Characterization Of Human Foreskin Langerhans Cells

By

Qumbelo Yamkela

A dissertation submitted in fulfillment of the requirements for the degree of

MSc (Med) in Clinical Sciences and Immunology

Department of Pathology

Division of Immunology

Faculty of Health Science

University of Cape Town

April 2020

Acknowledgements

I would like to thank sincerely my supervisor Doctor Nyaradzo Chigorimbo-Tsikiwa, for believing that am worthy of being a masters student and take on her project. I would like to thank her for making sure that I am not only a better scientist but also a better person and how to grow as an individual. She became more than a supervisor but a life mentor.

I would also like to thank the Gray/Jaspan and Dr Musa Mhlanga's laboratory team for assisting me with both laboratory work and writing up, more especially my co-supervisors Prof Clive Grey and Dr. Kyle O'Hagan.

I appreciate the support I had from my friends and family through all the hardships especially Ncumisa Qumbelo, my aunt and my grandmother who have been the wind beneath my wings. I appreciate your constant support and motivation.

Mostly, I want to thank my funders National Research Fund (NRF), Poliomyelitis Research Foundation (PRF) and University of Cape Town (UCT) Masters fund for settling all my study bills for this degree.

This thesis is dedicated to my grandmother, Mamlamli Qumbelo

Declaration

I, *Yamkela Qumbelo*, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature:

Signed by candidate

Date: 13 April 2020

Table of Contents

List of Figures

List of Tables

Abreviations

^{0}C: Degree(s) Celsius
AIDS: Acquired Immunodeficiency Syndrome

APCs: Antigen Presenting Cells

CCR5: Chemokine Receptor 5

CD: Cluster of Differentiation

cm: centimeter

CXCR4: C-X-C chemokine receptor 4

DC-SIGN: Dendritic cell-specific intercellular adhesion molecule 3-Grabbing non-integrin

DCs: Dendritic Cells

DMSO: Dimethyl Sulfoxide

DNA: Deoxyribonucleic Acid

FS: Foreskin

g: Gram

GM-CSF: Granulocyte macrophage colony stimulating factor

Gp120: 120kDa Envelope glycoprotein

Gp140: 140kDa Envelope glycoprotein

HIV: Human Immunodeficiency Virus

HSV: Herpes Simplex Virus

ICAM: Intercellular adhesion molecule

IL-4: Interleukin-4

LC-MS/MS: Liquid Chromatography with tandem Mass spectrometry

Min: Minutes

ml: Milliliter

MMC: Medical Male Circumcision

n: number of samples

PBS: Phosphate Buffered Saline

PCR: Polymerase Chain Reaction

RNA: Riboncleic Acid

RPMI: Roswell Park Memorial Institute Medium

RT: Room TemperatureLCs: Langerhans Cells

scRNA-seq: Single Cell Ribonucleic acid sequencing

SPICE: Simplified Presentation of incredibly complex evaluations

T cell: Thymus cell

TGF b1: transforming growth factor beta 1

TNF- α: Tumor Necrosis Factor alpha

UCT: University of Cape Town

UNAIDS: Joint United Nations Programme on HIV/AIDS

VIVID: Violet Fluorescent reactive dye

vs: Versus

µ: Micro

Abstract

Background: It is known that medical male circumcision (MMC) decreases HIV acquisition by up to 60%. One hypothesis is that MMC removes a foreskin (FS) that harbors different immune cells that are HIV target cells such as CD4+ macrophages, T, Langerhans (LCs), and dendritic cells (DCs). However, there have been different reports on whether the inner FS or outer FS has more HIV target cells. While LCs have been implicated in HIV transmission, their role remains controversial. Studies have shown that LCs can transmit the virus to T cells, which increases infection. On the contrary, others have reported that LCs prevent infection by degrading the virus through a langerin-dependent pathway. One of the factors that plays a major role in HIV transmission is their state of maturity and activation, which can be influenced by co-infection and other immunological processes. The aim of this study was to isolate, quantify and characterize Langerhans cells in the inner FS and outer FS from men undergoing MMC and to evaluate the phenotype of matured and activated LCS. Differences in the proteome of the inner FS and outer FS tissues were further investigated.

Methodology: FS were obtained from men undergoing voluntary MMC from clinics and hospitals in the Western Cape (Age 18 years or older). Epidermal FS cells were extracted using crawl (migratory) assay and liberase enzyme digestion. Langerhans cells were isolated by density gradient centrifugation, sorted, quantified and immune-profiled by flow cytometry. CD1a and CD207 were used to identify Langerhans cells while HLA-DR, CD80, CD86 and CD40 were used as markers of maturity and activation. The gene expression profile of sorted LCs was also examined by single-cell sequencing with seq-well. Lastly, the differences in the proteome of the inner FS and the outer FS migrated epidermal cells were assessed using liquid chromatography with tandem mass spectrometry (LC-MS/MS).

Results: Langerhans cells were an average of 85% pure post-sorting. The numbers of Langerhans cells between the inner FS vs. outer FS were not statistically different (mean: 0.56% vs. 0.68% (SD=0.37) from migratory cells and 0.28% vs. 0.45% (SD=0.18) from enzyme digest, p-value >0.05, n=9). Sequencing showed that the sorted cells pooled from 5 participants (inner and outer FS) had different gene expression profiles. Furthermore, two groups of cells were identified from the sorted LCs based on their gene expression profile. The identified cells were monocyte-like and melanocyte-like cells. The monocyte-like cells were

identified as LCs based on their gene expression profile while the melanocyte-like cells were identified as the contaminating cells as the cell purity was not 100%.

Upon activation with tumor necrosis factor alpha (TNF-α), activated LCs isolated by the migration assay had similar proportions of cells expressing surface maturity and activation markers (HLA-DR, CD40 and CD80/86) when compared to the unstimulated controls (inactivated) (mean: 73.53% vs. 75.66%, n=9, p-value >0.05, SD=4.4) However LCs that were isolated by the migration assay expressed markers of activation at a higher level compared to LCs isolated by Liberase enzyme digestion (mean: 79.4% vs. 40%, p-value < 0.05 n=9, SD=23).

Proteomics showed that the inner FS had an over-abundance of proteins involved in the interleukin 7 response and mRNA catabolic processes, while the outer FS had more spindle zones and cornified envelope proteins that were over-abundant when comparing inner and outer FS from 5 participants.

Discussion and Conclusion: The study successfully extracted, sorted and immunoprofiled Langerhans cells using different methods and from different FS compartments (inner FS versus outer FS). When LCs were spontaneously migrated and isolated using the "crawl method", they showed a more mature and activated phenotype compared to non-migrating "skin resident" immune cells. No differences were found between cells that were stimulated with inflammatory cytokines relative to unstimulated migratory cells in proportion of cells expressing activation markers. However, it was observed that cells isolated by liberase enzyme digestion showed significantly lower proportions of activation markers relative to migratory cells.

Using LC-MS/MS-based proteomics; the inner FS exhibited high expression of proteins involved in the interleukin 7 response while the outer FS exhibited high expression of structural proteins, which suggests that the inner FS might be more involved in immunity as interleukins can stimulate immune response while the outer FS has a more structural role than the inner FS.

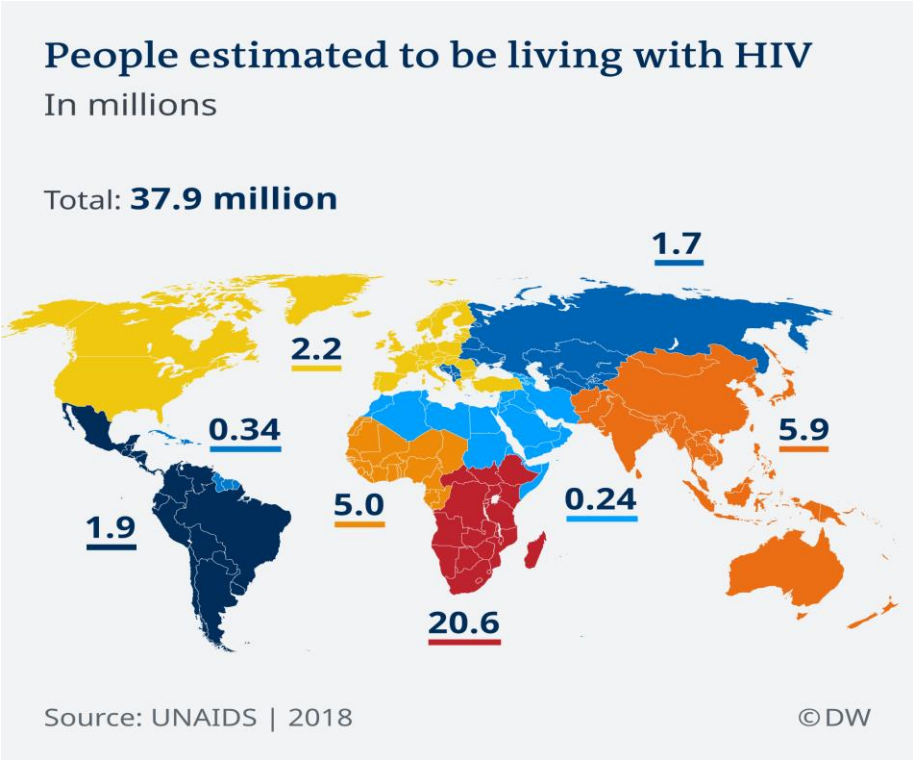
Chapter 1

Introduction

Human immunodeficiency virus (HIV)

The discovery of HIV as the causative agent of Acquired Immunodeficiency Syndrome (AIDS) gave hope to an end of the AIDS epidemic over 35 years ago (Ganor & Bomsel, 2011). However, to date there is still no cure or a fully efficacious preventative vaccine against the virus (Ahmed & Piguet, 2015). East and Southern African countries carry the highest burden; by 2018, there were 20.6 million people living with HIV in these regions as depicted in Figure 1.1. South Africa had over 240,000 new infections in 2018 alone and currently has a population of 7.7 million people living with the virus based on UNAIDS statistics 2018/2019.

A



B

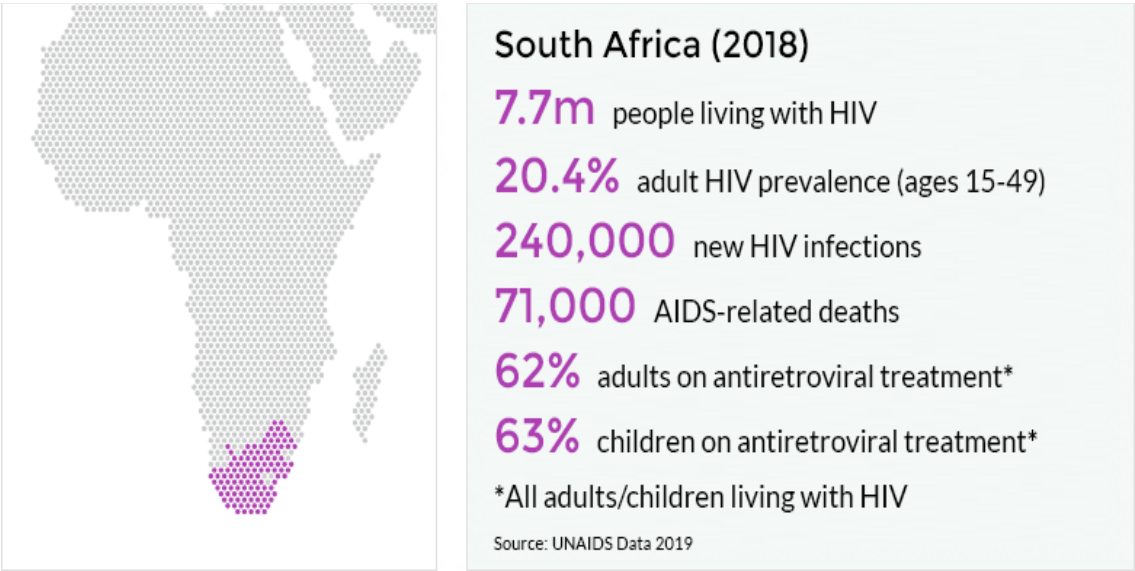


Figure 1.1: Proportion of people living with HIV and AIDS globally and in South Africa. (A)The number of people estimated to be living with HIV worldwide. There were 37.9 million people estimated to be living with HIV and 20.6 million are from African countries. Africa is more affected compared to other continents as over a half of the affected population is from this region. Figure (B) depicts prevalence of HIV in South Africa. There were 7.7 million people living with HIV in 2018 (UNAIDS, 2018/2019).

HIV life cycle and acquisition

To better understand HIV and discover ways to combat it, it is important to understand acquisition and the life cycle of the virus. Figure 1.2 below illustrates what is currently known about the pathway that leads to HIV infection. HIV is a Lentivirus that uses the host cells machinery to establish an infection.

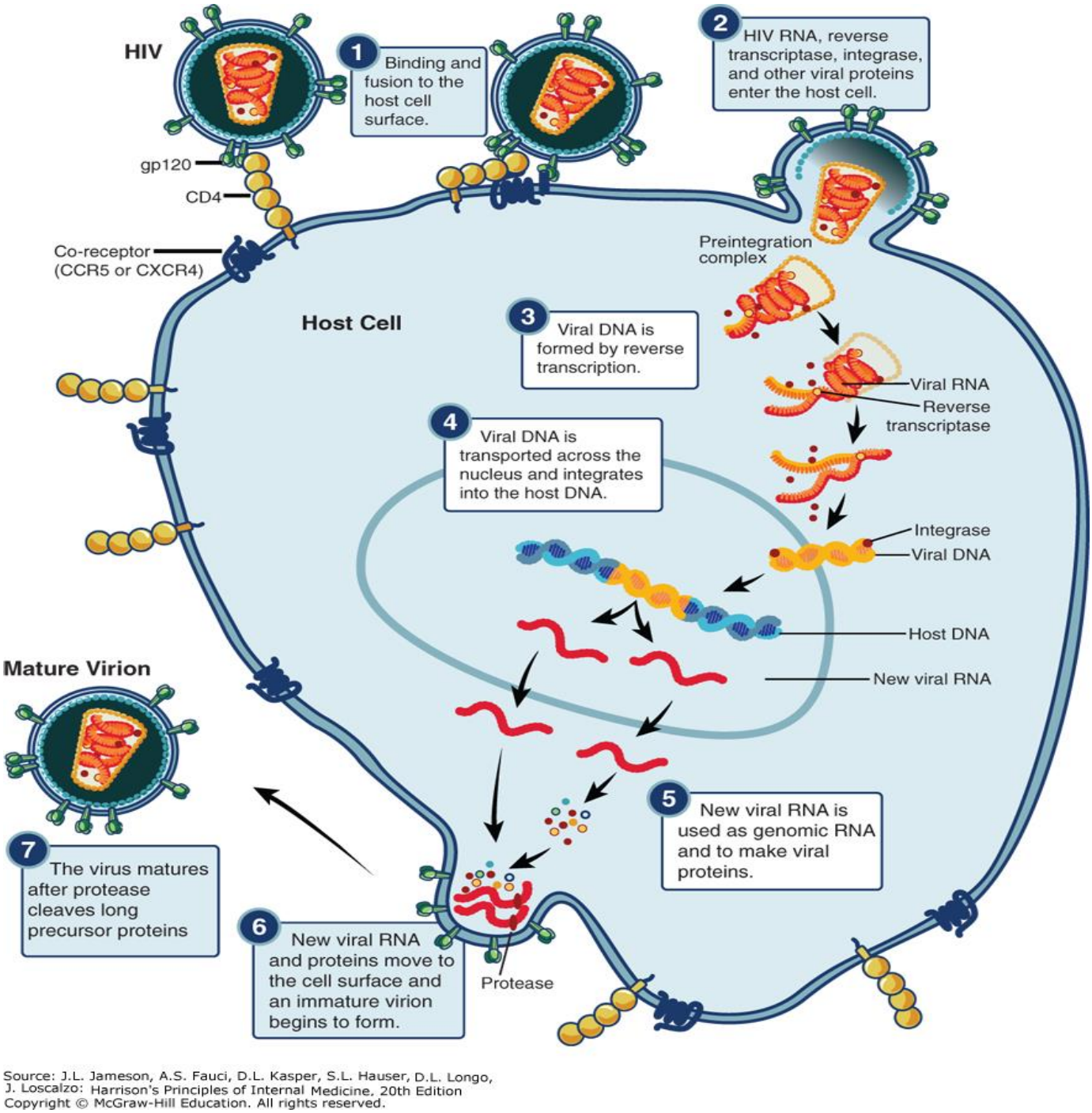


Figure 1.2: Human immunodeficiency virus (HIV) life cycle. Schematic representation of HIV life cycle from host cell entry to a mature infectious virion (taken from https://www.niaid.nih.gov/sites/default).

HIV-host cell interaction

The first step of the virus life cycle is fusion and entry as shown in Figure 1.2 step 1. In-order for HIV to enter the cell, it binds to host cell receptor cluster of differentiation 4 (CD4). CD4 is a glycoprotein expressed by cells of the immune system such as macrophages, T, Langerhans (LCs), and dendritic cells (DCs), which makes them HIV target cells. Typically, CD4 binds to the major histocompatibility class II (MHC II) complex on antigen presenting cells (APC) as a co-receptor to T cell receptors. It also activates T cells and is crucial for establishing the T cell repertoire during development in the thymus. Cell activation is a process of differentiation and proliferation of a naïve cell usually triggered by an antigen. CD4 also assists in prolonging the interaction with MHC II on antigen presenting cells (Abbas & Herbein, 2013). For viral fusion to occur, the viral envelope binds with the trimetric glycoproteins gp120 and gp140. This protein-to-protein interaction leads to a conformational rearrangement in both proteins, which helps the virus to bypass the host checkpoints. Following this, the virus binds to another host cell receptor, C-C Chemokine Receptor 5 (CCR5) or C-X-C Chemokine Receptor 4 (CXCR4), depending on the virus strain (Koppensteiner & Wu, 2012). The R5-tropic strains target the CCR5 co-receptor, whereas the R4-tropic strains target the CXCR4 co-receptor (Ahmed et al., 2015). Viral envelope fusion with the host cell membrane allows the virus to transfer the viral RNA and encapsulated enzymes, such as reverse transcriptase, to the cell (Février et al., 2011).

Once the virus enters the cell (step 2 in Figure 2.1), the reverse transcriptase enzyme converts viral RNA into DNA. The viral DNA is subsequently integrated into the host DNA in the nucleus, using the enzyme integrase. Following this, the virus begins to use the host cell's machinery to create long chains of viral proteins (Hindmarsh, 1999). The new immature viral RNA and proteins move to the surface of the cell, where the contents can bud off from the cell membrane. At this stage, the virus cannot infect other cells. However, once the virus has been released from the host cell, the viral protease cuts the long protein chains. These short proteins form virions that are mature enough to infect other cells using the same cycle (Piguet & Steinman, 2007).

HIV transmission

There are different routes for HIV transmission, and sexual intercourse is a major cause as it contributes over 70% of reported cases of HIV incidence (Shaw & Hunter, 2012). Even though

sexual intercourse has been reported to be the major route, it has also been reported that there is a very low chance of transmission via this route. The chance of transmission is estimated to be 1 in 200 to 1 in 3000 exposures during male-to-female and female-to-male exposure (Hladik et al., 2007). Since there is a low chance of HIV transmission during sexual intercourse, how is it the major route of transmission? The hypothesis remains that there must be other factors that play a role that may increase the chances of transmission and infection. Co-infection of sexually transmitted infections (STIs), genital ulceration, microbiome dysbiosis, the stage of HIV in the infected person, partner viral load, transmission route (penile-vaginal or penile-anal) (Shaw & Hunter, 2012) are all suggested to play a role in HIV transmission. Even though there are different routes of HIV transmission, once transmission is established the events that lead to productive infection are similar (Fiebig et al., 2003)

Although there are treatment strategies available for HIV infection, the infection still remains incurable to date. Different routes of HIV transmission and different viral strains form part of an obstacle in curing or designing vaccines for HIV (Sciences, 2016). Therefore, different and improved prevention strategies are needed. One such prevention strategy is medical male circumcision (MMC).

MMC as a prevention strategy against HIV

MMC, which is a complete removal of the foreskin (FS), has been established to be one of the prevention methods to HIV acquisition and infection by decreasing chances of acquisition by up to 60% (Auvert et al., 2005; Bailey et al., 2007; Gray et al., 2007).

MMC

Circumcision was initially employed as a religious and cultural act. Currently, MMC is used for medical benefits such as hygiene and prevention of infection with STIs such as HIV (Gray et al., 2007). One of the first studies suggesting that MMC is a preventative measure against HIV was published in 1989 (Bongaarts et al., 1989). Since then, many studies have investigated how the FS is susceptible to HIV. Through these studies, it was noted that most Eastern and Southern-African men, where HIV is most prevalent, were not circumcised (Auvert et al., 2005). However, circumcision does not provide direct protection of women from HIV infection from circumcised men (Haberland et al., 2016). Although it does not protect women

against HIV, it has been shown to provide protection against other STIs, cervical cancer and other associated conditions (Morris et al., 2019)

Other sites of the penis also play a role in HIV acquisition such as the urethra, but it is suggested that the bigger surface area of the FS increases the chances of acquiring the virus (Kigozi et al., 2008). Circumcision is now a highly recommended HIV prevention method (Liu et al., 2014). The initial trials that demonstrated HIV prevention through circumcision were observational studies, which raised the question of how the FS is susceptible to HIV infection. The skin is known to be an immune competent organ containing various immune cells (Clayton et al., 2017), therefore hinting at a possible role of the FS during HIV acquisition.

The Foreskin

The male genital tract is partially covered by FS tissue also known as the prepuce. The FS covers the tip of the penis and the urethra, an opening for urine and semen. When the FS is pulled back, it exposes the glans of the penis. The FS maintains the penis' sensitivity as it cushions it from damage and rubbing against abrasive agents (Patterson et al., 2002b). The FS, like skin on other parts of the body, is made of 3 parts: the epidermis, dermis and the subcutaneous tissue also known as the hypodermis. The FS primarily protects the male genitalia from environmental pathogens (Szabo & Short, 2000).

Foreskin immunity

FS immunity consists of components of both the innate and adaptive immune response. This provides defense against pathogens. During pathogen encounter, the innate immune response is the first line of defense. It consists of molecules that recognize the pathogens genetic material and send signals to the cells to release effector molecules, such as cytokines. The effector molecules are the first step in eliminating the pathogen. Following this, the adaptive immune response forms a more robust defense. It kills the pathogen, the infected cells and clears any toxins produced. The adaptive immune response also consists of memory, which is an ability to recognize previously encountered pathogens. Once a pathogen is neutralized, memory cells are formed to help neutralize the pathogen more rapidly upon future encounter (Bailey et al., 2007).

The two arms of the immune response are composed of unique cells. These cells all communicate with each other and the site of infection by means of circulation via the lymph node depicted on Figure 1.3. Skin antigen presenting cells (LCs and DCs) capture the pathogen

and present it on the cell surface through the MHC class II complex. Subsequently, interleukin-1, tumor necrosis factor alpha (TNF-α) and other cytokines are secreted. This leads to high expression of activation markers on the antigen presenting cells (Marsden et al., 2015). Expression of migratory receptors, such as CCR5 and CCR7 are also increased to help aid in the migration of cells to the lymph node. T-lymphocytes also become activated and express cell recruitment receptors. The T cells interact with the processed antigen via MHC II, proliferate and develop effector function then neutralize or eliminate the antigen (Koppensteiner & Wu, 2012). FS does not only act as a physical barrier but plays a major role in the immune response (Février et al., 2011).

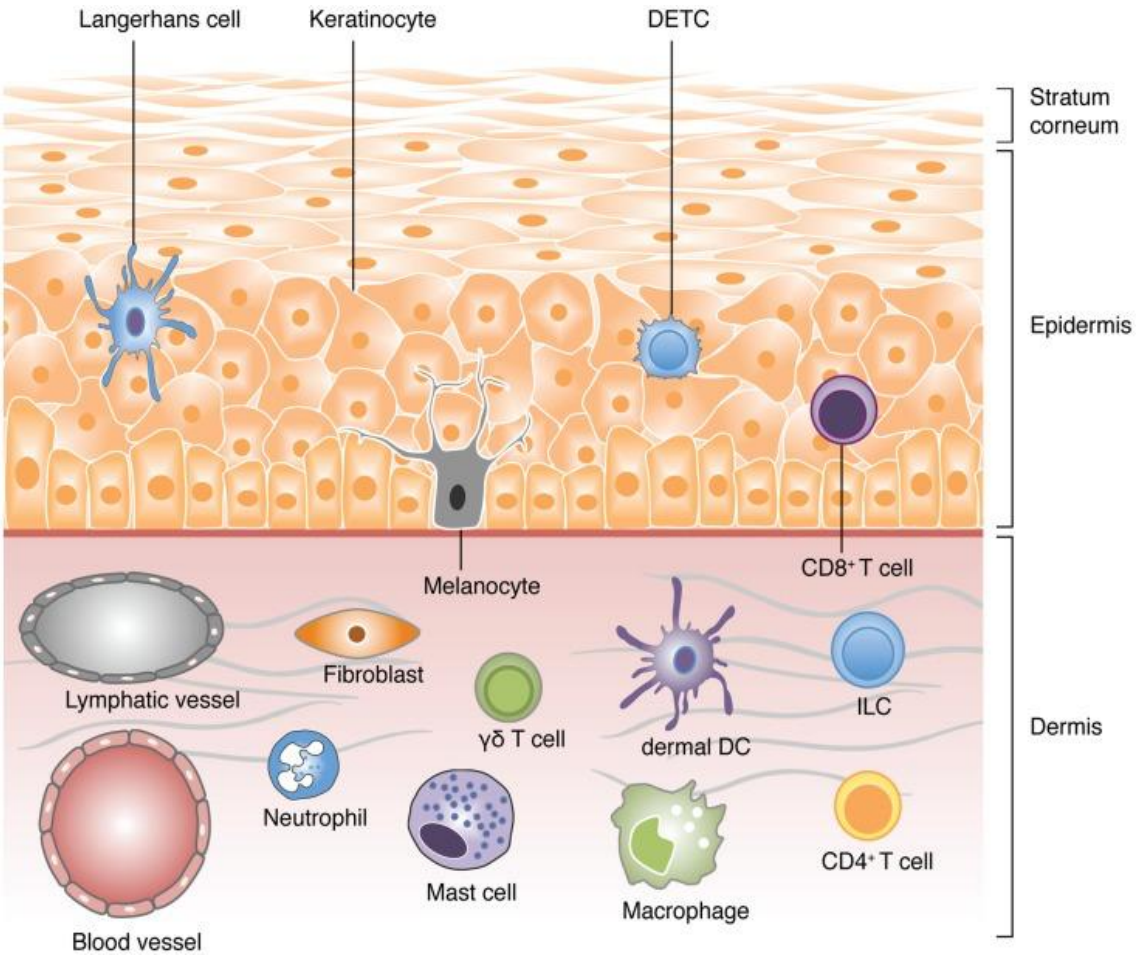


Figure 1.3: Different skin compartments and cells of the immune system in the skin. The skin has two major components which are the epidermis and the dermis. It also has a stratum corneum on top of the epidermis which is composed of dead keratin cells acting as barrier. The epidermis consists of Langerhans cells and dendritic epidermal T cells (DETC) which is a subtype of γδ T-cells. These are the major immune cells of the epidermis. The dermis is composed of many different cell types, fibroblasts, neutrophils etc. All these compartments work together to prevent pathogen invasion.

Inner and outer aspects of the FS

The FS is composed of inner FS and the outer FS components. They are named so because of their location, the inner FS is usually unexposed on a flaccid penis while the outer FS is a continuation of the skin on the shaft of the penis and is visible as shown in Figure 1.4 (Ganor & Bomsel, 2011). They are differentiated by the abundance of melanocytes with the outer FS containing more melanocytes than the inner FS (Hussain & Lehner, 1995) and also differ in the amount of keratinization (Botting et al., 2017). However, there are conflicting reports on the levels of keratinization of the inner FS and outer FS. Some studies have shown that the inner FS is more keratinized than the outer FS (Qin et al., 2009); others have shown no difference (Wang et al., 2008) while some suggest that the outer FS is more keratinized compared to the inner FS (Patterson et al., 2002a). This highlights the need for more robust studies on inner FS and outer FS keratinization. Keratin protects the skin epithelial cells therefore, if any aspects of the FS possess less keratin, the barrier will likely be easily breached making it more susceptible to pathogens and HIV (Shaw & Hunter, 2012).

Density of HIV target cells in the inner FS and the outer FS

There have been contradicting results regarding HIV target cell density between the inner FS and the outer FS. An *ex vivo* study to demonstrate the susceptibility of the FS to HIV showed that the inner FS possesses more immune cells that are potential HIV target cells compared to the outer FS including CD4+ T cells and LCs (Liu et al., 2014). On the contrary, Donoval et al., (2006) reported no significant difference in the density of inner FS and outer FS LCs. However, differences were observed in different compartments of the skin. There were more CD4+ T cells in the epidermis than the dermis and macrophages were found in the dermis. Patterson et al., (2002) evaluated distribution of cells in men who were circumcised due to inflammation and other FS conditions, such as phimosis, balanitis, etc. They found significantly higher numbers of LCs and CD4+ T cells in the inner FS compared to the outer FS. Furthermore, they found no differences in density of macrophages. Their findings suggest that the inner FS is more susceptible to the virus compared to the outer FS, as more HIV target cells were present in the inner FS compared to the outer FS (Patterson et al., 2002).

Table 1.1: List of studies that investigated the abundance of T and Langerhans cells (LCs) in the foreskin (FS) and their findings.

Author	**Health status, age and method of cell detection used**	**Findings**
McCoombe (2006)	*30 participants, 21 from MMC and 9 were post-mortem, all HIV negative. *Age 18 years and older *Immunohistochemistry (IHC) microscopy (confocal) was used for cell visualization	LCs were significantly more abundant in the outer FS compared to the inner FS. Inner FS: 61.3 cells/mm^2 and Outer FS: 85.5 cells/mm^2 There were differences in abundance of CD4+ T cells and macrophages only between the dermis and the epidermis with the dermis bearing more of both cell types.
Fischetti (2009)	* Health and STI status not specified *Age unspecified *Density and distribution of CD4 and CD1a cells were visualized by microscopy (IHC)	The study found more Langerhans cells in the glans than the inner FS and outer FS while the inner FS had more LCs than the outer FS. There were 230 LCs Cells/mm^2 found in the inner FS and 170 cells/mm^2 in outer FS The same results were observed in CD4+ Cells.
Donoval et al (2006)	*39 participants from randomized trial, 20 had no STI history while 19 had history of treated STI. *The participants were aged 18-24 years. *IHC analysis was performed by Enhanced V-Red detection kit	Differences between inner FS and the outer FS were not specified. LCs were reported to be more abundant in the epidermis compared to the dermis for both inner FS and outer FS, more macrophages in the dermis and no differences in CD4+ T cells were reported between the dermis and the epidermis. Men with history of STI had more Langerhans cells and macrophages compared to those with no history of infection. Epidermis Dermis LCs 1.23% 0.30% CD4 T cells 0.08% 0.02% Macrophages 0.075% 0.04%
Ganor et al (2010)	* FS explants with no history of pathological infections. *Age 18 to 87 years *The study used IHC microscopy	LCs were only found in the epidermis on both inner FS and outer FS. There were 62% more LCs in the inner FS compared to the outer FS.

Author	Health status, age and method of cell detection used	Findings
Hirbod (2010)	*33 STI negative males. *Age: 18-24 years *Samples were analyzed by confocal in situ imaging microscopy.	No differences between the inner FS and the outer FS LCs (average of 150 cells/mm^2 from the inner FS and the outer FS).
Qi Qin (2009)	*60 participants were 2-7 years and 20 were 20-29 years *Young participants were separated according to presence of urinary tract infections (UTI) and no UTI. The adults were healthy with no sign of UTI. * Staining was done with H&E, cells were examined with light microscopy	Inner FS LCs of infected children: 132.2 cells/mm^2. Outer FS LCs of infected children: 131.7 cells/mm^2 (inner FS=outer FS) Healthy children inner FS LCs: 87.5 cells/mm^2, outer FS LCs: 123.7 cells/mm^2 (inner FS <outer FS by 40%) Healthy adults inner FS LCs: 87.5 cells/mm^2, outer FS LCs: 88.3 cells/mm^2 (inner FS=outer FS)
Liu et al., 2014	* Mean age 27 years (between 10 and 53 years) * Samples were obtained from healthy youth and adults undergoing circumcision due to phimosis and FS redundancy. Only participants without detectable STIs and other conditions such as UTI were qualified to participate. *IHC staining and flow cytometry was used to quantify cells.	LCs were significantly higher in the epidermis of the inner FS than in the outer FS (Mean: 63.1% in the inner FS vs. 52.8% in the outer FS. There were no significant differences observed in the proportions of CD4+ T cells (34.6% vs. 37.7%)
Patterson (2002)	*Participants aged from 10 months to 69 years. *Circumcised due to phimosis and redundant FS. * Cells were quantified using immunofluorescence and image analysis	Both LCs and CD4+ T cells were significantly higher in the inner FS than the outer FS.

There are many attributes that may have led to the differences in these findings. A major factor that needs to be considered is the methodology used to quantify the cells. Some studies counted absolute cell numbers while others reported the proportion of stained cells. The inherent differences in participants also likely plays a major role such as age, asymptomatic STIs and other health conditions affecting the FS (Jayathunge et al., 2014).

Even though there were differences in the levels of keratinization and density of HIV target cells in the inner FS and outer FS, there were no significant differences in HIV infection (Liu et al., 2014) when measuring p24 (an HIV protein) antigen from virus inoculated inner FS and the outer FS. The FS culture explants were incubated with different pseudo and heat inactivated strains of HIV particles for 4 hours. In contrast, Ganor & Bomsel, (2011) observed HIV infection in the inner FS within one hour after infection, but the same was not observed for the outer FS. The inner FS and outer FS CD4+ LCs and T cells expressed more CCR5 compared to CXCR4 (Szabo & Short, 2000). This explains why there are more infections caused by R5-tropic strains compared to R4-tropic strains. Piguet and colleagues (2007) reported that the R5-tropic strains infected the inner FS more efficiently compared to R4-tropic strains. This suggests that the inner FS likely has a higher number of cells expressing CCR5 (Piguet & Steinman, 2007), resulting in greater infection incidence.

