Cellular Proteomic Analysis of Highly Exposed HIV Seronegative Women from the Pumwani Sex Worker Cohort in Nairobi, Kenya

by

Derek Riley Stein

A thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

In partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Department of Medical Microbiology

University of Manitoba

Winnipeg

Abstract

The HIV pandemic is well into its fourth decade and researchers have yet to develop a protective vaccine. The research presented in this thesis aims to identify and characterize correlates of HIV protection by studying highly exposed HIV seronegative (HESN) sex workers from Nairobi, Kenya. HESN women appear to have the natural ability to resist HIV acquisition giving the research community some hope that one-day a vaccine, microbicide or other prevention tool can be developed. Previous studies have indicated an overall immune quiescence phenotype characterized by lower immune activation in T cells and increased factors that limit HIV infection, which may play a significant role in protecting HESN women from HIV acquisition. The central hypothesis of this thesis is that quantitative shotgun proteomics of systemic and mucosal mononuclear cells from HESN women will demonstrate a proteomic profile characteristic of the immune quiescent phenotype, described by the altered expression of pathways important in immune activation, cell recruitment / migration, and proteins involved in HIV replication pathways. This hypothesis was addressed by quantifying the proteomic profile of immune cells isolated from the systemic and mucosal compartments of HESN women. In these studies HESN women were shown to have a proteomic profile representative of immune quiescence in genital tract immune cells but not systemically. Additionally, Mx2 a novel HIV restriction factor was found to be over-expressed in systemic immune cells from HESN women. These data support the role of immune quiescence in the genital tract as a potential mechanism for HIV resistance and has identified a novel HIV restriction factor that may contribute to

HIV resistance in HESN women. Strategies to mimic immune quiescence at the site of HIV acquisition and regulate HIV host restriction factors such as Mx2 should be considered during future vaccine and microbicide development.

Dedication

This dissertation is dedicated to my son, Riley. You have inspired me in ways that you may never understand. The day that you came into my life was the happiest day of my life. I look forward to watching you grow and achieve your dreams.

Acknowledgements

First and foremost I would like to thank my supervisors Drs. Blake Ball and Frank Plummer for giving me the opportunity to learn and conduct research in the HIV group. I would especially like to thank both of you for giving me the freedom to explore my own research questions and expertise over the years. It may have taken me a while but I think with your guidance I have finally gained the confidence and skills to become a successful research scientist.

I would also like to thank my committee members, Dr. Xi Yang, Dr. Michael Carpenter, and Dr. Neeloffer Mookherjee. Your feedback over the years has always aimed to keep my research on track and focused and will always be greatly appreciated.

I would like to thank the department of medical microbiology for all the administrative support and opportunities given to me to participate in departmental governance as a student representative. A special thanks to Angela Nelson, for keeping me up to date on all things pertaining to the Faculty of Graduate studies no matter how many of her emails I deleted. Additionally the support and technical staff: Leslie, John, Sue, Steve, Jude, Christine, and Max were essential in the completion of this thesis research. I would also like to thank Mike Carpenter for always taking the time to talk science no matter what the topic and no matter how busy you were. A big thanks goes to my fellow graduate students Phil, Melissa, Meika, Caitlin, Were, Aida and the rest of the HIV group who's commiseration over the years has kept me going no matter how many failed experiments I was faced with.

To Garrett Westmacott and the proteomics core: you have been in my corner from day one and given me unprecedented access to all the state of the art equipment and support the NML has to offer, I owe you a big thank you. Without your unwavering commitment this thesis would not be possible.

Last, but certainly not least, to my parents (Ken and Pat) and the rest of my family: I am so thankful to have all of you supporting me throughout the years and all the opportunities you gave me. Finally, I would like to thank my son Riley for reminding me of what is truly important in life, and for inspiring me to continue to learn and grow now matter what age you are.

Table of Contents

List of Tables

List of Figures

List of Copyrighted Material / Permission Obtained

Chapter 1. Introduction

1.1 The HIV Pandemic

Since the beginning of the pandemic the Human Immunodeficiency Virus-1 (HIV-1) has infected approximately 70 million people and lead to the deaths of over 35 million globally. In 2013, the WHO estimated that 35.3 million people were living with HIV and that 2.3 million new HIV infections occurred that year alone [1]. While the rollout of treatment and prevention strategies has resulted in a 33% decrease in new infections since 2001 it is clear that the WHO will not meet its targets of a 50% reduction in new infections by 2015 [1]. Sub-Saharan Africa has taken the brunt of the epidemic accounting for 70% of people living with HIV and over 75% of the HIV related deaths worldwide. Even more striking is that the pandemic disproportionately affects women, who account for more than half of HIV infected individuals in sub-Saharan Africa. Globally, female sex workers are 13.5 times more likely to be infected with HIV than other women [2] and have an estimated prevalence rate of 36.9% in sub-Saharan Africa [3]. Clearly the world, and women in sub-Saharan Africa are in desperate need of HIV prevention strategies such as a vaccine or microbicide to halt the HIV epidemic. The aim of this thesis is to identify and characterize novel mechanisms of HIV protection working with an observational cohort in Nairobi, Kenya that may lead to new knowledge that can inform novel HIV vaccine or microbicide approaches.

1.1.1 Kenyan Epidemic

At the peak of the epidemic in Kenya HIV prevalence hit 13.4% in the year 2000 [4]. However, HIV prevalence in Kenya has dropped significantly over the last decade from 7.2% in 2007 to 5.6% in 2012 in adults aged 15 to 64 years old due in large part to the scale up of antiretroviral, behavioral, and circumcision prevention programs [5]. Recent scale up of voluntary circumcision in Kenya has also lead to the proportion of men circumcised going from 85% in 2007 to 91% in 2012 which should enhance the continual downward trend of HIV incidence in Kenya. However, prevention strategies such as a vaccine or microbicide would be the preferred approach in terms of sustainability, cost, and ease of deployment.

1.1.2 Origins of HIV-1

Even though there is no direct epidemiological evidence to indicate cross-species transmission of HIV, considerable phylogenetic analysis of HIV genomes has resulted in the general acceptance of the zoonotic origin theory. Strains of simian immunodeficiency virus (SIV) isolated from *Pan troglodytes troglodytes* chimpanzees in western equatorial Africa bear the closest resemblance to the current group M (main) epidemic strain of HIV-1 [6]. While the exact date of transmission to humans is unknown the evidence points to initial transmission to humans well before 1940 [7]. It is believed that SIV cross species transmission was likely facilitated by the hunting of primates for consumption (bush meat) leading to human exposure to SIV infected animal fluids either during the butchering process or from uncooked meat. Another competing hypothesis for the initial introduction

of HIV into humans involved contaminated oral polio vaccine preparations. However, this hypothesis has been refuted by numerous sources [7,8] and is extremely unlikely to be the source of human transmission. The rapid spread and dissemination of HIV across Africa and worldwide is likely due to industrialization of parts of the African continent as well as increasing global travel and trade [9].

The first cases of acquired immunodeficiency syndrome (AIDS) were reported in men who have sex with men (MSM) from Los Angeles and New York presenting with rare forms of Kaposi's sarcoma and Pneumocystis pneumonia in 1981 [10]. In the following years the virus HTLV-III, later renamed HIV-1 was cultured and isolated as the etiological agent causing AIDS [11,12]. Subsequently, as awareness and scale-up of HIV surveillance efforts grew, it was clear that a global HIV pandemic was on the rise. Initial studies discovered subgroups with high levels of prevalence and incidence. In fact, two thirds of commercial sex workers from Nairobi, Kenya were found to be infected with HIV-1 during the early stages of the pandemic [13].

1.2 HIV Virology and Pathogenesis

1.2.1 HIV Diversity and Structure

HIV-1 is a Retrovirus from the genus Lentivirinae and is the causative agent of AIDS [14]. Due in large part to the error prone replication of HIV there is considerable diversity in the geographic distribution of circulating HIV strains [15]. For example, HIV-1 is classified into 11 different subtypes within group M alone, which accounts

for the majority of infections world wide (A1, A2, B, C, D, F1, F2, G, H, J, and K) [16]. Circulating recombinant forms (CRF) of the virus are generated by super-infection and recombination of the main group subtypes further contributing to the overall diversity of HIV-1. While clade C dominates the epidemic, accounting for over 50% of HIV infections worldwide, clades A1, D, G and some CRF forms are predominant in Kenya [17,18]. In North America and Europe, clade B is the major circulating virus. This genetic diversity is a major reason why an effective vaccine has been difficult to develop. A potential vaccine would need to induce immunity across numerous clades or be tailored to a specific geographical region in order to have an effective impact on the pandemic.

The HIV-1 genome encodes for two-single stranded positive sense RNA molecules [19,20]. The genome consists of 9 open reading frames from which 15 viral proteins are produced (Figure 1.1). Three poly-proteins Gag, Pol and Env are produced following proteolytic cleavage by host cell enzymes to make up the key structural and replication proteins needed by the virus. There are four Gag proteins, which include MA (matrix; p17), CA (capsid; p24) and NC (nucleocapsid; p6 and p7). The nucleocapsid p6 and p7 proteins bind and protect the two ssRNA molecules while the capsid protein p24 forms a characteristic cone-shaped structure around the nucleocapsid. A second p17 matrix protein surrounds the capsid further protecting the virion. The Envelope protein (gp160) is cleaved into two proteins; gp120 and gp41. A trimer of gp120 is anchored to the membrane of the virus by a trimer of gp41 molecules making up the viral spike used for entry into target cells. The Pol

protein is cleaved into 3 important enzymes; reverse transcriptase (RT), protease (PR) and integrase (IN), all three of which are packaged within the capsid of the virus. RT is needed to reverse transcribe the RNA genome into DNA in preparation for integration into host DNA by IN, while PR is need for cleavage of the poly-proteins Gag-Pol in order to produce mature virions. There are also several accessory proteins some of which are packaged within the virus (Vif, Vpr, Nef) while the rest aid in viral replication (Tat, Rev) and assembly (Vpu).

1.2.2 HIV Life Cycle

The initial steps in HIV replication involve the virus entering its target cells. In the case of HIV, the $CD4^+$ T cell is the major target, however other CD4+ cells such as macrophages and dendritic cells can also be targeted. Using its envelope glycoprotein, gp120, HIV binds the CD4 receptor on susceptible cells. Use of a second chemokine co-receptor on target cells (CCR5 or CXCR4) is necessary to achieve gp41 mediated virion fusion with the cell membrane. The indispensable requirement for HIV binding to the CCR5 molecule has been well documented. A polymorphism in the host CCR5 molecule (CCR5Δ32) results in a defective CCR5 surface protein. This leads to the inability of the virus to infect target cells, effectively making the host relatively resistant to HIV infection. Additionally, in the event of infection, individuals with the CCR5Δ32 mutation have significantly delayed disease progression [21,22]. Once the viral core has gained entry to the cell, uncoating of the viral capsid occurs (within the cytoplasm) by processes that are relatively unclear [23]. Then, reverse transcription of the viral RNA occurs to

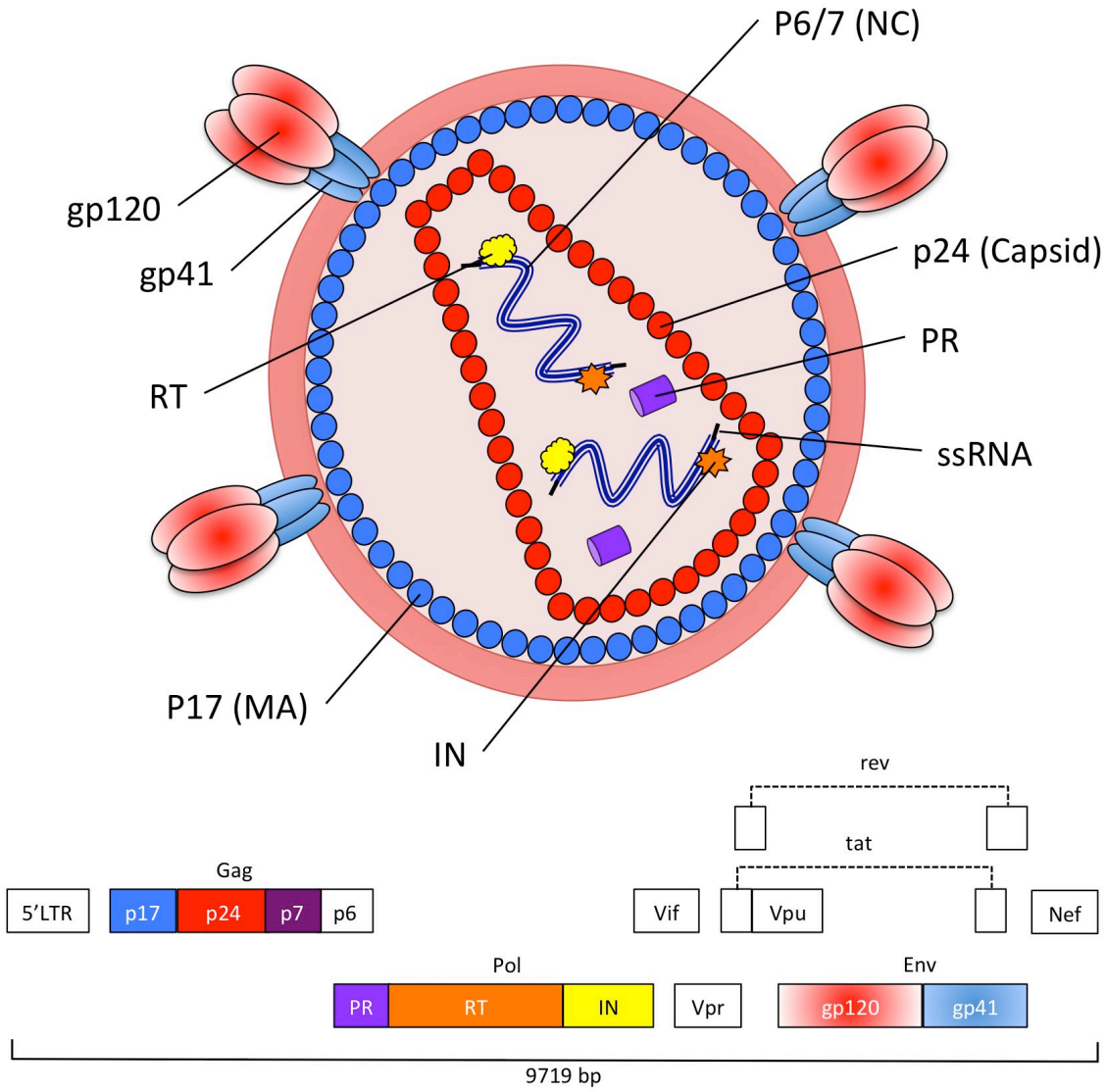


Figure 1.1. Structure and genomic makeup of HIV-1. The HIV-1 structure is adapted from [20].

generate double stranded DNA. It is this step where considerable genomic variation occurs, as the HIV reverse transcriptase is extremely error prone, introducing a mutation every $3.4x10^5$ base pairs [24]. Once reverse transcription is complete a pre-integration complex is formed consisting of IN, viral dsDNA, RT, MA, and Vpr [25]. This complex crosses the nuclear envelope and allows for integration of the virus into the host genome. In non-dividing cells the HIV genome is not transcribed, effectively hiding the virus from antiretroviral (ARV) drugs and the immune system, giving rise to a latent viral reservoir that current HIV treatment strategies are unable to eradicate. In the case of actively proliferating cells, HIV is efficiently transcribed via the 5´ long terminal repeat (LTR) which acts as a promoter/enhancer for virus replication. Viral mRNA's are transported to the cytoplasm where they are translated into viral poly-proteins and cleaved by viral / host proteases. These proteins are eventually shuttled and packaged at the host plasma membrane where new virions bud from the cell.

1.2.3 HIV Pathogenesis

1.2.3.1HIV Transmission

HIV is mainly transmitted through contact with bodily fluids. It is most often acquired by heterosexual or homosexual intercourse as well as through parenteral (injection drug use or blood transfusion) and vertical transmission (mother to child transmission thru breast milk or in-utero) [26]. However, the majority of new HIV infections worldwide are acquired through heterosexual intercourse. While transmission rates can be as low as 0.0082 per coital act, the rates are highly

dependent on stage of HIV disease, the current viral load of the infecting individual, as well as concomitant STI's, or genital infections [27,28]. Interestingly, placing a patient on ART has been suggested to lower HIV shedding in the genital tract which has been conclusively demonstrated to reduce HIV transmission rates [29]. This has led to the adoption of ART treatment as prevention (TasP) and HIV pre-exposure prophylaxis (PreP) as intervention strategies currently being examined in clinical trials [30].

1.2.3.2 HIV Disease Progression

There are essentially two main phases of HIV infection, an initial acute phase (weeks) and a prolonged chronic phase (years) that if left untreated, will eventually lead to the depletion of CD4+ cells and the development of AIDS. After a transmission event occurs, the first few weeks of infection set the stage for a gradual depletion of immune function that ultimately culminates in the onset of AIDS. The acute phase is characterized by a short eclipse phase when systemic infection is established followed by a large initial burst in plasma viremia (viral load). Viremia is rapidly brought under control by the host immune system, and is thought to be primarily mediated by the onset of HIV-specific CD8+ CTL responses (Figure 1.2) [31]. In addition, during acute infection there is a marked drop in systemically circulating CD4+ T-cells, which eventually rebounds after viral replication is contained. The nadir set point where plasma viral load stabilizes is termed the viral set point. This is important in that a lower viral set point has been associated with improved

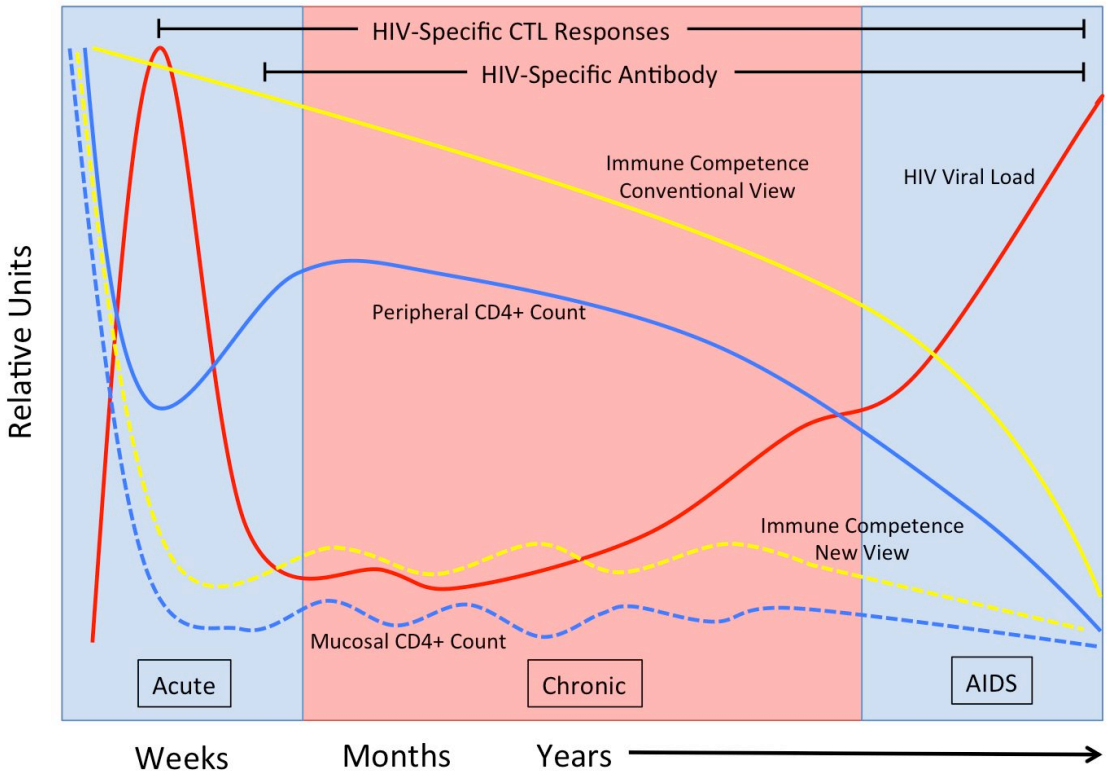


Figure 1.2. HIV pathogenesis and disease progression. Markers of disease course including peripheral and mucosal CD4+ T cell counts, HIV viral load, immune competence, and the onset of adaptive immune responses during the acute and chronic stages of HIV infection. Adapted from [32,33].

disease outcome [33,34] and is an important indicator of vaccine efficacy in T cell based HIV vaccine strategies. Newer observations in SIV infected rhesus monkeys [35,36] and later in humans [37,38] indicated that the primary phase of HIV infection drastically affects the gut associated lymphoid tissue (GALT) where irreversible CD4+ T cell depletion occurs early on in acute infection. By targeting CD4+CCR5+ T cells in the gut, HIV has a massive reservoir of target cells from which to replicate. The initial spike in viral replication during acute infection combined with HIV error prone replication also allows for rapid escape from adaptive immune responses during acute infection. These observations have significantly changed the traditional view of HIV pathogenesis from slow progressive loss of immune function over several years to one where the initial damage occurs early in acute infection and the host never quite recovers, leading to eventual immune failure [32]. The chronic phase is now considered to be a period of increasingly less-than effective immune function where CD4+ T cell turnover is extremely high, eventually leading to exhaustion and a steady decline in systemic CD4+ T cells over several years. In untreated individuals, once CD4+ T cell counts fall below 200 cell/μl of blood, the immune system is unable to deal with opportunistic infections and ongoing immune assaults, leading to AIDS and eventual death [39,40].

1.2.3.3 Atypical HIV Disease Progression

In some instances HIV disease progression does not follow the typical course described above and results in delayed or accelerated progression to AIDS. The former is a model of success against HIV and provides hope that by identifying the

mechanisms behind delayed progression this may inform the type of immune responses required for successful vaccine strategies. Long term non-progressors (LTNP) are able to maintain stable CD4+ T cell counts (typically above 500 cells/ul) and as a result do not progress to AIDS as quickly as the average person. In some cases elite controllers (EC) not only maintain high CD4+ counts but also are able to suppress viral replication to under 50 RNA copies/ml [41,42]. The study of rapid HIV progressors (RP) who are defined by a rapid CD4+ T cell decline to <300 cells/ml within 3 years of being infected also serve as a model in which to identify undesirable host responses during HIV infection [41]. Epidemiological models of immune success and failure in HIV infection such as these can inform future vaccine and microbicide design.

1.3 HIV Vaccine and Microbicide Prevention Trials

The most successful HIV vaccine clinical trial to date was the RV144 "Thai trial". All other large scale (Phase III) vaccine trails have either failed to be efficacious [43,44] or were halted early in the case of the phase IIb STEP trial (Ad5 HIV-1 gag/pol/nef) due to a possible increased risk of HIV infection in vaccinees [45]. The RV144 trial showed a 31.2% protective effect in the modified intent-to-treat analysis [46]. The partial protective effect induced by this vaccine is still not fully understood, however some post-trial studies have indicated non-neutralizing antibodies may be playing a partial protective role, reigniting calls for antibody based HIV vaccine strategies [47]. In addition, the discovery of new broadly neutralizing [48-50] antibodies has driven a new "vaccine gene-therapy" approach where antibodies are

delivered by adeno-associated virus (AAV) gene therapy, side stepping the need for the immune system to develop these extremely rare antibodies [51,52]. These failures and moderate successes in the HIV vaccine field have highlighted the need for further studies that clearly define correlates of HIV protection.

A microbicide is generally described as a topical gel or cream that can be applied to either the genital or rectal compartments before sexual intercourse providing protection from HIV. There have also been some moderate successes on the microbicide front with the CAPRISA 004 trial, which contained a topically applied vaginal gel with 1% Tenofovir (antiretroviral). The CAPRISA trial reduced HIV acquisition by 39% while women with high adherence (>80%) were protected by up to 54% [53]. The CAPRISA trial results suggest that, while promising, these approaches (similar to condoms) ultimately rely on personal adherence and that other approaches that do not depend on individual adherence like a vaccine may also be needed. Other preventative methods have also showed some success in clinical trials, the most successful intervention to date being circumcision. Circumcision is believed to reduce the number of target cells that would otherwise be present in un-circumcised men, as well as increased keratinization of the glans penis making it difficult for virus to penetrate [54,55]. Three trials in Kenya, Uganda, and South Africa were halted early, all due to success showing over 50% protection from HIV acquisition [56-58]. While a microbicide, like condom use is clearly going to require strict adherence, the challenges with circumcision

acceptance and scale up are going to be key issues going forward making a protective vaccine an important goal in the HIV prevention field.

It is clear that the HIV prevention field is finally gaining traction with these moderately successful trials however a fully effective single HIV prevention strategy is still desperately needed. Vaccine trials to date have yet to clearly define strong correlates of HIV protection, and an understanding of the host immune response is critically needed for developing additional protective approaches against HIV acquisition.

1.4 Basic Immunology

There are two main arms that make up the hosts immune response to invading pathogens, the innate and adaptive immune systems. Innate or natural immunity lacks immunological memory and essentially has the same response to pathogen challenge every time. The adaptive immune system learns from repeated exposure to a pathogen, improving over time and is highly specific.

1.4.1 Innate Immunity

There are three main components that make up the human innate immune system which include physical barriers, cellular mediators and soluble factors. The main physical barriers to invading pathogens are the skin and mucous membranes in the body. Mucus traps foreign bacteria, viruses, and objects non-specifically which are then cleared from the body. When pathogens get through this defense, several

specialized immune cells can engage them. Macrophages and neutrophils (phagocytes) engulf and internally digest pathogens. They are also capable of recognizing cells coated in antibody (generated by the adaptive immune response) or complement proteins enhancing phagocytosis [59]. Dendritic cells (DC) facilitate the link between the innate and adaptive immune system. DCs and other cell types can become activated through pattern-recognition receptors (PRR) such as toll-like receptors (TLR) present intracellularily and on the surface of the cell. TLRs can recognize a wide range of microbial patterns such as lipopolysaccharide (TRL4), dsRNA (TLR3), flagellin (TLR5), ssRNA (TLR7) and unmethylated CpG DNA (TLR9) [60]. Once activated DCs become antigen presenting cells (APC) and present antigen on their surface using major-histocompatibility-complex (MHC) molecules. T and B cells can recognize this antigen and become activated and/or proliferate, producing cells specifically capable of recognizing foreign antigens. Natural killer (NK) cells are also another important component of the innate response and are able to recognize and kill infected cells by releasing perforin and granzymes. NK cells recognize antibody (IgG) coated cells through Fc receptors (FcγR) and kill cells through a process called antibody dependent cell-mediated cytotoxicity (ADCC) or by the lack of MHC I molecules on the cell surface coupled with a balance between activating and inhibiting NK receptors on a target cell [59,61]. Finally, there are soluble factors such as complement that allow for antibody dependent and independent (classical and alternative pathways respectively) perforation of infected cell membranes through a cascade of complement proteins that bind the target cell [62]. Other soluble factors of the innate response such as cytokines can

enhance and activate nearby cells while chemokines can further attract immune cells to the site of infection. Interferons, such as IFN-α, IFN-β, and IFN-γ, released by infected or bystander cells, are also potent activators of various cellular pathogen restriction factors that promote viral resistance. A better understanding of innate immune responses amongst different epidemiological models of HIV protection and acquisition could lead to better microbicide or vaccine strategies.

1.4.2 Adaptive Immunity

1.4.2.1 Cellular Immunity

The adaptive immune system can be generally broken down into cellular and humoral immune responses. The cellular response consists of two major cells termed CD4+ T "helper" (T_H) cells and CD8+ cytotoxic T cells (CTL) [62]. The specificity of the adaptive immune response is due to the T and B cell receptors (TCR and BCR) on the surface of the respective cells [63]. These receptors are somatically generated and recombined to produce the extreme diversity needed to recognize any pathogen. TCR diversity has the theoretical potential to generate $1x10^{15}$ different receptors making them a highly diverse and adaptive weapon against constantly changing pathogens [63]. Antigen is presented to T cells by APCs via MHC class II molecules in conjunction with the correct co-receptor (CD80 or CD86). If the TCR is able to recognize the antigen, it signals the T cell to proliferate. This proliferation allows for the expansion of T cells specific to the pathogen. These T helper cells, as the name suggests are able to help prime B cells as well as CTL responses through the production of various cytokines. CD8+ cytotoxic T cells are

the effector arm of cell mediated immunity and are induced by antigen presenting cells in the context of MHC class I. MHC class I molecules are found on the surface of all nucleated cells. When cells become infected they present antigen from the virus or invading pathogen on the surface of the cell with MHC class I. CD8+ T cells with the correct TCR recognize the MHC I / antigen complex are induced to proliferate and kill the infected cells.

1.4.2.2 Humoral Immunity

The adaptive immune system also has a humoral arm consisting of B cells that produce antibodies. Antigen specific T_H cells can activate B cells through their B-cell receptor (BCR) or by exogenous antigen crosslinking membrane bound antigen specific antibodies [64]. Activation of B cells causes affinity maturation whereby antibody genes inside the cell are rearranged selecting for optimal antigen binding. B cells produce many different classes of antibody that include IgG1-IgG4, IgA1, IgA2, IgM, IgD and IgE [65]. These different classes are tissue and function specific. For example, IgA is the dominant class of antibody found at most mucosal surfaces while IgE is an antibody produced specifically to combat parasitic infections. B cell activation (either by CD4+ helper T cells or foreign antigen binding membrane bound antibody receptors) also stimulates the formation of plasma cells (antibody secreting cells) and memory B cells. A deeper understanding of the innate and adaptive immunological correlates of protection against HIV could be used to inform novel vaccine and microbicide approaches.

1.5 Models and Correlates of Natural Protection to HIV

1.5.1 Highly Exposed HIV-seronegative

There are many cases of individuals who are highly exposed to HIV on a regular basis but never become infected. Studies of these individuals have sought to identify correlates of protection against HIV in the hopes of applying this knowledge to the development of an HIV preventative vaccine or microbicide strategy. These cohorts are generally comprised of individuals with documented high risk behaviors which include discordant couples, men who have sex with men (MSM), injection drug users (IDU), infants born to HIV infected mothers, hemophiliacs and commercial sex workers (CSW) [66]. To date the genetic CCR5Δ32 polymorphism has been the only consistently replicated mechanism of HIV protection [21]. Additionally, this mutation is rare, and mostly limited to Caucasian populations [67].

For over 25 years, studies of highly exposed HIV-seronegative (HESN) sex workers in Nairobi, Kenya have been conducted to understand the correlates of HIV protection. The Pumwani sex worker cohort was established in 1984 to study the natural history of sexually transmitted infections, including HIV. Women, upon enrolling in the cohort, were at high risk of becoming HIV infected within the first three years of enrollment. In 1996 Fowke et al. made the observation that despite repeated exposure to HIV, 5-10% of the women never became infected with HIV (Figure 3) [68]. HESN women in the Pumwani cohort are HIV-negative by serology and PCR. They are also constantly active in sex work for over seven years and can be epidemiologically defined as relatively resistant to HIV infection. The identification of factors important in preventing HIV infection in HESN women from

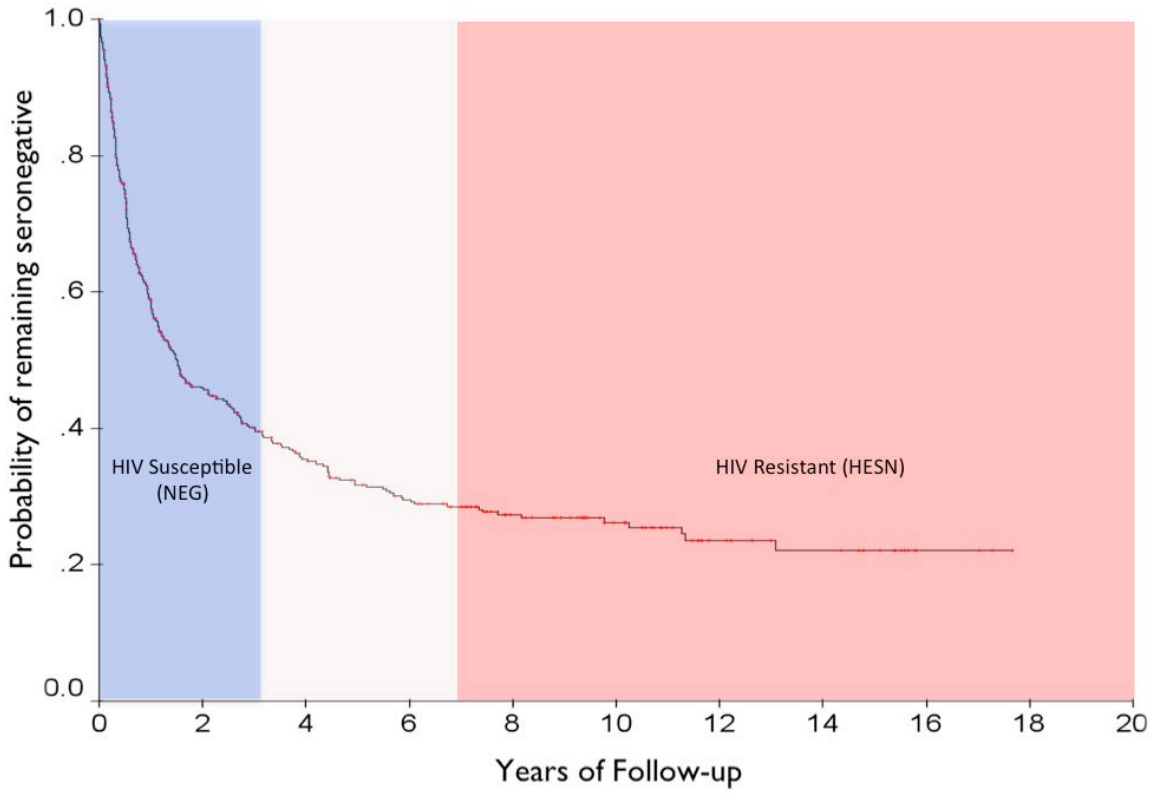


Figure 1.3. Kaplan-Meier survival analysis indicating reduced probability of HIV infection over time. In blue are the women defined as HIV susceptible (NEG) and are expected to seroconvert while HIV resistant (HESN) individuals are in red. Adapted with permission from Lancet [68].

the Pumwani cohort may lead to important insights into future vaccine and microbicide design.

1.5.2 Adaptive Immunity in HESN

Adaptive responses in HESN populations have been widely reported but their role in HIV resistance remains hotly disputed. Despite the clear protective and possible therapeutic effect of humoral antibodies to HIV/SIV in non-human primate studies, the protective role of antibody responses in HESN cohorts have yet to be well defined. In the Pumwani cohort, HIV-specific IgA has been identified as being over expressed in the genital tract of HESN women [69], however a second study in the same women found no difference in levels of HIV specific IgA [70]. In addition, no HIV-specific antibody responses were detected in a different Gambian cohort of HESN individuals [71]. However, several other HESN cohorts have identified HIV-specific IgA in the female genital tract (FGT) with the potential to neutralize HIV primary isolates as well as inhibit transytosis in-vitro [72-74]. These studies clearly indicate a lack of consensus surrounding HIV-specific antibody responses in HESN populations, and are in need of further study.

HIV-specific CD4+ and CD8+ T cell responses have been identified in the genital tract and the blood of HESN women from the Pumwani cohort [75-77]. The frequency of HIV-specific CD8+ T cells was found to be higher in HESN women than from HIV+ or HIV- negative controls in the genital tract [78]. However, these response do not appear to be sufficient for protection as late seroconversion in

HESN women has been documented [79]. It is likely that these responses may have some role to play in protection from disease progression since post-seroconversion pre-existing HIV-specific CD8+ T cell proliferative responses have been correlated with a lower viral set point [80]. Additionally, proliferative CD4+ T cell responses have also been associated with the HESN phenotype compared to HIV+ individuals. These studies indicate a possible qualitative and quantitative difference in the adaptive cellular responses in HESN individuals compared to susceptible controls [81]. Importantly, detection of IFN-γ responses (thought to be a correlate of a protective cellular response) did not correlate with HIV protection in HESN [81,82]. HIV-specific T cell responses have also been identified in other HESN cohorts including MSM [83,84] and serodiscordant couples [85,86]. However, it is now known that polyfunctional T cell responses need to be evaluated in a more comprehensive manner in these studies instead of using reductionist approaches that evaluate only singular measures of cellular responses [87]. Observations such as these highlight the lack of sensitivity in current immune assays for detecting correlates of protection in HESN individuals, making newer tools such as proteomic systems biology a promising alternative approach.

1.5.3 Genetic Correlates of HIV Protection

One of the best-characterized correlates of HIV protection is the CCR5Δ32 mutation. The expression of the non-functional HIV co-receptor molecule on the surface of HIV target cells inhibits the initial fusion of the HIV virus with the host cell, effectively blocking the establishment of a productive infection [21,22]. This mechanism has

lead to the only known case of an HIV cure. The "Berlin patient", as this case has come to be known, was a man who had end-stage acute myeloid leukemia and HIV infection. A bone marrow transplant was needed to treat the advancing cancer. Physicians decided to look for transplant matches that also had the CCR5Δ32 mutation. A subsequent stem cell transplantation with cells containing the CCR5Δ32 mutation has so far lead to the patient being HIV and cancer free [88,89]. Maraviroc, a new antiretroviral CCR5 fusion inhibitor has been developed in response to our understanding of this correlate of protection [90]. Increased knowledge of novel genetic correlates of protection such as the CCR5Δ32 mutation could lead to further advances in drug, vaccine, and microbicide development.

Polymorphisms in human leukocyte antigen (HLA) genes, which are important in adaptive immunity, have also been described in HESN individuals. Within the Pumwani cohort several HLA-DQ and DP polymorphisms have been associated with susceptibility and protection from HIV infection [91,92]. The HLA A2/6802 supertype has also been strongly associated with the HESN phenotype [93]. Additionally, certain allelic polymorphisms of killer immunoglobulin receptors (KIR) on NK cells have also been reported in HESN [94-96], specifically KIR3DS1 was shown to be over represented in a HESN population[95]. These data suggest that perhaps the adaptive immune response, which depends on HLA and KIR receptors, may contribute to the HESN phenotype.

Polymorphisms in the interferon regulatory factor 1 (IRF-1) gene have been described in the Pumwani cohort [97]. IRF-1 is an innate transcription factor that activates and suppresses target genes through binding of interferon stimulated response elements (ISRE). IRF-1 alleles (179 microsatellite, 619A, and 6516G) are associated with the HESN phenotype and lower levels of IRF-1 expression. Additionally, IRF-1 protective alleles allow for reduced long terminal repeat (LTR) transcription of HIV suggesting a mechanism for protection in these individuals [98]. These genetic factors could all be important targets for therapeutic and vaccine/microbicide prevention, however, host genetics does not account for all HESN women.

1.5.4 Innate Immunity in HESN

Studies on the innate arm of the immune system has produced some of the most interesting observations about HESN individuals and the potential for targeted HIV prevention approaches. A growing body of evidence suggests that only slight perturbations of the innate immune response in HESN individuals may account for greater HIV protection [99].

Altered innate TLR signaling in the PBMC of HESN individuals has been demonstrated by an increase in pro-inflammatory cytokines and chemokines in response to several TLR agonists, such as poly I:C, LPS, imiquimod, and ssRNA40 [100]. While this study also found relatively stable levels of TLR receptors between HESN and controls at the systemic level, another study of cervical mononuclear cells

(CMC) isolated from the genital tract in HESN women from the Pumwani cohort described lower basal levels of some TLR's with increased responsiveness to ssRNA40 [101]. These observations indicate that innate genital tract responses may be more important when identifying the correlates of HESN protection.

Alternative expression of secreted innate factors has also been observed in HESN individuals and may play an important role in innate resistance to HIV infection. The over-expression of MIP-1α, MIP-1β, and regulated upon activation normal T cell expressed and secreted (RANTES) have been demonstrated in HESN women to play an important role in competing for CCR5 co-receptor binding [102,103]. Additionally, other β-chemokines have been associated with resistance to HIV infection in salivary secretions of MSM populations [104]. In the Pumwani cohort, RANTES was also found to be over-expressed in the FGT of HESN [105], however other β-chemokine levels were comparable with controls [106]. Interluekin-22 was also found to be over-expressed in a different HESN population through serum amyloid A mediating the down-regulation of CCR5 and reducing susceptibility to HIV infection [107]. These studies provide additional evidence that HESN populations may be an important tool in confirming known correlates of protection (CCR5) and the possible discovery of new correlates that could be targeted for further intervention strategies.

The importance of the genital mucosa has led to several proteomic studies of genital secretions in HESN women from the Pumwani cohort. Trappin 2 / ELAFIN, an anti-

infammatory / anti-protease has been shown to be overexpressed in HESN women and has been shown to reduce HIV infection in-vitro [108-110]. Other anti-proteases such as cystatin A, the Serpin B family, and alpha-2 macroglobulin like-1 have also been shown to be overexpressed in the genital secretions of HESN women from the Pumwani cohort, which may play a role in their anti-HIV innate immunity [111,112]. These studies have discovered valuable candidates for future microbicide interventions, however, to date, the expression of intracellular factors in immune cells from HESN women have yet to be fully examined.

1.5.5 Innate HIV Cellular Restriction Factors in HESN

There are several innate HIV restriction factors that have been identified to have a significant impact on intracellular HIV replication in-vivo. These include APOBEC, TRIM5-α, BST-2/Tetherin, SAMHD1, and the newly discovered Mx2 / MxB protein [113,114]. APOBEC3G is a cytidine deaminase that is packaged into HIV particles in the absence of the Vif accessory protein, facilitating deamination of negative strand viral cDNA. APOBEC3G has been identified to be significantly over-expressed systemically as well as in the genital tract tissue of some HESN individuals [115]. TRIM5-α is a restriction factor that affects the capsid un-coating process through a Cyclophilin A / HIV capsid interaction. A single nucleotide polymorphism (SNP) in exon 2 of TRIM5-α of HESN women from the Majengo cohort has been associated with HIV protection [116]. BST-2 / tetherin, as the name suggests, tethers newly budding HIV virus to the cell inhibiting further infection of neighboring cells. However, in the presence of the HIV accessory protein Vpu HIV is successfully

released from infected cells [117]. SAMHD1 depletes nucleotides available for HIV reverse transcription thus blocking HIV replication and is antagonized by the HIV viral protein Vpx [118]. Finally, the newly discovered interferon induced restriction factor MxB has been proposed to inhibit integration of HIV provirus, however due to its presence in both the cytoplasm and nuclease, the mechanism of action is still unclear [119-121]. Not all restriction factors are over-expressed in HESN populations. The expression of APOBEC3G, TRIM5-α and Tetherin have been found not to associate with the HESN phenotype in a Senegalse cohort [122]. The expression of SAMHD1 and MxB has yet to be evaluated in the context of the HESN phenotype. Thus it appears some HESN populations overexpress factors that have the intrinsic ability to limit HIV replication however, the variability in intracellular restriction factors and the scarce amount of literature on the subject warrants further studies.

1.5.6 Immune Quiescence in HESN

The large number of correlates of protection identified in HESN women from the Pumwani cohort has recently been coalesced into a single overarching hypothesis that may explain the HESN phenotype. Several studies in HESN women from the Pumwani cohort have lead to an interest in the role of immune activation in the context of HIV acquisition. RNA expression analysis by micro-array analysis of CD4+ T cells isolated from HESN women indicated an overall decreased immune activation profile [123]. In the same study several HESN women produced significantly less pro-inflammatory cytokines at baseline compared to negative

controls. A second micro-array analysis of whole blood has also confirmed this gene expression pattern in HESN women [124]. An analysis by flow cytometry further indicated reduced expression of activation markers (CD69) on the surface of T cells as well as an elevated frequency of T-regulatory cells in HESN women [125]. While the proteome of immune cells from the systemic and mucosal compartments has never been characterized, proteomic analysis of cervical secretions from HESN women indicated high levels of anti-proteases which have anti-HIV and anti-inflammatory properties [111,112]. These observations have lead us to believe that HESN women posses an immune quiescent profile. This profile is represented by altered expression of pathways important in immune activation, cell recruitment / migration, as well as proteins involved in HIV replication processes. As this immune quiescent phenotype appears to be very complex and involves numerous, innate, adaptive and other immunologically relevant pathways, systems biology approaches may be a useful tool to identify correlates of protection in HESN women.

1.6 Mass Spectrometry and Quantitative Techniques

Thus far the majority of studies on HESN cohorts are mostly reductionist, resulting in a need for new approaches with a more global view of how systems interact. Systems Biology approaches are becoming an integral part of basic research and proteomics is a major tool that can be used in this approach. Proteomics is the study of the entire proteome within a given cell, tissue, or biological fluid which can contain any number of protein isoforms and post-translational modifications such as phosphorylation or glycosylation. Unfortunately, any one of todays mass

spectrometry techniques falls short in attaining the goal of studying the proteome in its entirety, leading to the establishment of multiple sub-fields such as phospho-proteomics, glyocomics, or metabalomics. RNA expression analysis thru micro arrays or newer RNA-Seq approaches has largely driven the majority of biomarker research and host response studies with respect to infectious diseases. However, RNA expression rarely correlates with protein expression directly, with correlations as low as r=0.27 allowing for substantial error in perceived host responses or biomarkers of infectious disease [126-131]. Approaches capable of examining proteins, which are the ultimate end product of gene expression, may offer a more accurate reflection of a particular biological state or phenotype such as that observed in HESN subjects.

The standard proteomics workflow is detailed in (Figure 1.4). Generally, cells, tissue or biological fluids of interest are harvested and lysed to release proteins. Isolated proteins are then digested in most cases with trypsin. The digested peptides are subjected to an "offline" fractionation method to reduce the sample complexity, which can include but is not limited to chromatography techniques such as strong cation exchange (SCX), strong anion exchange (SAX), reversed-phase (C18), and isoelectric focusing (IEF). Each fraction is then run through a second "online" fractionation typically involving a low-pH reversed-phase procedure, which further reduces the sample complexity allowing for direct injection into the mass spectrometer for analysis.

Mass spectrometers have substantially evolved in the last 10 to 15 years with four main types in use today. They are quadripole (Q), ion trap (IT), time of flight (TOF) and hybrid (Orbi) mass spectrometers. The hybrid instruments like the Orbitrap-VELOS or Q-Exactive (ThermoFisher) take advantage of two mass analyzers and a parallel data acquisition method to significantly increase the speed, and mass accuracy compared to conventional mass spectrometers. In large part the advances made by hybrid instruments have allowed for the identification and quantification of well over 4000 proteins in complex biological samples [132]. While mass spectrometers are capable of determining the masses of charged peptides, how those peptides are introduced into the MS for analysis is extremely important. Ionization of peptides is induced by two main methods, matrix assisted laser desorption (MALDI) or electrospray ionization (ESI). As the names of each technique suggest, MALDI makes use of a laser to ionize peptides embedded in a matrix while the ESI approach is achieved by applying a high voltage to a chromatography column with an emitter allowing for the ionization of peptides as they are eluted into the MS. The mass spectrometers mentioned along with an ionization method and various quantitative approaches are combined in different ways to achieve quantitative mass spectrometry (qMS).

