimulation with CEF, was 16 hours as measured by IFNγ qRT-PCR, although 6 hours still showed an appreciable IFNγ output.

4.3.2. PBMC stimulations with HIV-1 gag peptide pool

Two HIV-1 infected blood donors were recruited to test how well the gag peptide pool tested in *ex vivo* samples. A representative trial is shown in Figure 20c. PBMCs were again stimulated for 6 and 16 hours. The greatest fold changes in IFNγ expression were found in the 16-hour stimulated samples for both the gag peptide pool and PHA. The data from these trials and the CMC stimulations indicated the best incubation time to maximize HIV-1 gag-induced effects on gene expression on cells was 16 hours.

4.3.3. HIV-1 gag stimulation trials on Kenyan CMCs

CMCs were isolated and stimulated in Kenya in May and June 2008 and were successfully shipped to Winnipeg in TRIzol®. In order to perform baseline and

stimulations on the same patient, a minimum of 1.2 million CMCs were required from each CMC donor. HIV-1 susceptible women, represented by women newly enrolled in the Pumwani cohort (NN), as well as HIV-1 infected women (HIV+) were included in this study as control groups. CMC counts were significantly lower in Kenya than in Winnipeg and very few donors had 1.2 million CMCs; consequently, not all stimulation conditions were included for all donors. Furthermore, RNA yields from samples isolated in the resurvey were much lower than from CMCs in Winnipeg. None of the stimulated samples reached the minimum RNA concentration (150 ng in 11 ul, or 14 ng/ul) required to be run on the Illumina HumanRef-8 v2 Expression BeadChips and therefore this arm of the study was abandoned.

5. Discussion

5.1. Objective 1: Comparing gene expression patterns of CMCs and PBMCs

These experiments were conducted to address the first objective of this study, which was to show the unique immunological genetic environment at the level of the FGT as compared to the circulatory immune system. The data obtained from this part of the study will be published in combination with flow cytometry data phenotyping cells present in the FGT (Kaefer and Horton, submitted [196]).

5.1.1. CMC isolations

Through the course of this study, several months were spent trying to improve upon published protocols to maximize CMC yields from cervical scrapings. The protocol as described by Iqbal *et al* [132] was used as a starting point to maximize CMC yields from cervical scrapings. The final protocol led to significantly higher CMC counts.

The primary challenge in isolating and purifying CMCs from endocervical smears was cervical mucus. Cervical mucus is a semi-solid liquid present in the female genital tract and has several important biological roles [206]. Its primary function is to prevent pathogens from travelling into the uterus while allowing sperm to travel up the FGT for fertilization [206]. Cervical mucus is a complex mixture of a variety of proteins, the majority of which are heavily-glycosylated mucins [206]. Although there are a variety of products available to eliminate mucus, such as guanifesin and n-acetyl cysteine, most do not have direct mucolytic activity [207]. Although treating the cytobrushes with TrypLE (a trypsin cocktail commonly used in the lab to detach adherent cells) led to the highest

CMC yields of the isolation protocols attempted, there were no data available in the literature to determine the possible effects of TrypLE on gene expression in the cells. Incubating the samples at 37°C in media alone did not lead to significant increases in CMC yields, and again could possibly change gene expression patterns in the cells. Alternative methods of CMC separation, namely using a 40 um cell filter or nylon wool resulted in CMC yields lower than those obtained by Ficol. Mechanical disruption of the mucus by repeated pipetting followed by layering onto Ficol was found to significantly increase CMC yields and decrease the amount of contaminating epithelial cells present while minimizing known changes in cellular gene expression. Consequently, this protocol was adapted for all subsequent studies presented in this thesis. It might be valuable, however, to revisit these protocols in the future, especially the TrypLE modification. This protocol may be of use in downstream protocols where high cell counts are needed and that are less sensitive to changes in gene expression that may be induced by this treatment.

5.1.2. CMC vs. PBMC microarray results

Based on the gene expression data collected, the female genital tract and circulatory immune system reveal quite divergent gene expression patterns in each compartment. Over 5300 genes were differentially expressed between matched CMCs and PBMCs, and CMCs were shown to cluster independently of PBMCs by unsupervised hierarchical clustering analyses. While much of this difference in gene expression can be explained by the differences in CD14+, CD3+, and lineage negative HLA-DR+ cell makeup present in the FGT (Kaefer and Horton, submitted [196]), several important immunological and

cellular process pathways were identified by the detailed gene expression analyses that would not have been discovered by cell phenotyping alone. Three pathways identified by DAVID were selected for the follow-up qRT-PCR confirmation study. A discussion of these pathways and their relevance to understanding HIV-1 transmission across the endocervix will be discussed in the following sections. It should be emphasized that the women donating CMCs for this study were otherwise healthy with no reported genital tract infections, so the alteration of the pathways is occurring at baseline immune levels and under normal physiological conditions.

5.1.2.1. Cytokine-Cytokine receptor interaction

There was an increase in the expression level of several cytokines in the FGT when compared to the periphery according to the data obtained in this study (section 4.1.7.1). Over expressed cytokines include 29 well-described pro-inflammatory mediators, such as MIP-1α, MIP-1β and TNF [208,209], and two inhibitors of inflammation: IL1R2 and IL10RB [210,211]. Thus, there was not a bias toward up regulation of pro-inflammatory genes in CMCs over PBMCs (22 cytokines down-regulated by CMCs: 20 pro-inflammatory and two anti-inflammatory), but an overall global increase in cytokine gene expression. This global increase in cytokine expression suggests increased cytokine signalling and possibly a more active immune response within the FGT. Several factors regulate cytokine expression. Both apoptosis [212] and TLR signalling [76,92] are documented regulators of cytokine expression and were also found to be differentially regulated by CMCs in our study, suggesting a complex but relevant interplay between these two distinct pathways.

OSM was the cytokine gene selected for follow up confirmation from the arrays because it was 18-fold over expressed by CMCs (Figure 11). It belongs to the hematopoietin cytokine family and is closely related to IL6, a well-characterized pro-inflammatory cytokine [198]. OSM is secreted by activated monocytes, macrophages and T cells and triggers the further recruitment of immune cells to the site of infection [198]. OSM works through binding one of two receptor complexes. OSM first binds surface protein gp130, and the OSM-gp130 complex then either associates with the leukocyte inhibitory factor (LIF) receptor and is referred to as the Type I receptor complex, or associates with the OSM receptor and is referred to as the Type II receptor complex [213]. Activation of receptor complexes leads to the transcription of CCL1, CCL7 and CCL8, important chemokines involved in the chemoattraction of monocytes and T cells to the site of an infection [214]. Therefore, a high expression level of this cytokine suggests a basal level of inflammation in the FGT as compared to the circulatory system. OSM has also been shown to act synergistically with the TLR signalling pathway to promote the induction of other inflammatory cytokines [215]. Therefore, the over expression of OSM, along with other inflammatory cytokines and the MyD88-dependent TLR signalling pathway is logical and suggests these factors are working together to promote an inflammatory state in the FGT. Interestingly, OSM expression was also found to be significantly decreased in women with IRF-1 phenotypes with decreased susceptibility to HIV-1 (Aida Sivro, personal communication), suggesting this cytokine is a potential target to prevent HIV-1 infection. Unfortunately, the expression of OSM was not confirmed by either qRT-PCR or Western Blot. Further discussion of expression confirmation studies will follow in sections 5.1.3 and 5.1.4

5.1.2.2. TLR signalling

TLRs are an important part of innate immune surveillance [76], and increased innate immune signalling in epithelial cells of the FGT has been thoroughly documented [216-218]. The data presented in this thesis shows an increased expression of TLRs 1, 2, 4, 5, 6 and 8 in CMCs as well as their downstream signalling products, complete with the induction of inflammatory cytokines in the absence of active infection. In general, genes involved in the MyD88-dependent pathway arm of the TLR signalling pathway, which leads to the induction of NFκB and cytokine transcription, were over expressed in CMCs.

Although TLR signalling has been well described in the literature, very little of this work has been done on CMCs. Cervical epithelial cells are the first line of defence in the endocervix and express a breadth of TLRs. Characterizing TLRs expressed by CMC has only recently been attempted (Meyers *et al.*, Lester *et al.*, in preparation). This is important knowledge when trying to understand the immune response at the cervical mucosal level.

The most striking aspect of the TLR expression pattern in CMCs presented in this thesis is the apparent biased increase in TLRs recognising bacterial pathogenic patterns: TLRs 1, 2, 4, 5 and 6, and a down regulation of ssRNA-binding TLR7 [92]. As the endocervix has occasional contact with normal vaginal flora during the menstrual cycle, and mucosal surfaces are sites of immunotolerance, one could assume leukocytes isolated from this site are less responsive to bacterial PAMPs. Furthermore, the lack of ssRNA recognition could make the FGT more susceptible to infection by RNA viruses. On the other hand, the decreased expression of TLR7 in the FGT might be protective in the case of HIV-1 infection, as there may be less immune recognition of the virus and, therefore,

no increase in inflammatory cytokine expression upon HIV-1 exposure. Less cytokine production and inflammation may be beneficial in reducing immune cell recruitment, leading to fewer HIV-1 target cells in the FGT being infected. The over expressed TLRs signal via the MyD88-dependent pathway to activate NFκB transcription, and consequently MyD88 was chosen for confirmation by qRT-PCR. However, due to issues with poor RNA quality, this gene was not confirmed in this study. NFκB, which was selected, but remains unconfirmed by preliminary Western Blot confirmation studies, goes on to activate inflammatory cytokines. Along with an increase in MyD88 and NFKB2 gene expression, there was an increase in chemotactic molecules (such as MIP-1α/β) but no change in interferon expression according to the gene expression arrays.

One possible explanation for this increase in TLR signalling is the presence of natural flora. The lower female reproductive tract (vagina and ectocervix) is a non-sterile site and has diverse microbial populations, with the dominant species being gram positive *Lactobacillus* spp., while the uterus is a sterile site [85,116]. The endocervix represents a transitional environment, however, and the presence of natural flora at this site is potentially linked to the menstrual cycle [116]. Instead of decreasing TLR expression in CMCs to make the immune system of the FGT less responsive to normal flora, perhaps the close proximity of vaginal natural flora actually leads to increased TLR signalling. If the FGT is indeed tolerogenic, however, there must be another tolerance mechanism present. One such tolerance mechanism may be the presence of tolerogenic DCs. Although results from flow cytometry indicate there is an increase in DCs present in the FGT over numbers detected in blood, there is no increase in DC activation markers such as CD86, possibly suggesting non-activated DCs (Kaefer and Horton, submitted [196]).

Antigen presentation by non-activated DCs have been shown to induce peripheral tolerance at gut mucosal surfaces, and it is also possible that the same process is occurring in the FGT [99]. However, the data presented in this thesis also show an increase in cytokine expression, which would be inconsistent with a tolerant state. Further studies into the activation state of DCs need to be performed before making any conclusions regarding increased TLR signalling and how it relates to tolerogenic DCs in the FGT.

5.1.2.3. Apoptosis

In CMCs, although apoptosis-inducing cytokines and upstream signalling molecules of both the intrinsic and extrinsic apoptotic pathways were over expressed, there was concurrent expression of apoptosis inhibitors and a decrease in the downstream pathway products, suggesting an anti-apoptotic state. As the induction of apoptosis would lead to a decrease in cytokine expression [212], this downstream suppression of apoptosis along is consistent with high cytokine expression levels. Figure 12 shows the genes differentially expressed by CMCs as described in this thesis and their relative position in the apoptosis pathways. Apoptosis plays an important role in immune regulation [212]. There are three apoptotic pathways – the intrinsic mitochondrial pathway, the extrinsic pathway and caspase-independent cell death [219]. The intrinsic and extrinsic pathways have unique signals that trigger their activation, but both lead to the activation of the caspase cascade and consequently cell death [219]. The caspase-independent pathway is triggered by upstream apoptotic stimuli similar to the other pathways, but does not activate the caspase cascade [220]. A brief outline and discussion of the regulation of apoptosis and how it relates to the immune system follows.

The extrinsic apoptosis pathway is mediated by environmental signals TNFα, TNF-related apoptosis inducing factor (TRAIL), and Fas ligand (CD95L), which were found to be over expressed by CMCs [219]. Activation-induced cell death (AICD) is a process whereby T cells that have become activated without appropriate co-stimulation are eliminated via the extrinsic apoptosis pathway, typically once an infection has been cleared by the immune system [219]. TNFα, TRAIL and other apoptotic signalling molecules bind to and activate cell surface receptors, which then bind FADD via their intracellular death domains [221]. FADD then binds procaspase 8 to form the Death-Inducing Signalling Complex (DISC), which then activates the caspase cascade [222]. The caspase cascade ends with the cleavage of procaspase 3 into caspase 3, which in turn disinhibits DNA fragmentation factors and triggers other final steps in apoptosis [223]. Gene expression analysis showed significantly lower FADD expression in CMCs, suggesting overall inhibition of the extrinsic apoptotic pathway.

CFLAR is an important regulator of the extrinsic apoptotic pathway, and was significantly over expressed by CMCs, as would be expected by the decrease in FADD expression [224]. CFLAR can block activation of caspase 8 by competitively binding to DISC at high expression levels, stopping the caspase cascade before apoptosis is induced [225]. This increase in CFLAR expression suggests DISC formation is decreased in CMCs. There are three isoforms of CFLAR described in the literature, designated as Cellular FLICE-(caspase 8)-like inhibitory protein, long ($cFLIP_L$), $cFLIP_S$ (short) and $cFLIP_R$ (Raji) [225]. All three isoforms contain two death domains that bind DISC. $cFLIP_L$ is the longest CFLAR isoform and has a catalytically inactive caspase-like domain at its carboxy terminus [225]. When expressed at physiological levels, $cFLIP_L$

inhibits the activation of procaspase 8 by DISC by competitive binding and thus is an inhibitor of apoptosis [226,227]. However, very high concentrations of $cFLIP_L$ have been shown to form heterodimers with procaspase 8, enhancing its processing to caspase 8, and consequently $cFLIP_L$ is pro-apoptotic in this case [224,227]. Both $cFLIP_S$ and $cFLIP_R$ are mRNA splice variants of $cFLIP_L$ and are strictly involved in inhibition of apoptosis by competitively binding DISC [228]. $cFLIP_S$ has been shown to rescue T cells from AICD, demonstrating the potency of this protein [219]. Further studies into the CFLAR isoforms differentially expressed by CMCs must be done to confirm there is an inhibition of apoptosis. It should also be noted that the expression of CFLAR is over expressed by NFκB following CD3 activation with CD28 costimulation and by Fas ligand stimulation [226,227,229,230]. There was a significant over expression of the NFKB2 gene (a subunit of NFκB) observed in CMCs, which was consistent with higher expression levels of CFLAR. This increased expression of CFLAR, combined with a decrease in the expression of DNA fragmentation factors and no significant difference in the expression of caspases, suggests a suppression of the extrinsic apoptotic pathway in CMCs as compared to PBMCs. The prevention of the extrinsic apoptotic pathway in CMCs could be a mechanism exploited by CMCs to persist in the FGT to respond to pathogens in the presence of increased cytokine signalling. CFLAR was chosen for qRT-PCR and Western Blot confirmation, and although it was not confirmed by qRT-PCR, increased expression of $cFLIP_S$ was confirmed in CMCs by Western Blot analysis. There was no significant increase in $cFLIP_L$ expression in CMCs.