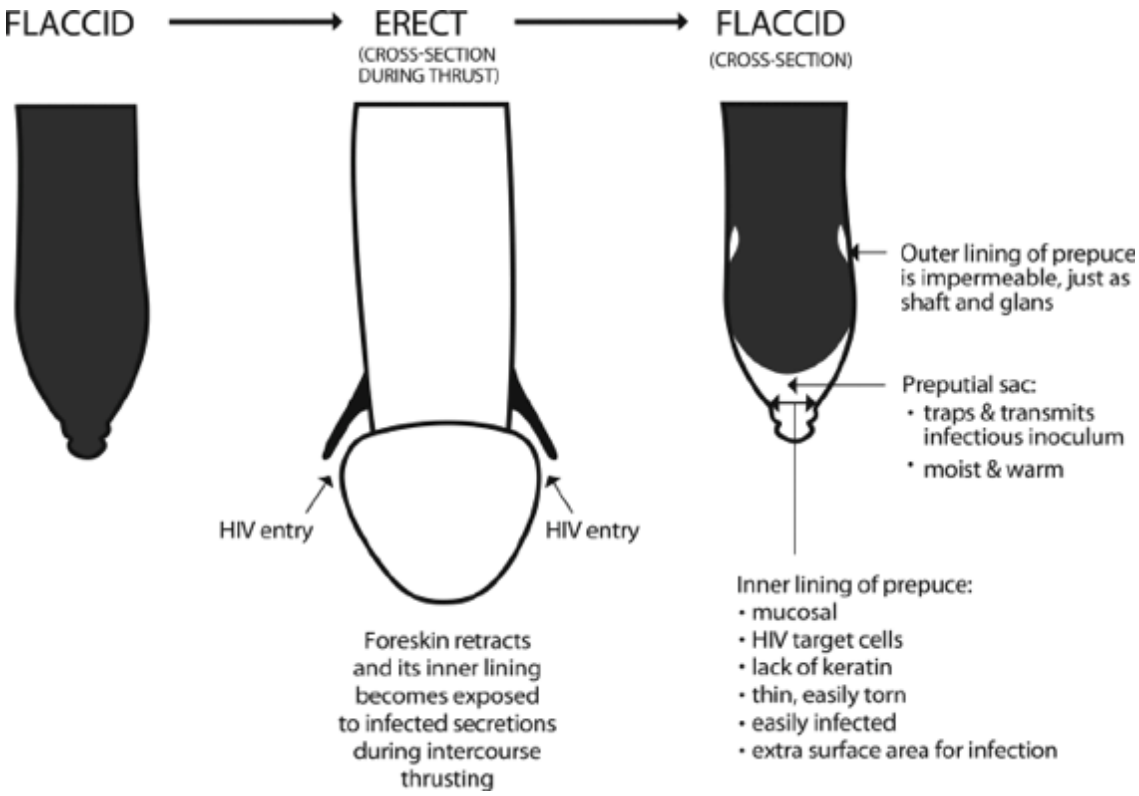


Figure 1.4 Susceptibility of the foreskin (FS) to human immunodeficiency virus (HIV). During intercourse, the inner FS is exposed with its mucosa that consists of HIV target cells, extra surface area and potentially infected with other sexually transmitted infections (STIs) making it easier for it to trap and transmit HIV. During circumcision, the outer FS which is exposed on both flaccid and erect penis and the inner FS exposed during erection, are completely removed (Morris & Wamai, 2012).

With the controversies regarding differences between the inner FS and outer FS including abundance of target cells and keratinization, it is still unknown which aspect of the tissue plays a more significant role in reducing HIV infection following MMC. However, as it stands, it appears that the inner FS may be more susceptible as majority of studies have shown that the inner FS has more target cells and is less keratinized (Szabo & Short, 2000, Ganor et al., 2010, Fischetti, 2009). Further studies need to be done to clear the controversies between the differences in these parts of the FS and to help conclude which aspect of the FS is more susceptible to HIV and other infections (Jayathunge et al., 2014).

Immune cells of the FS and their role in HIV

CD4 T cells

The hallmark of HIV is the depletion of CD4+ T cells, which weakens the immune system (Klatzmann et al., 1984). This subsequently allows infection with other opportunistic pathogens since there are fewer CD4+ T cells to induce a robust adaptive immune response

(Février et al., 2011; Koppensteiner and Wu, 2012) . After the discovery of HIV/AIDS, laboratory studies that were done on patients presenting with HIV/AIDS symptoms showed that HIV selectively infects CD4+ T cells (Okoye & Picker, 2013). This was concluded following a significant decrease in the number of CD4+ T cells of infected individuals. As a result, low CD4 cell count was used to predict the onset of overt infection and immunodeficiency as reviewed by Okoye & Picker, (2013). Additional evidence that HIV infects and kills T cells was shown when after suppression of the virus with anti-retroviral therapy, CD4+ T cell counts increased and the immunodeficiency was reversed (Février et al., 2011)

T cells get infected by HIV during antigen presentation of the virus antigen by antigen presenting cells. This is since antigen presenting cells sometimes bind the virus without being degraded. Another way T cells get infected by HIV is by direct contact of T cells with the virus after acquisition (Abbas & Herbein, 2013). The way in which HIV depletes T cells is through programmed cell death (PCD) and necrosis. Necrosis is a result of cell damage that leads to macrophage activation and the release of pro-inflammatory cytokines. Cytokines are danger signaling molecules that can activate T cells and immunity PCD is a controlled process that eliminates damaged and misplaced cells during organogenesis and tissue remodeling for development and homeostasis (reviewed by Février et al., 2011 and Okoye & Picker, 2013). There are other mechanisms through which CD4+ T cells are lost during HIV infection. These include: elimination by HIV specific cytotoxic CD8+ T cells and elevated levels of human Leukocyte antigen DRs (HLA-DRs) (MHC II surface receptors) due to hyper activated immunity, impaired production of T-lymphocytes by the thymus, changes in membrane permeability, syncytium formation etc. (Piguet and Steinman, 2007; Février et al., 2011; Abbas and Herbein, 2013).

Dendritic cells

DCs such as LCs are one of the first cells to interact with HIV during sexual transmission given their localization and function (Piguet & Steinman, 2007). They play a vital role in HIV pathogenesis as they capture pathogens and present their antigens to T cells (De Jong & Geijtenbeek, 2009). One of the questions about the relationship between dendritic cells and HIV was whether the virus infects DCs and to what extent do they get infected. It is known that the virus can cross the epithelium of the male and female genitals to DCs (Piguet & Steinman, 2007). Dendritic cells have been shown to get infected by HIV when co-cultured with T cells and that there was a more robust infection of the T cells than the dendritic cells.

These experiments were successful when using both X4 and R5 tropic strains of the virus. This led to the conclusion that dendritic cells transmit the virus to T cells and can become infected but at a lesser extent than the T cells. It is possible that this might happen during antigen presentation as this is when direct contact between T cells and DCs occurs (Piguet & Steinman, 2007). Therefore, it was proposed that the DCs transfer the virus during immunological synapse formation (when an interaction forms between antigen presenting cells and lymphocytes) (Piguet & Steinman, 2007). During this interaction, DC-captured HIV is transferred to target T cells as illustrated in Figure 1.5. Following this, T cells can infect surrounding T cells through virological synapse formation, where an infected T cell transfers the virus to uninfected T cells through T cell-T cell interactions (Abbas & Herbein, 2013; Ahmed et al., 2015; van den Berg et al., 2015).

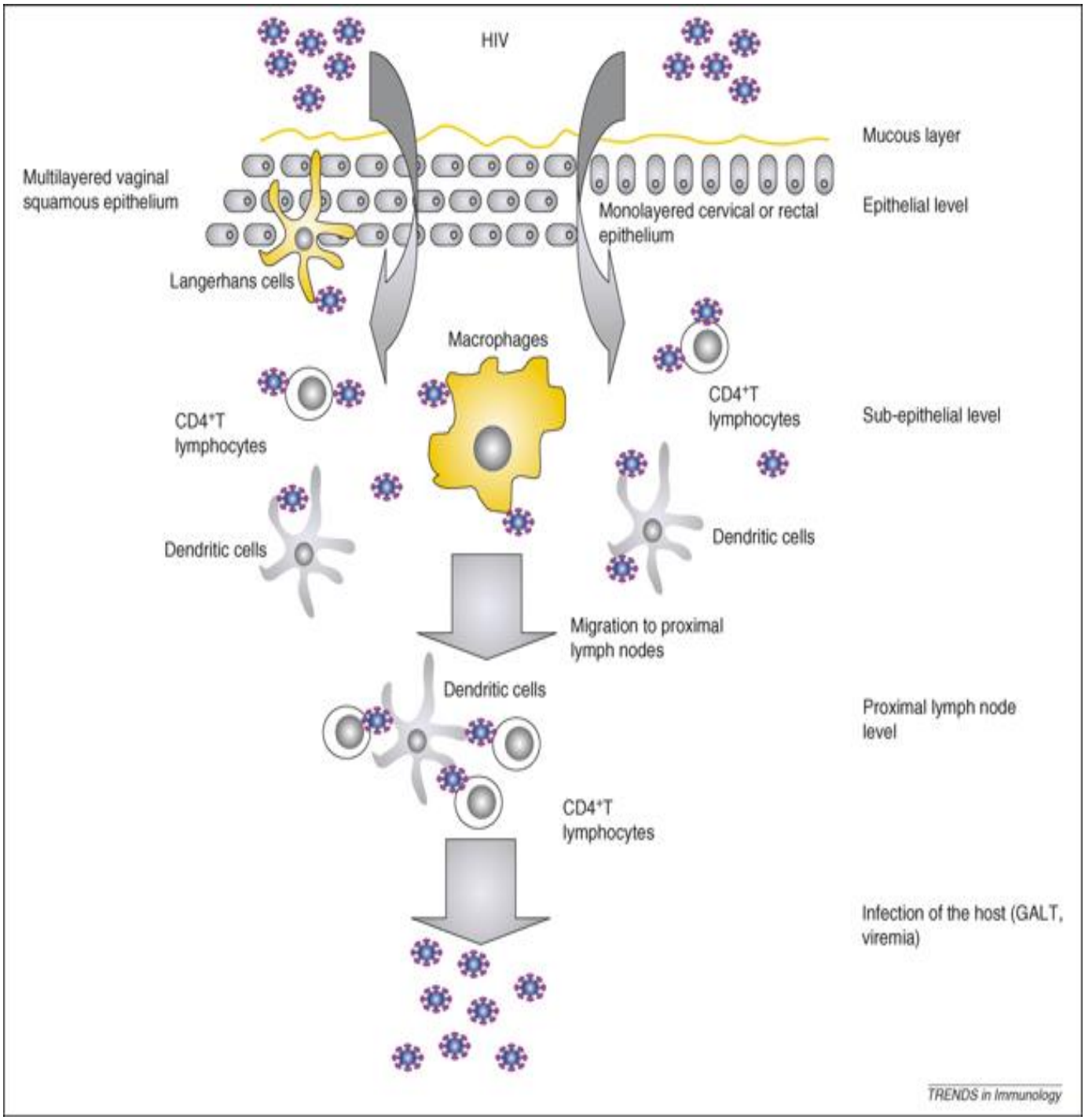


Figure 1.5: The role of dendritic cells (DCs) in human immuno-deficiency virus (HIV) infection. The virus crosses the epithelia and encounter DCs, Langerhans cells, (LCs) macrophages and other CD4+ cells. The DCs capture the virus, migrate to lymph node and transfer it to T cells through immunological synapse. In T cells, the virus multiplies and other cells also get infected leading to rapid increase in viral load. Adapted from (Piguet & Steinman, 2007).

Langerhans cells

What are Langerhans cells

LCs are a key feature of the skin immune response. Paul Langerhans discovered LCs in 1868. He used a gold chloride stain and from his results, he concluded that LCs were nerve cells. Gold chloride was a known marker for nerve cells at the time. After years of the cells being thought to be nerve cells, Steinman Ralph discovered that they were antigen-presenting cells. He was awarded a Nobel Prize in 2011 for this discovery (Botting et al., 2017; Piguet & Steinman, 2007). He described them as dendritic cells that form interconnecting networks on the skin, detecting infecting pathogens. They are mostly found on the stratum spinosum of the epidermis and as migratory cells on the dermis and lymph node.

Features and characteristics of LCs

Features that are often used to define LCs are C-lectin langerin (CD207), CD1a and HLA-DR (Peña-Cruz et al., 2001). Collectively these markers accurately distinguish LCs from other cells. This is since other antigen presenting cells express CD1a, whereas langerin has been found on dermal cells and cells of the lamina propria that are not LCs (Preza et al., 2014). HLA-DR was shown to be expressed by inflammatory dendritic epidermal cells (IDEC) that did not express langerin (Wollenberg et al., 1996) However, LCs have been shown to down regulate langerin after antigen encounter (van den Berg et al., 2015).

Langerhans cells uniquely express tennis racquet like intracellular structures called Birbeck granules. LCs express E-Cadherin and C-lectin mannose receptor EpCAM but do not express DC specific ICAM 3–grabbing non- integrin (DC-SIGN) (Peña-Cruz et al., 2001). All these markers are used to differentiate between LCs and other DCs as they are closely related to each other since they are all DCs performing similar functions (antigen recognition and presentation) and they are from the same origin (Botting et al., 2017)

Langerin

Langerin is a C-type lectin receptor that binds to mannose and fructose. It has a cytosolic tail, which is used for motion, vesicular trafficking and endocytosis. It also consists of a carbohydrate recognition domain (CRD) and a domain involved in oligomerization called the neck domain shown in Figure 1.6. The carbohydrate domain is used for binding carbohydrates. It can recognize and has high affinity for numerous carbohydrates that are expressed on different viruses. This gives it an advantage to bind to different viruses (van den Berg et al., 2015). Even though langerin has been found on other dermal DCs other than LCs,

LCs express greater levels of langerin (Romani et al., 2010). Langerin has been reported to perpetuate immune response by forming adhesions with other dendritic cells and plays a major role in HIV infection (de Witte et al., 2007).

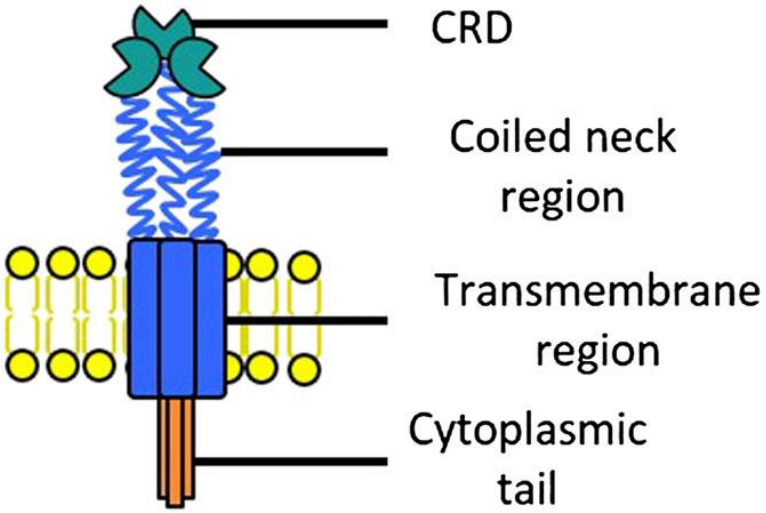


Figure 1.6: Structure of langerin. Langerin is made of the carbohydrate recognition domain, the neck, transmembrane regions and cytoplasmic tail.

Langerhans cells maturation and activation upon pathogen encounter

Langerhans cells that are found in tissue are in an immature state specializing in pathogen detection (Botting, et al., 2017). Maturation and activation of these cells is initiated by environmental stimuli such as pathogen encounter through pathogen associated molecular patterns (PAMPs) and cross talk with other cells and tissues. These include pro-inflammatory cytokines such as interferon gamma, TNF-α, growth factor-B, thymic stromal lymphopoietin (Polak et al., 2013). Once they encounter an antigen or when stimulated by a cytokine, they mature, become activated and down regulate antigen-detecting receptors shown in Figure 1.7. Subsequently, antigen presenting capacity, activation and maturation markers are up-regulated (De Jong & Geijtenbeek, 2009). These include markers such as HLA-DR, CD80, CD86, CD83, CD40 and CD70 (Ahmed et al., 2015). CD80 and CD86 are both found in the same antigen presenting cells and they both bind to the same receptors. They are CD28 and CD152 ligands. CD28 is responsible for promoting immune response while CD152 inhibits immune response. It is difficult to ascertain whether CD80 or CD86 enhance or inhibit responses since they both bind to these receptors, CD28 and CD152 (Larsen et al., 1990; Randolph, 2002). CD40 is also responsible for immune activation (Geissmann et al., 2002). It binds to CD154, a marker of activation of T helper cells and thus activates antigen-presenting cells leading to different downstream effects. HLA-DR is an MHC II cell surface receptor that presents antigen

peptides to T cells to initiate immune response or suppression (Geissmann et al., 2002; Randolph, 2002; Segura et al., 2013).

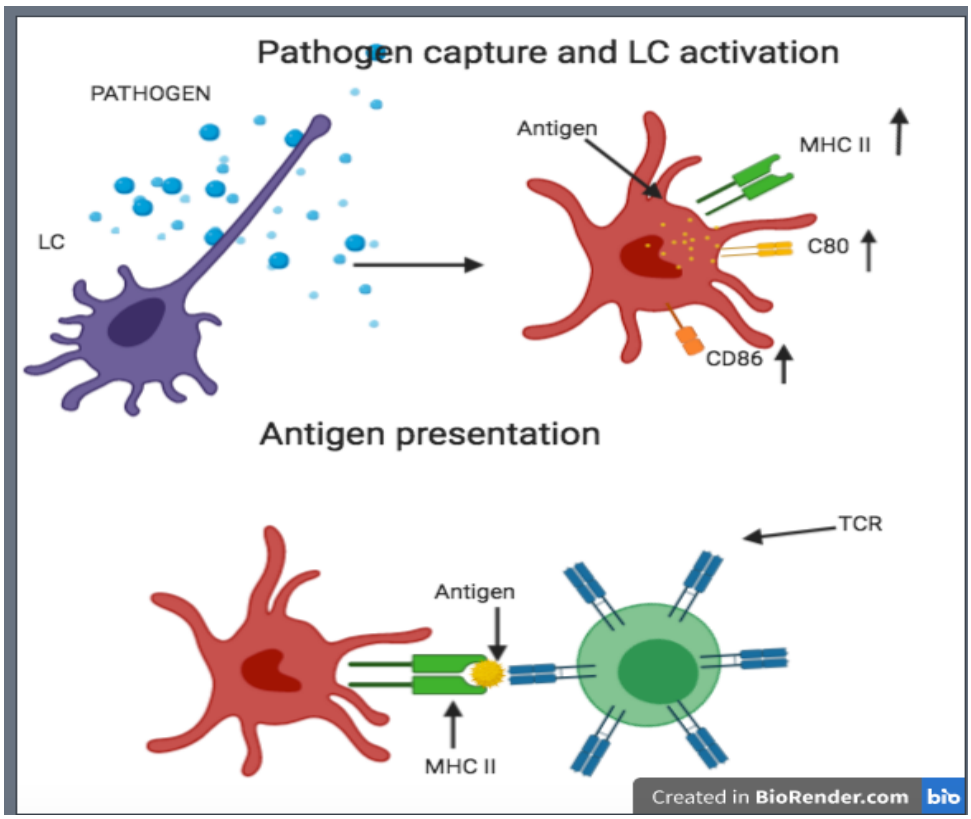


Figure 1.7: Langerhans cell (LCs) activation and antigen presentation: LCs detect pathogens with their long dendrites. As a result, they mature and are activated leading to high expression of antigen presenting molecules (MHC II) and activation markers (CD80 and CD86). The cell migrates to T cells to initiate antigen presentation and adaptive immune response.

Isolation of Langerhans cells

Since their discovery as DCs, they have been one of the most studied cells (Ganor & Bomsel, 2011). However, there have been many challenges in studying Langerhans cells. This is since they are embedded deeply in the skin epidermis and only make up 1-3% of the total population of skin cells (Botting et al., 2017). Another challenge is that researchers have not found the most suitable conditions to culture LCs. As a result, they do not have a long-life span *ex-vivo*. LCs have been found to be more difficult to isolate when working with genital tissue, such as FS, labia or penis glands and type II mucosa such as the anal canal, and vagina (Berg et al., 2014). Some researchers have used skin from abdominoplasty since the surface area is larger in size, helping to yield high quantities of cells (Polak et al., 2013).

There are differences in cell phenotype depending on isolation methods, which have led to different results for their role in pathogen interaction and transmission. To date, investigators have tried to derive methods to mimic true LCs *in-vitro*, with poor success. Most studies have used CD34+, CD14+ monocytes, cells from a leukemia cell line and blood dendritic cells, reviewed comprehensively by Botting et al., (2017). A lot of consideration needs to be taken when using these cells as it has been shown that they differ from LCs through genomic and functional studies (Harman et al., 2019). Other studies have investigated LCs using murine LCs. Even though these have shed light on LCs, there are some limitations with using murine models, including the fact that human tropic pathogens like HIV do not infect mice (Ahmed et al., 2015).

Two of the most popular methods of extracting LCs from tissue include enzyme digestion and spontaneous migration also termed the crawl assay. During enzyme digestion, cells are liberated out of the skin using proteases that digest the extracellular matrix, releasing cells. During spontaneous migration, skin epithelial sheets are placed on culture medium and the cells are allowed to spontaneously migrate out of the tissue into the culture media. The migratory cells are thought to be in a mature state as migration is coupled to maturity (Peña-Cruz et al., 2001). Therefore, spontaneous migration is not very suitable to study differences between mature and immature LCs as migratory cells are thought to be mature. Enzyme digestion, in such cases, can be used but the associated challenges need to be considered. For example, trypsin has been shown to cleave some receptors on cells, including the main HIV target receptor, CD4 (Nasr et al., 2014). Other enzymes such as collagenase and liberase have been used instead of trypsin (Artyomov et al., 2015). Therefore, the method of extraction is important and may depend on the questions of the study.

Role of Langerhans cells in HIV

Even though it has been proved that LCs are antigen-presenting cells, there has been a great deal of controversy on their role in HIV infection and transmission. There are currently three theories on the role of Langerhans cells during HIV acquisition i.e. transfection, infection or prevention as shown in Figure 1.8 and 1.9. Langerhans cells either act as carriers of the virus to T cells, consequently inducing infection (Ballweber et al., 2011), or they can get infected themselves, since LCs express CD4, an HIV target receptor (Nasr et al., 2014). Furthermore, Langerhans cells can also degrade the virus via the langerin pathway within the Birbeck granules, which restricts the virus, thus preventing infection ((De Witte et al., 2007).

When looking at preventative HIV strategies, studies have examined the events that occur during the viral acquisition phase leading to productive infection of the virus. Even though CD4+ T cells have been shown to be the main targets of HIV (Fahrbach et al., 2007, Cardinaud et al., 2018) it was unclear how the virus reaches these cells. Hladik's study is one of the first studies to show that the virus does not productively infect the Langerhans cells. The group used an *ex-vivo* model where they co-cultured infected epithelium of the vagina with purified Langerhans cells. There was no viral DNA in the cell genome. This suggests that there was no integration of the viral RNA, which shows no infection because viral DNA is found on infected cells. Therefore, it was concluded that there is no productive infection that occurs in LCs (Hladik et al., 2007).

This led to the hypothesis that LCs transport the virus to T cells without being infected. When LCs were co-cultured with lymphoblast derived from blood all the cultures were positive for the virus. This suggests that Langerhans cells can still transport the virus to the T cells even though they do not get infected themselves. The authors suggested that vaginal LCs can aid in protecting the virus against viral innate immunity through endocytosis as it has been shown that the LCs in the vaginal epithelia can endocytose the virus without progressive infection (Ballweber et al., 2011).

Some studies have reported that Langerhans cells respond differently to HIV as compared to other dendritic cells. De Witte (2007) showed that, at low viral concentrations similar to conditions during *in vivo* infection, T cells were infected in the presence of DC-SIGN+ DCs but not LCs (de Witte et al., 2007). At high viral concentrations, T cells alone were infected but there was a significant inhibition of the infection in the presence of LCs. This suggests that Langerhans cells do not enhance HIV infection as DC-SIGN+ DCs (Cardinaud et al., 2018; van den Berg et al., 2015). In these studies, it was proposed that the LC binds the virus using langerin allowing it to enter the Birbeck granules (De Witte et al., 2007). Other studies argue that LCs get infected by the virus since they express the HIV target receptors, CCR5 and CD4. Nasr *et al.* challenged De Witte's study by suggesting that the reason there were differences between Langerhans cells and other DCs was due to trypsin being used instead of collagenase to extract the cells. Trypsin is known to cleave the HIV target receptor CD4. Infection also depends on viral loads as there was no infection at low viral loads compared to high viral loads (Nasr et al., 2014).

The events that occur in the Birbeck granules during HIV restriction are still undefined. Only a few studies have investigated this pathway. Ahmed *et al.,* 2015 is one of the few studies that have shed light on the Langerin HIV degradation pathway. They elucidated the role of human E3-ubiquitin ligase tri-partite-containing motif 5 alpha (TRIM5-α). It is a host restriction factor against retroviral infection that binds to the viral capsule, which subsequently prevents uncoating and the reverse transcription process (Ahmed et al., 2015; Clayton et al., 2017). TRIM5- α prevents HIV infection in LCs but not in DC-SIGN DCs of the sub-epithelia (Ribeiro et al., 2016). To evaluate the role of TRIM5- α, it was knocked down using RNA interference and there was an increase in HIV transmission and infection of CD4+ T cells when compared to the control. TRIM5- α mediates the assembly of an autophagy-activating scaffold to langerin, which targets HIV-1 for autophagic degradation, thus preventing the infection of LCs. This mechanism is dependent on binding of langerin to the virus as it directs the virus to the TRIM5-α after the virus has been captured (Ribeiro et al., 2016).

Additional proteins, caveolin and clathrin have also been implicated in LCs HIV restriction through the langerin pathway. It has been shown that langerin and caveolin-1 co-localize at the cell membrane, intracellular vesicles and that Birbeck granules are langerin/caveolin-1-positive vesicles linked to the lysosomal degradation pathway in LCs. Furthermore, stopping caveolar endocytosis in primary LCs prevented HIV-1 sequestering into langerin+ caveolar structures. Both inhibition of caveolar uptake and silencing of caveolar structure protein resulted in increased HIV-1 integration and subsequent infection. In contrast, inhibition of clathrin-mediated endocytosis did not affect HIV-1 integration, even though HIV-1 uptake was decreased, suggesting that clathrin-mediated endocytosis is not involved in HIV-1 restriction in LCs (Berg et al., 2014)

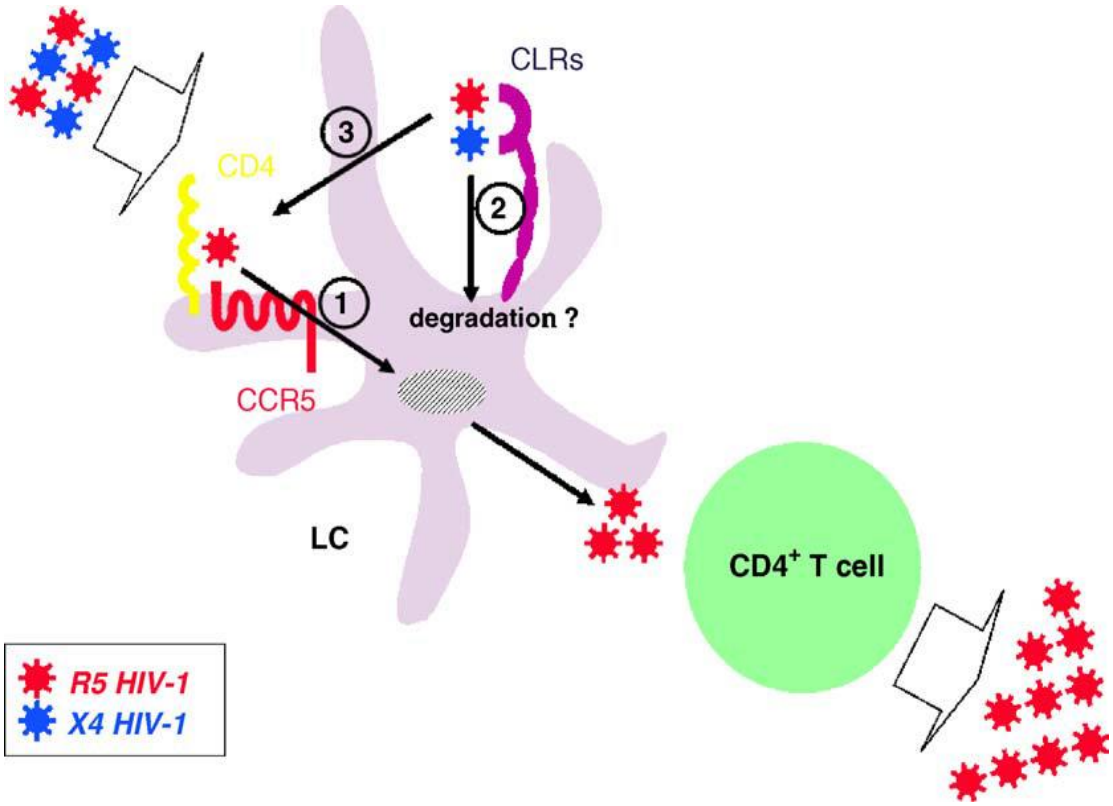


Figure 1.8: Role of Langerhans cells (LCs) in human immunodeficiency virus (HIV) transmission. The fate of HIV during LCs encounter is governed by several environmental conditions. There are three possible pathways upon HIV encounter with LCs. Infection, where LCs get infected by the virus as they express CD4 receptor as suggested by Nasr et al.,2014. The virus infects LCs via the CD4 receptor leading to LC infection. (pathway shown by arrow 1), transmission where the virus is transmitted to CD4+ T cells by the infected LCs during antigen presentation encouter or direct contact with mature virions (shown by arrow 3) or the virus strains are recognized and captured by C-lectin receptors (CLRs) such as langerin. The virus is then directed and degraded to Birbeck granules (shown by arrow 2). R5 and X4 HIV-1 are different HIV strains (Kawamura et al, 2005).

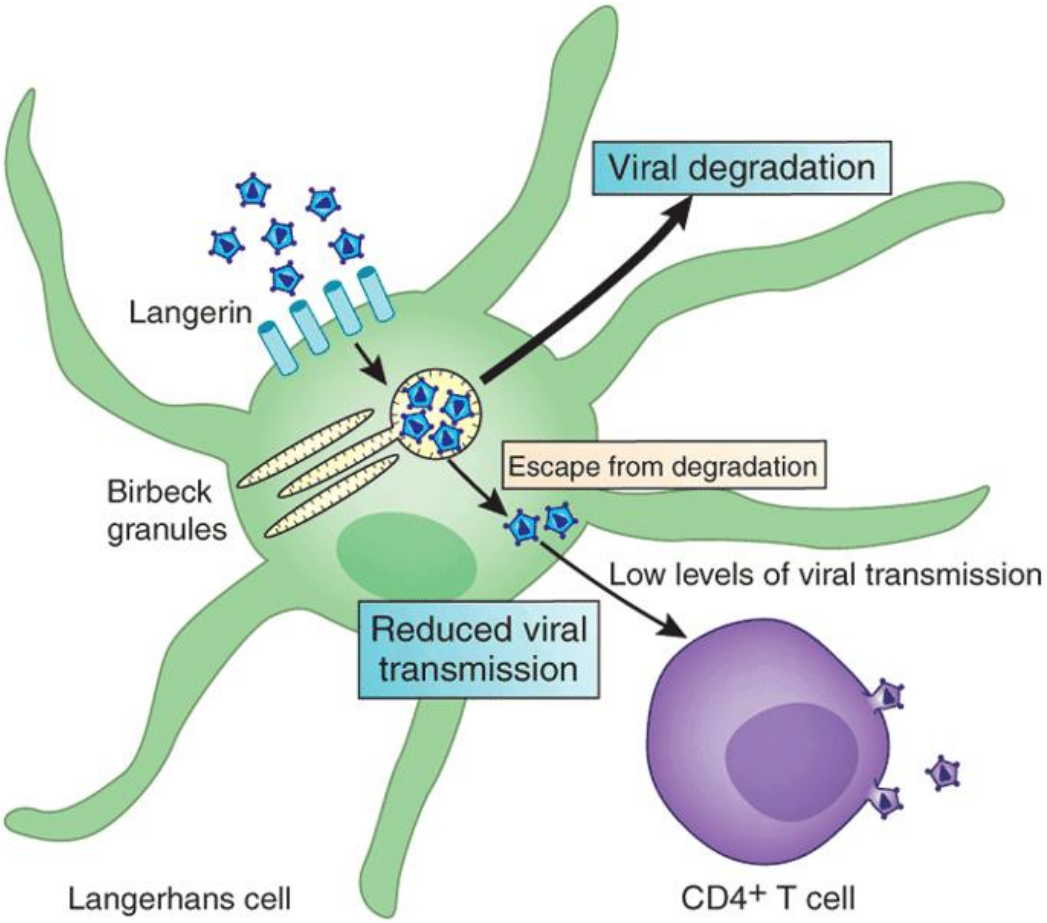


Figure 1.9: Human Immunodeficiency Virus (HIV) degradation pathway by Langerhans cells (LCs): If there is no immune activation upon pathogen encounter prior to HIV acquisition, HIV is likely to be captured by langerin and degraded in Birbeck granules, provided there is low viral load. At high viral loads, the virions may escape and multiply at low rates in the LCs. These may be transmitted to T cells during LC-T cell encounter such as antigen presentation. (Schwartz, 2007).

Impact of STI co-infection on the role of LCs on HIV

The interaction between HIV and immune target cells is influenced by inflammation and ulceration which is increased during STI infection (Freeman et al., 2006) Therefore, when there is co-infection of HIV and another STI, the rate of HIV infection will be increased (Ganor & Bomsel, 2011). Studies have shown that during co-infection of HIV and another STI, there is at least a 3 fold higher rate of HIV infection (Freeman et al., 2006). It has been demonstrated that other STIs account for over 30% of new HIV infections in Sub-Saharan Africa and there is a high frequency of individuals who have other STIs and are HIV positive compared to those who are HIV positive and negative for

other STIs (Freeman et al., 2006). Marsden and colleagues (2015) showed that infecting LCs and monocyte derived DCs with HSV-2 results in secretion of TNF-α by the infected cells. TNF-α is a cytokine that recruits HIV target cells and leads to immune activation, which are known susceptibility factors for HIV infection (Marsden et al., 2015)." Subsequently, there was an increase in CCR5 expression on the surface of target cells leading to increased productive infection of the HIV by 5 fold (Marsden et al., 2015).