One of the first quantitative proteomics techniques and still in use today is 2 dimensional-difference gel electrophoresis (2D-DIGE) [133,134]. 2D-DIGE is conducted by individually labeling control and experimental samples with florescent dyes (Cy2,Cy3, Cy5). The samples are mixed and then co-resolved first

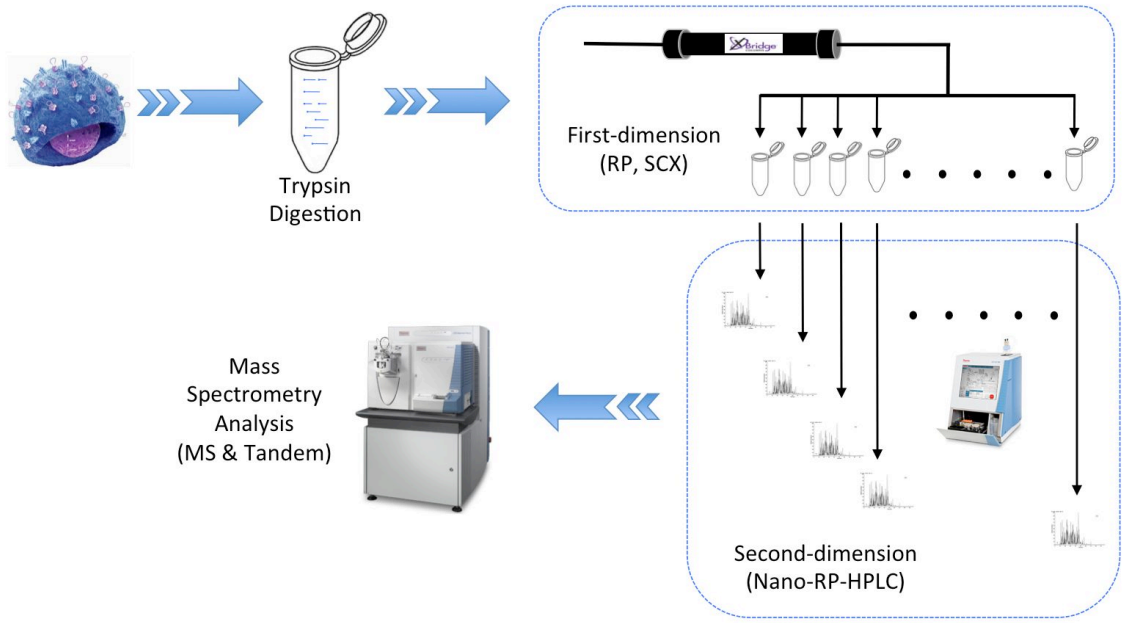


Figure 1.4. A general shotgun proteomics workflow is depicted for the identification and quantification of proteins from immune cells. Briefly, protein is isolated from cells and subjected to trypsin digestion. Peptides are fractionated using two different chromatography techniques and then analyzed by a mass spectrometer.

using a pI gradient (isoelectric focusing) followed by molecular weight separation on an acrylamide gel. Software is used to detect differentially labeled proteins between the different dyes and the protein spots are subsequently excised from the gel and identified by mass spectrometry. This technique is mainly limited to pair wise comparison and it can be difficult to extract proteins of interest for identification from acrylamide, particularly with respect to hydrophobic or low abundance proteins. While 2D-DIGE is still useful, the introduction of state-of-the-art hybrid mass spectrometers, has led to the technique falling out of common use.

More recent techniques such as stable isotope labeling of amino acids in cell culture (SILAC) which requires dividing cells in culture, allows for the incorporation of heavy labeled amino acids (typically arginine or lysine) into experimental cells [135]. Similar to 2D-DIGE, protein samples are mixed, however in the case of SILAC proteins are then digested (typically with trypsin), resolved by LC/MS/MS, followed by differential peptide analysis. Other labeling techniques such as differential ^{18}O labeling during peptide digestion are also employed. Although SILAC has been applied in multiplexed systems, both SILAC and ^{18}O qMS techniques are more often used in pair wise comparisons not lending themselves to the analysis of the multiple patient samples needed to address the proposed studies in this thesis.

Additional techniques involve either label-free quantitation [136] or post-trypsin digestion labeling which includes isotope coded affinity tags (iCAT) [137], isotope tags for relative and absolute quantitation (iTRAQ) [138], and tandem mass tags

(TMT) [139]. Label-free quantitation is one of the newest techniques to date and is achieved by resolving each biological sample sequentially on a high-resolution mass spectrometer. Each resulting chromatogram is aligned based on peptide mass and retention. Overlapping peptide intensities are then used to perform a relative quantitation measurement between biological samples. Theoretically, there is no limit to the number of samples that can be analyzed however, software capabilities to align data coupled with retention time drift due to column wear can significantly limit the practicality of this technique. iCAT and the iTRAQ/TMT labeling techniques make use of isobaric tags to label digested peptides. iCAT labels the cysteine residues of peptides however not all peptides contain cysteine and as a result this method was rapidly replaced in favor of the iTRAQ and TMT approaches [140]. iTRAQ and TMT, which allow for the labeling of Lysines and the N-termi of peptides would theoretically allow for complete labeling of all tryptic peptides in a sample. iTRAQ can allow for the labeling of up to 8 individual samples or conditions that can be analyzed in a single mass spectrometry run. Even though each iTRAQ tag is isobaric (identical in weight) this high multiplexing capacity is achieved through the addition of a reporter group. Each reporter group has a different mass and can be differentially detected by the mass spectrometer during MS/MS delineating which sample the peptide came from. In addition, the intensity of this reporter group allows for the relative quantitation of the peptide compared to a control sample. The flexibility and multiplexing capability of iTRAQ allows for the identification and relative quantitation of thousands of proteins within a given biological sample. A detailed list of the most relevant qMS techniques in the field

Technique	Advantages	Disadvantages	Selected References
2D-DIGE Two-dimensional difference gel electrophoresis	Samples are processed ex-vivo & Proteoform information may be retained	Spot identification is challenging	Unlü et al.1997[133]; Minden et al.2012[134]
SILAC Stable isotope labeling by amino acids in cell culture	Stable incorporation of heavy labeled amino acids into all proteins	Generally limited to dividing cells in cell culture	Ong et al. 2002[141],2012[142]
Stable ^{18}O Labeling	Samples can be labeled ex-vivo	Limited to pair-wise comparisons	Capelo et al.2010[143]
ICAT Isotope coded affinity tags	Samples can be labeled ex-vivo	Not all peptides contain cystine residues, targeted by ICAT	Gygi et al.1999[137]
iTRAQ Isobaric tags for relative and absolute quantitation	Multiplex of 8 samples into a single experiment	Co-selection of peptides can lead to underestimation of peptide abundance	Ross et al. 2004[138]; Thompson et al.2003[139];Christoforou et al.2012[144]
Label - Free Quantitation	No labeling required	Alignment and normalization requires sophisticated software	Zhu et al.2010[136]
MRM/SRM Multiple/Selective Reaction Monitoring	Absolute targeted quantitation & High through-put analysis	Difficult to identity consistent peptides to track for quantitation	Jaffe et al.2008[145]; Pan et al.2009[146]

Table 1.1. Potential quantitative proteomic techniques for profiling host immune responses and biomarkers of interest.

along with strengths and weaknesses are described in Table 1. Based upon these strengths and weaknesses iTRAQ was the technique chosen to characterize immune cells using a systems biology approach from HESN women (Chapters 6-8).

1.7 Mass Spectrometry in HIV Infection and Pathogenesis

In recent years qMS has developed to the point where it can be reliably used to identify markers of protection, pathogenesis and host responses to infection. However, perhaps more importantly, these MS proteomic technologies have the potential to increase our understanding of the interplay between HIV and host ex-vivo responses. Several studies have made use of qMS methodologies to study HIV in the context of in-vitro experimentation [147-151], however until recently MS has only been sporadically used to study HIV / host interactions and identify markers of HIV protection.

1.7.1 Proteomic Responses in HIV-infected individuals

qMS is now being used to broaden our understanding of critical HIV specific molecular events in the systemic and mucosal compartments. Considerable gaps remain with respect to understanding early acute phase HIV responses leading up to the establishment of chronic HIV infection and these new studies have the potential to direct future vaccine and microbicide strategies [152]. To this end, a study by Kramer et. al [153] was able to obtain plasma samples spanning the eclipse and expansion phases of HIV infection. Using MALDI-TOF label-free mass spectrometry they identified several novel classes of over-expressed proteins during acute HIV

infection including serum amyloid A; A-SAA, alpha-1-antitrypsin; AAT, and a C-terminal peptide derived from AAT. In particular A-SAA was induced by HIV infection several days before detectable viral RNA. These factors are all acute phase response proteins with antiviral properties, and again, it is possible that these may be part of the early host innate response to HIV. The other possibility of course, is that these responses are involved in HIV pathogenesis and their expression is induced by infection leading to eventual systemic dissemination. Clearly, additional proteomic and traditional studies like these will allow for a greater understanding of early innate responses that lead to HIV pathogenesis.

Another study making use of iTRAQ sought to determine if methamphetamine use affected inflammatory protein expression in the plasma of HIV infected patients [154]. Pottiez et. al. were able to quantify 390 distinct protein species in patient plasma of which 28 were shown to be differentially regulated. These proteins were not surprisingly involved in coagulation, complement and oxidative stress; all-important inflammatory markers that may lead to increased progression observed in these patients. Finally a considerable amount of research in the area of HIV-associated dementia (HAD) and cognitive impairment has made use of proteomics techniques to examine sera and cerebral spinal fluid (CSF) of HIV infected subjects with and without HAD [155-158]. These studies identified various biomarkers in association with HAD such as complement C3, cystatin C, MIF-1, and SOD-1. Unfortunately, the complexity with which HAD is accurately diagnosed made it difficult for the authors to confirm these markers using a multi-platform proteomic

approach. Studies of ex-vivo responses to HIV infection like these underscore the utility of proteomic techniques to measure potential markers of the host immune response to HIV and/or immune markers associated with HIV disease. They also stress the important role an unbiased proteomic approach can play in understanding the complex interplay between host and virus during all stages of HIV infection as well as the identification of novel biomarkers in various HIV associated diseases.

1.7.2 Quantitative mass spectrometry in innate and adaptive responses

Recently our group evaluated label-free quantitative techniques for the measurement of mucosal innate factors. qMS was able to quantify over 350 mucosal proteins in cervical vaginal lavage (CVL) from health women, including serpins, cystatins, defensins and soluble adaptive immune factors (IgA and IgG) all of which are important antimicrobial immune factors[159]. Proteomic approaches like these have a distinct advantage over traditional reductionist methods that often miss critical immune responses that may correlate with HIV protection.

Several groups including our own [147-151,160-162] are now expanding proteomic Systems Biology approaches to quantify ex-vivo immune responses to antigen stimulation. The potential of this approach for the measurement of immunological responses to a given immune stimulant could be amended to gauge recall responses to HIV vaccine antigen in clinical trails. In addition, this approach could be used in ongoing studies characterizing immune responses to HIV in the HESN individuals

described in this thesis. With correlates of protection or a specific immune signature being identified as protective (ie. immunological pathways), proteomics could be used to validate a desired response garnered from an HIV vaccine.

1.7.3 Proteomic Responses in HIV-1 Exposed Uninfected Individuals

The site of HIV transmission in women worldwide is primarily the female genital tract and efforts to develop prevention strategies require some prior knowledge about the type of immune response or interaction that is occurring genital mucosa. However, a number of qMS approaches have now been used which are capable of identifying biomarkers of interest that may not have been considered using traditional methodologies.

Initial studies of cervical vaginal lavage (CVL) by TOF-MS (label-free) identified Trappin-2/Elafin to be over-expressed in HESN women [108]. Expanding these findings with two additional studies in the same HESN cohort using 2D-DIGE and label-free mass spectrometry, Burgener et al identified a family of similarly over-expressed proteins (Elafin, RANTES, SLPI, MIP-α/β, Serpin A,B,C Families, and Cystatin B) [111,112]. Members of this protease inhibitor family play an important role in the acute phase immune response, inflammation, as well as having anti-HIV functions. Interestingly, at that time these factors had not been described to be present in the female genital tract. Discordant HESN men who have sex with men (MSM) have also been studied using qMS techniques [163]. iTRAQ proteomics was able to identify and quantify 337 unique proteins in saliva from HESN MSM. Basic

salivary proline-rich proteins 2 and 3 as well as Histatin-3, Lysozyme C, and SLPI were all found to be overabundant in HESN men. These studies highlight the ability of qMS techniques to identify differentially abundant immune factors at the site of HIV transmission while the identification of novel factors capable of inhibiting HIV in multiple studies by various mass spectrometry platforms confirms the utility of qMS as a reliable tool to identify correlates of immune protection.

As qMS methodologies advance, these techniques will make an even greater contribution to the search for novel intervention and treatment strategies, against HIV. However modern proteomic techniques have now advanced considerably in the past few years adding an important tool to a researcher's toolbox. Quantitative mass spectrometry techniques can now be utilized to study various aspects of mucosal immunology and ultimately correlates of protection against HIV.

1.8 Study Rationale, Hypothesis, and Objectives

Current studies to identify and characterize correlates of protection in HESN women from the Pumwani cohort have been primarily reductionist and/or genomic in nature (micro-array). Additionally these approaches have yet to completely examine the immune quiescence phenotype at the mucosal or cellular level in HESN women. Proteomic analysis of cervical fluid yielded some evidence to indicate this profile was mirrored in the genital tract however a proteomic analysis of immune cells has yet to be conducted at the systemic or genital tract level. In order to address these gaps in knowledge the **central hypothesis of this thesis is that**

quantitative shotgun proteomics of systemic and mucosal mononuclear cells from HESN women will demonstrate a proteomic profile characteristic of the immune quiescent phenotype, represented by the altered expression of pathways important in immune activation, cell recruitment / migration, and proteins involved in HIV replication pathways.

This hypothesis is addressed in the following objectives:

1. Optimize and develop an iTRAQ proteomics workflow capable of identification and quantification of proteins isolated from immune cells in order to characterize the cellular proteomic profile in the systemic and mucosal compartments of HESN women (Chapter 3: Shotgun proteomics development).

2. Identify an optimal peptide fractionation technique comparing high-pH reversed phase and off-gel isoelectric focusing as well as optimize fractionation and identification of iTRAQ labeled protein digest by mass spectrometry (Chapter 4: Quantitative proteomics development).

3. Evaluate the ability of the optimized iTRAQ proteomics platform to identify and quantify proteins relevant to immune activation and immune quiescence in peripheral blood mononuclear cells (PBMC) (Chapter 5: iTRAQ proteomics in peripheral blood mononuclear cells).

4. Utilize a qMS approach to determine the proteomic profile of ex-vivo PBMC from HESN women and validate individual protein biomarkers by Western blot analysis (Chapter 6: Comparative proteomic analysis of PBMC from HESN women)

5. Utilize a qMS approach to determine the proteomic profile of ex-vivo cervical mononuclear cells (CMC) from HESN women (Chapter 7: Comparative proteomic analysis of CMC from HESN women)

6. Validate potential protein biomarkers relevant to the immune quiescence phenotype and address any epidemiological factors that may play a cofounding role in HESN women (Chapter 8: Epidemiology of Mx1 and Mx2 in HESN women).

Chapter 2. Material and Methods

2.1 General Reagents

2.1.1 Buffers / Solutions

Phosphate Buffered Saline (PBS) – 48.5g PBS powder (Gibco), 137.93mM NaCl, 2.67mM KCl, 8.1mM Na2HPO4, 1.47mM KH2PO4 in 1L of ddH_2O

PBMC Wash Buffer – Phosphate Buffered Saline (PBS) with 2% Fetal Calf Serum (heat inactivated for 1hr at 56^oC), (Gibco)

R10 Cell Culture Media – RPMI-1600 (Sigma) supplemented with 10% Fetal Calf Serum (heat inactivated for 1hr at 56^oC) (Gibco), Penicillin, Streptomycin, Fungizone (Life Technologies)

DMEM Cell Culture Media – DMEM (Sigma) supplemented with heat inactivated 10% Fetal Calf Serum, Penicillin, Streptomycin, Fungizone (Life Technologies)

Tris Lysis Buffer – 115mM NaCl, 50mM Tris, 0.5% NP-40, +/- Roche Protease Inhibitor Tablet

Radio-Immunoprecipitation (RIPA) Lysis Buffer – 50mM Tris-HCl pH 8, 150mM sodium chloride, 1% NP-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate, +/- Roche Protease Inhibitor Tablet

SDS Lysis Buffer – 50mM HEPES, pH 8.8, 100mM Dithiothreitol (DTT; SIGMA) supplemented with 4% sodium dodecyl sulfate (GE Healthcare), +/- Roche Protease Inhibitor Tablet

Urea Exchange Buffer (UEB) – 8M urea with 50mM Tris or 50mM HEPES, pH 8.0

Freezing Medium – 90% Heat inactivated Fetal Calf Serum, 10% Dimethyl sulfoxide (DMSO, tissue culture grade, Sigma)

Odyssey Blocking Buffer (OBB) – 1:1 Dilution of Odyssey blocking buffer (LI-COR) in PBS

p24 ELISA Coupling Buffer - 1.4g Na_2CO_3 (Sigma), 2.93g $NaHCO_3$ (Sigma) dissolved in 1L of ddH_2O, pH 9.6.

p24 ELISA Blocking Buffer - 2% Goat Serum (Sigma), 0.01% Tween-20 (BDH Chemicals) in PBS.

Diethanolamine (DEA) p24 ELISA Buffer – 1:10 dilution of stock DEA (Sigma), 100mg $MgCl_2$ (Sigma), pH 9.8 w/ HCl.

2.1.2 Peptide and Protein Standards

6 Peptide Standard – 6 synthetic peptides obtained from American Peptide Company (0.8µg/ml/peptide): Peptide 1; LGGGGGGDGSR, Peptide 2; LGGGGGGDFR, Peptide 3; LLGGGGDFR, Peptide 4; LLLGGDFR, Peptide 5; LLLLDFR, Peptide 6; LLLLLDFR [164].

6 Protein Standard – Bovine serum albumin (22µg), α-lactalbumin (10µg), β-galactosidase (38µg), Lysozyme (10µg), Apotransferrin (25µg), and α-lactoglobulin (24µg). (Applied Biosystems)

2.2 Methods

2.2.1 Pumwani Commercial Sex Worker (CSW) Cohort

The Pumwani commercial sex worker cohort was established in 1985 and has >3800 female sex workers (FSW) currently enrolled. Upon enrollment in the cohort approximately 50% of women are found to be HIV positive. Despite intense HIV

prevention counseling 85% of initially HIV negative women are expected to seroconvert, becoming HIV positive within 3 years (HIV-N), and can be epidemiologically considered as HIV susceptible. HIV-N women in the cohort have the same selective pressures and exposure to HIV. Highly exposed seronegative (HESN) women are epidemiologically defined as relatively resistant to HIV acquisition and represent ~3% of the women in the cohort. HESN women are described as being HIV negative at the time of enrollment, continue to be negative for at least seven years of active follow-up (Serology and PCR negative), and are actively involved in sex work.

2.2.2 Ethics

Written informed consent was obtained from all study participants of the Pumwani cohort as well as local volunteers. Approval for all studies was obtained from research ethics board of the University of Manitoba and the Kenyatta National Hospital.

2.2.3 Sample Collection and Preparation

2.2.3.1 Blood Collection and Peripheral Blood Mononuclear Cell Isolation

Whole blood was collected from consenting donors into heparin containing Vacutainers. The remaining blood plasma was subsequently removed after a 7-minute spin at 1600rpm. The remaining blood was then layered onto Ficoll-Paque and centrifuged at 1400 rpm for 25 minutes. The peripheral blood mononuclear cell layer (PBMC) was collected and washed twice by centrifuging the cells at 1200 rpm

with PBMC Wash Buffer. Cells were centrifuged at 1600 rpm for 10 minutes and re-suspended in R10 media. PBMCs were counted by trypan blue exclusion using a hemocytometer and diluted with R10 media to the desired concentration.

2.2.3.2 Cervical mononuclear cell (CMC) isolation

Cells were obtained from the cervix by using a standard cervical brush and cell scraping, then transferred into a 50 ml Falcon tube containing 5 ml of PBS. CMCs were transferred to the lab on ice as soon as possible. The cervical cytobrush was vortexed briefly for 30 seconds and all cells and mucos were washed from the brush. The collected cells were then isolated by Ficoll-Paque density centrifugation as described in Section 2.2.3.1 above.

2.2.3.3 SDS Lysis of Ex-vivo Immune Cells and Cell Lines

Cells were washed 3 times of excess media with ice cold PBS and lysed in 0.2-0.5 ml of SDS lysis buffer. Cells were then immediately boiled for 5 min to inactivate any residual protease activity. Lysates were centrifuged at 15,000 g for 10 minutes to clear any insoluble cellular debris. Supernatants were collected and frozen at -80^{o}C.

2.2.3.4 Total Protein Determination of SDS Lysed Immune Cells

The protein concentration of samples lysed in SDS lysis buffer was determined using a 2D quant kit (GE HealthCare) capable of handling interfering chemical substances (SDS, DTT etc.).

2.2.3.5 Protein digestion and iTRAQ Labeling

Trypsin digestion was performed by the filter-aided sample preparation (FASP) method [165] with modifications, briefly described as follows: approximately 500 µg of protein per digestion was mixed with 7× its volume in UEB, then passed through a 10 kDa Nanosep spin filter (Pall Corp.). Two additional washes with UEB were performed to remove excess SDS. Protein was then alkylated with 50 mM iadoacetamide for 20 minutes at room temperature in the dark. Filters were washed twice with UEB, treated with 1000 units of benzonase (Novagen) in 50 mM Tris or 50mM HEPES for 30 min at room temperature, then rinsed again with Tris or HEPES buffer. Each filter was treated with 1.5 µg of trypsin gold (Promega) per 100 ug protein overnight at 37^{o}C. Filter cartridges were inverted and the peptides were spun off the filter with three washes of 50 mM Tris or HEPES. Peptides were lyophilized, re-suspended in water, and quantified by UV absorbance at 280 nm. Samples were then frozen at -80^{o}C for future analysis or processing. When iTRAQ labeling was required, digested peptides were re-suspended in 30µl of 100mM HEPES. The contents of a single iTRAQ reagent tube (AB Sciex) were re-suspended in 70µl of ethanol and added to the peptides. The pH of each reaction was verified and adjusted to pH 8 if necessary for optimal labeling efficiency. After labeling, 100µl of H2O was added to quench any further labeling.

2.2.4 Sample Fractionation and Mass Spectrometry

2.2.4.1 iTRAQ pre-scan normalization

Typical iTRAQ workflows call for immediate multiplexing of labeled samples before

fractionation and mass spectrometry analysis. Even though equal amounts of whole protein were added to each Nanosep filter, the recovery is typically not 100% which can bias the multiplexing of samples and result in an unnecessarily high normalization bias during data analysis. In order to ensure equal, one-to-one ratios between all iTRAQ reporter ions and verification of efficient labeling, equal volumes (1µl) of each labeled sample were mixed with Buffer A (2% acetonitrile, 0.1% formic acid) to a final volume of 20µl, and analyzed by online nLC/MS/MS (see below). Ratios from the most abundant proteins were then used to make minor corrections to the final mixing volumes ensuring equal ratios for all reporter channels.

2.2.4.2 High-pH reversed phase chromatography

After adjusting for mixing bias, samples were multiplexed to ensure an overall equal ratio for all the iTRAQ reporter channels. 75-µg of peptide was lypholized and re-suspended in 200 mM ammonium formate, pH 10 and loaded onto an offline High-pH reversed phase column for fractionation (Waters, XBridge C18 3.5um, 2.1x100 mm) described previously [166]. Samples were fractionated using an Agilent 1200 series microflow pump with Buffer A (20mM Ammonium Formate, pH 10) and Buffer B (90% ACN, 20mM Ammonium Formate, pH 10). The samples from each patient were then subjected to one of two linear gradients. A non-iTRAQ labeled gradient (0-60% buffer B at 150ul/min for 66 minutes) or a iTRAQ specific gradient (3% B – 10 min, 8-11.5% B – 7 min, 11.5-60% B – 58 min, at 150ul/min). A total of 24 fractions were collected over the peptide elution profile and then concatenated into

12 final fractions, maximizing the hyrdophobicity range. Fractions were then lyophilized and stored at -80°C for future MS analysis.

2.2.4.3 Off-gel Isoelectric Focusing

An Agilent 3100 OFFGEL fractionator was used to focus 75µg of PBMC peptide digest. Peptides were re-suspended and applied to an immobilized pH gradient 3–10, 24-cm strips (GE Healthcare) in a solution of 5% glycerol, and 1% ampholyte (GE Healthcare). Peptides were focused for 24-36 hours at 500 V for 5000 Vh. Fractions were isolated and pooled according to the mixing scheme outlined in Figure 3.9A followed by C18 stage tip purification in order to remove glycerol and ampoholytes. Fractions were lyophilization and stored at –80°C.

2.2.4.4 QStar XL nLC/MS/MS

Peptide fractions were re-suspended in nano-LC buffer A (0.1% formic acid, 5% ACN) and approximately 1.5 µg from each fraction was injected onto a C18- pre-column (Zorbax 300SB-C18, 5 µm, 5mm x 0.3 mm, Agilent) using a Agilent 1100 nanoflow LC system (Agilent). Peptides were resolved using a 15-cm analytical column (Zorbax 300SB-C18, 3.5 µm, 15cm x 75 µm, Agilent) using a 50-minute gradient (Buffer B: 0.1% formic acid, 95% ACN) at 300 nl/min (1 to 30%B; 30min, 40-95%B; 5min, 95%B; 5min, 1%B; 10min). Mass spectra were acquired using a QStar XL (Applied Biosystems) mass spectrometer in data dependant acquisition mode with a 10 second cycle time (1 sec: MS acquisition, 3 sec CID fragmentation on the top 3 most intense ions). The MS range was set from 350-100 m/z while the

MS/MS range was 70-2000 m/z. The CID collision energy used was automatically determined by acquisition software (Analyst QS 1.1).

2.2.4.5 Orbitrap XL and Orbitrap Velos nLC/MS/MS

Peptide fractions were re-suspended in nano-LC buffer A (0.1% formic acid, 2% ACN) and approximately 1.5 µg from each fraction was injected onto a 2-cm pre-column (id = 100 µm) consisting of 5-µm particle-sized ReproSil-Pur C18- AQ resin (Dr Maisch GmbH) using an Easy-nLC system (Thermo Fisher Scientific). Peptides were then resolved using a 10-cm analytical column (id = 75 µm) consisting of 3-µm particle-sized ReproSil-Pur C18-AQ resin (Dr Maisch GmbH) using a linear gradient from 0 to 40% buffer B (0.1% formic acid, 98% ACN) over 120 min at a flow rate of 300 nL/min. All experiments were run using this LC format with the exception of the study outlined in Chapter 7 (Comparative proteomic analysis of CMC from HESN women) which made use of a single 15-cm analytical column (id = 75 µm) consisting of 3-µm particle-sized ReproSil-Pur C18-AQ resin (Dr Maisch GmbH) and no pre-column. The total runtime was 160 min, which included sample loading, gradient elution, ACN wash, and column equilibration. Mass spectra were acquired using a data-dependent method using an LTQ Orbitrap XL or Orbitrap Velos mass spectrometer (Thermo Fisher Scientific). When unlabeled peptides (non-iTRAQ) were being analyzed, selection of the top five abundant precursor ions from each survey scan were isolated (2.0 m/z isolation width) and fragmented by CID (35% normalized collision energy, with 30 ms activation time). Experiments involving iTRAQ made use of a hybrid acquisition method involving the isolation and

subsequent fragmentation of the top 3 peptides in each survey scan by CID and HCD when using the Orbitrap XL. iTRAQ experiments ran on the Orbitrap Velos made use of HCD fragmentation only (40% normalized collision energy) for identification and quantitation. The survey scans, irrespective of the mass spectrometer used were acquired in the Orbitrap using a mass window of 300–1700 m/z at a target resolution of 60 000, and the subsequent fragment ion scans were acquired over a dynamic m/z range. The lower threshold for selecting a precursor ion for fragmentation was 1000 ions. Dynamic exclusion was enabled using a list size of 500 features, an m/z tolerance of 15 ppm, a repeat count of 1, a repeat duration of 30 s, and an exclusion duration of 120 s, with early expiration disabled.

2.2.4.6 Data Processing & Statistics

Wiff (Applied Biosystems) and RAW (Thermo Fisher) data files from the Qstar, Orbitrap XL or Orbitrap Velos were processed using the Mascot search engine (v2.3 Matrix Science). All data were searched against the International Protein Index (IPI version 3.78) human database. The following search parameters were used: monoisotopic precursor (MS) tolerance was set to 10 ppm while the MS/MS tolerance was set to 0.5 Da. Enzyme specificity was set to trypsin with a maximum of one missed cleavage. Carbamidomethyl of cysteine, iTRAQ labeling of N-termini and Lysine's were all applied as fixed modifications. Oxidation of methionine, and iTRAQ labeling of Tyrosine were set as variable modifications. All donors were imported into an instance of the Trans Proteomic Pipeline implemented in Scaffold v3.4.9 (Proteome Software). X!Tandem was used within Scaffold as an additional

measure for confident peptide and protein identifications (thegpm.org; version CYCLONE, 2010.12.01.1). Quantitative results were exported from Scaffold with the following criteria: minimum peptide probability of 80% (Peptide Prophet), a minimum protein probability of 99% (Protein Prophet), and a minimum of 2 peptides per protein. These setting resulted in a 0.1% protein and 1.5% peptide false discovery rate for the entire dataset. Statistical significance between the respective time points or study groups was evaluated using either a one-way ANOVA or students T test for each protein followed by correction for multiple comparisons using the Benjamini and Hochberg FDR method ($q<0.05$). Network and functional analyses were generated through the use of IPA (Ingenuity Systems, www.ingenuity.com). Cluster matrix generation was done using Pearson correlation via GraphPad Prism 5.0. In addition, hierarchical cluster analysis and dendrograms were generated using Cluster 3.0 and Java TreeView 1.1.6. All dendrograms were produced using an un-centered correlation similarity metric and complete linkage clustering.

2.2.5 Protein Gel Electrophoresis

Typically no more than 25 µg of protein was loaded on to a NuPAGE 4-12% Bis-Tris pre-cast acrylamide gel running MOPS buffer (Life Technologies). Affinity MagicMark™ XP (Life Technologies) or pre-stained Precision Plus™ Kaleidoscope™ (BioRad) were also loaded as molecular weight markers. Gels were resolved to completion using an XCell SureLock® Mini-Cell electrophoresis tank running at 150 V. When visualization of protein bands was required, staining of the gel was done

following the manufacturers instructions using the SilverQuest™ Silver Staining Kit (Life Technologies).

2.2.6 Licore Western Blot

After protein gel electrophoresis, proteins were transferred onto a nitrocellulose membrane using the Invitrogen iBlot transfer system. Membranes were directly placed in blocking buffer (OBB) for 1 hour at room temperature with gentle mixing. Primary antibodies for the protein target of interest were diluted in OBB with the addition of 0.1% Tween-20 (Sigma) and incubated overnight at 4^{o}C on an orbital shaker. The next day a loading control (b-actin antibody) was added to the blot for 1 hour at room temperature followed by four 5-minute washes with PBST (PBS, 0.1% Tween-20). A secondary infrared dye conjugated antibody (diluted in OBB, 0.1% Tween 20, and 0.01% SDS) was then added and incubated at room temperature for 1 hour in the dark followed by another four 5 minute washes with PBST. A final wash was performed for 5 minutes with PBS to remove any residual detergent. Blots were subsequently dried at 4^{o}C and imaged using an Odyssey CLx infrared imaging system (LI-COR). All primary and secondary antibodies used are listed in Table 2.1.

Antibodies	Company	Product Number
MX1	Epitomics	S1995
HLA-C	Epitomics	5472-1
Alpha 2 Macroglobulin	Epitomics	3458-1
RGS6	Epitomics	5758-1
β-actin	AbCam	ab8226
IRDye 800CW GαR	LI-CORE	926-32211
IRDye 680R GαM	LI-CORE	926-68020

Table 2.1. List of antibodies obtained for conformation of protein biomarkers in HESN women by Western blot analysis.

2.2.7 Generation of Overexpression Cell Lines

2.2.7.1 Stable overexpression SUPT1 cell lines

A third generation Lentivirus packaging system was used to create VSV-pseudotyped lentiviral particles containing an Mx1 overexpression plasmid, and the appropriate controls (GFP and empty plasmid). Details are provided below.

2.2.7.2 Plasmid Construction

An Mx1 cDNA plasmid was obtained (Open Biosystems, Accession #: BC032602) and the insert PCR amplified for cloning into a dual-promoter lentiviral packaging plasmid (SBI, CD514B-1). In addition, empty plasmid and a plasmid expressing green fluorescent protein (GFP) were prepared as controls. For the purposes of plasmid generation all plasmids were amplified in TOP10 *E.coli* cells (Invitrogen). All plasmids were sequenced to ensure correct amplification and insertion of the MxA sequence. The primers used to clone Mx1 into the Lentivirus packaging plasmid were 5'-ACATTCTAGAGCCACCATGGTTGTTTCCGAAGTGGACATC-3' and 5'-ATCAGGATCCTTAACCGGGGAACTGGGCAAG-3'. The forward primer contained a XbaI site (underlined) for cloning and a Kozak sequence for efficient translation while the reverse primer was designed with a BamHI (underlined) restriction site.

2.2.7.3 Lentivirus Expression and Concentration

A low passage number of 293TN cells (SBI, LV900A-1) grown in DMEM were plated onto 10 cm culture dishes to achieve 70% Confluency for the following day. The following day, 1 ml of serum free DMEM, 2.8 μg of each packaging plasmid, 1.7 μg of

overexpression expression vector, and 6µl of X-tremeGene HP transfection reagent (Roche) were mixed and added drop-wise to the 293TN cells. Cells were incubated for a further 48 hours at 37°C. Supernatants were collected after 48 hours and filtered through a 45 µm cellulose acetate filter. Virus concentration was achieved by adding PEG-it precipitation reagent (Systems Biosciences) to the supernatants which were then incubated overnight at 4°C. Virus was centrifuged the following day at 1500 g for 40 minutes. Viral particles were re-suspended in PBS, aliquoted, frozen at -80°C and titered as described below.

2.2.7.4 Lentivirus Titer Determination

Each Lentivirus construct was tittered using the Global UltraRapid Lentiviral Titer Kit (System Biosciences, LV961A-1) by quantifying the integration of the woodchuck hepatitis post-transcriptional regulatory element (WPRE) in SUPT1 cells. Briefly, SUPT1 cells were infected with a given volume of Lentivirus and cultured for 3 days followed by genomic DNA extraction. A master mix (23µl/rxn) containing both a WPRE or UCR1 (a house keeping gene for normalization) primer, and SYBRTaq supplied with the kit were added to a MicroAmp Optical 96-well reaction plate (Applied Biosystems). Either control DNA (supplied in the kit for the standard curve) or SUPT1 cell lysate was added to each well of the 96-well plate (2µl). The plates were sealed and a thermo-cycling program was conducted using StepOnePlus real time PCR system (Applied Biosystems) as follows (50°C for 2min, 95°C for 10min, (95°C for 15 sec; 60°C for 1 min) for 40 cycles). A dissociation step was also added to the program to rule out contamination of WPRE in the negative controls.

The MOI of the initial volume of Lentivirus construct used was then determined from the linear equation of the standard curve with an r^2 value no less than 0.99.

2.2.7.5 SUPT1 stable cell line generation

SUPT1 cells (CD4, CXCR4, lymphoblast cell line) were cultured in R-10 media and plated out into 96-well plates at $2x10^5$ cells/well. Cells were infected at an MOI of 1 using the specific Lentiviruses generated above. Cells were cultured for a further 48 hours and then the media was replaced under antibiotic selection (RPMI supplemented with 600ug of G418 or 1ug of Puromycin). Cells were cultured under these conditions for 5-10 days until the control cells died. Growth of the SUPT1 cell lines were then scaled up in T75 flasks with a splitting density of $5x10^6$ cells/ml, placed in 1 ml of freezing medium ($1x10^7$ cells/vial) and placed in liquid nitrogen.

2.2.8 HIV Virus Generation and Infections

2.2.8.1 HIV Virus Propagation

The HIV strain chosen for propagation and for use in the in-vitro infections of Mx1 expressing SUPT1 cells (Section 2.2.8.3) was a lab-adapted virus, HIV_{IIIB} (NIH AIDS Reference and Reagent Program). Briefly, PBMCs were isolated from a healthy donor for propagation of HIV_{IIIB} by Ficoll-Paque density centrifugation (Section 2.2.3.1). PBMCs were stimulated with PHA (5µg/ml) for 3 days at 37°C, 5% CO_2 and maintained in R-10 media supplemented with 20U/ml of IL-2 (NIH Aids Reference and Reagent Program). Cells were periodically checked for viability and fresh media was replaced when needed. PBMCs were washed twice with R-10 media and

counted by trypan blue exclusion. Cells were re-suspended at a concentration of $30x10^6$ cells/ml and taken to containment level 3 for viral inoculation. Cells were infected with HIV_{IIIB} at an MOI of 2 and incubated for 4 hours at 37°C, 5% CO_2. After incubation cells were diluted with R-10 media to a concentration of $3x10^6$ cells/ml and further incubated overnight at 37°C, 5% CO_2. Aliquots of supernatant were periodically measured by p24 ELISA. At peak p24 production (typically day 7), cell free media was harvested and aliquoted into 1ml cyrovials and stored at -80°C for infection and $TCID_{50}$ determination. An additional $30x10^6$ 'feeder' PHA-stimulated PBMCs from the same donor were added to the infection and incubated for an additional 7 days and monitored by p24 ELISA. A second harvest of the virus was done on day 14 and was also stored at -80°C.

2.2.8.2 $TCID_{50}$ Virus Titer Determination

HIV_{IIIB} was titered by the 50% tissue culture infectious dose assay ($TCID_{50}$), which determines the amount of virus necessary to infect 50% of inoculated wells. SUPT1 cells were cultured in T75 culture flasks with R-10 media and re-suspended at a concentration of $4x10^5$ cells/ml. In containment level 3, R-10 media was added to a 96-well-flat bottom tissue culture microtiter plate at 150 µl/well. Six wells received 133µl of R-10 media as these would be inoculated with undiluted stock virus. Sixty-seven microliters of undiluted HIV_{IIIB} stock was added to these 6 wells and 4-fold serial dilutions (50µl) were prepared across the microtiter plate (1:4 to 1:419430). This resulted in six replicates of the virus titration. Fifty-microleters of SUPT1 cells at a concentration of $4x10^5$ cell/ml were added to each well of microtiter plate

(20,000 cells/well). The infection was incubated for 4 hours at 37°C, 5% CO_2. Following incubation the microtiter plate was centrifuged at 1200rpm for 5 minutes and the viral inoculum (150µl) was removed and replaced with 150µl of fresh R-10 media. The plate was incubated for 3 days at 37°C, 5% CO_2. On day 3, 100µl of spent media was removed and replaced with 110µl of fresh R-10 media.

On day 7 110µl of media was removed and placed into a 96-well-round bottom microtiter plate containing 11µl of 10% Triton-X100 per well. The Triton-X100 disrupts the HIV envelope ensuring inactivation of the virus. Plates were then sealed and frozen at -80°C. Virus production for all the wells was determined by p24 ELISA (Section 2.2.9) and the 50% endpoint dilution was calculated using the Reed and Muench algorithm [167].

2.2.8.3 HIV infection of Mx1 expressing SUPT1 cells

SUPT1 cells are engineered to overexpress the CD4 receptor and also express the CXCR4 co-receptor. HIV_{IIIB} is a CXCR4 receptor dependent virus and was chosen for the infections for this reason. Prior to use in HIV infection assays, SUPT1 cells were thawed in a 37°C water bath, then washed in R-10 media by centrifugation at 1400 rpm for 10 minutes. Cells were passaged twice before use in HIV infection assays. 20,000 SUPT1 cells/well over-expressing Mx1 as well as controls were plated in triplicate into a 96-well flat bottom microtiter plate. HIV_{IIIB} virus was immediately thawed and diluted to achieve an MOI of 1, 0.5, and 0.25. Each dilution was performed in triplicate for each stably transfected SUPT1 cell line. After 4 hours of

incubation at 37°C, the viral inoculum was removed and replaced with fresh media (150µl). Culture supernatants were collected (100µl) and replaced with fresh media on Days 3, 5, and 7 of the infection and frozen at -80°C for future p24 ELISA analysis (Section 2.2.9).

2.2.9 p24 ELISA

2.2.9.1 p24 ELISA Reagents

IgG1 HIV-1 p24-specific monoclonal antibodies were produced in hybridoma cells (183-H12-5C, NIH AIDS Research and Reference Reagent Program), and served as the p24 ELISA coating antibody. Antibody purification from cell supernatant was carried out using Protein-G sepharose. Eluted antibody from the Protein-G column was then dialyzed against PBS and stored at -80°C. The detection antibody used in this ELISA was a polyclonal rabbit anti-p24 antibody (Advanced Biotechnologies) diluted in p24 ELISA blocking buffer (20µl of antibody in 1ml of p24 ELISA blocking buffer). Additionally, an equal volume of glycerol (1ml, Sigma) was added to the antibody mixure making a final dilution of 1:100, then aliquoted and stored at -80°C for future use. The secondary antibody used in this ELISA was a biotin anti-rabbit IgG goat antibody (Sigma-Aldrich), diluted 1:100 using the same method as above with storage at -80°C.

2.2.9.2 p24 ELISA Protocol

This in house ELISA protocol was used to detect p24 in the supernatants of SUPT1 infection assays. The coating antibody described above was thawed and diluted to

2µg/ml in ELISA coupling buffer. One-hundred microliters per well of this antibody solution was added to 96 well flat-bottom maxisorb microtiter plates (Nunc) and incubated overnight at 4°C. This buffer was removed the next day and coated plates were stored at -80°C for future use.

When needed, p24 ELISA plates were thawed and 100µl / well of p24 blocking buffer was added to the plates and incubated for 2 hours at 37°C. The blocking buffer was removed from the plates and washed twice with double distilled H_2O. Frozen supernatants from SUPT1 infection cultures were thawed and added in various dilutions (depending on the MOI used) to the ELISA plates (the final volume per well was 100µl). Additionally a standard curve of p24 was serially diluted across the plate from 20ng/ml to 0.02ng/ml. The plates were then sealed and incubated overnight at 4°C for maximum antigen capture.

Before use, p24 plates were washed six times with 300µl of 1X ELISA plate wash solution (Amplicor) using a microplate washer (ELx405, Biotek). All antibody dilutions were prepared in p24 ELISA blocking buffer. A 1:10,000 dilution of rabbit polyclonal anti-24 antibody was prepared in and 100ul added to each well. The plates were again sealed and incubated for 90 minutes at 37°C in a moisture box. After incubation the wells were washed again six times. Biotin anti-rabbit IgG antibody was diluted 1:13,200 and 100µl/well was added to the plates. The plates were then sealed and incubated for 90 minutes at 37°C in a moisture box. After

incubation, the plates were washed a final 6 times with the plate washer and rinsed twice with ddH_2O.

Two phosphatase substrate tablets (Sigma) were dissolved in 10ml of DEA buffer (10ml / ELISA plate). This substrate solution was added to the plates (100µl/well), and incubated at 37°C for 30min. The plates were read at 30, 60, and 90 minutes with a wavelength of 405nm using a SpectraMax Plus spectrophotometer (Molecular Devices). The p24 standard curve was subsequently fitted using a 4-parameter curve.

2.2.10 Expression and Epidemiological Analysis of Mx Proteins

2.2.10.1 PBMC Sample Preparation for Mx1 and Mx2 ELISA

All PBMC samples for the epidemiological studies (Chapter 8) conducted on Mx1 and Mx2 in the Pumwani cohort were isolated in Nairobi, Kenya (Section 2.2.1) from 1995 to 2009. PBMC were frozen in freezing media and shipped back to Winnipeg and stored in liquid nitrogen. Cells were immediately removed from the LN2 tank and rapidly thawed in a 37°C water bath. Cells were transferred from the cryovial and immediately placed in 10ml of R-10 media in order to dilute the DMSO and give the cells a chance to recover. Cell viability and counts were conducted on every sample to ensure a viability of at least 80% using the Invitrogen Countess® Automated Cell Counter by trypan blue staining. After counting the cells were centrifuged at 800g for 10 minutes. Cells were re-suspended in PBS and washed three times. The third time cells were re-suspended in 500ul of PBS and subjected

to three rapid freeze thaw cycles on dry ice. The resulting PBMC lysate was centrifuged at 1500g to remove any excess cellular debris. The PBMC lysate was transferred to a new cryovial and stored at -80^{o}C for future use in the Mx1 and Mx2 ELISA. The total protein concentration of PBMC lysate in PBS was determined by BCA assay (Millipore).

2.2.10.2 Mx1 & Mx2 ELISA

A commercial ELISA was purchased from USCN Life Science Inc. for the quantification of Mx1 and Mx2 in PBMC lysate. Briefly, a standard doubling dilution series was prepared for either Mx1 or Mx2 (20ng/ml to 0.02 ng/ml) and a 100ul/well of PBMC lysate in PBS from each study participant was added to the plates in duplicate. The plates were then sealed and incubated for 2 hours at 37^{o}C. After incubation the samples and standard curve were aspirated from the plate using the Bioteck plate washer. A biotin-conjugated antibody specific for Mx1 or Mx2 was added (100µl/well of a 1:100 dilution) followed by incubation at 37^{o}C for another 60 minutes. After incubation the plates were manually washed three times with the wash buffer provided in the ELISA kit (300ul/well). A secondary avidin conjugated to Horseradish Peroxidase (HRP) was then added (100µl/well of a 1:100 dilution) and incubated for 30 minutes at 37^{o}C. The plates were again washed five times with the provided wash buffer. TMB substrate (90µl/well) was then added to the plates and incubated at 37^{o}C for 15-30 minutes. Stop solution (sulphuric acid, 50µl/well) was then added to terminate the reaction. The plates were read using the SpectraMax Plus spectrophotometer at a wavelength of 450nm. As with the p24

ELISA a 4-parameter curve was fit the standard curve for accurate quantitation of Mx1 and Mx2.

2.2.10.3 Epidemiological Analysis

The Mx1 and Mx2 expression levels measured by the commercial ELISA were transformed using their natural logarithms. The log-transformed values of Mx1 and Mx2 were subsequently used as outcome variables in separate multiple linear regression models. Where appropriate, t-tests and ANOVA's were used in bivariate analyses with logged values of Mx1/Mx2 used as the outcome variables. Multiple linear regression models were fitted to the data and used to assess the independent association of HESN status and the logged values of Mx1 and Mx2, adjusted for age, use of medication and Depo-Provera. Given their theoretical importance, all confounders were kept in full linear regression models; however, each confounder was entered separately in the following order: use of Depo-Provera, age, and use of medication in the last 6 months. In addition to the above confounders, the interaction between Depo-Provera use and HESN status was added as an additional variable in the final full model. The interaction was considered statistically significant at $p<0.05$. Stata 12 (College Station, TX) was used for all analyses. Age, use of medication and Depo-Provera were included *a priori* due to the impact of these variables on Mx1/Mx2 levels observed in a crude analysis (Figure 8.3). The interaction between HESN status and Depo-Provera was planned a priori due to our interest in whether or not levels of Mx1/Mx2 between HESN and high-risk negative controls differed by use of Depo-Provera. From an initial sample of 276 women who

had values for both Mx1 and Mx2, 86 were excluded if they were HIV-positive, leaving 190 women. For all models, menopausal (n=2), pregnant (n=5) and missing values for contraception use (n=22) were excluded from the analyses, leaving a total sample size of 161 women.

Chapter 3. Shotgun proteomics development

3.1 Introduction & Rationale

We hypothesize that cells from the systemic and mucosal compartments of HESN women will recapitulate an immune quiescent phenotype with altered expression of pathways involved in immune function and HIV restriction factors at the proteomic level. In order to address this hypothesis a proteomics based systems biology pipeline for the identification and quantification of proteins in a complex biological sample needed to be developed. The protocols for cell lysis, optimization of offline peptide fractionation methods, choice of mass spectrometer, MS parameters during data acquisition and data analysis all had to be developed before this study could be undertaken Figure 3.1. The following objectives were established to develop an optimal proteomics discovery platform to be utilized in addressing our central hypothesis.