The intrinsic apoptosis pathway also plays an important role in mediating the immune system by triggering activated cell-autonomous death (ACAD). ACAD

eliminates T cells and is different from AICD as it is triggered by the absence of pro-survival cytokines (such as IL2 and IL13), as opposed to being induced by activation signals [219]. Mitochondria regulate the intrinsic apoptotic pathway [231]. This pathway is activated by cell stress such as hypoxia, by the loss of cell-survival signals, or by the dysregulation of the cell cycle as signalled by p53, and leads to the induction of the caspase cascade [231]. These signals lead to the transcription of mediators of apoptosis (Figure 12). The B cell leukemia (Bcl) family of proteins are the primary mediators of the intrinsic apoptotic pathway [231]. p53 up-regulated modulator of apoptosis (PUMA), which was over expressed by CMCs, and *Noxa* (Latin for damage, not significantly differentially expressed by CMCs) have p53-binding domains in their promoter regions, and an increase in cytoplasmic p53 triggers their transcription [232,233]. Upon translation, these proteins localize to mitochondria and induce apoptosis by antagonizing the activity of anti-apoptotic proteins Bcl2, myeloid cell leukemia sequence 1 (MCL1), and Bcl-like 1 ($BclX_L$) [231]. Bcl2 (under expressed by CMCs) and $BclX_L$ inhibit the induction of apoptosis by preventing the permeabilization of the mitochondrial membrane by Bcl2-associated X protein (BAX) and Bcl2-antagonist/killer 1 (BAK) as well as by sequestration of PUMA, *Noxa* and related proteins [234]. Both BAX and BAK were over expressed by CMCs. When PUMA and *Noxa* bind and sequester Bcl2 and $BclX_L$, BAX and BAK become active, form homo-oligomers and increase mitochondrial membrane permeability [234]. MCL1, which is significantly over expressed by CMCs, binds and inhibits BAX and BAK [235]. Increases in mitochondrial membrane permeability by BAX and BAK (as well as other) proteins allows for the release of the pro-apoptotic intermediary signalling molecules cytochrome c and Direct

IAP-binding protein with low pI (DIABLO) [236]. Cytochrome c (significantly under expressed by CMCs) binds apoptotic peptidase activating factor-1 (APAF-1) and forms the apoptosome, which goes on to activate procaspase 9 and the caspase cascade [237]. DIABLO inhibits the activity of inhibitors of apoptosis (IAPs) such as XIAP, which, under normal circumstances, directly block the activation of the caspase cascade [236]. XIAP was also found to be significantly over expressed by CMCs.

As seen with the extrinsic apoptotic pathway, there is an increase in the molecules that induce apoptosis. However, there is again an increase in downstream inhibitors of apoptosis and a decrease in downstream apoptosis-inducing genes, suggesting a shift to an anti-apoptotic state. Why CMCs would suppress the intrinsic apoptotic pathway is not clear, but perhaps it is to ensure leukocytes persist in the genital tract to respond to an infection.

The final apoptosis pathway is caspase-independent cell death (CICD). This pathway occurs when upstream pro-apoptotic signalling and cell death occurs without the activation of the caspase cascade or chromatin condensation [220]. Cell culture studies with caspase inhibitors showed marginal induction of apoptosis in the presence of a caspase inhibitor and in mice lacking caspase 9 [238,239]. There does not appear to be evidence in the literature for CICD in humans *in vivo*, although cardiomyocytes and certain neuronal cells express low level of APAF1 and thus could be candidates for CICD [220]. It has been suggested T cells undergoing CD95-independent cell death are also undergoing CICD, but more research must be done to confirm this process is occurring [240]. If CICD is indeed occurring in CMCs, this would be the first published *in vivo* occurrence of this process.

Differential expression of apoptotic pathways in CMCs suggests two possible scenarios. The first is that CMCs are indeed undergoing apoptosis and that CFLAR and XIAP are overwhelmed by the apoptosis-inducing molecules. The other scenario is that, despite the increase in TNF, TRAIL and other apoptosis marker expression, CFLAR and XIAP are able to suppress this signalling and prevent apoptosis under normal physiological conditions. These cells then complete the apoptosis pathway when apoptosis signals overwhelm CFLAR and XIAP when the cell is stressed. Either scenario must be confirmed with functional studies, but decreased expression of apoptosis end products such as DFFB does strongly suggest an anti-apoptotic state. Inhibition of apoptosis would then allow CMCs to persist in the FGT to respond in the event of an infection. As previously discussed, apoptosis is important to quench inflammation; thus, inhibiting this process could lead to increased inflammation. Inhibition of apoptosis by CMCs can, in part, explain some of the increased pro-inflammatory cytokine expression, mediated by the interplay between CFLAR, NFκB, and increased expression of inflammatory cytokines. However, the data presented in this thesis also shows there is a delicate balance between pro- and anti-apoptotic factors present in the FGT, the balance of which could be tipped toward a pro-apoptotic, anti-inflammatory state. Although the induction of apoptosis is risky, the safe removal of inappropriately activated cells, triggered by the presence of natural flora, might be one consideration when designing methods to prevent transmission of HIV-1.

5.1.2.4. Additional differentially expressed pathways

Although only three pathways were selected for analysis and discussion, there were many other pathways of immunobiological interest elucidated by this array analysis.

T and B cell receptor signalling were found to be altered in CMCs; however, this may be a reflection of a decrease in these cell types in CMCs as determined by flow cytometry (Kaefer and Horton, submitted [196]), and has been previously described [195]. There was also an apparent decrease in NK cell-mediated cytotoxicity; however this may actually be more of a reflection in the decreased CD8+ cytotoxic lymphocyte population in CMCs as there were no differences in NK cell populations between CMCs and PBMCs (Kaefer and Horton, submitted [196]). Several cellular metabolic pathways were differentially expressed, including oxidative phosphorylation, galactose and aminosugars metabolism, the citrate cycle, and fructose and mannose metabolism (Table 3). Alternative regulation of cellular metabolic pathways have been implicated in different activation states of cells of the immune system [241,242], and consequently would be of interest in follow up studies. The interplay between immune activation and cellular metabolism are currently under investigation (E.M. Songok, personal communication). The cell cycle was identified as a pathway differentially expressed by CMCs, as well as the synthesis of macromolecules, including purines and pyrimidines, aminoacyl-tRNAs, the ribosome, and DNA and RNA polymerases. There were a variety cellular pathways and factors differentially expressed by CMCs and demonstrate the uniqueness of gene expression in the FGT. Further analysis with respect to the interplay between the identified pathways and their contribution to the immunobiology of the FGT is crucial to understanding HIV-1 transmission across the FGT.

5.1.3. qRT-PCR confirmation

Despite selecting genes with highly significant p-values and fold-changes, qRT-PCR was unable to confirm the gene expression array data. All 5 samples tested were originally included in the initial gene expression analysis, and it is unclear how such a change in expression is possible. All samples went from being significantly over expressed in CMCs to being either not significantly differentially expressed or significantly under expressed in CMCs as compared with PBMCs. Although it is possible this poor correlation is due to problems with the initial gene expression analysis, other groups using this analysis method have shown good correlation between Illumina gene expression arrays and qRT-PCR data [243] and therefore other factors may explain this poor correlation of gene expression between the two methods.

One possible explanation of this contradictory data is the fact that samples hybridized to the BeadChips underwent linear transcriptional amplification. To address this concern, the Sigma TransPlex® Complete Whole Transcriptome Amplification Kit was tested to determine if mRNA amplification would increase the quantity and quality of starting material for gene expression verification. This protocol was suggested to linearly amplify mRNA from RNA preparations and the kit was tested on standard brain RNA in this study. However, after testing two genes, this claim could not be verified. The linear amplification decreased the amount of CFLAR detected as compared to unamplified controls, while 18S amplified as expected.

Another possible reason for the poor correlation between the qRT-PCR and microarray data was RNA quality. The RNA samples had spent 14 months at -80°C before cDNA was made for qRT-PCR gene expression confirmation. Consequently, it is

possible that the qRT-PCR data was not completely reliable as many of the samples had gene expression levels close to, or below the limit of detection of the LightCycler. Upon re-quantification of the RNA used in qRT-PCR, the CMC samples appeared to have undergone degradation, with one sample showing a 95% decrease in intact RNA concentration, while the PBMC RNA stayed relatively intact, thus casting further doubt on the reliability of the qRT-PCR data. There were no studies in the literature suggesting CMC RNA is more susceptible to RNA degradation than other tissue types. These qRT-PCR reactions should be repeated in the near future on freshly obtained matched CMC and PBMC samples to confirm the observed differences in gene expression. These samples should also be stored with RNase inhibitors to maintain sample quality. Further protein expression studies between CMCs and PBMCs would be of additional benefit to confirm these results, as proteins are the ultimate executors of the described cellular functions.

5.1.4. Protein confirmations of gene expression data

Although the CMC gene expression data was not confirmed by qRT-PCR, CMC phenotyping performed by Rachel Horton showed a significant decrease in CD3+ cells and an increase in CD14+, which corresponded well with the gene expression data (Kaefer and Horton, submitted [196]), lending strength to the array data presented in this thesis. Cell phenotyping also showed a trend toward a significant decrease in CD19+ cells in CMCs, and CD19 was found to be significantly under expressed by CMCs according to the gene expression array data (Kaefer and Horton, submitted [196]). Thus,

it was determined that protein confirmation using fresh samples of the important genes identified in this thesis may prove to be more successful that qRT-PCR alone.

To that end, Western blots on cFLIP (the protein product of CFLAR), OSM and p52 (the protein product of NFKB2) were done in an attempt to confirm the Illumina gene expression data. Unfortunately, the results from the gene expression arrays remain unconfirmed by the Western blot analyses. OSM was the first protein tested; however, the antibody used did not detect an isoform of the correct size. It is possible that the active form of OSM is below the detection limit of this protocol. Many groups have measured OSM by Western blot using this antibody and others [244-246], but the literature does not show any data for OSM expression levels in unstimulated PBMCs. Perhaps a more sensitive protein detection method would be more appropriate for confirming expression of this protein, such as an enzyme-linked immunosorbent assay (ELISA), an immunoprecipitation (IP) followed by western blotting, or a mass spectroscopy protocol.

The antibody selected for confirmation of p52 was also very non-specific, showing high levels of background binding. Indeed, only one CMC sample was shown to express the expected 52 kDa isoform, and the remaining samples were negative. Although this antibody has been used by another group for western blotting [247], they measured p52 expression after cytokine stimulation. A more sensitive assay would be recommended for determining levels of this protein for future studies such as those recommended for OSM.

The antibody used to measure levels of cFLIP was more successful than the antibodies used to measure OSM or p52. There was a trend to higher expression of both

cFLIP isoforms in CMCs as compared to PBMCs, which is what was expected from the gene expression arrays; however, the results failed to reach statistical significance. The sample sizes in each group were rather small (4 CMCs and 5 PBMCs) and increasing the number of samples tested would be beneficial to increasing the statistical power of this experiment. Some PBMC samples actually expressed $cFLIP_S$ at detectable levels; however, another confirmation method (such as an ELISA or an IP) should be used before concluding this protein isoform was indeed significantly over expressed by CMCs.

5.1.5. Summary of the CMC vs. PBMC gene expression studies

There was an apparent increase in inflammatory processes in the genital tract of healthy women compared to their circulating immune system based on this gene expression data (Figures 10 – 12). The TLR signalling pathway, complete with the increase in expression of the NFKB2 gene, has been shown to be significantly over expressed in CMCs when compared with PBMCs. NFκB is an important transcription factor in the immune response, and triggers the activation of pro-inflammatory cytokines. Several NFκB target cytokines, such as MIP-1α and MIP-1β, were also increased at the mRNA level in CMCs. OSM was another cytokine that was significantly over expressed by CMCs, and plays a role in the regulation of the TLR signalling pathway.

NFκB plays a role in the regulation of CFLAR, an important mediator of the apoptotic pathway. Both CFLAR and the apoptosis pathway in general were differentially expressed by CMCs. The possible suppression of apoptosis in CMCs, as suggested by the data in this thesis, has also been linked to increased inflammation as

apoptosis triggers rapid engulfment of neutrophils by macrophages without inducing macrophage cytokine expression [248,249].

Apoptosis is known to eliminate activated T cells by AICD and ACAD, thereby decreasing cytokine levels [212] and interferes with the induction of cytokine transcription by NFκB [250]. These pathways together suggest a chronic level of inflammation present in the genital tract of healthy female donors. Figure 21 demonstrates the relationship between the pathways.

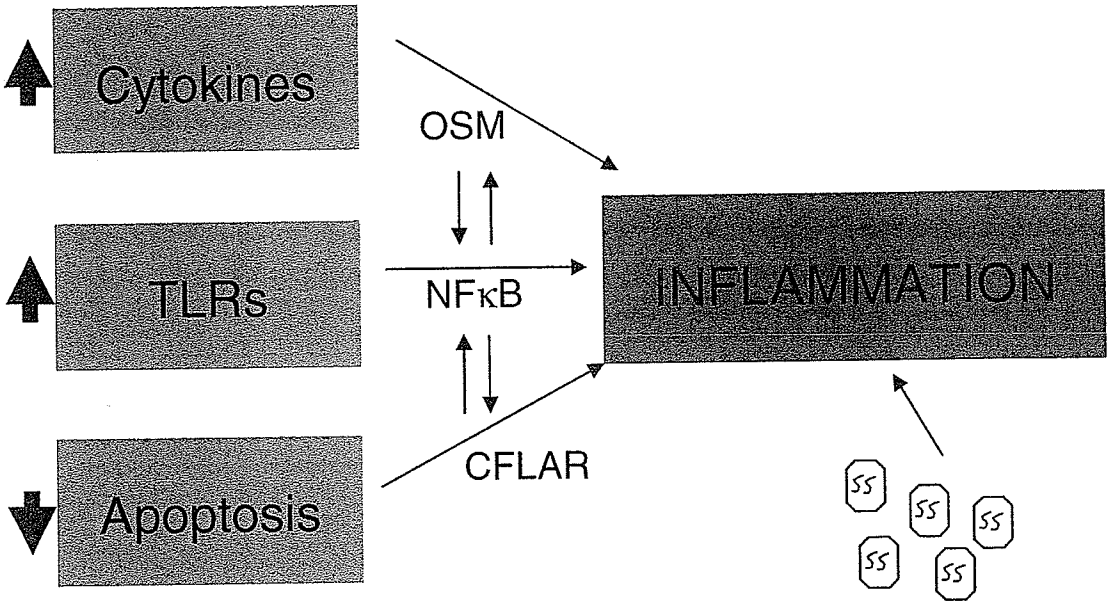


Figure 21: Relationship between three pathways differentially expressed by CMCs and their relation to inflammation. OSM is an inflammatory cytokine that works synergistically with the TLR signalling pathway to induce the transcription of inflammatory cytokines by NFκB. NFκB regulates the expression of CFLAR, an inhibitor of apoptosis. Apoptosis quenches inflammation, and the inhibition of apoptosis, combined with increased TLR signalling and proinflammatory cytokine expression suggests an increase in inflammation in the FGT as compared to the circulatory immune system. This increased inflammation is important to understand in the context of susceptibility to HIV-1 transmission.

The relevance of the identified pathways in terms of HIV-1 transmission is clear, as HIV-1 is known to have high rates of transmission at sites of inflammation [108]. The data gathered in this thesis suggest that reduction of inflammation in the FGT is one possible means of inhibiting HIV-1 transmission. One possible mechanism is via antagonisation of the activities of TLR2. This molecule primarily recognizes gram-positive bacteria (along with some gram negative, viral and fungal PAMPs), and there are no documented gram-positive bacterial agents causing STIs. Another possible mechanism to reduce the inflammatory state present in the FGT is to suppress the transcriptional activity of NFκB and thus decrease the expression of pro-inflammatory cytokines and increase apoptosis. Designing a microbicide to suppress the activity of potent inflammatory molecules, such as OSM, is another novel idea to reduce inflammation present in the FGT. Finally, the disinhibition of the apoptosis pathway by blocking the activity of CFLAR could also decrease inflammation in the FGT. However, any of these suggestions might interfere with the capability of women to fight other infections, as the activation of the immune response when challenged with other STIs is important to clearing infection.

In conclusion, CMCs and PBMCs differentially express cytokines, apoptotic factors, and the TLR signalling pathway, three factors that regulate inflammation and suggest a chronic inflammatory state present in the genital tract. This increased, chronic level of inflammation is important to understand in the context of susceptibility to STIs, especially viruses like HIV-1 that show increased transmission at inflammatory sites [108]. The immunology of the FGT is significantly different from that of the circulatory system, and understanding the identified differences is vital when designing future

microbicides or candidate mucosal vaccines. These results have been submitted to PLoS One for publication.