STIs do not only lead to immune activation and inflammation, but they also affect the langerin-dependent degradation of the virus. The STI, Herpes Simplex Virus (HSV), was shown to compete with HIV for binding to langerin. After HSV uptake, langerin was down regulated (Marsden et al., 2015). Therefore, if langerin can act as a barrier against HIV, as reported by De Witte et al., (2007), LCs may lead to HIV infection by transferring it to T cells as the natural barrier langerin is bound by the HSV and down regulated.

The glans and urethral epithelium as sites of transmission

The urethra and the glans penis are one of the most targeted sites by STIs in men, including HIV-1 (Pudney & Anderson, 1995). Ideally, the urethra and the glans would be one of the most studied sites as they affect both circumcised and uncircumcised men. However, due to difficulties in obtaining the penile tissue, they are one of the least studied HIV susceptible penile sites (reviewed by Anderson et al., 2012). Even so, there are some studies that evaluated their role on HIV acquisition and transmission. Some of the studies used the penile tissue post mortem such as (Pudney & Anderson, 1995) and others examined explants post transgender operations (Fischetti, et al., (2015) and Ganor et al., (2013)). Through these studies, it was established that the glans and urethra are one of the penile sites that are enriched with HIV-1 target cells. In uncircumcised men, in a flaccid penis, the glans and urethra are covered by the inner foreskin and therefore consists of mucosal epithelium, which has been shown to be infected by HIV-1 in vitro (Fischetti et al., 2015). This area also consists of microflora, which is suspected to enhance infection by recruiting and activating HIV-1 target cells (Price et al., 2010). The urethra ranges from squamous, keratinized epithelia to less keratinized pseudo-stratified columnar epithelia from the opening to the end. The urethra has been shown to harbor LCs

throughout while other HIV-1 target cells such as macrophages have been found to be abundant at the columnar site of the urethra (Ganor et al., 2013). These findings suggest that the less keratinized area could potentially be the site of HIV-1 infection. However, (Pudney & Anderson, 1995) showed T lymphocytes (CD4+) to be spread out equally throughout the urethra. Therefore, it is unclear which site is more susceptible as these studies contradict. Another study found that there is an abundance of cells expressing both CCR5 and CXCR4 at the opening of the urethra and were susceptible to R5 HIV-1 strain (Fischetti et al., 2015).

It is clear that the glans penis and the urethra are susceptible to HIV-1 mostly due to the mucosal epithelium especially in an uncircumcised male. The removal of the FS during MMC minimizes the risk of infection by this site, the glans and the urethra. This is achieved by eliminating the inner FS mucosa which consists of HIV-1 target cells (McCoombe & Short, 2006), the microflora (Price et al., 2010) which recruits and activates target cells and mucosa which is a suitable environment for pathogen survival. This action leaves a dry, more keratinized epithelial surface. This condition provides resistance and protection against HIV-1 acquisition and infection (reviewed by Anderson et al., (2012)).

Rationale of the study

It is evident that the FS is susceptible to HIV (Auvert et al., 2005; Bailey et al., 2007; Gray et al., 2007). This is because a FS harbors different immune cells that are HIV target cells such as CD4+ macrophages, T, LCs, and DCs. However, there have been different reports on whether the inner FS or outer FS has more HIV target cells. Therefore, it is not known which part of the FS is more susceptible to HIV and it would also depend on the degree of keratinization, surface area or other factors. While LCs have been implicated in HIV transmission, their role remains controversial. Studies have shown that LCs can transmit the virus to T cells, which increases infection. On the contrary, others have reported that LCs prevent infection by degrading the virus through a langerin-dependent pathway. Their states of maturity and activation have a major effect in their role in HIV transmission. This is since activation leads to recruitment of LCs and other HIV target cells, up-regulation of HIV binding receptors such as CCR5 and down regulation of antigen capture. Co-infection of STIs and other immunological processes such as

secretion of pro-inflammatory cytokines like TNF-α influence this process. Therefore, studies of LCs activation may shed light on changes that occur during activation and their association with HIV infection. The lack of methods to extract and isolate pure LCs has caused paucity in the study of these cells. Due to these findings, the current study aims were:

Aims

1. The aim of the study was to evaluate cell extraction methods that yield more cells from the foreskin, isolate LCs and compare the cell numbers and viability between the inner FS and outer FS.
2. To characterize the LCs of the FS.
3. To investigate the differences in protein relative abundance between the inner FS and the outer FS.

Objectives

1. To compare cell yield and viability from epidermal FS cells between migratory assay and enzyme digestion.
2. To retrieve pure LCs population from the extracted FS cells by enrichment using density gradient centrifugation and isolate pure LCs by a flow cytometer cell sorter.
3. To characterize LCs by assessing changes in maturation and activation markers such as CD80/86, CD40 and HLA-DR through antibody staining and immunophenotyping using flow cytometry upon stimulation with TNF-α.
4. To identify the gene expression profile of sorted enzyme-digested LCs by looking at whole transcriptome using single cell RNA sequencing.
5. To compare and identify differences in protein relative abundance between the inner FS and outer FS by identifying the proteome of the epidermal inner FS and outer FS cells using LC-MS/MS.

Chapter 2

Methodology

Foreskin sample collection

Fresh FS specimens were obtained from different MMC sites in Cape Town operated by the Southern African Clothing and Textile Worker's Union (SACTWU) MMC program. The samples were collected from healthy adults within an ongoing study at the University of Cape Town. Written informed consent under Human Research Ethics Committee (HREC) number 566/2012 was obtained from each individual before specimen collection. Participant inclusion and exclusion criteria (Table 2.1), were applied before enrolment to the study. Plain Dulbecco's Modified Eagle's Medium (DMEM) stored at 4 ^{0}C in 50ml falcon tubes was used to collect samples from the MMC sites. The samples were transported to UCT, Division of Immunology Laboratory on ice, where FS dissection and cell isolation took place within 4 hours after the procedure was performed.

Table 2.1: Inclusion and exclusion criteria applied for the selection of participants for the study.

Criteria	**Selection**
HIV status	Negative
Other STI status	Not symptomatic
Age	18 and older
Consent	Signed consent form

Foreskin dissection and digestion with dispase

Trimming

FS samples obtained were separated into the inner FS and outer FS by the surgeon. Using serological pipette, 10ml of phosphate buffered saline (PBS) and 1% penicillin streptomycin (P/S) (PBS+P/S) was aliquoted into each half of a partitioned bacteriological petri dish (enough to rinse tissue). Using sterile forceps, either inner FS or outer FS tissue was placed into an aliquot of PBS+P/S in the petri dish for rinsing. After rinsing, the tissue was placed on a lid of a petri dish. Sterile forceps and fixed

scalpel blades were used to dissect the FS including trimming subcutaneous hypodermal tissue. The remaining tissue was cut into 0.5 cm X 0.5 cm pieces, and placed into 5ml or 8ml of 5mg/ml dispase enzyme solution prepared in Hank's buffered salt solution (HBSS) for inner FS and outer FS respectively, as shown in Figure 2.1. The skins were incubated overnight at 4^0C for 18 hours to allow separation of the dermis from the epidermis. After the incubation period was over, the epidermal sheets were stripped off the dermis with fine edged sterile forceps and rinsed in plain PBS to rinse off dispase and processed further.

Cell extraction

Migration assay

An 8ml and 5ml aliquot of Roswell Park Memorial Institute (RPMI) with 10% heat inactivated Fetal Bovine Serum (FBS) and 1% P/S (RPMI++) were measured into the 10cm and 6cm culture dishes for crawling outer FS and inner FS cells respectively. The rinsed epidermal sheets were transferred into the aliquot of RPMI++ ensuring that the epidermis was facing right way up and incubated at 37°C (5% CO_2) for 48 hours.

Spontaneously migrating cells were harvested from the media using a Pasteur pipette and transferred to a 50 mL Universal conical tube through a 70µm cell strainer. The cells were centrifuged at 400xg speed for 7 minutes at room temperature, and the cell pellet was resuspended in RPMI++ medium. The cells were counted using a haemocytometer either automatically using a TC20 (supplied by Bio Rad laboratories) or manually using a light microscope. The remaining sheets in the Petri dish were rinsed with PBS+P/S, and the remaining tissue digested by enzymes to release non-migratory cells as detailed in the following section.

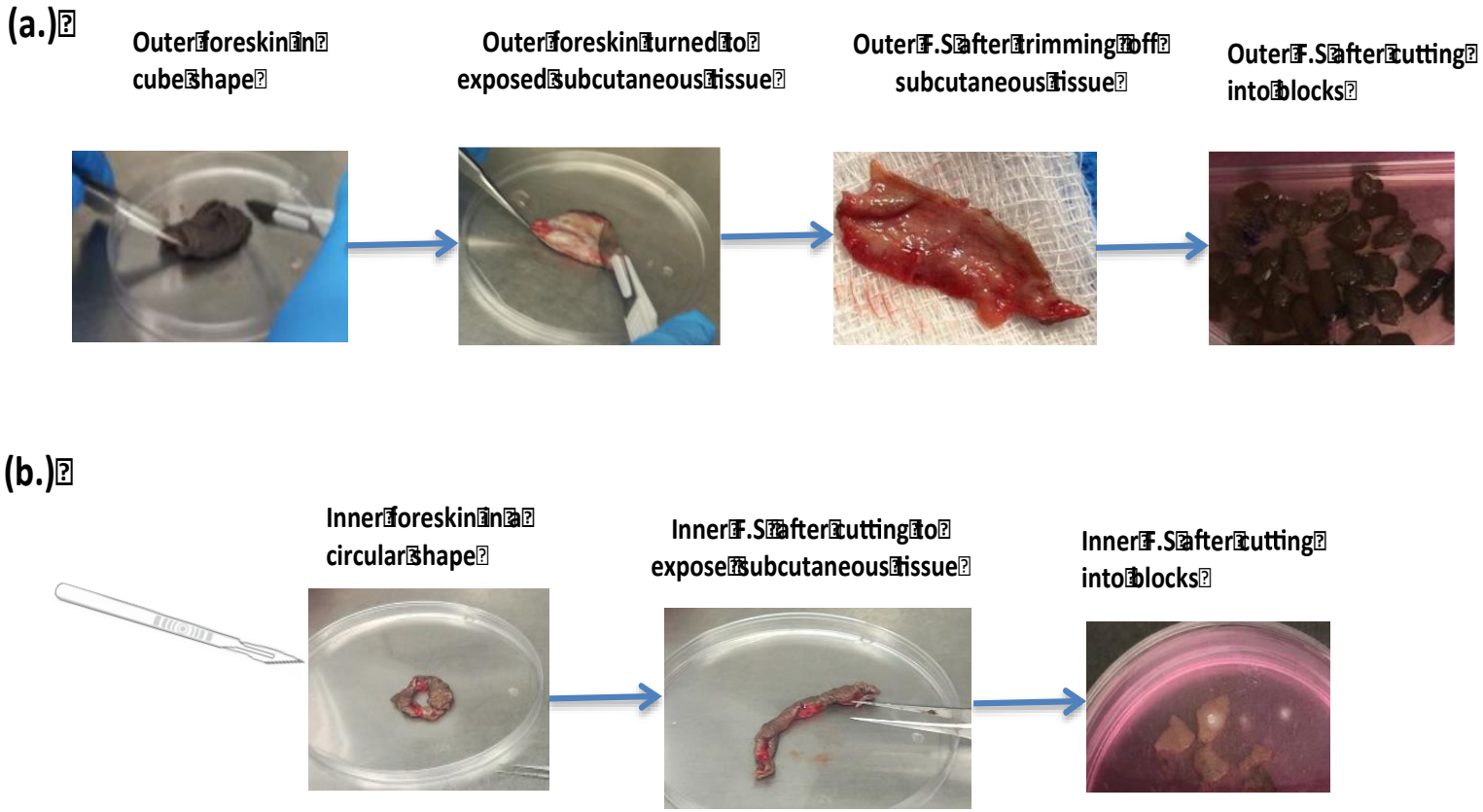


Figure 2.1: Epithelia sheet preparation. (a) Outer foreskin and (b) Inner foreskin tissues were rinsed and excess hypodermis was removed. Both the inner foreskin and the outer foreskin were cut into small blocks of 0.5cm X 0.5cm to prepare for dispase digestion.

Extraction of cells using enzymatic digestion

The remaining tissue after migration was digested using 25ug/ml of liberase enzyme (TH) (prepared in DMEM) to release tissue resident cells remaining in the epidermal sheets. The epidermis of inner FS and outer FS tissue were chopped into small pieces and transferred to a 50ml tube containing 3ml of liberase for inner FS and 5ml for outer FS. The tissue was incubated at 37°C for 3 hours to allow the liberase to digest the tissue and release cells into the medium (shaking gently on a rocker). Subsequently, the skins were gently triturated using a pre-wet 1ml pipette tip, and 0.5 mL of heat inactivated FBS was added into each tube to deactivate the liberase enzyme activity. The tissue suspension was transferred using a Pasteur pipette and filtered through a 70µm cell strainer into a 50ml tube. Cells were pelleted by centrifugation at 400 xg for 7 minutes and suspended in media. Cells were counted as previously described.

For liberase-digestion of epithelia sheets without crawl assay, liberase digestion of dispase prepared epithelia sheets was performed after washing off the dispase as stated above.

Enrichment of Langerhans cells using density gradient

Density gradient centrifugation of the extracted epidermal cell population was performed to enrich for LCs before their isolation using OptiPrep (Axis shield UK) shown in Figure 2.2. The density gradient was prepared with 2 solutions to make two layers. Solution 1, the underlying solution was made with 3 parts of plain PBS and 1 part of OptiPrep (25%) (3:1 ratio). Solution 2 was prepared by adding 2ml of a 1:1 RPMI++ and OptiPrep in a cell suspension of 3.3ml in RPMI++, (19% OptiPrep). Solution 2 was added on a universal tube then 5ml of solution 1 was under laid under solution 2 using a Pasteur pipette. An aliquot of 2ml RPMI++ was overlaid with a Pasteur pipette on top of the gradient. The prepared gradient was centrifuged at 600 xg for 15 min at 4^{o}C without brakes or acceleration. After centrifugation the cells were separated according to density, the cells with low density such as LCs float at the top and the ones with high density such as lymphocytes in the middle and dead cells at the bottom. The cells from the upper band and cells from the lower band were harvested separately and washed in RPMI++ and spun at 400 xg, for 7 minutes. The cells were counted then stained for flow cytometry as described below.

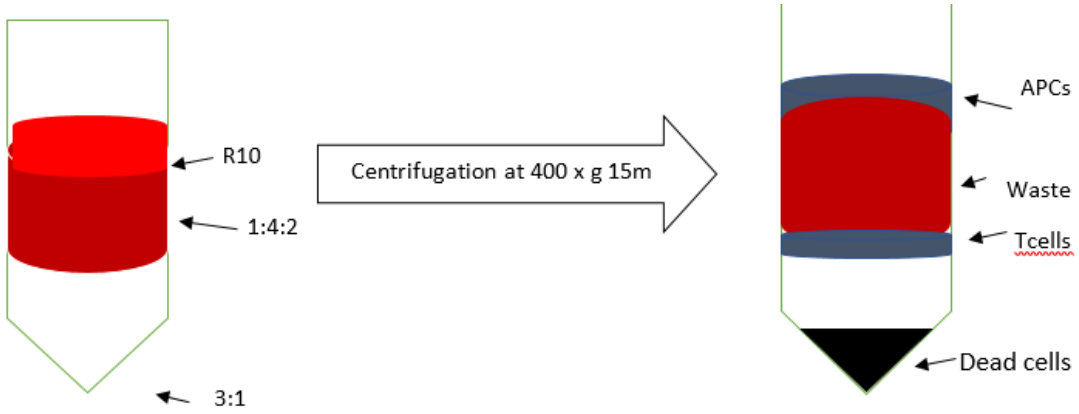


Figure 2.2: Cell enrichment with density gradient. A schematic presentation of density gradient with cell localization before and after centrifugation.

Cell flow cytometry

Flow cytometry was used to study protein expression of LCs (immunophenotyping) and separate/ isolate LCs by sorting. The first step was to identify the optimum volume of antibodies to be used. This was achieved by titrating the antibodies using the following method.

Antibody titration

After counting the cells, a total of one million epidermal FS cells were pipetted out from the cell suspension and transferred to a 96 well plate for staining. The plate was centrifuged at 400 xg for 3 min. The supernatant was discarded. All the centrifugation steps were carried at 400 xg for 3 minutes at room temperature (RT) for this section. The cells were washed by adding 250 µL of PBS buffer in each well, the plate was centrifuged and the supernatant was discarded. The cells were stained with 1μl of 1:40 vivid, which was used to distinguish between live and dead cells. After staining, the cells were incubated at RT for 20 minutes in the dark then centrifuged. The cells were rinsed by washing three times with PBS (centrifugation was done in between the washes). After the last wash, antibodies to be used for the titrations were prepared by serial dilution. The initial dilution was 1:10 for each antibody and then serial dilutions were made. Prior to addition of the antibodies, the cells were blocked with FACs buffer (PBS+1% FBS) for 20 min at RT then centrifuged and the supernatant was discarded. The required antibody volumes were added to the cells in FACs buffer. After adding the

antibodies, the total volume was adjusted to 50μl by adding FACs buffer. The cell/antibody solution was mixed gently using a pipette and incubated for 20 min at RT. The cells were washed with 200 µL of wash buffer, centrifuged and the supernatant was discarded. The wash was repeated with 250 µL of wash buffer and the supernatant was discarded. The cells were suspended in 200 µL of FACS fixing buffer containing 10% FACs lysis buffer prepared in distilled water then incubated for 10 min at RT. The cells were washed 3 times and the supernatant was discarded. The stained cell pellet was resuspended in FACs buffer and transferred to FACS tubes. Cells were acquired in a flow cytometer. In the case where the cells/ samples were going to be acquired later, the samples were stored at 4^{o}C in the dark until the time of flow cytometry analysis. The appropriate volumes of antibodies were used based on the analysis. The volume with the best separation and fluorescence was used, determined by a stain index.

Preparation of compensation controls

Compensation beads are single stained beads that are used to avoid overlap of the fluorochromes as they stain brighter and more uniform than cells. For this protocol, negative compensation beads (supplied by BD Science) were prepared by mixing through hand shaking adequately and ~1-2 drops were added in clean FACS tubes for unstained beads and for each marker in the panel. An aliquot of 1μl of antibody was added into FACs tubes and mixed by vortexing (separate tube for each antibody). The mixture was incubated at RT for 20 min, after the incubation period lapsed; 1 mL of wash buffer was added to wash the beads. The beads were then centrifuged at 400 xg for 3 min, after the wash step, the supernatant was discarded and the rim was blot dried using paper towel. Following that was the addition of 150μl of FACs buffer to the pellet. Then the beads were acquired with the samples using a Becton Dickinson LSRII flow cytometer, BD Science US.

Multicolor antibody staining

Cell staining was conducted using similar methodology to antibody titration with a few differences. The optimum antibody volume was used as identified by antibody titration for both sorting and immunoprofiling. For sorting, the cells were not fixed to avoid cell lysis as they were going to be used later and 0.5mg/ml of EDTA was used on the last wash to avoid cell clustering during sorting.

Flow cytometry assay for sorting cells

Cell sorting was used to isolate LCs from the epidermal FS cells extracted. To do the flow cytometry-based assay, the individual stained samples were analyzed on a Becton Dickinson FACSAria flow cytometer (BD Bioscience, US). The software used, was BD FACSDiva version 6.0. The flow cytometer has a three-laser system, i.e. blue laser (488nm), a red laser (633nm), and a violet laser (405nm) using a 100micron nozzle. The cell populations were identified and gated on a forward scatter (FSC) vs. side scatter (SSC) dot plot, using a linear scale, in acquisition mode. Doublet discrimination was obtained by using an FSC-area vs. FSC-width dot plot. A negative control was used to set up the PMT (photomultiplier tube) voltages as well as instrument optimization. Compensation of overlapping fluorochromes was achieved by using cells that are singly labeled with various fluorochromes. Dot plots of various fluorescent channels were used to enumerate various populations and sub populations. The following fluorochromes were used: CD1A-APC, CD3-PE, and CD207- PE-Cy7 (supplied by Biocom Africa). The combined parameters FSC, SSC, and fluorescent channels display the results. A threshold of 5000 on the FSC channels was set to remove sample debris. Langerhans cells were sorted as double positive CD207 and CD1a cells; CD3 was used to exclude the lymphocytes. A post sort was done to ensure purity of the cells; the cells had to be at least 85% CD1a+ and CD207+ to be considered pure Langerhans cells. Cells that fit the established criteria were sorted for further investigation, such as RNA sequencing.

Langerhans cell Immunophenotyping using multiparameter flow cytometry

Immunophenotyping was performed to evaluate the differences in the phenotype of LCs from the inner FS and the outer FS during immune activation and homeostatic condition. For this purpose, the medium for crawl assay as described above was aided with 0.25ug/mL of TNF-α to mimic immune activated environment, as it is a known cytokine that induces immune activation following the method of Berthier-vergnes et al., (2005) who showed that activation by TNF- α is best at 0.25ug/mL. The physiological range of TNF-α is 8.1pg/mL (0.0000081ug/mL) based on (G. Li et al., 2018) The current study used 0.2499919 ug/mL more TNF-α compared to the physiological range. The participant samples i.e. inner FS and outer FS; were split in half resulting into two samples per tissue (either inner FS or outer FS). One half was stimulated with TNF-α while the other half acted as a control. The cells were incubated for 48hrs in RPMI++

medium at 37^{0}C (5% CO_2) to allow for migration. After the incubation time lapsed the epidermis was chopped into pieces for enzyme digestion (as previously described) to release any tissue resident cells that did not migrate out of the tissue.

LSRII

To immunophenotype cells, a Becton Dickinson LSRII flow cytometer (BD biosciences, US) was used instead of cell sorter FACSAria. The following fluorochromes were used: CD1A-APC, CD40-PE, and CD207- PE-Cy7, Vivid-BV421 and HLA-DR-PerCP-Cy5.5 (supplied by Biocom Africa) CD80/86-FITC (supplied by BD science) shown in Table 2.2. The combined parameters FSC, SSC, and fluorescent channels display the results. A threshold of 5000 on the FSC channels was set to remove sample debris. After the parameters were applied, cell acquisition was done.

Table 2.2: Antibodies used and their fluorochromes.

Antibody	Fluorochrome
CD207	Pe-Cy7
CD1a	APC
CD40	PE
Cd80/86	FITC
HLA-DR	PerCp Cy5.5
Vivid	BV421

Fluorescent minus one (FMO)

FMOs were used as a spillover control. Multicolor panels may lead to fluorochromes spilling over or spreading into a channel of another fluorochrome leading to false positive cells. To control for this, all the antibodies in the panel were used except for the one in question. For example, if a spread of other antibodies on HLA-DR channel was measured, all the other antibodies were included except HLA-DR. Thereafter; the spread of these antibodies on the HLA-DR channel/gate was measured. Any fluorescence found on the HLA-DR gate from an HLA-DR negative sample was regarded as a false positive and subtracted from the population when calculating the number of positive cells.

Table 2.3: Layout of FMO control

Antibody	APC	FITC	PE	Pe-Cy7	PerCp Cy5.5
CD1a FMO	------	CD80/86	CD40	CD207	HLA-DR
CD80/86 FMO	CD1a	------	CD40	CD207	HLA-DR
CD40 FMO	CD1a	CD80/86	------	CD207	HLA-DR
CD207 FMO	CD1a	CD80/86	CD40	------	HLA-DR
HLA-DR FMO	CD1a	CD80/86	CD40	CD207	------

Data visualization and Statistics

FCS files were analyzed and visualized using FlowJo version 10.5.2 (https://www.flowjo.com). All the statistical analysis for this aim was done using GraphPad Prism 7 (https://www.graphpad.com/scientific-software/prism/), Pestle and Simplified Presentation of Incredibly Complex Evaluation (SPICE) tools version (5.3033) (https://niaid.github.io/spice/). *, ** and ***indicate $p<0.05$, $p<0.01$ and $p<0.001$, respectively.

RNA sequencing

Single cell RNA sequencing of sorted LCs using Seq-Well (a platform for single cell RNA sequencing) was performed in collaboration with Dr Musa Mhlanga at UCT. After cell sorting as previously described, Dr Yutaka Negishi prepared Seq-Well arrays by injecting Sygard 184 to the mould and baking at 70°C for 2.5 hr. Array surface was functionalized by chitosan (top surface) and poly (glutamic) acid (inside walls). Oligonucleotide conjugated beads and sorted LCs were loaded onto the array and sealed with semi-permeable membrane. The arrays were sub-merged in lysis buffer for cell lysis to occur and beads in the well captured RNAs. RNA-captured beads were collected and cDNA was synthesized by reverse transcriptase. The synthesized cDNA was amplified by PCR and then sequencing libraries were constructed by using Illumina Nextera XT kit. Sequencing library was sequenced by Illumina NextSeq 500 with high-output mode in paired end mode. Briefly, primer sequence and poly-A sequences in the second reads were trimmed and aligned to human genome hg38 using STAR version 2.2.5b. Uniquely mapped reads were grouped by cell barcode sequence and a list of

unique molecular identifiers (UMIs) in each gene was assembled (UMIs within edit distance = 1 were merged) to generate digital count matrix. After sequencing the library by the Illumina NextSeq 500, there were 543.9 million reads. Drop-seq tool version 1.12 was used to process the sequencing data from raw data to an expression measurement for each gene expressed by each cell

The cells with low abundant genes (with detected genes that were less than 200, or less than 2,500 numbers of detected unique molecular identifiers (UMIs)) were filtered out using R package 'Seurat'. Furthermore, the cells with a high ratio of mitochondrial transcripts (> 0.25%) were removed. Data from 827 cells passed the above quality control step and were subject to further gene expression analysis. Gene count data was normalized with the total number of transcripts, multiplied by 10,000 and log-transformed to generate a gene expression matrix. Lowly expressed genes (mean expression is less than 0.125units) were removed and 2,854 variably expressed genes (dispersion > 0.5) were identified and subjected to principal component analysis. To visualize data, tSNE analysis was performed with first 16 principal components and the dimension was reduced to two.

Proteomics

To differentiate between the inner FS and the outer FS, protein relative abundance from epidermal migrated FS cells was investigated. To achieve this aim, two different protein extraction and digestion methods i.e. Acetone protein precipitation coupled with in-solution digest (Borchert et al 2012) and sodium dodecyl sulfate (SDS) protein extraction coupled with filter aided sample preparation (FASP) were used (Zhou et al., 2013). Generally, both methods involve extraction of protein either from cells or tissue, digestion of protein to peptides, identification of protein using mass spectrometry and data analysis as shown in Figure 2.3. For this section, migratory epidermal FS cells were used. The cells were cryopreserved in freezing medium containing 90% FBS and 10% dimethyl sulfoxide (DMSO) at -80^{0}C after migration.

Radioimmunoprecipitation assay (RIPA) protein extraction

One million cells were lysed by adding 150μl lysing solution for protein extraction. The lysing solution was prepared by mixing 4% SDS, 100mM Tris/HCl, pH 7.6 and 0.1M DTT. The cells were incubated at 95°C for 3 minutes after the incubation period lapsed, the resulting DNA was sheared by sonication to reduce sample viscosity. Important to note is that, before adding the lysate to the cells, the lysate was clarified by centrifugation at 16,000 xg for 5 min to allow for protein separation. After, the protein was cleaned and digested using FASP (Zhou et al., 2013).

Protein digestion

Filter aided sample preparation

To digest the protein 200μg of protein extract was transferred into a 500 µL of Ultracel 30 000 MWCO centrifugal unit (Amicon Ultra, Merck) in triplicates. Protein extract was washed with three rounds of 200 µL of UT buffer (8 M urea, 0.1 M Tris, pH 8.5). Subsequently, alkylation of reduced cysteine bonds was carried out by incubation in the dark for 20 min in 200 µL of UT buffer containing 0.05 M iodacetamide (Sigma). Two washes of 200 µL UT were used to remove the alkylating agent, followed by three buffer exchanges with 100 µL of ABC buffer (50 mM ammonium bicarbonate buffer (ABC), pH 8). 40 µL of ABC buffer containing a volume of 1:50 ratio of sequencing-grade modified trypsin (Promega) to the amount of protein used was added for proteolysis. Thereafter, proteolysis was carried out at 37°C for 18 hr in a humidified chamber. After 18 hours, the peptides were washed three times with 40µL of ABC buffer to elute the peptide-rich solution. For this protocol, buffer exchanges were carried out by centrifugation at 14 000 xg for 15 min. Then the salts were removed from the peptides through desalting method described below.

Acetone protein precipitation

Previously frozen epidermal migrated FS cells as described above were thawed on ice in 4 volumes of extraction buffer [4M Guanidine HCl, 100mM NaCl, 5mM TCEP (tris-carboxyethyl phosphine), 2mM EDTA, 1% OGP (octylglucopyranoside) in 100mM TEAB (triethelammonium bicarbonate)], together with glass beads. Once thawed the samples were vortexed for one minute before placing them back on ice. The samples were sonicated for 30s and the process was repeated three times. The samples were centrifuged for 10min at 12 000 xg, after centrifugation the supernatant was pipetted

into 2ml tubes and stored on ice. The process was repeated 3 times, each time pipetting out the supernatant. The proteins were precipitated from the solution with 5 volumes of ice-cold acetone overnight at -20^{0}C.

Insolution protein digestion

The resulting protein precipitate was resuspended in 4M guanidine and 100mM TEAB. Then the protein was quantified using Pierce Bicinchoninic Acid (BCA) assay (kit supplied by Thermo Fischer Scientific), after quantification 100ug sample was used. The volume was reduced in a speed vac to approximately 5 µl. The sample volume was then adjusted to 10µl with 50mM TEAB. A concentration of 0.5 mM TCEP was added for reduction and samples incubated for 1h at 37^{o}C, 1 μL of 100mM methyl methane thiosulfate (MMTS) prepared in 2-propanol was added to a final concentration of 0,1 mM to the cooled samples. The samples were incubated at room temperature for 30 minutes. After which the volume was adjusted with 5 mM TEAB to 95 µl, 5 µl of trypsin (1:20) was added in each sample. All samples were vortexed and wrapped with parafilm to prevent evaporation then incubated for 18 hours at 37˚C for protein digestion into peptides to occur. The resulting peptides were dried using a SpeedVac and resuspended in 30µl of 0.1% formic acid (FA). The salts were removed through desalting method described below.

Desalting/Protein clean-up

Peptide eluate obtained from either FASP or in-solution was desalted using a homemade stage tip containing Empore Octadecyl C18 solid-phase extraction disk (Supelco) (supplied by sigma Aldrich). For this protocol; activation, equilibration, peptide wash and elution was carried out using centrifugation at 5000 xg for 5 min. Activation and equilibration of the C18 disk was carried out using three rinses with 80% acetonitrile (ACN), followed by three rinses with 2% ACN, respectively. Peptide solution was then loaded onto the disk and centrifuged. Desalting was done by washing the peptides three times with 2% ACN, followed by additional three washes with 2% ACN containing 0.1% formic acid. After the final wash step the desalted peptides were eluted into glass capillary tubes, this was done with three rounds of 100 µL of 60% ACN, 0.1% formic acid. Peptides were dried in a vacuum and resuspended in a solution containing 2% ACN and 0.1% formic acid with a concentration of 250 ng/µL. The peptides were then

analyzed using Liquid Chromatography with tandem Mass Spectrometry LC-MS/MS (method described below).

Liquid chromatography tandem mass spectrometry (LC-MS/MS)

LC–MS/MS Analyses Liquid chromatography separation was performed using a home packed 100 µM ID × 20 mm precolumn connected to a 75 Mm × 500 mm analytical column packed with C18 Luna beads (5 µm diameter, 100 Å pore size; Phenomenex 04A-5452). The columns were connected to an Ultimate 3500 RS nano UPLC system (Dionex). One microgram of desalted peptides was loaded onto the column with starting mobile phase of 2% ACN, 0.1% formic acid. Peptides were eluted with the following gradient of 10 min at 2% ACN, increase to 25% ACN for 115 min, to 35% ACN over 5 min, to 80% ACN over 5 min, followed by a column wash of 85% for 20 min. Mass spectra were acquired by the Central Analytical Facilities at Stellenbosch medical school, with an Orbitrap Q-Exactive mass spectrometer (Thermo Fischer Scientific) in a data-dependent manner, with automatic switching between MS and MS/MS scans using a top-10 method.

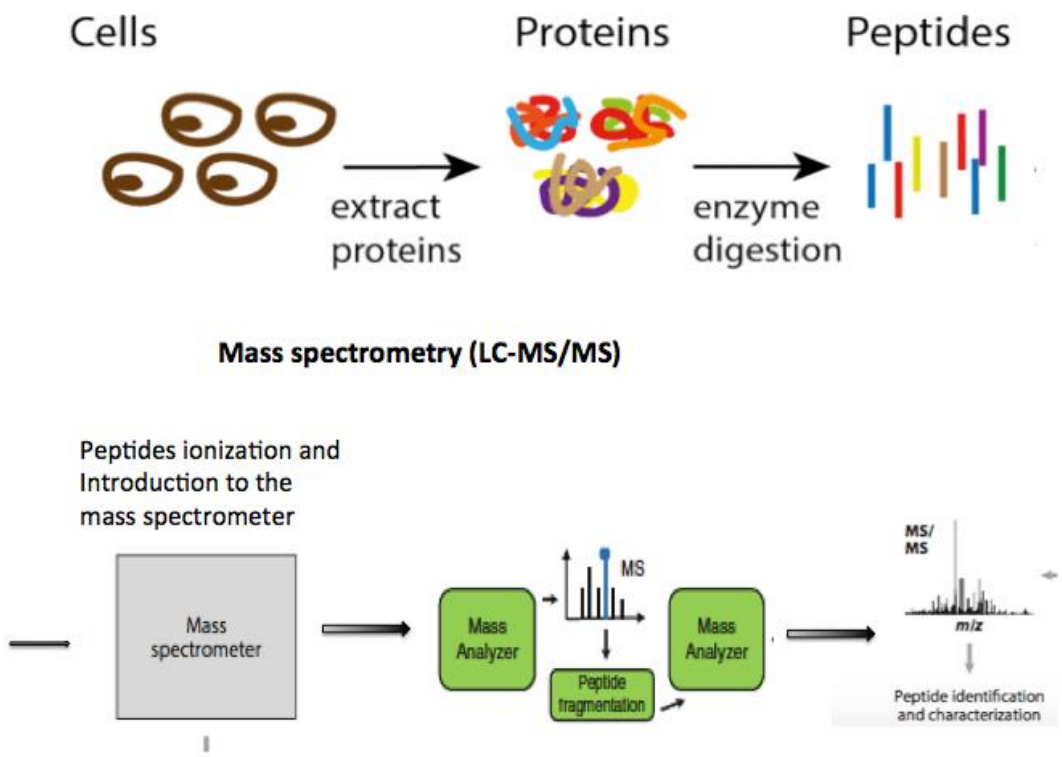


Figure 2.3 Workflow of proteomics using LC/MS-MS: Protein is extracted from cells and digested into peptides. The peptides are quantified and identified by mass spectrometry.