3.2 Objectives

1. Develop a method for the identification of proteins from whole cell lysates

2. Identify optimal lysis, peptide digestion, and peptide fractionation, techniques for a proteomics biomarker analysis pipeline

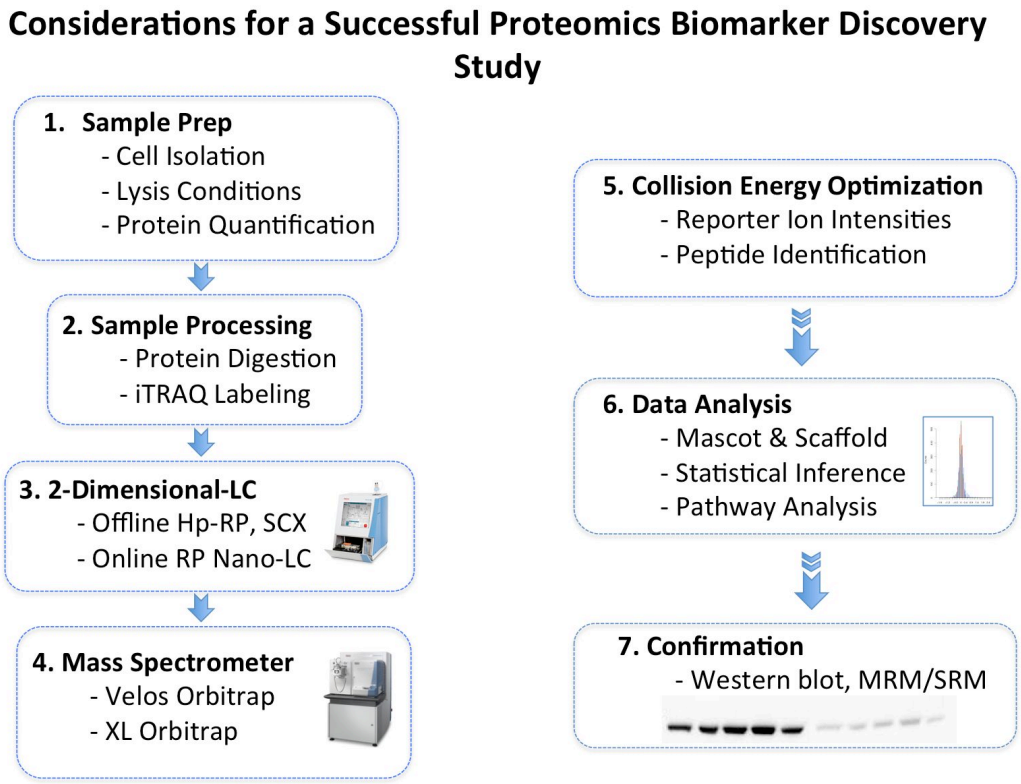


Figure 3.1. Criteria and considerations for an optimal systems biology and biomarker discovery shotgun proteomics pipeline.

3.3 Results

3.3.1 Study Population

All optimizations in this chapter were performed with ex-vivo isolated peripheral blood mononuclear cells (PBMC) from healthy consenting local donors by Ficoll density gradient centrifugation described in Materials and Methods 2.2.2

3.3.2 Cell Lysis & Digestion

To obtain the maximum amount of protein that could be isolated from freshly isolated PBMC, three whole cell lysis buffers were tested, a Tris buffer, a RIPA buffer and a 4% SDS buffer (Section 2.1.1). The protein content was evaluated by loading 15µg of whole cell lysate onto a 4-12% SDS – PAGE gel followed by silver stain analysis to visualize the resulting cell lysate. An examination of the protein band distribution obtained from each lysis buffer indicated that the 4% SDS lysis buffer extracted the most protein bands over the largest molecular weight range (Figure 3.2). As a result the 4% SDS lysis buffer was chosen for total protein isolation in future studies.

Typically, shotgun proteomics relies on the analysis of tryptic peptides by mass spectrometry resulting in the need for optimal digestion of isolated protein. To optimize this protocol, digestion of PBMC lysate in solution with various concentration of trypsin was carried out overnight and the resulting digestion products were resolved by SDS-PAGE and visualized by silver stain analysis (Section 2.2.5). The results indicated that increasing amounts of trypsin did not improve the

digestion efficiency as indicated by numerous proteins with an apparent molecular weight well above 25 kDa (Figure 3.3A). To evaluate if improvements could be made to the original digestion protocol we adjusted the temperature of the disulfide bond reduction reaction from room temperature to 56^{o}C. We reasoned that increasing the temperature of the reaction would increase the energy available for the reducing agent (DTT) to break the disulfide bonds. We observed considerable improvements in the digestion efficiency after increasing the temperature of the reaction as well as increasing the trypsin to protein ratios, however a significant amount of proteins above 25 kDA were still present (Figure 3.3B). To further optimize the digestion efficiency and achieve complete digestion of our PBMC lysate sample a filter aided digestion protocol (FASP) was employed (Section 2.2.3.5) [165] and tested with all three lysis buffers on PBMC lysate. The FASP method allowed for complete removal of chemicals from subsequent steps (DTT, IAA, SDS) that could interfere with complete digestion. When we tested the enhanced FASP protocol we found that it lead to the efficient digestion of PBMC lysate irrespective of the lysis buffer used (Figure 3.2; digests). The FASP protocol in conjunction with 4% SDS lysis of PBMC was subsequently chosen for all future studies.

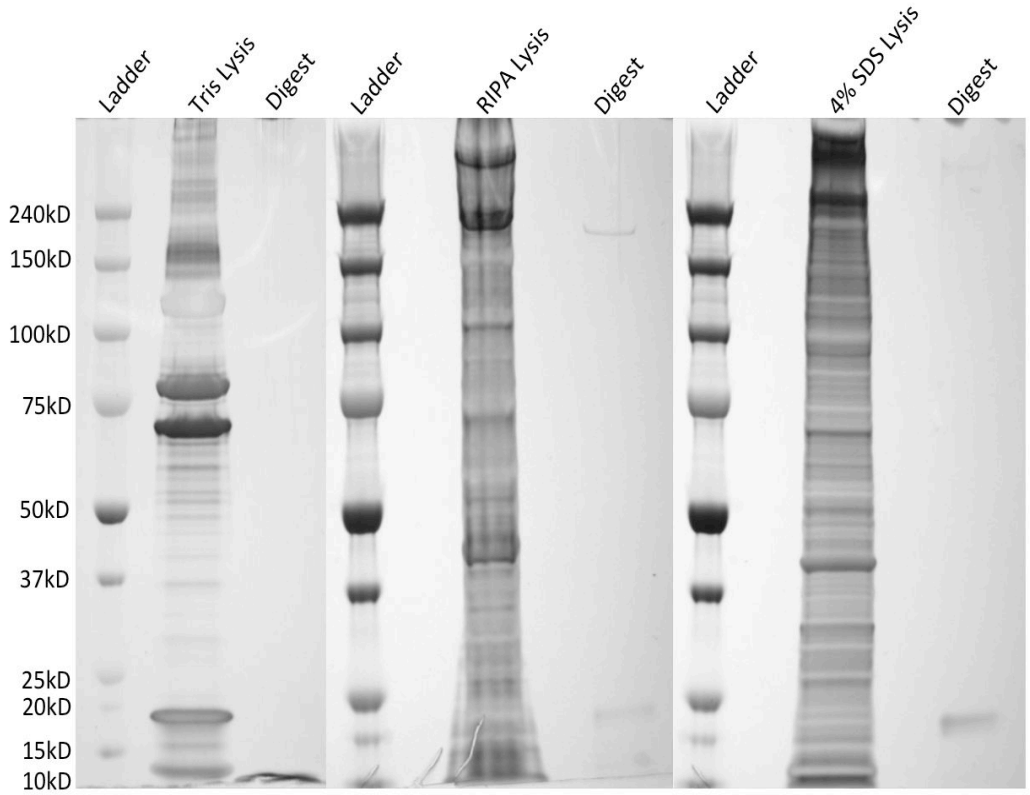


Figure 3.2. PBMC lysis buffer and digestion analysis. Three lysis buffers were evaluated for their potential to isolate a complex mixture of proteins from PBMCs (Tris, RIPA, and 4% SDS). Fifteen micrograms of protein was loaded into each well of a 4-12% Bis-Tris SDS gradient gel as well as trypsin digest (using a modified filter aided digestion) of the respective lysates. Visualization was performed by silver stain analysis.

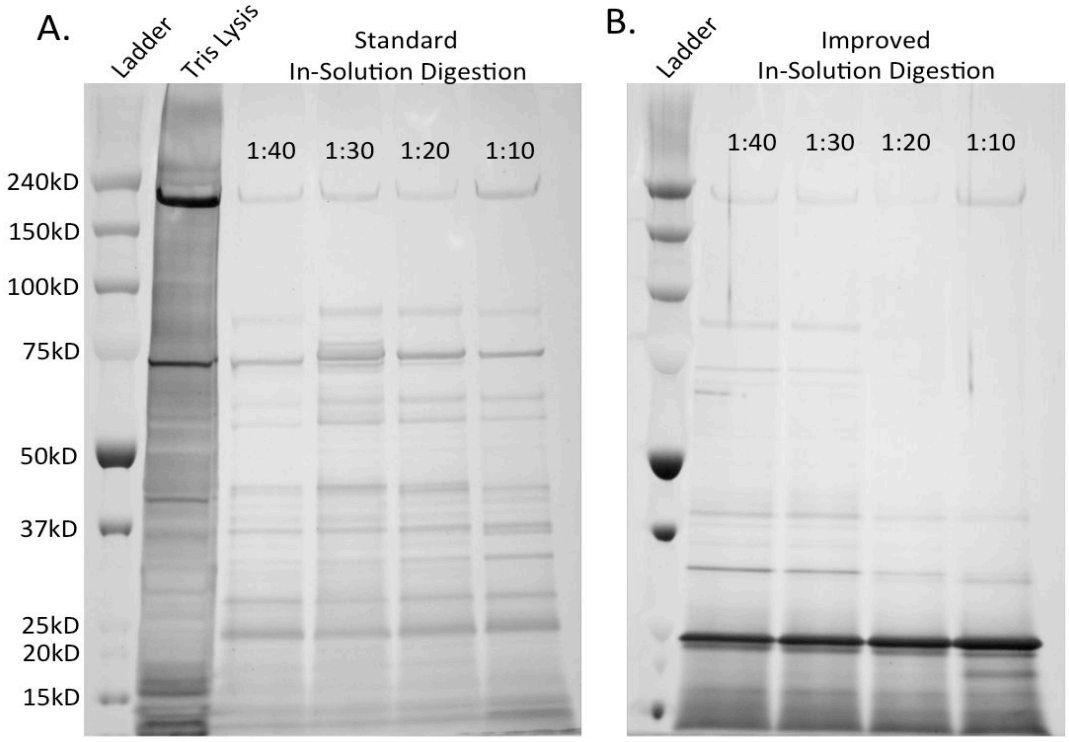


Figure 3.3. Optimization of trypsin digestion conditions. Fifty micrograms of PBMC lysate was subjected to various trypsin:protein ratios (in-solution) and 25ug loaded into each well of 4-12% Bis-Tris SDS gradient gel followed by silver stain visualization. The (A) standard digestion procedure called for reduction of protein at room temperature while the (B) improved protocol increased the temperature to 56^{o}C.

3.3.3 Offline Peptide Fractionation

Fractionation of peptide samples is a common method for reducing the complexity of a sample in order to allow the mass spectrometer the best chance to identify as many peptides as possible. A common technique known as MuDPIT (Multi-Dimensional Protein Identification Technology) was used to pre-fractionate the sample and then run each individual fraction through a second fractionation technique and then directly into a mass spectrometer for analysis. This second dimension of fractionation is typically referred to as "online" liquid chromatography due to the direct coupling to the mass spectrometer. There are numerous techniques available for the first fractionation step, including strong cation exchange (SCX), strong anion exchange (SAX), reversed phase (RP), isoelectric focusing (IEF), and hydrophobic interaction chromatography (HILIC). We chose to optimize a high-pH reversed phase fractionation method (Hp-RP). Hp-RP has several advantages over standard fractionation techniques in that it is highly compatible with subsequent analysis steps. Other techniques such as SCX, in practice, require significant desalting of the samples before being injected into a mass spectrometer. This step can lead to sample loss and could significantly impact the analysis of clinical samples that often contain small amounts of material. In addition, the orthogonality of Hp-RP has been shown to be as good as the more typical methods such as SCX [168].

We made use of the optimized protein isolation and peptide digestion protocol above and used an Agilent 1200 microflow pump to inject 75 μg of PBMC protein

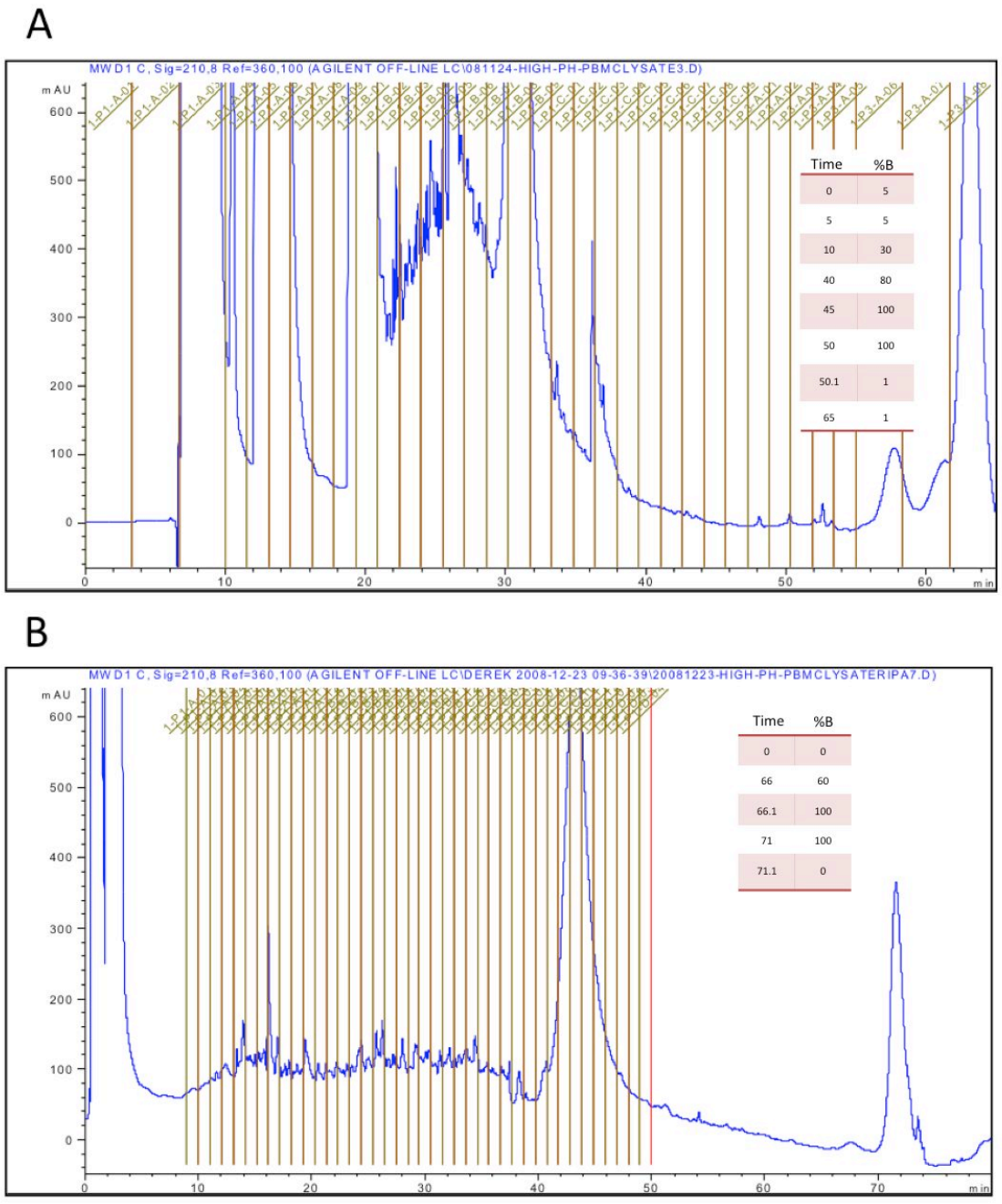


Figure 3.4. Optimization of high-pH reversed phase chromatography fractionation of PBMC whole cell digest. In each case 75 µg of digested PBMC lysate was loaded onto an Xbrige C18 column and subjected to (A) a rapid semi-linear gradient of ACN (1.6%/min) or (B) a slow linear gradient of ACN (0.9%/min) in order to achieve optimal peptide fractionation.

digest onto an Xbridge C18 column for fractionation by high-pH reversed phase (Section 2.2.4.2). We first tested a semi-linear acetonitrile (ACN) gradient, increasing the percentage of acetonitrile at a rate of 1.6%/min, based upon an initial lab protocol (Figure 3.4A). We found that peptide eluted off the column from 19 minutes to 34 minutes. An optimal gradient should allow for the maximum amount of fractionation over a large time frame without loosing substantial peptide signal. Subsequently a new gradient was designed with a much slower rate of acetonitrile addition (0.9%/min). This new gradient allowed for peptide elution over a much wider elution profile spanning 30 minutes (Figure 3.4B). Visual analysis of the LC chromatogram after fractionation indicated a strong peak eluting between 40 and 50 minutes. Fractions containing large peaks of this intensity are not normally analyzed by mass spectrometry due to the risk of column clogging and failure. However, this peak may also include peptides of interest that would otherwise be excluded from mass spectrometry analysis.

To identify and eliminate the strong peak eluting from 40-50 minutes in Figure 3.4B as well as determine if there were peptides to be collected in this time frame we conducted several sets of experiments where various components of the sample preparation protocol were eliminated. This was followed by Hp-RP chromatography to isolate and identify the peak. In order to test the hypothesis that the peak was due to excess undigested trypsin in the sample, a mock digestion protocol was prepared with or without trypsin on a blank sample and subjected to Hp-RP liquid chromatography (LC). This experiment also allowed us to rule out

other chemicals included in the digestion procedure as being responsible for the interfering peak. The absence of a peak eluting between 40-50 minutes in Figure 3.5A indicates that the trypsin, dithiothreitol (DTT), and iodoacetamide (IA) used in the optimized digestion protocol were not responsible for the peak. To test if the SDS lysis buffer was responsible, SDS lysis buffer was injected alone and the buffer was determined to be the source of the interfering peak (Figure 3.5B). Additionally, SDS lysis buffer (Figure 3.3C) was prepared with and without the Roche protease inhibitor cocktail (a standard component in most lysis buffer formulations) to rule out this component of the lysis buffer. Removing the protease tablet from the lysis buffer resulted in a loss of the interfering peak eluting between 40 and 50 minutes (Figure 3.5C; red). Lastly, to determine if peptides were eluting during the 40-50 minute window and being masked by the protease inhibitor cocktail peak, a new PBMC sample was lysed with SDS buffer sans the protease inhibitor cocktail and digested using the optimized protocol. Seventy-five micrograms of PBMC digest was injected onto the HPLC and subjected to Hp-RP fractionation. The protease inhibitor contamination peak was successfully removed from the chromatogram and revealed a significant amount of hydrophobic peptides eluting in the 40-50 minute time frame (Figure 3.6).

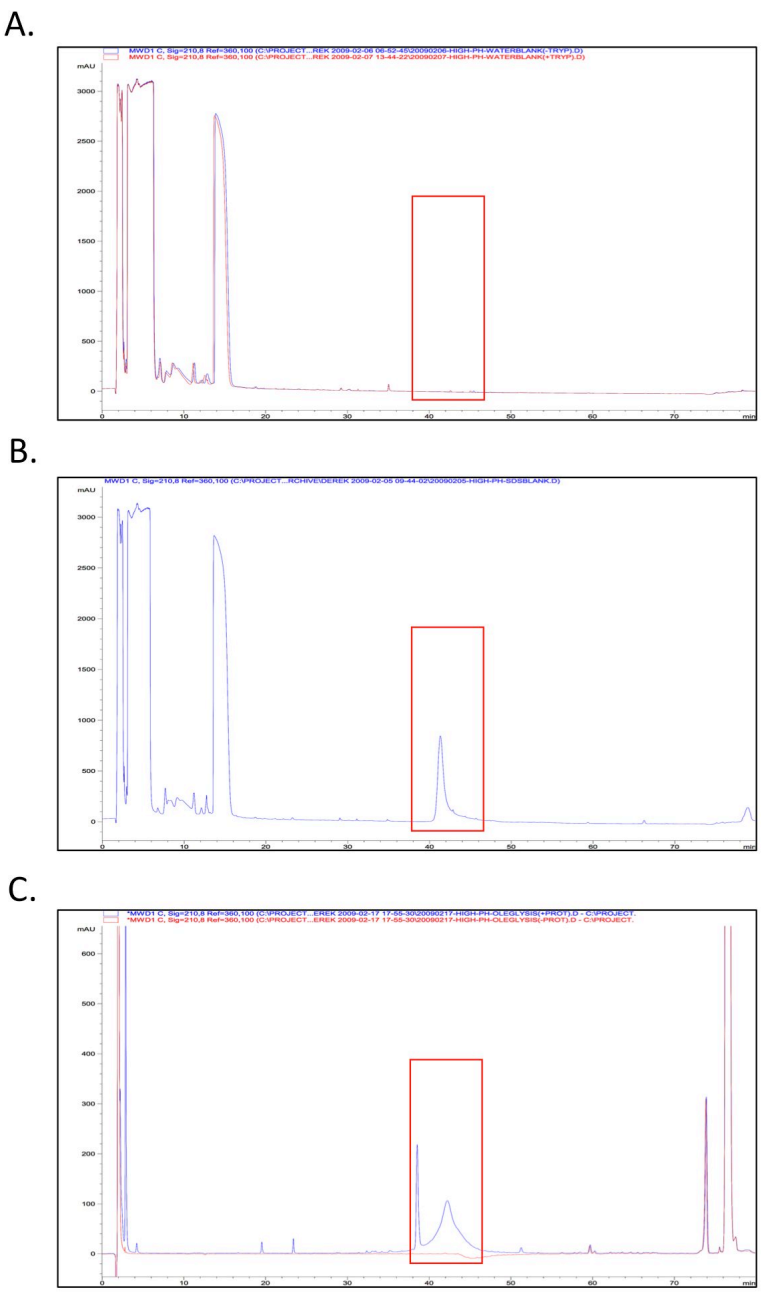


Figure 3.5. Isolation and removal of protease inhibitor cocktail contamination peak in peptide fractionation. Injection of (A) a mock digestion procedure with (red) and without (blue) trypsin, (B) a single injection of SDS lysis buffer, and (C) SDS lysis buffer with (blue) and without (red) the addition of protease inhibitors.

Fractions collected from the chromatogram illustrated in Figure 3.6 were submitted for mass spectrometry analysis using a QSTAR mass spectrometer (Section 2.2.4.4). Scaffold analysis of all 25 fractions resulted in the identification of 238 and 879, unique proteins and peptides respectively. Only being able to identify 238 proteins from the QSTAR analysis would likely be insufficient to address our hypothesis. To optimize and maximize the number of identified proteins we made use of the newly available hybrid Orbitrap XL mass spectrometer (ThermoFisher). Using the optimal sample preparation protocol a second PBMC sample was prepared and analyzed using the Orbitrap XL. While the initial QSTAR analysis examined 25 separate fractions we decided to employ a fraction mixing strategy to reduce the number of fractions from 25 down to 12. During the initial fractionation, 24 fractions are collected however these fractions were then mixed in order to efficiently span the maximal hydrophobicity range (Section 2.2.4.2; Figure 4.1). This strategy reduced the mass spectrometer run time in half which improves the number of clinical samples that could ultimately be evaluated for future experiments. We then compared the performance of the QSTAR vs. the Orbitrap XL mass spectrometer in identifying proteins and peptides from PBMC digest using the optimized sample processing technique (Figure 3.7). We found that by analyzing 12 fractions, the Orbitrap XL was capable of identifying 3152 proteins and 24,343 peptides. This resulted in a more than 13-fold and 27-fold increase in protein and peptide identifications respectively, in half the time using the optimized sample preparation protocol. The peptide identification rate from the spectra collected for the QSTAR was only 3% while the Orbitrap XL achieved 41%. The percent protein

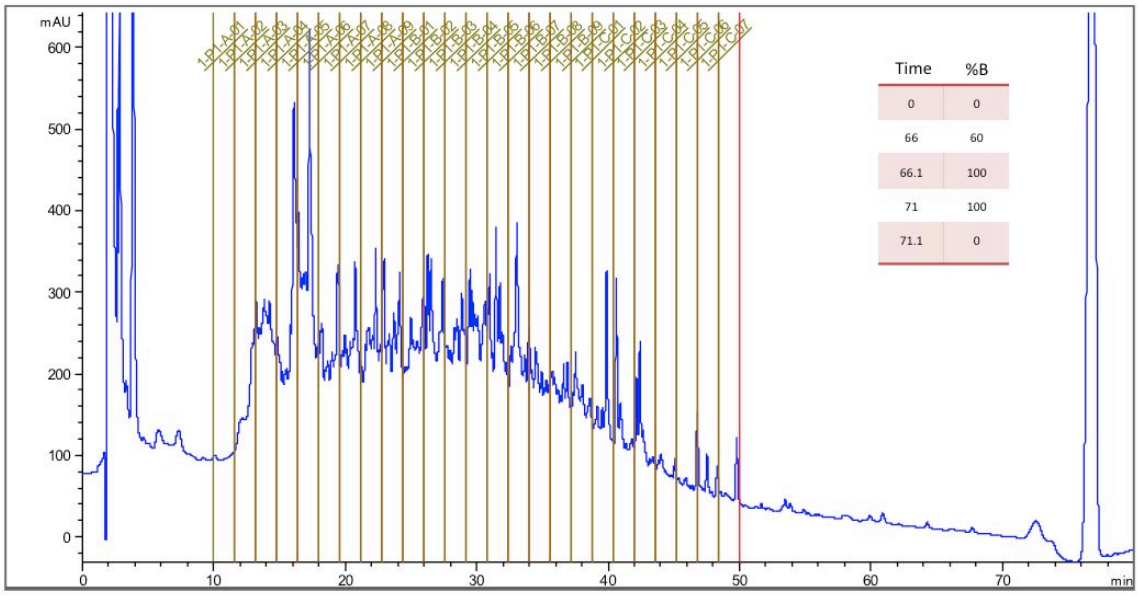


Figure 3.6. Optimized PBMC lysis, digestion and peptide fractionation for analysis by mass spectrometry.

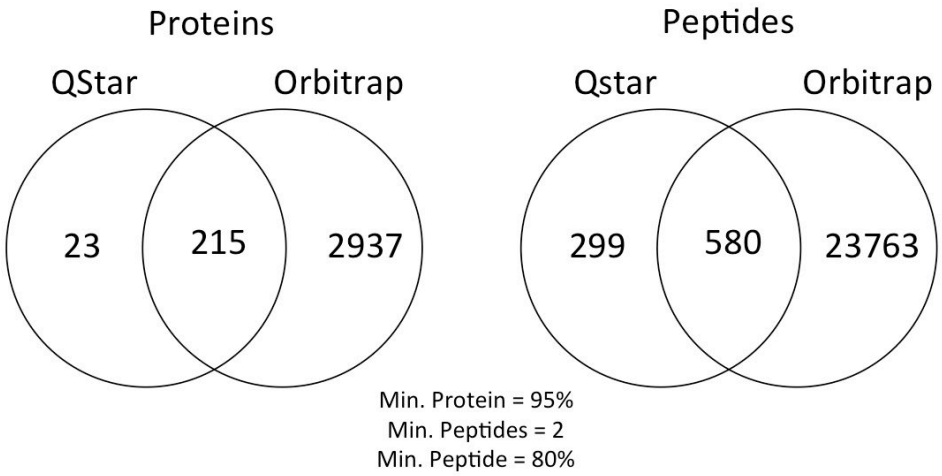


Figure 3.7. Venn diagrams depicting the proteins and peptides identified by the Qstar and Orbitrap mass spectrometers using the optimized sample processing protocol. Twenty-five fractions were injected onto the Qstar mass spectrometer while only 12 were injected onto the Orbitrap mass spectrometer.

sequence coverage also doubled for the Orbitrap XL (21%) compared to the QSTAR (11%). This data showed that the filter aided digestion protocol resulted in a considerable increase in protein and peptide identifications on the Orbitrap XL. One issue that arose from this optimized protocol was that the 10 kDa filter employed for the digestion is extremely time consuming. Switching to a 30 kDa filter allows for the loading and washing of the initial protein sample at a much faster rate but at the cost of allowing species smaller than 30 kDa to pass through the filter and be lost in subsequent wash steps. In an attempt to reduce the sample preparation time that would be needed to process hundreds of samples from clinical isolates a comparative analysis of PBMC fractionated digest prepared with the two different filters was undertaken. We then prepared matched PBMC digests using the 10 kDa and 30 kDa filter followed by Hp-RP fractionation and Orbitrap XL analysis. The resulting data (Figure 3.8) indicated that the 10 kDa filter identified considerably more proteins and peptides than the 30 kDA filter (2901 vs. 1763 and 21,914 vs. 10,663 respectively). In fact the 10kDa filter was able to identify well over 10,000 more peptides than the 30kDa filter. We also expected that the 30 kDa filter would have difficulty identifying proteins below 30 kDa however both techniques were comparable in terms of the theoretical size of proteins identified. As a result of this experiment the reduction in sample processing time using the 30 kDA filter did not warrant the significant loss in protein and peptides identified, indicating that the 10 kDa filter method, while time consuming, would be the best choice for any future studies.

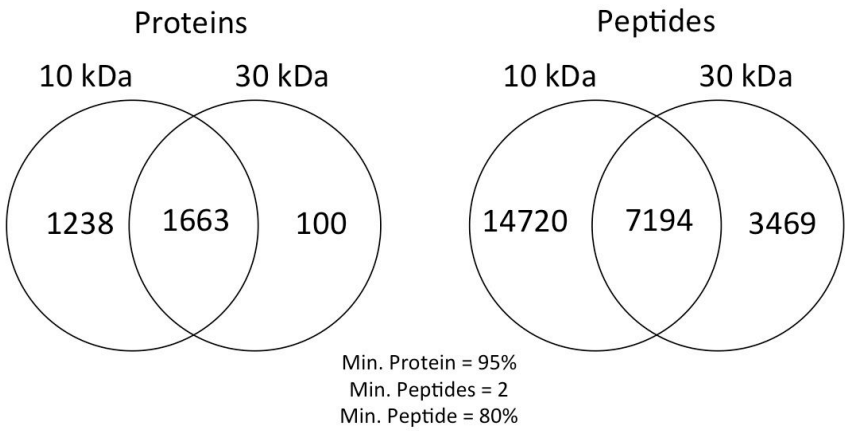


Figure 3.8. Proteins and peptides identified by the Orbitrap XL mass spectrometer by filter aided digestion using 10 kDa and 30 kDa molecular weight filters.

3.4 Summary

The development of optimal cell lysis, trypsin digestion, and Hp-RP fractionation was achieved for the analysis of PBMC samples. The 4% SDS lysis with a modified filter-aided digestion was shown to be the optimal combination for the isolation and digestion of ex-vivo PBMC samples. The Hp-RP fractionation method was optimized to yield the maximum amount of protein and peptide identifications by mass spectrometry. In addition, the Orbitrap based hybrid mass spectrometer was shown to identify significantly more proteins and peptides than the QSTAR mass spectrometer. In the following chapter, the evaluation of a different fractionation technique compared to the optimized Hp-RP protocol is conducted. Additionally, conditions for iTRAQ fractionation and quantitation are assessed to identify an optimal quantitative proteomics platform.

Chapter 4. Quantitative proteomics development

4.1 Introduction and Rationale

It was important that other fractionation techniques also be considered for the identification and eventual quantitation of proteins from PBMC lysate. One such technique that has similar capabilities to the Hp-RP method and would be highly orthogonal is Off-gel isoelectric focusing (OG-IEF). This technique was extensively optimized in Dr. Carpenter's lab and was available for a detailed comparison of the two techniques. OG-IEF has been evaluated extensively in the literature in comparison to SCX fractionation [169,170] but had yet to be comprehensively compared with regards to Hp-RP. Additionally, fractionation of iTRAQ labeled peptides needed to be optimized as well as mass spectrometer conditions for the eventual identification and quantitation of proteins and peptides in order to address our main hypothesis.

4.2 Objectives

1. Identify an optimal peptide fractionation technique for the identification of proteins and peptides from PBMC lysate comparing Hp-RP and OG-IEF fractionation.

2. Develop optimized fractionation and mass spectrometry parameters for the identification and quantification of iTRAQ labeled protein digest.

4.3 Results

4.3.1 Study Population

All optimizations in this chapter were performed with ex-vivo isolated peripheral blood mononuclear cells (PBMC) from healthy consenting local donors by Ficoll density gradient centrifugation described in Materials and Methods 2.2.2

4.3.2 Hp-RP and OG-IEF Fractionation

PBMC lysate was isolated and prepared from a single healthy donor as described in Sections 2.2.3 - 2.2.4. Seventy-five microgram aliquots of PBMC digest were prepared and subjected to fractionation either by Hp-RP (Section 2.2.4.2) or OG-IEF (Section 2.2.4.3). As our previous work showed that Hp-RP could identify over 3000 proteins with just 12 fractions we also limited the OG-IEF fractionation scheme to the same number of fractions in order to ensure a fair comparison. Both fractionation techniques produced 24 fractions however as described, an optimized mixing scheme was employed for both techniques to try and maximize the number of peptide and protein identifications (Figure 4.1). Running 12 fractions on the mass spectrometer resuled in roughly 32 hours of mass spectrometer analysis per sample. Quadruplicate aliquots of PBMC digest (75 µg) were subjected to either Hp-RP or OG-IEF fractionation followed by Orbitrap XL peptide and protein identification. Several factors such as the number of proteins and peptides identified, the peptide distributions over the 12 fractions collected, peptide resolution, and biophysical properties were analyzed to determine the best technique for future shotgun proteomics discovery and systems biology studies.

The results from the peptide and protein identifications as well as the overlap between the four individual replicates are depicted in Figure 4.2. The Hp-RP approach lead to 17% more protein identifications (average ± SD, 2378 ± 169) compared to OG-IEF (2024 ± 130) ($p = 0.0021$). An average MS/MS identification rate of 31% was achieved for both the OG-IEF and HP-RP. For OG-IEF, 85% (1716±81) of proteins were detected in at least two replicates, 78% (1570±54) in at least three replicates, and 73% (1477) in all four replicates (Figure 4.2A). For Hp-RP, 84% (2002±109) proteins were identified in at least two replicates, 77% (1824±73) in at least three replicates, and 72% (1711) in all four replicates (Figure 4.2B). A combination of the protein results from all replicates indicated an overlap of 2391 proteins for OG-IEF and Hp-RP together, while 216 proteins were identified by OG-IEF and 693 proteins were uniquely identified using Hp-RP (Figure 4.2C).

The Hp-RP fractionation with respect to peptide identifications also identified 16% (14,179 ± 752) more peptides compared to the OG-IEF technique (12 180 ± 603), $p = 0.0010$ (Fig. 4.2B). For OG-IEF, 68% (8310 ± 437) peptides were found in two replicates, 55% (6756 ± 260) were in three replicates and 48% (5848) were found in all four replicates. The Hp-RP, identified 65% (9186 ± 620) of peptides in two replicates, 51% (7245 ± 427) in three replicates and 43% (6151) in all four replicates. A combination of the peptide results from all replicates indicated an overlap of 13,984 peptides for OG-IEF and Hp-RP together, while 6053 were identified by OG-IEF and 10,445 peptides were uniquely identified by Hp-RP (Figure 4.2C). Even though the Hp-RP method outperformed the OG-IEF technique in terms

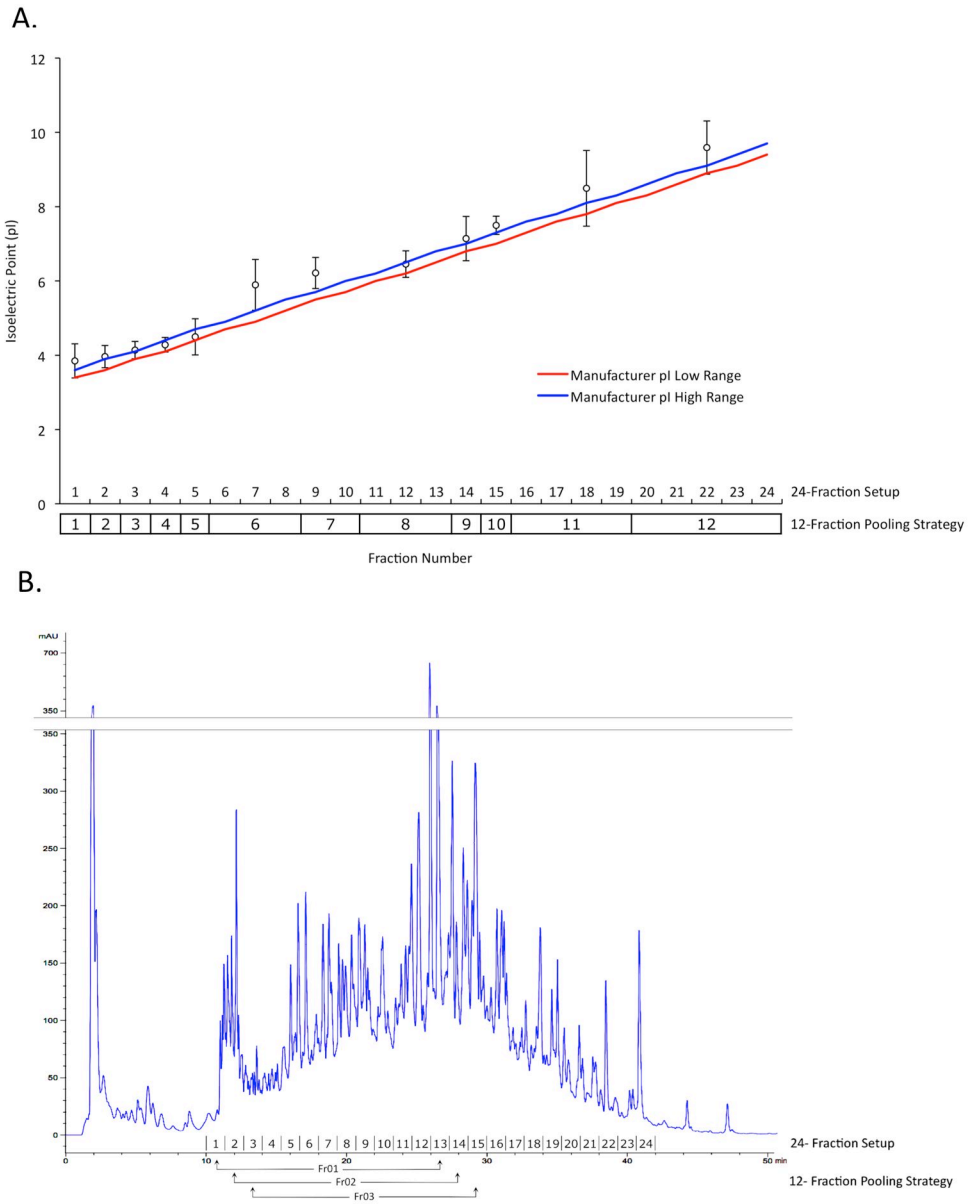


Figure 4.1. Pooling strategy for both OG-IEF and Hp-RP. (A) Initially a 24 well IPG strip was used to fractionate the PBMC samples followed by sequential mixing into 12 fractions. (B) Twenty-four fractions were also collected for the Hp-RP technique and also mixed into 12 fractions (Fr1+13, Fr2+14 etc.).

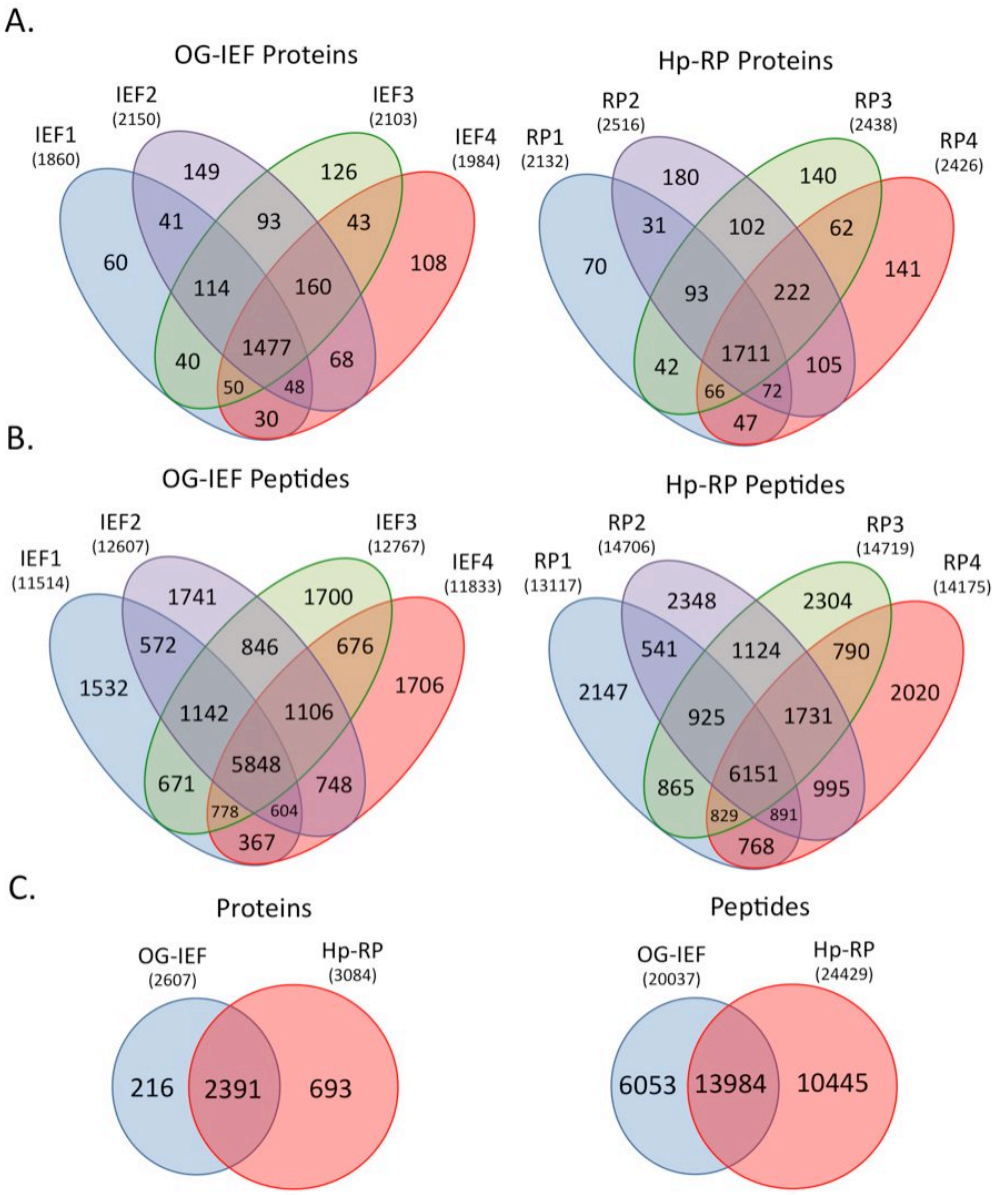


Figure 4.2. Quadruplicate analysis of OG-IEF or Hp-RP fractionated trypsin-digested PBMC lysate (75 ug/replicate). The overlap for (A) proteins identified and (B) peptides identified for OG-IEF and Hp-RP are shown. (C) Indicates the overlap of the total proteins and peptides from the combined replicates for OG-IEF and Hp-RP.

of protein and peptide identifications they both resulted in a similar average protein sequence coverage; 16.3% for each (OG-IEF SD = 12.7, Hp-RP SD = 11.53).

In order to measure the fractionation performance we the distribution of total peptides and peptides specific to a particular fraction for each technique was plotted (Figure 4.3). It is clear that the range of total peptides for the OG-IEF did not follow a uniform distribution ranging from 1020 to 2725 peptides per fraction. This also resulted in a similar trend for the fraction specific peptides (26-65%). In contrast, Hp-RP had an even spread of total and fraction specific peptides for all the fractions collected (Figure 4.3B). Determining the number of times a peptide was found in one, two, three, and four or more fractions was used to analyze the resolving power of the two techniques (Figure 4.4). While both techniques confined the majority of peptides to a single fraction, Hp-RP had significantly superior resolving capacity compared to OG-IEF (91.2% vs 76.8%, $p<0.001$).

Since the two techniques also differed drastically in the way they fractionate peptides in terms of biochemical properties (pI vs. Hydrophobicity) we determined the theoretical pI and GRAVY scores for all the peptides identified by each technique (Figure 4.5). It is generally accepted that hydrophobic peptides are poorly released from IPG-IEF due to the acrylamide matrix and so it was thought that Hp-RP might have a distinct advantage at identifying hydrophobic peptides and proteins. However, in either case both techniques were equally good at identifying hydrophobic peptides as well as peptides with different isoelectric points.

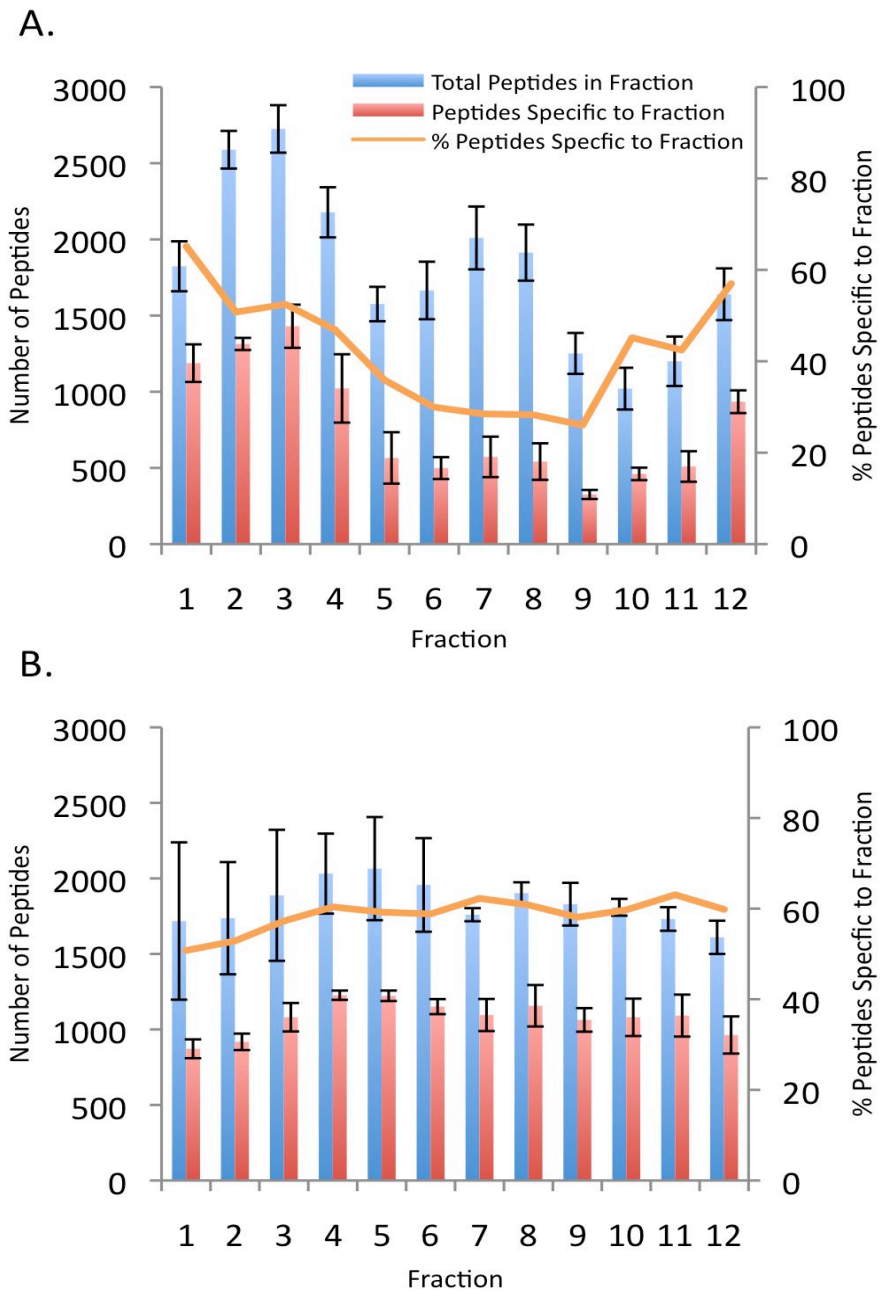


Figure 4.3. Distribution of peptide identifications from (A) OG-IEF and (B) Hp-RP. Total peptide identifications for each of the 12 fractions (blue bars) and the number of peptides identified in only one specific fraction (red bars) are shown. Error bars represent the standard deviation for a quadruplicate analysis. The average percent of peptides specific to a single fraction are plotted in orange.