5.2. Objective 2: Gene expression studies in HEPS CMCs

To address the second and third objectives of this study, CMCs were isolated from HEPS sex workers to conduct genetic analysis for mucosal markers of resistance and immune response to HIV-1. To test the hypothesis that there is altered gene expression in the FGT of HEPS women that reduces their susceptibility to HIV-1 infection, CMCs were isolated in Kenya during the Spring 2008 resurvey and treated on site at the University of Nairobi Institute of Tropical and Infectious Diseases Building in Nairobi, Kenya.

The original study included indentifying changes in gene expression in HEPS, NN and HIV+ donor CMCs after antigen stimulation, as well as identifying differences in baseline gene expression. As sample collection continued, it became clear that not all stimulation conditions would be possible for samples collected from each study participant. In order to have the required 300,000 CMCs for all stimulation conditions and for baseline studies, 1.2 million CMCs were required from every donor. This cell count was well above the average CMC count obtained during sample collection, and consequently not all samples were used for all stimulation conditions.

Furthermore, once the samples were shipped to Winnipeg, it was clear the RNA yields obtained from local donors in Winnipeg did not accurately predict the RNA yields obtained from samples isolated in Kenya. In fact, none of the samples that were shipped from Nairobi to Winnipeg had sufficient RNA to be run on arrays. Although the real cause for this poor RNA quality was never determined, quality RNA has been isolated

from PBMCs from the same donors, and there were no previous reports of poor RNA quality from Kenyan CMC samples. However, this was the first time CMC RNA samples from this cohort were quantified on the BioAnalyzer, which gives a better idea of the quality of an RNA sample over other methods. These RNA quality results suggest a need to develop new protocols and new approaches that are more suitable to RNA extraction from CMCs isolated in Kenya. Finally, the stimulation experiments performed in the Spring 2008 resurvey were abandoned and 56 frozen samples in TRIzol® collected during the Spring 2007 resurvey were used for baseline gene expression studies. Of the new samples, approximately 20% of samples had sufficient RNA for array experiments (as compared to 40% of samples from Winnipeg donors for the CMC vs. PBMC study). Unfortunately, there was no cell count data available for the samples and consequently the identification of any novel differentially expressed genes or pathways must be followed up with future studies.

5.2.1. Cell counts

Despite the lack of genetic analyses obtained from the Spring 2008 resurvey, CMCs were counted from over 120 donors and cell counts were compared across HIV-1 status. There was no significant difference in CMC counts between NN and HEPS; therefore, differences in the total number of CMCs present in the FGT are not likely contributing to the HIV-1 resistance phenotype.

There are a few possible explanations as to why the cell yields were lower than anticipated in the Spring 2008 resurvey. All women included in the analyses either douche daily or post-coitally, primarily with soap and water, which could potentially

decrease the number of cells present. Furthermore, it is unknown what effects douching can have on total cellular RNA expression, which could explain why there were such low RNA yields obtained from the CMCs isolated from these women. These confounding factors can play a role in the results of this section of the thesis.

5.2.2. HEPS vs. NN microarray results

The gene expression patterns of CMCs isolated from HEPS and NN were not strikingly different from each other. Approximately 11% (1268 of 11,857) of the genes expressed by CMCs were significantly differentially expressed between HEPS and NN with a p-value < 0.05 (Figure 16). None of the genes passed multiple test corrections, but as differences in gene expression patterns between two human phenotypes can be small [191], analysis was continued with the provision that there could be false positives in the dataset. The genes and pathways identified as significantly differentially expressed need to be thoroughly verified before drawing conclusions about their role in HIV-1 resistance.

Clustering analysis of the gene expression patterns between HEPS and NN CMCs indicated that, at a whole genome expression level, HEPS and NN did not cluster independently. Nor did the women cluster independently when only examining the significantly differentially expressed genes, nor when examining the 10 "most resistant" sex workers (i.e. engaged in sex work for the longest amount of time) and comparing them to the newest sex workers. Therefore, evidence suggests there may not be a global "resistance phenotype" detectable in CMCs from HEPS women, and that if there was a difference in the gene expression pattern between HEPS and NN CMCs it was more subtle than can be determined at the whole genome level. Furthermore, there have been

many published correlates of protection against HIV-1 infection, reviewed in [125] and [130], suggesting there is not one global resistance phenotype, and that many cellular and/or genetic mechanisms may confer protection against HIV-1 infection.

Brass *et al.*, Zhou *et al.* and Konig *et al.* published data in 2008 showing host genes HIV-1 necessarily exploits during its life cycle; however, there is very little overlap between the identified gene sets between the groups, highlighting the complexity of this issue [19-21]. One conclusion that may be drawn from these studies is that examining individual genes in the context of HIV-1 replication is impractical and therefore gene pathways may provide more beneficial knowledge to understand host factors required for the HIV-1 life cycle. To that end, extensive pathway and biological theme analysis was done on the gene expression data from HEPS and NN CMCs to identify differentially expressed biological processes.

In another study, whole genome expression patterns of whole blood were also examined between HEPS and lower risk blood donors from an antenatal clinic, seven of whom were included in the study presented in this thesis [167]. There, HEPS women did cluster independently from the low risk controls, and had genes that were still significantly differentially expressed after multiple test corrections (E.M Songok, personal communication). The greater significance of the whole blood results may be partially due to the increased sample size analysed (43 women in each group). Also, the whole blood analysis was comparing HEPS to a different control group who are not engaged in sex work and therefore do not completely control for the gene expression pattern changes potentially caused by sex work. Ficol separation also removes neutrophils and other granulocytes from the isolated cell population [251]; thus, the

whole blood gene expression analyses would include these cells while the CMC population would not. Also, a variety of confounding variables can significantly affect CMC expression data but not the whole blood data, such as differences in douching practices, the menstrual cycle and the variability inherent to collecting cervical scrapings (as opposed to a blood draw, which is fairly standardized). To that end, increasing the sample size to help control for behavioural and menstrual variables will be critical in follow up studies. In fact, a much larger sample size per group may identify significant gene expression trends that were present in this group of women. Other ways to control for these variables is to sample the women at the same point in their menstrual cycle (i.e. right after menses) to control for the hormonal fluctuations associated with the menstrual cycle, and to request they refrain from douching for several days prior to sampling (to control for the effects of douching, and hopefully increase CMC counts).

5.2.3. Pathways differentially expressed by HEPS CMCs

Two independent pathway analysis programs were used to identify biological themes differentially expressed by HEPS CMCs. Two analysis programs were used in this study to maximise the number of pathway hits and to partially confirm pathway data using different analysis techniques. There were many pathways identified by more than one database, including the over expression of oxidative phosphorylation (and related themes, such as mitochondrial genes and the electron transport chain) and under expression of pathways related to the immune response (such as inflammation and TLR signalling). A variety of cancer-related pathways were identified by DAVID and GSEA databases as being significantly under expressed by HEPS CMCs (not featured in Table

5, as they were not among the 5 most differentially expressed pathways). Clustering analyses of the pathways again do not show HEPS and NN completely clustering independently, but as the pathways were significantly enriched in HEPS, they do warrant further study as mechanisms for resistance against HIV-1 infection and could be used to generate new hypotheses.

5.2.3.1. Oxidative phosphorylation

Oxidative phosphorylation and related pathways were identified as significantly over expressed by HEPS CMCs by DAVID and both GSEA databases, and the oxidative phosphorylation pathway is shown in Figure 17. This pathway was also recognized as significantly differentially expressed by HEPS women from the Pumwani cohort by gene expression analysis done on whole blood (E.M. Songok, personal communication). Thus, this pathway is likely to be of significant importance in understanding HIV-1 resistance in women from the Pumwani cohort. A brief discussion of this pathway and its potential role in reducing susceptibility to HIV-1 infection follows.

Oxidative phosphorylation is the last step in aerobic respiration [252]. The co-enzyme products of glycolysis and the citrate cycle, NADH and flavin adenine dinucleotide ($FADH_2$) enter the ETC and their electrons are passed down sequential protein complexes, leading to the synthesis of three and two adenosine triphosphate (ATP) molecules, respectively [253]. Oxygen acts as the final electron acceptor [253]. The first entry point for the ETC is NADH dehydrogenase, the largest protein complex of the ETC consisting of 45 protein subunits [254]. The genes encoding seven of these proteins (15%) were significantly over expressed by HEPS CMCs (Figure 17). Another protein complex, succinate-Q oxidoreductase, acts as an alternative ETC entry point for

succinate, an intermediary product of the citrate cycle and fatty acid metabolism (Figure 17) [200]. This protein complex is made up of 4 protein subunits [255], and subunit A was found to be significantly increased in HEPS CCs. A third entry point for the ETC is electron transfer flavoprotein-Q oxidoreductase (not significantly differentially expressed). This complex oxidizes other redox proteins, such as those generated by amino acid metabolism [256]. All three ETC entry points lead to the generation of ubiquinol and a proton gradient.

Ubiquinol is then shuttled to the cytochrome b_1-c complex (also known as complex III), where it is oxidised and leads to the eventual reduction of cytochrome c [257]. This complex has 11 subunits [201], of which only one, UQCRFS1, is differentially expressed by HEPS CMCs. The final protein complex in the ETC is the 13 protein complex cytochrome c oxidase, of which four genes are over expressed by HEPS CMCs [258]. This complex uses oxygen as a final electron acceptor [258]. This transfer to the final ETC protein complex is very important, as leakage of electrons before this transfer can prematurely react with oxygen, forming reactive oxygen species (ROS) [259]. Excessive formation of ROS leads to tissue damage from oxidative stress [260].

Oxidative phosphorylation also plays a vital role in the immune response, as it is the means by which ROS are produced [260]. ROS are important to induce inflammation and destroy invading pathogens, and are produced by activated neutrophils and macrophages [261]. However, as discussed in previous sections, inflammation in the context of HIV-1 transmission is decidedly not protective [108]. Therefore, it is unclear how a process that is linked to inflammation can be a mechanism of HIV-1 resistance. Nevertheless, the observed decrease in inflammation-related genes would suggest that, if

there is indeed an increase in ROS production, this is not coupled with an increase in inflammatory cytokine expression. However, oxidative phosphorylation plays other important roles in the immune response. Oxidative phosphorylation has been shown to induce the maturation of "alternatively activated" macrophages that are anti-inflammatory [241]. However, such macrophages are activated by IL4 and IL13 and are identified by the expression of arginase I [241], and none of these genes are significantly over expressed by HEPS CMCs; therefore it is unlikely differences in gene expression are due to an increase in "alternatively activated" macrophages in HEPS women. Oxidative phosphorylation is also the preferred energy metabolic process used by resting T cells, as activated T cells would divert macromolecules for anabolic processes [242]. However, one would have to isolate the T cell fraction of CMCs specifically before concluding there was a more quiescent T cell state in the FGT of HEPS women as compared to NN. Finally, this difference in oxidative phosphorylation expression by HEPS women could be a function of differences in age between the two groups as HEPS women are generally older than NN; however age typically leads to a decrease in oxidative phosphorylation, not an increase [260].

Diabetes signalling was also found to be significantly decreased in HEPS CMCs (Table 4). Diabetes has been correlated with the induction of the inflammatory transcription factor NFκB [262], and the gene for one subunit of this protein (NFKB2) was shown to be significantly under expressed in HEPS CMCs. NFκB has been shown to be induced by ROS [263]; therefore, this decrease in NFκB further suggests increased oxidative phosphorylation is not leading to an inflammatory state and is possibly reducing the oxidative stress in HEPS women. Further studies into the interplay between

oxidative phosphorylation, inflammation, and diabetes signalling are needed before conclusions can be drawn from the data.

The genes identified by the gene expression arrays as being significantly differentially expressed by HEPS CMCs as compared to NN must be confirmed by qRT-PCR and Western blotting. It should also be noted that an increase in the genes involved in this pathway does not necessarily mean there was a productive increase in the pathway itself. Functional studies targeting the identified pathways must be performed to determine if HEPS women have an increase in oxidative phosphorylation or if this increase in gene expression was a sign of a less efficient ETC. The production of ROS should be measured to determine if there is increased production in HEPS women. A less effective ETC could lead to the production of more ROS and therefore inflammation; however, other pathways identified as differentially expressed by HEPS women regulate these factors as well and would suggest this was not the case. As the ultimate product of the ETC is ATP, testing the differences in ATP levels between the groups could also confirm the results of this pathway.

5.2.3.2. MAPK

Several cancer pathways were significantly under expressed by CMCs collected from HEPS women. DAVID and GSEA also showed an increase in the MAPK pathway, which is a signalling cascade that responds to external stimuli and triggers a variety of cellular processes, including cellular proliferation and apoptosis (Figure 18). MAPK is a factor in approximately 1/3 of human cancers [202], and many genes identified as differentially expressed in the MAPK pathway were also found in the cancer

pathways; therefore, the MAPK pathway will be reviewed for discussion of the numerous identified cancer pathways.

There are currently 5 groups of MAPKs described in the literature: extracellular signal-regulated kinases (ERKs) 1/2, ERK3/4, ERK7/8, ERK5, c-Jun N-terminal kinases (JNKs) and p38 MAPKs [202]. All of the MAPKs are phosphorylated by a specific MAPKK (MAPK kinase), which is phosphorylated by MAP3K (MAPKK kinase), respond to different stimuli, and trigger different downstream molecules, partially explaining the variety of cellular functions regulated by the MAPK signalling pathway (Figure 18) [202]. The ERK1/2 module is initiated by growth factors and ends in the activation of several transcription factors involved the regulation of the cell cycle, including Elk-1 and c-myc (no difference in expression by HEPS and NN CMCs) and ATF4 (significantly decreased) [202]. Several genes involved in this module, including MAPK1 are under expressed by HEPS CMCs (Figure 18). This MAPK module is inhibited by a variety of MAPK phosphatases, one of which, DUSP5, is over expressed by HEPS CMCs [203].

The JNK and p38 modules are activated in response to cell stresses such as hypoxia and inflammatory cytokines such as IL1, and play a role in regulating the cell cycle and apoptosis [264]. There is no difference in expression of JNK or p38 isoforms nor their downstream pathways, suggesting these two MAPK modules are not likely differentially expressed by HEPS CMCs and will not be discussed further. ERK5 has a similar method of induction to ERK1/2 and differs primarily by function [265]. This module plays a role in cardiovascular and neural development and differentiation [265], and ERK5 is not differentially expressed by HEPS CMCs. The ERK3/4 and ERK7/8

modules are not well characterized in the literature, and these two kinases are also not significantly differentially expressed by HEPS CMCs.

Using only gene expression data, it was impossible to make conclusions as to the role of any of the MAPK modules in HIV-1 resistance. However, a decrease in the MAPK pathway suggests HEPS CMCs are less responsive to external stimuli as baseline levels, giving them a more quiescent phenotype. This assertion could be supported by the observed increase in oxidative phosphorylation and decrease in immune response in CMCs from HEPS individuals. The MAPK cascade is propagated by phosphorylation, and thus protein phosphorylation studies must be done to confirm a suppression of this cascade in HEPS women. The ERK1/2 signalling cascade is the most logical pathway to confirm initially, as this module showed a difference in expression level in the kinases themselves. Furthermore, at least one target of this module, ATF4, was shown to be significantly decreased in HEPS CMCs but needs to be confirmed by qRT-PCR and Western blotting. Confirming the differential expression of MAPK1 and its activating receptors by qRT-PCR and Western blot should also be performed to strengthen the results of this study.

5.2.3.3. Pathways related to the immune response

Several pathways involved in regulating the immune response were identified as significantly under expressed by HEPS CMCs via analysis of both GSEA databases, including inflammation. All immune response terms were clustered, but once again HEPS and NN did not cluster independently, again suggesting there was no single resistance phenotype in the HEPS immune response. TLR signalling plays an important role in inflammation, and is discussed below.

The data in this thesis shows that the TLR signalling pathway was significantly over expressed by CMCs as compared to PBMCs and plays an important role in inflammation; therefore, understanding this process in HEPS CMCs is important for understanding HIV-1 resistance at the genital tract level. To review, TLRs recognize PAMPs and trigger a signalling cascade that results in the transcription of inflammatory cytokines [92]. Therefore, a decrease in this pathway could suggest a decrease in inflammation in HEPS women. There was a decrease in the expression of TLR6 and TLR10 and MyD88 trends toward significantly under expression. NFKB2, a subunit of NFκB, was significantly decreased in HEPS CMCs, and not only plays a role in inflammatory cytokine transcription, but also in HIV-1 genome transcription [18]. NFκB target genes such as TNFα, IL1, RANTES and IL6 were not significantly differentially expressed according to the gene expression data, but previous work done on samples from women in the Pumwani cohort showed increased RANTES expression from cervicovaginal lavages (CVLs) from HEPS women [132]. This discrepancy could suggest that protein studies are best to quantify differences in cytokine expression between HEPS and NNs.