Protein Identification and Quantitation

Raw data files from the Q-Exactive were processed in MaxQuant15 version 1.5.0.3 (https://www.maxquant.org). The complete human proteome from UniProt was used to define the search space for the built-in Andromeda search engine. The mass and intensity of the detected peptides were processed using default MaxQuant settings with a few changes (Giddey et al., 2017). Results were filtered by a limit of at-most two missed cleavages and a false discovery rate (FDR) of 1%, estimated using the reversed proteome in a target-decoy approach, used to restrict identifications at both the peptide spectrum matching and protein inference levels. Protein inference required at least one unique or razor peptide for identification of a protein group. The label free quantitation (LFQ) was enabled through the Maxquant LFQ algorithm.

Data analysis and Statistical Treatment

Perseus (**www.perseus-framework.org**) was used to differentiate protein expression between the inner FS and the outer FS using protein groups identified by MaxQuant. LFQ intensity was used as expression column. Contaminating and reverse proteins were removed then the data was log transformed into log2(x) to reduce skewness. The data was filtered for valid values where one-quantification events were removed (at least 3 valid values in one group). Statistical analysis was done using a Student's t-test on the LQF values in triplicate measures (at least three samples per group) from the inner FS and the outer FS. Proteins that were completely absent in one group while present in the other (either inner FS or outer FS) were considered dysregulated. A Cluego plug-in on Cytoscape version 3.6.1 (https://cytoscape.org) was used for biomolecular network analysis of the dysregulated proteins and STRING (https://string-db.org) was used for protein-protein interaction analysis of proteins that were dysregulated between the inner FS and the outer FS identified by perseus.

Chapter 3

Results

Cell extraction and immunophenotyping

Spontaneous cell migration and liberase enzyme digestion yield similar numbers of FS cells

FS epidermal sheets were prepared by digesting 1-2cm^2 (1cmx1cm -2cmx1cm) pieces of whole FS tissue with dispase enzyme. Dispase cleaves cellular junctions of the basal membrane allowing easy detachment of the epidermal sheet from the dermis (Peña-Cruz et al., 2001). The epidermal cells were extracted using migration assay (crawl) where cells migrate from the skin into the media or enzyme digestion where the epidermis is digested by an enzyme to liberate the cells. In order to determine which extraction method was better, the yield and viability of all cells isolated after spontaneous migration of FS epidermal sheets and all those extracted by enzymatic digestion with liberase were compared. From the FS of 21 men, an average of 167x10^6 (SD=26x10^6) cells from each outer FS using the spontaneous migration method was isolated. Similarly, from another set of 21 outer FS, 181x10^6 (SD=21x10^6) cells were isolated following liberase digestion (Figure 3.1.1). There was a higher cell yield when using enzyme digestion by 4% compared to migration assay. A Student's *t*-test was used to assess if the differences were significant. No statistical significance was found between the two isolation methods (p=0.06). When assessing viability rate, enzyme digestion had an average of 44.67% (SD=23) proportion of live cells while migration assay had an average of 57.64% (SD=16.64) proportion of live cells. A Student's t-test was used to assess the difference. There was a significant difference when comparing viability between the two extraction methods, p-value = 0.04 as shown in Figure 3.1.2.

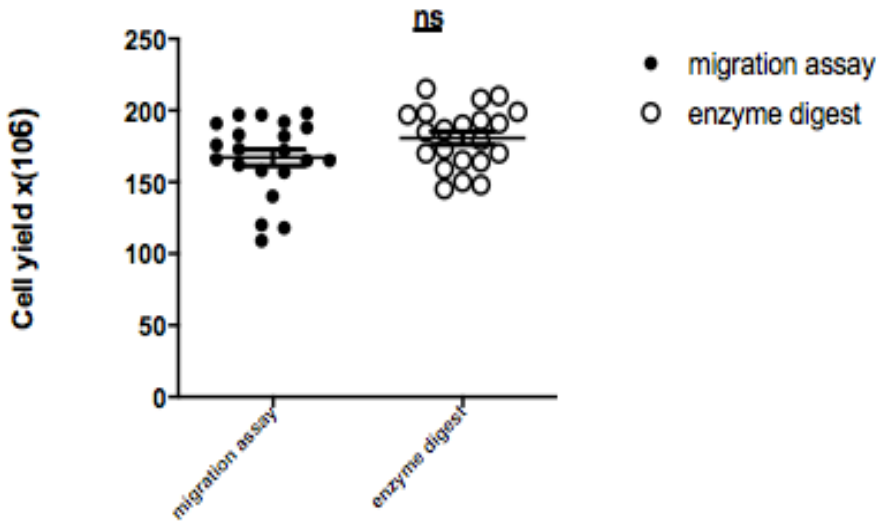


Figure 3.1.1 There was no significant difference in cell yield between migratory assay and enzyme digestion of outer foreskin (FS) epithelia: A statistically non-significant difference in cell yield between the 2 methods was observed; enzyme digestion compared to migratory assay, p=0.0556. Black circles represent the cell yield obtained following spontaneous migration with an average of 167x10^6 cells (SD=26x10^6) and white circles represent the cell yield obtained following liberase enzyme digestion with an average of 181x10^6 cells (SD=21x10^6). The comparison was between all the cell types that migrate and those that were enzymatically digested from the FS epidermis. Total number of participants used in this analysis was 21 for each extraction method.

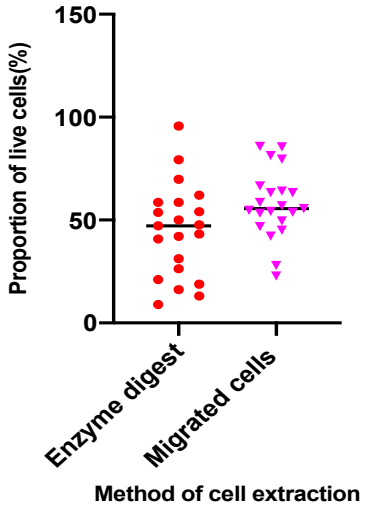


Figure 3.1.2 Cells extracted using migration assay are more viable than cells extracted using enzyme digestion. There was a significant difference in proportion of viable cells between enzyme digestion with an average of 44.67% (SD=23) and migration with an average of 57.65% (SD=16.64). p-value = 0.04, n=21 for each method.

When the cell yields between all migrated epidermal inner FS cells and migrated epidermal outer FS cells were compared, a significant difference between them was observed from 5 donors as depicted in Figure 3.1.3a. The outer FS had a significantly higher cell yield compared to the inner FS. There was an average of $148x10^6$ ($SD=32.3x10^6$) from the outer FS and $50x10^6$ ($SD=13.89x10^6$) from the inner FS. However, this was expected as the tissues differ in size. The outer FS is approximately three times bigger in surface area than the inner FS as shown in Figure 3.1.3b depicting a representative FS tissue. The Figure is a photograph of the inner FS (less pigmented) attached to the outer FS. A line has been drawn to show the demarcation between the two, and demonstrates the approximate three-fold difference in size. There were variations among participants that were due to differences in FS sizes.

A B

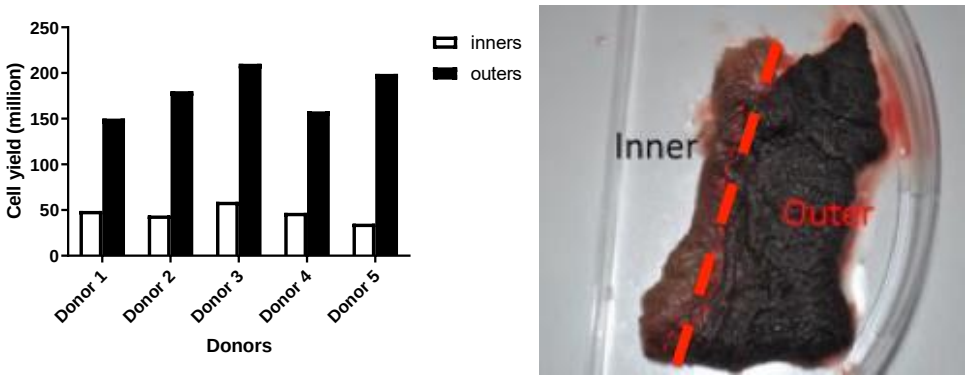


Figure 3.1.3 The outer foereskin (FS) has higher cell yield than the inner FS and the inner FS has a smaller surface area. (A.) There was an average of $148x10^6$ cells ($SD=32.3x10^6$) in total that were found in the outer FS per participant while there were $50x10^6$ ($SD=13.89x10^6$) cells in average obtained from the inner FS using the cell migration isolation method. There was a significant difference in the cell yield of outer FS vs. inner FS after statistical analysis using a Student's t-test, p-value=0.0001 (black bar is the outer FS, white represents the inner FS) n=5. (B)The FS tissue, inner FS is approximately 3 times smaller than the outer FS in surface area (separated by the dotted line). All cell types that migrate out of the epidermis were included for this comparison.

Density gradient centrifugation using OptiPrep enriches Langerhans cell population

The study investigated the enrichment of FS DCs from total epithelial cells obtained from enzyme digested and migrated FS tissue. Human epithelia tissue contains a diverse population of cells including keratinocytes, melanocytes and fibroblasts (Salmon & Ansel, 1994). Dendritic LCs are a minority subset making up 1-3% and we sought to exclude structural epithelial cells using density centrifugation. The cells were enriched using density gradient centrifugation where APCs are separated from other cells according to differences in their density as shown in the methodology (Figure 2.2). APC's, which are larger and have a bigger surface area, collect in a fraction above lymphocytes cell layer, which are smaller and denser. The enriched cells were analyzed by a flow cytometer to compare the frequencies of live LCs obtained from the respective fractions from 6 donors. After density gradient centrifugation using OptiPrep (a density gradient medium), there was a 15-25% yield of LCs compared to 0-1% before enrichment. LC population was enriched to an average of 20%. The results of this analysis are shown in Figure 3.1.4 below.

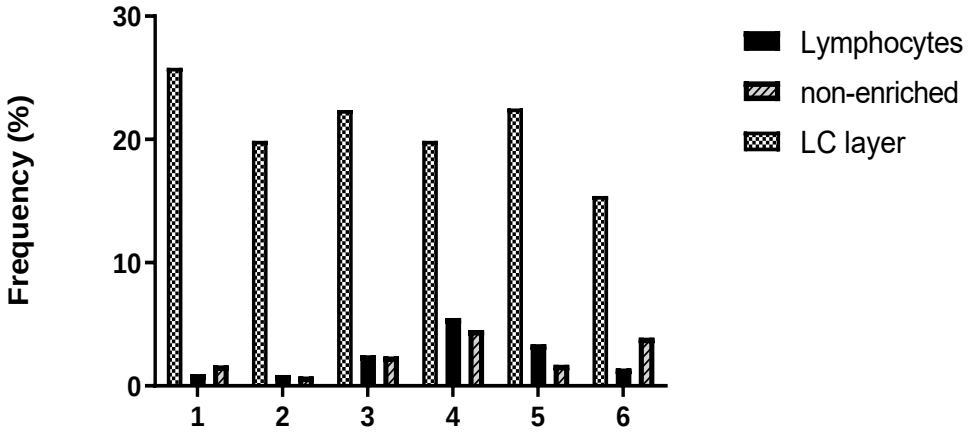


Figure 3.1.4 Cell enrichment by density gradient yields high percentage of Langerhans cell (LCs) population: Before enrichment by density gradient, 0.5-3% of the cell population were LCs shown by the bars with lines. After enrichment, 15-25% of the enriched cells were LCs from the LC-layer shown by the bars with squares. The lymphocytes layer had 0.1-5% LCs after centrifugation shown by the black bars. Y-axis is the frequency of LCs and X-axis shows the donors 1 -6. The key of the graph is shown on the top right. The LC layer is the layer of APCs after enrichment by centrifugation while the Lymphocytes layer is T-cells and potentially other undefined cells after enrichment by centrifugation. Before enrichment are the LCs from the pool of FS cells before enrichment by centrifugation. n=6.

Multiparameter flow cytometric analysis of FS LCs populations

A. Determination of antibody quantities for flow cytometry analysis

One of the study aims was to isolate pure LCs by sorting and better understand the biology of FS LCs by investigating the expression of activation and maturity markers of spontaneous migrating and liberase digested cells using flow cytometry. This was achieved by first titrating several LC-specific antibodies. This was done in order to determine the correct volume of antibody to use for each sample. The staining index, which is a way of measuring brightness of fluorochromes across different volumes of antibody in a sample, was used to determine the optimum volume, with the most optimal fluorescence and separation of the positive cells and negative cells as shown in Figure 3.1.5 and on appendices in Figure A1. The volumes of the titres used are summarized on Table 3.1 below.

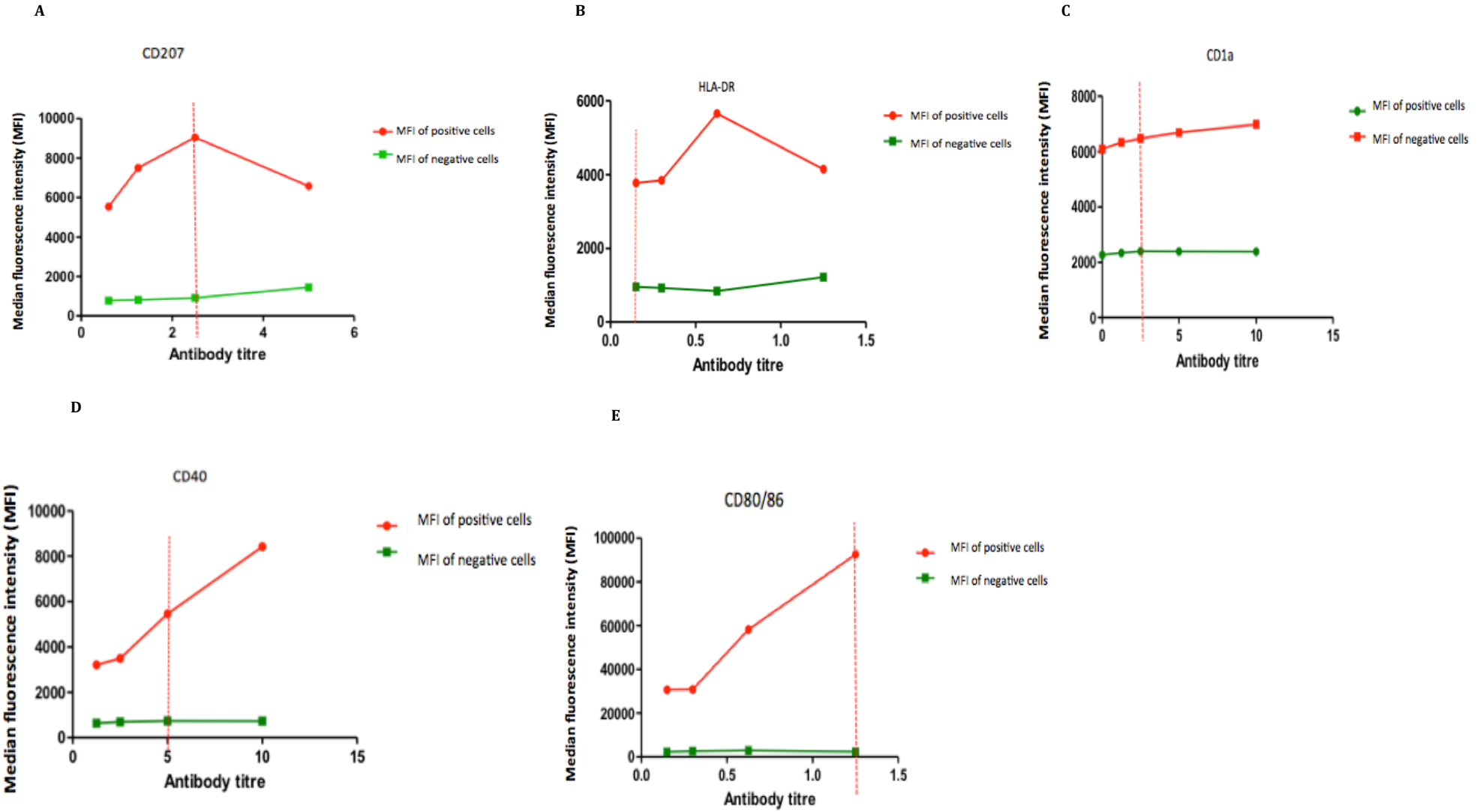


Figure3.1.5: Antibody titration for flow cytometry: The antibody titres used for immunophenotyping Langerhans cells CD207 (A), HLA-DR(B), CD1a (C), CD40 (D), CD80/86 (E). Red is the median fluorescence intensity of the positive population while green is the median fluorescence intensity of the negative population. The optimal titres are shown by the dotted red line.

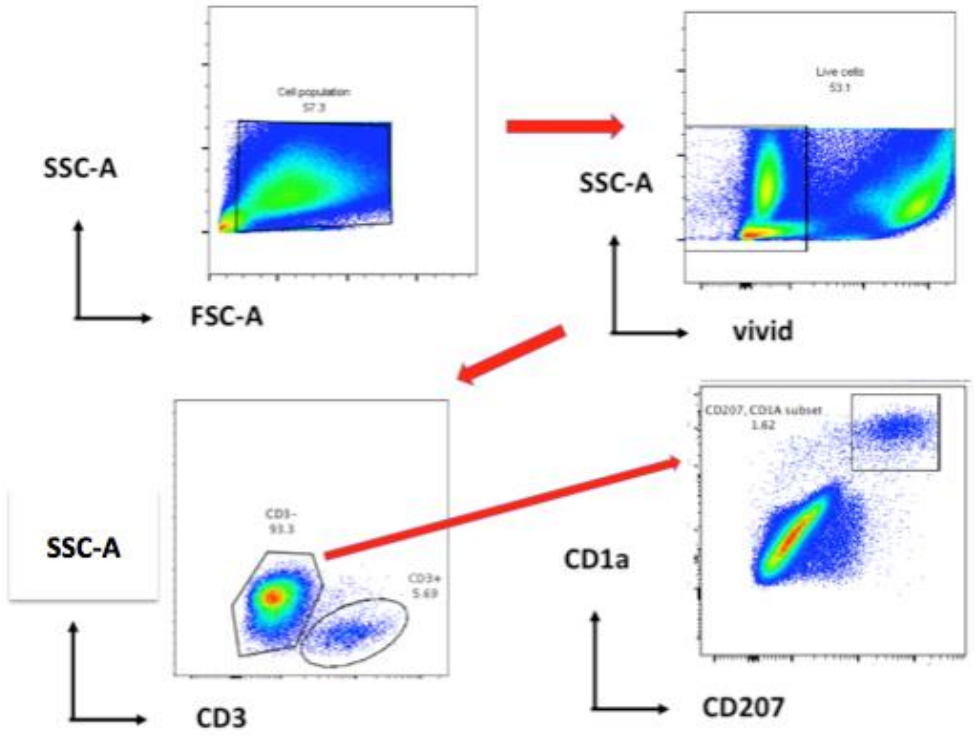
Table 3.1: Summary of antibody titers used during flow cytometry analysis of Langerhans cells

Antibody	Antibody concentration (μl)
CD207	2.5
CD1a	2.5
HLA-DR	0.15
CD40	5
CD80/86	1.25
Vivid	1

B. LC isolation by Fluorescence-activated cell sorting (FACS)

Cell enrichment of antigen-presenting cells (APCs) results in a heterogeneous population, including cells unwanted. As a result, FACS sorting was used to obtain a pure LC population. The enriched cells were stained with LC-specific fluorophore-conjugated antibodies against CD1a and CD207 and the cells were sorted using a fluorescence-activated cell sorter (FACS). The cells were also stained with an antibody against CD3 to exclude T cell contamination during sorting as shown in Figure 3.1.6a. To confirm if the sorted cells were pure LCs; the sorted cells were acquired in the cytometer to check the percentage of LCs (a post-sort). The cells were 85-98 percent pure, as assessed by CD1a and CD207 double positive cells shown in Figure 3.1.6b determined from flow cytometry of post-sorted cells.

A



B

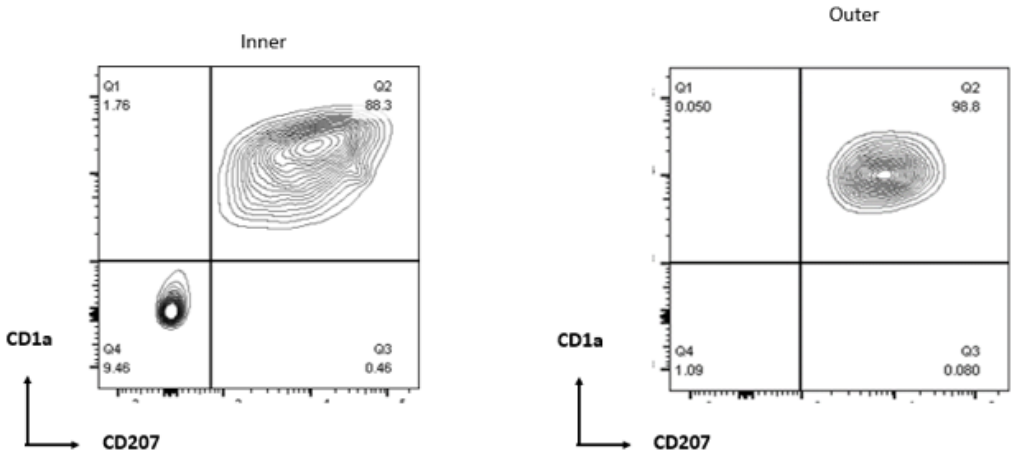


Figure 3.1.6 Isolation of Langerhans cells (LCs) by sorting yields a pure population of Langerhans cells. (A) Gating strategy used for sorting LCs. The dendritic cell population was identified and gated based on their characteristic forward vs side scatter. A live/dead gate to select live cells. CD3 positive cells were excluded from the live cells then CD1a+/CD207+ cells (LCs) were gated from the CD3- population. (B) A post-sort from inner FS cells and outer FS cells. High proportion of the cells were LCs (CD1a+, CD207+) after sorting with a percentage purity above 85%.

C. Quantification of Langerhans cells of the inner and outer aspects of the FS

Inner FS and Outer FS yielded equal frequencies of live LCs

There are contradicting results in literature regarding which aspect of the FS contains more Langerhans cells (Liu et al., 2014; Morris & Wamai, 2012). Due to these contradictions, it is difficult to predict which part plays a more significant role between the inner FS and the outer FS during the removal of the FS in HIV acquisition and infection with regards to LCs as quantity of cells is one of the factors that play a role in HIV infection. The part with more HIV target cells would potentially be more susceptible to HIV. Therefore, the study evaluated the differences in the density of Langerhans cells between the inner FS and outer FS. FS cells were extracted and stained with LC-specific antibodies to isolate LC as previously described. For this section, LC population was determined using the gating strategy illustrated in Figure 3.1.7. The position of positively staining cell gates for each antibody was determined using FMO as explained in the Methodology and shown in Appendices Figure A2.

There was a difference of 0.1% between the frequency of live LCs in the inner FS and outer FS tissues from 9 participants. The outer FS possessed more LCs, but it was not statistically significant, (mean: 0.56% (SD=0.37) vs. 0.68% (SD=0.79), p-value=0.29) after a Student's t-test on cells extracted using migration assay and 0.2% on cells extracted using enzyme digestion after migration (mean: 0.28% (SD=0.18) vs. 0.45% (SD=0.21), p-value=0.08, n=9). The study also compared the LC yield between migratory and enzyme digested cells post migration to determine whether all LCs migrate out of the tissue or whether there are some skin-resident LCs following migration. After spontaneous migration, CD1a+CD207+ cells from remaining enzyme-digested epithelial sheets were observed. There was a difference of 0.2% in frequency of LCs found during migration and enzyme digestion after migration with a p-value of 0.01, which was statistically significant (depicted in Figure 3.1.8). This suggests that the majority of the LCs were migratory as more LCs were isolated from the media than the tissue following the crawl assay.

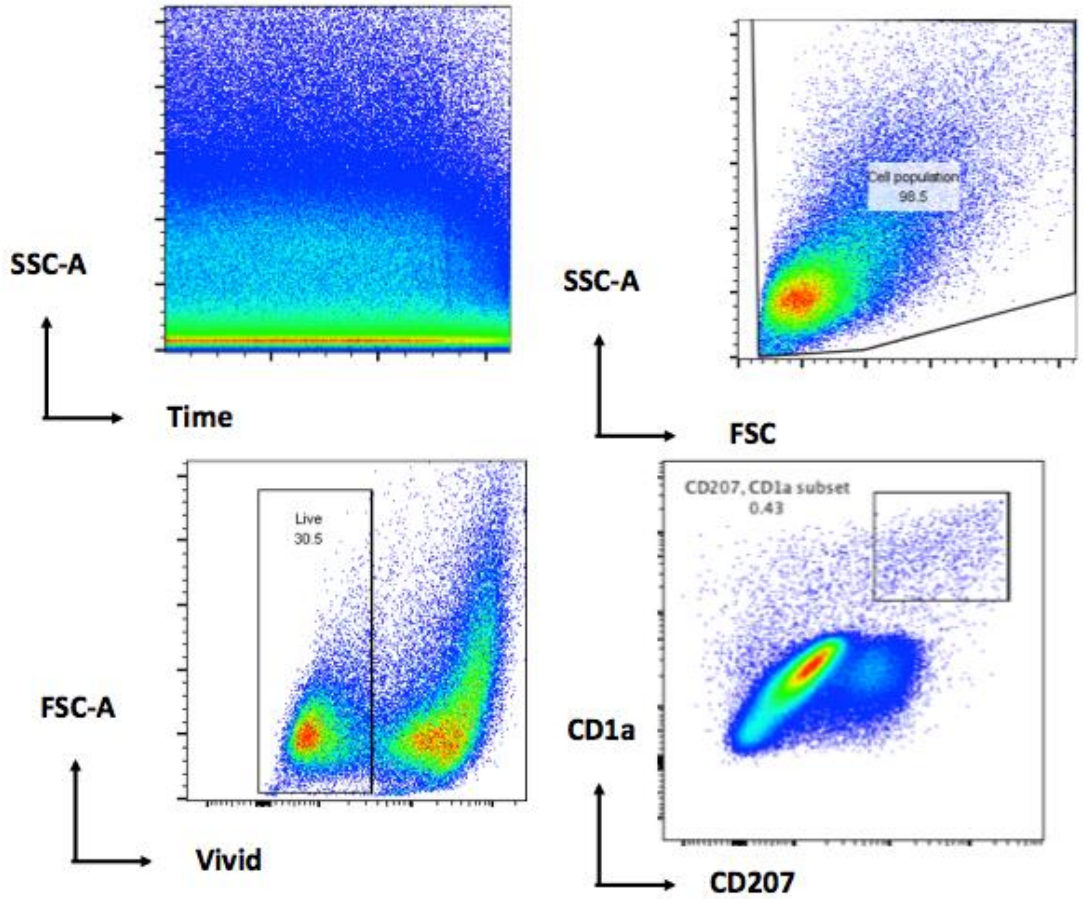


Figure 3.1.7 Gating strategy to identify Langerhans cells: (1) The cells were first gated from the time gate to assess errors during acquisition. (2) Then a forward and side scatter to include the granular LCs. (3) A live/dead gate was used to exclude dead cells. (4) LCs were gated from the live population and identified using CD207+, CD1a+ cells.

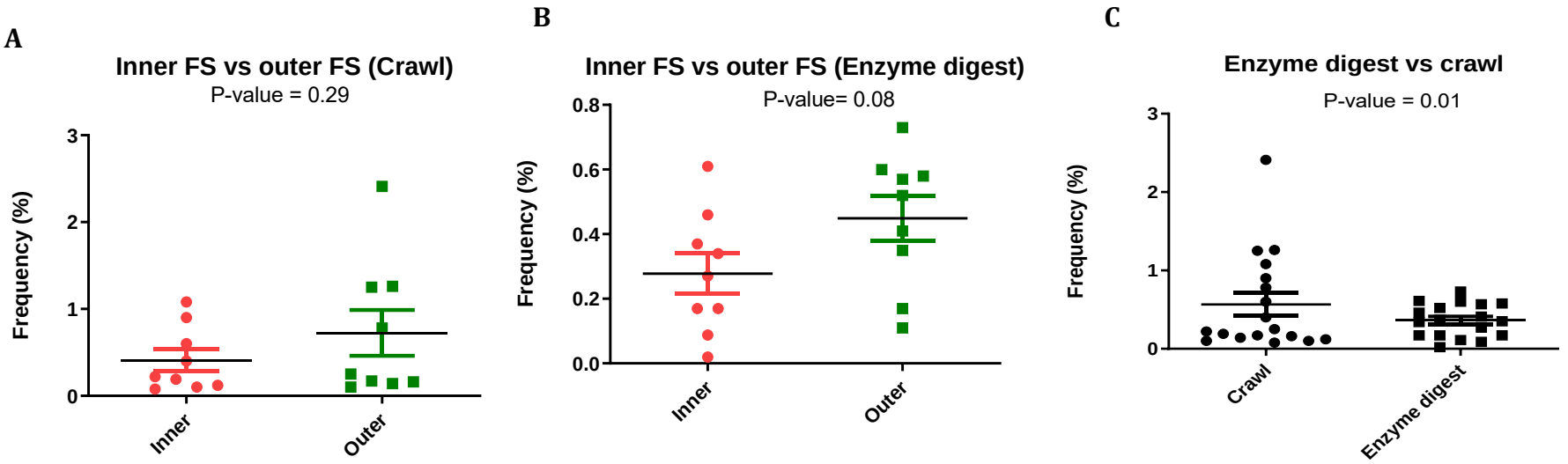


Figure 3.1.8 Differences in frequency of Langerhans cells (LCs) between the inner FS and the outer FS and migration versus enzyme digestion after migration. Frequency of LCs from the inner FS (red) and the outer FS (green). (A) there was a difference of 0.1% between the inner FS and the outer FS (mean: 0.56% (SD=0.37) vs. 0.68% (SD=0.79), p-value=0.29) extracted using migratory assay and (B) 0.2% (mean: 0.28% (SD=0.18) vs. 0.45% (SD=0.21), p-value=0.08). on cells extracted using enzyme digestion after migration (C) There was a significant difference in frequency of LCs between migratory and enzyme digested after migration cells (mean: 0.57% (SD=0.62) vs. 0.36% (SD=0.21), P=0.01, data includes nine donors both inner FS and outer FS.

D. Migratory Langerhans cells express a mature and activated phenotype compared to skin resident Langerhans cells.

Changes during cell maturity and immune activation of LCs are associated with increased HIV infection (van den Berg et al., 2015). Other studies reported that these changes lead to phenotypic changes such as increased expression of HIV receptors and immune cell recruitment to the site of activation which are mostly HIV targets, reviewed by (Botting et al., 2017). This study aimed to evaluate the differences that occur in the expression of markers of LCs during maturity and activation since LCs are a subset of HIV target cells. This was achieved by evaluating proportion of maturity and activation markers between stimulated cells and unstimulated (control) cells.

To mimic the immune active environment, cells were stimulated with TNF-α, as one of the known pro-inflammatory cytokines to induce immune activation during the 48 hours of spontaneous migration. The experimental cells were subject to TNF-α (stimulated) while the control contained media alone (unstimulated). Fluorophore-conjugated antibodies against CD80/86, CD40 and HLA-DR were used to investigate maturity and immune activation. Antibodies against CD207 and CD1a (CD207+CD1a+) were used to identify LCs then the markers were assessed from this population as shown in the gating strategy in Figure 3.1.9.

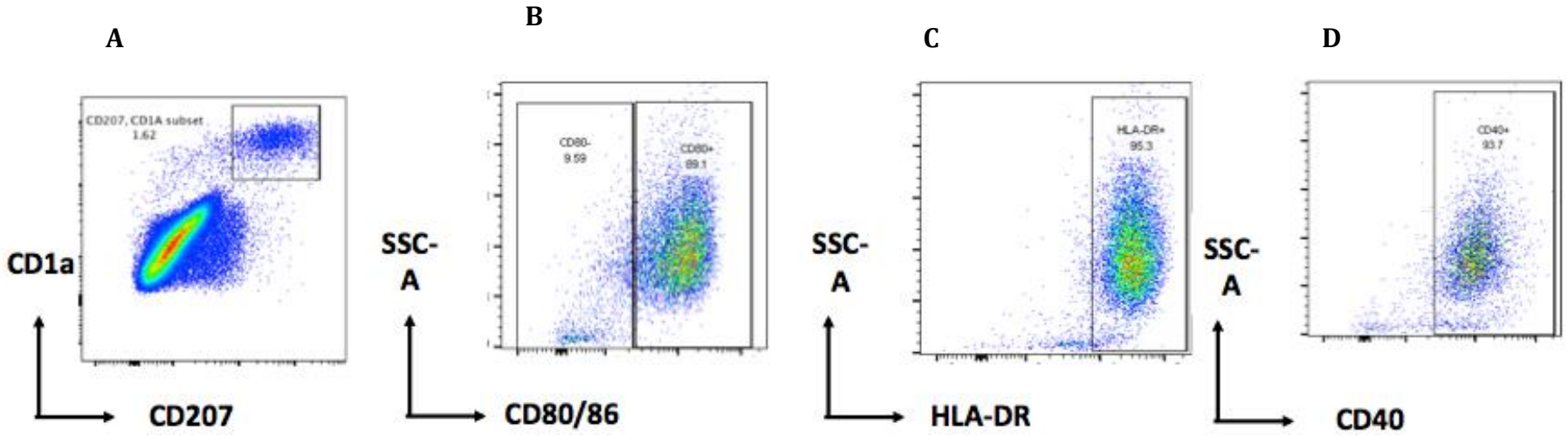


Figure 3.1.9: Gating strategy to identify maturity/activation markers from Langerhans cells (LCs). (A) CD1a, CD207 double positive LCs gated from live cells. (B) CD80/86 gated from LCs population. (C) HLA-DR+ populating gated from LCs. (D) CD40+ cells gated from LCs.