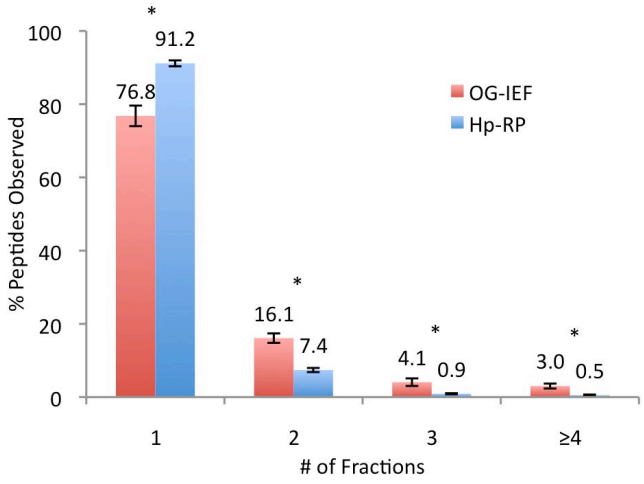


Figure 4.4. Peptides per number of fractions for OG-IEF and Hp-RP. The number of peptides identified was counted per the number of fraction(s) in which they were identified. Error bars represent the standard deviation of the mean for a quadruplicate analysis. Asterisks indicate p-values < 0.001 (Student's t-test).

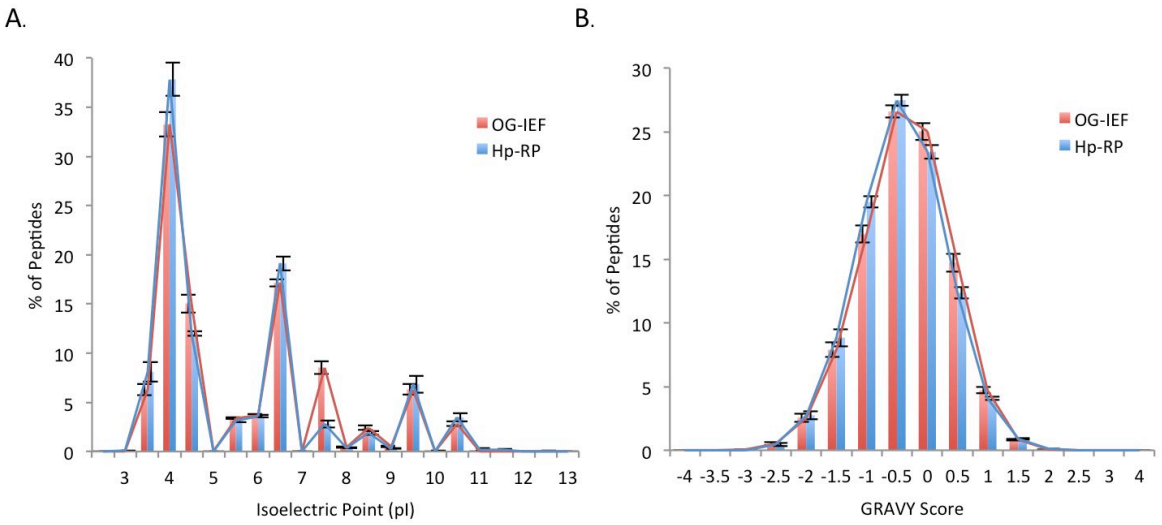


Figure 4.5. Biophysical properties for the percentage of peptides identified by nLC/MS/MS using the OG-IEF and Hp- RP techniques. The distributions for (A) theoretical pI and (B) GRAVY were determined for each peptide. Peptides were subsequently binned by pI or GRAVY score. Error bars represent the standard deviation of the mean for the quadruplicate analysis.

Due to the larger number of peptide and protein identifications, it was decided that Hp-RP would be the most logical fractionation approach to carry forward into subsequent studies.

4.3.3 Fractionation of iTRAQ Labeled Peptides

With the sample preparation and fractionation optimized we next tested the iTRAQ labeling approach. In order to make sure that iTRAQ labeling did not alter or interfere with the ability to reproducibly fractionate peptides, iTRAQ labeled and unlabeled PBMC digests were evaluated by Hp-RP fractionation. A PBMC digest was prepared with and without iTRAQ labeling (Section 2.2.3.5) and subjected to the optimized Hp-RP fractionation method. When we examined the resulting chromatograms we found that iTRAQ labeling drastically changed the elution profile of a complex PBMC peptide sample (Figure 4.6). The chromatogram was significantly compressed, altering the optimal peptide elution profile. A significantly compressed peptide elution profile would have undoubtedly reduced the number of peptides that could be identified and ultimately quantified by mass spectrometry. To examine how iTRAQ labels affected the elution profile, and by extension the hydrophobicity characteristics of peptides, we injected a six-peptide mix with and without iTRAQ labeling (Figure 4.7). Additionally this comparison would allow for a more precise analysis of the hydrophobic shift observed in the complex PBMC digest. The six-peptide mix was designed with a wide range of hydrophobicities, and is normally used for routine quality control of an Agilent 1200 LC system. The unlabeled 6 peptides (Figure 4.7; blue) indicate the expected elution times for this particular LC gradient, while the iTRAQ labeled peptides (Figure 4.7; red) show a

considerable amount of hydrophobic shift for the first four peptides. Based upon these findings a new gradient was designed by undertaking several successive experiments altering the gradient profile until the 6 iTRAQ labeled peptides achieved equal retention times to the non-iTRAQ labeled peptides. Another, iTRAQ labeled PBMC sample was prepared and subjected to the new LC gradient that was designed specifically for iTRAQ labeled PBMC digest. Figure 4.8 indicates that the new gradient efficiently corrected for the hydrophobicity shifts caused by labeling of peptides with iTRAQ reagents.

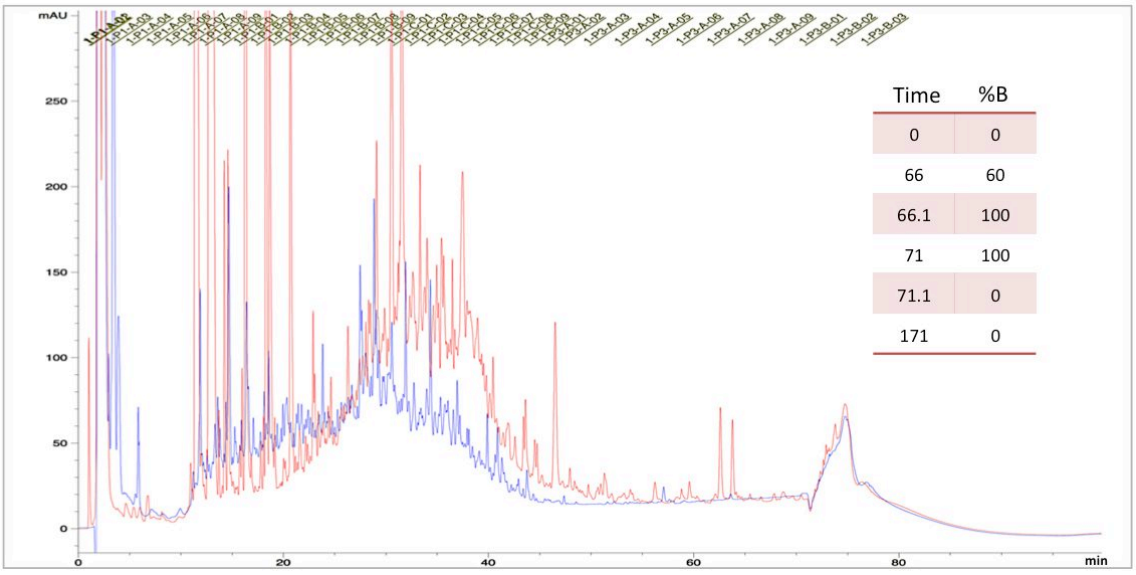


Figure 4.6. iTRAQ labeled (red) and unlabeled (blue) PBMC digest (75µg) was fractionation with the standard optimized LC protocol by Hp-RP .

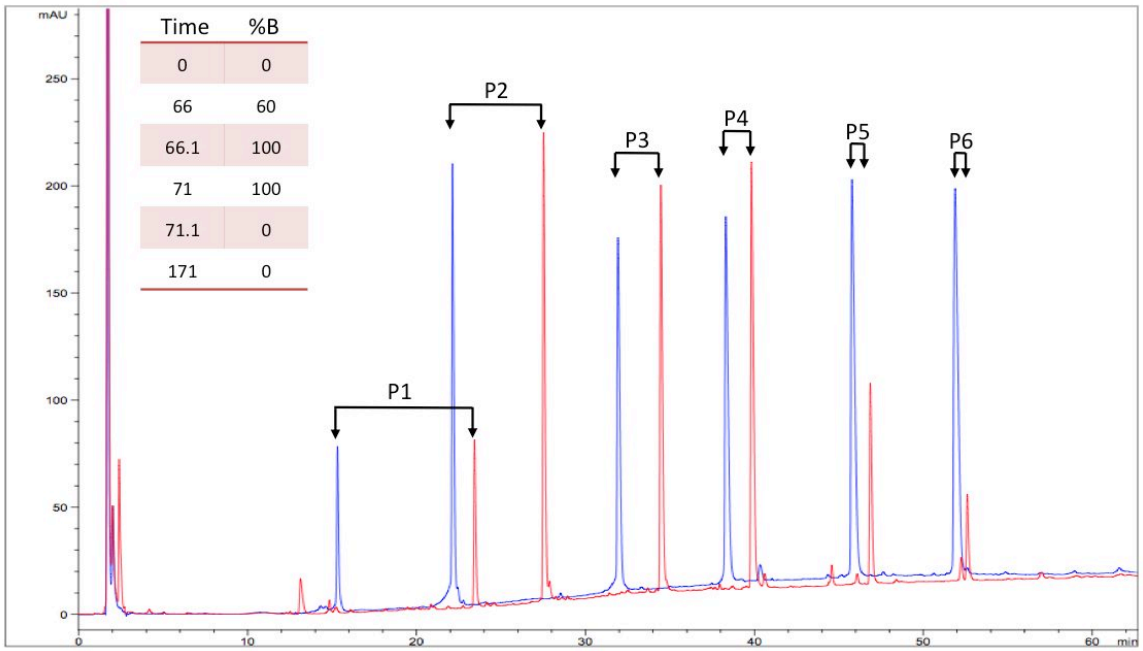


Figure 4.7. A comparison between iTRAQ labeled (red) and unlabeled (blue) 6 peptide standard using the optimized Hp-RP gradient for PBMC lysate.

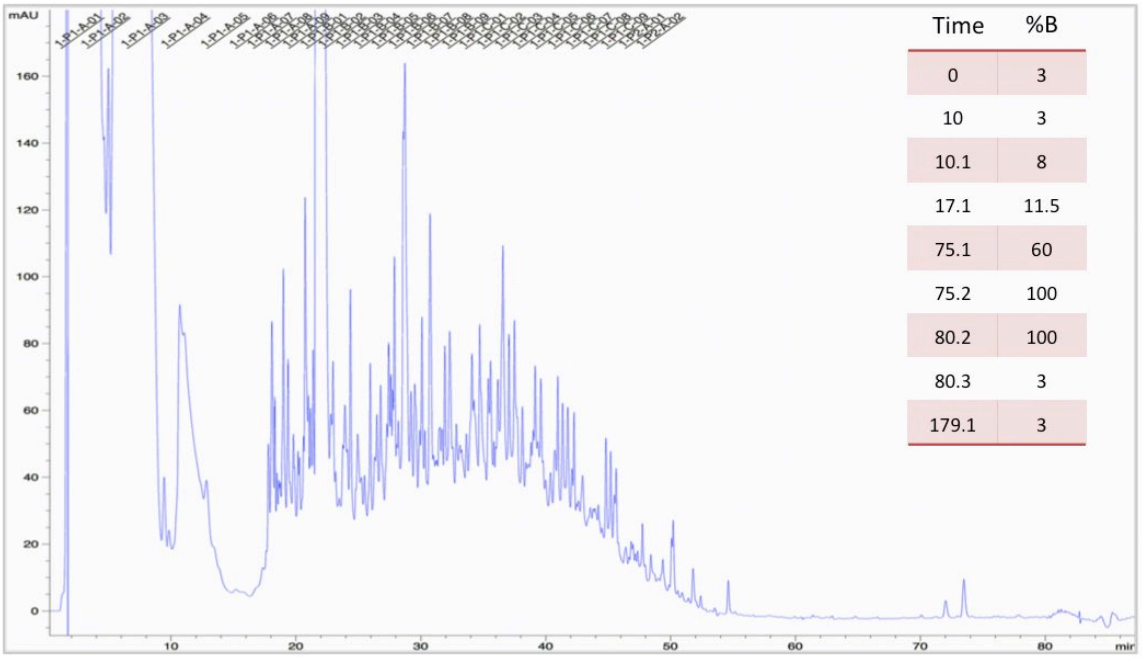


Figure 4.8. iTRAQ labeled PBMC digest (75µg) fractionated using the new iTRAQ specific Hp-RP LC gradient.

4.3.4 Mass Spectrometry Fragmentation of iTRAQ Labeled Peptide

The results obtained in Chapter 3 indicated that the Orbitrap XL mass spectrometer would be the clear choice for further analysis of complex immune cell digest in future studies. This decision proposed a unique problem in that the ion trap within the Orbitrap XL mass spectrometer, under normal operating conditions is unable to detect iTRAQ reporter ions during peptide fragmentation (1/3 rule). In order to solve this problem a specialized fragmentation process called high energy collision dissociation (HCD) was used to detect the iTRAQ reporter ions. In order for this method to work correctly a balance between optimal peptide identification and reporter ion intensity needed to be struck. To optimize this parameter we prepared a 6 protein mix iTRAQ labeled at a ratio of 2:1 (iTRAQ 117/116). This 6 protein mix was then analyzed in triplicate by nLC/MS/MS (Sections 2.1.2 + 2.2.4.5) on the Orbitrap XL mass spectrometer while varying the HCD collision energy (30-45%). Strong reporter ion intensities have been shown to be considerably more accurate than low intensity reporter ions [171] so we monitored the reporter ion intensity while varying the HCD collision energy. Figure 4.9A clearly indicates that as the HCD collision energy was increased the intensity of the reporter ions increased. Additionally, by calculating the reporter ratio while increasing the HCD collision energy (Figure 4.9B), successively more accurate reporter ratios were achieved (true ratio is 2:1). The major draw back to increasing the collision energy too much, is that it results in lower peptide identification rates. This is illustrated by plotting the number of peptides identified for each protein in the 6-protein mix. Figure 4.10 clearly indicates that

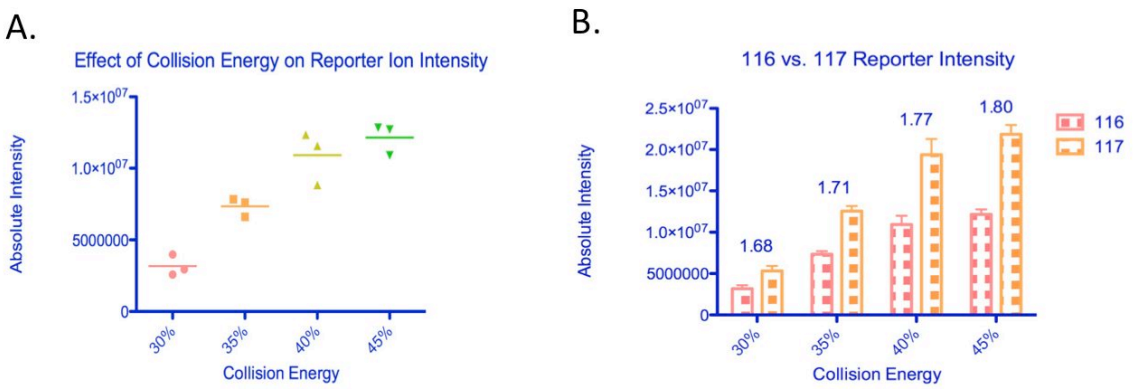


Figure 4.9. Effect of HCD normalized collision energy on the absolute intensity (A) and ratio accuracy (B) of iTRAQ reporter ions. This experiment was performed in triplicate with error bars representing the standard deviation of the mean.

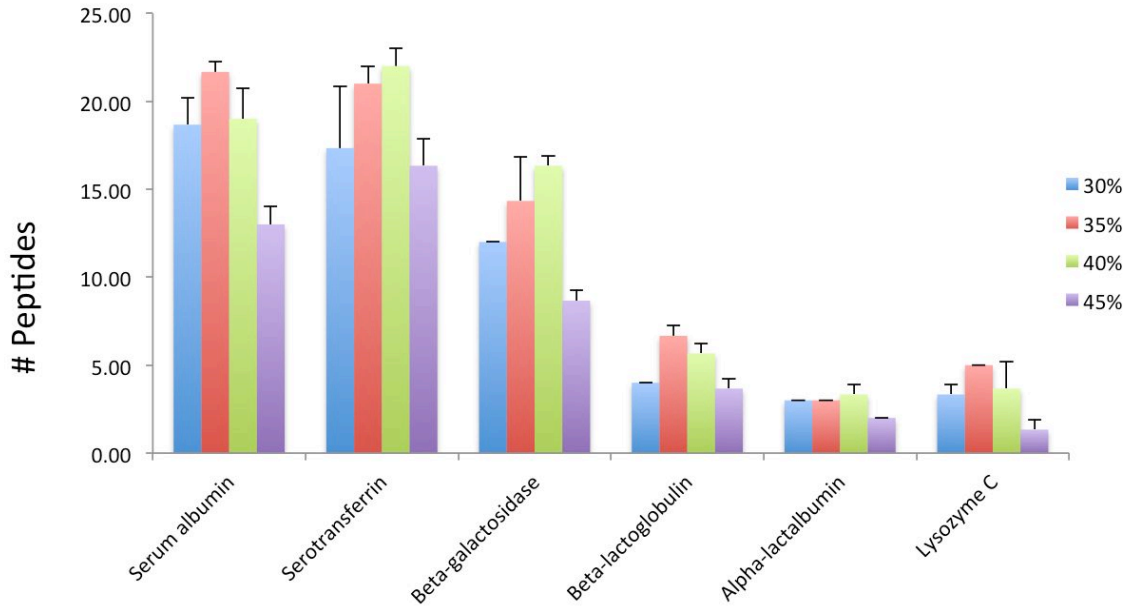


Figure 4.10. Peptides identified by mass spectrometry for six different proteins with varied HCD collision energies. This experiment was performed in triplicate with error bars representing the standard deviation of the mean.

once an HCD energy of greater than 40% was applied the number of peptides identified for each protein was significantly reduced. An HCD collision energy of ≤30% and ≥45% was clearly suboptimal leading to a loss in the number of peptides that could be effectively identify. These results indicated that an optimal HCD collision energy would be between 35-40%. We therefore, employed an HCD collision energy of 40% for all subsequent iTRAQ experiments that made use of the Orbitrap XL.

4.4 Summary

Comparative analysis of the Hp-RP and OG-IEF fractionation techniques illustrated that Hp-RP consistently led to the identification of more proteins and peptides, superior resolution and an improved peptide distribution profile. This data indicated that the Hp-RP fractionation method was superior to the OG-IEF method of fractionation. iTRAQ labeling of peptides was found to drastically shift the retention profile of PBMC digest and as a result, a new iTRAQ specific Hp-RP LC gradient was optimized to achieved maximum identification and quantification of proteins and peptides after mass spectrometry analysis. Finally, optimal collision energy parameters were achieved to maximize the quantification accuracy of the iTRAQ reporter ions while still allowing for optimal peptide identification. These techniques were subsequently applied to an iTRAQ proteomics approach to measure immunologically relevant proteins which may confer resistance to HIV infection in HESN women.

Chapter 5. iTRAQ proteomics in peripheral blood mononuclear cells.

5.1 Introduction & Rationale

Mass spectrometry based approaches for biomarker discovery have become an increasingly popular systems biology tool over the past decade. The capacity now exists to characterize immunological responses in a model antigen stimulation system and measure ex-vivo heterogeneity of PBMC responses from healthy donors to address the major objectives of this thesis. qMS has yet to be routinely used as a tool to measure specific host immune responses to antigenic or pathogen challenge [172]. iTRAQ was developed to multiplex biological samples into a single mass spectrometry experiment in a high-throughput manner [138]. The multiplexing of samples into a single experiment significantly reduces the systematic and technical error often seen in some mass spectrometry experiments. This high throughput capacity can now be used, for example, to compare host immune responsiveness under a variety of conditions, including antigen specific and mitogen responses, and allows for the relative quantitation of thousands of protein readouts at once in a given biological sample or experimental condition [173,174].

Vaccine efficacy and immunogenicity are often measured, using decades old assays like INF-γ ELISPOT or the measurement of single cytokine responses by ELISA. However, these assays are limited in that they assess only a single or at most a few correlates of protection previously identified in the literature. It has been known for some time that measuring a single parameter can miss important poly-functional responses that have been known to correlate with positive immune outcomes [87].

Deciding which proteins to measure by ELISA or bead-arrays can also be problematic in that there are numerous cytokines that may correlate with immunogenicity. Additionally, these approaches are reductionist, and based upon current dogma or understanding, which may not reflect the best, or most efficacious readout. Systems biology approaches such as RNA expression analysis using micro-arrays or RNA-Seq approaches have also been commonly used to study correlates of protection in the vaccine arm of clinical trials[175], however it has been well established that mRNA expression ultimately does not correlate well with protein expression with some reports showing extremely low correlation (as low as r = 0.27) [126-129]. Protein expression is the ultimate end product defining a specific biological phenotype, and for this reason, can accurately describe what is occurring within a cell or a specific tissue sample at a phenotypic level. Thus, proteomics approaches can now be applied to solve some of the limitations of previous immune assays.

In order to evaluate the iTRAQ quantitative shotgun proteomics approach as a tool for measuring immune responses the proteome of activated PBMC was assessed. Immune activation due to concurrent STI's and perturbation of the mucosal immune environment has been well described as a risk factor for HIV acquisition. Additionally, the altered susceptibility observed by us and others in HESN women correlates with decreased immune activation in target cells and a reduced ability to support HIV replication [176]. Thus, determining susceptibility to HIV infection requires an understanding of cellular immune activation. Further, one of the most

common models used for ex-vivo assessment of HIV replication uses the mitogen, phytohemagglutinin (PHA) to activate immune cells. Phytohemagglutinin is a plant-derived lectin that binds complex carbohydrates on the surface of lymphocytes and can act as a powerful immune inducer. PHA is commonly used in the laboratory to activate and expand immune cells in-vitro [177]. In this chapter PHA was used as a surrogate model for the activation state that occurs in-vivo during HIV acquisition and infection.

The utility of qMS using iTRAQ to define cellular immune activation pathways and to quantify the proteome changes due to mitogen (PHA) activation of primary lymphocytes over a 24-hour time period was determined. Activation of lymphocytes by PHA is known to trigger cell growth, numerous immune response pathways, and allows cells to become readily infected with primary and lab adapted HIV isolates [178]. This is a useful system to test the iTRAQ approach and its ability to measure ex-vivo immune pathways from HESN women described in (Chapter 6 & 7)

5.2 Objectives

1. Successfully identify and quantify immunologically relevant proteins using qMS in PHA stimulated peripheral blood mononuclear cells.

2. Identify HIV infection specific functional networks associated with PHA activation of peripheral blood mononuclear cells.

5.3 Results

5.3.1 Early antigen specific proteomic responses in PBMC

The primary goal of this chapter was to evaluate the ability of an iTRAQ proteomics platform to reliably measure immunologically relevant proteins and functional networks important in HIV replication. A quantitative shotgun proteomics analysis using an iTRAQ approach was conducted on freshly isolated PBMCs from 4 healthy donors stimulated with PHA for 0, 1, 3, and 24-hours (Figure 5.1). Analysis of detected peptides resulted in the identification of 2472 unique proteins across all 4 donors with an average of 1831 unique proteins identified per donor. The data was subsequently filtered to include any protein with quantitative data for 3 or more patients at any specific time point. The filtered data resulted in a total of 1816 proteins identified and quantified across 3 or more study participants. Of the 1816 proteins, 220 were significantly deregulated between at least two time points (Anova, Benjamini Hochberg, $q<0.05$). The twenty most significantly deregulated proteins at 24 hours, post-PHA stimulation as well as their expression at 1 and 3 hours are identified in Table 5.1.

At 24 hours, proteins involved in the immune response, cell growth, and anti-apoptosis were among the 10 most significantly up-regulated proteins including IDO1, SERPINB2, NAMPT, TRAF1, MARCKS, FTH1, APOL2, WARS, DDX21, and SOD2

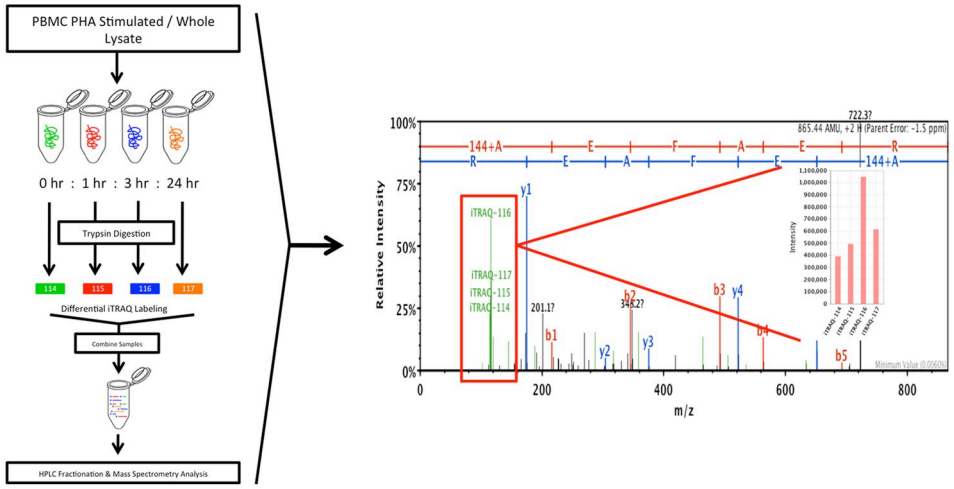


Figure 5.1. Experimental workflow for PHA stimulation of peripheral blood mononuclear cells. PBMCs were isolated from four healthy donors and stimulated for up 24 hours with phytohemagglutinin. Cells were collected at various time points, lysed, trypsin digested, labeled with iTRAQ followed by offline-HPLC fractionation. Fractions were collected and then analyzed by nLC/MS/MS to describe proteomic changes in response to immune activation over time.

IPI Accession	Protein Name	1hr*	3hr*	24hr*	Exp. Chart 1,3,24
IPI00028096	Indoleamine 2,3-dioxygenase 1 (IDO1)	0.400	0.100	3.450	
IPI00007117	Plasminogen activator inhibitor 2 (SERPINB2)	0.150	1.025	2.950	
IPI00018873	Nicotinamide phosphoribosyltransferase (NAMPT)	0.075	0.325	2.275	
IPI00011549	TNF receptor-associated factor 1 (TRAF1)	0.267	0.200	2.267	
IPI00219301	Myristoylated alanine-rich C-kinase substrate (MARCKS)	-0.075	0.425	1.925	
IPI00554521	Ferritin heavy chain (FTH1)	0.025	0.625	1.825	
IPI00877964	Apolipoprotein L2 (APOL2)	-0.025	-0.100	1.750	
IPI00295400	Isoform 1 of Tryptophanyl-tRNA synthetase (WARS)	-0.075	-0.025	1.700	
IPI00015953	Isoform 1 of Nucleolar RNA helicase 2 (DDX21)	0.125	0.050	1.600	
IPI00022314	Superoxide dismutase [Mn], mitochondrial (SOD2)	0.150	0.200	1.575	
IPI00012555	Ficolin-1 (FCN1)	-0.350	-0.650	-1.750	
IPI00019038	Lysozyme C (LYZ)	-0.400	-0.125	-1.450	
IPI00027410	Platelet glycoprotein V (GP5)	0.167	0.000	-1.367	
IPI00021885	Isoform 1 of Fibrinogen alpha chain (FGA)	0.075	0.125	-1.250	
IPI00298497	Fibrinogen beta chain (FGB)	0.200	0.275	-1.175	
IPI00022445	Platelet basic protein (PPBP)	0.175	-0.075	-1.175	
IPI00021891	Isoform Gamma-B of Fibrinogen gamma chain (FGG)	0.150	0.100	-1.150	
IPI00292532	Cathelicidin antimicrobial peptide precursor (CAMP)	-0.200	-0.300	-1.150	
IPI00914602	Amyloid beta A4 protein isoform f precursor (APP)	0.133	0.067	-1.133	
IPI00930124	Putative uncharacterized protein (THBS1)	-0.100	-0.150	-1.100	

*Log_2 ratio of each time point compared to the media control

Table 5.1. The top 20 differentially expressed proteins by 24 hours and their changes at 1 and 3 hours post-PHA stimulation. Expression levels for each time point are expressed as the Log_2 ratio normalized to the media control ($q<0.05$).

(q<0.05). The majority of these proteins were also moderately up-regulated at 1 and 3 hours followed by a larger increase at 24 hours. A number of the down-regulated proteins in response to PHA stimulation at 24 hours are also implicated in the immune response including inflammation, cell trafficking, and cell-cell signaling. The 10 proteins exhibiting the largest decrease compared to controls included FCN1, LYZ, GP5, FGA, PPBP, FGB, FGG, CAMP, APP, and THBS1 (q<0.05). Interestingly, some of these proteins were initially induced at 1 and 3 hours, however they become reduced in abundance by 24 hours. Overall, qMS was able to reliably detect considerable protein changes directly related to immune activation in response to PHA stimulation within 24 hours.

5.3.2 Cluster & Correlation of Proteome Responses to PHA

In order to describe and visualize the response to PHA stimulation across all donors for each time point, cluster analysis was performed on all 1816 quantified proteins (Figure 5.2 A). The resulting heat map indicates that individual donor responses to stimulation at 1 and 3 hours clustered closest together, and that responses within an individual at this time point were more similar than changes between individuals. However by 24-hours post-stimulation the proteome changes amongst all donors clustered tightly together regardless of the individual responses. It is interesting to note that the proteome changes in D02 at 1 and 3-hours clustered more closely to the 24-hour stimulation time point suggesting that we may have detected increased baseline immune activation in this individual. When we repeated the analysis only considering the 220 proteins that were statistically different when induced by

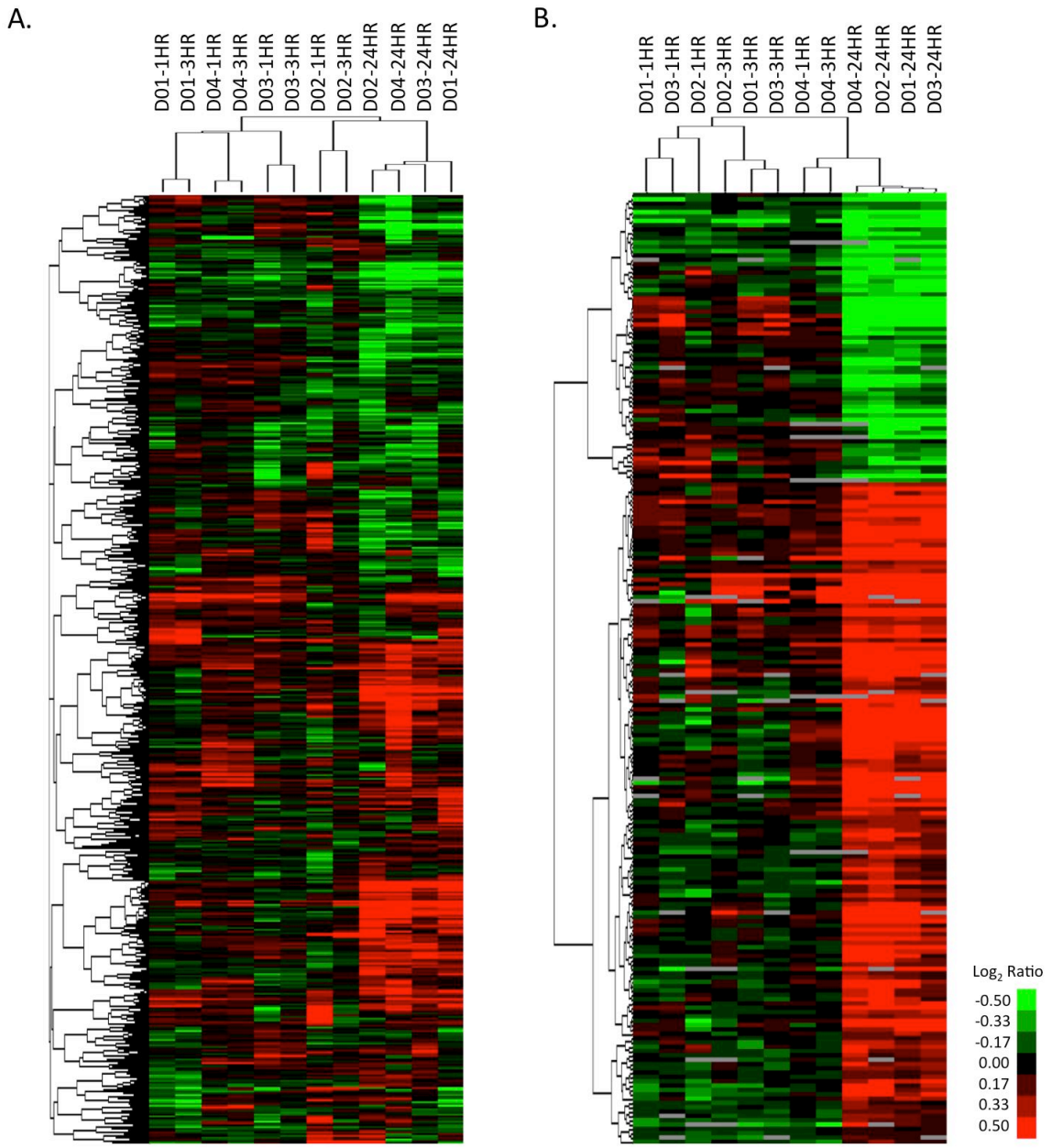


Figure 5.2. Hierarchical cluster analysis of (A) all 1816 identified and quantified proteins and (B) the 220 significantly deregulated proteins in response to 24-hour PHA stimulation of peripheral blood mononuclear cells in four donors. Dendrograms were generated using an un-centered correlation similarity metric and complete linkage clustering.

PHA (Figure 5.2 B) we observed tighter clustering for each of the sampled time points between donors with the exception of D04. This may suggest a more homogeneous response between individuals when looking at proteins that were the most strongly induced by immune activation. To further understand the correlation between individual donor responses to PHA, and the time of induction of these responses, a Pearson correlation matrix was generated comparing all 1816 proteins for each donor and time point (Figure 5.3 A). The analysis indicated that the proteome changes observed for the 1 and 3-hour time points did not correlate between individuals, for example no correlation was observed between D01 and D02 at 1 hour (Figure 5.3 B) indicating significant heterogeneity between patients at early time points. However, within any individual patient the 1 and 3-hour time points were most strongly correlated (Figure 5.3 C) and the strength of this correlation diminished between 1 hour and 24-hours and 3-hours and 24-hours. Once the 24-hour time point was reached the proteomic profile between subjects became more homogeneous and correlated strongly, making differences between individual patient responses relatively indistinguishable. For example, the correlation analysis between D02 and D04 (Figure 5.3 D) illustrates that after 24-hours of PHA stimulation the individual donor responses correlated strongly ($r = .66$; $p<0.0001$). Additionally, individual responses to PHA stimulation at early time points were more subject dependent than antigen dependent suggesting that individual subjects may take different pathways to reach a particular immune activation profile.

A

r-values

	D01-1hr	D02-1hr	D03-1hr	D04-1hr	D01-3hr	D02-3hr	D03-3hr	D04-3hr	D01-24hr	D02-24hr	D03-24hr	D04-24hr
D01-1hr												
D02-1hr	0.06											
D03-1hr	0.23	-0.07									r > 0.5	
D04-1hr	0.16	0.08	0.19								r > 0.3	
D01-3hr	0.79	0.03	0.12	0.19								
D02-3hr	-0.01	0.40	0.04	0.05	0.05							
D03-3hr	0.15	0.14	0.60	0.16	0.20	0.08						
D04-3hr	0.14	0.11	0.18	0.81	0.15	0.06	0.19					
D01-24hr	0.39	-0.02	0.02	0.15	0.51	0.09	0.07	0.12				
D02-24hr	-0.04	0.24	-0.07	0.06	-0.01	0.57	-0.04	0.10	0.54			
D03-24hr	0.06	0.05	0.32	0.17	0.01	0.07	0.32	0.20	0.57	0.59		
D04-24hr	0.11	0.20	0.07	0.45	0.06	0.20	0.00	0.42	0.62	0.66	0.66	

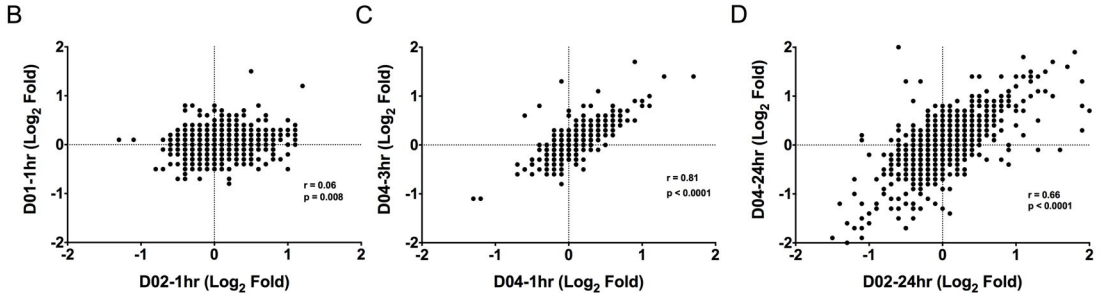


Figure 5.3. (A) Correlation matrix demonstrating the correlation strength of proteomic responses between all patients and all PHA stimulation time points. While all 1816 proteins had significant correlations ($p < 0.0001$) the strength of these associations differed considerably. Those denoted in blue had an r-value between 0.3 to 0.5, while those in yellow had a r-value > 0.5. Representative plots indicating the correlation and breadth of proteomic responses for (B) at the 1-hour time point between D01 and D02, (C) between 1 and 3 hour time point for D04 and (D) at the 24-hour time point between D02 and D04.

5.3.3 Biological Networks & Functions Induced by PHA Stimulation

We hypothesized that cellular functional pathways affecting cell proliferation and the replication of HIV would be the main pathways induced in response to stimulation with PHA. In order to test this hypothesis the 220 significantly differentially expressed proteins in the data set were subjected to Ingenuity Pathway Analysis (IPA). IPA allows for a detailed analysis of the significantly enriched pathways and functions deregulated by PHA stimulation of PBMC's. Our analysis focused on the 220 proteins that were statistically de-regulated since these proteins were responding to PHA stimulation. Also, since the 24-hour time point yielded the greatest protein expression differences, we focused our analysis on this time point.

Surprisingly, the analysis identified and predicted significant decreases in several biological functions in response to 24 hours of PHA stimulation. Biological functions involved in cell death and survival, hematological system development and function, cell to cell signaling, cellular movement/immune cell trafficking, inflammatory response and tissue development were significantly repressed by PHA stimulation (Blue). These functions are represented with a heat map generated using IPA in the center of Figure 5.4. However, as expected significant up-regulation was observed in functional pathways important in immunity to infectious diseases, cell growth and proliferation, and specific pathways in cell death and survival (orange). Next, we analyzed specific protein changes induced by PHA within a given sub-functional

network of interest for all donors and all time points (Figure 5.4 A-F). It is important to note that within a particular sub-function, proteins can be up-regulated and still have a significant predicted decrease in that function. Apoptosis (z=-3.365), cell movement (z=-2.275) and adhesion of immune cells (z=-2.534) all had negative z-scores of 2 or less indicating statistically significant down-regulation of these functions (Figure 5.4 A,B and E). Functions that were predicted by IPA to be up-regulated were cell survival (z=1.442), HIV-infection (z=2.906), and proliferation of cells (z=2.363) (Figure 5.4 C,D and F). While cell survival (Figure 5.4 D) did not reach a significant z-score (z=1.442) it was trending towards significance and fits into the overall effect seen within the broader category of cell death and survival. Thus, the systems biology approach described here was able to identify numerous immunological functions that we would expect to be induced by PHA stimulation (i.e. HIV infection or proliferation) as well as novel functions not expected to be decreased in response to PHA (cell-to-cell signaling or tissue development).

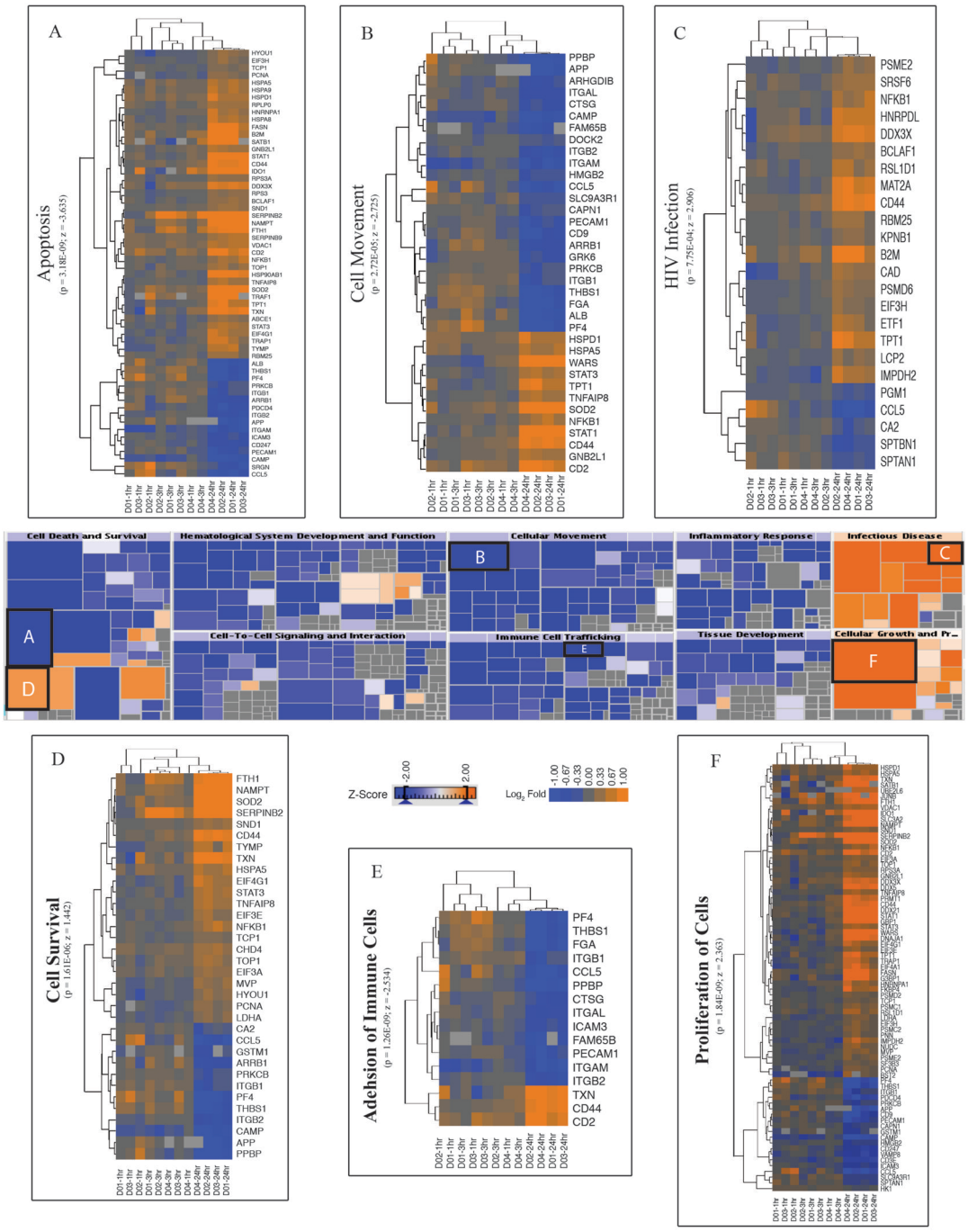


Figure 5.4. Functional map of proteins differentially expressed by 24 hours post-PHA stimulation. The center map was generated using IPA where the size of the squares are proportional to the number of proteins within the indicated functional network. The degree of blue/orange in the center map indicates the degree of over/under expression (z-score). Cluster analysis was performed on all the proteins within specific functional networks of interest (expanded clusters A-F) where expression levels represent Log_2 Fold change. Blue indicates a down-regulation of the indicated function or protein while orange indicates up-regulation.

5.4 Summary

Using a quantitative shotgun proteomics approach we were able to observe that PHA-stimulation of PBMC results in strong proliferative, anti-apoptotic specific responses, along with the induction of functions that may lead to increased viral replication, particularly with regards to HIV. This proliferative response complements the most widely accepted models of immune activation and the pathways utilized for HIV infection in-vivo, suggesting that early induction of immune activation pathways may vary between individuals. These observations also suggest that iTRAQ appears to be suitable for observing an immune quiescence phenotype at the proteomic level in PBMC of HESN women. During the early stages of immune activation significant heterogeneity existed between subjects tested indicating that individuals may utilize different functional pathways to reach a common immune response profile by 24-hours. In addition, several functional pathways expected to be overexpressed in actively proliferating cells like adhesion and migration were significantly suppressed, highlighting the ability of this approach to identify novel immune response pathways. These results suggest that iTRAQ mass spectrometry is a useful tool for monitoring host immune responses and lays the foundational groundwork for the use of quantitative shotgun proteomics in measuring the ex-vivo immuno-proteomic profile in HESN sex workers from Nairobi, Kenya. This study represents a comprehensive quantitative analysis of the proteome changes occurring in a mixed cell population (PBMC) in response to ex-vivo PHA stimulation, and is a useful tool for describing the global proteomic response in host immune activation. Furthermore, systems biology approaches such as the one described here have

the potential to examine multiple functional networks when studying immune activation and responses in the context of HIV vaccine or microbicide trials.

Chapter 6. Comparative proteomic analysis of PBMC from HESN women

6.1 Introduction & Rationale

The central hypothesis of this thesis is that HESN women will display an immune quiescent phenotype at the proteomic level similar to the genetic characterization studies we have previously published. The iTRAQ proteomics pipeline evaluated in the previous three chapters was shown to be able to identify immunologically relevant proteins in the context of HIV and is an appropriate platform for the characterization of immune cells from HESN women. This chapter presents the findings from a pilot study to identify and validate proteomic correlates of HIV protection in the PBMC of HESN sex workers from the Pumwani sex worker cohort.

6.2 Hypothesis

1. Quantitative shotgun proteomics of systemic immune cells from HESN women will demonstrate a proteomic profile characteristic of the immune quiescent phenotype, described by the altered expression of pathways important in immune activation, cell recruitment / migration, and proteins involved in HIV replication pathways.

6.3 Objective

1. Utilize a qMS approach to determine the proteomic profile of ex-vivo PBMC from HESN and susceptible controls and validate protein biomarkers of interest by Western blot analysis.

6.4 Results

To evaluate the cellular proteome of systemic immune cells, freshly isolated PBMC from HESN and HIV susceptible women of the Pumwani sex worker cohort were processed using the procedures developed in the first two chapters of this thesis (Section 2.2.3 – 2.2.4). A total of 12 HESN and 18 high-risk HIV susceptible individuals were available for analysis using the iTRAQ shotgun proteomics platform. A pool consisting of an equal amount of protein from each individual patient in the study was used as an internal control for which all samples were later compared. Each 4-plex experiment used 4 iTRAQ tags that correspond to the mass/charge ratio of the reporter generated during peptide identification (114-117). Each iTRAQ 4-plex experiment consisted of the pooled control (114), one HESN individual, an HIV susceptible control, and one HIV infected subject to minimize any run-to-run variability. Additionally, the remaining iTRAQ tags were randomized when labeling each particular study group, this for example would eliminate the possibility of all the HESN women being labeled with the same iTRAQ tag, thereby reducing any iTRAQ tag specific variability within the experiment. This experimental setup resulted in a total of 15 separate 4-plex iTRAQ experiments, each one of which was subjected to Hp-RP offline fractionation followed by mass spectrometry analysis on the Orbitrap Velos mass spectrometer, outlined in Section 2.2.4.5 (Figure 6.1). This analysis scheme resulted in a total of one month of online mass spectrometer sample run-time. The raw data was processed through the Mascot search algorithm and visualized in Scaffold Q+ as described in Section 2.2.4.6.