IL1R1 was identified as the most significantly differentially expressed gene in this set of arrays. IL1 is a potent inflammatory cytokine and plays a role in signal transduction. It binds to its receptor and triggers the transcription of NFκB, therefore inducing the expression of other inflammatory cytokines [205]. IL1 also signals via the MAPK cascades, and therefore can play a role in cell proliferation and differentiation. Although there is no significant decrease in IL1 between HEPS and NN, a decrease in the

expression of the receptor suggests that HEPS women are less responsive to this cytokine and its inflammatory effects.

CD86 was also found to be significantly decreased by HEPS CMCs by these gene expression arrays, and has been shown to be significantly under expressed by women with protective IRF1 phenotypes in this cohort (Aida Sivro, personal communication). CD86 is expressed by APCs and provides a co-stimulatory signal required to activate T cells [110]. Decreasing the expression of this receptor could suggest a lowered capacity for APCs to activate T cells at baseline levels.

The combination of a decrease in TLR signalling, as well as a decrease in the expression of IL1R1, CD86 and NFκB seems to suggest lowered inflammation in the CMCs from HEPS women. The implications of these findings are that decreasing inflammation in the FGT is one possible mechanism to prevent HIV-1 transmission. Decreases in the expression of TLRs, IL1R1 and CD86 would suppress inflammation, and would, therefore, protect the epithelial barrier against off-target effects (such as increased vascular permeability), and decrease the recruitment of target cells to the FGT.

The role of immune activation in susceptibility to HIV-1 infection has been much debated in the literature [125,128,266,267], but increasingly there is evidence showing that a lower immune activation state is more beneficial in prevention against and treating HIV-1 infection [268]. This decrease in immune activation acting as a protective mechanism against HIV-1 infection is reasonable since HIV-1 infects activated T cells [108]. Furthermore, activated macrophages and neutrophils recruit target cells to the site of infection, thus providing more target cells for HIV-1 to infect [72]. Activated neutrophils also secrete granzyme and cathepsins, which activate other immune cells and

kill infected host cells [269]. Granzyme and cathepsins can also have off-target effects, such as the damage to the cervical epithelial layer [270], thus providing a mechanism for HIV-1 to enter the host. However, a balance must be struck between protecting the host from invading pathogens and scaling back immune activation to prevent against HIV-1 infection. HEPS women do not appear to be at an increased risk of infection from other STIs; therefore, if they do have a decrease in their immune response, it does not seem to put them at a higher risk for other infections.

All the above mentioned genes must be confirmed to show there is indeed a difference in the immune response between HEPS and NN CMCs. Confirming the lowered expression of NFκB should be the first step, as this protein is an important transcription factor for inflammatory cytokines. Also, the decreased expression of IL1R1 must be confirmed, as IL1 plays a significant role in the induction of inflammation. NFκB should first be confirmed by qRT-PCR and then Western blot. The expression of IL1R1 and CD86 should be confirmed by flow cytometry to ensure differences in expression are not due to difference in cell populations between the two groups, as well as by Western blot and qRT-PCR. Once the decreased expression levels of IL1R1 and NFκB have been confirmed, they could prove to be an exciting new strategy to reduce inflammation and therefore also hopefully decrease HIV-1 transmission. Ultimately, however, the goal is to show there is a decrease in inflammation in HEPS women. Cytokine bead arrays on CVLs donated by women from the Pumwani cohort would provide an excellent means to evaluate the inflammatory cytokines secreted in the FGT by HEPS women. Repeating the CMC stimulations with HIV-1 peptides (i.e. gag) would

also be beneficial to determine if the HEPS women maintain a lower immune activation state when challenged with HIV-1 antigens.

5.2.3.4. Other pathways

Ribosomal genes were also significantly over expressed by HEPS CMCs, possibly indicating an increase in protein translation. Further research into the apparent up regulation of ribosomal proteins and ubiquination (Tables 4 and 5) in HEPS CMCs is warranted, as these pathways play a role on MHC Class I presentation [271].

5.2.4. Summary and future directions for HEPS vs. NN

Overall, CMCs from HEPS and NN women do not appear to have strikingly different global gene expression profiles. However, the data presented in this thesis does suggest HEPS women have decreased basal levels of inflammation as demonstrated by decreased expression by immune response-related genes, decreased MAPK signalling and increased oxidative phosphorylation.

There are many future investigations that should be followed based on the findings presented in this thesis. First, the sample size for these array studies should be increased to boost the statistical power and lower risk women should be included to make the results more comparable to the results from Songok *et al* [167], and to ensure the results from that study are not simply differences in gene expression caused by sex work. Also, a closer examination of the HEPS women who did cluster together for different pathways could elucidate novel resistance mechanisms in this subset of women. There are a number of improvements that can be instituted at the time of sampling to better

control for variable effects of menstrual cycle and douching, such as requesting the women refrain from douching for several days before the exam and to ask them to come in a set time after menses. Another explanation for the poor statistical significance of the results could be that there is no real difference in expression between HEPS and NN CMCs. However, as the FGT is still the site of HIV-1 challenge for the majority of the women in this cohort, additional gene expression arrays must be done on epithelial cells from the same groups to determine if HEPS women show unique epithelial gene profiles and to identify other markers of protection against HIV-1 infection. Furthermore, the reduction in inflammation hypothesized to be in HEPS women should be taken into consideration and further investigated. A number of anti-inflammatory proteins have already been identified by 2D gel electrophoresis of CVLs from HEPS women as compared to NN [158], and something similar could be done on CMC cell lysates from the same groups using mass spectroscopy. Once the decreased expression levels of IL1R1 and NFκB have been confirmed, finding a way to decrease the expression of these molecules *in vivo* could be an exciting HIV-1 prevention method, although caution must be taken to ensure no other cellular processes or pathways are affected. IL1 receptor antagonist (IL1RA) with an NFκB inhibitor such as acetyl-11-keto-beta-boswellic acid (AKBA) [272] at low levels in the form of a topical microbicide could be one prevention strategy.

6. Overall summary and conclusions

CMCs and PBMCs have unique gene expression profiles, mainly due to differences in cell populations between the two compartments but also due to differential expression of several pathways involved in inflammation. CMCs were shown to have increased inflammatory gene expression, which plays a role in HIV-1 transmission. Thus, it is important to consider the genital tract separately when studying HIV-1 transmission. CMCs from HEPS and NN women do not show global unique gene expression profiles, indicating that differences in gene expression patterns between the two groups either do not exist, are too subtle to be determined by whole genome expression analyses, or that there are many different mechanisms for HIV-1 resistance. There was an observed increase in oxidative phosphorylation and decreases in immune and MAPK pathways in HEPS women, potentially suggesting a more quiescent immune cell phenotype. It was not possible to study the effects of antigenic challenge on HIV-1 resistant women due to RNA quality issues.

This thesis is the first description of global gene expression differences between PBMCs and CMCs, showing increased inflammation in the FGT of healthy donors as compared to the periphery. This inflammation is extremely important to consider when studying HIV-1 infection, and significantly contributes to current knowledge of FGT immunobiology. This thesis is also the first to use whole genome expression arrays to attempt to identify markers of HIV-1 resistance in CMCs, and has contributed to the current understanding of HIV-1 resistance. Together, these projects suggest healthy women have higher inflammatory gene expression in the FGT relative to the periphery, and that HIV-1 resistant women have decreased levels of this inflammation in the FGT, suggesting lowered FGT inflammation as an HIV-1 resistance mechanism.

7. References

1. Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vezinet-Brun F, Rouzioux C, et al.: **Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS).** *Science* 1983, **220**:868-871.
2. Freed EO, Martin MA: **HIVs and Their Replication.** In *Fields Virology*, edn 4. Edited by Knipe DM HP: Lippincott Williams and Wilkins; 2001:1971-2041. vol 2.]
3. Gao F, Yue L, White AT, Pappas PG, Barchue J, Hanson AP, Greene BM, Sharp PM, Shaw GM, Hahn BH: **Human infection by genetically diverse SIVSM-related HIV-2 in west Africa.** *Nature* 1992, **358**:495-499.
4. UNAIDS: **2008 Report on the Global AIDS epidemic.** Edited by: Joint United Nations Programme on HIV/AIDS; 2008.
5. Poss M, Gosink J, Thomas E, Kreiss JK, Ndinya-Achola J, Mandaliya K, Bwayo J, Overbaugh J: **Phylogenetic evaluation of Kenyan HIV type 1 isolates.** *AIDS Res Hum Retroviruses* 1997, **13**:493-499.
6. Kiwanuka N, Laeyendecker O, Robb M, Kigozi G, Arroyo M, McCutchan F, Eller LA, Eller M, Makumbi F, Birx D, et al.: **Effect of human immunodeficiency virus Type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection.** *J Infect Dis* 2008, **197**:707-713.
7. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, et al.: **Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes.** *Nature* 1999, **397**:436-441.
8. Wolfe ND, Switzer WM, Carr JK, Bhullar VB, Shanmugam V, Tamoufe U, Prosser AT, Torimiro JN, Wright A, Mpoudi-Ngole E, et al.: **Naturally acquired simian retrovirus infections in central African hunters.** *Lancet* 2004, **363**:932-937.
9. **Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California.** *MMWR Morb Mortal Wkly Rep* 1981, **30**:305-308.
10. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, Muyembe JJ, Kabongo JM, Kalengayi RM, Van Marck E, et al.: **Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960.** *Nature* 2008, **455**:661-664.
11. Monini P, Sgadari C, Toschi E, Barillari G, Ensoli B: **Antitumour effects of antiretroviral therapy.** *Nat Rev Cancer* 2004, **4**:861-875.
12. Simon V, Ho DD: **HIV-1 dynamics in vivo: implications for therapy.** *Nat Rev Microbiol* 2003, **1**:181-190.
13. Rollins BJ: **Physiological coreceptor use by dual-tropic HIV-1: one plus one equals one.** *J Clin Invest* 1999, **104**:531-532.
14. Bleul CC, Wu L, Hoxie JA, Springer TA, Mackay CR: **The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes.** *Proc Natl Acad Sci U S A* 1997, **94**:1925-1930.
15. Beilhartz GL, Wendeler M, Baichoo N, Rausch J, Le Grice S, Gotte M: **HIV-1 reverse transcriptase can simultaneously engage its DNA/RNA substrate at both DNA polymerase and RNase H active sites: implications for RNase H inhibition.** *J Mol Biol* 2009, **388**:462-474.

16. Colicelli J, Goff SP: **Mutants and pseudorevertants of Moloney murine leukemia virus with alterations at the integration site.** *Cell* 1985, **42**:573-580.
17. Yang X, Chen Y, Gabuzda D: **ERK MAP kinase links cytokine signals to activation of latent HIV-1 infection by stimulating a cooperative interaction of AP-1 and NF-kappaB.** *J Biol Chem* 1999, **274**:27981-27988.
18. Nabel G, Baltimore D: **An inducible transcription factor activates expression of human immunodeficiency virus in T cells.** *Nature* 1987, **326**:711-713.
19. Brass AL, Dykxhoorn DM, Benita Y, Yan N, Engelman A, Xavier RJ, Lieberman J, Elledge SJ: **Identification of host proteins required for HIV infection through a functional genomic screen.** *Science* 2008, **319**:921-926.
20. Zhou H, Xu M, Huang Q, Gates AT, Zhang XD, Castle JC, Stec E, Ferrer M, Strulovici B, Hazuda DJ, et al.: **Genome-scale RNAi screen for host factors required for HIV replication.** *Cell Host Microbe* 2008, **4**:495-504.
21. Konig R, Zhou Y, Elleder D, Diamond TL, Bonamy GM, Irelan JT, Chiang CY, Tu BP, De Jesus PD, Lilley CE, et al.: **Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication.** *Cell* 2008, **135**:49-60.
22. Benko DM, Schwartz S, Pavlakis GN, Felber BK: **A novel human immunodeficiency virus type 1 protein, tev, shares sequences with tat, env, and rev proteins.** *J Virol* 1990, **64**:2505-2518.
23. Hallenberger S, Bosch V, Angliker H, Shaw E, Klenk HD, Garten W: **Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160.** *Nature* 1992, **360**:358-361.
24. Debouck C, Gorniak JG, Strickler JE, Meek TD, Metcalf BW, Rosenberg M: **Human immunodeficiency virus protease expressed in Escherichia coli exhibits autoprocessing and specific maturation of the gag precursor.** *Proc Natl Acad Sci U S A* 1987, **84**:8903-8906.
25. Mansky LM: **Retrovirus mutation rates and their role in genetic variation.** *J Gen Virol* 1998, **79 (Pt 6)**:1337-1345.
26. Delelis O, Carayon K, Saib A, Deprez E, Mouscadet JF: **Integrase and integration: biochemical activities of HIV-1 integrase.** *Retrovirology* 2008, **5**:114.
27. Brigati C, Giacca M, Noonan DM, Albini A: **HIV Tat, its TARgets and the control of viral gene expression.** *FEMS Microbiol Lett* 2003, **220**:57-65.
28. Foster JL, Garcia JV: **HIV-1 Nef: at the crossroads.** *Retrovirology* 2008, **5**:84.
29. Fischer U, Huber J, Boelens WC, Mattaj IW, Luhrmann R: **The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs.** *Cell* 1995, **82**:475-483.
30. 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, **418**:646-650.
31. Janket ML, Manickam P, Majumder B, Thotala D, Wagner M, Schafer EA, Collman RG, Srinivasan A, Ayyavoo V: **Differential regulation of host cellular genes by HIV-1 viral protein R (Vpr): cDNA microarray analysis using isogenic virus.** *Biochem Biophys Res Commun* 2004, **314**:1126-1132.
32. Miyagi E, Andrew AJ, Kao S, Strebel K: **Vpu enhances HIV-1 virus release in the absence of Bst-2 cell surface down-modulation and intracellular depletion.** *Proc Natl Acad Sci U S A* 2009, **106**:2868-2873.