After spontaneous cell migration, there was no significant difference in maturity or activation status between the control and experimental samples. This led to the hypothesis that migratory cells might already be matured immune-active cells as there was a high frequency of migratory LCs expressing the maturity/activation markers before stimulation. The average frequency of cells expressing the maturity markers was 66,7 to 80 percent for both stimulated and unstimulated for all the markers as shown in Figure 3.1.10. It was hypothesized that when using markers that are expressed on a high proportion of a particular cell type, the impact of stimulation often does not change overall % cells expressing the markers, but the levels of expression can change. Therefore, levels of expression were assessed using Mean Fluorescence intensity (MFI) between the stimulated and unstimulated migratory cells (Figure 3.1.11). There were still no statistical significant differences between the two populations, p-value>0.05.

Subsequently, in order to evaluate whether migratory cells are already active and matured, these cells were compared to cells that remained "tissue-resident" within the FS tissue. To test this, the remaining FS tissue was enzymatically digested after migration. Interestingly, there was a significant difference between the phenotype of migratory cells and tissue-resident FS cells in frequency of the activation/maturation markers. The frequency of migratory cells expressing these markers was greater compared to cells that remained in the tissue. The frequency of migratory cells expressing CD40+ cells was 32% greater, while CD80/86+ was 56% greater and HLA-DR+ was 35% greater than cells that remained in the tissue. The overall average proportion of positive cells was 73.53% (SD=6,28) vs. 75.66% (SD=3.32) These results are shown in Figure 3.1.12 and summarized on Table 3.2. This suggests that the cells that spontaneously migrate out of the tissue are mature activated cells, whereas most of the cells that remain resident in the FS tissue are immature/inactive cells. However, the possibility that the digestive enzyme might have an impact on expression of these markers was not assessed.

The next question was whether there were differences in expression levels of these markers between the enzyme digested skin resident cells and migratory cells. The MFI was used to determine levels of expression on a per cell basis. There was no significant

difference between the migratory and skin resident cells in fluorescence intensity per cell basis Figure 3.1.13.

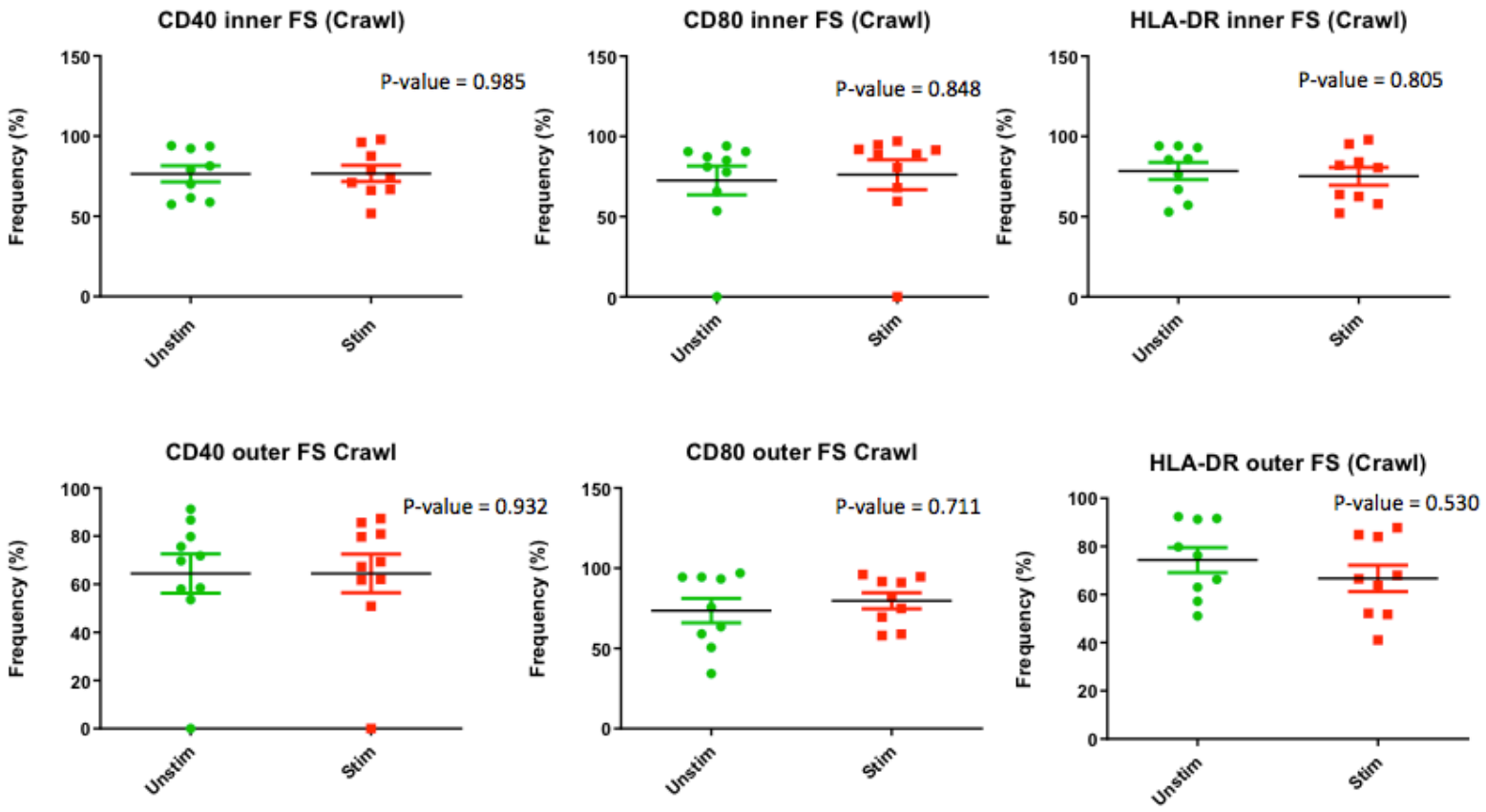


Figure 3.1.10 Frequency of maturity and activation markers on Langerhans cells (LCs) between stimulated (stim) and unstimulated (unstim) cells from migrated cells. There were no significant differences between inner FS (top) and outer FS (bottom) in expression of CD40 (A and D),CD80/86 (B and E) and HLA-DR (C and F) between unstimulated (green) and stimulated (red) cells. All the P-value> 0.05. The data presented had a sample size of n=9. The overall average proportion of positive cells was 73.53% (SD=6,28) vs. 75.66% (SD=3.32) between stimulated and unstimulated, respectively for all the markers (CD40, CD80/86 and HLA-DR).

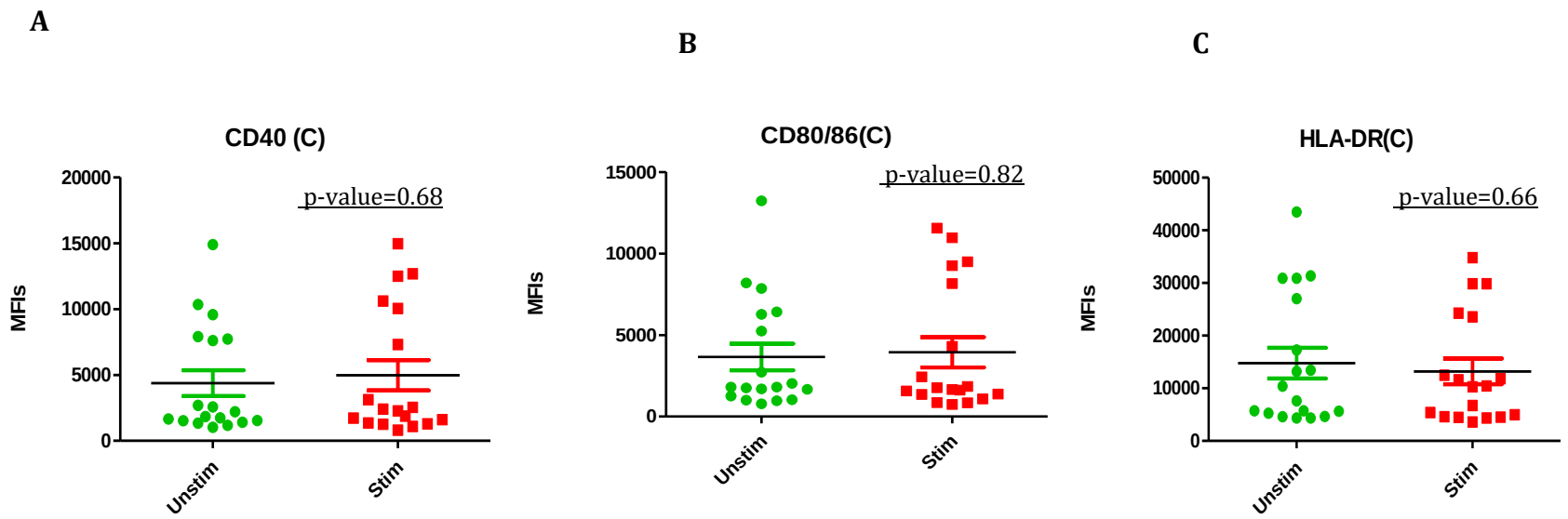


Figure 3.1.11 Level of expression of maturation and activation markers on per cell basis between stimulated and unstimulated migratory cells: Levels of expression of CD40, CD80/86 and HLA-DR on unstimulated migratory cells (green) vs. stimulated (red) measured by a flow cytometer. The data is presented as median fluorescence intensity (MFIs). The data includes 9 donors, both inner and outer foreskin. P-values >0.05 for A, B and C. There is no significant difference in expression of activation markers between stimulated and unstimulated migratory cells.

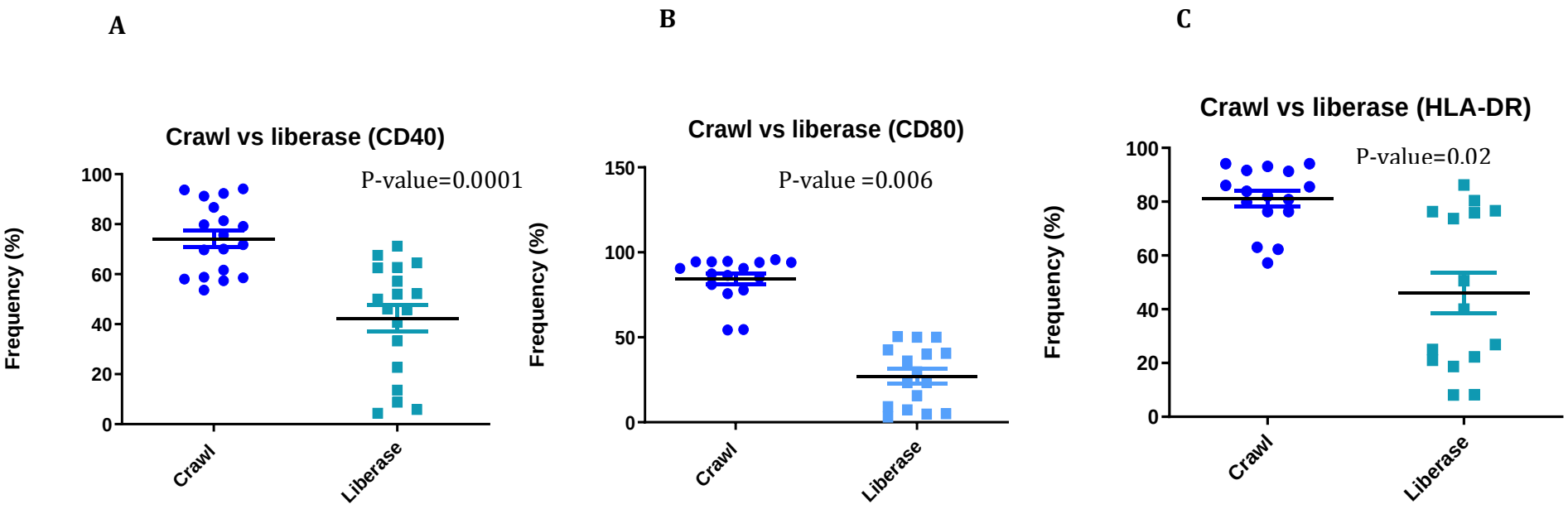


Figure 3.1.12 Frequency of migratory vs. skin resident cells expressing maturity and activation markers: Frequency of CD40 (A), CD80/86 (B) and HLA-DR (C) on Langerhans cells from migration assay (blue) and enzyme digest (turquoise) measured by a flow cytometer. The data includes 9 donors, both inner FS and outer FS. P-values were 0.0001, 0.006 and 0.02 for for A, B and C respectively. There was an average of 74% (SD=13.14) CD40+ cells from migratory cells, 81% (SD=11.68) HLA-DR+ cells and 82% (SD=12.26) cells CD80/86+ while there was an average proportion of 42% (SD=22.88) CD40+, 46% (SD=28.26) HLA-DR+ and 27% (17.68) CD80/86+ tisssue resident enzyme digested cells.

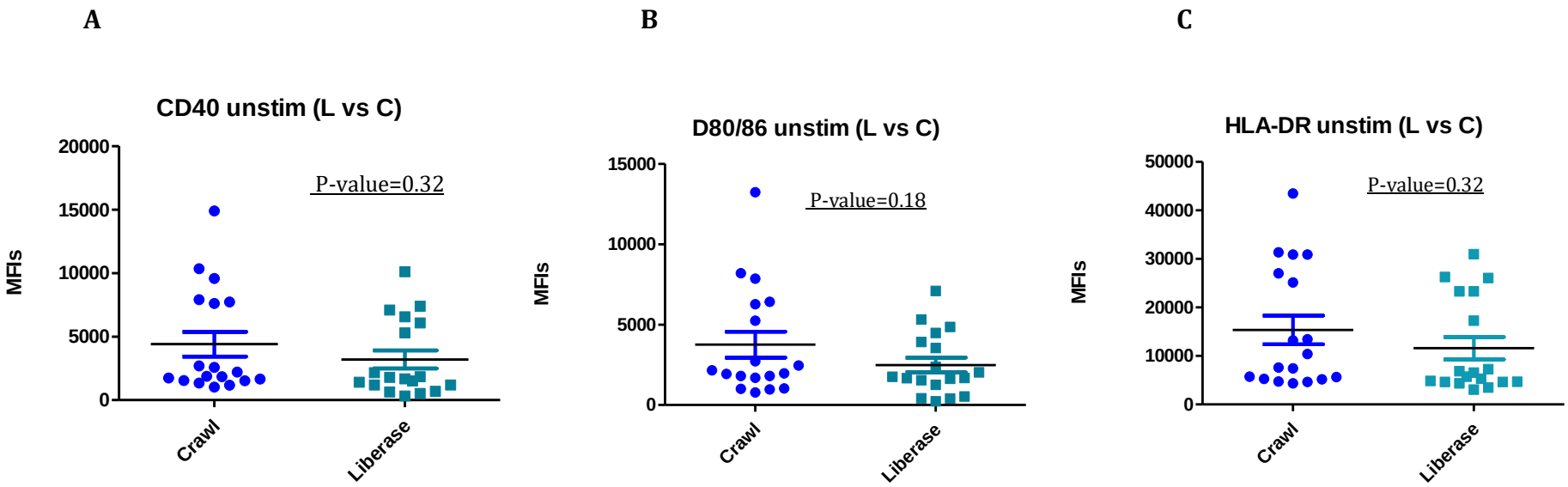


Figure 3.1.13 Level of expression of maturation and activation markers on per cell basis between migratory vs enzyme digested cells: Levels of expression of CD40, HLA-DR and CD80/86 on Langerhans cells from migration assay (blue) and enzyme digest (turquoise) measured by a flow cytometer. The data is presented as median fluorescence intensity (MFIs). The data includes 9 donors, both inner and outer foreskin. P-values >0.05 for A, B and C. There is no significant difference in expression of activation markers between enzyme digest and migratory cells.

Table 3.2: Average frequencies of migratory cells vs. skin resident cells expressing maturity and activation markers of inner FS and outer FS from 9 donors.

Marker	**Average frequency from migratory cells (%)**	**Average frequency from Tissue resident cells (%)**	**P-value**
CD40	74	42.27	0.0001
HLA-DR	81	46	0.02
CD80/86	82	26.95	0.006

E. SPICE visualisation showed that matured and activated Langerhans cells express CD80/86, CD40 and HLA-DR concurrently

To determine a sub-population of cells that co-expressed all markers, analysis of CD80/86+, HLA-DR+ and CD40+ (Figure 3.1.14) cells was performed using FlowJo boolean gating on migrated and tissue resident cells. Boolean gates produces a sub-population that expresses one, neither, some or all of the markers. The polychromatic flow cytometry immunophenotyping data set from boolean gating was subject to SPICE, a data mining software application to analyse FlowJo data sets from polychromatic flow cytometry immunophenotyping experiments for visualisation (Roederer, Nozzi, & Nason, 2011). Not surprisingly, there was significant co-expression of LC maturity and activation markers on migrated cells (CD40+CD80/86+HLA-DR+ population when compared to cells isolated by enzyme digestion following migration (this is designated by # in Figure 3.1.15). There were high proportions of enzyme-digested cells that did not express all three markers (CD40-CD80/86-HLA-DR-) population compared to cells that spontaneously migrated. There were no differences between these two groups when the expression of these markers was assessed without the other, for example; expression of CD40 and HLA-DR without CD80/86. This suggests that the cells highly express all these markers during the maturation and activation state. This suggests that there is a correlation between CD40, CD80/86 and HLA-DR as they co-express on these LCs.

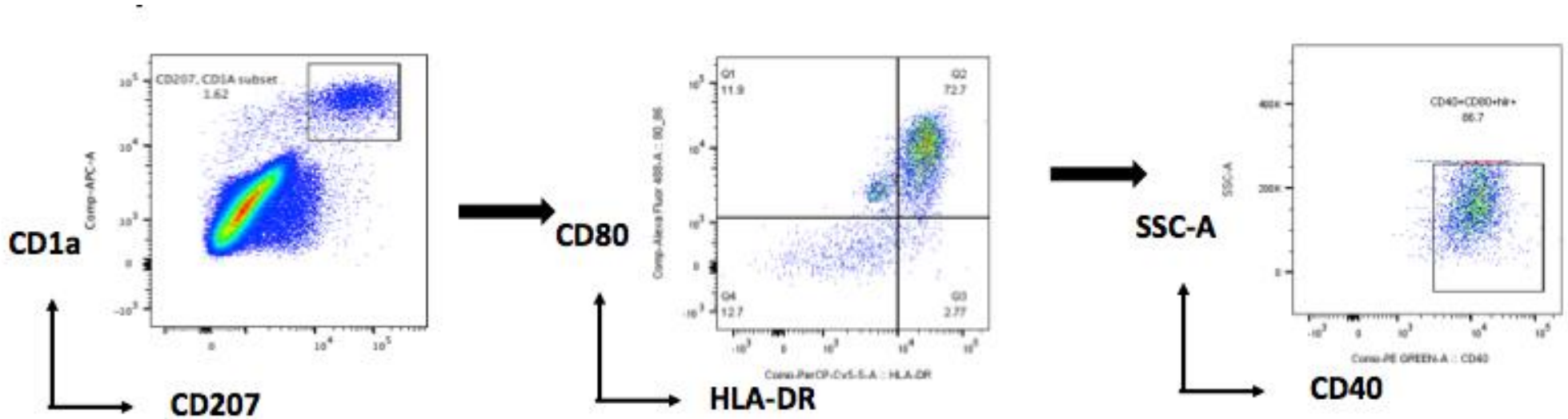
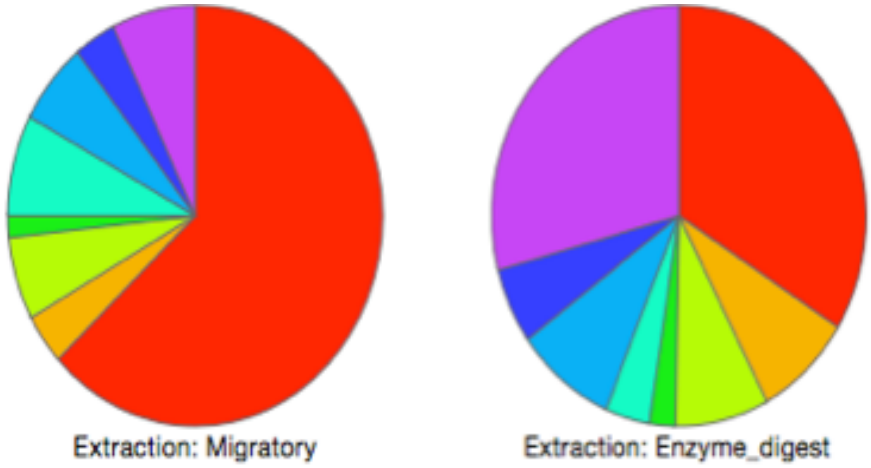


Figure 3.1.14: Langerhans cells co-express CD80/86, HLA-DR and CD40. Representation of co-expression of CD80/86+, HLA-DR+ and CD40+. CD1a+ and CD207+ cells were gated from live cells then CD0/86+ vs HLA-DR+ from that population, were gated for CD40+ cells.

A



Pie Category Test Results

Pie #	1	2
1		<0.0001
2	<0.0001	

B

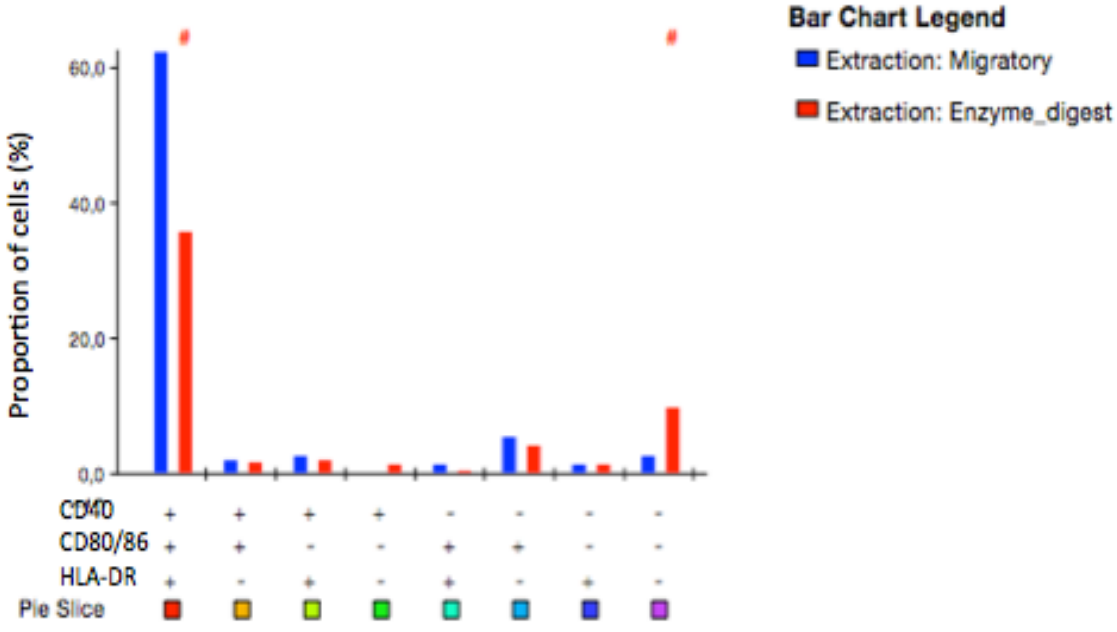


Figure 3.1.15 Co-expression of CD80/86, HLA-DR and CD40 altogether on tissue migratory cells and resident cells (enzyme digested after migration). The pie charts (A) and bar graph (B) show proportion of cells that express the different combinations of markers. The different combinations are differentiated by colors as seen from the key at the bottom of the bar graph as presented by the pie chart. Blue represents migratory cells while red represents enzyme digested. The "#" represents significant difference as shown on the bar graph.

Characterisation of Langerhans cells gene expression using single cell transcriptomic analysis

This section of work was performed in collaboration with Dr Musa Mhlanga and Dr Yutaka Negishi. Dr Negishi conducted the sequence library preparation and the bioinformatics analysis of the raw sequencing data. Yamkela Qumbelo, the applicant, isolated the cells, purified them through FACS sorting and performed the functional enrichment analysis on the expressed genes.

To characterize FS LCs further, the study investigated gene expression profile of LCs. FS epidermal sheets were digested with liberase enzyme to extract all types of LCs populations present in them. This was so that both cells capable and incapable of spontaneous migration would be isolated. Previous data shown in Figure 3.1.8 showed that tissue resident, non-migrating LC and spontaneously migrating LCs can be isolated from epidermal sheets. Therefore, the study determined the gene expression of sorted, enzyme-digested cells. LCs were purified from the pool of enzyme-digested FS cells using a FACSAria. CD1a/CD207 double positive cells were sorted to above 85% purity as described previously and were subjected to single cell RNA sequencing. Assessing gene expression of LCs provided insight in their transcriptome and their gene expression patterns.

tSNE revealed variable cell population from the sorted cells.

Data from 827 cells passed the quality control steps mentioned in the methodology and were subject to further gene expression analysis and 2,854 variably expressed genes (dispersion > 0.5) were identified and subjected to principal component analysis The study determined the cell heterogeneity of sorted, enzyme-digested cells. The scRNA-seq analysis showed that the sorted cells were unique populations, suggesting that they have different gene expression profiles represented in Figure 3.2.1 The PBMCs used in Figure 3.2.1 are from a publicly available PBMCs scRNA-seq data from a different experiment. The study used the data to compare a uniform population and the sorted cells with different populations using tSNE analysis. The results showed that there were different populations from the enzyme digested sorted FS cells. tSNE is a dimensionality reduction technique,

which is a method to graphically simplify very large datasets and display similarities and differences.

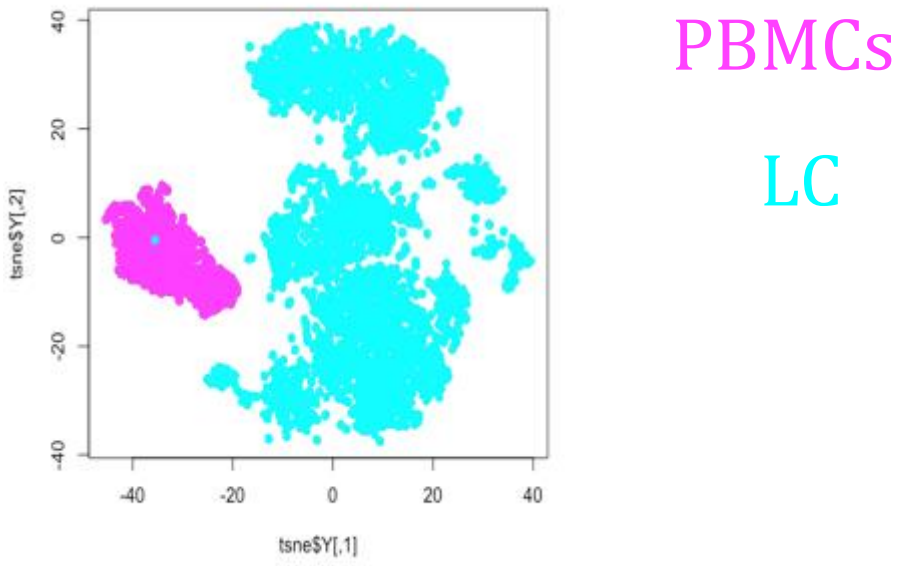


Figure 3.2.1 . A comparison between PBMCs and Langerhans cells*:* The data shows that Langerhans cells have distinct gene expression profile. Blue are the sorted Langerhans cells and pink are PBMCs. PBMCs form a tight cluster while Langerhans cells have separate groups.

The Shared Nearest Neighbor (SNN) clustering algorithm was then employed to cluster similar cells by gene expression shown in Figure 3.2.2. The different cell populations were identified by expression of known marker genes. Marker genes are the genes that were differentially expressed among the different cell populations. The results were presented by tSNE and violin plots in Figure 3.2.2 which showed that there are two clusters, Cluster 0 and cluster 1. Cluster 0 was differentiated from cluster 1 by expression of marker genes thymosine beta-4 (TMSB4X) which regulates actin polymerization and formin binding protein 1 (FNBP1) whilst cluster 1 cell population expressed 5,6-dihydroxyindole-2-carboxylic acid oxidase (TYRP1) which codes for tyrosinase enzyme responsible for melanin production and insulin-like growth factor (IGFBP7) which codes for production of insulin-like growth factor protein. The marker genes suggests that cluster 0 are monocyte-like cells which constitute 94.8% of the cells while cluster 1 are melanocyte-like cells which make up 5.2% of the total cells determined by Cap Analysis of Expressed Genes (CAGE) (Shiraki et al., 2003). LCs typically resemble monocyte cells in previous literature

(Otsuka, Watanabe, Shinya, Okura, & Saeki, 2018) and the melanocyte-like like cells were presumed to be contaminating melanocytes. Figure 3.2.3 shows distribution of marker genes on the 2 different populations.

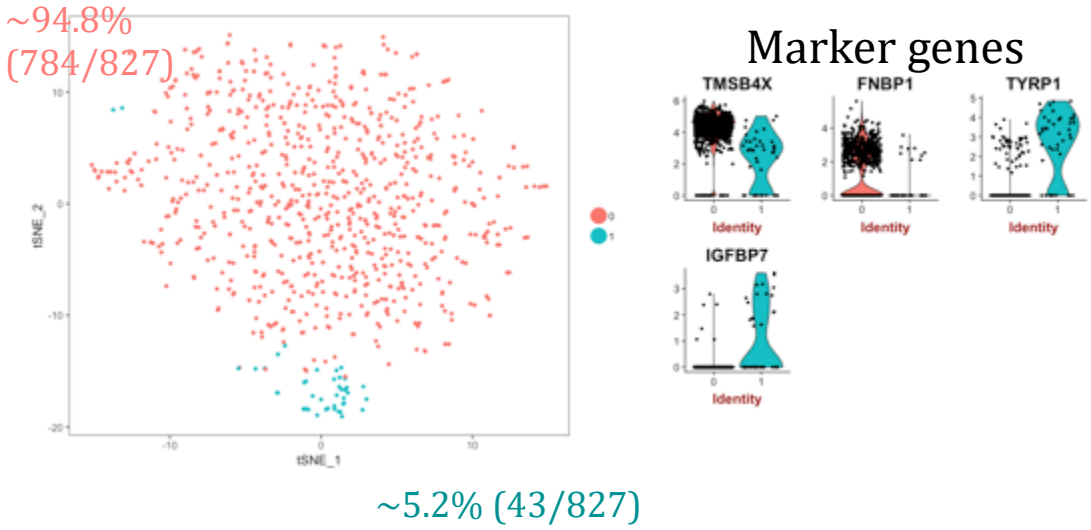


Figure 3.2.2. There are 2 cell populations within the sorted enzyme digested FS cells, melanocyte-like cells and monocyte-like cells*:* Different clustering algorithms (tSNE and SNN) show that the sorted cell population has 2 cell types, cluster 0 (pink) and cluster 1 (blue). The violin plots on the right show similarities of the cells by expression of marker genes on the two cell populations, cluster 0 and cluster 1. Cluster 0 are 94.8% of the population, the monocyte-like cells while cluster 1 are 5.2%, the melanocyte-like cells identified by the marker genes.

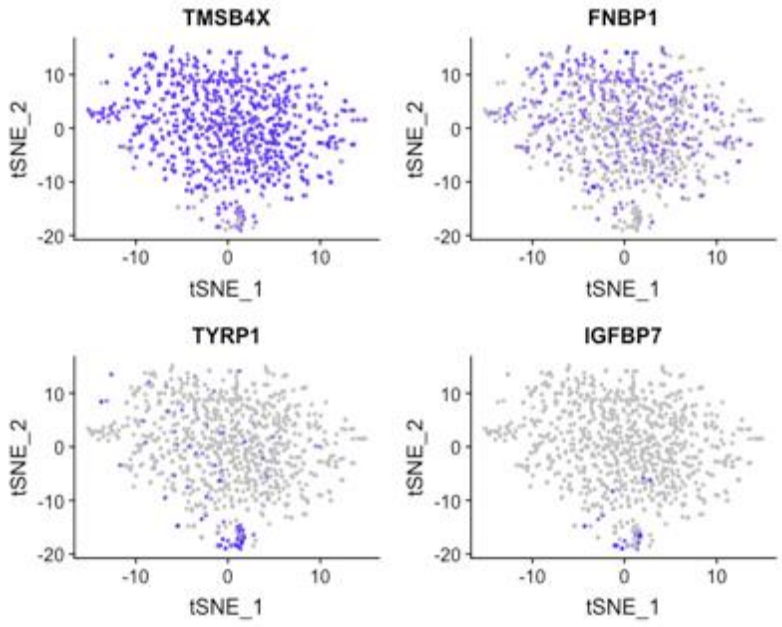


Figure 3.2.3 tSNE visualisation of marker genes from the 2 populations*:* Marker genes TMSB4X and FNBP1 were expressed on monocyte-like cells. Marker genes TYRP1 and IGFBP7 were expressed by the melanocyte-like cluster of possible contaminating melanocytes. The purple colour represents expression while grey shows no expression.

Further analysis was done on the data using different parameters to separate the different populations further. There were three groups identified. Cluster 0, cluster 1 and cluster 2 depicted by Figure 3.2.4. The violin plots and tSNE shows distribution of the marker genes among these groups. This analysis showed that the monocyte like population had 2 subtypes, cluster 0 and cluster 1.

FANTOM (Functional Annotation Of the Mammalian Genome) analysis was performed on the monocyte-like group 2 subtypes as shown (cluster 0 and cluster 1). One of the subtypes was mast cell-like (cluster 1) while the similarity of cluster 0 to other cell identities was not identified. This suggests that these cell groups are different even though they are all monocyte-like cells. Clusters 2 are melanocyte-like cells with no sub-populations based on this analysis (Kawaji et al., 2017).