Analysis of the proteomic data collected from PBMC resulted in an average of 1893 proteins for each individual iTRAQ experiment with an average of 42,556 peptides per run (Table 6.1). The overall protein false discovery rate was 0.1% while the peptide false discovery rate was 1.5%. However, combining all the results identified and quantified a total of 3105 proteins across all 45 study participants. Exporting these results we further restricted our analysis to include proteins that had at least 3 data points per study group (HESN, NEG). This resulted in a total of 2913 proteins for down-stream analysis and ensured that a standard deviation for each protein could be calculated when analyzing individual protein expression for targeted analysis.

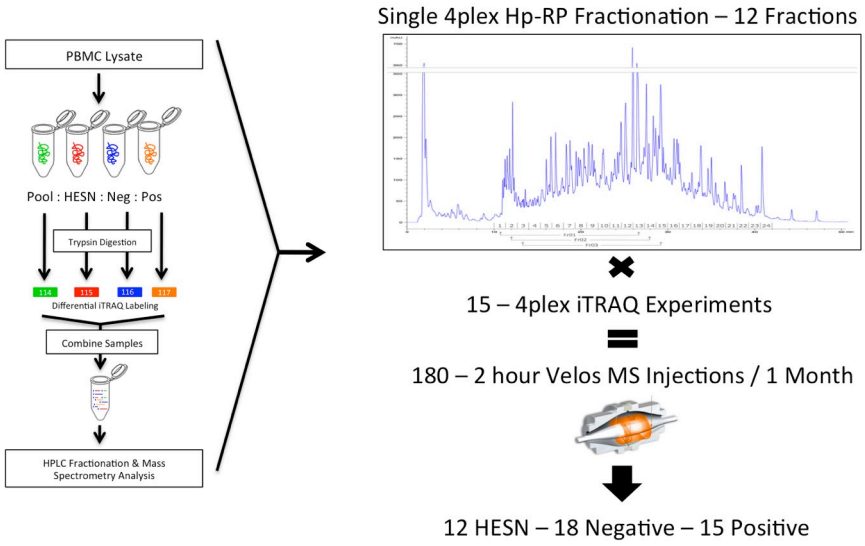


Figure 6.1 Experimental outline for the characterization of the immuno-proteomic profile of peripheral blood mononuclear cells from HESN women.

Bio-Sample	#Prot	#Peptides	#Spec	%IDs
Sample A	1720	42089	201970	20.84%
Sample B	1988	55206	227116	24.31%
Sample C	1616	43992	203505	21.62%
Sample D	1941	49282	219550	22.45%
Sample E	1568	40299	207683	19.40%
Sample F	1962	46585	213825	21.79%
Sample G	1825	36632	183368	19.98%
Sample H	2322	54504	216363	25.19%
Sample I	2135	38424	182100	21.10%
Sample J	1966	42840	205077	20.89%
Sample K	1882	34798	179333	19.40%
Sample L	1901	30357	178924	16.97%
Sample M	1808	33978	191924	17.70%
Sample N	1976	44764	208968	21.42%
Sample O	1778	44583	215236	20.71%
Average	1893	42556	202329	20.92%

Table 6.1 Cumulative statistics from HESN PBMC pilot study. Bio-Sample refers to the particular iTRAQ run (15 in total), while the number of proteins identified, peptides identified, total spectra (# of MS/MS scans), and the percentage of spectra (# of MS/MS scans identified as a peptide in the database) is also tabulated for each iTRAQ experiment.

6.4.1 Significantly deregulated proteins in PBMC from HESN women

A statistical comparison between HESN and susceptible controls for all 2913 proteins quantitatively identified was undertaken to identify immunological biomarkers of interest. The data was found to be normally distributed (residuals from nonlinear regression) and thus a student's T test was used for protein comparisons. However, after correcting for multiple comparisons by various methods (Bonferroni, Benjamini-Hochbeg, Storey) no proteins were able to survive the multiple comparison analysis. Not correcting for multiple comparisons we chose a p-value cutoff of 0.05, which led to the identification of 136 differentially, expressed proteins between HESN and high-risk negative controls. Additionally, a fold change cutoff of ±1.5 as our criteria for identifying proteins of further interest was applied (Figure 6.2). This cutoff was chosen mainly because this was the threshold in which we could confidently detect a protein expression difference by Western blot. We found that there were a total of 7 under-expressed and 19 over-expressed proteins in HESN women compared to negative controls that fit these cutoff criteria (Figure 6.2). The top 10 differentially regulated proteins of which six proteins were chosen for Western blot confirmation are listed in Table 6.2. Statistically significant proteins for validation by Western blot were chosen either because they had also been identified in past studies of HESN women or if they had a documented link in infection and immunity. HLA-C was previously identified to be under-expressed in a micro-array RNA expression analysis study of whole blood from HESN women [124], while alpha-2-HS glycoprotein and a similar protein to

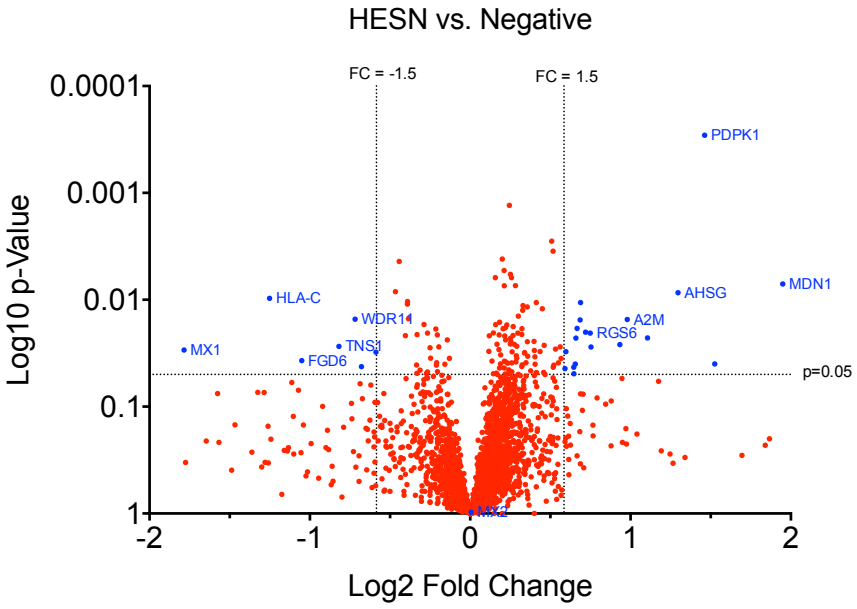


Figure 6.2. Immuno-proteomic profiling of PBMC from HESN women identified several proteins of statistical and biological significance with a cutoff of $p<0.05$ and a FC>1.5.

IPI Accession	Protein Name	Log_2 Fold HESN/NEG	p-value
IPI00167949	MX1, Interferon-induced GTP-binding protein Mx1	-1.78	0.0295
IPI00657976	HLA-C, Uncharacterized protein	-1.25	0.0097
IPI00419100	FGD6, Isoform 1 of FYVE, RhoGEF and PH domain-containing protein 6	-1.05	0.0371
IPI00307545	TNS1, Tensin-1	-0.82	0.0273
IPI00412224	WDR11, WD repeat-containing protein 11	-0.72	0.0152
IPI00167941	MDN1, Midasin	1.95	0.0071
IPI00002538	PDPK1, Isoform 1 of 3-phosphoinositide-dependent protein kinase 1	1.46	0.0003
IPI00022431	AHSG, Alpha-2-HS-glycoprotein	0.98	0.0153
IPI00478003	A2M, Alpha-2-macroglobulin	0.75	0.0276
IPI00216322	RGS6, Isoform 2 of Regulator of G-protein signaling 6	0.75	0.0205

*Log_2 ratio of each time point compared to the media control

Table 6.2. Top 10 proteins of interest over and under expressed in HESN women compared to high-risk negative controls for follow up in Western blot confirmatory analysis.

alpha-2 macroglobulin were identified to be over-expressed in the genital secretions of HESN women from the same cohort [111,112]. MX1, an innate interferon response protein is essential for clearance of influenza infection and other RNA viruses and had yet to be studied in the context of HIV [179].

6.4.2 Protein biomarker validation in HESN women

Due to a limited amount of protein sample left after conducting the qMS experiments, a multiplex Western blot strategy was undertaken to maximize the number of targets that could be validated at once. Antibodies from several commercial companies were obtained for validation of A2M, FGD6, PDPK1, AHSG, RGS6, HLA-C, and Mx1. Alpha-2 macroglobulin was shown to be a plasma contaminant and as it is not expressed in peripheral blood mononuclear cells it was omitted from further validation. Figure 6.3 clearly shows the presence of A2M in serum but not in purified PBMC cell lysate. Unfortunately, antibodies for three other proteins of interest (FGD6, PDPK1, AHSG) were not suitable for Western blot analysis. In the case of FGD6 and AHSG we were unable to obtain any binding/detection of the protein while the PDPK1 antibody produced non-specific binding. Antibodies from several companies were obtained in an attempt to validate these proteins however no reliable antibodies were found. A single western blot was optimized to quantify HLA-C, MX1, and RGS6 using Actin-β as a loading control with the Odyssey CLx infrared imaging system. The infrared dye coupled secondary antibodies allow for overlapping proteins to be quantified separately and is often used for phosphorylation analysis. To validate our makers of interest a total of 16μg

of PBMC lysate for each individual in the study was loaded onto a SDS-PAGE gel along with a standard serial dilution of

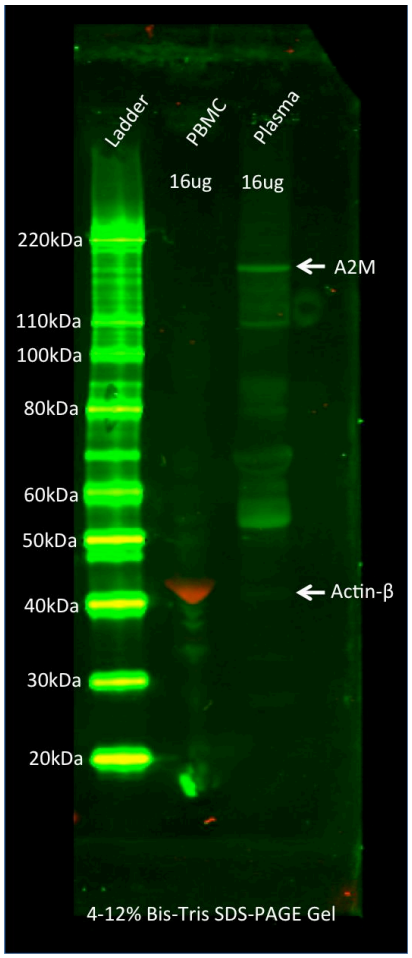


Figure 6.3. Western blot analysis of PBMC lysate and plasma for Alpha-2 macroglobulin (A2M). A2M is present in plasma however is not expressed in PBMC.

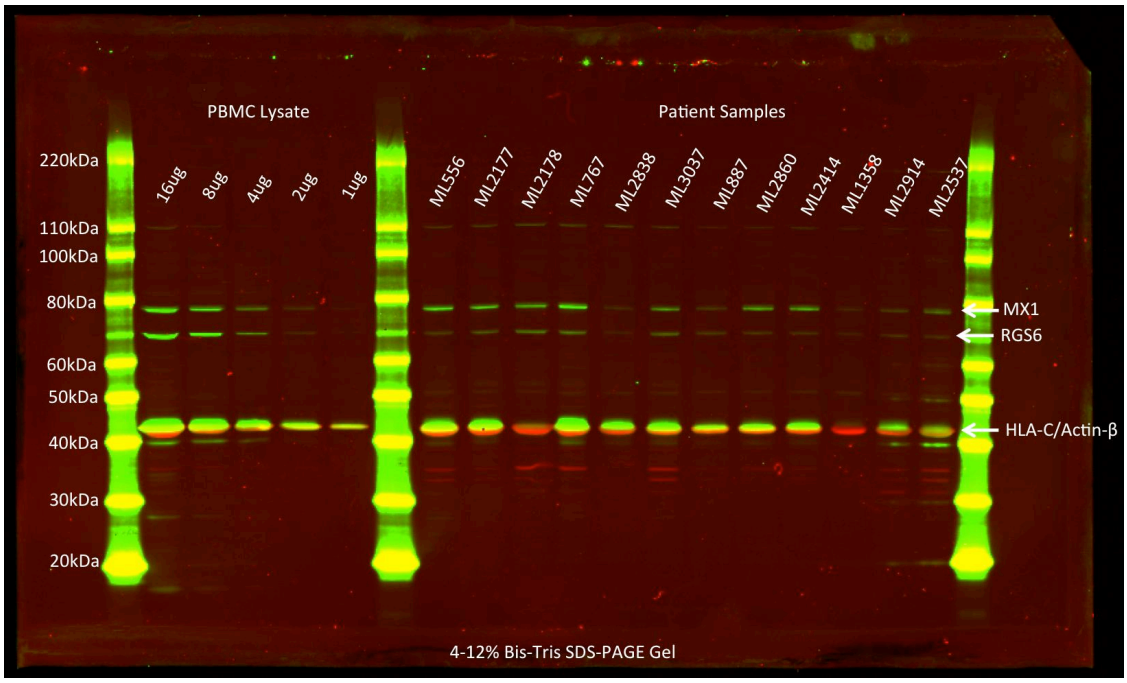


Figure 6.4. Western blot confirmation of MX1, RGS6 and HLA-C (Green) expression in PBMC from individual patient samples (HESN, NN, POS). All samples were normalized to Actin-β (Red).

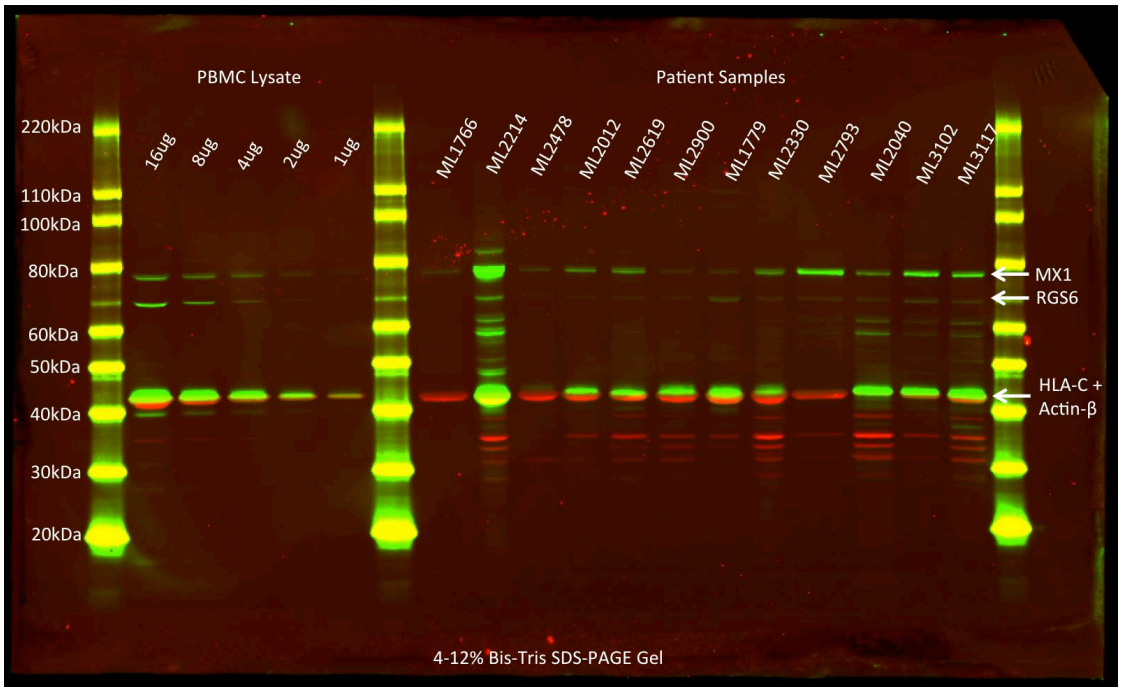


Figure 6.5. Western blot confirmation of MX1, RGS6 and HLA-C (Green) expression in PBMC from individual patient samples (HESN, NN, POS). All samples were normalized to Actin-β (Red).

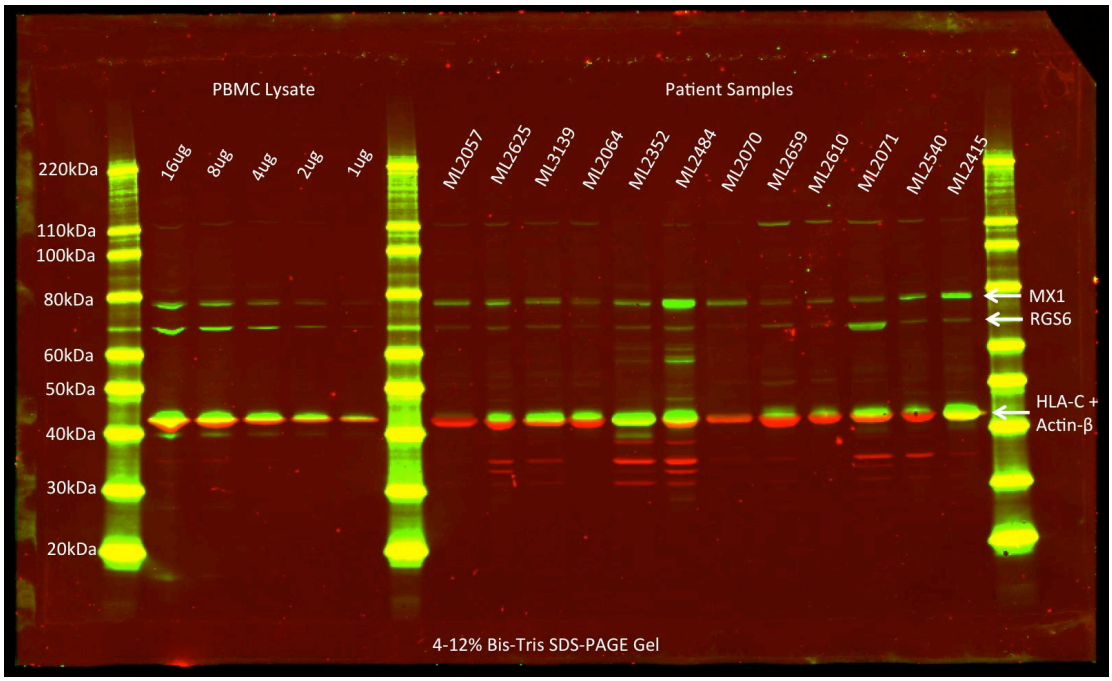


Figure 6.6. Western blot confirmation of MX1, RGS6 and HLA-C (Green) expression in PBMC from individual patient samples (HESN, NN, POS). All samples were normalized to Actin-β (Red).

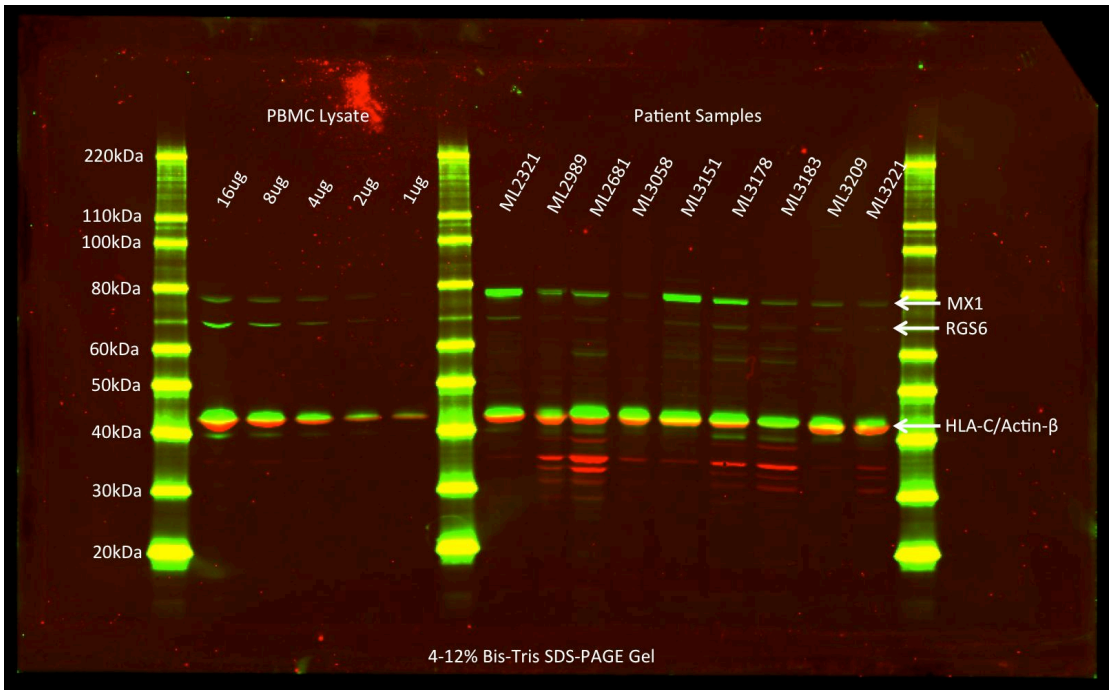
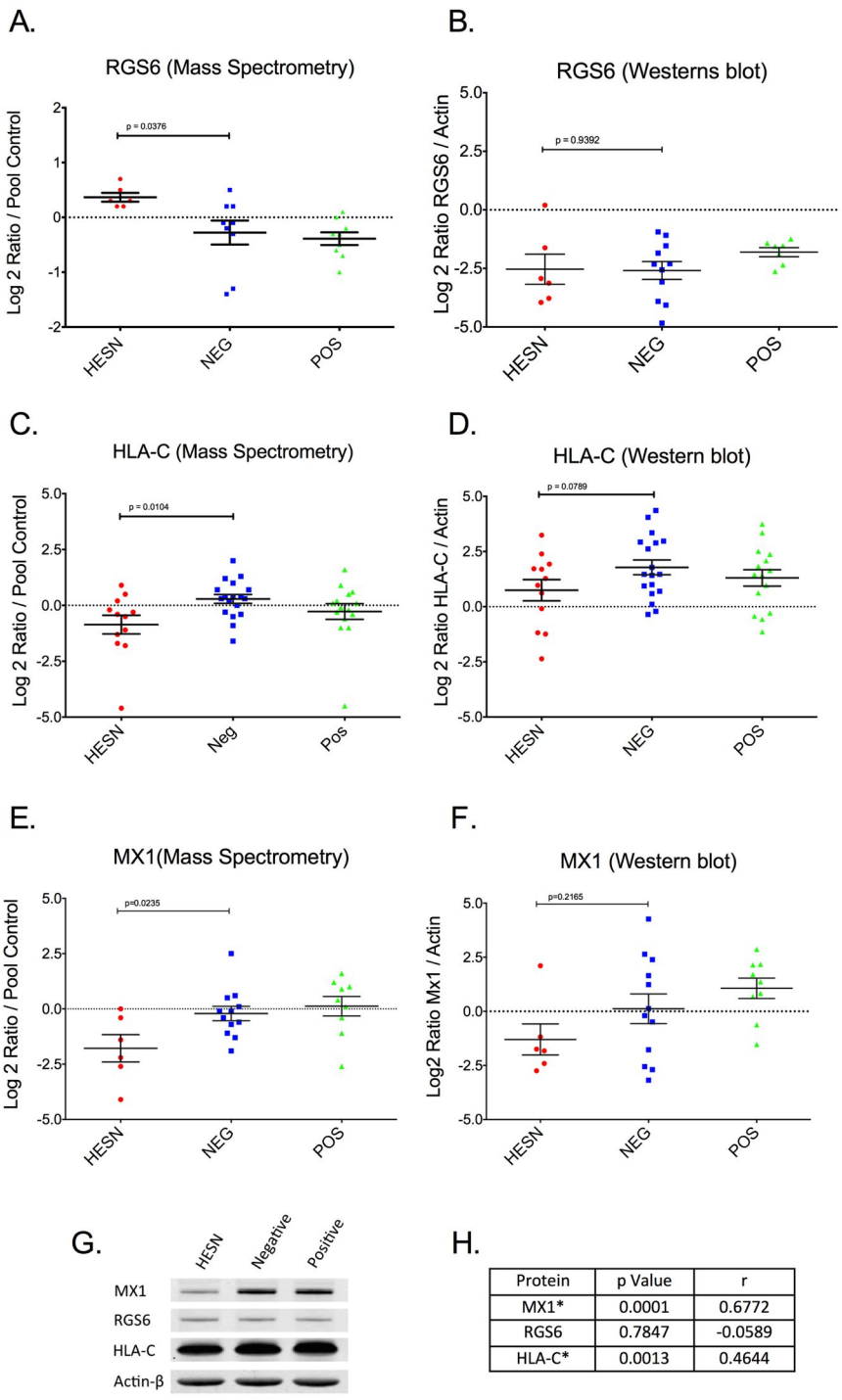


Figure 6.7. Western blot confirmation of MX1, RGS6 and HLA-C (Green) expression in PBMC from individual patient samples (HESN, NN, POS). All samples were normalized to Actin-β (Red).

PBMC lysate and subjected to Western blot analysis (Sections 2.2.5 + 2.2.6). A total of 4 gels were ran and are illustrated in Figure's 6.4-6.7. The band intensity of RGS6, HLA-C, and Mx1 were measured and normalized to the intensity of Actin-β and expressed as a ratio for comparative analysis. We compared the abundance of each protein measured by qMS expressed as log2 ratio/pool control and Western blot expressed as log2 ratio/Actin-β (Figure 6.8).

As can be observed in Figure 6.8A the qMS data indicates a significant overexpression of RGS6 in HESN women. However, when evaluating the Western blot data for RGS6 the differences between HESN and susceptible controls were no longer significant (Figure 6.8B). Additionally, there was no correlation between the expression values measured by qMS and the Western blot for RGS6 (Figure 6.8H). When we examined the expression of HLA-C by Western blot we found that similar to the qMS data (Figure 6.8C) the levels of HLA-C trended towards a decrease in HESN women however, the data was not statistically significant (Figure 6.8D; $p = 0.0789$). Additionally, the Western blot data correlated well with the original qMS data for HLA-C ($r = 0.4644$ $p = 0.003$). Similar to HLA-C, comparison of Mx1 expression using quantitative Western blot analysis trended towards a decrease in HESN women but was not deemed statistically significant (Figure 6.8F; $p = 0.2165$). However, correlation of the Mx1 expression measured by iTRAQ mass spectrometry with the Western blot expression (Figure 6.8H) showed that the two techniques significantly correlated ($r = 0.6772$ $p = 0.0001$).



Protein	p Value	r
MX1*	0.0001	0.6772
RGS6	0.7847	-0.0589
HLA-C*	0.0013	0.4644

Figure 6.8. Protein biomarker confirmation in HESN women. Mass spectrometry and Western blot expression levels are illustrated for (A/B) RGS6 (C/D) HLA-C and (E/F) Mx1. HIV+ individuals are plotted for comparative purposes only. (G) A representative Western blot visualizing the protein abundance in three individuals (H) Degree of correlation (Pearson) between the mass spectrometer and Western blot for MX1, RGS6, and HLA-C.

In general the protein expression levels measured by qMS correlated well with those measured by Western blot with the exception of RGS6. It is likely that expression of both HLA-C and Mx1 are decreased in the PBMC of HESN women. The correlation for HLA-C was not as strong as for Mx1 (r = 0.6772 vs. 0.4644) making the validation of this biomarker in a larger study more probable. Additionally, Mx1, an innate restriction factor, at the time of these studies had yet to be characterized in the context of HIV infection. These data allowed us to rationalize a much larger and more robust study to characterize Mx1 in the context of the HESN phenotype and HIV replication (Chapter 8: Epidemiology and function of Mx proteins in HESN women).

6.4.3 Ex-vivo proteome profile of PBMC from HESN sex workers.

While we successfully identified a candidate biomarker from the HESN cellular proteome, we also wanted to determine if the HESN phenotype could be predicted by the overall proteomic profile of PBMC from these women. To test this all 2913 proteins quantitative identified in HESN and HIV (negative) susceptible controls were subjected to unsupervised hierarchical cluster analysis described in Section 2.2.4.6. However, as can be observed in Figure 6.9, the cluster analysis failed to accurately assign whether an individual was classified as HESN or a susceptible control. However, distinct subject clusters with very different proteomic profiles could be observed. In addition, four individuals clearly had a drastically polarized proteomic profile compared to the rest of the individuals in the study. A closer look at the epidemiological and clinical data available for each of these individuals

indicated that ML2352, ML2040, ML3102, and ML3183 all had reported infections at the time the samples were collected (three had a cold, and one had Gonorrhea). Initially, these four individuals drastically skewed the statistical analysis of individual differentially regulated proteins between the two study groups. After removing these individuals from the analysis we were able to identify 136 significantly different protein biomarkers between HESN and susceptible controls (Section 6.4.1). To determine if only the significant proteins (136) were capable of differentiating between HESN and high-risk negative controls a second hierarchical cluster analysis was performed (Figure 6.10). Using only the significant proteins resulted in a clearer differentiation between HESN and susceptible controls. This clustering pattern was only achieved after we identified the four individuals with infections by the more comprehensive analysis using all the proteins quantified in the study. These findings provide strong evidence that this type of analysis approach is able to identify important immune responses that would otherwise be missed using a more biased approach.

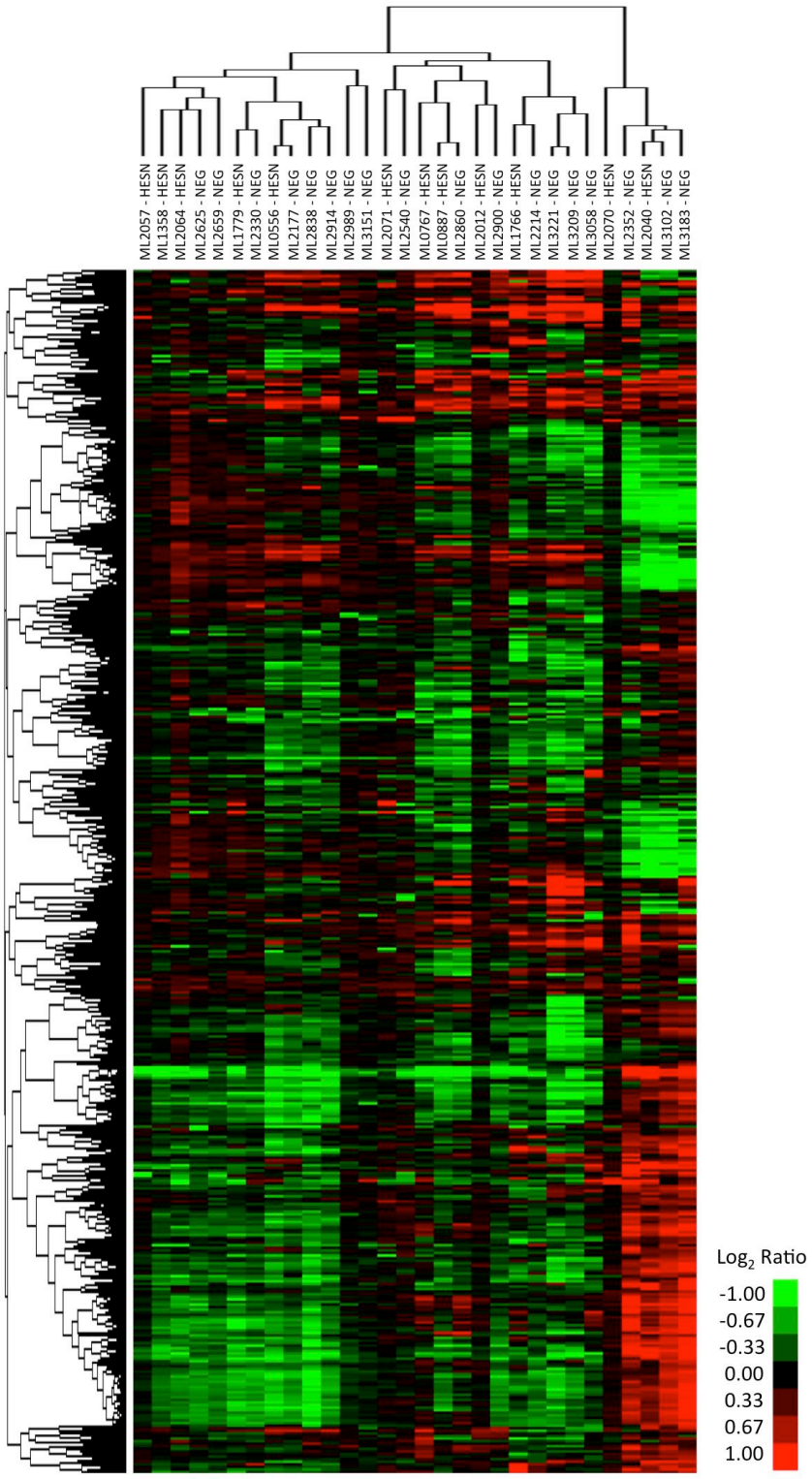


Figure 6.9. Hierarchical cluster analysis of all 2913 proteins identified and quantified in PBMC from HESN and susceptible controls. Dendrograms were generated using an un-centered correlation similarity metric and complete linkage clustering.

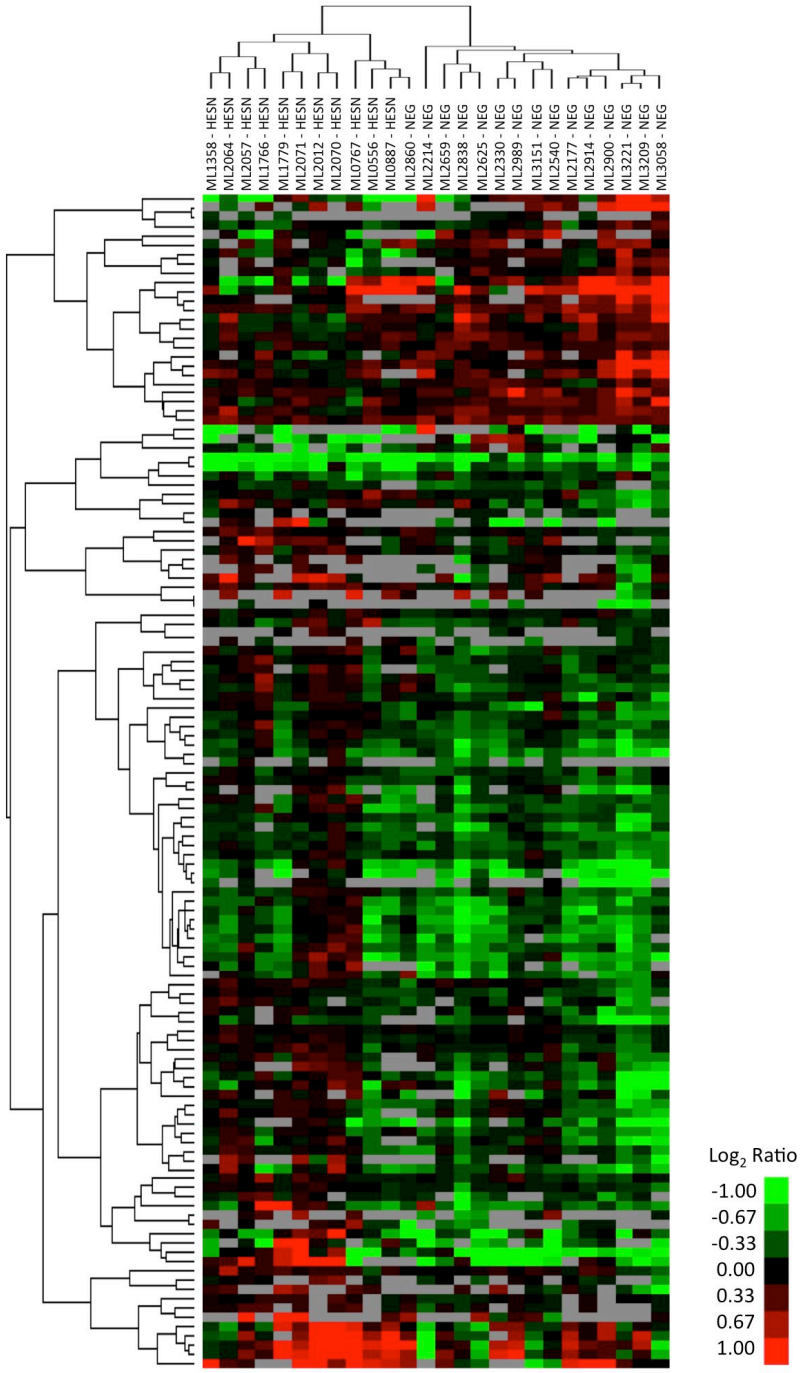


Figure 6.10. Hierarchical cluster analysis of the 136 significantly deregulated PBMC proteins in HESN and susceptible controls. Dendrograms were generated using an un-centered correlation similarity metric and complete linkage clustering.

6.4.4 Biological Functions in HESN sex workers

Based on previous descriptive studies at the systemic level in whole blood and isolated CD4+ T-cells from HESN women [123,124] we hypothesized that a similar immune quiescent profile would be seen at the proteomic level in PBMC. This profile would be reflected by down-regulation of proteins in immunological functions and pathways. To evaluate this hypothesis the expression ratios of all 2913 proteins that were quantified were subjected to IPA analysis described in the Section 2.2.4.6. Surprisingly IPA analysis identified significant up-regulation in most of the biological functions identified in HESN women. This included, hematological system development and function (p-value; 4.08E-36 – 1.06E-06), cellular function and maintenance (p-value; 1.01E-39 – 9.28E-07), cell death and survival (p-value; 2.08E-47 – 1.17E-06), cellular movement (p-value; 3.30E-25 – 1.06-06), inflammatory response (p-value; 4.08E-36 – 1.06E-06), cellular assembly and organization (p-value; 1.01E-39 – 9.28E-07), cell-to-cell signaling and interaction (p-value; 3.87E-19 – 1.06E-06), infectious disease (p-value; 4.12E-67 – 4.70E-07), immune cell trafficking (p-value; 1.19E-24 – 6.94E-07), and cell morphology (p-value; 2.43E-30 – 9.28E-07) which were all identified as functions significantly enriched in PBMC from HESN women (Figure 6.11). Additionally, specific sub-functions within these pathways were predicted to be significantly over-expressed with the exception of cell death and survival, which showed, mixed expression. However, sub-functions predicated to be under-expressed were associated with cell death while sub-functions shown to be over-expressed were associated with cell survival indicating a common phenotype. Sub-functions of particular interest found

to be overexpressed in HESN women included viral infection (Fig 6.11A; z = 5.257), HIV infection (Fig 6.11B; z = 5.867), lymphocyte migration (Fig 6.11C; z = 2.794), proliferation of lymphocytes (Fig 6.11D; z = 2.073), immune response of cells (Fig 6.11E; z = 2.134), and activation of leukocytes (Fig 6.11F; z = 3.068). This data suggests that HESN women may have increased immune activation at the systemic level and is not representative of the immune quiescent phenotype observed in our previous studies.

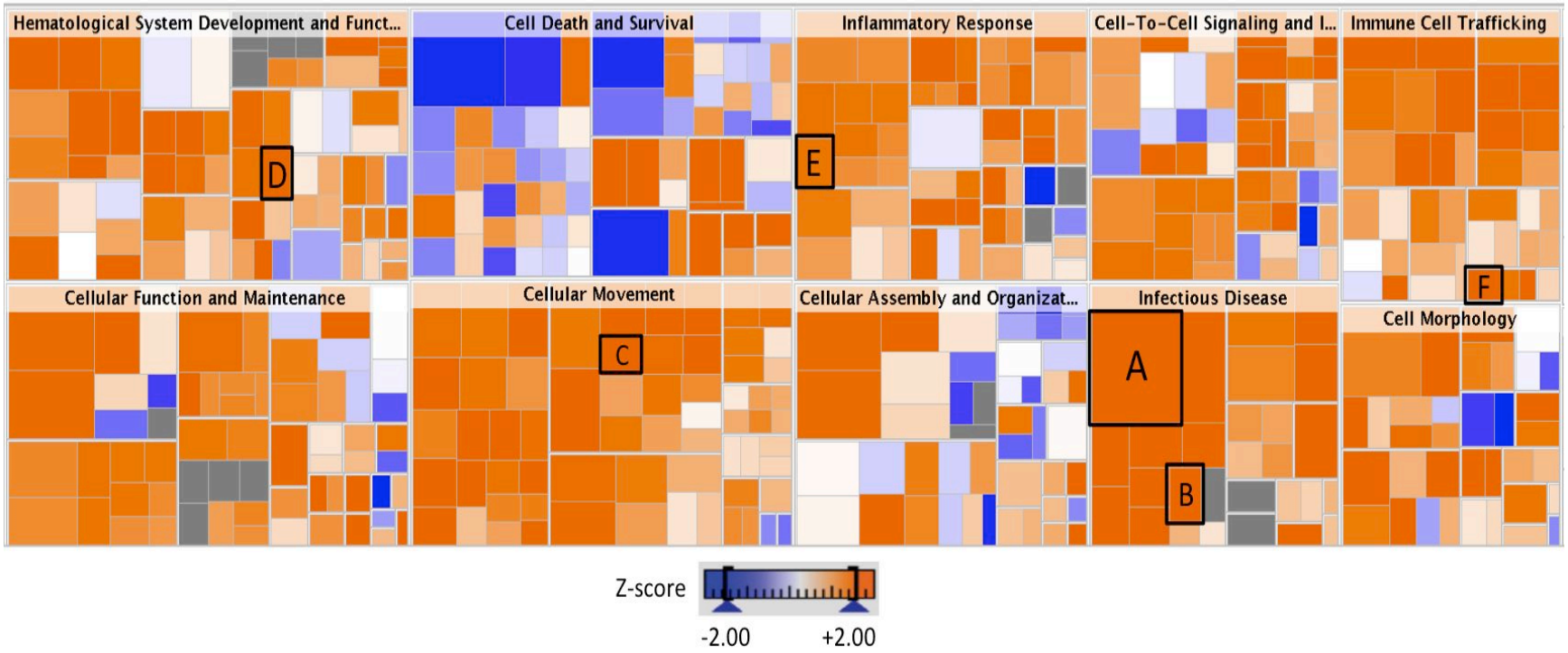


Figure 6.11. Functional analysis of proteins quantified in HESN women compared to high risk negative controls. Sub-functions of particular interest are highlighted A-F. This map was generated using IPA where the size of the squares is directly proportional to the p-value associated with each function. The degree of blue/orange indicates the under/over expression of the given function respectively (z-score).

6.5 Summary

In summary, we measured the immuno-proteomic profile of PBMC from HESN women compared to susceptible controls. We were also successful in identifying and validating Mx1 as an innate interferon restriction factor that may associate with the HESN phenotype for further study (Chapter 8). Cluster analysis initially did not allow us to distinguish between the two groups at the global proteomic level. However, it did allow us to identify individuals with ongoing infections that severely affected our clustering and statistical analysis. Once these confounders were removed from the analysis, clustering of the significantly different proteins allowed for segregation of the HESN and NEG patient groups. IPA functional analysis identified an increased immune activation profile in HESN women with increased expression by proteins associated with viral infection, proliferation, and activation compared to high-risk negative controls. In the following chapter we aim to identify if the systemic profile measure here is reflected at the genital tract level or if an immune quiescent profile can be observed in the same individuals.

Chapter 7. Comparative proteomic analysis of CMC from HESN women

7.1 Introduction & Rationale

In order to address the central hypothesis of this thesis, we wanted to characterize the proteomic profile of mucosal cervical mononuclear cells (CMC) in addition to PBMC from the systemic compartment of HESN women. The initial site of HIV-1 infection in women is the female genital tract and identifying an immune quiescent profile in CMC would lend further evidence to support the model natural immunity to HIV in HESN women. This chapter presents the results from the characterization of CMC from HESN women and compares the proteomic profiles of PBMC and CMC in HESN and high-risk negative controls.

7.2 Hypothesis

- HESN women will posses an immune quiescent proteomic profile when characterizing cervical mononuclear cells by shotgun proteomics.

7.3 Objective

- Utilize a qMS approach to determine the proteomic profile of ex-vivo cervical mononuclear cells (CMC) from HESN and susceptible controls and identify potential mucosal biomarkers of immune quiescence.

7.4 Results

Paired cervical mononuclear cells (CMC) were isolated at the same time from the majority of donors in the PBMC characterization study (Chapter 6) as described in Section 2.2.3.2. To examine the CMC proteome in a total of 13 HESN, and 17 high-

risk negative controls, CMC lysates were digested and labeled with iTRAQ (Sections 2.2.3.5). Of the 45 CMC samples included in this study, 28 of the samples were matched with samples characterized in the PBMC study (Chapter 6). As in the previous chapter a common pool for all the CMC samples was made and used as an internal control for each iTRAQ experiment with the addition of randomizing the patient groups with respect to iTRAQ reporter tags. Each 4-plex iTRAQ experiment was fractionated and peptides were identified and quantified by online-LC/MS/MS analysis using the Velos Orbitrap mass spectrometer (Section 2.2.4.5) As before, the raw mass spectrometer files were processed through the Mascot search algorithm and visualized using Scaffold Q+ as describe in Section 2.2.4.6 of the Materials and Methods.

An average of 2423 proteins and 39,062 peptides were identified from 15 iTRAQ experiments (Average: 194,522 spectra/run). When combining all of the data we were able to identify and quantify 4451 proteins across all the donors (Table 7.1). As in Chapter 6 the data was filtered to include proteins that had at least 3 data points for each experimental group (ie. HESN and NEG). This analysis yielded a total of 3972 proteins for further down stream analysis. As stated in Chapter 5 donors ML2352, ML2040, ML3102, and ML3183 were excluded due to on going infections at the time of sample collection.

Bio-Sample	#Prot	#IDs	#Spec	%IDs
Sample A	1300	20101	226753	8.86%
Sample B	2293	40805	188790	21.61%
Sample C	2333	30862	177548	17.38%
Sample D	2620	35323	195174	18.10%
Sample E	2781	51317	194704	26.36%
Sample F	2290	35404	175665	20.15%
Sample G	2105	37175	167829	22.15%
Sample H	2212	44447	194114	22.90%
Sample I	2362	29386	168316	17.46%
Sample J	2894	44060	209472	21.03%
Sample K	2542	41286	208058	19.84%
Sample L	2746	48292	209026	23.10%
Sample M	2433	38880	199428	19.50%
Sample N	2708	47765	207438	23.03%
Sample O	2736	40836	195527	20.89%

Table 7.1. Cumulative statistics from HESN CMC pilot study. Bio-Sample refers to the particular iTRAQ run (15 in total), while the number of proteins identified, peptides identified, total spectra (# of MS/MS scans), and the percentage of spectra (# of MS/MS scans identified as a peptide in the database) is also tabulated for each iTRAQ experiment.

7.4.1 Significantly deregulated proteins in CMC from HESN women

To identify novel biomarkers associated with HESN women a statistical comparison between HESN and high-risk negative controls for all 3972 proteins identified and quantified in this study was undertaken. Similar to the PBMC study the data was found to be normally distributed and a student's T test was used for each protein comparison. We chose a p-value cutoff of 0.05, which led to the identification of 193 differentially expressed proteins between HESN and high-risk negative controls. Additionally, a fold change cutoff of ±1.5 was applied as our criteria for identifying proteins of further interest (Figure 7.1). We found that there were a total of 74 under-expressed and 37 over-expressed proteins in HESN women compared to negative controls. A list of the top 10 deregulated proteins is listed in Table 7.2 with their biological significance and relation to HIV infection discussed in Section 9.4.1. Unfortunately, considerably less cells were isolated during CMC sample preparation, resulting in only enough protein lysate for mass spectrometry analysis, precluding the possibility of Western blot confirmation on the same samples. Additionally, CMC samples are not routinely banked for future studies / analysis making a large scale conformational study such as the one proposed in Chapter 8 impractical.