33. Hymes KB, Cheung T, Greene JB, Prose NS, Marcus A, Ballard H, William DC, Laubenstein LJ: **Kaposi's sarcoma in homosexual men-a report of eight cases.** *Lancet* 1981, **2**:598-600.
34. Microbiology and Immunology Online: Chapter 7: Part 3 - Virology on World Wide Web URL: http://pathmicro.med.sc.edu/lecture/HIV3.htm
35. Kahn JO, Walker BD: **Acute human immunodeficiency virus type 1 infection.** *N Engl J Med* 1998, **339**:33-39.
36. Cohen OJ, Fauci AS: **Pathogenesis and medical aspects of HIV-1 infection.** In *Fields Virology*, edn 4. Edited by Knipe DM, Howley PM: Lippincott Williams & Wilkins; 2001:2043-2094. vol 2.]
37. Hammer SM, Squires KE, Hughes MD, Grimes JM, Demeter LM, Currier JS, Eron JJ, Jr., Feinberg JE, Balfour HH, Jr., Deyton LR, et al.: **A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team.** *N Engl J Med* 1997, **337**:725-733.
38. Bazin H, Chattopadhyaya J, Datema R, Ericson AC, Gilljam G, Johansson NG, Hansen J, Koshida R, Moelling K, Oberg B, et al.: **An analysis of the inhibition of replication of HIV and MuLV by some 3'-blocked pyrimidine analogs.** *Biochem Pharmacol* 1989, **38**:109-119.
39. De Clercq E: **Chemotherapy of the acquired immune deficiency syndrome (AIDS): non-nucleoside inhibitors of the human immunodeficiency virus type 1 reverse transcriptase.** *Int J Immunopharmacol* 1991, **13 Suppl 1**:83-89.
40. Lambert DM, Bartus H, Fernandez AV, Bratby-Anders C, Leary JJ, Dreyer GB, Metcalf BW, Petteway SR, Jr.: **Synergistic drug interactions of an HIV-1 protease inhibitor with AZT in different in vitro models of HIV-1 infection.** *Antiviral Res* 1993, **21**:327-342.
41. Caldwell DJ, Evans JD: **Developing clinical role of a CCR5 co-receptor antagonist in HIV-1 infection.** *Expert Opin Pharmacother* 2008, **9**:3231-3242.
42. Melby T, Westby M: **Inhibitors of viral entry.** *Handb Exp Pharmacol* 2009:177-202.
43. Markowitz M, Nguyen BY, Gotuzzo E, Mendo F, Ratanasuwan W, Kovacs C, Prada G, Morales-Ramirez JO, Crumpacker CS, Isaacs RD, et al.: **Rapid and durable antiretroviral effect of the HIV-1 Integrase inhibitor raltegravir as part of combination therapy in treatment-naive patients with HIV-1 infection: results of a 48-week controlled study.** *J Acquir Immune Defic Syndr* 2007, **46**:125-133.
44. Shattock RJ, Moore JP: **Inhibiting sexual transmission of HIV-1 infection.** *Nat Rev Microbiol* 2003, **1**:25-34.
45. Quinn TC, Overbaugh J: **HIV/AIDS in women: an expanding epidemic.** *Science* 2005, **308**:1582-1583.
46. Lavreys L, Baeten JM, Martin HL, Jr., Overbaugh J, Mandaliya K, Ndinya-Achola J, Kreiss JK: **Hormonal contraception and risk of HIV-1 acquisition: results of a 10-year prospective study.** *Aids* 2004, **18**:695-697.
47. Kiddugavu M, Makumbi F, Wawer MJ, Serwadda D, Sewankambo NK, Wabwire-Mangen F, Lutalo T, Meehan M, Xianbin, Gray RH: **Hormonal contraceptive**

use and HIV-1 infection in a population-based cohort in Rakai, Uganda. *Aids* 2003, **17**:233-240.
48. Gray RH, Li X, Kigozi G, Serwadda D, Brahmbhatt H, Wabwire-Mangen F, Nalugoda F, Kiddugavu M, Sewankambo N, Quinn TC, et al.: **Increased risk of incident HIV during pregnancy in Rakai, Uganda: a prospective study.** *Lancet* 2005, **366**:1182-1188.
49. Hirbod T, Broliden K: **Mucosal immune responses in the genital tract of HIV-1-exposed uninfected women.** *J Intern Med* 2007, **262**:44-58.
50. Carias AM, McCoombe S, McRaven M, Anderson M, Dinh M, Veazey RS, Hope T: **Mechanism of entry of HIV into the squamous epithelium of the female genital tract.** In *Keystone Symposia: HIV Immunobiology: From Infection to Immune Control*; *Keystone, CO*: 2009.
51. Balzarini J, Van Damme L: **Microbicide drug candidates to prevent HIV infection.** *Lancet* 2007, **369**:787-797.
52. Bomsel M: **Transcytosis of infectious human immunodeficiency virus across a tight human epithelial cell line barrier.** *Nat Med* 1997, 3:42-47.
53. de Witte L, Nabatov A, Pion M, Fluitsma D, de Jong MA, de Gruijl T, Piguet V, van Kooyk Y, Geijtenbeek TB: **Langerin is a natural barrier to HIV-1 transmission by Langerhans cells.** *Nat Med* 2007, **13**:367-371.
54. Blauvelt A, Asada H, Saville MW, Klaus-Kovtun V, Altman DJ, Yarchoan R, Katz SI: **Productive infection of dendritic cells by HIV-1 and their ability to capture virus are mediated through separate pathways.** *J Clin Invest* 1997, **100**:2043-2053.
55. Haase AT: **Perils at mucosal front lines for HIV and SIV and their hosts.** *Nat Rev Immunol* 2005, **5**:783-792.
56. Hu J, Gardner MB, Miller CJ: **Simian immunodeficiency virus rapidly penetrates the cervicovaginal mucosa after intravaginal inoculation and infects intraepithelial dendritic cells.** *J Virol* 2000, **74**:6087-6095.
57. Zhang Z, Schuler T, Zupancic M, Wietgrefe S, Staskus KA, Reimann KA, Reinhart TA, Rogan M, Cavert W, Miller CJ, et al.: **Sexual transmission and propagation of SIV and HIV in resting and activated CD4+ T cells.** *Science* 1999, **286**:1353-1357.
58. Honey K: **Microbicide trial screeches to a halt.** *J Clin Invest* 2007, **117**:1116.
59. Check E: **Scientists rethink approach to HIV gels.** *Nature* 2007, **446**:12.
60. Ramjee G, Govinden R, Morar NS, Mbewu A: **South Africa's experience of the closure of the cellulose sulphate microbicide trial.** *PLoS Med* 2007, 4:e235.
61. Hillier SL, Moench T, Shattock R, Black R, Reichelderfer P, Veronese F: **In vitro and in vivo: the story of nonoxynol 9.** *J Acquir Immune Defic Syndr* 2005, **39**:1-8.
62. Skoler-Karpoff S, Ramjee G, Ahmed K, Altini L, Plagianos MG, Friedland B, Govender S, De Kock A, Cassim N, Palanee T, et al.: **Efficacy of Carraguard for prevention of HIV infection in women in South Africa: a randomised, double-blind, placebo-controlled trial.** *Lancet* 2008, **372**:1977-1987.
63. Maguire RA, Bergman N, Phillips DM: **Comparison of microbicides for efficacy in protecting mice against vaginal challenge with herpes simplex virus type 2,**

cytotoxicity, antibacterial properties, and sperm immobilization. *Sex Transm Dis* 2001, **28**:259-265.
64. Teleshova N, Chang T, Profy A, Klotman ME: **Inhibitory effect of PRO 2000, a candidate microbicide, on dendritic cell-mediated human immunodeficiency virus transfer.** *Antimicrob Agents Chemother* 2008, **52**:1751-1758.
65. Mumper RJ, Bell MA, Worthen DR, Cone RA, Lewis GR, Paull JR, Moench TR: **Formulating a Sulfonated Antiviral Dendrimer in a Vaginal Microbicidal Gel Having Dual Mechanisms of Action.** *Drug Dev Ind Pharm* 2008:1-10.
66. Li Q, Estes JD, Schlievert PM, Duan L, Brosnahan AJ, Southern PJ, Reilly CS, Peterson ML, Schultz-Darken N, Brunner KG, et al.: **Glycerol monolaurate prevents mucosal SIV transmission.** *Nature* 2009.
67. O'Keefe BR, Vojdani F, Buffa V, Shattock RJ, Montefiori DC, Bakke J, Mirsalis J, d'Andrea AL, Hume SD, Bratcher B, et al.: **Scaleable manufacture of HIV-1 entry inhibitor griffithsin and validation of its safety and efficacy as a topical microbicide component.** *Proc Natl Acad Sci U S A* 2009, **106**:6099-6104.
68. Saxena BB, Han YA, Fu D, Rathnam P, Singh M, Laurence J, Lerner S: **Sustained release of microbicides by newly engineered vaginal rings.** *Aids* 2009, **23**:917-922.
69. Neil SJ, Zang T, Bieniasz PD: **Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu.** *Nature* 2008, **451**:425-430.
70. 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, **424**:99-103.
71. 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, **427**:848-853.
72. Janeway CA, Jr. Travers, P. Walport, M. Shlomchik, M.: *Immunobiology* edn 5. New York, NY: Garland Publishing; 2001.
73. Mestecky J, Fultz PN: **Mucosal immune system of the human genital tract.** *J Infect Dis* 1999, **179 Suppl 3**:S470-474.
74. Sun L, Finnegan CM, Kish-Catalone T, Blumenthal R, Garzino-Demo P, La Terra Maggiore GM, Berrone S, Kleinman C, Wu Z, Abdelwahab S, et al.: **Human beta-defensins suppress human immunodeficiency virus infection: potential role in mucosal protection.** *J Virol* 2005, **79**:14318-14329.
75. Medzhitov R, Preston-Hurlburt P, Janeway CA, Jr.: **A human homologue of the Drosophila Toll protein signals activation of adaptive immunity.** *Nature* 1997, **388**:394-397.
76. Kawai T, Akira S: **TLR signaling.** *Cell Death Differ* 2006, **13**:816-825.
77. Yamamoto M, Sato S, Mori K, Hoshino K, Takeuchi O, Takeda K, Akira S: **Cutting edge: a novel Toll/IL-1 receptor domain-containing adapter that preferentially activates the IFN-beta promoter in the Toll-like receptor signaling.** *J Immunol* 2002, **169**:6668-6672.
78. Fitzgerald KA, Palsson-McDermott EM, Bowie AG, Jefferies CA, Mansell AS, Brady G, Brint E, Dunne A, Gray P, Harte MT, et al.: **Mal (MyD88-adapter-like) is required for Toll-like receptor-4 signal transduction.** *Nature* 2001, **413**:78-83.

79. Medzhitov R, Preston-Hurlburt P, Kopp E, Stadlen A, Chen C, Ghosh S, Janeway CA, Jr.: **MyD88 is an adaptor protein in the hToll/IL-1 receptor family signaling pathways.** *Mol Cell* 1998, **2**:253-258.
80. Chen F, Castranova V, Shi X, Demers LM: **New insights into the role of nuclear factor-kappaB, a ubiquitous transcription factor in the initiation of diseases.** *Clin Chem* 1999, **45**:7-17.
81. Schwandner R, Dziarski R, Wesche H, Rothe M, Kirschning CJ: **Peptidoglycan- and lipoteichoic acid-induced cell activation is mediated by toll-like receptor 2.** *J Biol Chem* 1999, **274**:17406-17409.
82. Takeuchi O, Kaufmann A, Grote K, Kawai T, Hoshino K, Morr M, Muhlradt PF, Akira S: **Cutting edge: preferentially the R-stereoisomer of the mycoplasmal lipopeptide macrophage-activating lipopeptide-2 activates immune cells through a toll-like receptor 2- and MyD88-dependent signaling pathway.** *J Immunol* 2000, **164**:554-557.
83. Takeuchi O, Sato S, Horiuchi T, Hoshino K, Takeda K, Dong Z, Modlin RL, Akira S: **Cutting edge: role of Toll-like receptor 1 in mediating immune response to microbial lipoproteins.** *J Immunol* 2002, **169**:10-14.
84. Takeuchi O, Kawai T, Muhlradt PF, Morr M, Radolf JD, Zychlinsky A, Takeda K, Akira S: **Discrimination of bacterial lipoproteins by Toll-like receptor 6.** *Int Immunol* 2001, **13**:933-940.
85. Redondo-Lopez V, Cook RL, Sobel JD: **Emerging role of lactobacilli in the control and maintenance of the vaginal bacterial microflora.** *Rev Infect Dis* 1990, **12**:856-872.
86. Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC, Goodlett DR, Eng JK, Akira S, Underhill DM, Aderem A: **The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5.** *Nature* 2001, **410**:1099-1103.
87. Alexopoulou L, Holt AC, Medzhitov R, Flavell RA: **Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3.** *Nature* 2001, **413**:732-738.
88. Heil F, Hemmi H, Hochrein H, Ampenberger F, Kirschning C, Akira S, Lipford G, Wagner H, Bauer S: **Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8.** *Science* 2004, **303**:1526-1529.
89. Nazli A, Yao XD, Smieja M, Rosenthal KL, Ashkar AA, Kaushic C: **Differential induction of innate anti-viral responses by TLR ligands against Herpes simplex virus, type 2, infection in primary genital epithelium of women.** *Antiviral Res* 2009, **81**:103-112.
90. Baenziger S, Heikenwalder M, Johansen P, Schlaepfer E, Hofer U, Miller RC, Diemand S, Honda K, Kundig TM, Aguzzi A, et al.: **Triggering TLR7 in mice induces immune activation and lymphoid system disruption, resembling HIV-mediated pathology.** *Blood* 2009, **113**:377-388.
91. Hoshino K, Takeuchi O, Kawai T, Sanjo H, Ogawa T, Takeda Y, Takeda K, Akira S: **Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product.** *J Immunol* 1999, **162**:3749-3752.
92. Takeda K, Akira S: **TLR signaling pathways.** *Semin Immunol* 2004, **16**:3-9.

93. Zhang SY, Jouanguy E, Sancho-Shimizu V, von Bernuth H, Yang K, Abel L, Picard C, Puel A, Casanova JL: **Human Toll-like receptor-dependent induction of interferons in protective immunity to viruses.** *Immunol Rev* 2007, **220**:225-236.
94. Madigan MT MJ, Parker J: *Brock Biology of Microorganisms*, vol 10. Edited by G C. Upper Saddle River, NJ: Prentice Hall; 2003.
95. Fowke KR, Nagelkerke NJ, Kimani J, Simonsen JN, Anzala AO, Bwayo JJ, MacDonald KS, Ngugi EN, Plummer FA: **Resistance to HIV-1 infection among persistently seronegative prostitutes in Nairobi, Kenya.** *Lancet* 1996, **348**:1347-1351.
96. Hemmi H, Takeuchi O, Kawai T, Kaisho T, Sato S, Sanjo H, Matsumoto M, Hoshino K, Wagner H, Takeda K, et al.: **A Toll-like receptor recognizes bacterial DNA.** *Nature* 2000, **408**:740-745.
97. Hasan UA, Bates E, Takeshita F, Biliato A, Accardi R, Bouvard V, Mansour M, Vincent I, Gissmann L, Iftner T, et al.: **TLR9 expression and function is abolished by the cervical cancer-associated human papillomavirus type 16.** *J Immunol* 2007, **178**:3186-3197.
98. Butterfield TA, Best TM, Merrick MA: **The dual roles of neutrophils and macrophages in inflammation: a critical balance between tissue damage and repair.** *J Athl Train* 2006, **41**:457-465.
99. Steinman RM, Hawiger D, Nussenzweig MC: **Tolerogenic dendritic cells.** *Annu Rev Immunol* 2003, **21**:685-711.
100. Colucci F, Di Santo JP, Leibson PJ: **Natural killer cell activation in mice and men: different triggers for similar weapons?** *Nat Immunol* 2002, **3**:807-813.
101. Fernandez EJ, Lolis E: **Structure, function, and inhibition of chemokines.** *Annu Rev Pharmacol Toxicol* 2002, **42**:469-499.
102. Medzhitov R: **Origin and physiological roles of inflammation.** *Nature* 2008, **454**:428-435.
103. Schluns KS, Lefrancois L: **Cytokine control of memory T-cell development and survival.** *Nat Rev Immunol* 2003, **3**:269-279.
104. Bryant VL, Ma CS, Avery DT, Li Y, Good KL, Corcoran LM, de Waal Malefyt R, Tangye SG: **Cytokine-mediated regulation of human B cell differentiation into Ig-secreting cells: predominant role of IL-21 produced by CXCR5+ T follicular helper cells.** *J Immunol* 2007, **179**:8180-8190.
105. Dinarello CA: **Proinflammatory cytokines.** *Chest* 2000, **118**:503-508.
106. Martin S, Maruta K, Burkart V, Gillis S, Kolb H: **IL-1 and IFN-gamma increase vascular permeability.** *Immunology* 1988, **64**:301-305.
107. Bevilacqua MP, Nelson RM, Mannori G, Cecconi O: **Endothelial-leukocyte adhesion molecules in human disease.** *Annu Rev Med* 1994, **45**:361-378.
108. Rottman JB, Ganley KP, Williams K, Wu L, Mackay CR, Ringler DJ: **Cellular localization of the chemokine receptor CCR5. Correlation to cellular targets of HIV-1 infection.** *Am J Pathol* 1997, **151**:1341-1351.
109. Opal SM, DePalo VA: **Anti-inflammatory cytokines.** *Chest* 2000, **117**:1162-1172.
110. Jeannin P, Magistrelli G, Aubry JP, Caron G, Gauchat JF, Renno T, Herbault N, Goetsch L, Blaecke A, Dietrich PY, et al.: **Soluble CD86 is a costimulatory molecule for human T lymphocytes.** *Immunity* 2000, **13**:303-312.