To evaluate these differences functional differences between the two groups were investigated. Functional analysis using FANTOM CAGE showed that marker genes in mast cell-like clusters are highly expressed in immune cells while marker genes in the unknown

group are highly expressed in stem cells as well as immune cells, implying that the mast cell-like group has immune cell-like properties, as previously described by (Cell et al., 2016).

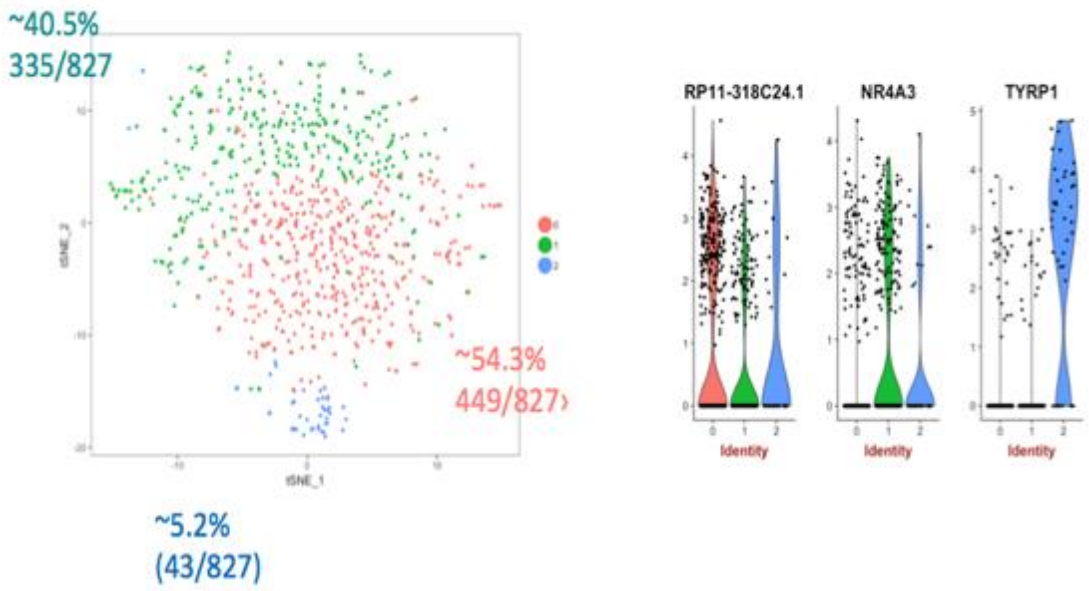


Figure 3.2.4 Different cell types and cell markers from the sorted enzyme digested cells: A tSNE on the left shows that there were three different cell groups after further analysis using different parameters. On the right are violin plots showing distribution of the marker genes across the different cell populations. The marker genes showed that cluster 0 and 1 are immune and stem cells while cluster 2 are melanocyte-like. Cluster 0 which are unidentified are 54.3% of the monocyte-like population while cluster 1 were 40.5% of the monocyte-like population.

The variable genes were subject to shared nearest neighbor clustering to identify subgroups according to co-expression (Figure 3.2.5). A functional analysis was subsequently performed with all the variable genes using ClueGo, a plug-in in cytoscape that visualizes the non-redundant biological terms for large clusters of genes in a functionally grouped network. There were varying gene ontologies, including genes involved in the metabolism and catabolism pathway, and other biological processes such as cytoskeletal reorganization. Interestingly, there were genes involved in immunological pathways such as innate immune response, immune regulation, leukocyte activation, etc. depicted by Figure 3.2.6. The figure shows different gene ontology terms in different colors

and percentages of genes involved in the designated gene ontology. Only genes involved in immunological processes are plotted on the figure. These results were expected, as Langerhans are known cells of the immune system.

An expression pattern analysis was performed to evaluate the expression of Langerhans cell markers from the sorted cells, which are CD1a, CD207 (langerin) and CDH1 (E-cadherin). A number of 11.4% of the sorted cells expressed langerin (CD207) while there was no CD1a expression and 16.35 E-cadherin (CDH1) shown in Figure A4.

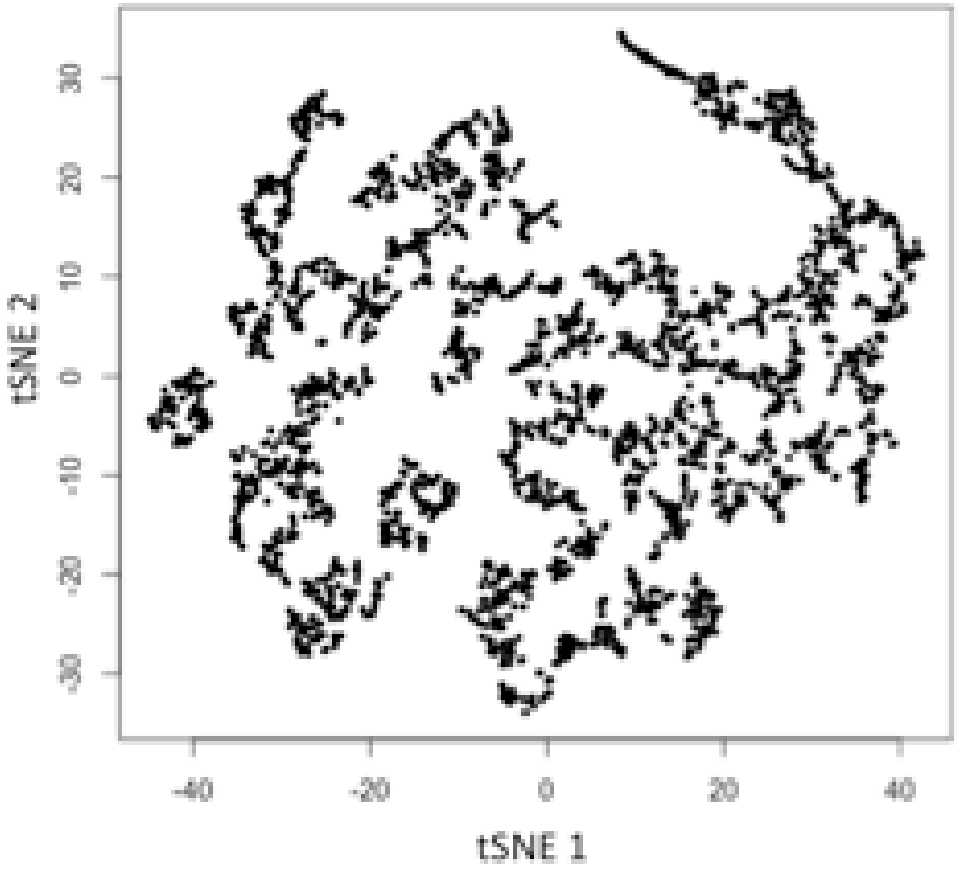


Figure 3.2.5 Network visualisation of variable genes: The 2584 variable genes were clustered according to co-expression using shared nearest neighbour clustering. The different clusters represent the different groups of genes that were co-expressed.

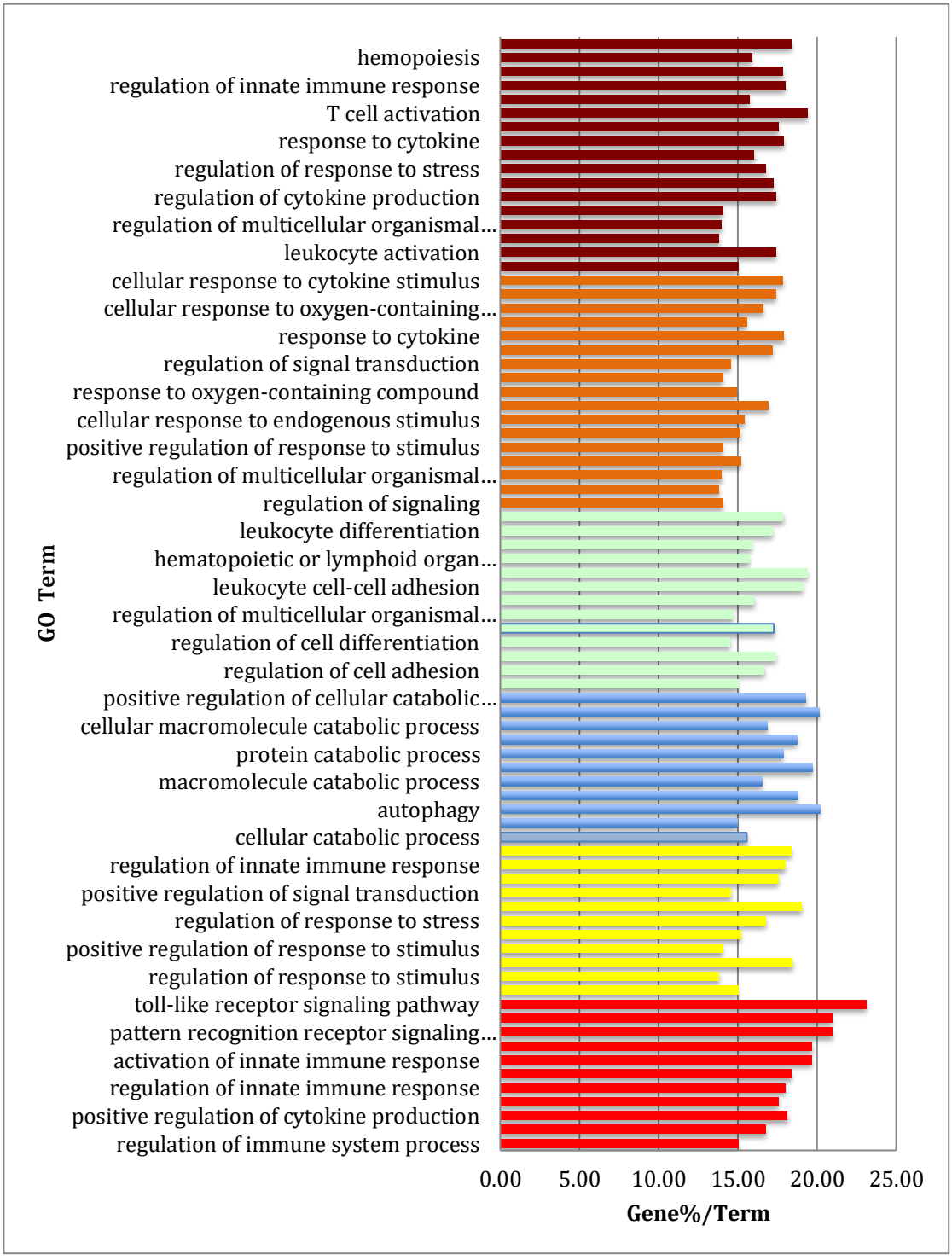


Figure 3.2.6 Enriched gene ontologies showed Langerhans cells (LCs) biological and molecular function characteristics of expressed LC genes: Functional analysis of variable genes from the sorted cells. The bars are percentages of genes found within each term. Related terms are grouped by colour. Only biological processes involved in the immune system were plotted on this graph as LCs are cells of the immune system. All ontologies had a p-value less that 0.05. Y –axis are the gene ontology terms and X-axis is % of the genes belonging to the gene ontology term.

Comparison of protein expression between the inner FS and the outer FS spontaneously migratory cells.

Optimization of protein extraction

LC-MS/MS based proteomics was used in order to comprehensively characterize the inner FS and outer FS by assessing cells that spontaneously migrate out of the FS epidermis. The proteome of these cells was investigated using two protein extraction methods (RIPA buffer with SDS and acetone precipitation, as described in the methodology). There was more protein extract obtained from the acetone precipitation method by 70% compared to the RIPA buffer method (Figure 3.3.1). RIPA has SDS and SDS has been shown to have a deleterious effect on reversed phase liquid chromatography and leads to ionization suppression during electrospray ionization (ESI). SDS also significantly affects trypsin digestion (Zhou et al., 2013).

Due to the above, complete removal of SDS is necessary for LC-MS/MS. While doing so, most of the protein is lost which results in low protein yield. In our study we used FASP, which involves the use of a filter to remove other reagents from the extraction in preparation for enzyme digestion after RIPA buffer protein extraction. In the FASP method, during digestion, some of the protein peptides are lost. Other studies have shown that most of the protein gets trapped in the filter membrane leading to few peptides detected (Zhou et al., 2013). Therefore, this method works best where there are copious amounts of starting material unlike those available with clinical human samples.

As a result of this, the acetone precipitation method was subsequently performed, coupled with in-solution digest (protein digestion is done in a solution as described in the methodology instead of a filter), which provided a high protein yield. Acetone easily evaporates and does not require removal steps, which often results in some of the protein being lost. Additionally, in-solution digestion does not involve filters or gels, which translates to less protein loss. Fewer missed cleavages (undigested proteins into peptides) were also obtained from in-solution compared to the filter, which suggests that it allows better trypsin digestion compared to the filter shown in appendices Figure A6. Due to these findings, the study used acetone precipitation to

extract protein and in-solution digestion to digest the protein to peptides using trypsin.

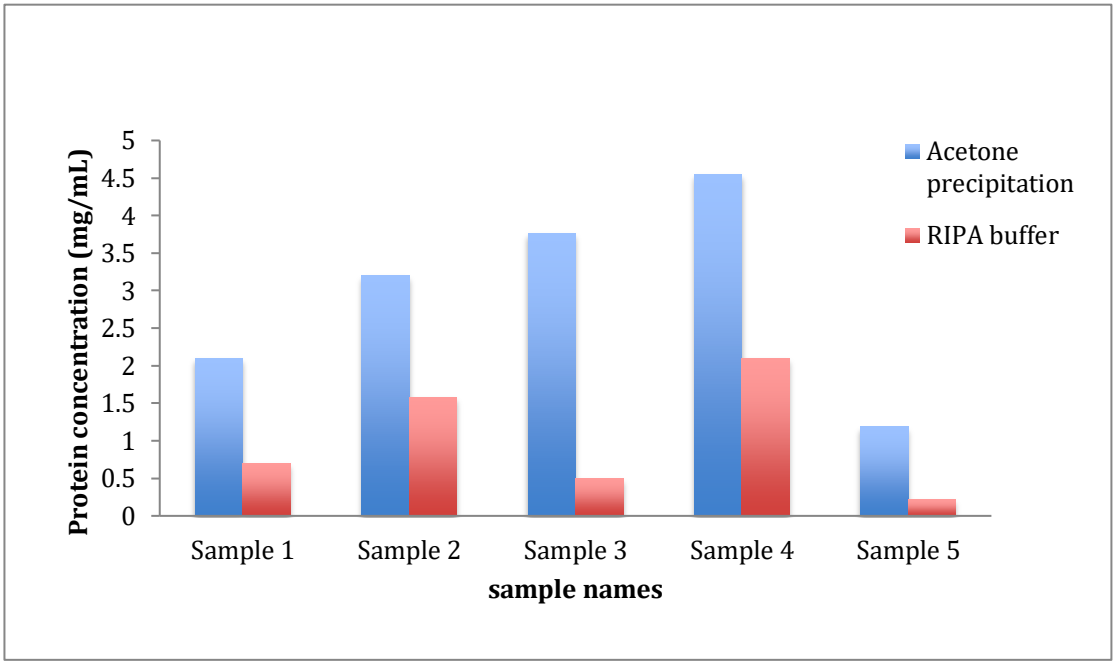


Figure 3.3.1 Protein concentration using RIPA buffer vs acetone precipitation: A comparison of the the protein concentration between the two protein extraction methods using cells extracted from FS epidermis by migration. Acetone precipitation (Blue bars) had an average of 3mg/mL from 5 samples while RIPA buffer (Red bars) had an average of 1mg/mL.

Differences in protein expression of inner FS and outer FS of migrated cells

As a part of the study aims, the inner FS and outer FS proteins from spontaneously migrated epidermal cells prepared by acetone precipitation and in-solution trypsin digestion were compared. This study investigated proteins that were over-abundant in the inner FS compared to the outer FS or exclusively expressed by inner FS or outer FS. Across 5 participants, trypsin digested peptides were subjected to a Q-Exactive mass spectrometer. An average of 5000 peptides was identified per sample with a total of 1657 proteins. An average of 10 peptides was unique to each sample and 2 peptides were used to identify proteins. About 74.5% of the peptides had zero missed cleavages, 24.1% had 1 while 4.1% had 2 (Figure A6 and A7). Reverse proteins and contaminants (except keratin) were excluded and the remaining 1516 quantified proteins were used for downstream analysis.

To look at the differences between the proteins found in the inner FS and those found in the outer FS, statistical analysis was performed using Perseus version 1.6.2.3. Proteins that were found in less than 3 out of the 5 replicates in each group were not considered as differentially expressed proteins and the missing/unidentified values were replaced by a constant of zero, only proteins expressed in 3 samples out of the 5 samples were used for the comparison for accuracy, statistical power and precision provided by biological replicates. After data filtration, 787 proteins met the criteria and were subjected to a two-tailed student's *t*-test and used for further analysis (i.e. non zero values from 3 replicates out of 5 for each protein in each group).

The proteins that met the criteria of the Student's t-test are plotted in a volcano plot, which shows proteins that were over-abundant between the inner FS and outer FS and those that were expressed uniformly between the two groups. There were 35 proteins that were over-abundant in the inner FS and 28 that were over-abundant in the outer FS (Figure 3.3.2).

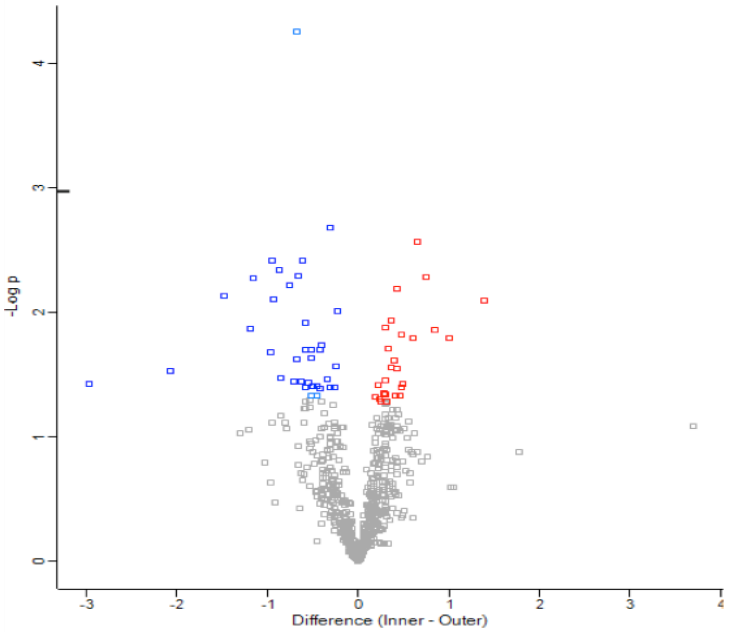


Figure 3.3.2 volcano plot showing significantly over-abundant proteins between the inner FS and outer FS. The blue squares are over-abundant proteins in the inner FS while the red circles are over-abundant proteins in the outer FS. The grey circles were not significantly different between the 2 groups ($p > 0.05$). The Y-axis is the log p-value of intensity change while the X-axis is fold change of the expressed proteins between the inner FS and the outer FS.

A principal component analysis (PCA) was performed to account for variation in the data and cluster the different samples according to sample type based on protein relative abundance from all the identified proteins. Figure 3.3.3 shows that indeed the protein relative abundance of the outer FS samples is similar as they clustered closely together and the inner FS samples had more variable protein relative abundance, as they did not cluster together.

Furthermore, an unsupervised hierarchical clustering was used to assess the relationship between the two groups according to protein relative abundance of the proteins between the inner FS and the outer FS. The two groups clustered closely together according to inner FS and outer FS. Figure 3.3.4 shows the 5 outer FS samples clustering together and 3 inner FS while 2 inner FS samples clustered together. The results are similar to the PCA as they both show that the outer FS are more related and uniform than the inner FS samples.

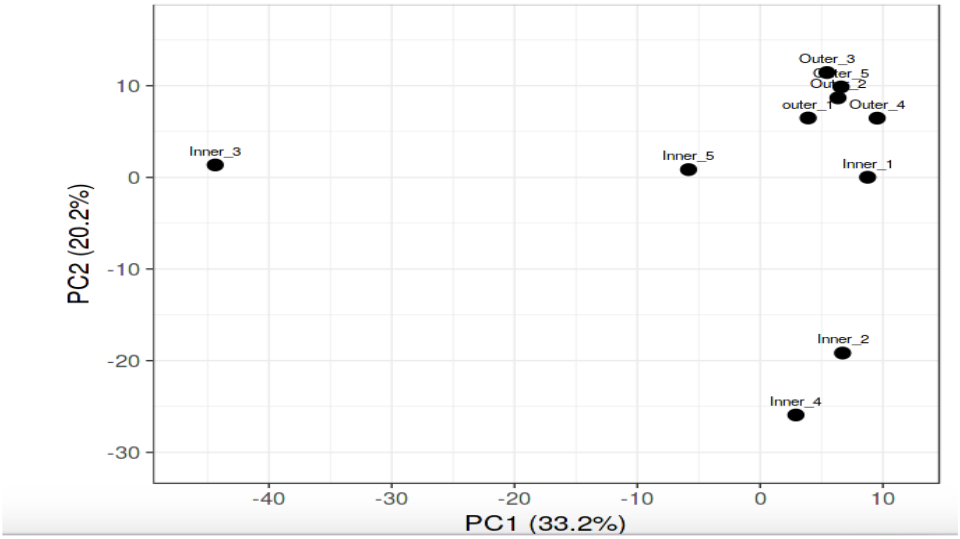


Figure 3.3.3: Principal component analysis of protein relative abundance in inner and outer migrated FS cells. The points represent the samples (inner FS and outer FS) and are labelled according to sample ID (inner FS or outer FS). There was a uniform protein relative abundance in the outer FS samples shown by the samples grouping together forming one cluster. The inner FS samples show distinct expression profiles as these samples did not group together. X and Y axis show principal component 1 and principal component 2 that explain 33.2% and 20.2% of the total variance. Number of samples is 10. The numbers next to the sample name represent sample number from 1-5 (participant 1-5). The numbers in the PID represents samples that match (inner 1 is the same participant as outer one).

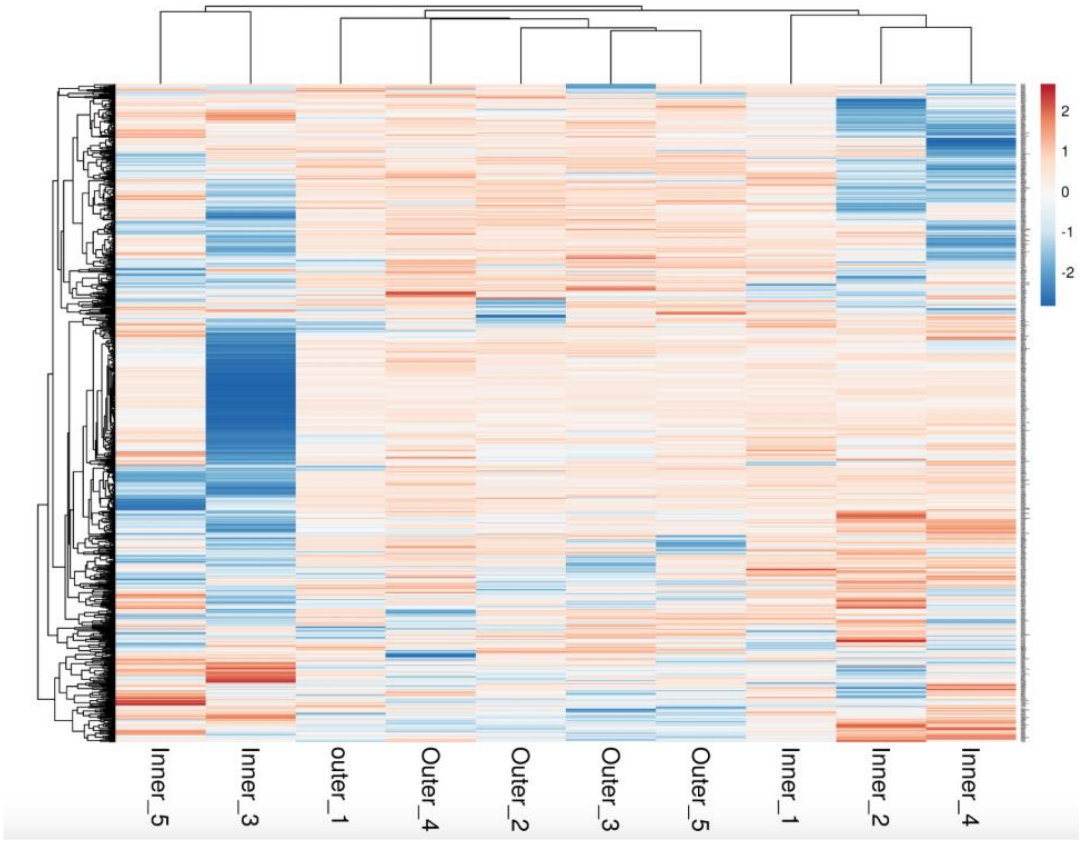


Figure 3.3.4 Unsupervised hierarchical clustering showing clustering according to protein relative abundance. The map shows protein relative abundance using LQF intensity of proteins from blue to red which is low to high respectively as shown by the scale key on the right. The 10 samples were clustered (5 inner FS and 5 outer FS) and 787 proteins using correlation distance and average linkage. All the 5 outer FS samples clustered together. The inner FS samples were more scattered as three samples clustered together and the other two together. The numbers in the PID represents samples that match (inner 1 is the same participant as outer one)

We performed a GO term analysis using cluego. It was observed that the proteins found to be over-abundant in the outer FS are mostly involved in cornified envelope, regulation of mRNA catabolic processes and the spindle midzone. Cornification is the last stage of keratinization where mammals form tough protective layers such as hair, nails and the outmost layer of the skin using squamous epithelial cells (Eckhart et al., 2013). Spindle midzones are involved in the process of cytokinesis where they assure equal chromosomal division between the two dividing cells. MRNA catabolic process is

a pathway involved in breaking down mRNA, which carries genetic codes from DNA to protein (Li et al., 2012).

The proteins over-abundant in the inner FS are mostly involved in response to interleukin-7 (IL-7) and Inosine monophosphate (IMP) biosynthesis processes. Interleukins are cytokines that mediate cell-to-cell communication. IL-7 is particularly important in regulating cell differentiation, growth and motility. They also stimulate differentiation and development of B and T cells. Other roles include stimulating cell immune responses.(Nguyen, Mendelsohn, & Larrick, 2017). IMP is formed from metabolism of adenosine and plays a major role in metabolism (Nowak & Shaw, 1997). Figures 3.3.5 and 3.3.6 show protein-protein interaction of the over-abundant outer FS proteins while 3.3.7 and 3.3.8 show those found to be over-abundant in the inner FS. The analysis of the over-abundant proteins is represented in the Tables below (Tables 3.3 and 3.4) for outer FS and inner FS, respectively.

The over-abundant proteins from both the inner FS and the outer FS were subject to STRING analysis (a biological database and web resource of known and predicted protein–protein interactions; Szklarczyk *et al.*, 2017) to evaluate protein-to-protein interaction networks according to functional analysis. Furthermore, STRING was used to identify the proteins of interest which are those involved in cornification found in the outer FS as it links to keratinization and those involved in interleukin response in the inner FS as it is linked to the immune system. These proteins include involucrin (IVL) and CDC42 small effector protein (CDC42) involved in cornification and interleukin response respectively. The proteins were associated according to co-expression, biological pathways and genetic interactions. Of the proteins over-abundant in the outer FS, 67,8% are associated by co-expression while 11% are in similar biological pathways and 0.8 had genetic interactions (Figure 3.3.4). Majority of the proteins that were over-abundant in the inner FS were associated based on biological pathways. Over 70% associates with each other according biological pathways to 13,5% according to co-expression and 5% according to physical interaction.

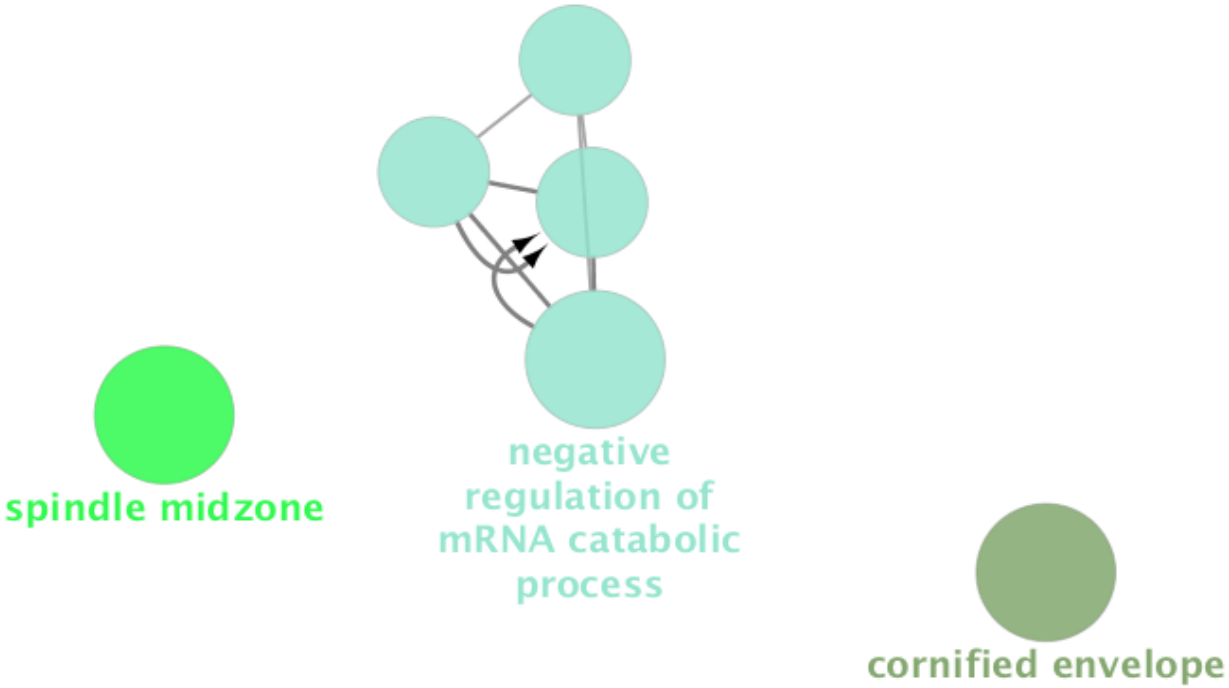


Figure 3.3.5 Gene ontology enrichment results for significantly over-abundant proteins from the outer FS ($p<0.05$). Functional analysis of the proteins over-abundant in the outer FS. Different colors present different functions/ gene ontology terms. Army-green depicts 14.29% of proteins involved in cornification, neon green are those involved in spindle midzone (14,2%) while blue are 71.14% involved in negative regulation of mRNA catabolic process. The circle size represents the number of proteins involved in the pathway, the bigger the circle, the more the proteins involved.

Table 3.3: Gene ontology analysis performed using Cluego on Cytoscape on proteins that were significantly over-abundant in cells isolated from the outer FS compared to the inner FS (number of proteins used = 28).

G0ID	GO Term	Term P-value	% Genes/Term
GO:000153	Cornified envelope	8.3E-6	4.5
GO:0051233	Spindle midzone	1.6E-4	6.5
GO:0009409	Response to cold	1.9E-4	5
GO:0043489	RNA stabilization	1.6E-4	4.7
GO: 1902373	Negative regulation of mRNA catabolic processes	9.0E-5	4.5
GO:004825	mRNA stabilization	2.2E-4	5.3

Gene ontologies were associated with cornified envelope, mRNA regulation and spindle midzone processes. %genes is the percentage of the genes in each G0 term out of the 28 proteins.

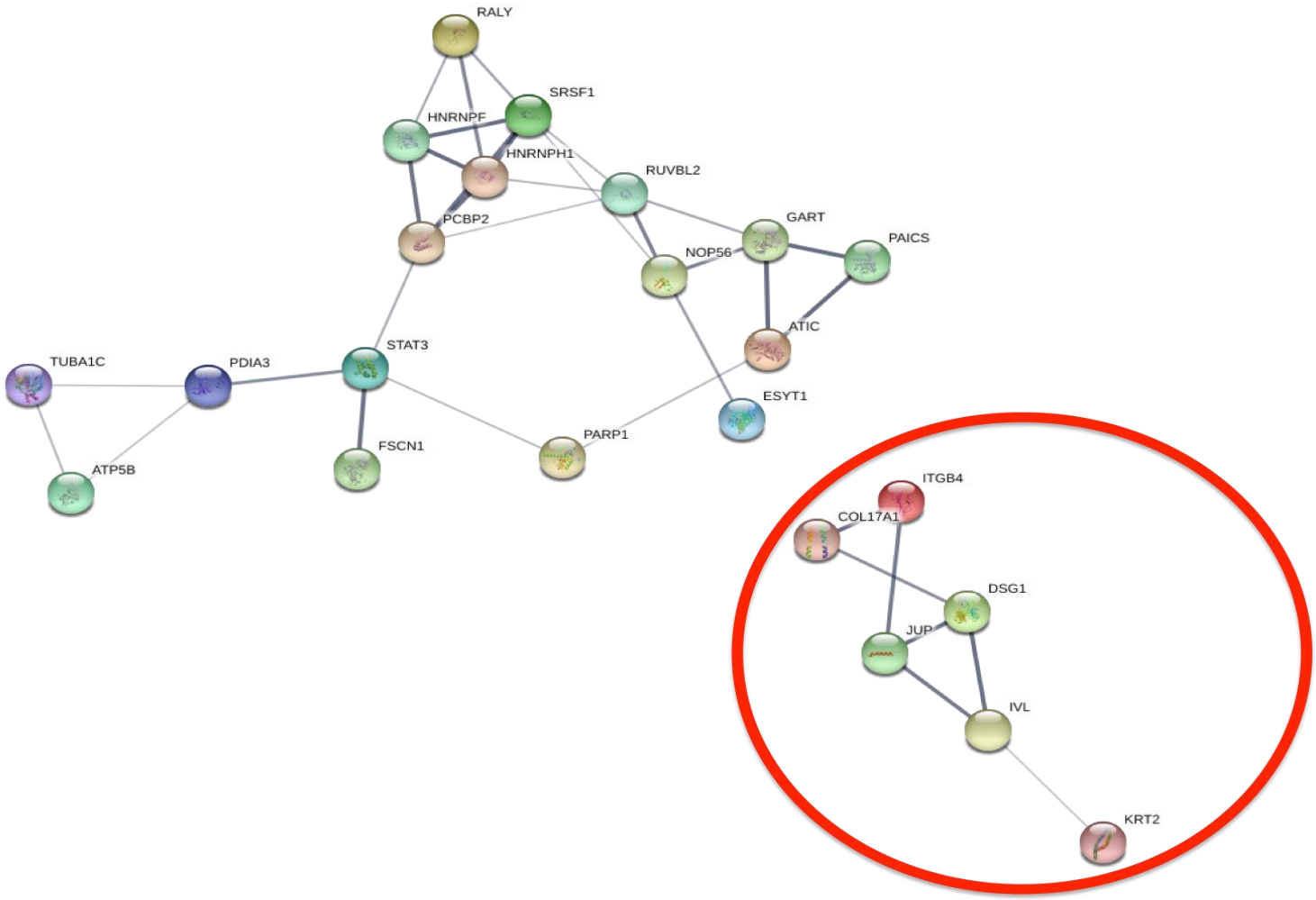


Figure 3.3.6 Network and Enrichment Analysis of the proteins over-abundant in the FS. An image from STRING website, showing results obtained upon entering a set of 28 proteins that were over-abundant in the outer FS. The image shows protein-to-protein interactions, each circle represents a protein while the thickness of the lines indicates the strength of the association/confidence in the interaction for STRING plots. The highlighted circles are proteins involved in cornification and keratinization.