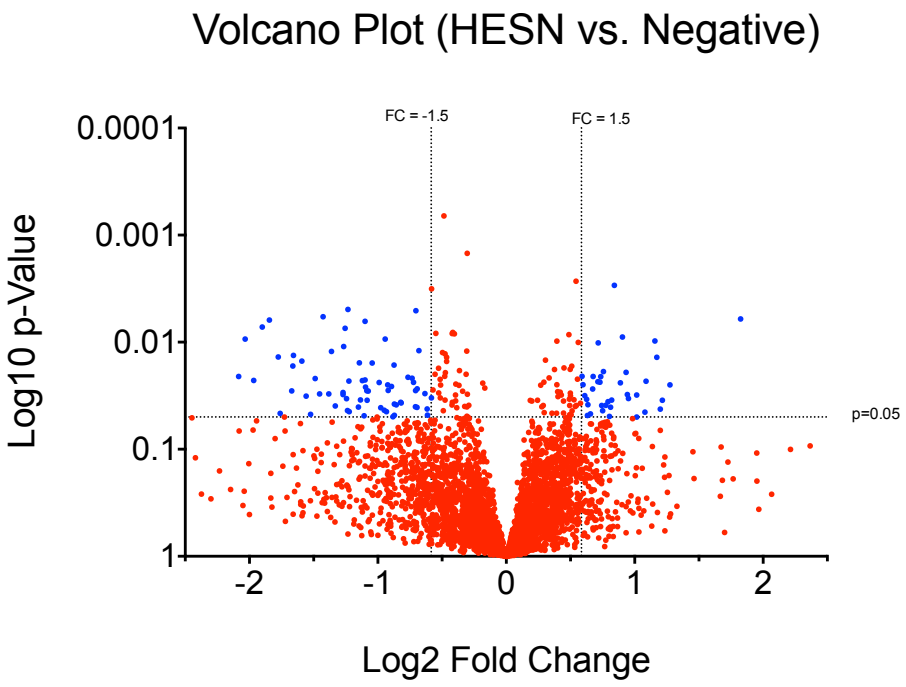


Figure 7.1 Immuno-proteomic profiling of CMC from HESN women identified several proteins of statistical and biological significance with a cutoff of $p<0.05$ and a FC>1.5

IPI Accession	Protein Name	Log_2 Fold HESN/NEG	p-value
IPI00384176	PBRM1, Isoform 1 of Protein Polybromo-1	-3.76	0.01
IPI00896452	PPP1R9A, Neurabin-1 Isoform 1	-3.41	0.02
IPI00299679	RGL1, Isoform B of Ral Guanine Nucleotide Dissociation Stimulator-Like 1	-2.98	0.04
IPI00028122	PSIP1, Isoform 1 of PC4 and SFRS1-Interacting Protein	-2.85	0.01
IPI00025879	MYH1, Myosin-1	-2.80	0.02
IPI00971069	PDZD7, Isoform 3 of PDZ Domain-Containing Protein 7	1.82	0.01
IPI00060414	CHMP4C, Charged Multivesicular Body Protein 4c	1.27	0.03
IPI00940125	TNFRSF21, Tumor Necrosis Factor Receptor Superfamily, Member 21 Variant	1.22	0.03
IPI00002824	CSRP2, Cysteine and Glycine-Rich Protein 2	1.20	0.04
IPI00305286	HPGD, 15-hydroxyprostaglandin dehydrogenase [NAD+]	1.17	0.01

*Log_2 ratio of each time point compared to the media control

Table 7.2. The top 10 significantly altered CMC proteins in HESN women compared to susceptible controls.

7.4.2 Ex-vivo proteome profile of CMC from HESN sex workers

To determine if the HESN phenotype could be predicted by the proteomic profile of CMC, all 3972 proteins for all HESN and HIV (negative) susceptible controls were subjected to unsupervised hierarchical cluster analysis described in Section 2.2.4.6. Unlike the scattered clustering observed when the PBMC proteome profile of HESN women was analyzed (Figure 6.9) there were two main groups of individuals clustering depending on protein expression (Figure 7.2). The largest cluster has a mix of negative and HESN individuals however most of the HESN individuals clustered to the left of the tree while the second smaller cluster consisted of only negative individuals, with the exception of ML2057 (a HESN individual). To determine if only the significant proteins (193) were capable of differentiating between HESN and high-risk negative controls, a second hierarchical cluster analysis was performed (Figure 7.3). Similar to the PBMC analysis in Chapter 6 (Figure 6.10), clustering based on significant proteins was generally able to differentiate between the two groups. While we observed similar clustering to Figure 7.2 where the majority of HESN were on the right of the dendrogram and the majority of high-risk negative controls were on the left, the clustering did not fall into two distinct groups. These results indicate that generally the global proteome profile we observed was able to define a HESN proteomic profile in CMC. Additionally, clustering based solely on the statistically significant proteins between HESN and susceptible controls did not further improve the differentiation between groups as was observed for the PBMC study.

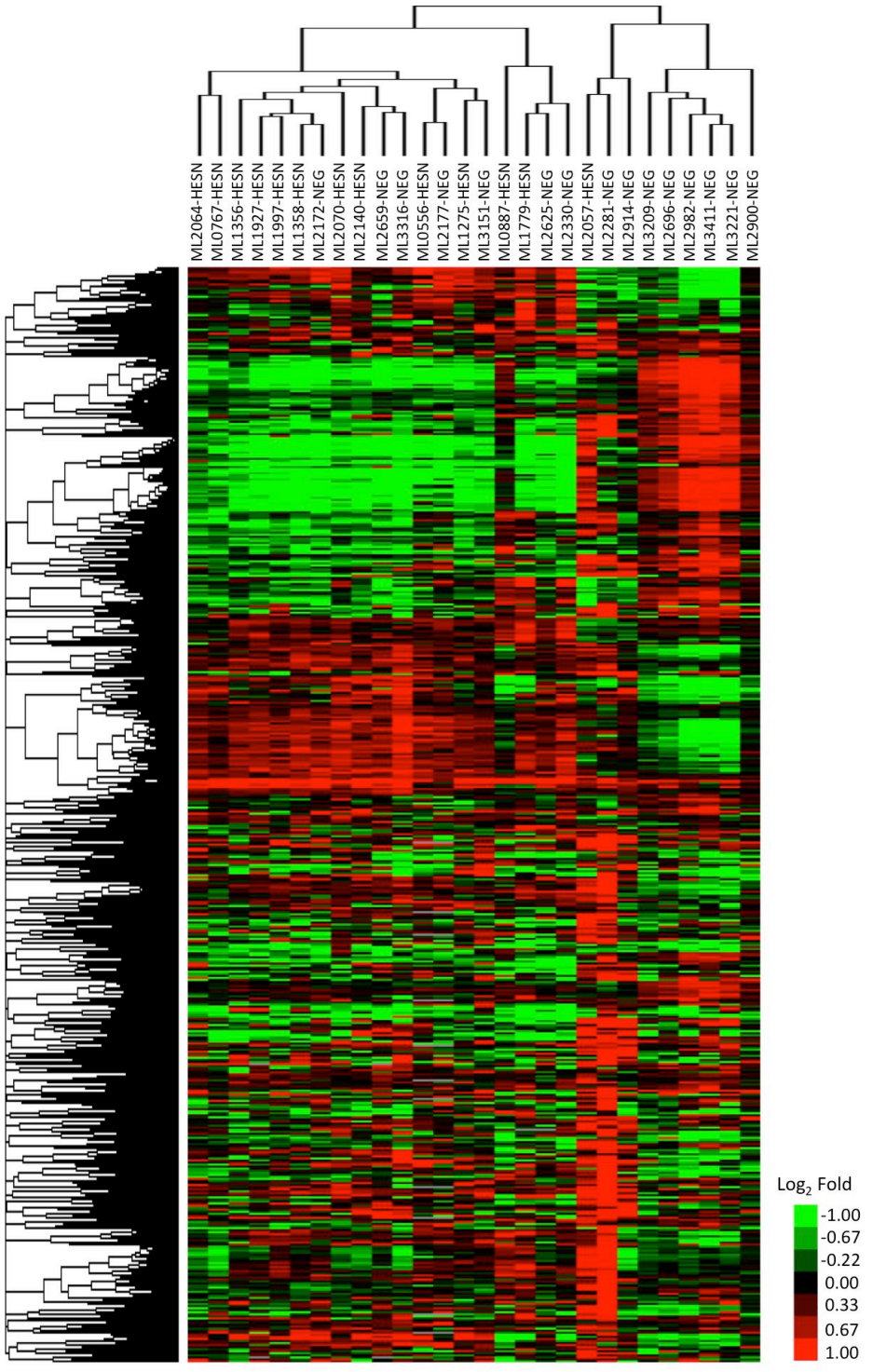


Figure 7.2. Hierarchical cluster analysis of all 3972 proteins identified and quantified in the CMC of HESN and susceptible controls. Dendrograms were generated using an un-centered correlation similarity metric and complete linkage clustering.

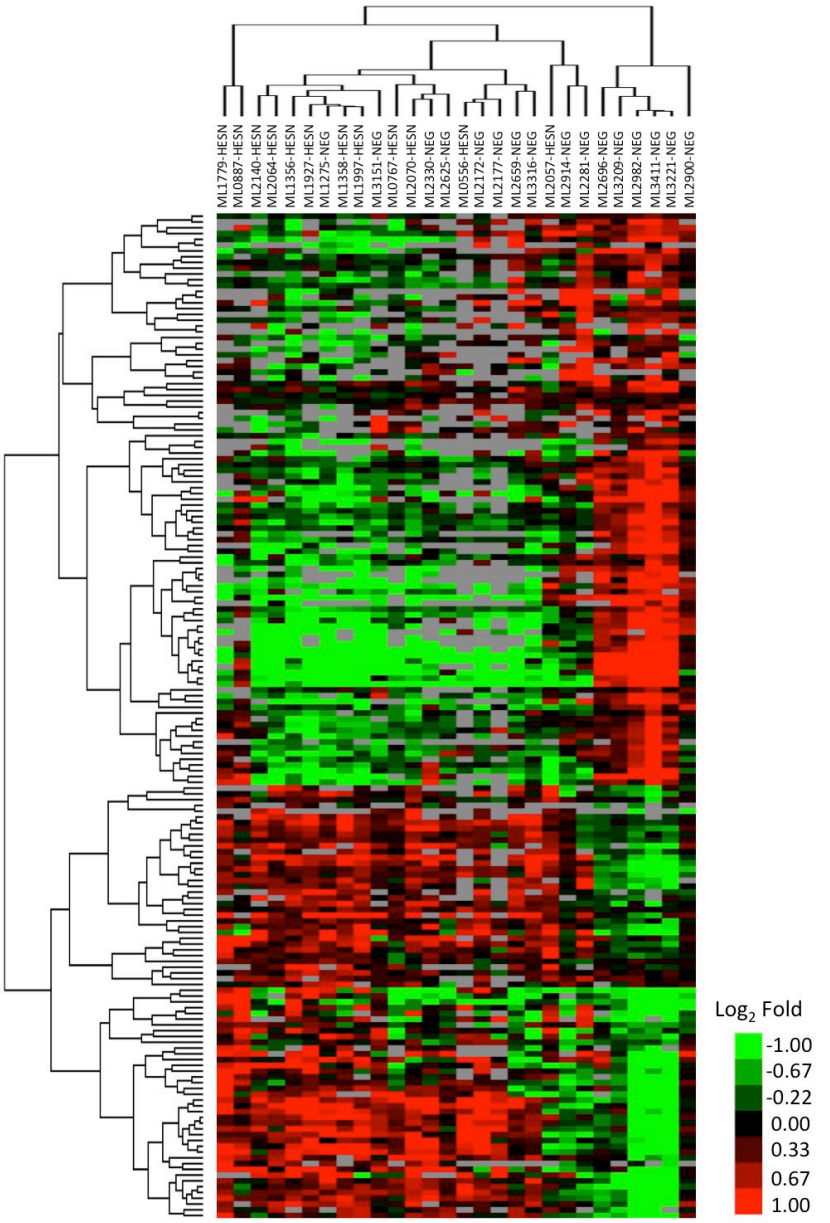


Figure 7.3. Hierarchical cluster analysis of the 193 significantly deregulated CMC proteins in HESN and susceptible controls. Dendrograms were generated using an un-centered correlation similarity metric and complete linkage clustering.

7.4.3 Biological functions in CMC of HESN sex workers

Based on previous descriptive studies of RNA expression at the systemic level in whole blood and CD4+ T-cells from HESN women [123,124] we hypothesized that a similar immune quiescent profile would be identified at the genital tract level in cervical mononuclear cells. We expected HESN women would have decreased expression of proteins important in immunological functions and pathways that would decrease the likelihood of HIV transmission occurring. To evaluate this hypothesis the expression ratios of all 3972 proteins quantified in the study were subjected to IPA analysis as described in Section 2.2.4.6. IPA analysis showed significant down-regulation in most of the biological functions identified in HESN women with the exception of cancer (p-value; 3.5E-26 – 6.5E-06), which was found to have increased expression. We found that proteins involved in cell death and survival (p-value; 8.42E-45 – 5.07E-06), cellular function and maintenance (p-value; 5.39E-37 – 6.75E-06), cellular movement (p-value; 1.84E-21 – 6.03E-06), cellular assembly and organization (p-value; 5.39E-37 – 6.75E-06), inflammatory response (p-value; 1.78E-19 – 5.03E-06), infectious disease (p-value; 1.88E-63 – 3.49E-06), hematological system development and function (p-value; 2.32E-16 – 3.63E-06), cell-to-cell signaling and interaction (p-value; 1.81E-13 – 6.75E-06), and immune cell trafficking (p-value; 6.55E-17 – 3.63E-06) were all identified as functions significantly underexpressed in CMC from HESN women compared to susceptible controls (Figure 7.3). Additionally, the majority of sub-functions within these pathways were also predicted to be significantly under-expressed. Sub-functions of particular interest, which were under-expressed in the CMC of HESN women,

included the cell death of epithelial cells (Fig 7.3A; z = -3.220), viral infection (Fig 7.3B; z = -1.647), HIV infection (Fig 7.3C; z = -0.740), immune response of cells (Fig 7.3D; z = -2.227), activation of cells (Fig 7.3E; z = -2.602), activation of leukocytes (Fig 7.3F; z = -2.974), homing of leukocytes (Fig 7.3G; z = -4.522), and leukocyte migration (Fig 7.3H; z = -3.098). Expression of functions associated with cancer and in particular cervical cancers are over-represented in the IPA database resulting in cancer being a significantly enriched function. The majority of the sub-functions in this category either had neutral expression (grey) or were not significant (z-score ≤ 2) with a trend in the positive direction. Unlike the functional analysis of PBMC, CMC clearly recapitulated the immune quiescent phenotype in immune cells from the female genital tract (FGT).

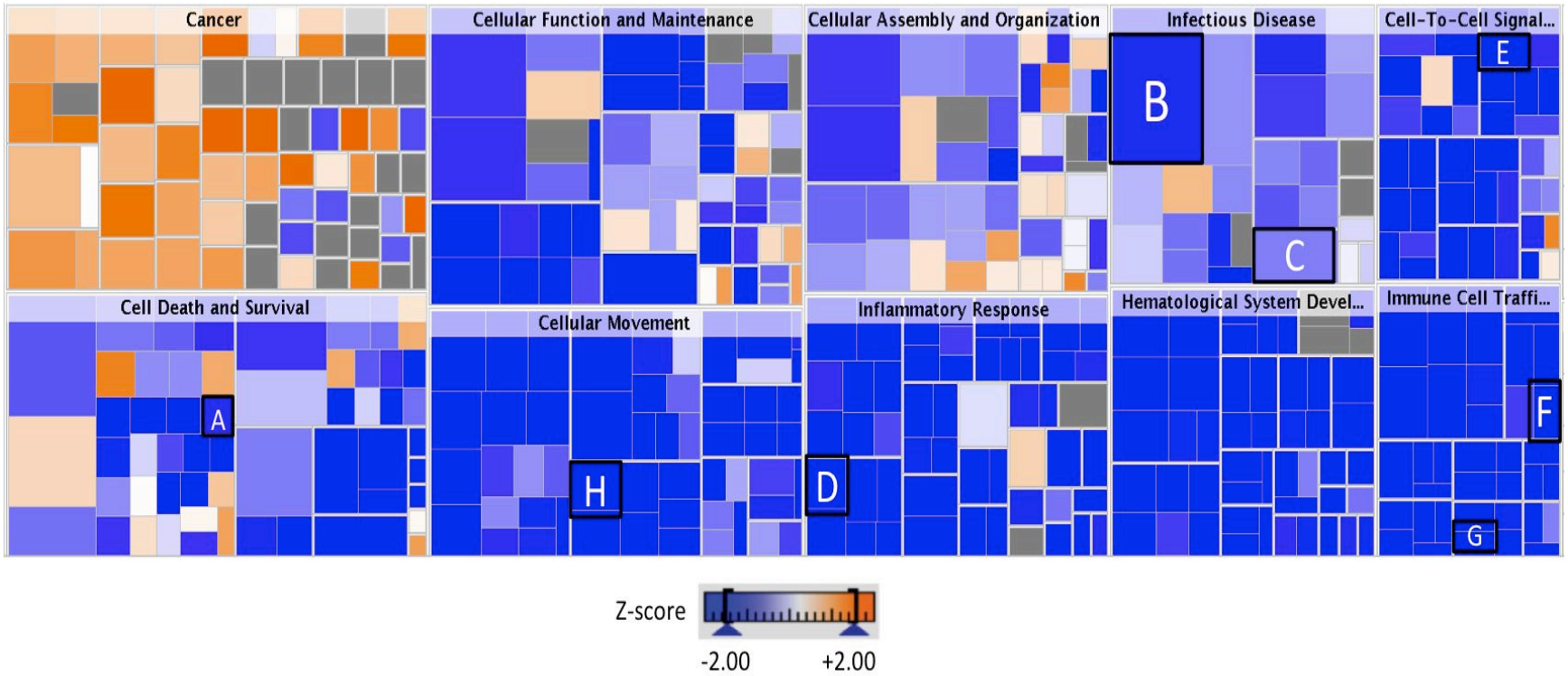


Figure 7.3. Functional analysis of proteins quantified in HESN women compared to high risk negative controls. Sub-functions of particular interest are highlighted A-H. This map was generated using IPA where the size of the squares is directly proportional to the p-value associated with each function. The degree of blue/orange indicates the under / over expression of the given function respectively (z-score).

7.4.4 Comparative Proteomic Signatures in PBMC and CMC from HESN women

Both the PBMC (Figure 6.11) and the CMC (Figure 7.3) functional profiles of HESN women illustrated extremely polarized expression. A comparative cluster analysis between PBMC and CMC from HESN women was undertaken to examine commonalities and differences in the proteomes from the two anatomical sites. The analysis included a total of 54 samples of which 28 were matched PBMC and CMC samples from the same study participants. Hierarchical cluster analysis of the 2080 proteins that were quantitatively identified and found in common between PBMC and CMC was conducted (Figure 7.4). As expected the PBMC and CMC proteomic profiles clearly predict if the lymphocyte sample was obtained from the genital tract or the systemic compartment indicating that similar immune cells from different compartments have very different immuno-proteomic profiles. Again, as seen in Sections 6.4.3 and 7.4.2 we were unable to perfectly differentiate HESN from susceptible controls with respect to PBMC or CMC. However, some clustering with HESN individuals to the right and susceptible controls to the far left of Figure 7.4 was observed in lymphocytes isolated from the genital tract. This data suggests that the FGT may be more important when studying immune quiescence and the HESN phenotype.

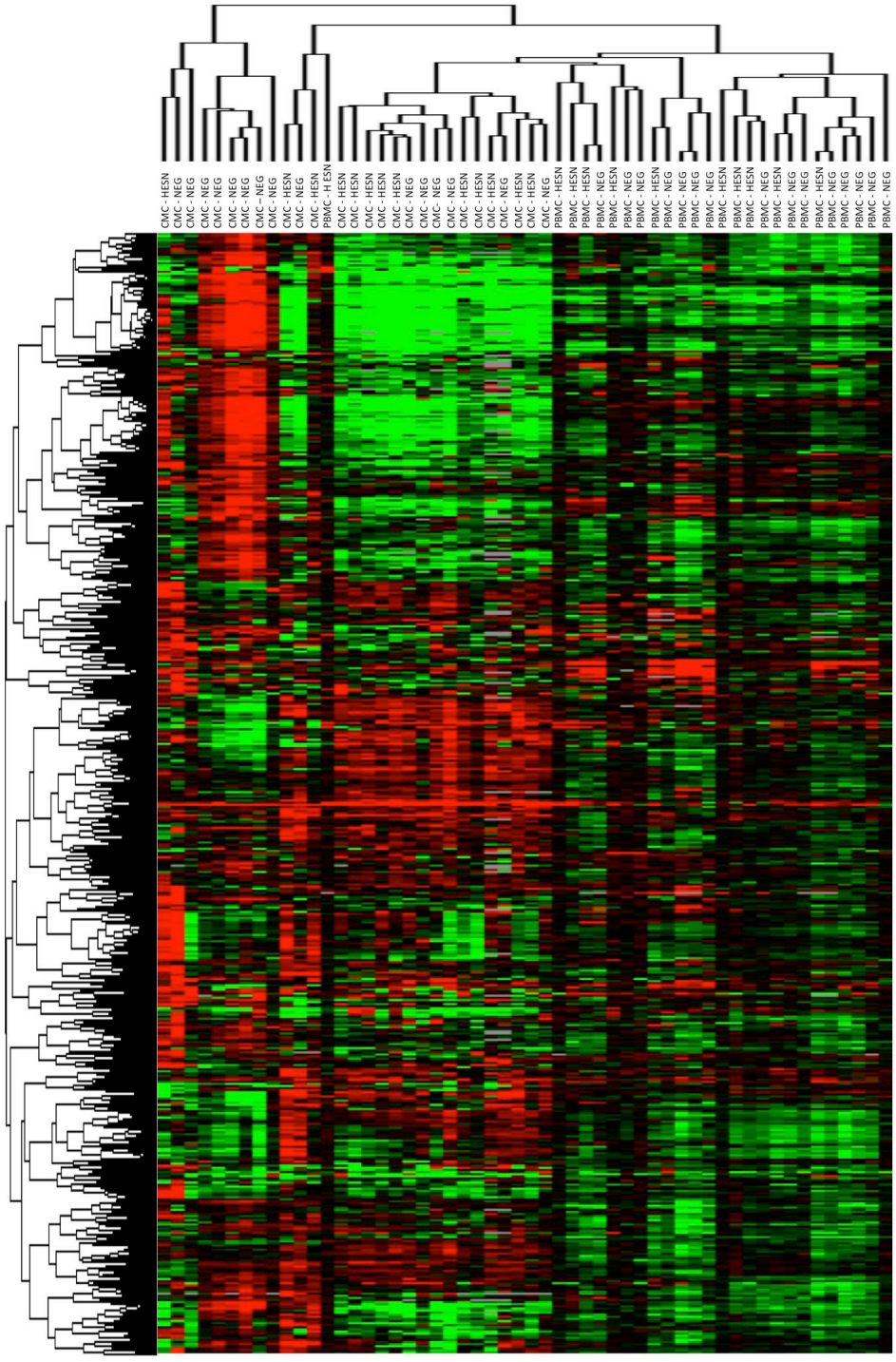


Figure 7.4. Dendrogram comparing the proteome profile of CMC and PBMC from HESN and susceptible controls.

7.5 Summary

In summary, we were successful in examining the proteomic profile of CMC from HESN and high-risk negative controls. We identified 37 over-expressed and 74 under-expressed proteins associated with the HESN phenotype as potential markers of HIV resistance. Functional analysis of proteins identified in PBMC from HESN women did not reflect a immune quiescent immune profile (Chapter 6) at the proteomic level however, CMC isolated from the same individuals did confirm significant down-regulation in biological functions such as viral replication, inflammation, and leukocyte activation / migration consistent with our immune quiescence hypothesis. The functional profile of HESN women is highly polarized when comparing the systemic and genital tract environments. Cluster analysis trended towards the identification of a proteomic signature that defines the HESN phenotype in CMC, however it was not conclusive. A comparative cluster analysis between PBMC and CMC clearly illustrated distinct proteomic signatures due to the anatomical site of immune cell isolation. These data suggest that the proteomic profile of immune cells isolated from the genital tract of HESN women may be more reflective of immune quiescence. Recapitulation of immune quiescence in the genital tract, where HIV acquisition most often occurs, lends strong evidence for our model of natural resistance to HIV infection in HESN women.

Chapter 8. Epidemiology and function of Mx proteins in HESN women.

8.1 Introduction & Rationale

It was shown in Chapter 6 that Mx1 appeared to be under-expressed in PBMC from HESN women compared to susceptible HIV uninfected controls. The expression of Mx1 was independently validated by Western blot and correlated strongly with the qMS analysis. For this reason we decided to undertake a large-scale population level analysis of Mx1 expression in PBMC from over 300 women in the Pumwani sex worker cohort. Mx1 has been well characterized with respect to Influenza infection and is recognized as an important innate factor in restricting the replication of DNA and RNA viruses [179]. However, a recent report suggested that Mx1 expression is capable of enhancing HSV-1 replication making its role in anti-viral immune responses unclear [180]. If true, and Mx1 plays multifunctional roles in the replication of different viruses, the expression of Mx1 observed in HESN women could be biologically plausible. Until recently, there have been no reports characterizing the effect Mx1 could have with respect to HIV replication. Three new studies show that Mx2 and not Mx1 restrict HIV-1 infection in-vitro [120,121,181]. However, due to the unclear role Mx1 plays as a pro or anti-HIV restriction factor an experiment was designed to evaluate the contribution Mx1 may have during in-vitro HIV infection. We also concluded that an evaluation of both Mx proteins at a population level in the Pumwani cohort was also warranted due to the lack of literature available for Mx1/Mx2 expression in large-scale human studies.

8.2 Hypotheses

1. HESN women will have lower expression of Mx1 in PBMC at a population level and over-expression of Mx1 will enhance HIV infection in-vitro.

2. The newly discovered HIV restriction factor, Mx2, will be over-expressed in the PBMC of HESN women compared to high-risk negative controls

8.3 Objectives

1. Demonstrate the ability of Mx1 to enhance HIV infection in an in-vitro model

2. Measure Mx1 levels at a population level in PBMC of HESN women and susceptible controls to investigate epidemiological confounders that may affect Mx1 expression in PBMC.

3. Measure Mx2 levels at a population level in PBMC of HESN women and susceptible controls to investigate epidemiological confounders that may affect Mx2 expression in PBMC

8.4 Results

8.4.1 Overexpression of Mx1 in HIV infection in-vitro

Based on our hypothesis that Mx1 expression was significantly decreased in HESN women we designed an experiment to determine if increased Mx1 expression using a stable cell line would enhance HIV replication. SUPT1 cell lines were engineered to overexpress Mx1, as well as 3 controls; including cells expressing GFP, the

expression vector only (no insert), and cells only, as described in the Sections 2.2.7. Western blot analysis of the engineered cells lines confirmed the proper expression levels of Mx1 (Figure 8.1D). The engineered SUPT1 cell lines were then infected with HIV_{IIIB} at various multiplicities of infection (MOI 1, 0.5, 0.25) while culture supernatants were harvested on Day 5 and viral replication was assessed by p24 ELISA (Sections 2.2.8.3 and 2.2.9). The over-expression of Mx1 resulted in no statistically significant differences when compared to the irrelevant protein control (GFP) during in-vitro HIV infection (Figure 8.1). Our data is in agreement with three recently published studies concluding that Mx1 has no effect on HIV replication [119-121]. Thus, while Mx1 may be decreased in HESN women our data and others demonstrate that overexpression of Mx1, does not affect HIV replication.

8.4.2 Population level analysis of Mx1 and Mx2

Our proteomics studies demonstrated that there is considerable individual variation in the cellular proteome between individuals and it is well described that epidemiological factors can play an influential role in affecting the immune response and in particular, HIV susceptibility. To examine Mx1 and Mx2 in a large scale study to control for individual variation and epidemiological factors we measured the basal expression of these proteins in ex-vivo PBMC from over 300 individuals in the Pumwami cohort. This analysis allowed us to control for and evaluate epidemiological factors such as age, contraception, menopause, pregnancy, infection

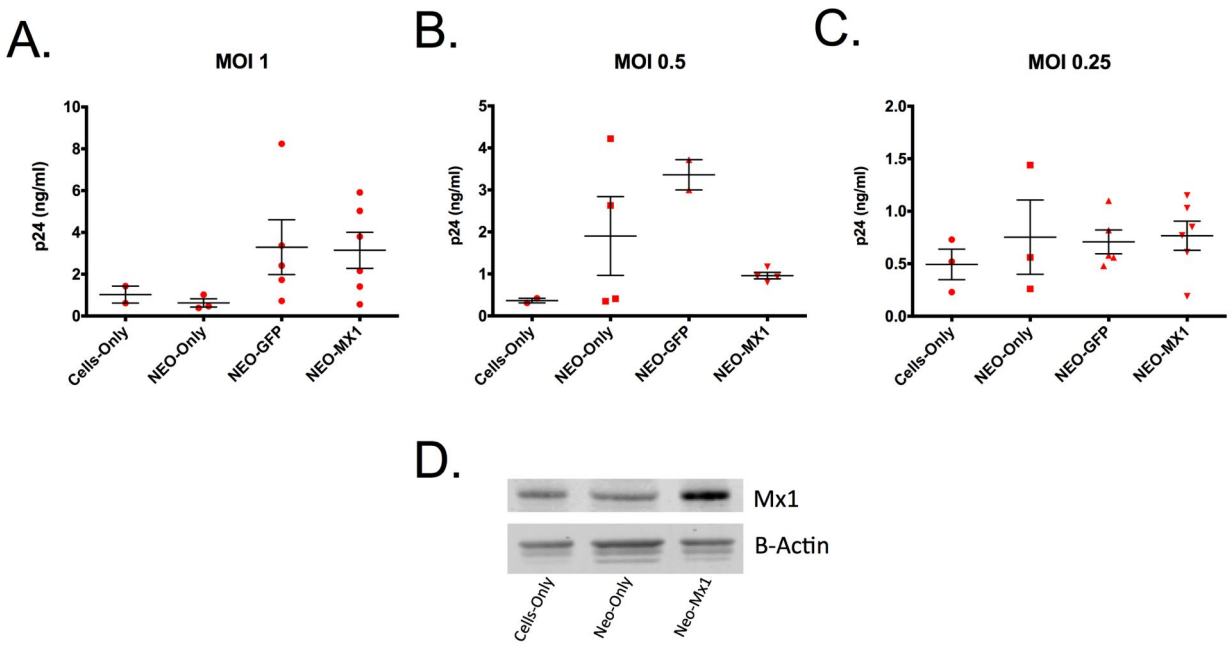


Figure 8.1. SUPT1 cell lines expressing vector only (Neo-Only), GFP vector (Neo-GFP), Mx1 (NEO-Mx1), and no vector (Cells-Only) were infected with HIV_{IIIB} (CXCR4) MOI of 1 (A), 0.5 (B), 0.25 (C). Cell culture supernatants were harvested on Day 5 of infection and HIV p24 was assayed by in an in house ELISA. The infections were performed on 6 separate occasions however some experiments were below the limit of detection for the ELISA explaining different numbers of data points for the various cell lines and MOI's. Mx1 overexpression was confirmed by Western blot analysis (D).

status, and medication usage and how they may affect Mx1 as well as Mx2 in the context of HESN women and HIV infection.

In total PBMC from 103 HESN, 100 susceptible controls and 104 HIV-infected subjects were available for analysis. All PBMC samples were obtained from the University of Manitoba / University of Nairobi Pumwani cohort biorepository. The samples chosen spanned a dates from 1995 to 2009. All PBMC samples were thawed, counted, lysed, and assessed for total protein content as outlined in Section 2.2.10.1. Only PBMC samples with a viability of 80% or higher as assessed by Trypan blue exclusion were included in the study. A commercial ELISA was then used to assess the expression of Mx1 and Mx2 in PBMC lysate from all study participants normalized to total cellular protein content (Section 2.2.10.2).

We compared the expression of Mx1 and Mx2 in the PBMC of HESN and negative controls (Figure 8.2). Using a non-parametric Mann-Whitney T test to compare the levels of each protein between the two study groups we found that HESN women had significantly higher levels of Mx2 expression in PBMC compared to high-risk negative controls ($p=0.0193$), contrary to our proteomics data (Figure 6.2). HESN women had similar levels of Mx1 expression compared to susceptible controls and were not statistically significant ($p=0.1299$) between the two groups.

To asses the effects age may have on Mx1 and Mx2 expression the study participants were further broken down into 3 separate age categories, as age is known to play a significant role in immune function. Figure 8.3 illustrates that there were no

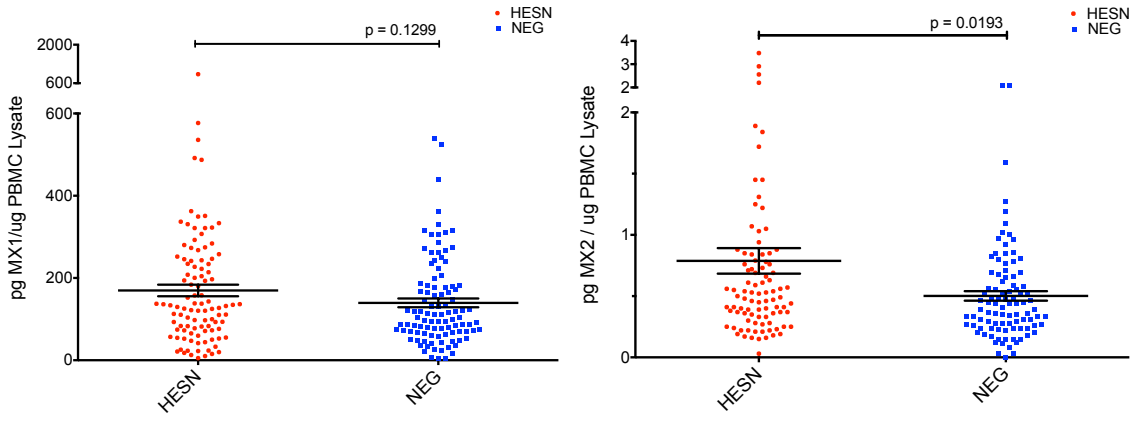


Figure 8.2 Quantitative levels of the (A) Mx1 and (B) Mx2 proteins in peripheral blood mononuclear cells isolated from HESN and HIV susceptible women. Error bars represent the standard error of the mean while a standard non-parametric Mann-Whitney T test was used for pair wise comparisons. * $p<0.05$

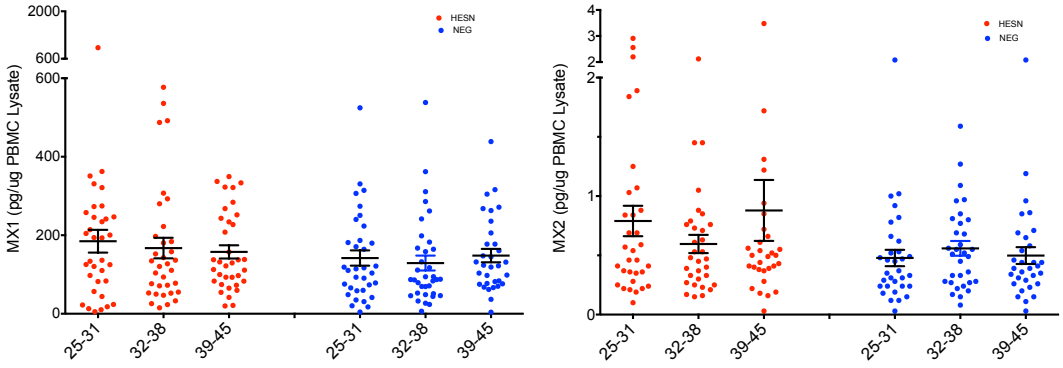


Figure 8.3. Quantitative levels of the (A) Mx1 and (B) Mx2 proteins in peripheral blood mononuclear cells separated by age. Error bars represent the standard error of the mean while a standard non-parametric Mann-Whitney T test was used for pair wise comparisons.

significant age related differences between and within groups in the expression of Mx1 and Mx2 in PBMC. Additionally, we were particularly interested in the effect contraception may have on Mx1 and Mx2, as hormonal contraception has long been recognized to have potent effects on immune regulation and has been associated with increased risk of HIV infection [182-184]. A comparative analysis between different methods of contraception including: no contraception, oral, Depo-Provera, and condoms was undertaken with regards to the levels of Mx1 and Mx2 in PBMC lysate (Figure 8.4A). Interestingly, we observed significant differences in the levels of Mx1 between women who used Depo-Provera and condoms as contraception ($p=0.0494$) as well as, women who did not use any form of contraception and those who used only condoms ($p=0.0486$). Identical comparisons were undertaken for the levels of Mx2 (Figure 8.4B) with regards to contraception use. Here we identified significant differences in expression between women who used Depo-Provera or a condom as contraception ($p=0.0243$). These data suggest that hormonal contraception use may play a role in Mx1/Mx2 expression.

To better define Mx1/Mx2 expression in systemic PBMC's we conducted bivariate and multiple linear regression (Section 2.2.10.3) analysis in an effort to identify and account for epidemiological confounders. Several patient characteristics including: HESN status, Depo-Provera, age, IRF-1 genotype, gravida, sex work duration, sex work break, recent infections, and medication usage, that we felt could have an immunological effect on Mx1/Mx2 were evaluated using a bivariate analysis. Table 8.1 indicates that none of these factors were found to be significantly associated

with Mx1 expression levels in PBMC with the exception of gravida (p = 0.001). This indicates that women who have never been pregnant tended to have higher expression of Mx1 compared to women who have had at least one or more pregnancies. We next assessed these factors using the same bivariate analysis with regards to Mx2 expression (Table 8.2). Unlike for Mx1, Mx2 expression was significantly higher in HESN women compared to negative controls (p = 0.027). Additionally, women who used Depo-Provera also had significantly higher expression of Mx2 (p = 0.005). Thus it appears that Mx2, but not Mx1 is associated with the HESN phenotype and Depo-Provera use.

To more precisely correct and adjust for the epidemiological confounders described in Tables 8.1 and 8.2 we conducted a secondary multiple linear regression analysis. This analysis allows us to add a single variable into the model in succession to evaluate how it affects either the Mx1 or Mx2 expression in HESN women. Four statistical models were constructed which included: Model 1 representing the addition of Depo-Provera, model 2; Depo-Provera and age, model 3; Depo-Provera, age and medication history and model 4; Depo-Provera use, age, medication history, and the HESN/Depo-Provera interaction. Table 8.3 illustrates that under all circumstances, Mx1 expression was not significant between HESN and susceptible controls. Table 8.4 indicates that with the addition of each variable into the model the increased expression of Mx2 observed in HESN women remained significant (Model 3; p = 0.04). However, when adding the interaction between HESN and Depo-Provera to the model the association was no longer significant (Model 4; p =

0.39), indicating that the interaction alone between HESN women and Depo-Provera usage was most likely significant (Model 4; p = 0.03). In order to explain this observation, the interaction model between Depo-Provera use and HESN status was further tested (Table 8.4, Model 4), and it was found that the HESN women who used Depo-Provera were largely responsible for the higher levels of Mx2 observed in this study (Figure 8.5). Compared to susceptible controls not using Depo-Provera, log Mx2 levels for HESN women also not using Depo-Provera were only 0.116 pg/ug PBMC lysate higher (p = 0.388, not shown). Only the log Mx2 expression levels for HESN women reporting Depo-Provera use was consistently higher at 0.781 pg/ug PBMC lysate. These data suggest that Mx2 is significantly overexpressed in the PBMC of HESN women and is associated with the use of Depo-Provera.

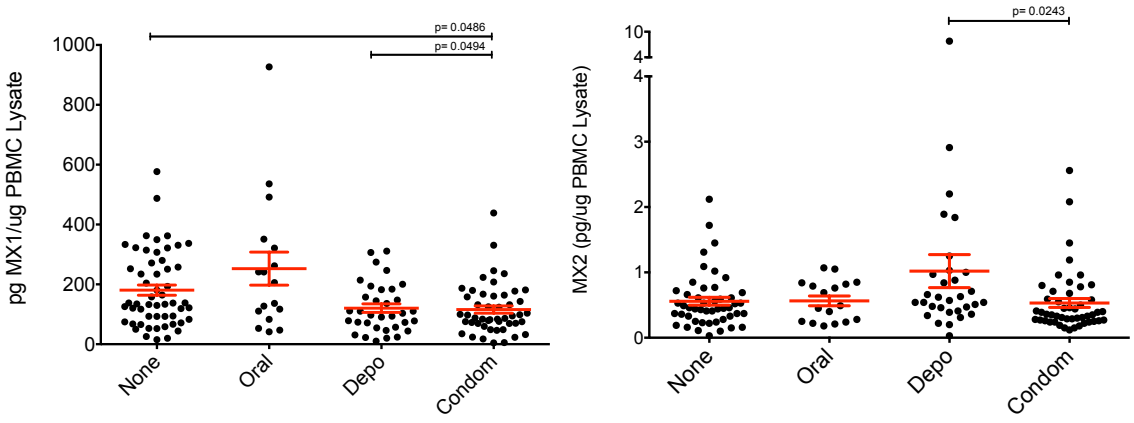


Figure 8.4. Expression of (A) Mx1 and (B) Mx2 in peripheral blood mononuclear cells isolated from HESN and HIV susceptible women separated by contraception usage. Error bars represent the standard error of the mean while a standard non-parametric Kruskal-Whallis test was used for multiple comparisons. * $p<0.05$

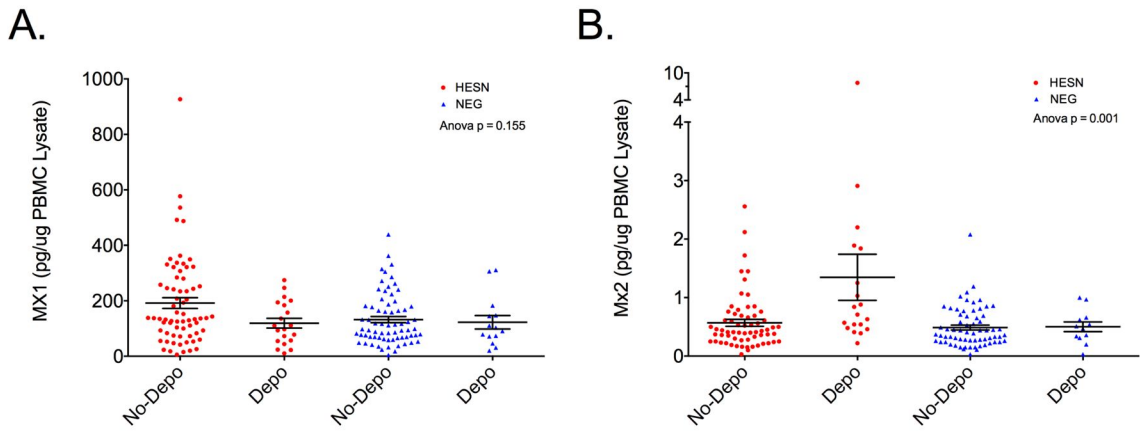


Figure 8.5. Interaction model describing the relationship between Depo-Provera use and HESN/NEG status for (A) Mx1 and (B) Mx2. The analysis excludes individuals who were menopausal or pregnant at the time of sampling.

	No.	Mean	SD	Median	IQR	p
TOTAL	161	4.7	0.9	4.8	4.2 - 5.3	
HESN vs. NEG (ref)						0.169
NEG	78	4.59	0.79	4.54	4.20 - 5.18	
HESN	83	4.78	0.96	4.88	4.28 - 5.49	
Depo-provera						0.132
Non-depo-user	130	4.74	0.89	4.82	4.24 - 5.46	
Depo-user	31	4.47	0.86	4.54	4.00 - 5.21	
Age						0.164
<34	81	4.56	1.01	4.70	3.97 - 5.30	
35-39	35	4.76	0.83	4.85	4.41 - 5.20	
40+	45	4.85	0.63	4.82	4.35 - 5.46	
IRF-1						0.233
Non-protective	48	4.63	0.84	4.74	4.14 - 5.33	
Protective	11	4.97	0.94	4.83	4.25 - 5.86	
Gravida						0.001
0	88	4.91	0.85	4.92	4.47 - 5.54	
1-2	30	4.46	0.79	4.52	4.09 - 5.06	
3+	43	4.37	0.89	4.33	3.97 - 4.98	
Sex work duration						0.568
2 or fewer years	39	4.61	0.88	4.55	4.19 - 5.41	
2-5 years	39	4.81	0.74	4.83	4.33 - 5.29	
6+ years	83	4.66	0.95	4.79	4.09 - 5.46	
Sex work break						0.144
No	104	4.76	0.82	4.81	4.32 - 5.33	
Yes	57	4.55	0.98	4.48	4.00 - 5.32	
Had infection, last 2 months						0.243
No	135	4.72	0.90	4.80	4.25 - 5.41	
Yes	26	4.50	0.77	4.42	3.89 - 5.08	
Had medication, last 6 months						0.255
No	135	4.72	0.89	4.83	4.24 - 5.41	
Yes	26	4.50	0.83	4.48	4.13 - 5.12	

Table 8.1. Patient characteristics are tabulated for the HESN and HIV susceptible women included in the study of Mx1 expression in PBMC. The data excludes individuals who were menopausal or pregnant at the time of sampling.

	No.	Mean	SD	Median	IQR	p
TOTAL	161	-0.8	0.8	-0.8	-1.3 - -0.3	
HESN vs. NEG (ref)						0.027
NEG	78	-0.95	0.75	-0.92	-1.34 - -0.38	
HESN	83	-0.67	0.82	-0.72	-1.12 - -0.24	
Depo-provera						0.005
Non-depo-user	130	-0.89	0.73	-0.89	-1.36 - -0.37	
Depo-user	31	-0.44	0.96	-0.57	-0.88 - -0.00	
Age						0.098
<34	81	-0.72	0.79	-0.77	-1.26 - -0.17	
35-39	35	-0.73	0.81	-0.67	-1.40 - -0.37	
40+	45	-1.02	0.78	-0.90	-1.27 - -0.62	
IRF-1						0.204
Non-protective	48	-0.72	0.85	-0.80	-1.38 - -0.19	
Protective	11	-0.38	0.43	-0.37	-0.82 - -0.04	
Gravida						0.045
0	88	-0.66	0.81	-0.70	-1.04 - -0.17	
1-2	30	-1.01	0.73	-1.03	-1.42 - -0.37	
3+	43	-0.95	0.78	-1.01	-1.34 - -0.48	
Sex work duration						0.194
2 or fewer years	39	-0.99	0.77	-1.06	-1.45 - -0.57	
2-5 years	39	-0.81	0.84	-0.79	-1.36 - -0.23	
6+ years	83	-0.71	0.78	-0.78	-1.12 - -0.28	
Sex work break						0.299
No	104	-0.76	0.81	-0.80	-1.30 - -0.21	
Yes	57	-0.89	0.78	-0.88	-1.24 - -0.54	
Had infection, last 2 months						0.387
No	135	-0.78	0.80	-0.79	-1.32 - -0.26	
Yes	26	-0.93	0.78	-0.98	-1.23 - -0.42	
Had medication, last 6 months						0.473
No	135	-0.78	0.84	-0.79	-1.32 - -0.26	
Yes	26	-0.91	0.50	-0.90	-1.26 - -0.60	

Table 8.2. Patient characteristics are tabulated for the HESN and HIV susceptible women included in the study of Mx2 expression in PBMC. The data excludes individuals who were menopausal or pregnant at the time of sampling.