111. Buchbinder SP, Mehrotra DV, Duerr A, Fitzgerald DW, Mogg R, Li D, Gilbert PB, Lama JR, Marmor M, Del Rio C, et al.: **Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial.** *Lancet* 2008, **372**:1881-1893.
112. Johnston MI, Fauci AS: **An HIV vaccine--challenges and prospects.** *N Engl J Med* 2008, **359**:888-890.
113. Pudney J, Quayle AJ, Anderson DJ: **Immunological microenvironments in the human vagina and cervix: mediators of cellular immunity are concentrated in the cervical transformation zone.** *Biol Reprod* 2005, **73**:1253-1263.
114. Hayday A, Theodoridis E, Ramsburg E, Shires J: **Intraepithelial lymphocytes: exploring the Third Way in immunology.** *Nat Immunol* 2001, **2**:997-1003.
115. Alexander R, Mestecky J: **Neutralizing antibodies in mucosal secretions: IgG or IgA?** *Curr HIV Res* 2007, **5**:588-593.
116. Cole AM: **Innate host defense of human vaginal and cervical mucosae.** *Curr Top Microbiol Immunol* 2006, **306**:199-230.
117. Pabst O, Bernhardt G, Forster R: **The impact of cell-bound antigen transport on mucosal tolerance induction.** *J Leukoc Biol* 2007, **82**:795-800.
118. Boris S, Barbes C: **Role played by lactobacilli in controlling the population of vaginal pathogens.** *Microbes Infect* 2000, **2**:543-546.
119. Deng L, Pan X, Wang Y, Wang L, Zhou XE, Li M, Feng Y, Wu Q, Wang B, Huang N: **Hemoglobin and its derived peptides may play a role in the antibacterial mechanism of the vagina.** *Hum Reprod* 2009, **24**:211-218.
120. Furci L, Lopalco L, Loverro P, Sinnone M, Tambussi G, Lazzarin A, Lusso P: **Non-cytotoxic inhibition of HIV-1 infection by unstimulated CD8+ T lymphocytes from HIV-exposed-uninfected individuals.** *Aids* 2002, **16**:1003-1008.
121. Rich KC, Siegel JN, Jennings C, Rydman RJ, Landay AL: **Function and phenotype of immature CD4+ lymphocytes in healthy infants and early lymphocyte activation in uninfected infants of human immunodeficiency virus-infected mothers.** *Clin Diagn Lab Immunol* 1997, **4**:358-361.
122. de la Torre MS, Torres C, Nieto G, Vergara S, Carrero AJ, Macias J, Pineda JA, Caruz A, Fibla J: **Vitamin D receptor gene haplotypes and susceptibility to HIV-1 infection in injection drug users.** *J Infect Dis* 2008, **197**:405-410.
123. Pinto LA, Sullivan J, Berzofsky JA, Clerici M, Kessler HA, Landay AL, Shearer GM: **ENV-specific cytotoxic T lymphocyte responses in HIV seronegative health care workers occupationally exposed to HIV-contaminated body fluids.** *J Clin Invest* 1995, **96**:867-876.
124. Koning FA, Jansen CA, Dekker J, Kaslow RA, Dukers N, van Baarle D, Prins M, Schuitemaker H: **Correlates of resistance to HIV-1 infection in homosexual men with high-risk sexual behaviour.** *Aids* 2004, **18**:1117-1126.
125. Miyazawa M, Lopalco L, Mazzotta F, Lo Caputo S, Veas F, Clerici M: **The 'immunologic advantage' of HIV-exposed seronegative individuals.** *Aids* 2009, **23**:161-175.
126. Kaul R, Rowland-Jones SL, Kimani J, Dong T, Yang HB, Kiama P, Rostron T, Njagi E, Bwayo JJ, MacDonald KS, et al.: **Late seroconversion in HIV-resistant Nairobi prostitutes despite pre-existing HIV-specific CD8+ responses.** *J Clin Invest* 2001, **107**:341-349.

127. Biasin M, Caputo SL, Speciale L, Colombo F, Racioppi L, Zagliani A, Ble C, Vichi F, Cianferoni L, Masci AM, et al.: **Mucosal and systemic immune activation is present in human immunodeficiency virus-exposed seronegative women.** *J Infect Dis* 2000, **182**:1365-1374.
128. Suy A, Castro P, Nomdedeu M, Garcia F, Lopez A, Fumero E, Gallart T, Lopalco L, Coll O, Gatell JM, et al.: **Immunological profile of heterosexual highly HIV-exposed uninfected individuals: predominant role of CD4 and CD8 T-cell activation.** *J Infect Dis* 2007, **196**:1191-1201.
129. Fowke K, McLaren PJ, Card CM, Kimani J, Luo M, Ball TB, Plummer FA: **The role of immune quiescence in resistance to HIV-1 infection.** In *CAHR*; *Montreal*: 2008.
130. Piacentini L, Biasin M, Fenizia C, Clerici M: **Genetic correlates of protection against HIV infection: the ally within.** *J Intern Med* 2009, **265**:110-124.
131. Martin MP, Gao X, Lee JH, Nelson GW, Detels R, Goedert JJ, Buchbinder S, Hoots K, Vlahov D, Trowsdale J, et al.: **Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS.** *Nat Genet* 2002, **31**:429-434.
132. Iqbal SM, Ball TB, Kimani J, Kiama P, Thottingal P, Embree JE, Fowke KR, Plummer FA: **Elevated T cell counts and RANTES expression in the genital mucosa of HIV-1-resistant Kenyan commercial sex workers.** *J Infect Dis* 2005, **192**:728-738.
133. Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, Nibbs RJ, Freedman BI, Quinones MP, Bamshad MJ, et al.: **The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility.** *Science* 2005, **307**:1434-1440.
134. Biasin M, Piacentini L, Lo Caputo S, Kanari Y, Magri G, Trabattoni D, Naddeo V, Lopalco L, Clivio A, Cesana E, et al.: **Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G: a possible role in the resistance to HIV of HIV-exposed seronegative individuals.** *J Infect Dis* 2007, **195**:960-964.
135. Misse D, Yssel H, Trabattoni D, Oblet C, Lo Caputo S, Mazzotta F, Pene J, Gonzalez JP, Clerici M, Veas F: **IL-22 participates in an innate anti-HIV-1 host-resistance network through acute-phase protein induction.** *J Immunol* 2007, **178**:407-415.
136. Javanbakht H, An P, Gold B, Petersen DC, O'Huigin C, Nelson GW, O'Brien SJ, Kirk GD, Detels R, Buchbinder S, et al.: **Effects of human TRIM5alpha polymorphisms on antiretroviral function and susceptibility to human immunodeficiency virus infection.** *Virology* 2006, **354**:15-27.
137. MacDonald KS, Fowke KR, Kimani J, Dunand VA, Nagelkerke NJ, Ball TB, Oyugi J, Njagi E, Gaur LK, Brunham RC, et al.: **Influence of HLA supertypes on susceptibility and resistance to human immunodeficiency virus type 1 infection.** *J Infect Dis* 2000, **181**:1581-1589.
138. Pastori C, Weiser B, Barassi C, Uberti-Foppa C, Ghezzi S, Longhi R, Calori G, Burger H, Kemal K, Poli G, et al.: **Long-lasting CCR5 internalization by antibodies in a subset of long-term nonprogressors: a possible protective effect against disease progression.** *Blood* 2006, 107:4825-4833.

139. Bomsel M, Pastori C, Tudor D, Alberti C, Garcia S, Ferrari D, Lazzarin A, Lopalco L: **Natural mucosal antibodies reactive with first extracellular loop of CCR5 inhibit HIV-1 transport across human epithelial cells.** *Aids* 2007, **21**:13-22.
140. Lopalco L, Barassi C, Paolucci C, Breda D, Brunelli D, Nguyen M, Nouhin J, Luong TT, Truong LX, Clerici M, et al.: **Predictive value of anti-cell and anti-human immunodeficiency virus (HIV) humoral responses in HIV-1-exposed seronegative cohorts of European and Asian origin.** *J Gen Virol* 2005, **86**:339-348.
141. Horton RE, Ball TB, Wachichi C, Jaoko W, Rutherford WJ, McKinnon L, Kaul R, Rebbapragada A, Kimani J, Plummer FA: **Cervical HIV-specific IgA in a population of commercial sex workers correlates with repeated exposure but not resistance to HIV.** *AIDS Res Hum Retroviruses* 2009, **25**:83-92.
142. Rowland-Jones SL, Dong T, Fowke KR, Kimani J, Krausa P, Newell H, Blanchard T, Ariyoshi K, Oyugi J, Ngugi E, et al.: **Cytotoxic T cell responses to multiple conserved HIV epitopes in HIV-resistant prostitutes in Nairobi.** *J Clin Invest* 1998, **102**:1758-1765.
143. Kaul R, Plummer FA, Kimani J, Dong T, Kiama P, Rostron T, Njagi E, MacDonald KS, Bwayo JJ, McMichael AJ, et al.: **HIV-1-specific mucosal CD8+ lymphocyte responses in the cervix of HIV-1-resistant prostitutes in Nairobi.** *J Immunol* 2000, **164**:1602-1611.
144. Makedonas G, Bruneau J, Lin H, Sekaly RP, Lamothe F, Bernard NF: **HIV-specific CD8 T-cell activity in uninfected injection drug users is associated with maintenance of seronegativity.** *Aids* 2002, **16**:1595-1602.
145. Pallikkuth S, Wanchu A, Bhatnagar A, Sachdeva RK, Sharma M: **Human immunodeficiency virus (HIV) gag antigen-specific T-helper and granule-dependent CD8 T-cell activities in exposed but uninfected heterosexual partners of HIV type 1-infected individuals in North India.** *Clin Vaccine Immunol* 2007, **14**:1196-1202.
146. Goh WC, Markee J, Akridge RE, Meldorf M, Musey L, Karchmer T, Krone M, Collier A, Corey L, Emerman M, et al.: **Protection against human immunodeficiency virus type 1 infection in persons with repeated exposure: evidence for T cell immunity in the absence of inherited CCR5 coreceptor defects.** *J Infect Dis* 1999, **179**:548-557.
147. Kaul R, Dong T, Plummer FA, Kimani J, Rostron T, Kiama P, Njagi E, Irungu E, Farah B, Oyugi J, et al.: **CD8(+) lymphocytes respond to different HIV epitopes in seronegative and infected subjects.** *J Clin Invest* 2001, **107**:1303-1310.
148. Kebba A, Kaleebu P, Rowland S, Ingram R, Whitworth J, Imami N, Gotch F: **Distinct patterns of peripheral HIV-1-specific interferon- gamma responses in exposed HIV-1-seronegative individuals.** *J Infect Dis* 2004, **189**:1705-1713.
149. Furci L, Scarlatti G, Burastero S, Tambussi G, Colognesi C, Quillent C, Longhi R, Loverro P, Borgonovo B, Gaffi D, et al.: **Antigen-driven C-C chemokine-mediated HIV-1 suppression by CD4(+) T cells from exposed uninfected individuals expressing the wild-type CCR-5 allele.** *J Exp Med* 1997, **186**:455-460.

150. Jennes W, Sawadogo S, Koblavi-Deme S, Vuylsteke B, Maurice C, Roels TH, Chorba T, Nkengasong JN, Kestens L: **Cellular human immunodeficiency virus (HIV)-protective factors: a comparison of HIV-exposed seronegative female sex workers and female blood donors in Abidjan, Cote d'Ivoire.** *J Infect Dis* 2003, **187**:206-214.
151. Schenal M, Lo Caputo S, Fasano F, Vichi F, Saresella M, Pierotti P, Villa ML, Mazzotta F, Trabattoni D, Clerici M: **Distinct patterns of HIV-specific memory T lymphocytes in HIV-exposed uninfected individuals and in HIV-infected patients.** *Aids* 2005, **19**:653-661.
152. Card CM, McLaren PJ, Wachihi C, Kimani J, Plummer FA, Fowke KR: **Decreased Immune Activation in Resistance to HIV-1 Infection Is Associated with an Elevated Frequency of CD4(+)CD25(+)FOXP3(+) Regulatory T Cells.** *J Infect Dis* 2009, **199**:1318-1322.
153. Kaul R, Rowland-Jones SL, Kimani J, Fowke K, Dong T, Kiama P, Rutherford J, Njagi E, Mwangi F, Rostron T, et al.: **New insights into HIV-1 specific cytotoxic T-lymphocyte responses in exposed, persistently seronegative Kenyan sex workers.** *Immunol Lett* 2001, **79**:3-13.
154. Kaul R, Thottingal P, Kimani J, Kiama P, Waigwa CW, Bwayo JJ, Plummer FA, Rowland-Jones SL: **Quantitative ex vivo analysis of functional virus-specific CD8 T lymphocytes in the blood and genital tract of HIV-infected women.** *Aids* 2003, **17**:1139-1144.
155. Plummer FA, Ball TB, Kimani J, Fowke KR: **Resistance to HIV-1 infection among highly exposed sex workers in Nairobi: what mediates protection and why does it develop?** *Immunol Lett* 1999, **66**:27-34.
156. Kaul R, Trabattoni D, Bwayo JJ, Arienti D, Zagliani A, Mwangi FM, Kariuki C, Ngugi EN, MacDonald KS, Ball TB, et al.: **HIV-1-specific mucosal IgA in a cohort of HIV-1-resistant Kenyan sex workers.** *Aids* 1999, **13**:23-29.
157. Berry JD, Rutherford J, Silverman GJ, Kaul R, Elia M, Gobuty S, Fuller R, Plummer FA, Barbas CF, 3rd: **Development of functional human monoclonal single-chain variable fragment antibody against HIV-1 from human cervical B cells.** *Hybrid Hybridomics* 2003, **22**:97-108.
158. Burgener A, Boutilier J, Wachihi C, Kimani J, Carpenter M, Westmacott G, Cheng K, Ball TB, Plummer F: **Identification of differentially expressed proteins in the cervical mucosa of HIV-1-resistant sex workers.** *J Proteome Res* 2008, **7**:4446-4454.
159. Jayakumar A, Kang Y, Frederick MJ, Pak SC, Henderson Y, Holton PR, Mitsudo K, Silverman GA, AK EL-N, Bromme D, et al.: **Inhibition of the cysteine proteinases cathepsins K and L by the serpin headpin (SERPINB13): a kinetic analysis.** *Arch Biochem Biophys* 2003, **409**:367-374.
160. Schick C, Kamachi Y, Bartuski AJ, Cataltepe S, Schechter NM, Pemberton PA, Silverman GA: **Squamous cell carcinoma antigen 2 is a novel serpin that inhibits the chymotrypsin-like proteinases cathepsin G and mast cell chymase.** *J Biol Chem* 1997, **272**:1849-1855.
161. Galliano MF, Toulza E, Gallinaro H, Jonca N, Ishida-Yamamoto A, Serre G, Guerrin M: **A novel protease inhibitor of the alpha2-macroglobulin family expressed in the human epidermis.** *J Biol Chem* 2006, **281**:5780-5789.