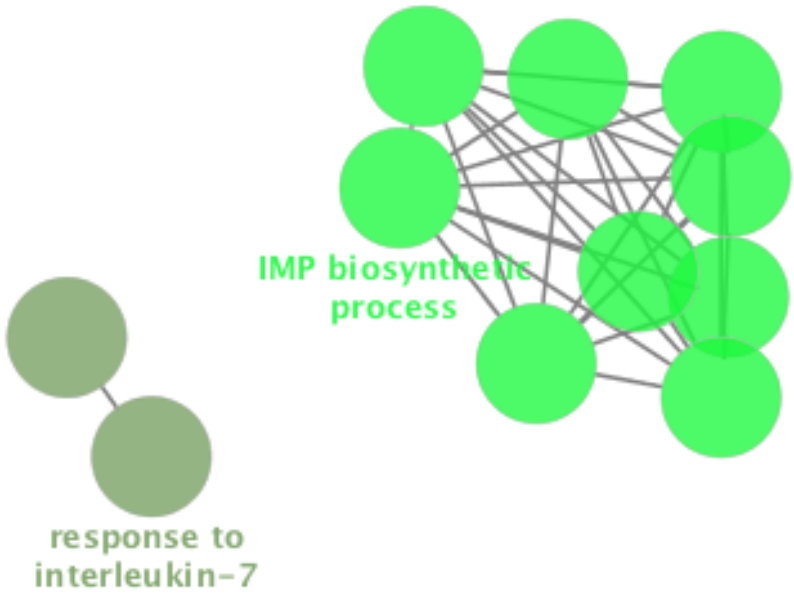


Figure 3.3.7 gene ontology enrichment results for significantly over-abundant proteins from the outer FS (p<0.05). Functional analysis of the proteins over-abundant in the inner FS. Different colors present different functions/ gene ontology terms. Green depicts proteins involved in response to interleukin-7, Lemon are those involved in IMP biosynthesis process. There are 18.18% proteins involved in response to interleukin-7 and 82% in IMP biosynthesis process.

Table 3.4: Gene ontology analysis of proteins that were over-abundant in the inner FS compared to the outer FS using Cluego on cytoscape. (number of proteins used = 35)

G0ID	GO Term	Term P-value	% Genes/Term
GO:0098760	Response to interleukin-7	2.5E-7	9
GO:0098761	Cellular response to Interleukin-7	2.56E-7	9
GO:0009124	Nucleoside monophosphate biosynthetic process	4.7E-5	5.2
GO:0009126	Purine nucleoside monophosphate metabolic process	2.9E-5	6
GO:00091	Ribonucleoside monophospate metabolic process	7.03E-5	4.9
GO:0009127	Purine nucleoside monophosphate biosynthetic process	3.1E-5	12.3
GO:0009156	Ribonucleoside monophosphate	3.5E-5	7.5
GO:0009167	Purine ribonucleoside monophosphate metabolic process	2.9E-5	6.2
GO:0009168	Purine ribonucleoside monophosphate biosynthetic process	1.7E-5	12.5
GO:0006188	IMP biosynthetic process	3.8E-7	25

Gene ontologies were linked to interleukin response and IMP biosynthetic processes. %genes is the percentage of the genes in each G0 term out of the 28 proteins.

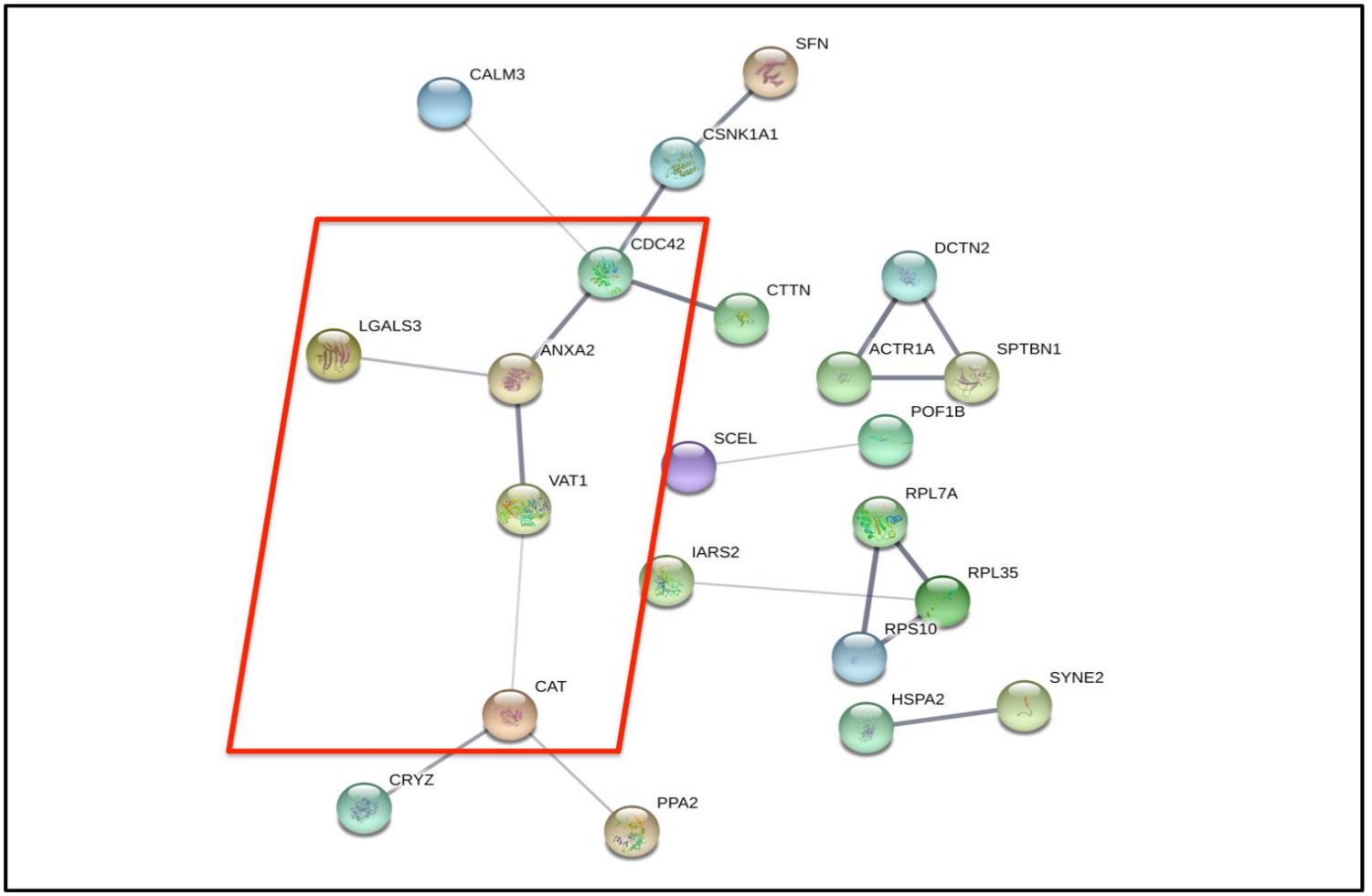


Figure 3.3.8 Network and Enrichment Analysis of the proteins over-abundant in the inner FS. An image from STRING website, showing results obtained upon entering a set of 35 proteins that were over-abundant in the inner FS. The image shows protein-to-protein interactions, each circle represents a protein while the thickness of the lines indicates the strength of the association/confidence in the interaction for STRING plots. The highlighted circles are proteins involved in response to interleukin and other immune processes.

Chapter 4

Discussion

Discussion

FS epidermal cell extraction and Langerhans cells isolation

It has been shown by different studies that LCs undergo changes upon stimulation from various sources. These changes include maturation, activation and migration, which then lead to differences in function and phenotype differentiating them from immature LCs in the epidermis. Different methods used during extraction and isolation of these cells have been shown to aid in these changes (Botting et al., 2017). Other challenges posed by different established extraction methods are contaminants by other cells and low viable cell yield. Most studies have developed useful methods for the isolation of keratinocytes, which are one of the most abundant cells in the epidermis. However, contamination by other dendritic cells has been poorly investigated. These cell types can be misleading in studies aimed at LCs as they are also cells of the immune system and closely related to LCs (Peña-Cruz et al., 2001). Therefore, a reliable method to study and analyze pure LCs in their innate status is required.

In the present study, LCs were successfully extracted and isolated from other FS cells using various methods, cell migration, enzyme digestion, cell enrichment and cell sorting. The epidermis and the dermis were separated by means of enzyme digestion of the basal membrane, which links them. Therefore, experiments were carried without dermal cell contamination such as macrophages and dermal dendritic cells. The study also compared two different cell extraction methods, migratory assay and enzyme digestion by liberase as explained in the Methodology. Both methods yielded a similar number of cells. Our data shows that the migratory extraction methods yield cells that are significantly matured and activated shown by high frequency of maturity and activation markers CD40, CD80/86 and HLA-DR. Therefore, migratory assay is not suitable for studies that investigate the role of other molecules on maturity and activation or LCs in their innate state as the cells are already matured and activated. Previous studies have shown different enzymes to cleave some cell receptors such as trypsin cleaves CD4 receptor. Even though the impact of liberase enzyme digest was not assessed directly, there were significant numbers of cells expressing the markers in question from liberase enzyme digested samples suggesting insignificant cleavage by liberase. Liberase was also chosen as it is a milder enzyme and was shown by previous studies to not have an effect on cleavage of receptors (Dolmans et al., 2006; Schmidt et al., 2018). Dolmans *et al.*, (2006)

showed that liberase retrieves significantly high viable ovarian cells without altering their morphology. Even though liberase has not been shown to affect morphology of cells, it has been shown to have low viability rate of cells compared to other enzymes such as Tumor Dissociation Enzyme (TDE) (Schmidt et al., 2018). It has also been shown to have a low number of isolated cells compared to collagenase (Dolmans et al., 2006). There was a variation on the data set as some donors had high proportion of cells expressing maturation markers as shown by Figure 3.1.12. There were low frequencies of positive cells because the cells were skin resident cells, which have been reported to be in an immature inactive state (Larsen et al., 1990).

Furthermore, LCs were separated from other epidermal cells by density gradient after cell extraction as described in the Methodology. Density gradient centrifugation enriched the LC population by 25% from 1-5%. This method does not require antibody labeling which may result in changes in the phenotype of the cell e.g. by triggering maturity as suggested by Peña-Cruz et al., (2001) and it separates live cells from dead cells. However, further optimization of the method is necessary for high LC enrichment. Another method that was used in the study to isolate LCs from other epidermal cells is cell sorting using a flow cytometer. Using cell sorting, a percentage of >85% purity was obtained post sorting and the sorted LCs were live. Therefore, these cells can be used for downstream experiments that require live cells targeting LCs only. Other ways to confirm if the isolated cells were LCs include microscopic detection of Birbeck granules, as they are only present in LCs and other surface markers such as E-cadherin (Peña-Cruz et al., 2001).

Immunoprofile of Langerhans cells maturity and activation

The immune response of LCs is dependent on their state of maturity and activation especially in HIV as reviewed in Chapter 1. According to literature, steady state LCs are tissue resident until stimulated by an antigen or other immunological stimuli such as TNF- α (Botting et al., 2017; Marsden et al., 2015; van den Berg et al., 2015).
This study aimed to evaluate the phenotypic changes that occur during LCs activation and maturation. The study observed that, upon stimulation with TNF-α, there were no significant differences between the stimulated and unstimulated migratory cells. A greater proportion of migratory cells expressed markers of maturity and activation as indicated by a significantly high frequency of cells expressing maturity and activation

markers CD40, HLA-DR and CD80/86 for both stimulated and unstimulated cells; shown in Figure 3.1.10 in Chapter 3. It is a widely accepted idea that LCs migration and activation/maturation are coupled events and that migration always occur upstream of maturation (Randolph, 2002). Studies on DCs usually define DC maturation and activation as an ability to induce immunity. It is coupled with expression of CD80, CD86 and MHC receptors, the main receptors of T cell activation (Randolph, 2002; Schubel et al., 1999). Therefore, the study speculates that the effect of stimulation did not change the overall percentage of cells expressing the markers as there were already high proportions of cells expressing the markers as they were migratory cells.

Furthermore, the study compared MFIs of the stimulated and unstimulated cells to assess the levels of expression of the activation and migration markers. Unxpectedly, there were no differences between the stimualted and unstimulated migratory cells. This suggets that TNF-α had no effect on the level of expression of the activation/maturation markers. The results were expected to show a shift in levels of expression as TNF- α is a known inducer of LCs activation (Otsuka et al., 2018). This is contrary to other studies who showed that stimulation of LCs with TNF-α increases levels of expression of activation, maturity and magration markers (Berthier-vergnes et al., 2005 (used trypsinized LCs and not migrated LCs), Polak et al., 2013).

To validate whether migratory cells are matured/activated cells and the TNF- α had no effect on the cells; the study compared unstimulated migratory LCs to skin resident cells (extracted by enzyme digestion after migration). There was an average frequency of 41% more migratory cells expressing maturity and activation markers compared to the skin resident cells, supporting the hypothesis that migratory LCs are matured activated LCs. Another study corroborating these findings is by Botting and Bertram, et al., (2017) who showed that LCs extracted through migration consistently showed an activated mature phenotype compared to the enzyme digested cells. Other authors corroborating these findings are Schubel et al., (1999) and Yanagihara et al.(2019) who showed that the maturity of LCs is associated with migration. They showed that matured LCs expressed migratory markers CCR5 and CCR7, which are markers associated with migration. The knowledge of the extent to which migration and maturation are connected is still questionable. For example, immature DCs are known to induce toleregenic response of T

cells (reviewed by Clayton et al., 2017). Geissmann et al., (2002) also showed accumulation of immature LCs in the lymph nodes from patients with dermatopathic lymphadenitis. If immature LCs are found in the lymph nodes and migration is coupled to maturation, it is still in question how the cells migrate efficiently to the lymph node, as they are immature. Which raises the question that, are these two processes really dependent; if they are, to what extent are they dependent. Studies on LCs interaction with T cells may shed light on such questions.

Even though there were differences in proportion of cells expressing activation/maturation markers between the enzyme digested cells and migratory cells, there were no differences reported when investigating the levels of expression using MFIs. Cell receptors behave differently when it comes to expression. Some receptors are expressed in low levels and up regulated upon stimulation. Therefore, they show changes in MFIs. However, some receptors are either not expressed (absent) or expressed (present). This is governed by many factors including migration and stimulation by cytokines or an antigen.(Botting and Bertram, et al., 2017) The non-significant differences in levels of expression of activation/maturity markers on the migratory vs. enzyme digested LCs suggest that they behave similarly; they are expressed during migration and not expressed when immature and inactivated. This could explain why there was an increase in proportion of positive cells but not levels of expression.

This may also be due to technicalities of the flow cytometry method. Flow cytometry data has been reported to give different results when investigating cell frequency compared to MFIs as fluorescence is sensitive to changes in experimental conditions and may be infuenced by distribution of the data especially across multiple experiments. Currently, there are no studies that invastigated levels of expression of these markers on LCs using MFIs on unstimulated cells or studies that compared results using both percentage of positive cells and levels of expression.

In relation to risk of HIV infection, matured and activated DC's have been reported to be more infectable than immature LCs (Morris et al., 2019). Upon maturity, LCs release cytokines which lead to recruitment of other immune cells that are HIV target cells leading to HIV infection (Creus et al., 2020). CCR5, which is one of the primary HIV target

receptors, has also been shown to increase during maturation, migration and during the activation of LCs (Schubel et al., 1999).

This study also showed that there are langerin positive cells that were positive for activation markers (CD80/86, CD40, HLA-DR) from the FS cells (NB: Even though the cells were langerin positive, levels of langerin expression were not quantified). This contradicts findings by (Geissmann et al., 2002; Larsen et al., 1990; Randolph, 2002) who showed that mature dendritic Langerhans cells have rare to no langerin detection in inflamed lymph nodes and those that were langerin abundant lacked CD83, a marker of LCs activation. It is noteworthy that some of these studies used antigen-infected cells or cells from unhealthy individuals, which potentially resulted in down regulation of langerin in LCs. This could be the reason that we found langerin positive cells expressing activation markers simultaneously as LCs were not infected by an antigen in our experiments. Differences in methodology may also influence the differences. For example some studies used microscopy to measure LCs markers while some used LCs extracted from mice instead of human and monocyte derived cell lines.

Characterization of Langerhans cells by RNA sequencing

Single cell RNA-seq is one of the established methods to study cell heterogeneity of cell populations that may seem identical. The study utilized scRNA-seq to characterize LCs from the enzyme digested epidermal sheets after sorting. The molecular networks of the sorted cells showed distinctiveness in cell population and the gene expression from RNA-seq data. Gene analysis of the different populations showed that the genes were associated with monocyte-like cells which were presumed to be LCs and melanocytes which were presumed to be contaminating cells. Melanocytes are one of the major cell types found in the epidermis. Therefore, it was expected that they might be one of the contaminating cells during epidermal LCs isolation. Monocytes have been shown to differentiate into LCs by different studies (Baek et al., 2012; Otsuka et al., 2018) Baek et al., (2012) showed two types of LCs derived from distinct ontogenies, short term and long term. They speculated that monocytes are precursors for short term LCs during inflammation (monocytes are recruited to the site of inflammation and develop into LCs) unlike the long term LCs that originate from the bone marrow into LCs. (Baek et al., 2012). In another study, it was shown that monocytes differentiate into DCs and

Langerhans cells when co-cultured with GM-CSF, IL-4, and TGF-b1(Otsuka et al., 2018). This established relationship between LCs and monocytes could be the reason some of the marker genes in cluster 0 (Figure 3.3.1) showed the LCs population to be monocyte-like. Therefore, in this study, it is speculated that the monocyte-like cells are LCs and may have been identified by the software analysis as monocytes due to the relationship between these cells as mentioned above. The monocyte like group had 2 subsets suggesting different LCs subsets identified by FANTOM CAGE. Figure 4.2.1 shows a summary of the different cell populations from the scRNA-seq data of sorted enzyme digested cells.

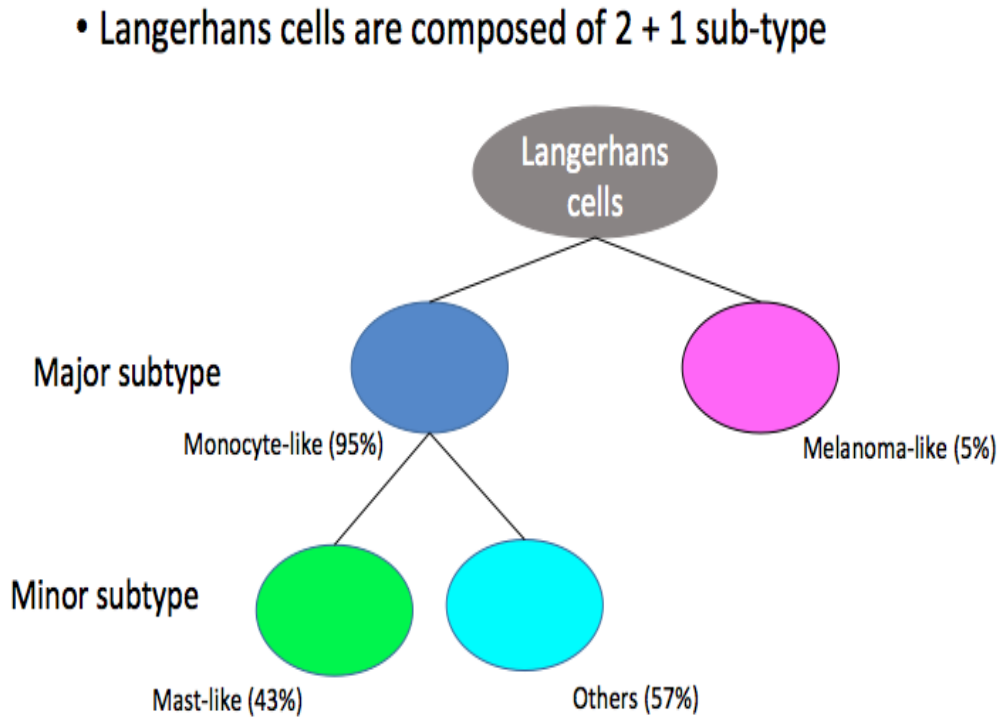


Figure 4.2.1 Representation of different cell types found on sorted enzyme digested Langerhans cells: The sorted Langerhans cell are composed of 3 sub-types, the melanocyte-like-like and monocyte-like as major types. Within the monocyte like cells, there were 2 minor populations, the mast-like and other cells. The monocyte like are presumed to be LCs while the melanocyte-like-like are the contaminating melanocytes.

In addition, using transcriptomics the study confirmed that Langerhans cells are antigen-presenting cells that induce immune response as genes involved in antigen presentation and T cell activation were found, (Figure 3.3.7.) T cell activation, activation of innate immune response and pattern recognition receptor signalling are some of the pathways

involved in antigen presentation that were identified in this study and corroborated in literature. (Artyomov et al., 2015; Segura et al., 2013; van den Berg et al., 2015).

LCs have long been recognized as sentinels of the immune system. They are responsible for cutaneous immune response, either tolerance or inflammation (Clayton et al., 2017). There have been different reports on the role of Langerhans cells but most studies and our study agree that LCs are cells of the immune system as genes involved in pathways such as innate immune activation.

Proteomic analysis of the FS

There are debates surrounding the differences between the inner FS and the outer FS. It is still unclear which part has more immune cells more specifically LCs. The differences in keratin levels between these two sites are still debatable (Ganor et al., 2010; Liu et al., 2014; Patterson et al., 2002). This has led to the question of, which part of the FS is more susceptible to HIV acquisition. Some studies suggest that it is the inner FS as it has been shown to have more HIV target cells and is less keratinized therefore its barrier is easily compromised. Others have proved otherwise as they showed no differences between the inner FS and outer FS in terms of keratinization and/or number of HIV target cells (Reviewed comprehensively by Jayathunge et al., 2014 and Morris & Wamai, 2012).

Using mass spectrometry, the study evaluated the differences between the cells of the epidermal inner FS and outer FS. This study showed that both inner FS and outer FS express the same proteins however the quantities expressed are different.

Functional enrichment analysis revealed proteins involved in interleukin-7 response and IMP biosynthesis were over-abundant in the inner FS versus the outer FS while proteins involved in cornified envelope, spindle midzone and negative regulation of mRNA catabolic processes were over-abundant in the outer FS. This suggests that the outer FS is more cornified which may relate to previous findings of the thicker keratin layer as cornification is linked with keratinization (Patterson et al., 2002b) . Il-7 is linked to immunity as it stimulates differentiation and development of B and T cells (Chazen et al., 1989). It is still unclear what leads to these differences in protein relative abundance. It is possible that they might be naturally different with different physiology or certain proteins are over-abundant in the inner FS or the outer FS due to biological and

environmental requirements. Differences in proteomics profiles may also be due to differences in cell compositions between the inner and outer FS.

The study found more proteins involved in cornification in the outer FS suggesting that the outer FS is a more of a barrier protective layer and maybe the more keratinized since cornification is the last stage of keratinization process. The inner FS had high levels of proteins involved in interlukin-7 response that is produced by most immune cells including dendritic cells and keratinocytes (Chazen, et al 1989). The cytokine is responsible for cell-to-cell signaling therefore is important for initiating an immune response (Zaitseva et al., 1997). This suggests that the inner FS potentially plays a differential role in immune responses against pathogens once it has crossed the outer FS barrier. The outer FS forms a tough protective layer of keratin than the inner FS and the inner FS plays a role in the immune response. However, there were no proteins that were found to be exclusive to either the inner FS or the outer FS in our study. The limitation to the present study was that the protein composition in the tissue samples may not be representative of the total protein content found in either inner FS or outer FS, because the LC-MS/MS utilized could only detect proteins expressed in high concentrations thus may have excluded proteins that are found in low concentrations. Non-detection of expression does not equate to the protein absence.

Conclusion and future work

A limitation is that, while a large variety of different experiments were conducted, the sample sizes were at times very small, limiting statistical power, especially for omics-type analysis

The present study has shown that FS tissue harbors distinct Langerhans cell subsets. There are no differences in frequency of LCs between the inner FS and the outer FS after flow cytometry acquisition from both migratory and enzyme digested cells. Migratory cells are matured activated cells and tissue resident cells are immature/inactive cells as frequency of cells expressing maturity/activation markers (CD40, CD80/86, HLA-DR) was >30% for each marker on migrating cells compared to the skin resident cells. These findings suggest that migration; maturity and activation are coupled events. The differences between these cell types may also be due to the different extraction methods. For future studies, migratory markers such as CCR5 and CCR7 to validate this relationship

can extend the panel. Other markers such as CD4 can be added to assess the effect of activation on these receptors. This would shed light on the relationship between HIV and cell activation, as CD4 is an HIV target receptor.

The study also showed that scRNA-seq is a good method to distinguish cell populations that seem to be similar. There were three different populations from a sorted population with 2 subsets of LCs.

Through proteomics studies, we showed that the inner FS potentially plays an immunological role while the outer FS is more of a physical barrier to pathogen. This is due to high levels of proteins involved in immunological processes in the inner FS and those involved in cornification were high in the outer FS. LC-MS/MS used in the study is quantity dependent meaning that it detects proteins that are at high quantities; as a result, some of the proteins of interest expressed in low concentrations may have been missed. Therefore, using more sensitive assays such as targeted proteomics or western blot will be more accurate in determining the protein composition in the FS tissue.

LCs isolation and quantification using flow cytometry suggests that the inner FS and outer FS play similar roles in HIV in the context of LCs as similar quantities of LCs were found and there were no differences in expression of markers associated with activation (CD80) or maturation (CD40). As reviewed in Chapter 1, HIV infection is dependent but not limited to the density of target cells such as LCs. Our study contradicts findings by (Ganor et al., 2010) who found more LCs in the inner FS than the outer FS and (McCoombe & Short, 2006) who found more LCs in the outer FS compared to the inner FS. Differences in quantification methods may account to these differences and health status of the participants.

Protein studies suggest that the outer FS may be less susceptible to the virus due to high keratin levels while the inner FS is potentially more susceptible as it may constitute more immune cells that are HIV target cells since IL-7 is produced by most immune cells. Differences between the inner FS and outer FS were not seen in LCs density using quantification by flow cytometry but were observed when using LC/MS-MS. This indicates the benefit of multiple ways of analysis to get whole biological information.

Bibliography

Abbas, W., & Herbein, G. (2013). T-Cell Signaling in HIV-1 Infection. *The Open Virology Journal*, *7*, 57–71. https://doi.org/10.2174/1874357920130621001

Ahmed, Z., Kawamura, T., Shimada, S., & Piguet, V. (2015). The Role of Human Dendritic Cells in HIV-1 Infection. *J Invest Dermatol*, *135*(5), 1225–1233. https://doi.org/10.1038/jid.2014.490

Anderson, D., Politch, J. A., & Pudney, J. (2012). NIH Public Access, *65*(3), 220–229. https://doi.org/10.1111/j.1600-0897.2010.00941.x.HIV

Artyomov, M. N., Munk, A., Gorvel, L., Korenfeld, D., Cella, M., Tung, T., & Klechevsky, E. (2015). Modular expression analysis reveals functional conservation between human Langerhans cells and mouse cross-priming dendritic cells. *The Journal of Experimental Medicine*, *212*(5), 743–757. https://doi.org/10.1084/jem.20131675

Auvert, B., Taljaard, D., Lagarde, E., Sobngwi-Tambekou, J., Sitta, R., & Puren, A. (2005). Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 Trial. *PLoS Medicine*, *2*(11), e298. https://doi.org/10.1371/journal.pmed.0020298

Baek, J., Ober-blo, J., Tacke, F., Yokota, Y., Zenke, M., & Hieronymus, T. (2012). Article Two Distinct Types of Langerhans Cells Populate the Skin during Steady State and Inflammation, *2*, 905–916. https://doi.org/10.1016/j.immuni.2012.07.019

Bailey, R. C., Moses, S., Parker, C. B., Agot, K., Maclean, I., Krieger, J. N., … Ndinya-Achola, J. O. (2007). Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. *Lancet*, *369*(9562), 643–656. https://doi.org/10.1016/S0140-6736(07)60312-2

Ballweber, L., Robinson, B., Kreger, A., Fialkow, M., Lentz, G., Mcelrath, M. J., & Hladik, F. (2011). Vaginal Langerhans Cells Nonproductively Transporting HIV-1 Mediate Infection of T Cells □, *85*(24), 13443–13447. https://doi.org/10.1128/JVI.05615-11

Berg, L. M. Van Den, Ribeiro, C. M. S., Zijlstra-willems, E. M., Witte, L. De, Fluitsma, D., Tigchelaar, W., … Geijtenbeek, T. B. H. (2014). Caveolin-1 mediated uptake via langerin restricts HIV-1 infection in human Langerhans cells, 1–9. https://doi.org/10.1186/s12977-014-0123-7

Berthier-vergnes, O., Bermond, F., Flacher, V., Massacrier, C., Schmitt, D., Pe, J., … Lyon,

I. (2005). TNF- a enhances phenotypic and functional maturation of human epidermal Langerhans cells and induces IL-12 p40 and IP-10 / CXCL-10 production, *579*, 3660–3668. https://doi.org/10.1016/j.febslet.2005.04.087

Bongaarts, J., Reining, P., Way, P., & Conant, F. (1989). The relationship between male circumcision and HIV infection in African populations. *AIDS*, *3*(6). Retrieved from https://journals.lww.com/aidsonline/Fulltext/1989/06000/The_relationship_between_male_circumcision_and_HIV.6.aspx

Botting, R. A., Bertram, K. M., Baharlou, H., Sandgren, K. J., Fletcher, J., Rhodes, J. W., … Harman, A. N. (2017). Phenotypic and functional consequences of different isolation protocols on skin mononuclear phagocytes, *101*(June), 1393–1403. https://doi.org/10.1189/jlb.4A1116-496R

Botting, R. A., Rana, H., Bertram, K. M., Rhodes, J. W., Baharlou, H., Nasr, N., … Harman, A. N. (2017). Langerhans cells and sexual transmission of HIV and HSV. *Reviews in Medical Virology*, (September 2016), e1923. https://doi.org/10.1002/rmv.1923

Cardinaud, S., Aar, V. Der, Sprokholt, J. K., Jong, M. A. W. P. De, Zijlstra-willems, E. M., Moris, A., & Alerts, E. (2018). Langerhans Cell − Dendritic Cell Cross-Talk via Langerin and Hyaluronic Acid Mediates Antigen Transfer and Cross-Presentation of HIV-1. https://doi.org/10.4049/jimmunol.1402356

Cell, M., Wood, J. G., & Cell, M. M. (2016). Mast Cell : A Multi-Functional, *6*(January), 1–12. https://doi.org/10.3389/fimmu.2015.00620

Chazen, G. D., Pereira, G. M. B., Legros, G., Gillist, S., & Shevach, E. M. (1989). Interleukin 7 is a T-cell growth factor, *86*(August), 5923–5927.

Clayton, K., Vallejo, A. F., Davies, J., Sirvent, S., Polak, M. E., & Bennett, C. L. (2017). Langerhans Cells — Programmed by the epidermis, *8*(November), 1–14. https://doi.org/10.3389/fimmu.2017.01676

Creus, A. De, Beneden, K. Van, Taghon, T., Debacker, V., Plum, J., & Leclercq, G. (2020). Langerhans Cells That Have Matured In Vivo in the Absence of T Cells Are Fully Capable of Inducing a Helper CD4 as Well as a Cytotoxic CD8 Response. https://doi.org/10.4049/jimmunol.165.2.645

De Jong, M. A. W. P., & Geijtenbeek, T. B. H. (2009). Human immunodeficiency virus-1 acquisition in genital mucosa: Langerhans cells as key-players. *Journal of Internal Medicine*, *265*(1), 18–28. https://doi.org/10.1111/j.1365-2796.2008.02046.x

de Witte, L., Nabatov, A., Pion, M., Fluitsma, D., de Jong, M. a W. P., de Gruijl, T., … Geijtenbeek, T. B. H. (2007). Langerin is a natural barrier to HIV-1 transmission by Langerhans cells. *Nature Medicine*, *13*(3), 367–371. https://doi.org/10.1038/nm1541

Dolmans, M., Michaux, N., Camboni, A., Martinez-madrid, B., Langendonckt, A. Van, Nottola, S. A., & Donnez, J. (2006). Evaluation of Liberase , a purified enzyme blend , for the isolation of human primordial and primary ovarian follicles, *21*(2), 413–420. https://doi.org/10.1093/humrep/dei320

Eckhart, L., Lippens, S., Tschachler, E., & Declercq, W. (2013). Biochimica et Biophysica Acta Cell death by corni fi cation ☆, *1833*, 3471–3480. https://doi.org/10.1016/j.bbamcr.2013.06.010

Fahrbach, K. M., Barry, S. M., Ayehunie, S., Lamore, S., Klausner, M., & Hope, T. J. (2007). Activated CD34-Derived Langerhans Cells Mediate Transinfection with Human Immunodeficiency Virus □, *81*(13), 6858–6868. https://doi.org/10.1128/JVI.02472-06

Février, M., Dorgham, K., & Rebollo, A. (2011). CD4+ T Cell Depletion in Human Immunodeficiency Virus (HIV) Infection: Role of Apoptosis. *Viruses*, *3*(12), 586–612. https://doi.org/10.3390/v3050586

Fiebig, E. W., Wright, D. J., Rawal, B. D., Garrett, P. E., Schumacher, R. T., Peddada, L., … Busch, M. P. (2003). Dynamics of HIV viremia and antibody seroconversion in plasma donors : implications for diagnosis and staging of primary HIV infection, (August 2002). https://doi.org/10.1097/01.aids.0000076308.76477.b8

Fischetti, L., Barry, S. M., Hope, T. J., Shattock, R. J., & Terrace, C. (2015). HHS Public Access, *23*(3), 319–328. https://doi.org/10.1097/QAD.0b013e328321b778.HIV-1

Freeman EE, Weiss HA, Glynn JR, Cross PL, Whitworth JA, H. R. (2006). Herpes simplex virus 2 infection increases HIV acquisition in men and women: systematic review and meta-analysis of longitudinal studies. *AIDS (London, England)*, *20*(1), 73–83. https://doi.org/00002030-200601020-00011 [pii]

Ganor, Y, Zhou, Z., Bodo, J., Tudor, D., Leibowitch, J., Mathez, D., … Bomsel, M. (2013). The adult penile urethra is a novel entry site for HIV-1 that preferentially targets resident urethral macrophages. *Mucosal Immunology*, *6*(4). https://doi.org/10.1038/mi.2012.116

Ganor, Y, Zhou, Z., Tudor, D., Schmitt, a, Vacher-Lavenu, M.-C., Gibault, L., … Bomsel, M.