	Crude	Model 1	Model 2	Model 3	Model 4
HESN vs. NN (ref)	0.192	0.214	0.210	0.209	0.241
	(0.17)	(0.13)	(0.13)	(0.13)	(0.12)
Depo-provera		-0.292	-0.223	-0.241	-0.142
		(0.10)	(0.22)	(0.19)	(0.61)
Age: 35-39 years (<35 yrs as ref)			0.175	0.164	0.154
			(0.33)	(0.36)	(0.39)
Age: 40+ years (<35 yrs as ref)			0.251	0.266	0.257
			(0.14)	(0.11)	(0.13)
Had medication, last 6 months				-0.256	-0.246
				(0.18)	(0.20)
HESN & Depo Interaction					-0.170
					(0.64)
R^2	0.01	0.03	0.04	0.06	0.06
N	161	161	161	161	161

Table 8.3. Multiple linear regression correcting for epidemiological confounders associated with HESN women and log Mx1 expression in PBMC. * $p<0.05$; ** $p<0.01$

	Crude	Model 1	Model 2	Model 3	Model 4
HESN vs. NN (ref)	0.277	0.246	0.249	0.248	0.116
	(0.03)*	(0.05)*	(0.04)*	(0.04)*	(0.39)
Depo-provera		0.418	0.369	0.365	-0.046
		(0.01)**	(0.02)*	(0.02)*	(0.85)
Age: 35-39 years (<35 yrs as ref)			0.036	0.034	0.074
			(0.82)	(0.83)	(0.64)
Age: 40+ years (<35 yrs as ref)			-0.225	-0.222	-0.184
			(0.13)	(0.14)	(0.21)
Had medication, last 6 months				-0.055	-0.094
				(0.74)	(0.57)
HESN & Depo Interaction					0.711
					(0.03)*
R^2	0.03	0.07	0.09	0.09	0.12
N	161	161	161	161	161

Table 8.4. Multiple linear regression correcting for epidemiological confounders associated with HESN women and log Mx2 expression in PBMC.* $p<0.05$; ** $p<0.01$

8.4.3 IRF-1 protective genotype and Mx1 Expression

Mx1 as well as Mx2 are important interferon response effector proteins and it has been shown that interferon regulatory factor 1 (IRF-1) expression can lead to increases in Mx1 RNA expression [185]. Our group has characterized specific IRF-1 genotypes that are associated with reduced acquisition of HIV and are enriched in the HESN population. It is proposed that the protective genotypes lead to lower levels of IRF-1 making it difficult for HIV to replicate [97,98]. If IRF-1 levels are low in women with the protective genotype then it stands to reason that this genotype could have also been driving the lower levels of Mx1 seen in our pilot PBMC study in HESN women (Chapter 6). While our population level analysis of Mx1 in PBMC clearly showed that the levels of Mx1 are not different between HESN and high-risk controls, specific IRF-1 genotypes may still contribute to the expression of Mx1. However, when we compared Mx1 expression in the PBMC of individuals with different IRF-1 genotypes using a univarient (Table 8.5) analysis, we found there were no differences between IRF-1 protective and non-protective genotypes ($p=0.280$). Other epidemiological factors such as age, medication history, and recent infections did not illustrate any significant differences. However, while taking into account status (HESN or NEG) with respect to having the IRF-1 protective or non-protective genotype using an interaction model (Table 8.5), it was found that HESN women with the protective IRF-1 genotype have significantly higher levels of log Mx1 protein compared to HESN women with the IRF-1 non-protective genotype ($p=0.025$). There was no association found with IRF-1 genotypes and Mx2 expression (data not shown).

	No.	Mean	SD	Median	IQR	
TOTAL	63	4.7	0.9	4.8	4.2 - 5.5	
HESN vs. NEG (ref)						0.390
NEG	22	4.6	0.6	4.6	4.2 - 4.9	
HESN	41	4.8	1.0	4.9	4.3 - 5.5	
IRF-1						0.280
Non-protective	51	4.7	0.8	4.8	4.2 - 5.4	
Protective	12	5.0	0.9	4.8	4.3 - 5.7	
Age						0.258
<34	36	4.6	0.9	4.6	4.0 - 5.3	
35-39	14	4.9	0.9	5.0	4.6 - 5.6	
40+	13	4.9	0.6	4.9	4.4 - 5.5	
Had medication, last 6 months						0.370
No	53	4.8	0.8	4.8	4.3 - 5.4	
Yes	10	4.5	0.9	4.5	4.0 - 5.5	
Had infection, last 2 months						0.513
No	53	4.7	0.9	4.8	4.3 - 5.5	
Yes	10	4.6	0.9	4.6	3.9 - 5.3	
Status & IRF1 Interaction Indicator						0.025
NN, Non-Protective	17	4.7	0.6	4.8	4.3 - 5.2	
NN, Protective	34	4.6	0.9	4.8	4.0 - 5.5	
HESN, Non-Protective	5	4.1	0.5	4.2	4.0 - 4.3	
HESN, Protective	7	5.5	0.6	5.6	4.8 - 6.2	

Table 8.5. Patient characteristics / interaction model analysis associating Mx1 expression and IRF-1 genotypes. The data excluded individuals who were menopausal or pregnant at the time of sampling. * $p<0.05$; ** $p<0.01$

8.5 Summary

In summary, by using a combination of multiple approaches (basic science and epidemiological principles) we have shown that HESN women express the same levels of Mx1 as susceptible controls, and we concluded that Mx1 has no effect on HIV-1 replication in-vitro in our current experimental model. While Mx1 levels could not be attributed to the HESN phenotype, Mx1 expression levels were found to be significantly lower in women who reported condom use and Depo-Provera as a contraceptive method compared to nothing at all (Figure 8.4A). We also showed that IRF-1 protective polymorphisms associated with increased Mx1 in HESN individuals. While Mx1 levels were overall unchanged, the levels of Mx2 were significantly higher in HESN women. Additionally, HESN women who reported Depo-Provera use as a contraceptive were found to be responsible for the high expression of Mx2 seen in the HESN group. For the first time we demonstrate that Mx2 and not Mx1 could be a potent anti-HIV interferon molecule that contributes to the HESN, HIV-1 resistance phenotype.

Chapter 9. Discussion

Understanding the correlates of natural immunity to HIV infection will likely be a key component of HIV vaccine and microbicide design. The aim of the research presented in this thesis was to confirm at the cellular level an immune quiescent profile in PBMC and CMC from HESN women, which had yet to be confirmed at a proteomic level. Additionally, we aimed to identify novel correlates of protection that could be used in the development of future vaccine or microbicide strategies. **The central hypothesis of this thesis is that: Quantitative shotgun proteomics of immune cells from HESN women will demonstrate an immune quiescent profile, represented by the altered expression of pathways in immune activation, cell recruitment / migration, and proteins involved in HIV replication pathways.**

9.1 Shotgun Proteomics Development

To characterize the proteome of PBMC and CMC from HESN women a proteomics / systems biology platform needed to be developed. The major objective of Chapter 3 was to develop a method for the identification of proteins from ex-vivo lymphocyte (PBMC) protein lysate. The identification of optimal lysis buffers, peptide digestion, and peptide fractionation techniques was also needed to identify and quantify thousands of proteins in a complex sample.

9.1.1 Cell Lysis and Digestion

Lysis buffers (Tris or RIPA) using a conventional in-solution digest were fairly well optimized (Figure 3.3) however we were unable to obtain complete digestion of the sample using this method. Optimization of the cell lysis (4% SDS) and protein digestion protocol (FASP) (Figure 3.2) allowed for complete digestion of the sample [165]. While this method was time consuming it increased our ability to isolate a wider range of proteins with regards to molecular weight and hydrophobicity and was used for all future studies outlined in this thesis.

9.1.2 Offline Peptide Fractionation

Considerable effort went into optimizing the Hp-RP chromatography to maximize the fractionation of digested peptides, which would translate into more peptides and proteins identified by mass spectrometry. Adjusting the organic gradient to maximize the elution of peptides from the reversed phase column and identify and eliminate the protease inhibitor tablet interference was a key milestone. To compensate for the removal of the protease inhibitors all samples were immediately boiled following lysis to inactivate protease activity as soon as possible. The addition of this step replaced the need for the protease inhibitor cocktail, which masked the elution of hydrophobic peptides from the reversed phase column (Figure 3.6). A comparative analysis of identically processed samples between the Qstar and Orbitrap XL mass spectrometer also clearly indicated that the Orbitrap XL would be the best mass spectrometer for identifying as many peptides and proteins as possible (Figure 3.8).

To increase the throughput of our digestion protocol we evaluated the difference between using a 30kDa or 10kDa molecular weight filter during the on filter digestion method. However this lead to a substantial decrease in the number of peptides and proteins identified when using the 30kDa filter, leading us to retain the use of the 10kDa filter for future studies. While this was essentially a single experiment with no replicates we believed the large difference between the two methods (39% and 51% more proteins and peptides respectively) warranted the continued use of the 10kDA filter.

9.2 Quantitative Proteomics Development

The major objectives of Chapter 4 were to identify an optimal peptide fractionation technique as well as optimize the fractionation, identification and quantification of iTRAQ labeled peptides. These experiments finalized the proteomics discovery pipeline allowing us to address the main hypothesis of this thesis.

9.2.1 Hp-RP and OG-IEF Fractionation

A suitable proteomics pre-fractionation technique should maximize protein/peptide identifications while striving for reproducibility of those identifications across biological replicates. We compared in-solution IEF (OG-IEF) and high-pH reversed-phase LC (Hp-RP) by in depth replicate analysis in order to ascertain the best method for proteomic characterization of ex-vivo PBMC.

This comparison demonstrated that Hp-RP lead to the identification of 17% more proteins and 16% more peptides than the OG-IEF (Figure 4.2). Peptide sampling inefficiencies and overall identification rates between the two techniques were the same, making potential MS based method confounders unlikely. In addition, there was no distinct biophysical bias for either of the two techniques despite the drastically different properties by which the peptides were fractionated (Figure 4.5). One explanation for the difference in protein identifications is that isoelectric focusing of peptides results in distinct clusters due to similar charge states whereas Hp-RP has a wider range of elution values (Figure 4.1). This is illustrated by the uniform distribution of identified peptides across all fractions by Hp-RP (59%) compared to OG-IEF (Figure 4.3A vs. B). However, it is possible that alternative strategies to further expand certain pI ranges or alter the linearity of the IPG strip could be attempted to increase the peptide and protein identification capacity of the OG-IEF technique [169,186].

Hp-RP may have also out preformed OG-IEF in protein and peptide identifications at least in part due to its significantly higher resolving power, as illustrated in Figure 4.4. Ninety-one percent of all peptides identified were limited to single fractions for the Hp-RP, whereas OG-IEF had 78% of all peptides in single fractions. This is consistent with previous reports for OG-IEF [169,187-189]. There was a possibility that our fraction mixing scheme for the Hp-RP may have improved our perceived resolving power however, running all 24 fractions out separately only lead to a 1% decrease in apparent resolution (data not shown and Figure 4.4).

Additionally, there were important practical/operational differences that we considered when ultimately choosing the optimal fractionation technique. One advantage offered by Hp-RP LC over OG-IEF is the ability to visualize a sample's chromatographic profile prior to committing to a lengthy series of nLC/MS/MS runs. Visualizing the chromatogram enables sample quality screening and provides an approximate measure of peptide relative abundance per fraction. The OG-IEF readout is less informative, providing only a voltage, which, when stabilized, indicates that a run is complete. Another practical difference between OG-IEF and Hp-RP is the runtime per sample. The fractionation process takes less than two hours for Hp-RP but takes over 24 hours for OG-IEF. While OG-IEF does provide the ability to fractionate multiple samples in parallel, Hp-RP is easily automatable for running samples in series. With these observations we concluded that the Hp-RP fractionation method would be the most appropriate chromatography technique for use in the proteomic characterization of immune cells from HESN women. While the gold-standard chromatography technique for pre-fractionation of samples (SCX) is still the norm for most MudPit based experiments Hp-RP is now becoming a more common approach [166,190-194] lending evidence to support the use of this chromatography setup in our studies.

9.2.2 iTRAQ labeled peptides and Mass spectrometry

During the optimization of our iTRAQ protocol we discovered that iTRAQ labeled PBMC digest had a drastically different elution profile when subjected to Hp-RP

fractionation. As a result the peptide elution profile was severely compressed compromising our ability to maximize the spread of peptide elution. A new gradient was generated to adjust for these retention time shifts (Figure 3.16). This new gradient was then used for all future experiments that involved PBMC and CMC peptide fractionation. At the time these experiments were conducted there was no literature suggesting iTRAQ labeled peptides drastically shift retention times. More recently, Pichler et. al. have observed this phenomenon but did not attempt to adjust their chromatography to account for the iTRAQ retention time shifts [195].

Finally, the HCD collision energy was optimized on the Orbitrap XL in order to maximize the intensity of the iTRAQ reporter ions and peptide identifications. Unfortunately, due to the 1/3 rule, the ion trap in the Orbitrap XL is unable to detect ions below 1/3 the mass of the peptide, making detection of the iTRAQ reporter ions impossible. To circumvent this challenge an additional HCD collision cell was used. Our optimization experiments found that an HCD energy of 40% was optimal for obtaining good identification without compromising iTRAQ reporter intensities (Figure 4.9 and 4.10). The optimization of the HCD energy as well as reducing the isolation width for peptide fragmentation was also done in an attempt to reduce the effects of peptide co-isolation which leads to an under estimation of iTRAQ reporter ratios [171]. Unfortunately, the method utilized here will suffer from this ratio under estimation and can only be corrected using altered mass spectrometer acquisition parameters that would significantly reduced the numbers of proteins and peptides identified [196]. Other methods such as pulsed Q dissociation (PQD)

are capable of measuring iTRAQ reporter ions in the ion trap. However, PQD makes use of a slower duty cycle on the mass spectrometer resulting in less proteins and peptides identified [197]. At the onset of the PBMC and CMC characterization studies in HESN women the newer Orbitrap Velos mass spectrometer became available. The Orbitrap Velos overcomes the need for an ion trap during peptide identification and has significantly improved scan times leading to more peptide identifications [198]. Additionally, recent studies also suggest that an optimized HCD energy of 40% on the Orbitrap Velos is ideal for identification and quantification of iTRAQ labeled peptides, making our optimized parameters interchangeable between mass spectrometers [199].

9.3 iTRAQ Proteomics in PHA Stimulated Peripheral Blood Mononuclear Cells.

In Chapter 5 we examined the utility of a proteomics approach to measure immune activation using PHA stimulation of primary human cells as a model for immune activation and by extrapolation, antigen specific responses in future studies. We were particularly interested in evaluating in a qualitative way the ability of our optimized iTRAQ proteomics platform to identify proteins and functional networks with immunological relevance to HIV and infectious disease. PBMCs from multiple donors were stimulated with PHA at multiple time points, and changes in the cellular proteome were measured over a 24-hour period. In total 1816 unique proteins were quantified making this study one of the most comprehensive proteomic analyses of PBMCs under immune activation conditions to date [160,161]. Of the 1816 proteins, 220 were identified as significantly altered in at least one time-point. Cluster and correlation

analysis indicated considerable heterogeneity between patients at early stages in response to PHA followed by a common activation pattern by 24 hours. A functional pathway analysis allowed us to verify previously known as well as unexpected networks induced by PHA stimulation of PBMC.

Of the 220 individual proteins (Table 5.1) that we found to be significantly over-expressed by 24-hours, IDO1, SERPINB2, and NAMPT, were the most differentially regulated. These proteins were also found to be differentially expressed in a similar proteomic study, however unlike in our study, the authors were unable to report the extent of the expression level changes [161]. IDO1 is thought to have significant immunosuppressive effects [200] and is known to be expressed in response to HIV infection of pDC's [201] and induced by HIV accessory proteins, demonstrating that this proteomics platform can detect proteins and pathways relevant in HIV infection processes [202,203]. Further examination of the 1816 proteins examined showed that well over 75 proteins with proliferative functions were up-regulated by PHA stimulation (Figure 5.4 F). It is interesting to note that the most highly overexpressed protein in the entire dataset (IDO1) seems to have significant anti-proliferative functions. On the other hand, NAMPT is known to be a strong activator of proliferation and fits well into the overall PHA proliferative responses observed [204]. SERPINB2 has been implicated in cell survival as an anti-apoptotic protein (Figure 5.4 A&D) and has also been associated with enhanced HIV replication and transcription [205]. These three proteins with the exception of IDO1 indicate a strong proliferative and anti-apoptotic proteomic signature in activated PBMCs, which could allow for increased infection and virus production. It

also suggests that some of the earliest signatures of immune activation are the induction of negative regulators of such, which clearly in the case of PHA stimulation, is not able to limit full immune activation and cellular proliferation.

Other differentially expressed proteins including numerous acute phase response proteins (FGA, FGB, and FGG), were among the top 10 down-regulated proteins by 24-hours (Table 5.1). Most of these proteins are found within the extracellular matrix and are involved in adhesion and migration pathways, both of which were found to be significantly suppressed in response to PHA stimulation (Figure 5.4 B & E). This was a surprising finding given that one of the in-vitro phenotypic hallmarks of stimulating PBMC with PHA is significant cell adhesion and clumping [206]. Overall, proteins that were in favor of viral infection and proliferation had much higher levels of activation compared to the adhesive and migratory proteins (Table 5.1) demonstrating a possible tug of war between an activated and naïve state, possibly suggesting that different thresholds of protein induction are required to elicit specific cellular pathways such as activation, and / or adhesion. In the case of PHA stimulation the activating effect clearly overcomes the suppressive migratory/adhesive effect, allowing for cellular proliferation and HIV infection. Furthermore, we note that the majority of pathways depicted in Figure 4 are suppressed, highlighting the specific expression of functions/proteins as key components in the balance between an activated or naïve state. It is clear from this data that proteomics is a useful tool for monitoring immune activation and identifies proteins one may not expect to be induced in a model of immune activation allowing for a more global understanding of the immune activation process.

Potential pathways affected by PHA stimulation were visualized using cluster analysis of all the proteins identified in this study (1816) in addition to the significantly deregulated proteins (220). Figure 5.2 A shows that the early stage PBMC response to PHA stimulation between patients was very heterogeneous although a specific immune activation profile common to all subjects was reached at 24-hours. This data suggests that patients may use different mechanisms/pathways during PHA stimulation to arrive at a similar cellular state by 24-hours. However, analysis of later time points, not examined in this study (8 or 16 hours) would be helpful in clearly defining this. When only the significant proteins are investigated (Figure 5.2 B) this heterogeneity disappeared and data from specific time points cluster more closely than individual subjects, suggesting that while the most differentially regulated proteins are common in their induction, there may be differences in the underlying activation profiles. The majority of proteomic analyses in the literature implement cluster analysis after filtering for significantly regulated proteins, however if we had exclusively done this type of analysis we would have missed out on what could be a very interesting observation of individual patient heterogeneity. Using a typical reductionist approach, we may have missed these potentially very important differences in individual responses to immune activation, even to that of a mitogen such as PHA. It is for this reason that subsequent cluster analyses in this thesis were generating using both differentially regulated proteins as well as the complete protein data set.

The results of this study have practical implications with respect to ex-vivo studies. Cluster analysis as well as the correlations indicated that by 24-hours PHA alters the ex-vivo state of PBMC to the point whereby heterogeneity between individual patients is no longer identifiable (Figure 5.2A and Figure 5.3D). This has relevance to HIV transmission in that HIV infection across mucosa is relatively infrequent (one to a few founder viruses per transmission) suggesting that minor changes in immune activation may be critical in susceptibility to HIV infection [207]. Studies in HESN individuals seeking to evaluate HIV susceptibility in ex-vivo infection experiments have in the past made use of PHA stimulation (standard lab practice) and have not been able to detect differences between HESN and controls [106]. However, in later studies when PHA stimulation was not used, differences in susceptibility were readily measured [176,208]. Thus, it may be necessary to rethink standard laboratory practices when studying certain populations and their susceptibility to HIV infection ex-vivo.

Finally, the IPA functional network analysis complements our findings made at the individual protein level whereby cell pathways involved in survival, proliferation and HIV infection were all predicted to be overexpressed at 24 hours post PHA stimulation (Figure 5.4). In addition, we were able to observe decreased expression of apoptosis pathways (Figure 5.4 A) further corroborating our findings of an increased anti-apoptotic, proliferative, and HIV replication response. These findings highlight the use of a quantitative proteomics approach in not only measuring single correlates of activation but also allowing for a global functional network analysis of stimulated PBMC.

One challenge that was not addressed in this study pertains to the culture / PHA stimulation of PBMC and how it affects particular subsets of cells within a mixed lymphocyte population. PBMCs are a mixture of neutrophils, lymphocytes and granulocytes, however during the 24-hour stimulation, granulocytes die off and monocytes begin to differentiate into macrophages adhering to the tissue culture flasks. Additionally, lymphocytes become enriched during the stimulation as they begin to proliferate (~90% of the cells are lymphocytes by 24-hours) [209]. This shift in the proportion of monocytes/macrophages was somewhat mitigated by extensively washing adhered cells off the plastic before each sampling time point, however it does not negate the fact that these cells have significantly changed over time. One way to control for this would have been to add additional media controls for each time point or to further purify a specific subset of cells (CD4+ T cells only) in order to control for the differing proteome expression contributions from multiple cell subsets.

The proliferative response observed in this study complements the most widely accepted models of immune activation and the pathways utilized for HIV infection in-vivo, suggesting that early induction of immune activation pathways may vary between individuals. This methodology is an excellent tool for describing the global proteomic response in host immune activation. Furthermore, a systems biology approach such as the one described is ideal for validating the immune quiescence hypothesis at the proteomics level and to identify possible markers of natural resistance to HIV acquisition.

9.3 Systemic Immune Quiescence in HESN Women

The central hypothesis of this thesis was that HESN women would have an immune quiescent proteome signature in ex-vivo isolated PBMC and CMC. The objectives of Chapter 6 were to determine the proteomic profile of ex-vivo PBMC from HESN women and susceptible controls using the optimized qMS approach and validate protein biomarkers of interest.

9.4.1 Individual deregulated proteins in PBMC from HESN women

We were able to quantitatively identify 2913 proteins from PBMC in HESN and high-risk negative controls. Of those proteins, 7 were found to be under-expressed and 19 over-expressed in HESN women compared to susceptible negative controls (Figure 6.2). From this list we chose to validate the expression of 7 proteins by Western blot from the original samples (Mx1, HLA-C, FGD6, PDPK2, AHSG, A2M, and RGS6). Due to limitations in the amount of sample, and available reagents, several markers could not be tested further. As a result we were able to examine three proteins of interested (Mx1, HLA-C, and RGS6). As can be observed from the Western blots (Figures 6.5-6.7) there was considerable protein degradation. This is most easily observed by counting the number of Actin-β (red) bands in each patient sample compared to the standard PBMC control lysate on each blot. All samples were limited to 3 freeze thaw cycles however, long term storage of proteins in 4% SDS buffer may have lead to considerable protein degradation. This observation likely had a major effect on our ability to correctly validate the protein biomarkers from our mass spectrometry results. Unfortunately we were unable to compensate for this degradation since all three protein biomarkers were detected at the same

time making it difficult to determine which breakdown product belongs to which protein. The results in Figure 6.8 indicate that none of the biomarkers measured by Western blot were significantly different between HESN and negative controls, although Mx1 and HLA-C trended towards significance. Interestingly, when correlating the values measured by mass spectrometry and by Western blot, both Mx1 and HLA-C levels significantly correlated. This tells us that the levels measured by both techniques agree with each other in some cases (RGS6 did not correlate). Since Mx1 had the strongest correlation, was biologically plausible and both techniques indicated a trend towards lower expression of Mx1 in HESN women this protein was chosen for further testing as a potential biomarker of immune quiescence at a population level in the Pumwani cohort (Chapter 8).

These results highlight some important strengths and weaknesses in these two approaches (qMS and Western blot). In this case we would tend to believe the results measured by the mass spectrometer more so than the Western blot analysis due to the inherent nature of how the samples are processed before analysis. Proteins need to be digested into peptides before mass spectrometry analysis and are thus able to tolerate some protein degradation in the sample. While the intensity of a degraded target protein analyzed by Western blot will be split up over several bands, making it extremely difficult to accurately quantify a particular protein. However, by digesting the sample into peptides for mass spectrometry it becomes almost impossible to gather any quantitative information on protein isoforms or splice variants, where as Western blot is better suited for this analysis.

Three of the biomarkers identified have been previously identified in HESN women from the same cohort, strengthening the validity of our technique. HLA-C under-expression was observed in whole blood by micro array analysis in the same HESN women [124]. Additionally, a form of A2M and AHSG were also shown to be up-regulated in the genital secretions of HESN women [111,112]. These initial mass spectrometry results are very encouraging and solidify the qMS approach as a valid method for identifying biomarkers of the HESN phenotype however, it also highlights some weaknesses, with a need for increased samples sizes and solid epidemiological records.

9.3.2 Ex-vivo proteome profile of PBMC from HESN sex workers

Unsupervised hierarchical clustering using all the proteins quantified by iTRAQ mass spectrometry in the PBMC of the individuals in our study was conducted to identify and characterize novel proteome signatures that could distinguish HESN and susceptible controls. However, while there was clustering of specific individuals a clear proteomic profile that describes the HESN phenotype was not observed. Interestingly, a strong clustering of four individuals was observed with a unique proteomic signature. Further analysis indicated that these individuals had ongoing infections at the time of sampling indicating that our analysis is capable of defining ongoing disease states. These findings highlight the challenges and opportunities associated with these types of studies. The ability to identify active

infections by this qMS approach has significant potential, and should be further examined.

During sample collection for this pilot study, epidemiological factors such as age, subclinical infections, contraception could not be controlled for. The effects these factors may have on the larger proteomic profile to define a specific signature need to be appropriately controlled for in future studies. Post-hoc analysis at the individual protein level for Mx1 indicated that at least 32 individuals in each group would be needed to detect statistically significant differences ($p < 0.05$). However, proteome profiling studies deal with several thousand proteins, which will inevitably increase the number of patients needed to correctly define a proteomic signature in HESN women. Additionally, the p-values obtained from the study were not able to stand up to any multiple correction technique further highlighting the need for much larger and appropriately powered studies. The majority of proteomic studies in the literature suffer greatly from his issue. This is most likely due to the practical (cost, core facility, shared time constraints) and technical (linking individual iTRAQ runs, or chromatogram alignment in the case of label free) issues associated with proteomics experiments.

While we were not able to identify a specific proteomic profile that describes the HESN phenotype, we were able to identify potential biomarkers that were able differentiate between the HESN and susceptible controls (Figure 6.10). However, as seen in Chapter 5 and now Chapter 6, clustering with only the significantly

expressed proteins intentionally biases the analysis making it difficult to identify more subtle immunological responses or states that likely represent true biological phenotype. These data provide strong evidence to support the use of qMS to define different immunological states when important epidemiological confounders can be appropriately controlled for.

9.3.3 Biological Functions in the PBMC of HESN sex workers

An analysis of the functional expression differences between HESN and high-risk negative controls was also undertaken. Figure 6.11 clearly indicates that HESN women do not appear to have a functional proteomic profile that reflects an immune quiescent phenotype. There was clear overexpression in numerous immunological functions that we hypothesized would be down-regulated or suppressed in HESN women (viral infection, HIV infection, lymphocyte migration, proliferation, and activation). This analysis excluded the four individuals with on going infections ruling them out as potentially biasing the functional signature. This analysis suggests that HESN women do not have an immune quiescent phenotype at the systemic level in PBMC, contrary to our previous genetic studies. Our proteomics technique was unable to detect an immune quiescent state in HESN women. One likely reason is the lack of sensitivity of modern proteomic techniques. Typically we identify only 4000 of the most abundant proteins in a proteomics experiment where RNA expression analysis (micro array) measures the expression of tens of thousands of genes biasing the information generated by the two techniques and their ability to identify an immune quiescent signature. Additionally, mass

spectrometry suffers from an extreme case of missing data while the micro array technologies are able to measure quantitative levels for all the genes expressed in a given sample. The ability to draw on tens of thousands of genes to define a particular state such as the one seen in HESN women is arguably much easier to achieve with fewer study participants than the mass spectrometry approach that allows for approximately 4000 protein signatures to define the same state. For this reason it would be reasonable to conclude that the pilot studies performed were drastically under-powered making it difficult to definitively prove our hypotheses. Finally, it is also possible that in this case the RNA expression does not equal the protein expression, giving rise to the question about which technique more accurately identifies a particular biological state.

It is clear that the HESN phenotype is extremely complex and a significantly larger / more robust study controlling for basic epidemiological factors is clearly needed to make any final conclusions. Additionally, while these results suggest that HESN women at the systemic level do not have an immune quiescence phenotype a proteomic study of immune cells from the initial site of acquisition (genital tract) may prove drastically different.

9.4 Genital Tract Immune Quiescence in HESN Women

The hypothesis of this chapter was that cervical mononuclear cells (CMC) isolated from HESN women would posses an immune quiescent proteomic profile indicated by a lower expression of proteins related to immunity and infection. Our objectives

were to define a proteomic profile in CMC that distinguishes between HESN and susceptible women in the cohort.

9.4.1 Individual deregulated proteins in CMC from HESN women

Statistical analysis of individual proteins (Figure 7.1) yielded 74 significantly under-expressed and 37 over-expressed proteins in the CMC of HESN women compared to high-risk negative controls. This is considerably different from the 7 under-expressed and 19 over-expressed proteins observed in the PBMCs from the same individuals (Figure 6.2). While both studies were treated almost identically in terms of sample processing and mass spectrometry analysis, there were some differences. Mainly, the CMC isolations in general yield fewer cells and so less protein was available for down stream processing. As a result, in an attempt to compensate for this deficiency a slightly different online nano-LC setup was used to analyze the CMC samples. With the PBMC study an analytical and trap column were used however with the CMC samples a slightly longer analytical column and no trap column was used. The reasoning behind this was that we had reduced sample amounts in terms of micrograms of protein to work with and did not want to loose sample when loading and washing each fraction for analysis onto the trap column. This subtle change in the method also resulted in a general increase in total proteins identified and quantified between the PBMC and CMC studies even though we had less sample to work with (PBMC: 2913, and CMC: 3972 proteins).

Of interest both PBRM1 and PSIP1 (LEDGF) were both highly down regulated in CMC from HESN women. PBRM1 allows for Tat dependent activation of the HIV LTR making its decreased expression level in the CMC of HESN women biologically relevent [210]. PSIP1 also known as LEGDF/p75 is an important co-factor for HIV integrase and is required for integration into the host genome [211]. Additionally, lower LEGDF levels have been described in HESN individuals from a Senegalese cohort and may contribute to HIV resistance [212]. Of the proteins overexpressed in CMC from HESN women both CHMP4C and TNFRSF21 were of particular interest. While CHMP4C is associated with HIV budding [213], tumor necrosis factor (TNF) receptors are important in causing apoptotic cell death and may be a promising target for new therapeutic approaches [214]. Unfortunately, there was not enough sample left after mass spectrometry analysis to confirm these markers by Western blot. Further studies in particular with respect to the expression of LEGDF/p75 in the genital tract are warranted.

9.4.2 Ex-vivo proteome profile of CMC from HESN sex workers

Unsupervised hierarchical clustering using all the proteins identified and quantified by in the CMC (3972) of the individuals in our study was conducted to identify novel proteome signatures that could distinguish HESN and susceptible controls. The same four individuals from the PBMC study (Chapter 6) were also excluded in this analysis due to active infections. Unlike the almost completely random clustering observed in the PBMC study we observed a more structured clustering (Figure 7.2). In general we were still unable to establish a perfect, unequivocal proteome profile

that distinguishes HESN women from high-risk negative controls. However, in general most HESN women clustered to the left of the dendrogram while some susceptible women exclusively clustered to the far right. As with the PBMC study it is likely that we could not clearly differentiate between the two groups of women, as we were unable to control for the epidemiological factors that could affect the overall proteome profile. A much larger study taking into account basic factors such as age, infection, and contraception may allow us to discover a more comprehensive profile that describes the HESN phenotype.

9.4.3 Biological functions in CMC of HESN sex workers

We undertook a high level analysis of the biological functions that were differentially regulated in CMC between HESN and high-risk negative controls. We observed an extremely polarized profile compared to that in the PBMC from the same women (Figure 6.11 and Figure 7.3). Sub-functions such as viral infection, HIV infection, immune responses of cells, leukocyte activation, homing, and migration were all predicted to be reduced in the CMC of HESN women compared to high-risk negative controls. Decreases in immunological functions such as these may be the first indications of an immune quiescent profile at the proteomic cellular level in HESN women. Reduced homing and activation of leukocytes is key in limiting establishment of HIV infection foci and systemic dissemination of the virus. An observed reduction in the sub-function of viral infection could also reduce the ability of HIV to establish an ongoing local infection. Of some concern is the identification of decreased expression in the cell death of epithelial cells. While functionally this would fit well with a strengthening of the epithelial barrier making

HIV target cells more difficult to infect it may also signal that the CMC isolation may have contained some epithelial cell contamination and was not as homogeneous a population as the PBMC study.

This analysis suggests that immune quiescence may be more of a mucosal phenomenon than systemic at the proteomic level in HESN women. It also corroborates our findings for the first time at the genital tract level of lower expression in important immune activation pathways from previous systemic studies in HESN women [123-125]. Studies in other HESN populations such as MSM [215] and sero-discordant couples [216] have also identified lower activation markers on systemic target cells but until now a study identifying a proteomic quiescent phenotype in immune cells from the genital tract had yet to be presented.

9.4.4 Comparative Proteomic Signatures in PBMC and CMC from HESN women

We also conducted a cluster analysis of all the proteins that were identified and quantified in common between the CMC and PBMC studies. The unsupervised hierarchical cluster analysis (Figure 7.4) clearly showed that the proteome profile is able to distinguish between PBMC and CMC in the same individuals. However as we have seen in the previous two chapters the proteome profile was not able to define HESN or high-risk negative control status in the Pumwani cohort. Additionally, while the CMC study identified more proteins it could have affected our cluster analysis (Figure 7.2) leading to an increase in the number of significantly different proteins. However, the functional and comparative analyses discussed make use of proteins that were found in common between both studies making it unlikely that

our adjustment to the nano-LC columns had an effect on protein expression. The ability of proteomics to detect only the most abundant proteins in a sample compared to RNA expression analysis is a clearly a major drawback to this approach. Additionally, the ability to control for important epidemiological confounders, mixed with the practical aspects (number study subjects) of employing mass spectrometry technology needs to be addressed in future studies.

9.5 Mx1 and Mx2 in the systemic compartment of HESN Women

9.5.1 Mx1 expression during in-vitro HIV infection

Based upon the potential data that Mx1 expression is decreased in HESN women, we hypothesized that Mx1 would have an enhancing effect on HIV replication in-vitro, which in part could have explained our initial observations. Our results indicated that at varying multiplicities of infection, SUPT1 cells over-expressing Mx1 neither enhance nor minimized HIV infection in-vitro (Figure 8.1). These results complement the findings from three other recently conducted studies [119-121] strongly suggesting no role for Mx1 during HIV infection in-vitro. For the most part our data supports these new findings however, Figure 8.1A indicates a strong difference in HIV replication between SUPT1 cell only / vector only controls and GFP/Mx1 expressing cells at a MOI of 1. Unfortunately, our Western blot analysis of Mx1 expression in the various cells lines used in these experiments (Figure 8.1D) is missing the Mx1 expression in GFP expressing cells. It is possible that our irrelevant GFP protein control is inadvertently activating Mx1 or some other infection enhancing response making this an inappropriate control for this experiment. If this

is the case then there could still be a case for Mx1 enhancing HIV infection at higher MOI's. Final experiments with all the appropriate controls need to be conducted to determine, in our hands, if Mx1 enhances HIV infection. However, based on recently conducted studies in the literature, a role for Mx1 in increased HIV replication does not seem likely.

9.5.2 Population level analysis of Mx1 and Mx2

The major hypothesis of Chapter 8 was that HESN women have low levels of Mx1 and high levels of Mx2 compared to susceptible controls at a population level in the Pumwani sex worker cohort. With a large number of women in this study we could also control for major epidemiological confounders that we were unable to control for in our pilot studies (Chapters 6 and 7). Mx1, which was found to be associated with the HESN phenotype in Chapter 6 has been well characterized with respect to Influenza infection and is recognized as an important innate factor in clearing a variety of DNA and RNA viruses [179]. The Mx1 gene encodes a protein referred to as MxA while the Mx2 gene encodes a protein referred to as MxB in humans [179,217,218]. The expression of both proteins is induced by IFN-α and IFN-β and is thought to form ring structures through oligomerization [219,220]. These rings structures allow binding of nucleocapsid at virus replication sights disrupting early replication events [221]. Additionally, MxA is exclusively a cytoplasmic protein while Mx2 is present in both the cytoplasm and the nuclease. At the onset of the studies proposed in this chapter the prevailing dogma indicated that Mx2/MxB had no antiviral properties while Mx1/MxA was solely responsible for antiviral activity. Additionally, no studies had been done to

ascertain the role either of these proteins plays with regards to HIV infection. However, three recent high profile studies have been published that convincingly show MxB inhibits HIV replication while MxA was not shown to have an effect [119-121]. These studies coupled with the close amino acid homology between MxA and MxB warranted further examination of both proteins in the context of the HESN phenotype.

We showed that in a population level analysis of over 200 women from the Pumwani cohort that Mx1/MxA exhibited no significant expression differences between HESN women and susceptible controls (Figure 8.2A). However there were some interesting expression differences in Mx1/MxA between women who reported using a condom or nothing at all for contraception (Figure 8.3A). Several studies have suggested the use of MxA as a marker for general inflammation, viral infection, HIV therapy failure, [222-224] and has even been linked to HIV seroconversion [225]. It is possible that we are observing higher levels of MxA in women (due to sex / allo induced inflammation) who do not use condoms as contraception, perhaps making MxA a marker or risk factor for HIV acquisition because of increased inflammation. However, our in-vitro infection studies (Figure 8.1) as well as three other published studies indicate that Mx1/MxA have no direct enhancing or inhibiting effect on HIV replication. While Mx1 is most likely precluded from having any direct effect on infection it may still be a good indicator for women who are at increased risk of HIV acquisition possibly allowing for tailored and intensified risk reduction counseling.

The Mx2/MxB expression in PBMC isolated from HESN women and high-risk negative controls was also evaluated and was found to be significantly overexpressed by HESN

women (Figure 8.2B). Adjusting for major epidemiological confounders like age, medication usage, recent infection, Depo-Provera use, menopause, and pregnancy the association was still significant (Table 8.4). An interaction analysis showed that HESN women who use Depo-Provera as contraception were largely responsible for the increased expression of Mx2 in PBMC from HESN women (Figure 8.5). There is still considerable controversy in the HIV field as to the effects of contraception on HIV acquisition. Some studies have shown that Depo-Provera use can lead to a higher risk of HIV acquisition while others claim no effect [184,226-232]. Many early non-human primate HIV infection studies made use of Depo-Provera to increase the success of establishing HIV infection. HESN women who use Depo-Provera have much higher levels of Mx2 indicating that HESN women may have distinct differences in how they respond to contraception and that these responses at least in terms of Mx2 may protect against HIV acquisition. It is also possible, however unlikely, that a selection bias of women overexpressing Mx2 in HESN Depo-Provera users could have occurred. The use of Depo-Provera in the Sub-Saharan setting is quite high and is a relatively inexpensive alternative to oral and IUD contraception. It is possible that Depo-Provera use could be a proxy for some other socio-economical factor we were unable to capture in our study. However, given that Depo-Provera has been shown to have a suppressive effect with respect to interferon responses there is likely a meaningful biological interaction between Depo-Provera and Mx2 in HESN women [233,234].

Of particular interest are the observations by Kane et. al. which observed higher potency of Mx2 mediated HIV-specific restriction in non-dividing cells compared to

actively dividing cells. [120] This observation fits well with our model of immune quiescence where resistant women have lower baseline activation leading to less HIV target cells at the site of HIV acquisition. This observation suggests a specific HIV restriction factor (Mx2) may actually play a role in enhancing the immune quiescence phenotype and warrants serious investigation. The discovery of Mx2 overexpression in HESN women (an important HIV restriction factor) is clearly a strong marker of immune quiescence and further studies to characterize the importance of this protein in HIV acquisition are necessary.

9.5.3 IRF-1 protective genotype and Mx1 Expression

While we conclusively found no significant differences in Mx1 protein expression between HESN and susceptible controls, interferon regulatory factor-1 has been shown to regulate the gene expression of Mx1. As such we attempted to correlate the levels of Mx1 with the IRF-1 protective genotypes that have been associated with HESN women and reduced susceptibility to HIV acquisition. We did not observe any differences in Mx1 expression between protective and non-protective genotypes of the women in our study. However, when taking into account HESN status we did see an association with increased Mx1 expression in HESN women with the protective IRF-1 genotype. Since IRF-1 genotype is independent of HESN status and we observed no expression differences between protective and non-protective genotypes in terms of Mx1 protein expression this is likely a spurious result.

9.6 Contributions to the Field of Proteomics

A highly detailed studied on the differences between two very prominent chromatography techniques (Hp-RP and OG-IEF) were carried out showing that Hp-RP was the better option for shotgun proteomics studies of complex ex-vivo immune cells (Chapter 4). This comprehensive analysis was lacking in the literature and has now made it clear which technique should be used as well as the advantages and disadvantages depending on the proteomic application.

Studies using the iTRAQ proteomics pipeline developed in this thesis were able to identify pathways and functions associated with immune activation in the response to PHA stimulation (Chapter 5). In addition, the typical analysis that focuses only on statistically significant proteins between study groups or conditions may result in the overlooking of important immunological responses (Chapters 5 + 6). The approach we have undertaken has allowed us to identify individuals with ongoing infections that significantly skewed the protein expression analysis, providing important proof of concept and further evidence to support iTRAQ proteomics as a new technique for measuring global immune responses to infection.

9.7 Major Findings

Functional network analysis of PBMC from HESN women compared to high-risk negative controls illustrated that there was no clearly observed immune quiescent phenotype. However, this analysis was able to detect and identify individuals with subclinical infections. Proteomic analysis of CMCs from the same individuals did

indicate significant decreases in functional pathways associated with viral infection, HIV infection, cellular immune responses, leukocyte activation, migration, and homing. Additionally, essential host factors for HIV replication such as LEDGF/p75 were identified as promising biomarkers in the CMC of HESN women. Altered expression of these proteins and pathways contained therein suggest an immune quiescent phenotype in the genital tract of HESN women. The genital tract is most often the sight of initial HIV infection highlighting the importance of the immunobiology of this compartment in natural immunity to HIV. The novel HIV restriction factor Mx2 was shown to have increased expression in PBMC of HESN women compared to susceptible controls. Further, the interaction between a contentious contraceptive (Depo-Provera) with regards to HIV acquisition and Mx2 expression in HESN women warrants further study. HESN women clearly respond differently to contraception on many different levels, quite possibly to complement epidemiological resistance to HIV infection.

9.8 Concluding Remarks and Future Directions

Based on the findings in this thesis there are several existing avenues of study to pursue. The observation of elevated Mx2 levels in the PBMC of HESN women and its link with Depo-Provera use is of most interest. The immunogenetics of HESN women regarding Mx2 and other upstream genes in the interferon pathway should be compared to shed light on the mechanism resulting in overall elevated levels of Mx2 in HESN women. In addition, experiments to evaluate the Mx2 levels in HESN women in response to Depo-Provera treatment in-vitro may also allow for some

insight into the epidemiological association we observed. Also, possible genetic polymorphisms in progesterone receptor genes of HESN women should be examined to determine their role in the altered expression of Mx2.

This thesis has highlighted the massive heterogeneity in the proteome of PBMC and CMC from all the women studied. It is clear that in order to evaluate epidemiological confounders on a large scale our initial pilot studies were significantly underpowered. Even with over one hundred subjects in each group, measuring a single protein marker, it was difficult to account for all the factors we were interested in. A much larger full scale proteomic analysis of PBMC and CMC from HESN women using a label-free mass spectrometry approach may shed light on the true proteomic profile of HESN women allowing us to create a proteomic signature of what immune quiescence should look like in the immune cells that are most often targeted by HIV during both the initial stages of acquisition (CMC) and those targeted systemically (PBMC).

One major question is how can we use this knowledge to our advantage in preventing HIV acquisition and informing HIV vaccine and microbicide design. Secretion or uptake of Mx2 by cells has yet to be shown, making it difficult to imagine a microbicide containing this protein would have any benefit with respect to reducing acquisition. However, gene therapy approaches are under consideration as HIV intervention strategies and with a better understanding of the possible genetic factors at play with respect to Mx2, therapies to permanently

enhance the expression or potency of Mx2 could lead to some novel strategies. In addition, Mx2 could also be used as a prognostic marker for assessing an immune quiescent state induced by drugs or other strategies. With this information in hand I believe the goal of an effective vaccine or microbicide for the prevention of HIV infection is within reach, putting a halt to the spread of HIV across the world.

Chapter 10. References

[1] WHO, U., UNAIDS report on the global AIDS epidemic 2013 2013, 1–272.

[2] Baral, S., Beyrer, C., Muessig, K., Poteat, T., et al., Burden of HIV among female sex workers in low-income and middle-income countries: a systematic review and meta-analysis. *The Lancet Infectious Diseases* 2012, 12, 538–549.

[3] Kerrigan, D., Wirtz, A., Semini, I., Stanciole, A., Butler, J., The Global HIV Epidemics among Sex Workers. *The World Bank* 2010.

[4] National AIDS and STD Control Programme (NASCOP), Sentinel Surveillance of HIV & STDs in Kenya REPORT 2006 2009, 1–39.

[5] National AIDS and STI Control Programme, Kenya AIDS Indicator Survey 2012 2013, 1–32.

[6] Gao, F., Bailes, E., Robertson, D.L., Chen, Y., Origin of HIV-1 in the chimpanzee : Pan troglodytes troglodytes : Abstract : Nature. *Nature* 1999.

[7] Sharp, P.M., Bailes, E., Chaudhuri, R.R., Rodenburg, C.M., et al., The origins of acquired immune deficiency syndrome viruses: where and when? *Philosophical Transactions of the Royal Society B: Biological Sciences* 2001, 356, 867–876.

[8] Worobey, M., Santiago, M.L., Keele, B.F., Ndjango, J.-B.N., et al., Origin of AIDS: contaminated polio vaccine theory refuted. *Nature* 2004, 428, 820.

[9] De Cock, K.M., Jaffe, H.W., Curran, J.W., Reflections on 30 Years of AIDS. *Emerging Infect. Dis.* 2011, 17, 1044–1048.

[10] Centers for Disease Control (CDC), Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *MMWR Morb. Mortal. Wkly. Rep.* 1981, 30, 305–308.

[11] Barré-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., et al., Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 1983, 220, 868–871.

[12] Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., et al., Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* 1984, 224, 500–503.

[13] Kreiss, J.K., Koech, D., Plummer, F.A., Holmes, K.K., et al., AIDS virus infection in Nairobi prostitutes. Spread of the epidemic to East Africa. *N Engl J Med* 1986, 314, 414–418.

[14] Goto, T., Nakai, M., Ikuta, K., The life-cycle of human immunodeficiency virus type 1. *Micron* 1998, 29, 123–138.

[15] Requejo, H.I.Z., Worldwide molecular epidemiology of HIV. *Rev. Saúde Pública* 2006, 40, 331–345.

[16] Buonaguro, L., Tornesello, M.L., Buonaguro, F.M., Human Immunodeficiency Virus Type 1 Subtype Distribution in the Worldwide Epidemic: Pathogenetic and Therapeutic Implications. *J. Virol.* 2007, 81, 10209–10219.

[17] Khoja, S., Ojwang, P., Khan, S., Okinda, N., et al., Genetic analysis of HIV-1 subtypes in Nairobi, Kenya. *PLoS ONE* 2008, 3, e3191.