162. Voltersvik P, Bostad L, Dyrhol-Riise AM, Eide GE, Rosok BI, Olofsson J, Asjo B: **Cystatin A and HIV-1 p24 antigen expression in tonsillar lymphoid follicles during HIV-1 infection and during highly active antiretroviral therapy.** *J Acquir Immune Defic Syndr* 2006, **41**:277-284.
163. Hardie RA, Knight E, Bruneau B, Semeniuk C, Gill K, Nagelkerke N, Kimani J, Wachihi C, Ngugi E, Luo M, et al.: **A common human leucocyte antigen-DP genotype is associated with resistance to HIV-1 infection in Kenyan sex workers.** *Aids* 2008, **22**:2038-2042.
164. Hardie RA, Luo M, Bruneau B, Knight E, Nagelkerke NJ, Kimani J, Wachihi C, Ngugi EN, Plummer FA: **Human leukocyte antigen-DQ alleles and haplotypes and their associations with resistance and susceptibility to HIV-1 infection.** *Aids* 2008, **22**:807-816.
165. Lacap PA, Huntington JD, Luo M, Nagelkerke NJ, Bielawny T, Kimani J, Wachihi C, Ngugi EN, Plummer FA: **Associations of human leukocyte antigen DRB with resistance or susceptibility to HIV-1 infection in the Pumwani Sex Worker Cohort.** *Aids* 2008, **22**:1029-1038.
166. Ball TB, Ji H, Kimani J, McLaren P, Marlin C, Hill AV, Plummer FA: **Polymorphisms in IRF-1 associated with resistance to HIV-1 infection in highly exposed uninfected Kenyan sex workers.** *Aids* 2007, **21**:1091-1101.
167. Songok EM, Ball TB, Luo M, Liang B, Kaefer N, Apidi W, McLaren P, Kimani J, Wachichi C, Samson BS, et al.: **Genome wide expression analysis of HIV Exposed uninfected commercial sex workers reveal novel biomarkers of HIV resistance and a metabolic signature pattern of a quiescent immune state.** In *Keystone Symposia: Prevention of HIV/AIDS*; *Keystone*: 2009.
168. Bryant PA, Smyth GK, Robins-Browne R, Curtis N: **Detection of gene expression in an individual cell type within a cell mixture using microarray analysis.** *PLoS ONE* 2009, **4**:e4427.
169. McLaren PJ, Mayne M, Rosser S, Moffatt T, Becker KG, Plummer FA, Fowke KR: **Antigen-specific gene expression profiles of peripheral blood mononuclear cells do not reflect those of T-lymphocyte subsets.** *Clin Diagn Lab Immunol* 2004, **11**:977-982.
170. Hirschhorn K, Bach F, Kolodny RL, Firschein IL, Hashem N: **Immune Response and Mitosis of Human Peripheral Blood Lymphocytes in Vitro.** *Science* 1963, **142**:1185-1187.
171. Qiagen: **RNeasy® MinElute® Cleanup Handbook.** 2007.
172. Scientific T: **NanoDrop 1000 Spectrophotometer V3.7 User's Manual.** 2008.
173. Agilent: **Agilent RNA 6000 Nano Kit Guide.** 2006.
174. Qiagen: **QuantiTect® SYBR® Green PCR Handbook.** 2008.
175. Sigma: **TransPlex® Whole Transcriptome Amplification Kit.** Edited by; 2008.
176. Qiagen: **RNeasy® Plus Mini Handbook.** 2005.
177. SABiosciences: **Technical Note: Roche LightCycler 480 Instrument Setup Instructions.** Edited by; 2008.
178. Pfaffl MW: **A new mathematical model for relative quantification in real-time RT-PCR.** *Nucleic Acids Res* 2001, **29**:e45.
179. Ambion: **Illumina® TotalPrep RNA Amplification Kit.** Edited by; 2006.

180. Illumina: **Human-6 v2 and HumanRef-8 v2 Expression BeadChips**. Edited by; 2006.
181. Illumina: **Whole-Genome Gene Expression Direct Hybridization Assay Guide**. 2007.
182. Illumina: **BeadArray Reader User Guide**. 2005.
183. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, et al.: **Bioconductor: open software development for computational biology and bioinformatics**. *Genome Biol* 2004, 5:R80.
184. Benjamini Y HY: **Controlling the False Discovery Rate - A Practical and Powerful Approach**. *The Royal Statistical Society Series B - Methodological* 1995, **57**:289-300.
185. Eisen M: **Cluster and TreeView Manual**. Edited by. Stanford University; 1998.
186. Huang da W, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC, et al.: **DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists**. *Nucleic Acids Res* 2007, **35**:W169-175.
187. Hosack DA, Dennis G, Jr., Sherman BT, Lane HC, Lempicki RA: **Identifying biological themes within lists of genes with EASE**. *Genome Biol* 2003, **4**:R70.
188. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes**. *Nucleic Acids Res* 2000, **28**:27-30.
189. Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, Hirakawa M: **From genomics to chemical genomics: new developments in KEGG**. *Nucleic Acids Res* 2006, **34**:D354-357.
190. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, et al.: **KEGG for linking genomes to life and the environment**. *Nucleic Acids Res* 2008, **36**:D480-484.
191. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al.: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles**. *Proc Natl Acad Sci U S A* 2005, **102**:15545-15550.
192. Hiesinger M LD, Ritt C, Oelmuller U: **The effects of phenol on nucleic acid preparation and downstream applications**. *Qiagen News* 2001:4.
193. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T: **The RIN: an RNA integrity number for assigning integrity values to RNA measurements**. *BMC Mol Biol* 2006, **7**:3.
194. Kuhn K, Baker SC, Chudin E, Lieu MH, Oeser S, Bennett H, Rigault P, Barker D, McDaniel TK, Chee MS: **A novel, high-performance random array platform for quantitative gene expression profiling**. *Genome Res* 2004, **14**:2347-2356.
195. Prakash M, Patterson S, Kapembwa MS: **Evaluation of the cervical cytobrush sampling technique for the preparation of CD45+ mononuclear cells from the human cervix**. *J Immunol Methods* 2001, **258**:37-46.
196. Kaefer N, Horton RE, Songok EM, Guijon FB, Kettaf N, Boucher G, Sekaly RP, Ball TB, Plummer FA: **A comparative analysis of gene expression patterns and cell phenotypes between cervical and peripheral blood mononuclear cells**. *PLoS One* 2009 (submitted).

197. Huang da W, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.** *Nat Protoc* 2009, **4**:44-57.
198. Tanaka M, Miyajima A: **Oncostatin M, a multifunctional cytokine.** *Rev Physiol Biochem Pharmacol* 2003, **149**:39-52.
199. Arikawa E, Sun Y, Wang J, Zhou Q, Ning B, Dial SL, Guo L, Yang J: **Cross-platform comparison of SYBR Green real-time PCR with TaqMan PCR, microarrays and other gene expression measurement technologies evaluated in the MicroArray Quality Control (MAQC) study.** *BMC Genomics* 2008, **9**:328.
200. Cecchini G: **Function and structure of complex II of the respiratory chain.** *Annu Rev Biochem* 2003, **72**:77-109.
201. Iwata S, Lee JW, Okada K, Lee JK, Iwata M, Rasmussen B, Link TA, Ramaswamy S, Jap BK: **Complete structure of the 11-subunit bovine mitochondrial cytochrome bc1 complex.** *Science* 1998, **281**:64-71.
202. Dhillon AS, Hagan S, Rath O, Kolch W: **MAP kinase signalling pathways in cancer.** *Oncogene* 2007, **26**:3279-3290.
203. Keyse SM: **Dual-specificity MAP kinase phosphatases (MKPs) and cancer.** *Cancer Metastasis Rev* 2008, **27**:253-261.
204. Chaussabel D, Quinn C, Shen J, Patel P, Glaser C, Baldwin N, Stichweh D, Blankenship D, Li L, Munagala I, et al.: **A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus.** *Immunity* 2008, **29**:150-164.
205. Baek SH, Ohgi KA, Rose DW, Koo EH, Glass CK, Rosenfeld MG: **Exchange of N-CoR corepressor and Tip60 coactivator complexes links gene expression by NF-kappaB and beta-amyloid precursor protein.** *Cell* 2002, **110**:55-67.
206. Andersch-Bjorkman Y, Thomsson KA, Holmen Larsson JM, Ekerhovd E, Hansson GC: **Large Scale Identification of Proteins, Mucins, and Their O-Glycosylation in the Endocervical Mucus during the Menstrual Cycle.** *Mol Cell Proteomics* 2007, **6**:708-716.
207. Rubin BK: **An in vitro comparison of the mucoactive properties of guaifenesin, iodinated glycerol, surfactant, and albuterol.** *Chest* 1999, **116**:195-200.
208. Bradley JR: **TNF-mediated inflammatory disease.** *J Pathol* 2008, **214**:149-160.
209. Wolpe SD, Davatelis G, Sherry B, Beutler B, Hesse DG, Nguyen HT, Moldawer LL, Nathan CF, Lowry SF, Cerami A: **Macrophages secrete a novel heparin-binding protein with inflammatory and neutrophil chemokinetic properties.** *J Exp Med* 1988, **167**:570-581.
210. Smith DE, Hanna R, Della F, Moore H, Chen H, Farese AM, MacVittie TJ, Virca GD, Sims JE: **The soluble form of IL-1 receptor accessory protein enhances the ability of soluble type II IL-1 receptor to inhibit IL-1 action.** *Immunity* 2003, **18**:87-96.
211. Kotenko SV, Krause CD, Izotova LS, Pollack BP, Wu W, Pestka S: **Identification and functional characterization of a second chain of the interleukin-10 receptor complex.** *Embo J* 1997, **16**:5894-5903.
212. Birge RB, Ucker DS: **Innate apoptotic immunity: the calming touch of death.** *Cell Death Differ* 2008, **15**:1096-1102.

213. Mosley B, De Imus C, Friend D, Boiani N, Thoma B, Park LS, Cosman D: **Dual oncostatin M (OSM) receptors. Cloning and characterization of an alternative signaling subunit conferring OSM-specific receptor activation.** *J Biol Chem* 1996, **271**:32635-32643.
214. Hintzen C, Haan C, Tuckermann JP, Heinrich PC, Hermanns HM: **Oncostatin M-induced and constitutive activation of the JAK2/STAT5/CIS pathway suppresses CCL1, but not CCL7 and CCL8, chemokine expression.** *J Immunol* 2008, **181**:7341-7349.
215. Zhang Q, Hui W, Litherland GJ, Barter MJ, Davidson R, Darrah C, Donell ST, Clark IM, Cawston TE, Robinson JH, et al.: **Differential Toll-like receptor-dependent collagenase expression in chondrocytes.** *Ann Rheum Dis* 2008, **67**:1633-1641.
216. Wira CR, Fahey JV, Sentman CL, Pioli PA, Shen L: **Innate and adaptive immunity in female genital tract: cellular responses and interactions.** *Immunol Rev* 2005, **206**:306-335.
217. Andersen JM, Al-Khairy D, Ingalls RR: **Innate immunity at the mucosal surface: role of toll-like receptor 3 and toll-like receptor 9 in cervical epithelial cell responses to microbial pathogens.** *Biol Reprod* 2006, **74**:824-831.
218. Fichorova RN, Cronin AO, Lien E, Anderson DJ, Ingalls RR: **Response to Neisseria gonorrhoeae by cervicovaginal epithelial cells occurs in the absence of toll-like receptor 4-mediated signaling.** *J Immunol* 2002, **168**:2424-2432.
219. Krammer PH, Arnold R, Lavrik IN: **Life and death in peripheral T cells.** *Nat Rev Immunol* 2007, **7**:532-542.
220. Tait SW, Green DR: **Caspase-independent cell death: leaving the set without the final cut.** *Oncogene* 2008, **27**:6452-6461.
221. Chinnaiyan AM, O'Rourke K, Tewari M, Dixit VM: **FADD, a novel death domain-containing protein, interacts with the death domain of Fas and initiates apoptosis.** *Cell* 1995, **81**:505-512.
222. Muzio M, Chinnaiyan AM, Kischkel FC, O'Rourke K, Shevchenko A, Ni J, Scaffidi C, Bretz JD, Zhang M, Gentz R, et al.: **FLICE, a novel FADD-homologous ICE/CED-3-like protease, is recruited to the CD95 (Fas/APO-1) death--inducing signaling complex.** *Cell* 1996, **85**:817-827.
223. Thornberry NA, Lazebnik Y: **Caspases: enemies within.** *Science* 1998, **281**:1312-1316.
224. Chang DW, Xing Z, Pan Y, Algeciras-Schimnich A, Barnhart BC, Yaish-Ohad S, Peter ME, Yang X: **c-FLIP(L) is a dual function regulator for caspase-8 activation and CD95-mediated apoptosis.** *Embo J* 2002, **21**:3704-3714.
225. Budd RC, Yeh WC, Tschopp J: **cFLIP regulation of lymphocyte activation and development.** *Nat Rev Immunol* 2006, **6**:196-204.
226. Srinivasula SM, Ahmad M, Ottilie S, Bullrich F, Banks S, Wang Y, Fernandes-Alnemri T, Croce CM, Litwack G, Tomaselli KJ, et al.: **FLAME-1, a novel FADD-like anti-apoptotic molecule that regulates Fas/TNFR1-induced apoptosis.** *J Biol Chem* 1997, **272**:18542-18545.
227. Irmler M, Thome M, Hahne M, Schneider P, Hofmann K, Steiner V, Bodmer JL, Schroter M, Burns K, Mattmann C, et al.: **Inhibition of death receptor signals by cellular FLIP.** *Nature* 1997, **388**:190-195.

228. Golks A, Brenner D, Fritsch C, Krammer PH, Lavrik IN: **c-FLIPR, a new regulator of death receptor-induced apoptosis.** *J Biol Chem* 2005, **280**:14507-14513.
229. Kataoka T, Budd RC, Holler N, Thome M, Martinon F, Irmler M, Burns K, Hahne M, Kennedy N, Kovacsovics M, et al.: **The caspase-8 inhibitor FLIP promotes activation of NF-kappaB and Erk signaling pathways.** *Curr Biol* 2000, **10**:640-648.
230. Micheau O, Lens S, Gaide O, Alevizopoulos K, Tschopp J: **NF-kappaB signals induce the expression of c-FLIP.** *Mol Cell Biol* 2001, **21**:5299-5305.
231. Karst AM, Li G: **BH3-only proteins in tumorigenesis and malignant melanoma.** *Cell Mol Life Sci* 2007, **64**:318-330.
232. Nakano K, Vousden KH: **PUMA, a novel proapoptotic gene, is induced by p53.** *Mol Cell* 2001, **7**:683-694.
233. Oda E, Ohki R, Murasawa H, Nemoto J, Shibue T, Yamashita T, Tokino T, Taniguchi T, Tanaka N: **Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis.** *Science* 2000, **288**:1053-1058.
234. Breckenridge DG, Xue D: **Regulation of mitochondrial membrane permeabilization by BCL-2 family proteins and caspases.** *Curr Opin Cell Biol* 2004, **16**:647-652.
235. Zhai D, Jin C, Huang Z, Satterthwait AC, Reed JC: **Differential regulation of Bax and Bak by anti-apoptotic Bcl-2 family proteins Bcl-B and Mcl-1.** *J Biol Chem* 2008, **283**:9580-9586.
236. Li S, Zhao Y, He X, Kim TH, Kuharsky DK, Rabinowich H, Chen J, Du C, Yin XM: **Relief of extrinsic pathway inhibition by the Bid-dependent mitochondrial release of Smac in Fas-mediated hepatocyte apoptosis.** *J Biol Chem* 2002, **277**:26912-26920.
237. Zou H, Li Y, Liu X, Wang X: **An APAF-1.cytochrome c multimeric complex is a functional apoptosome that activates procaspase-9.** *J Biol Chem* 1999, **274**:11549-11556.
238. Ohta T, Kinoshita T, Naito M, Nozaki T, Masutani M, Tsuruo T, Miyajima A: **Requirement of the caspase-3/CPP32 protease cascade for apoptotic death following cytokine deprivation in hematopoietic cells.** *J Biol Chem* 1997, **272**:23111-23116.
239. Kuida K, Haydar TF, Kuan CY, Gu Y, Taya C, Karasuyama H, Su MS, Rakic P, Flavell RA: **Reduced apoptosis and cytochrome c-mediated caspase activation in mice lacking caspase 9.** *Cell* 1998, **94**:325-337.
240. Jaattela M, Tschopp J: **Caspase-independent cell death in T lymphocytes.** *Nat Immunol* 2003, **4**:416-423.
241. Vats D, Mukundan L, Odegaard JI, Zhang L, Smith KL, Morel CR, Wagner RA, Greaves DR, Murray PJ, Chawla A: **Oxidative metabolism and PGC-1beta attenuate macrophage-mediated inflammation.** *Cell Metab* 2006, **4**:13-24.
242. Fox CJ, Hammerman PS, Thompson CB: **Fuel feeds function: energy metabolism and the T-cell response.** *Nat Rev Immunol* 2005, **5**:844-852.
243. Gaucher D, Therrien R, Kettaf N, Angermann BR, Boucher G, Filali-Mouhim A, Moser JM, Mehta RS, Drake DR, 3rd, Castro E, et al.: **Yellow fever vaccine**

induces integrated multilineage and polyfunctional immune responses. *J Exp Med* 2008, **205**:3119-3131.
244. Garcia-Tunon I, Ricote M, Ruiz A, Fraile B, Paniagua R, Royuela M: **OSM, LIF, its receptors, and its relationship with the malignance in human breast carcinoma (in situ and in infiltrative).** *Cancer Invest* 2008, **26**:222-229.
245. Cross A, Edwards SW, Bucknall RC, Moots RJ: **Secretion of oncostatin M by neutrophils in rheumatoid arthritis.** *Arthritis Rheum* 2004, **50**:1430-1436.
246. Royuela M, Ricote M, Parsons MS, Garcia-Tunon I, Paniagua R, de Miguel MP: **Immunohistochemical analysis of the IL-6 family of cytokines and their receptors in benign, hyperplasic, and malignant human prostate.** *J Pathol* 2004, **202**:41-49.
247. Baeza-Raja B, Munoz-Canoves P: **p38 MAPK-induced nuclear factor-kappaB activity is required for skeletal muscle differentiation: role of interleukin-6.** *Mol Biol Cell* 2004, **15**:2013-2026.
248. Savill JS, Wyllie AH, Henson JE, Walport MJ, Henson PM, Haslett C: **Macrophage phagocytosis of aging neutrophils in inflammation. Programmed cell death in the neutrophil leads to its recognition by macrophages.** *J Clin Invest* 1989, **83**:865-875.
249. Meagher LC, Savill JS, Baker A, Fuller RW, Haslett C: **Phagocytosis of apoptotic neutrophils does not induce macrophage release of thromboxane B2.** *J Leukoc Biol* 1992, **52**:269-273.
250. Cvetanovic M, Ucker DS: **Innate immune discrimination of apoptotic cells: repression of proinflammatory macrophage transcription is coupled directly to specific recognition.** *J Immunol* 2004, **172**:880-889.
251. Nauseef WM: **Isolation of human neutrophils from venous blood.** *Methods Mol Biol* 2007, **412**:15-20.
252. Smeitink J, van den Heuvel L, DiMauro S: **The genetics and pathology of oxidative phosphorylation.** *Nat Rev Genet* 2001, **2**:342-352.
253. Tortora GJ, Grabowski SR: **Metabolism.** In *Principles of Anatomy and Physiology*, edn 10th. Edited by Roesch B: John Wiley & Sons; 2003:914.
254. Carroll J, Fearnley IM, Skehel JM, Shannon RJ, Hirst J, Walker JE: **Bovine complex I is a complex of 45 different subunits.** *J Biol Chem* 2006, **281**:32724-32727.
255. Horsefield R, Iwata S, Byrne B: **Complex II from a structural perspective.** *Curr Protein Pept Sci* 2004, **5**:107-118.
256. Zhang J, Frerman FE, Kim JJ: **Structure of electron transfer flavoprotein-ubiquinone oxidoreductase and electron transfer to the mitochondrial ubiquinone pool.** *Proc Natl Acad Sci U S A* 2006, **103**:16212-16217.
257. Crofts AR: **The cytochrome bc1 complex: function in the context of structure.** *Annu Rev Physiol* 2004, **66**:689-733.
258. Tsukihara T, Aoyama H, Yamashita E, Tomizaki T, Yamaguchi H, Shinzawa-Itoh K, Nakashima R, Yaono R, Yoshikawa S: **The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 A.** *Science* 1996, **272**:1136-1144.
259. Young TA, Cunningham CC, Bailey SM: **Reactive oxygen species production by the mitochondrial respiratory chain in isolated rat hepatocytes and liver**

mitochondria: studies using myxothiazol. *Arch Biochem Biophys* 2002, **405**:65-72.

260. Genova ML, Pich MM, Bernacchia A, Bianchi C, Biondi A, Bovina C, Falasca AI, Formiggini G, Castelli GP, Lenaz G: **The mitochondrial production of reactive oxygen species in relation to aging and pathology.** *Ann N Y Acad Sci* 2004, **1011**:86-100.
261. Forman HJ, Torres M: **Reactive oxygen species and cell signaling: respiratory burst in macrophage signaling.** *Am J Respir Crit Care Med* 2002, **166**:S4-8.
262. Dandona P, Aljada A, Mohanty P, Ghanim H, Hamouda W, Assian E, Ahmad S: **Insulin inhibits intranuclear nuclear factor kappaB and stimulates IkappaB in mononuclear cells in obese subjects: evidence for an anti-inflammatory effect?** *J Clin Endocrinol Metab* 2001, **86**:3257-3265.
263. Bowie A, O'Neill LA: **Oxidative stress and nuclear factor-kappaB activation: a reassessment of the evidence in the light of recent discoveries.** *Biochem Pharmacol* 2000, **59**:13-23.
264. Wada T, Penninger JM: **Mitogen-activated protein kinases in apoptosis regulation.** *Oncogene* 2004, **23**:2838-2849.
265. Nishimoto S, Nishida E: **MAPK signalling: ERK5 versus ERK1/2.** *EMBO Rep* 2006, **7**:782-786.
266. Begaud E, Chartier L, Marechal V, Ipero J, Leal J, Versmisse P, Breton G, Fontanet A, Capoulade-Metay C, Fleury H, et al.: **Reduced CD4 T cell activation and in vitro susceptibility to HIV-1 infection in exposed uninfected Central Africans.** *Retrovirology* 2006, **3**:35.
267. Jennes W, Evertse D, Borget MY, Vuylsteke B, Maurice C, Nkengasong JN, Kestens L: **Suppressed cellular alloimmune responses in HIV-exposed seronegative female sex workers.** *Clin Exp Immunol* 2006, **143**:435-444.
268. Lori F: **Treating HIV/AIDS by reducing immune system activation: the paradox of immune deficiency and immune hyperactivation.** *Curr Opin HIV AIDS* 2008, **3**:99-103.
269. Speer CP, Pabst MJ, Hedegaard HB, Rest RF, Johnston RB, Jr.: **Enhanced release of oxygen metabolites by monocyte-derived macrophages exposed to proteolytic enzymes: activity of neutrophil elastase and cathepsin G.** *J Immunol* 1984, **133**:2151-2156.
270. Nahori MA, Renesto P, Vargaftig BB, Chignard M: **Activation and damage of cultured airway epithelial cells by human elastase and cathepsin G.** *Eur J Pharmacol* 1992, **228**:213-218.
271. Princiotta MF, Finzi D, Qian SB, Gibbs J, Schuchmann S, Buttgereit F, Bennink JR, Yewdell JW: **Quantitating protein synthesis, degradation, and endogenous antigen processing.** *Immunity* 2003, **18**:343-354.
272. Takada Y, Ichikawa H, Badmaev V, Aggarwal BB: **Acetyl-11-keto-beta-boswellic acid potentiates apoptosis, inhibits invasion, and abolishes osteoclastogenesis by suppressing NF-kappa B and NF-kappa B-regulated gene expression.** *J Immunol* 2006, **176**:3127-3140.

Appendix A: List of abbreviations

ACAD	Activated cell-autonomous death
AFAF	v-raf murine sarcoma 3611 viral oncogene homolog
AICD	Activation-induced cell death
AIDS	Acquired Immune Deficiency Syndrome
AKBA	Acetyl-11-Keto-β-Boswellic Acid
ANOVA	Analysis of variance
APAF-1	Apoptotic peptidase activating factor-1
APC	Antigen presenting cell
APOBEC3G	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G
ARAF	v-raf murine sarcoma 3611 viral oncogene homolog
ART	Antiretroviral therapy
ARV	Antiretroviral
ATF4	Activating transcription factor 4
ATP	Adenosine triphosphate
BAK	Bcl2-antagonist/killer 1
BAX	Bcl2-associated X protein
Bcl	B cell leukemia
$BclX_L$	Bcl-like 1
Bp	Base pair
BRAF	v-raf murine sarcoma viral oncogene homolog B1
BSA	Bovine serum albumin
Cat. No.	Catalogue number

CCR5	Chemokine (C-C motif) receptor 5
CD	Cluster of differentiation
cDNA	Complimentary DNA
CEF	Cytomegalovirus, Epstein Barr and influenza
CFLAR	Caspase 8 and FADD –like apoptosis regulator
cFLIP	Cellular FLICE-(caspase 8)-like inhibitory protein
CHUK	Conserved helix-loop-helix ubiquitous kinase
CICD	Caspase-independent cell death
CMC	Cervical mononuclear cell
CRFs	Circulating recombinant forms
cRNA	Complimentary RNA
CTL	Cytotoxic lymphocyte
CVL	Cervicovaginal lavages
CXCR4	Chemokine (CXC motif) receptor 4
DAVID	Database for Annotation, Visualization and Integrated Discovery
DC	Dendritic cell
DFFB	DNA fragmentation factor B
dH_2O	Distilled water
DIABLO	Direct IAP-binding protein with low pI
DISC	Death-Inducing Signalling Complex
°C	Degrees Celsius
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid

dNTP	Deoxyribonucleotide triphosphate
DUSP5	Dual specificity phosphatase 5
EASE	Expression Analysis Systematic Explorer
e.g.	Example given
Env	Envelope
ERK	Extracellular signal-regulated kinases
ES	Enrichment score
ETC	Electron transport chain
EtOH	Ethanol
FADD	Fas-associated death domain
$FADH_2$	Flavin adenine dinucleotide
FC	Fold change
FCS	Fetal calf serum
FDR	False Discovery Rate
FGT	Female genital tract
FU	Fluorescence units
Gag	Group specific antigen
gDNA	Genomic DNA
Gp	Glycoprotein
GPCRs	G protein-coupled receptors
GSEA	Gene Set Enrichment Analysis
HAART	Highly Active Antiretroviral Therapy
HCB	Humidity Control Buffer

HCl	Hydrochloric acid
HEPS	Highly exposed, persistently seronegative
HIV	Human Immunodeficiency Virus
HIV+	HIV-positive
HLA	Human Leukocyte Antigen
HPV	Human Papillomavirus
HSV-2	Herpes simplex virus type 2
HYB	Hybridization buffer
IAPs	Inhibitors of apopotosis
i.e.	*id est*
IEL	Intraepithelial lymphocyte
IFN	Interferon
IgA	Immunoglobulin A
IgM	Immunoglobulin M
IgG	Immunoglobulin G
IL10RB	Interleukin 10 receptor beta
IL13RA1	Interleukin 13 receptor alpha 1
IL1R1	Interleukin 1 receptor 1
IL1R2	Interleukin 1 receptor 2
IRF	Interferon regulatory factor
IRF	Interferon regulatory factor
ITGAL	Integrin, alpha L
IVT	*In-vitro* transcription

JNK	c-Jun N-terminal kinase
Kb	Kilobase pair
KCl	Potassium chloride
KEGG	Kyoto Encyclopaedia of Genes and Genomes
KH_2PO_4	Potassium dihydrogen phosphate
LIF	Leukocyte inhibitory factor
LPS	Lipopolysaccharide
LTNP	Long term non-progressor
LTR	Long terminal repeat
MAP3K	MAPKK Kinase
MAPK	Mitogen-activated protein kinase
MAPKK	MAPK Kinase
MCL1	myeloid cell leukemia sequence 1
mg	Milligrams
MHC	Major Histocompatibility Complex
Min	Minute
MIP-1α	Macrophage inflammatory protein-1 alpha
MIP-1β	Macrophage inflammatory protein-1 beta
ML	Malaya
ml	Millilitre
mRNA	Messenger ribonucleic acid
MSigDB	Molecular Signatures Database
MyD88	Myeloid differentiation primary response gene 88

N9	Nonoxynol-9
Na_2HPO_4	Sodium phosphate
NaCl	Sodium chloride
NADH	Nicotinamide adenine dinucleotide
Nef	Negative regulation factor
NES	Normalized expression score
NFκB	Nuclear factor kappa-light-chain-enhancer of activated B cells
ng	Nanogram
NK	Natural killer
NN	New negative
NNRTI	Non-nucleoside reverse transcriptase inhibitor
NRTI	Nucleoside reverse transcriptase inhibitor
OSM	Oncostatin-M
PAMP	Pathogen-associated molecular pattern
p52	Protein 52
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
pH	Potential of hydrogen
PHA	Phytohemagglutinin
Phe	Phenylalanine
PI	Protease Inhibitor
PP2A	Protein phosphatase 2A

PUMA	p53 up-regulated modulator of apoptosis
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
r.c.f.	Relative centrifuge force
r.t.	Room temperature
RANTES	Regulated on Activation Normal T Cell Expressed and Secreted
Rev	Regulator of virion
RIN	RNA integrity number
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RRE	Rev-response elements
RT	Reverse Transcriptase
s	Seconds
Sec	Second
SIV	Simian Immunodeficiency Virus
ssRNA	Single-stranded RNA
STI	Sexually Transmitted Infection
Tat	Transactivator
TE	Tris- ethylenediaminetetraacetic acid
T_H	T helper
TIR	Toll/interleukin-1
TIRAP	TIR domain-containing adaptor protein
TLR	Toll-like receptor
TNF(α)	Tumor necrosis factor (alpha)

TNFR1	Tumor Necrosis Factor Receptor 1
TRAIL	TNF-related apoptosis inducing factor
TRIF	TIR-domain-containing adapter-inducing interferon-β
TRIM5α	Tripartite motif-containing 5
Trp	Tryptophan
Tyr	Tyrosine
TZ	Transformation zone
ul	Microlitre
UNAIDS	United Nations Programme on HIV/AIDS
UQCRFS1	Ubiquinol-cytochrome c reductase iron-sulphur subunit
v	Version
VEGF	Vascular endothelial growth factor
VHR	Vaccinia H1-related
Vif	Viral infectivity factor
Vpr	Viral protein r
Vpu	Viral protein u
XIAP	X-linked inhibitor of apoptosis

Appendix B: HIV-1 peptide sequences

Number	Sequence	Length
1	MGARASVLSGGKLDA	15
2	SVLSGGKLDAWEKIR	15
3	GKLDAWEKIRLR	12
4	LDAWEKIRLRPGGKK	15
5	KIRLRPGGKKKYRLK	15
6	PGGKKKYRLKHLVWA	15
7	KYRLKHLVWASREL	14
8	KHLVWASRELERFAL	15
9	ASRELERFALNPGLL	15
10	ERFALNPGLLETSEG	15
11	NPGLLETSEGCQQII	15
12	ETSEGCQQIIGQL	13
13	CQQIIGQLQPAL	12
14	QIIGQLQPALQ	11
15	IIGQLQPALQTGTEEL	16
16	PALQTGTEELRSLY	14
17	TGTEELRSLYNTVA	14
18	ELRSLYNTVATLYCV	15
19	YNTVATLYCVHQKI	14
20	ATLYCVHQKIEVK	13
21	YCVHQKIEVKDTKEA	15

22	KIEVKDTKEALDKI	14
23	KDTKEALDKIEEEQNK	16
24	LDKIEEEQNKSQQK	14
25	EEEQNKSQQKTQQAA	15
26	KSQQKTQQAAADTGN	15
27	TQQAAADTGNSSQV	14
28	AADTGNSSQVSQNY	14
29	GNSSQVSQNYPIV	13
30	QVSQNYPIVQNL	13
31	IVQNLQGQMVHQAI	14
32	LQGQMVHQAISPRTL	15
33	VHQAISPRTLNAWVK	15
34	SPRTLNAWVKVIEEK	15
35	NAWVKVIEEKAF	12
36	WVKVIEEKAFSPEVI	15
37	EEKAFSPEVIPMFSA	15
38	SPEVIPMFSALSEGA	15
39	PMFSALSEGATPQDL	15
40	PQDLNTMLNTVGGH	14
41	GHQAAMQMLKDTI	13
42	AAMQMLKDTINEEAA	15
43	LKDTINEEAAEWDRL	15
44	NEEAAEWDRLHPVHA	15

45	EWDRLHPVHAGPI	13
46	IPPGQMREPRGSDIA	15
47	MREPRGSDIAGTTSTL	16
48	SDIAGTTSTLQEQI	14
49	STLQEQIGWMTSNPPI	16
50	IGWMTSNPPIPVGEI	15
51	SNPPIPVGEIYKRWI	15
52	PVGEIYKRWIILGL	14
53	IYKRWIILGLNKIVR	15
54	IILGLNKIVRMYSPV	15
55	NKIVRMYSPVSILDI	15