[18] Abecasis, A.B., Wensing, A.M.J., Paraskevis, D., Vercauteren, J., et al., HIV-1

subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. *Retrovirology* 2013, 10, 7.
[19] Ganser-Pornillos, B.K., Yeager, M., Sundquist, W.I., The structural biology of HIV assembly. *Current Opinion in Structural Biology* 2008, 18, 203–217.
[20] Frankel, A.D., Young, J.A., HIV-1: fifteen proteins and an RNA. *Annu. Rev. Biochem.* 1998, 67, 1–25.
[21] Liu, R., Paxton, W.A., Choe, S., Ceradini, D., et al., Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 1996, 86, 367–377.
[22] Samson, M., Libert, F., Doranz, B.J., Rucker, J., et al., Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 1996, 382, 722–725.
[23] Arhel, N., Revisiting HIV-1 uncoating. *Retrovirology* 2010.
[24] Mansky, L.M., Temin, H.M., Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.* 1995, 69, 5087–5094.
[25] Greene, W.C., Peterlin, B.M., Charting HIV's remarkable voyage through the cell: Basic science as a passport to future therapy. *Nat Med* 2002, 8, 673–680.
[26] Simon, V., Ho, D.D., Karim, Q.A., HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. *The Lancet* 2006.
[27] Wawer, M.J., Gray, R.H., Sewankambo, N.K., Serwadda, D., et al., Rates of HIV-1 Transmission per Coital Act, by Stage of HIV-1 Infection, in Rakai, Uganda. *J INFECT DIS* 2005, 191, 1403–1409.
[28] Hollingsworth, T.D., Anderson, R.M., Fraser, C., HIV-1 Transmission, by Stage of Infection. *J INFECT DIS* 2008, 198, 687–693.
[29] Cohen, M.S., McCauley, M., Gamble, T.R., HIV treatment as prevention and HPTN 052. *Curr Opin HIV AIDS* 2012, 7, 99–105.
[30] Hankins, C.A., Dybul, M.R., The promise of pre-exposure prophylaxis with antiretroviral drugs to prevent HIV transmission: a review. *Curr Opin HIV AIDS* 2013, 8, 50–58.
[31] ROWLAND-JONES, S.L., Timeline: AIDS pathogenesis: what have two decades of HIV research taught us? *Nature Reviews Immunology* 2003, 3, 343–348.
[32] Picker, L.J., Watkins, D.I., HIV pathogenesis: the first cut is the deepest. *Nat. Immunol.* 2005.
[33] Hel, Z., McGhee, J.R., Mestecky, J., HIV infection: first battle decides the war. *Trends in Immunology* 2006, 27, 274–281.
[34] Stevenson, M., HIV-1 pathogenesis. *Nat Med* 2003, 9, 853–860.
[35] Li, Q., Duan, L., Estes, J.D., Ma, Z.-M., et al., Peak SIV replication in resting memory CD4+ T cells depletes gut lamina propria CD4+ T cells. *Nature* 2005, 434, 1148–1152.
[36] Mattapallil, J.J., Douek, D.C., Hill, B., Nishimura, Y., et al., Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection. *Nature* 2005, 434, 1093–1097.

[37] Mehandru, S., Poles, M.A., Tenner-Racz, K., Horowitz, A., et al., Primary HIV-1 infection is associated with preferential depletion of CD4+ T lymphocytes from effector sites in the gastrointestinal tract. *J. Exp. Med.* 2004, 200, 761–770.

[38] Brenchley, J.M., Schacker, T.W., Ruff, L.E., Price, D.A., et al., CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *J. Exp. Med.* 2004, 200, 749–759.

[39] Crowe, S.M., Carlin, J.B., Stewart, K.I., Lucas, C.R., Hoy, J.F., Predictive value of CD4 lymphocyte numbers for the development of opportunistic infections and malignancies in HIV-infected persons. *J. Acquir. Immune Defic. Syndr.* 1991, 4, 770–776.

[40] Hoover, D.R., Rinaldo, C., He, Y., Phair, J., et al., Long-term survival without clinical AIDS after CD4+ cell counts fall below 200 x 10(6)/l. *AIDS* 1995, 9, 145–152.

[41] Gurdasani, D., Iles, L., Dillon, D.G., Young, E.H., et al., A systematic review of definitions of extreme phenotypes of HIV control and progression. *AIDS* 2013.

[42] Deeks, S.G., Walker, B.D., Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. *Immunity* 2007, 27, 406–416.

[43] Group, R.H.V.S., Placebo-Controlled Phase 3 Trial of a Recombinant Glycoprotein 120 Vaccine to Prevent HIV-1 Infection. *Journal of Infectious …* 2005.

[44] Pitisuttithum, P., Gilbert, P., Gurwith, M., Heyward, W., et al., Randomized, Double-Blind, Placebo-Controlled Efficacy Trial of a Bivalent Recombinant Glycoprotein 120 HIV-1 Vaccine among Injection Drug Users in Bangkok, Thailand. *J INFECT DIS* 2006, 194, 1661–1671.

[45] Buchbinder, S.P., Mehrotra, D.V., Duerr, A., Fitzgerald, D.W., Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. *The Lancet* 2008.

[46] Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., Kaewkungwal, J., et al., Vaccination with ALVAC and AIDSVAX to Prevent HIV-1 Infection in Thailand. *N Engl J Med* 2009, 361, 2209–2220.

[47] Haynes, B.F., Gilbert, P.B., McElrath, M.J., Zolla-Pazner, S., et al., Immune-correlates analysis of an HIV-1 vaccine efficacy trial. *N Engl J Med* 2012, 366, 1275–1286.

[48] Wu, X., Yang, Z.-Y., Li, Y., Hogerkorp, C.-M., et al., Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* 2010, 329, 856–861.

[49] Zhou, T., Georgiev, I., Wu, X., Yang, Z.-Y., et al., Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* 2010, 329, 811–817.

[50] Walker, L.M., Phogat, S.K., Chan-Hui, P.Y., Wagner, D., et al., Broad and Potent Neutralizing Antibodies from an African Donor Reveal a New HIV-1 Vaccine Target. *Science* 2009, 326, 285–289.

[51] Johnson, P.R., Schnepp, B.C., Zhang, J., Connell, M.J., et al., Vector-mediated gene transfer engenders long-lived neutralizing activity and protection against SIV infection in monkeys. *Nat Med* 2009, 15, 901–906.
[52] Balazs, A.B., Chen, J., Hong, C.M., Rao, D.S., et al., Antibody-based protection against HIV infection by vectored immunoprophylaxis. *Nature* 2012, 481, 81–84.
[53] Abdool Karim, Q., Abdool Karim, S.S., Frohlich, J.A., Grobler, A.C., et al., Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. *Science* 2010, 329, 1168–1174.
[54] Szabo, R., How does male circumcision protect against HIV infection? *BMJ* 2000, 320, 1592–1594.
[55] McCoombe, S.G., Short, R.V., Potential HIV-1 target cells in the human penis. *AIDS* 2006, 20, 1491–1495.
[56] Auvert, B., Taljaard, D., Lagarde, E., Sobngwi-Tambekou, J., et al., Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 Trial. *PLoS Med.* 2005, 2, e298.
[57] Bailey, R.C., Moses, S., Parker, C.B., Agot, K., et al., Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. *Lancet* 2007, 369, 643–656.
[58] Gray, R.H., Kigozi, G., Serwadda, D., Makumbi, F., et al., Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. *Lancet* 2007, 369, 657–666.
[59] Delves, P.J., Roitt, I.M., The immune system. First of two parts. *N Engl J Med* 2000, 343, 37–49.
[60] Medzhitov, R., TOLL-LIKE RECEPTORS AND INNATE IMMUNITY. *Nature Reviews Immunology* 2001, 1, 135–145.
[61] Rosen, F.S., MEDZHITOV, R., Janeway, C., Jr, Innate Immunity — NEJM. *New England Journal …* 2000.
[62] Abbas, A.K., Lichtman, A.H., Pober, J.S., Cellular and molecular immunology 1991.
[63] Nikolich-Žugich, J., Slifka, M.K., Messaoudi, I., The many important facets of T-cell repertoire diversity. *Nature Reviews Immunology* 2004, 4, 123–132.
[64] Owen, J., Punt, J., Stranford, S., Stranford, S.A., Kuby, J., *Kuby Immunology*, W H Freeman & Company, 2013.
[65] Delves, P.J., Roitt, I.M., The immune system. Second of two parts. *N Engl J Med* 2000, 343, 108–117.
[66] Horton, R.E., McLaren, P.J., Fowke, K., Kimani, J., Ball, T.B., Cohorts for the Study of HIV-1–Exposed but Uninfected Individuals: Benefits and Limitations. *J INFECT DIS* 2010, 202, S377–S381.
[67] Martinson, J.J., Chapman, N.H., Rees, D.C., Liu, Y.T., Clegg, J.B., Global distribution of the CCR5 gene 32-basepair deletion. *Nat. Genet.* 1997, 16, 100–103.
[68] Fowke, K.R., Nagelkerke, N.J., Kimani, J., Simonsen, J.N., et al., Resistance to HIV-1 infection among persistently seronegative prostitutes in Nairobi, Kenya. *The Lancet* 1996, 348, 1347–1351.

[69] Kaul, R., Trabattoni, D., Bwayo, J.J., Arienti, D., et al., HIV-1-specific mucosal IgA in a cohort of HIV-1-resistant Kenyan sex workers. *AIDS* 1999, 13, 23.
[70] Horton, R.E., Ball, T.B., Wachichi, C., Jaoko, W., et al., Cervical HIV-Specific IgA in a Population of Commercial Sex Workers Correlates with Repeated Exposure But Not Resistance to HIV. *AIDS Research and Human Retroviruses* 2009, 25, 83–92.
[71] Dorrell, L., Hessell, A.J., Wang, M., Whittle, H., et al., Absence of specific mucosal antibody responses in HIV-exposed uninfected sex workers from the Gambia. *AIDS* 2000, 14, 1117–1122.
[72] Devito, C., Hinkula, J., Kaul, R., Kimani, J., et al., Cross-clade HIV-1-specific neutralizing IgA in mucosal and systemic compartments of HIV-1-exposed, persistently seronegative subjects. *J. Acquir. Immune Defic. Syndr.* 2002, 30, 413–420.
[73] Devito, C., Hinkula, J., Kaul, R., Lopalco, L., et al., Mucosal and plasma IgA from HIV-exposed seronegative individuals neutralize a primary HIV-1 isolate. *AIDS* 2000, 14, 1917–1920.
[74] Devito, C., Broliden, K., Kaul, R., Svensson, L., et al., Mucosal and plasma IgA from HIV-1-exposed uninfected individuals inhibit HIV-1 transcytosis across human epithelial cells. *J. Immunol.* 2000, 165, 5170–5176.
[75] Fowke, K.R., Kaul, R., Rosenthal, K.L., OYUGI, J., et al., HIV-1-specific cellular immune responses among HIV-1-resistant sex workers. *Immunol Cell Biol* 2000, 78, 586–595.
[76] Alimonti, J.B., Kimani, J., MATU, L., Wachihi, C., et al., Characterization of CD8+ T-cell responses in HIV-1-exposed seronegative commercial sex workers from Nairobi, Kenya. *Immunol Cell Biol* 2006, 84, 482–485.
[77] Kaul, R., ROWLAND-JONES, S.L., Kimani, J., Fowke, K., et al., New insights into HIV-1 specific cytotoxic T-lymphocyte responses in exposed, persistently seronegative Kenyan sex workers. *Immunology Letters* 2001, 79, 3–13.
[78] Kaul, R., Plummer, F.A., Kimani, J., Dong, T., 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.
[79] Kaul, R., Rowland-Jones, S.L., Kimani, J., Dong, T., et al., Late seroconversion in HIV-resistant Nairobi prostitutes despite pre-existing HIV-specific CD8+ responses. *J. Clin. Invest.* 2001, 107, 341–349.
[80] Kaul, R., MacDonald, K.S., Nagelkerke, N.J., Kimani, J., et al., HIV viral set point and host immune control in individuals with HIV-specific CD8+ T-cell responses prior to HIV acquisition. *AIDS* 2010, 24, 1449–1454.
[81] Alimonti, J.B., Koesters, S.A., Kimani, J., MATU, L., et al., CD4+ T cell responses in HIV-exposed seronegative women are qualitatively distinct from those in HIV-infected women. *J INFECT DIS* 2005, 191, 20–24.
[82] Hirbod, T., Kaul, R., Reichard, C., Kimani, J., et al., HIV-neutralizing immunoglobulin A and HIV-specific proliferation are independently associated with reduced HIV acquisition in Kenyan sex workers. *AIDS* 2008, 22, 727–735.
[83] Clerici, M., Giorgi, J.V., Chou, C.C., Gudeman, V.K., et al., Cell-mediated

immune response to human immunodeficiency virus (HIV) type 1 in seronegative homosexual men with recent sexual exposure to HIV-1. *J INFECT DIS* 1992, 165, 1012–1019.

[84] Erickson, A.L., Willberg, C.B., McMahan, V., Liu, A., et al., Potentially exposed but uninfected individuals produce cytotoxic and polyfunctional human immunodeficiency virus type 1-specific CD8(+) T-cell responses which can be defined to the epitope level. *Clin. Vaccine Immunol.* 2008, 15, 1745–1748.

[85] Makedonas, G., Bruneau, J., Alary, M., Tsoukas, C.M., et al., Comparison of HIV-specific CD8 T-cell responses among uninfected individuals exposed to HIV parenterally and mucosally. *AIDS* 2005, 19, 251–259.

[86] Bienzle, D., MacDonald, K.S., Smaill, F.M., Kovacs, C., et al., Factors contributing to the lack of human immunodeficiency virus type 1 (HIV-1) transmission in HIV-1-discordant partners. *J INFECT DIS* 2000, 182, 123–132.

[87] Richmond, M., McKinnon, L.R., Kiazyk, S.A.K., Wachihi, C., et al., Epitope mapping of HIV-specific CD8+ T cell responses by multiple immunological readouts reveals distinct specificities defined by function. *J. Virol.* 2011, 85, 1275–1286.

[88] Lisziewicz, J., Rosenberg, E., Lieberman, J., Jessen, H., et al., Control of HIV despite the discontinuation of antiretroviral therapy. *N Engl J Med* 1999, 340, 1683–1684.

[89] Hütter, G., Nowak, D., Mossner, M., Ganepola, S., et al., Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med* 2009, 360, 692–698.

[90] Dorr, P., Westby, M., Dobbs, S., Griffin, P., et al., Maraviroc (UK-427,857), a Potent, Orally Bioavailable, and Selective Small-Molecule Inhibitor of Chemokine Receptor CCR5 with Broad-Spectrum Anti-Human Immunodeficiency Virus Type 1 Activity. *Antimicrobial Agents and Chemotherapy* 2005, 49, 4721–4732.

[91] Hardie, R.-A., Luo, M., Bruneau, B., Knight, E., et al., Human leukocyte antigen-DQ alleles and haplotypes and their associations with resistance and susceptibility to HIV-1 infection. *AIDS* 2008, 22, 807–816.

[92] Hardie, R.-A., Knight, E., Bruneau, B., Semeniuk, C., 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.

[93] MacDonald, K.S., Fowke, K.R., Kimani, J., Dunand, V.A., et al., Influence of HLA supertypes on susceptibility and resistance to human immunodeficiency virus type 1 infection. *J INFECT DIS* 2000, 181, 1581–1589.

[94] Jennes, W., Verheyden, S., Demanet, C., Adjé-Touré, C.A., et al., Cutting edge: resistance to HIV-1 infection among African female sex workers is associated with inhibitory KIR in the absence of their HLA ligands. *J. Immunol.* 2006, 177, 6588–6592.

[95] Boulet, S., Sharafi, S., Simic, N., Bruneau, J., et al., Increased proportion of KIR3DS1 homozygotes in HIV-exposed uninfected individuals. *AIDS* 2008,

22, 595–599.
[96] Tomescu, C., Duh, F.-M., Lanier, M.A., Kapalko, A., et al., Increased plasmacytoid dendritic cell maturation and natural killer cell activation in HIV-1 exposed, uninfected intravenous drug users. *AIDS* 2010, 24, 2151–2160.
[97] Ball, T.B., Ji, H., Kimani, J., Mclaren, P., et al., Polymorphisms in IRF-1 associated with resistance to HIV-1 infection in highly exposed uninfected Kenyan sex workers. *AIDS* 2007, 21, 1091–1101.
[98] Ji, H., Ball, T.B., Ao, Z., Kimani, J., et al., Reduced HIV-1 long terminal repeat transcription in subjects with protective interferon regulatory factor-1 genotype: a potential mechanism mediating resistance to infection by HIV-1. *Scand. J. Infect. Dis.* 2010, 42, 389–394.
[99] Card, C.M., Ball, T.B., Fowke, K.R., Immune Quiescence: a model of protection against HIV infection. *Retrovirology* 2013, 10, 141.
[100] Biasin, M., Piacentini, L., Caputo, Lo, S., Naddeo, V., et al., TLR activation pathways in HIV-1-exposed seronegative individuals. *The Journal of Immunology* 2010, 184, 2710–2717.
[101] Yao, X.-D., Omange, R.W., Henrick, B.M., Lester, R.T., et al., Acting locally: innate mucosal immunity in resistance to HIV-1 infection in Kenyan commercial sex workers. *Mucosal Immunology* 2013.
[102] Suresh, P., Wanchu, A., Bhatnagar, A., Sachdeva, R.K., Sharma, M., Spontaneous and antigen-induced chemokine production in exposed but uninfected partners of HIV type 1-infected individuals in North India. *AIDS Research and Human Retroviruses* 2007, 23, 261–268.
[103] Furci, L., Scarlatti, G., Burastero, S., Tambussi, G., 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. *Journal of Experimental Medicine* 1997, 186, 455–460.
[104] Tomescu, C., Abdulhaqq, S., Montaner, L.J., Evidence for the innate immune response as a correlate of protection in human immunodeficiency virus (HIV)-1 highly exposed seronegative subjects (HESN). *Clinical & Experimental Immunology* 2011, 164, 158–169.
[105] Iqbal, S.M., Ball, T.B., Kimani, J., Kiama, P., et al., 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.
[106] Fowke, K.R., DONG, T., ROWLAND-JONES, S.L., OYUGI, J., et al., HIV type 1 resistance in Kenyan sex workers is not associated with altered cellular susceptibility to HIV type 1 infection or enhanced β-chemokine production. *AIDS Research and Human Retroviruses* 1998, 14, 1521–1530.
[107] Missé, D., Yssel, H., Trabattoni, D., Oblet, C., et al., IL-22 participates in an innate anti-HIV-1 host-resistance network through acute-phase protein induction. *J. Immunol.* 2007, 178, 407–415.
[108] Iqbal, S.M., Ball, T.B., Levinson, P., Maranan, L., et al., Elevated elafin/trappin-2 in the female genital tract is associated with protection against HIV acquisition. *AIDS* 2009, 23, 1669–1677.
[109] Drannik, A.G., Nag, K., Yao, X.-D., Henrick, B.M., et al., Anti-HIV-1 activity of

elafin is more potent than its precursor's, trappin-2, in genital epithelial cells. *J. Virol.* 2012, 86, 4599–4610.
[110] Drannik, A.G., Nag, K., Yao, X.-D., Henrick, B.M., et al., Anti-HIV-1 activity of elafin depends on its nuclear localization and altered innate immune activation in female genital epithelial cells. *PLoS ONE* 2012, 7, e52738.
[111] Burgener, A., Boutilier, J., Wachihi, C., Kimani, J., et al., Identification of Differentially Expressed Proteins in the Cervical Mucosa of HIV-1-Resistant Sex Workers. *J. Proteome Res.* 2008, 7, 4446–4454.
[112] Burgener, A., Rahman, S., Ahmad, R., Lajoie, J., et al., Comprehensive proteomic study identifies serpin and cystatin antiproteases as novel correlates of HIV-1 resistance in the cervicovaginal mucosa of female sex workers. *J. Proteome Res.* 2011, 10, 5139–5149.
[113] Strebel, K., Luban, J., Jeang, K.-T., Human cellular restriction factors that target HIV-1 replication. *BMC Med* 2009, 7, 48.
[114] Malim, M.H., Bieniasz, P.D., HIV Restriction Factors and Mechanisms of Evasion. *Cold Spring Harb Perspect Med* 2012, 2, a006940.
[115] Biasin, M., Piacentini, L., Caputo, Lo, S., Kanari, Y., 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.
[116] Price, H., Lacap, P., Tuff, J., Wachihi, C., et al., A TRIM5α exon 2 polymorphism is associated with protection from HIV-1 infection in the Pumwani sex worker cohort. *AIDS* 2010, 24, 1813–1821.
[117] Neil, S.J.D., Zang, T., Bieniasz, P.D., Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 2008, 451, 425–430.
[118] Laguette, N., Sobhian, B., Casartelli, N., Ringeard, M., et al., SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* 2011, 474, 654–657.
[119] Liu, Z., Pan, Q., Ding, S., Qian, J., et al., The Interferon-Inducible MxB Protein Inhibits HIV-1 Infection. *Cell Host Microbe* 2013, 14, 398–410.
[120] Kane, M., Yadav, S.S., Bitzegeio, J., Kutluay, S.B., et al., MX2 is an interferon-induced inhibitor of HIV-1 infection. *Nature* 2013.
[121] Goujon, C., Moncorgé, O., Bauby, H., Doyle, T., et al., Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature* 2013.
[122] Mous, K., Jennes, W., Camara, M., Seydi, M., et al., Expression analysis of LEDGF/p75, APOBEC3G, TRIM5alpha, and tetherin in a Senegalese cohort of HIV-1-exposed seronegative individuals. *PLoS ONE* 2012, 7, e33934.
[123] McLaren, P.J., Ball, T.B., Wachihi, C., Jaoko, W., et al., HIV-exposed seronegative commercial sex workers show a quiescent phenotype in the CD4+ T cell compartment and reduced expression of HIV-dependent host factors. *J INFECT DIS* 2010, 202 Suppl 3, S339–44.
[124] Songok, E.M., Luo, M., Liang, B., Mclaren, P., et al., Microarray analysis of HIV resistant female sex workers reveal a gene expression signature pattern reminiscent of a lowered immune activation state. *PLoS ONE* 2012, 7, e30048.
[125] Card, C.M., McLaren, P.J., Wachihi, C., Kimani, J., et al., 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.

[126] Vogel, C., Marcotte, E.M., Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* 2012, 13, 227–232.

[127] Nagaraj, N., Wiśniewski, J.R., Geiger, T., Cox, J., et al., Deep proteome and transcriptome mapping of a human cancer cell line. *Molecular Systems Biology* 2011, 7, 548.

[128] Ghazalpour, A., Bennett, B., Petyuk, V.A., Orozco, L., et al., Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genet.* 2011, 7, e1001393.

[129] Miranda, H.C., Herai, R.H., Thomé, C.H., Gomes, G.G., et al., A quantitative proteomic and transcriptomic comparison of human mesenchymal stem cells from bone marrow and umbilical cord vein. *Proteomics* 2012, 12, 2607–2617.

[130] Gygi, S.P., Rochon, Y., Franza, B.R., Aebersold, R., Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* 1999, 19, 1720–1730.

[131] Greenbaum, D., Colangelo, C., Williams, K., Gerstein, M., Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol.* 2003, 4, 117.

[132] Kelstrup, C.D., Young, C., Lavallee, R., Lund Nielsen, M., Olsen, J.V., Optimized Fast and Sensitive Acquisition Methods for Shotgun Proteomics on a Quadrupole Orbitrap Mass Spectrometer. *J. Proteome Res.* 2012.

[133] Unlü, M., Morgan, M.E., Minden, J.S., Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 1997, 18, 2071–2077.

[134] Minden, J.S., DIGE: past and future. *Methods Mol. Biol.* 2012, 854, 3–8.

[135] Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., et al., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell Proteomics* 2002, 1, 376–386.

[136] Zhu, W., Smith, J.W., Huang, C.-M., Mass Spectrometry-Based Label-Free Quantitative Proteomics. *J. Biomed. Biotech.* 2010, 2010, 1–7.

[137] Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., et al., Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 1999, 17, 994–999.

[138] Ross, P.L., Multiplexed Protein Quantitation in Saccharomyces cerevisiae Using Amine-reactive Isobaric Tagging Reagents. *Molecular & Cellular Proteomics* 2004, 3, 1154–1169.

[139] Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., et al., Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. *Anal. Chem.* 2003, 75, 1895–1904.

[140] Cho, W., Proteomics Technologies and Challenges. *Genomics* 2007.

[141] Ong, S.E., Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as

a Simple and Accurate Approach to Expression Proteomics. *Molecular & Cellular Proteomics* 2002, 1, 376–386.
[142] Ong, S.-E., The expanding field of SILAC. *Anal Bioanal Chem* 2012, 404, 967–976.
[143] Capelo, J.L., Carreira, R.J., Fernandes, L., Lodeiro, C., et al., Latest developments in sample treatment for 18O-isotopic labeling for proteomics mass spectrometry-based approaches: a critical review. *Talanta* 2010, 80, 1476–1486.
[144] Christoforou, A.L., Lilley, K.S., Isobaric tagging approaches in quantitative proteomics: the ups and downs. *Anal Bioanal Chem* 2012, 404, 1029–1037.
[145] Jaffe, J.D., Keshishian, H., Chang, B., Addona, T.A., et al., Accurate inclusion mass screening: a bridge from unbiased discovery to targeted assay development for biomarker verification. *Molecular & Cellular Proteomics* 2008, 7, 1952–1962.
[146] Pan, S., Aebersold, R., Chen, R., Rush, J., et al., Mass Spectrometry Based Targeted Protein Quantification: Methods and Applications. *J. Proteome Res.* 2009, 8, 787–797.
[147] Douglas, J.L., Viswanathan, K., McCarroll, M.N., Gustin, J.K., et al., Vpu directs the degradation of the human immunodeficiency virus restriction factor BST-2/Tetherin via a {beta}TrCP-dependent mechanism. *J. Virol.* 2009, 83, 7931–7947.
[148] Gautier, V.W., Gu, L., O'Donoghue, N., Pennington, S., et al., In vitro nuclear interactome of the HIV-1 Tat protein. *Retrovirology* 2009, 6, 47.
[149] He, F., Zeng, Y., Wu, X., Ji, Y., et al., Endogenous HIV-1 Vpr-mediated apoptosis and proteome alteration of human T-cell leukemia virus-1 transformed C8166 cells. *Apoptosis* 2009, 14, 1212–1226.
[150] Naji, S., Ambrus, G., Cimermancic, P., Reyes, J.R., et al., Host cell interactome of HIV-1 Rev includes RNA helicases involved in multiple facets of virus production. *Mol. Cell Proteomics* 2012, 11, M111.015313.
[151] Jäger, S., Cimermancic, P., Gulbahce, N., Johnson, J.R., et al., Global landscape of HIV-human protein complexes. *Nature* 2012, 481, 365–370.
[152] McMichael, A.J., Borrow, P., Tomaras, G.D., Goonetilleke, N., Haynes, B.F., The immune response during acute HIV-1 infection: clues for vaccine development. *Nature Reviews Immunology* 2009, 10, 11–23.
[153] Kramer, H.B., Lavender, K.J., Qin, L., Stacey, A.R., et al., Elevation of intact and proteolytic fragments of acute phase proteins constitutes the earliest systemic antiviral response in HIV-1 infection. *PLoS Pathogens* 2010, 6, e1000893.
[154] Pottiez, G., Jagadish, T., Yu, F., Letendre, S., et al., Plasma proteomic profiling in HIV-1 infected methamphetamine abusers. *PLoS ONE* 2012, 7, e31031.
[155] Rozek, W., Ricardo-Dukelow, M., Holloway, S., Cerebrospinal Fluid Proteomic Profiling of HIV-1-Infected Patients with Cognitive Impairment - Journal of Proteome Research (ACS Publications). *… of proteome research* 2007.
[156] Rozek, W., Horning, J., Anderson, J., Ciborowski, P., Sera proteomic

biomarker profiling in HIV-1 infected subjects with cognitive impairment. *Proteomics Clin Appl* 2008, 2, 1498–1507.
[157] Wiederin, J., Rozek, W., Duan, F., Ciborowski, P., Biomarkers of HIV-1 associated dementia: proteomic investigation of sera. *Proteome Sci* 2009, 7, 8.
[158] Laspiur, J.P., Anderson, E.R., Ciborowski, P., Wojna, V., et al., CSF proteomic fingerprints for HIV-associated cognitive impairment. *J. Neuroimmunol.* 2007, 192, 157–170.
[159] Birse, K.M., Burgener, A., Westmacott, G.R., McCorrister, S., et al., Unbiased proteomics analysis demonstrates significant variability in mucosal immune factor expression depending on the site and method of collection. *PLoS ONE* 2013, 8, e79505.
[160] Sheng, W.-Y., Wang, T.-C.V., Proteomic analysis of the differential protein expression reveals nuclear GAPDH in activated T lymphocytes. *PLoS ONE* 2009, 4, e6322.
[161] Haudek-Prinz, V.J., Klepeisz, P., Slany, A., Griss, J., et al., Proteome signatures of inflammatory activated primary human peripheral blood mononuclear cells. *Journal of Proteomics* 2012, 76, 150–162.
[162] Groessl, M., Luksch, H., Rösen-Wolff, A., Shevchenko, A., Gentzel, M., Profiling of the human monocytic cell secretome by quantitative label-free mass spectrometry identifies stimulus-specific cytokines and proinflammatory proteins. *Proteomics* 2012.
[163] Burgener, A., Mogk, K., Westmacott, G., Plummer, F., et al., Salivary basic proline-rich proteins are elevated in HIV-exposed seronegative men who have sex with men. *AIDS* 2012, 1.
[164] Krokhin, O.V., Spicer, V., Peptide retention standards and hydrophobicity indexes in reversed-phase high-performance liquid chromatography of peptides. *Anal. Chem.* 2009, 81, 9522–9530.
[165] Wiśniewski, J.R., Zougman, A., Nagaraj, N., Mann, M., Universal sample preparation method for proteome analysis. *Nat Meth* 2009, 6, 359–362.
[166] Stein, D.R., Hu, X., McCorrister, S.J., Westmacott, G.R., et al., High pH Reversed-phase Chromatography as a Superior Fractionation Scheme Compared to Off-Gel Isoelectric Focusing for Complex Proteome Analysis. *Proteomics* 2013.
[167] Reed, L.J., Muench, H., A SIMPLE METHOD OF ESTIMATING FIFTY PER CENT ENDPOINTS, 1938.
[168] Spicer, V., Yamchuk, A., Cortens, J., Sousa, S., et al., Sequence-Specific Retention Calculator. A Family of Peptide Retention Time Prediction Algorithms in Reversed-Phase HPLC: Applicability to Various Chromatographic Conditions and Columns. *Anal. Chem.* 2007, 79, 8762–8768.
[169] Slebos, R.J.C., Brock, J.W.C., Winters, N.F., Stuart, S.R., et al., Evaluation of Strong Cation Exchange versus Isoelectric Focusing of Peptides for Multidimensional Liquid Chromatography-Tandem Mass Spectrometry. *J. Proteome Res.* 2008, 7, 5286–5294.
[170] Essader, A.S., Cargile, B.J., Bundy, J.L., Stephenson, J.L., A comparison of

immobilized pH gradient isoelectric focusing and strong-cation-exchange chromatography as a first dimension in shotgun proteomics. *Proteomics* 2005, 5, 24–34.

[171] Ow, S.Y., Salim, M., Noirel, J., Evans, C., et al., iTRAQ Underestimation in Simple and Complex Mixtures: "The Good, the Bad and the Ugly." *J. Proteome Res.* 2009, 8, 5347–5355.

[172] Stein, D.R., Burgener, A., Ball, T.B., Proteomics as a novel HIV immune monitoring tool. *Curr Opin HIV AIDS* 2013, 8, 140–146.

[173] Choe, L., D'Ascenzo, M., Relkin, N.R., Pappin, D., et al., 8-Plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease. *Proteomics* 2007, 7, 3651–3660.

[174] Pierce, A., Unwin, R., Evans, C., Griffiths, S., Eight-channel iTRAQ enables comparison of the activity of six leukemogenic tyrosine kinases. *Molecular & Cellular …* 2008.

[175] Gaucher, D., Therrien, R., Kettaf, N., Angermann, B.R., et al., Yellow fever vaccine induces integrated multilineage and polyfunctional immune responses. *J. Exp. Med.* 2008, 205, 3119–3131.

[176] Card, C.M., Rutherford, W.J., Ramdahin, S., Yao, X., et al., Reduced Cellular Susceptibility to In Vitro HIV Infection Is Associated with CD4+ T Cell Quiescence. *PLoS ONE* 2012, 7, e45911.

[177] Hammarström, S., Hammarström, M.L., Sundblad, G., Arnarp, J., Lönngren, J., Mitogenic leukoagglutinin from Phaseolus vulgaris binds to a pentasaccharide unit in N-acetyllactosamine-type glycoprotein glycans. *Proc. Natl. Acad. Sci. U.S.A.* 1982, 79, 1611–1615.

[178] Koup, R.A., Ho, D.D., Poli, G., Fauci, A.S., Isolation and quantitation of HIV in peripheral blood. *Curr Protoc Immunol* 2001, Chapter 12, Unit 12.2.

[179] Verhelst, J., Hulpiau, P., Saelens, X., Mx proteins: antiviral gatekeepers that restrain the uninvited. *Microbiol. Mol. Biol. Rev.* 2013, 77, 551–566.

[180] Ku, C.-C., Che, X.-B., Reichelt, M., Rajamani, J., et al., Herpes simplex virus-1 induces expression of a novel MxA isoform that enhances viral replication. *Immunol Cell Biol* 2011, 89, 173–182.

[181] Liu, Z., Pan, Q., Ding, S., Qian, J., et al., The interferon-induced MxB protein inhibits an early step of HIV-1 infection. *Retrovirology* 2013, 10, P48.

[182] Stringer, E., Antonsen, E., Hormonal contraception and HIV disease progression. *CLIN INFECT DIS* 2008, 47, 945–951.

[183] Huijbregts, R.P.H., Helton, E.S., Michel, K.G., Sabbaj, S., et al., Hormonal contraception and HIV-1 infection: medroxyprogesterone acetate suppresses innate and adaptive immune mechanisms. *Endocrinology* 2013, 154, 1282–1295.

[184] Morrison, C., Fichorova, R., Mauck, C., Chen, P., et al., Cervical Inflammation and Immunity Associated with Hormonal Contraception, Pregnancy and HIV-1 Seroconversion. *J. Acquir. Immune Defic. Syndr.* 2014, 1.

[185] Harman, A.N., Lai, J., Turville, S., Samarajiwa, S., et al., HIV infection of dendritic cells subverts the IFN induction pathway via IRF-1 and inhibits type 1 IFN production. *Blood* 2011, 118, 298–308.

[186] Eriksson, H., Lengqvist, J., Hedlund, J., Uhlén, K., et al., Quantitative membrane proteomics applying narrow range peptide isoelectric focusing for studies of small cell lung cancer resistance mechanisms. *Proteomics* 2008, 8, 3008–3018.
[187] Horth, P., Miller, C.A., Preckel, T., Wenz, C., Efficient Fractionation and Improved Protein Identification by Peptide OFFGEL Electrophoresis. *Molecular & Cellular Proteomics* 2006, 5, 1968–1974.
[188] Hubner, N.C., Ren, S., Mann, M., Peptide separation with immobilized pI strips is an attractive alternative to in-gel protein digestion for proteome analysis. *Proteomics* 2008, 8, 4862–4872.
[189] Elschenbroich, S., Ignatchenko, V., Sharma, P., Schmitt-Ulms, G., et al., Peptide Separations by On-Line MudPIT Compared to Isoelectric Focusing in an Off-Gel Format: Application to a Membrane-Enriched Fraction from C2C12 Mouse Skeletal Muscle Cells. *J. Proteome Res.* 2009, 8, 4860–4869.
[190] Dwivedi, R.C., Spicer, V., Harder, M., Antonovici, M., et al., Practical implementation of 2D HPLC scheme with accurate peptide retention prediction in both dimensions for high-throughput bottom-up proteomics. *Anal. Chem.* 2008, 80, 7036–7042.
[191] Tenorio-Laranga, J., Valero, M.L., Männistö, P.T., Sánchez del Pino, M., García-Horsman, J.A., Combination of snap freezing, differential pH two-dimensional reverse-phase high-performance liquid chromatography, and iTRAQ technology for the peptidomic analysis of the effect of prolyl oligopeptidase inhibition in the rat brain. *Anal. Biochem.* 2009, 393, 80–87.
[192] Yang, Y., Qiang, X., Owsiany, K., Zhang, S., et al., Evaluation of different multidimensional LC-MS/MS pipelines for isobaric tags for relative and absolute quantitation (iTRAQ)-based proteomic analysis of potato tubers in response to cold storage. *J. Proteome Res.* 2011, 10, 4647–4660.
[193] Cao, Z., Tang, H.-Y., Wang, H., Liu, Q., Speicher, D.W., Systematic Comparison of Fractionation Methods for In-depth Analysis of Plasma Proteomes. *J. Proteome Res.* 2012.
[194] O'Brien, D.P., Timms, J.F., Employing TMT Quantification in a Shotgun-MS Platform. *Methods Mol. Biol.* 2014, 1156, 187–199.
[195] Pichler, P., Köcher, T., Holzmann, J., Mazanek, M., et al., Peptide Labeling with Isobaric Tags Yields Higher Identification Rates Using iTRAQ 4-Plex Compared to TMT 6-Plex and iTRAQ 8-Plex on LTQ Orbitrap. *Anal. Chem.* 2010, 82, 6549–6558.
[196] Ting, L., Rad, R., Gygi, S.P., Haas, W., MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat Meth* 2011, 8, 937–940.
[197] Köcher, T., Pichler, P., Schutzbier, M., Stingl, C., et al., High Precision Quantitative Proteomics Using iTRAQ on an LTQ Orbitrap: A New Mass Spectrometric Method Combining the Benefits of All. *J. Proteome Res.* 2009, 8, 4743–4752.
[198] Pichler, P., Köcher, T., Holzmann, J., Möhring, T., et al., Improved precision of iTRAQ and TMT quantification by an axial extraction field in an Orbitrap HCD cell. *Anal. Chem.* 2011, 83, 1469–1474.
[199] Chiva, C., Sabidó, E., HCD-only fragmentation method balances peptide

identification and quantitation of TMT-labeled samples in hybrid linear ion trap/orbitrap mass spectrometers. *Journal of Proteomics* 2014, 96, 263–270.

[200] Takamatsu, M., Hirata, A., Ohtaki, H., Hoshi, M., et al., IDO1 plays an immunosuppressive role in 2,4,6-trinitrobenzene sulfate-induced colitis in mice. *The Journal of Immunology* 2013, 191, 3057–3064.

[201] Boasso, A., Herbeuval, J.-P., Hardy, A.W., Anderson, S.A., et al., HIV inhibits CD4+ T-cell proliferation by inducing indoleamine 2,3-dioxygenase in plasmacytoid dendritic cells. *Blood* 2007, 109, 3351–3359.

[202] Campbell, G.R., Watkins, J.D., Singh, K.K., Loret, E.P., Spector, S.A., Human immunodeficiency virus type 1 subtype C Tat fails to induce intracellular calcium flux and induces reduced tumor necrosis factor production from monocytes. *J. Virol.* 2007, 81, 5919–5928.

[203] Samikkannu, T., Rao, K.V.K., Gandhi, N., Saxena, S.K., Nair, M.P.N., Human immunodeficiency virus type 1 clade B and C Tat differentially induce indoleamine 2,3-dioxygenase and serotonin in immature dendritic cells: Implications for neuroAIDS. *J. Neurovirol.* 2010, 16, 255–263.

[204] Garten, A., Petzold, S., Körner, A., Imai, S.-I., Kiess, W., Nampt: linking NAD biology, metabolism and cancer. *Trends Endocrinol. Metab.* 2009, 20, 130–138.

[205] Darnell, G.A., Schroder, W.A., Gardner, J., Harrich, D., et al., SerpinB2 is an inducible host factor involved in enhancing HIV-1 transcription and replication. *Journal of Biological Chemistry* 2006, 281, 31348–31358.

[206] NOWELL, P.C., Phytohemagglutinin: an initiator of mitosis in cultures of normal human leukocytes. *Cancer Res.* 1960, 20, 462–466.

[207] Li, Q., Estes, J.D., Schlievert, P.M., Duan, L., et al., Glycerol monolaurate prevents mucosal SIV transmission. *Nature* 2009, 458, 1034–1038.

[208] Bégaud, E., Chartier, L., Marechal, V., Ipero, J., et al., Reduced CD4 T cell activation and in vitro susceptibility to HIV-1 infection in exposed uninfected Central Africans. *Retrovirology* 2006, 3, 35.

[209] O'Donovan, M.R., Johns, S., Wilcox, P., The effect of PHA stimulation on lymphocyte sub-populations in whole-blood cultures. *Mutagenesis* 1995, 10, 371–374.

[210] Rafati, H., Parra, M., Hakre, S., Moshkin, Y., et al., Repressive LTR nucleosome positioning by the BAF complex is required for HIV latency. *PLoS Biol.* 2011, 9, e1001206.

[211] Christ, F., Debyser, Z., The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy. *Virology* 2013.

[212] Mous, K., Jennes, W., Camara, M., Seydi, M., et al., Expression Analysis of LEDGF/p75, APOBEC3G, TRIM5alpha, and Tetherin in a Senegalese Cohort of HIV-1-Exposed Seronegative Individuals. *PLoS ONE* 2012, 7, e33934.

[213] Strack, B., Calistri, A., Craig, S., Popova, E., Göttlinger, H.G., AIP1/ALIX Is a Binding Partner for HIV-1 p6 and EIAV p9 Functioning in Virus Budding. *Cell* 2003, 114, 689–699.

[214] Kumar, A., Abbas, W., Herbein, G., TNF and TNF receptor superfamily members in HIV infection: new cellular targets for therapy? *Mediators*

Inflamm. 2013, 2013, 484378.
[215] Koning, F.A., Otto, S.A., Hazenberg, M.D., Dekker, L., et al., Low-level CD4+ T cell activation is associated with low susceptibility to HIV-1 infection. *J. Immunol.* 2005, 175, 6117–6122.
[216] Camara, M., Dieye, T.N., Seydi, M., Diallo, A.A., et al., Low-level CD4+ T cell activation in HIV-exposed seronegative subjects: influence of gender and condom use. *J INFECT DIS* 2010, 201, 835–842.
[217] Martens, S., Howard, J., The Interferon-Inducible GTPases - Annual Review of Cell and Developmental Biology, 22(1):559. *Annu Rev Cell Dev Biol* 2006.
[218] Taylor, G.A., Feng, C.G., Sher, A., p47 GTPases: regulators of immunity to intracellular pathogens. *Nature Reviews Immunology* 2004, 4, 100–109.
[219] Gao, S., Malsburg, von der, A., Dick, A., Faelber, K., et al., Structure of myxovirus resistance protein a reveals intra- and intermolecular domain interactions required for the antiviral function. *Immunity* 2011, 35, 514–525.
[220] Sadler, A.J., Williams, B.R.G., Dynamiting viruses with MxA. *Immunity* 2011, 35, 491–493.
[221] Xiao, H., Killip, M.J., Staeheli, P., Randall, R.E., Jackson, D., The human interferon-induced MxA protein inhibits early stages of influenza A virus infection by retaining the incoming viral genome in the cytoplasm. *J. Virol.* 2013.
[222] Chieux, V., Hober, D., Harvey, J., Lion, G., et al., ScienceDirect.com - Journal of Virological Methods - The MxA protein levels in whole blood lysates of patients with various viral infections. *J. Virol. Methods* 1998, 70, 183–191.
[223] Badolato, R., Ghidini, C., Facchetti, F., Serana, F., et al., Type I interferon-dependent gene MxA in perinatal HIV-infected patients under antiretroviral therapy as marker for therapy failure and blood plasmacytoid dendritic cells depletion. *J Transl Med* 2008, 6, 49.
[224] Kawamura, M., Kusano, A., Furuya, A., Hanai, N., et al., New sandwich-type enzyme-linked immunosorbent assay for human MxA protein in a whole blood using monoclonal antibodies against GTP-binding domain for recognition of viral infection. *J. Clin. Lab. Anal.* 2012, 26, 174–183.
[225] Singh, R., Gaiha, G., Werner, L., McKim, K., et al., Association of TRIM22 with the type 1 interferon response and viral control during primary HIV-1 infection. *J. Virol.* 2011, 85, 208–216.
[226] Murphy, K., Irvin, S.C., Herold, B.C., Research Gaps in Defining the Biological Link between HIV Risk and Hormonal Contraception. *American Journal of Reproductive Immunology* 2014, n/a–n/a.
[227] PhD, D.C.B.P., PhD, K.M.C., Use of hormonal contraceptives and HIV acquisition in women: a systematic review of the epidemiological evidence. *The Lancet Infectious Diseases* 2013, 13, 797–808.
[228] Tomasicchio, M., Avenant, C., Toit, Du, A., Ray, R.M., Hapgood, J.P., The Progestin-Only Contraceptive Medroxyprogesterone Acetate, but Not Norethisterone Acetate, Enhances HIV-1 Vpr-Mediated Apoptosis in Human CD4+ T Cells through the Glucocorticoid Receptor. *PLoS ONE* 2013, 8, e62895.

[229] Hapgood, J.P., Immunosuppressive Biological Mechanisms Support Reassessment of Use of the Injectable Contraceptive Medroxyprogesterone Acetate. *Endocrinology* 2013, 154, 985–988.
[230] McCoy, S.I., Zheng, W., Montgomery, E.T., Blanchard, K., et al., Oral and injectable contraception use and risk of HIV acquisition among women in sub-Saharan Africa. *AIDS* 2013, 27, 1001–1009.
[231] Heffron, R., Donnell, D., Rees, H., Celum, C., et al., Use of hormonal contraceptives and risk of HIV-1 transmission: a prospective cohort study. *The Lancet Infectious Diseases* 2012, 12, 19–26.
[232] Morrison, C.S., Chen, P.-L., Kwok, C., Richardson, B.A., et al., Hormonal contraception and HIV acquisition: reanalysis using marginal structural modeling. *AIDS* 2010, 24, 1778–1781.
[233] Hughes, G.C., Thomas, S., Li, C., Kaja, M.-K., Clark, E.A., Cutting edge: progesterone regulates IFN-alpha production by plasmacytoid dendritic cells. *J. Immunol.* 2008, 180, 2029–2033.
[234] Tayel, S.S., Helmy, A.A., Ahmed, R., Esmat, G., et al., Progesterone suppresses interferon signaling by repressing TLR-7 and MxA expression in peripheral blood mononuclear cells of patients infected with hepatitis C virus. *Arch. Virol.* 2013.

Chapter 11. Appendices

11.1 Abbreviations

HESN	highly exposed HIV seronegative
HIV-1	human immunodeficiency virus 1
SIV	simian immunodeficiency virus
AIDS	acquired immunodeficiency syndrome
CRF	circulating recombinant forms
ARV	antiretroviral
LTR	long terminal repeat
TasP	treatment as prevention
PreP	pre-exposure prophylaxis
STI	sexually transmitted infection
GALT	gut associated lymphoid tissue
LTNP	long-term non-progressors
EC	elite controllers
RP	rapid progressors
AAV	adeno-associated virus
DC	dendritic cell
TLR	toll-like receptor
APC	antigen presenting cell
MHC	major histocompatibility complex
NK	natural killer cell
ADCC	antibody dependent cellular cytotoxicity
CTL	CD8+ cytotoxic T cell
TCR	T cell receptor
MSM	men who have sex with men
IDU	injection drug user
CSW	commercial sex worker
FGT	female genital tract
HLA	human leukocyte antigen
KIR	killer immunoglobulin receptor
IRF	interferon regulatory factor
CMC	cervical mononuclear cells
SNP	single nucleotide polymorphism
SCX	strong cation exchange
SAX	strong anion exchange
RP	reversed phase
IEF	isoelectric focusing
MALDI	matrix assisted laser desorption ionization
ESI	electrospray ionization
qMS	quantitative mass spectrometry
2D-DIGE	two-dimensional difference gel electrophoresis
IT	ion trap

TOF	time of flight
Q	quadrapole
MS	mass spectrometer
SILAC	stable isotope labeling by amino acids in cell culture
iCAT	isotope coded affinity tag
iTRAQ	isotope coded tags for relative and absolute quantitation
TMT	tandem mass tags
HAD	HIV associated dementia
CSF	cerebral spinal fluid
CVL	cervical vaginal lavage
PBMC	peripheral blood mononuclear cells
HIV-N	HIV negative/susceptible controls
LC	liquid chromatography
MuDPIT	multi-dimensional protein identification technology
DTT	dithiothreitol
FASP	filter aided sample preparation
HILIC	hydrophobic interaction liquid chromatography
Hp-RP	high-pH reversed phase chromatography
HCD	high energy collision dissociation
PHA	phytohemagglutinin