(2010). Within 1 h, HIV-1 uses viral synapses to enter efficiently the inner, but not outer, foreskin mucosa and engages Langerhans-T cell conjugates. *Mucosal Immunology*, *3*(5), 506–522. https://doi.org/10.1038/mi.2010.32

Ganor, Yonatan, & Bomsel, M. (2011). HIV-1 transmission in the male genital tract. *American Journal of Reproductive Immunology (New York, N.Y. : 1989)*, *65*(3), 284–291. https://doi.org/10.1111/j.1600-0897.2010.00933.x

Geissmann, F., Dezutter, C., Valladeau, J., Kayal, S., Leborgne, M., Brousse, N., … Davoust, J. (2002). Accumulation of Immature Langerhans Cells in Human Lymph Nodes Draining Chronically Inflamed Skin, *196*(4). https://doi.org/10.1084/jem.20020018

Giddey, A. D., Kock, E. De, Nakedi, K. C., Garnett, S., Nel, A. J. M., Soares, N. C., & Blackburn, J. M. (2017). A temporal proteome dynamics study reveals the molecular basis of induced phenotypic resistance in Mycobacterium smegmatis at sub-lethal rifampicin concentrations. *Nature Publishing Group*, (October 2016), 1–13. https://doi.org/10.1038/srep43858

Gray, R. H., Kigozi, G., Serwadda, D., Makumbi, F., Watya, S., Nalugoda, F., … Wawer, M. J. (2007). Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. *Lancet*, *369*(9562), 657–666. https://doi.org/10.1016/S0140-6736(07)60313-4

Haberland, N. A., Kelly, C. A., Mulenga, D. M., Mensch, B. S., & Hewett, C. (2016). Women ' s Perceptions and Misperceptions of Male Circumcision : A Mixed Methods Study in Zambia, *2014*(Who 2015), 1–15. https://doi.org/10.1371/journal.pone.0149517

Harman, A. N., Bye, C. R., Nasr, N., Kerrie, J., Kim, M., Mercier, S. K., … Paul, U. (2019). Identification of Lineage Relationships and Novel Markers of Blood and Skin Human Dendritic Cells. https://doi.org/10.4049/jimmunol.1200779

Hindmarsh, P. (1999). Retroviral DNA Integration, *63*(4), 836–843.

Hladik, F., Sakchalathorn, P., Ballweber, L., Lentz, G., Fialkow, M., Eschenbach, D., & Mcelrath, M. J. (2007). Article Initial Events in Establishing Vaginal Entry and Infection by Human Immunodeficiency Virus Type-1, *26*(2), 257–270. https://doi.org/10.1016/j.immuni.2007.01.007

Hussain, L. A., & Lehner, T. (1995). Comparative investigation of Langerhans' cells and potential receptors for HIV in oral, genitourinary and rectal epithelia. *Immunology*, *85*(3), 475–484.

Jayathunge, P. H. M., Mcbride, W. J. H., Maclaren, D., Kaldor, J., Vallely, A., & Turville, S. (2014). Male Circumcision and HIV Transmission ; What Do We Know ?, 31–44.

Kawaji, H., Kasukawa, T., Forrest, A., & Carninci, P. (2017). Comment : The FANTOM 5 collection , a data series underpinning mammalian transcriptome atlases in diverse cell types, (August), 2016–2018.

Kawamura, T., Kurtz, S. E., Blauvelt, A., & Shimada, S. (2005). The role of Langerhans cells in the sexual transmission of HIV, 147–155. https://doi.org/10.1016/j.jdermsci.2005.08.009

Kigozi, G., Watya, S., Polis, C. B., Buwembo, D., Kiggundu, V., Wawer, M. J., … Gray, R. H. (2008). The effect of male circumcision on sexual satisfaction and function, results from a randomized trial of male circumcision for human immunodeficiency virus prevention, Rakai, Uganda. *BJU International*, *101*(1), 65–70. https://doi.org/10.1111/j.1464-410X.2007.07369.x

Klatzmann, D., Barre-Sinoussi, F., Nugeyre, M. T., Danquet, C., Vilmer, E., Griscelli, C., … et, al. (1984). Selective tropism of lymphadenopathy associated virus (LAV) for helper-inducer T lymphocytes. *Science*, *225*(4657), 59 LP – 63. https://doi.org/10.1126/science.6328660

Koppensteiner, H., & Wu, L. (2012). Macrophages and their relevance in Human Immunodeficiency Virus Type I infection. *Retrovirology*, *9*(1), 82. https://doi.org/10.1186/1742-4690-9-82

Larsen, B. C. P., Steinman, R. M., Witmer-pack, M., Hankins, D. F., Morris, P. J., & Austyn, J. M. (1990). Migration and Maturation of Langerhans Cells in Skin Transplants and Explants, *172*(November).

Li, F., Zheng, Q., Ryvkin, P., Dragomir, I., Desai, Y., Aiyer, S., … Cherry, S. (2012). Resource Global Analysis of RNA Secondary Structure in Two Metazoans. *CellReports*, *1*(1), 69–82. https://doi.org/10.1016/j.celrep.2011.10.002

Li, G., Wu, W., Zhang, X., Huang, Y., Wen, Y., Li, X., & Gao, R. (2018). Serum levels of tumor necrosis factor alpha in patients with IgA nephropathy are closely associated with disease severity, 1–9.

Liu, A., Yang, Y., Liu, L., Meng, Z., Li, L., Qiu, C., … Zhang, X. (2014). Differential compartmentalization of HIV-targeting immune cells in inner and outer foreskin tissue. *PLoS ONE*, *9*(1), 1–10. https://doi.org/10.1371/journal.pone.0085176

Marsden, V., Donaghy, H., Bertram, K. M., Harman, A. N., Nasr, N., Keoshkerian, E., … Cunningham, A. L. (2015). Herpes Simplex Virus Type 2–Infected Dendritic Cells Produce TNF-α, Which Enhances CCR5 Expression and Stimulates HIV Production from Adjacent Infected Cells. *The Journal of Immunology*, *194*(9), 4438–4445. https://doi.org/10.4049/jimmunol.1401706

McCoombe, S. G., & Short, R. V. (2006). Potential HIV-1 target cells in the human penis. *AIDS (London, England)*, *20*(11), 1491–1495. https://doi.org/10.1097/01.aids.0000237364.11123.98

Morris, B J, & Wamai, R. G. (2012). Biological basis for the protective effect conferred by male circumcision against HIV infection. *International Journal of STD & AIDS*, *23*(3), 153–159. https://doi.org/10.1258/ijsa.2011.011228

Morris, Brian J, Hankins, C. A., Banerjee, J., Lumbers, E. R., Mindel, A., Klausner, J. D., & Krieger, J. N. (2019). Does Male Circumcision Reduce Women ' s Risk of Sexually Transmitted Infections , Cervical Cancer , and Associated Conditions ?, *7*(January), 1–21. https://doi.org/10.3389/fpubh.2019.00004

Nasr, N., Lai, J., Botting, R. a., Mercier, S. K., Harman, a. N., Kim, M., … Cunningham, a. L. (2014). Inhibition of Two Temporal Phases of HIV-1 Transfer from Primary Langerhans Cells to T Cells: The Role of Langerin. *The Journal of Immunology*, *193*(5), 2554–2564. https://doi.org/10.4049/jimmunol.1400630

Nguyen, V., Mendelsohn, A., & Larrick, J. W. (2017). Review Article Interleukin-7 and Immunosenescence, *2017*.

Nowak, I., & Shaw, L. M. (1997). Effect of Mycophenolic Acid Glucuronide on Inosine Monophosphate Dehydrogenase Activity. *Therapeutic Drug Monitoring*, *19*(3). Retrieved from https://journals.lww.com/drug-monitoring/Fulltext/1997/06000/Effect_of_Mycophenolic_Acid_Glucuronide_on_Inosine.18.aspx

Okoye, A. A., & Picker, L. J. (2013). CD4+T-Cell Depletion In Hiv Infection: Mechanisms Of Immunological Failure. *Immunological Reviews*, *254*(1), 54–64. https://doi.org/10.1111/imr.12066

Otsuka, Y., Watanabe, E., Shinya, E., Okura, S., & Saeki, H. (2018). Differentiation of Langerhans Cells from Monocytes and Their Specific Function in Inducing IL-22–Specific Th Cells, (13). https://doi.org/10.4049/jimmunol.1701402

Patterson, B. K., Landay, A., Siegel, J. N., Flener, Z., Pessis, D., Chaviano, A., & Bailey, R. C. (2002a). Susceptibility to human immunodeficiency virus-1 infection of human

foreskin and cervical tissue grown in explant culture. *The American Journal of Pathology*, *161*(3), 867–873. https://doi.org/10.1016/S0002-9440(10)64247-2

Patterson, B. K., Landay, A., Siegel, J. N., Flener, Z., Pessis, D., Chaviano, A., & Bailey, R. C. (2002b). Susceptibility to human immunodeficiency virus-1 infection of human foreskin and cervical tissue grown in explant culture. *American Journal of Pathology*, *161*(3), 867–873. https://doi.org/10.1016/S0002-9440(10)64247-2

Peña-Cruz, V., Ito, S., Oukka, M., Yoneda, K., Dascher, C. C., Von Lichtenberg, F., & Sugita, M. (2001). Extraction of human Langerhans cells: A method for isolation of epidermis-resident dendritic cells. *Journal of Immunological Methods*, *255*(1–2), 83–91. https://doi.org/10.1016/S0022-1759(01)00432-X

Piguet, V., & Steinman, R. M. (2007). The interaction of HIV with dendritic cells: outcomes and pathways. *Trends in Immunology*, *28*(11), 503–510. https://doi.org/10.1016/j.it.2007.07.010

Polak, M. E., Thirdborough, S. M., Ung, C. Y., Elliott, T., Healy, E., Freeman, T. C., & Ardern-jones, M. R. (2013). Distinct Molecular Signature of Human Skin Langerhans Cells Denotes Critical Differences in Cutaneous Dendritic Cell Immune Regulation. *Journal of Investigative Dermatology*, *134*(3), 695–703. https://doi.org/10.1038/jid.2013.375

Preza, G. C., Tanner, K., Elliott, J., Yang, O. O., Anton, P. A., & Ochoa, M. (2014). Antigen-Presenting Cell Candidates for HIV-1, *30*(3), 241–249. https://doi.org/10.1089/aid.2013.0145

Price, L. B., Liu, C. M., Johnson, K. E., Aziz, M., Lau, M. K., Bowers, J., … Gray, R. H. (2010). The Effects of Circumcision on the Penis Microbiome, *5*(1), 1–12. https://doi.org/10.1371/journal.pone.0008422

Pudney, J., & Anderson, D. J. (1995). Immunobiology of the Human Penile Urethra, *147*(1), 155–165.

Qin, Q., Zheng, X.-Y., Wang, Y.-Y., Shen, H.-F., Sun, F., & Ding, W. (2009). Langerhans' cell density and degree of keratinization in foreskins of Chinese preschool boys and adults. *International Urology and Nephrology*, *41*(4), 747–753. https://doi.org/10.1007/s11255-008-9521-x

Randolph, G. J. (2002). Is Maturation Required for Langerhans Cell Migration ?, *196*(4), 5–8. https://doi.org/10.1084/jem.20021240

Ribeiro, C. M. S., Sarrami-forooshani, R., Setiawan, L. C., Zijlstra-willems, E. M., Hamme, J. L. Van, Tigchelaar, W., … Geijtenbeek, T. B. H. (2016). TRIM5 a in dendritic cell subsets. *Nature Publishing Group*, *540*(7633), 448–452. https://doi.org/10.1038/nature20567

Roederer, M., Nozzi, J. L., & Nason, M. C. (2011). SPICE : Exploration and Analysis of Post-Cytometric Complex Multivariate Datasets, (11), 167–174. https://doi.org/10.1002/cyto.a.21015

Romani, N., Clausen, B. E., & Stoitzner, P. (2010). Europe PMC Funders Group Langerhans cells and more : langerin-expressing dendritic cell subsets in the skin, *234*(1), 120–141. https://doi.org/10.1111/j.0105-2896.2009.00886.x.Langerhans

Salmon, J. K., Armstrong, C. A., & Ansel, J. C. (n.d.). Conferences and Reviews The Skin as an Immune Organ, 181–183.

Schmidt, V. M., Isachenko, V., Rappl, G., Rahimi, G., Hanstein, B., Morgenstern, B., … Isachenko, E. (2018). Comparison of the enzymatic efficiency of Liberase TM and tumor dissociation enzyme : effect on the viability of cells digested from fresh and cryopreserved human ovarian cortex, 1–14.

Schubel, A., Breitfeld, D., Kremmer, E., Renner-mu, I., Wolf, E., Lipp, M., & Tumorgenetics, M. (1999). CCR7 Coordinates the Primary Immune Response by Establishing Functional Microenvironments in Secondary Lymphoid Organs, *99*, 23–33.

Schwartz, O. (2007). Langerhans cells lap up HIV-1 Long-distance affair with adrenal GRK2 hangs up heart failure, *13*(3), 245–246. https://doi.org/10.1016/j.immuni.2007.01.007

Sciences, C. (2016). Characterisation of Mucosal Tissue in the Foreskin after ve rs ity of e To w ve rs ity e To w.

Segura, E., Durand, M., & Amigorena, S. (2013). Similar antigen cross-presentation capacity and phagocytic functions in all freshly isolated human lymphoid organ–resident dendritic cells. *The Journal of Experimental Medicine*, *210*(5), 1035–1047. https://doi.org/10.1084/jem.20121103

Shaw, G. M., & Hunter, E. (2012). HIV Transmission, 1–23.

Shiraki, T., Kondo, S., Katayama, S., Waki, K., Kasukawa, T., Kawaji, H., Hayashizaki, Y. (2003). Cap analysis gene expression for high-throughput analysis of

transcriptional starting point and identification of promoter usage.

Szabo, R., & Short, R. V. (2000). How does male circumcision protect against HIV How does male circumcision protect against HIV infection ?, *320*(October 2008). https://doi.org/10.1136/bmj.320.7249.1592

Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., … Mering, C. Von. (2017). The STRING database in 2017 : quality-controlled protein – protein association networks , made broadly accessible, *45*(October 2016), 362–368. https://doi.org/10.1093/nar/gkw937

van den Berg, L. M., Cardinaud, S., van der Aar, A. M. G., Sprokholt, J. K., de Jong, M. A. W. P., Zijlstra-Willems, E. M., … Geijtenbeek, T. B. H. (2015). Langerhans Cell-Dendritic Cell Cross-Talk via Langerin and Hyaluronic Acid Mediates Antigen Transfer and Cross-Presentation of HIV-1. *Journal of Immunology (Baltimore, Md. : 1950)*, *195*(4), 1763–1773. https://doi.org/10.4049/jimmunol.1402356

Wald, A., & Link, K. (2002). Risk of Human Immunodeficiency Virus Infection in Herpes Simplex Virus Type 2–Seropositive Persons: A Meta-analysis. *The Journal of Infectious Diseases*, *185*(1), 45–52. https://doi.org/10.1086/338231

Wang, J., Wells, C., & Wu, L. (2008). NIH Public Access, *381*(1), 143–154. https://doi.org/10.1016/j.virol.2008.08.028.Macropinocytosis

Wollenberg, A., Kraft, S., Hanau, D., & Bieber, T. (1996). Immunomorphological and Ultrastructural Characterization of Langerhans Cells and a Novel, Inflammatory Dendritic Epidermal Cell (IDEC) Population in Lesional Skin of Atopic Eczema. *Journal of Investigative Dermatology*, *106*(3), 446–453. https://doi.org/10.1111/1523-1747.ep12343596

Yanagihara, S., Komura, E., Nagafune, J., Watarai, H., Yamaguchi, Y., & Alerts, E. (2019). EBI1/CCR7 Is a New Member of Dendritic Cell Chemokine Receptor That Is Up-Regulated upon Maturation.

Zaitseva, M., Blauvelt, A., Lee, S., Lapham, C. K., Klaus-Kovtun, V., Mostowski, H., … Golding, H. (1997). Expression and function of CCR5 and CXCR4 on human Langerhans cells and macrophages: implications for HIV primary infection. *Nature Medicine*, *3*(12), 1369–1375.

Zhou, J., Dann, G. P., Shi, T., Wang, L., Gao, X., Su, D., … Qian, W. (2013). NIH Public Access, *84*(6), 2862–2867. https://doi.org/10.1021/ac203394r.A

Appendix

Formula for counting cells using hemocytometer

$$\text{Total cells/ml} = \text{Total cells counted} \times \frac{\text{dilution factor}}{\text{\# Squares}} \times 10\,000 \text{ cells / ml}$$

Table 1: List of reagents used in the laboratory

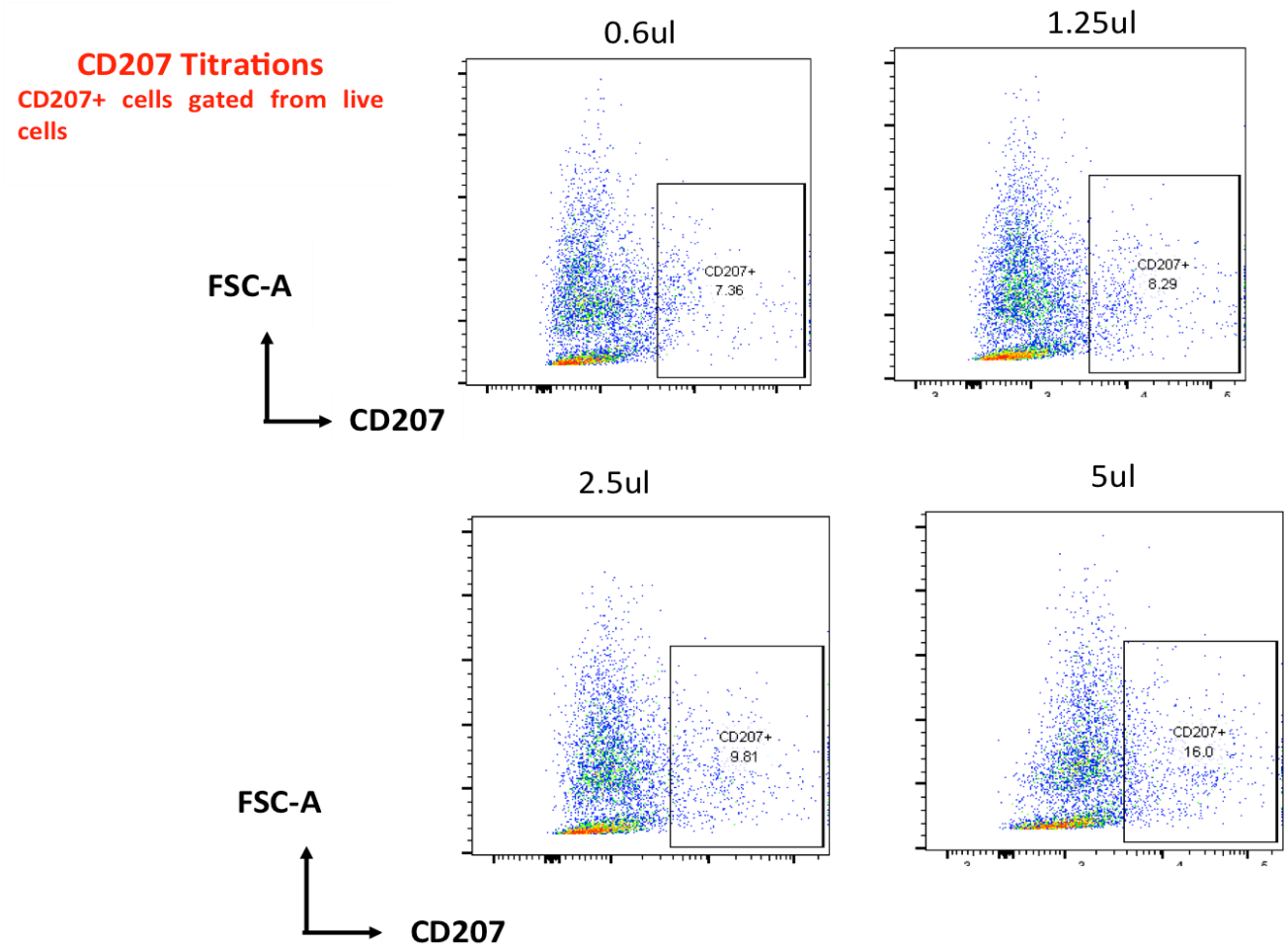
Reagent	Composition	Supplier details	Other details
1 x RPMI-1640 Medium (RPMI)	200mM L-glutamine	Sigma Aldrich 500ml	Cat. No.: R8758
1 x Phosphate buffered Saline (PBS)	0.138M NaCl, 0.0027M KCl (pH 7.2)	Sigma Aldrich 500ml	Cat. No. D8537
BD Facs lysing solution		BD Bioscience	Cat. No. 349202
Dimethly Sulfoxide (DMSO)		Sigma-Aldrich	Cat. No. D2650
Biochrom AG Fetal Bovine Serum (FBS)		Scientific Group 100ml	Cat. No.: S0613
Ficoll- Histopaque®		Sigma-Aldrich 500ml	Cat. No. 10771
Dispase	1g	Sigma	D4693-1G
DMEM	500ml	Thermofischer	Cat No. 31966021
Liberase enzyme			
Liberase solution	25ug/ml + DMEM	*	
Dispase solution	5mg/ml + HBSS	*	
1% wash buffer	1% FBS + 99% PBS	*	
Freezing medium	20% DMSO + 80% FBS	*	
R10	90% RPMI + 10% FBS	*	
RPMI++	90% RPMI + 10% FBS+1%PEN/STREP	*	

*Mixed in the laboratory

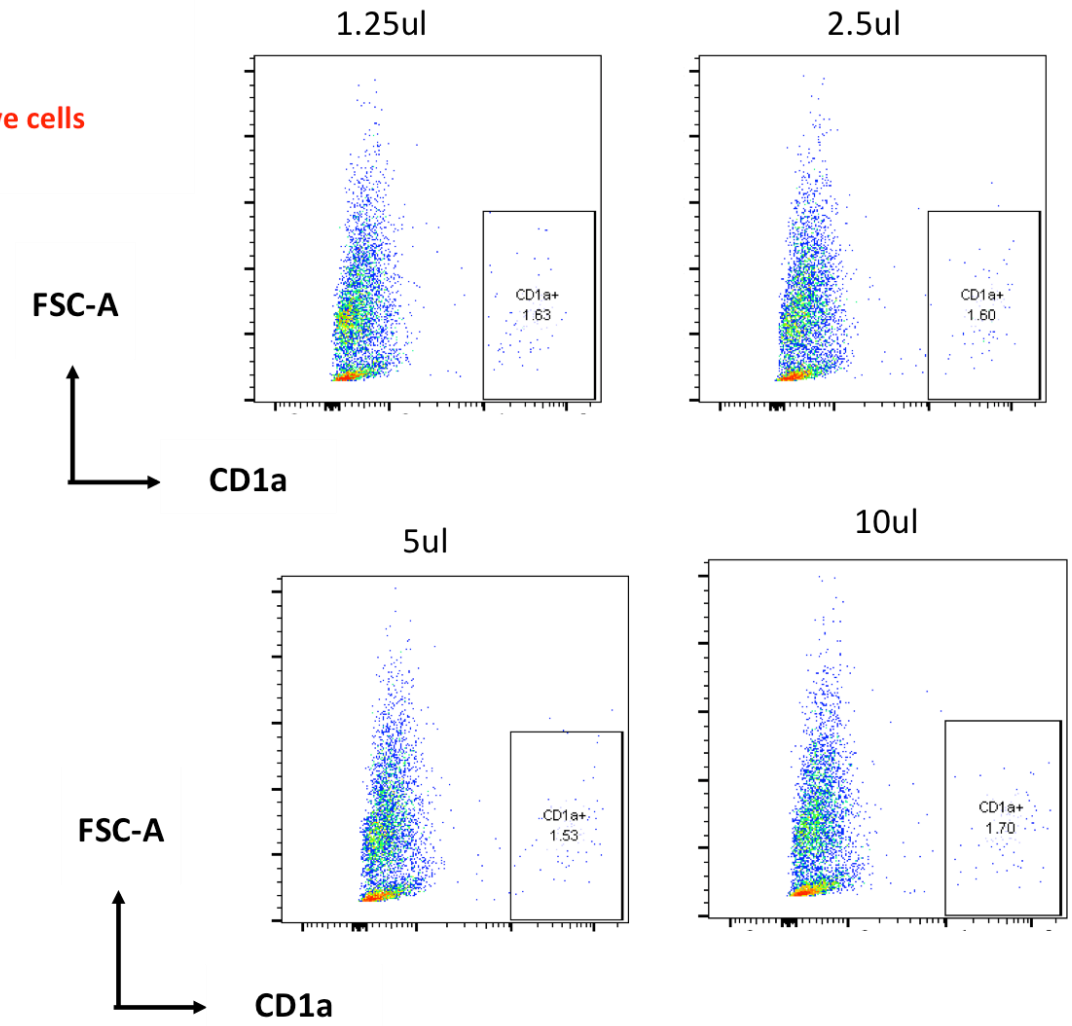
Table 2: List of antibodies used in the study

Antibody	Conjugate/Isotype/clone	Company	Catalogue number
CD207	Pe-Cy7/recombinant human IgGq/REA770	Biocom Africa	130-112-213
CD1a	APC/mouse IgG1/HI149	Biocom Africa	1A-364-T100
CD80	FITC/Armenian hampster/16-10A1	BD biosciences	557226
CD40	PE/mouse/IgG1/5c3	Biocom Africa	12-0409-42
HLA-DR	Per-Cp-Cy5.5/IgG2a/L243	Biocom Africa	307630
CD3	PE Green/mouse IgG1/6B10.2	Biocom Africa	317323
CD86	FITC/rat IgG2a/GL1	BD Biosciences	555657

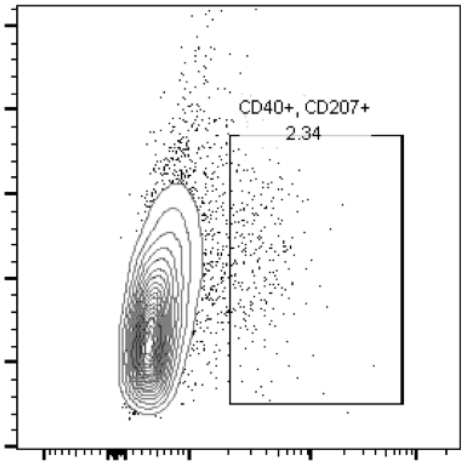
Figure A1: Antibody titration plots: Flow plots of antibody titrations for the different antibodies used for the study. For this analysis, frequency of positive cells was used.

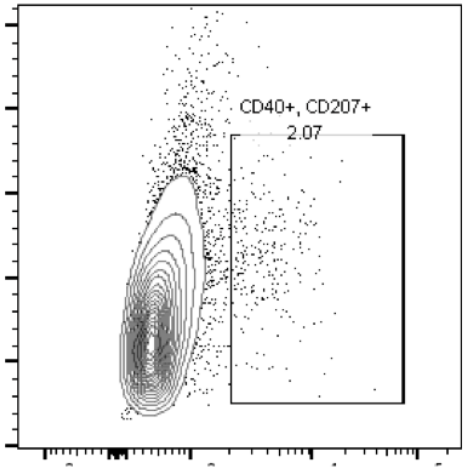


CD1a titration
Gated from live cells



CD40 Titration
Gated from live
CD207+ cells

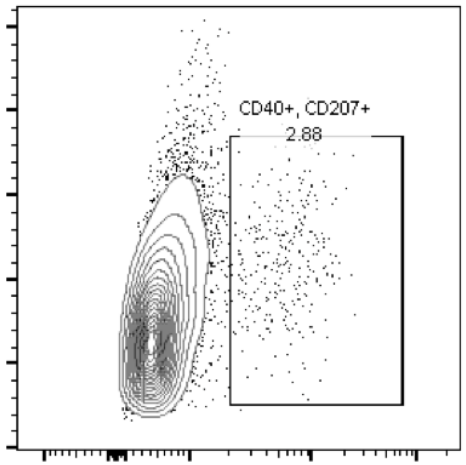




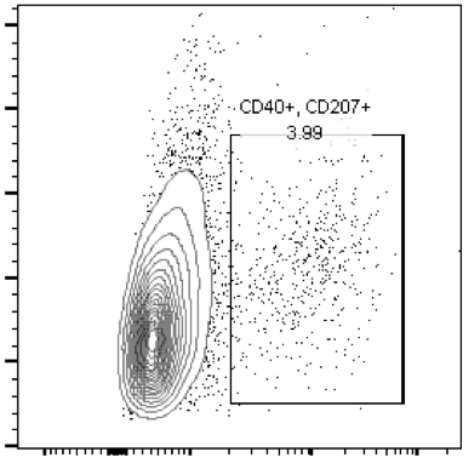
FSC-A

CD40

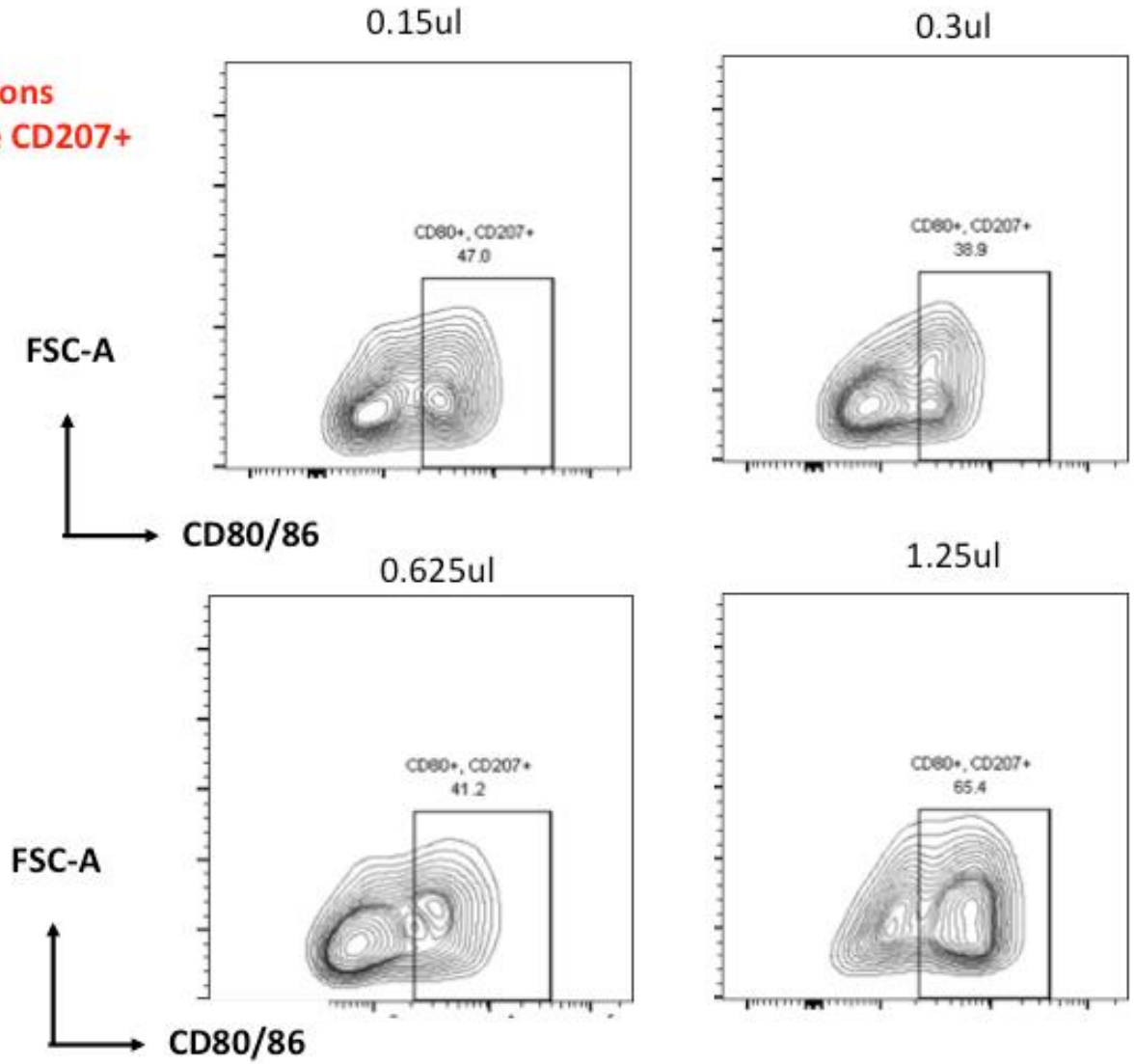
5ul



10ul



CD80/86 Titrations
Gated from live CD207+ cells



HLA-DR Titration
Gated from live CD207 cells

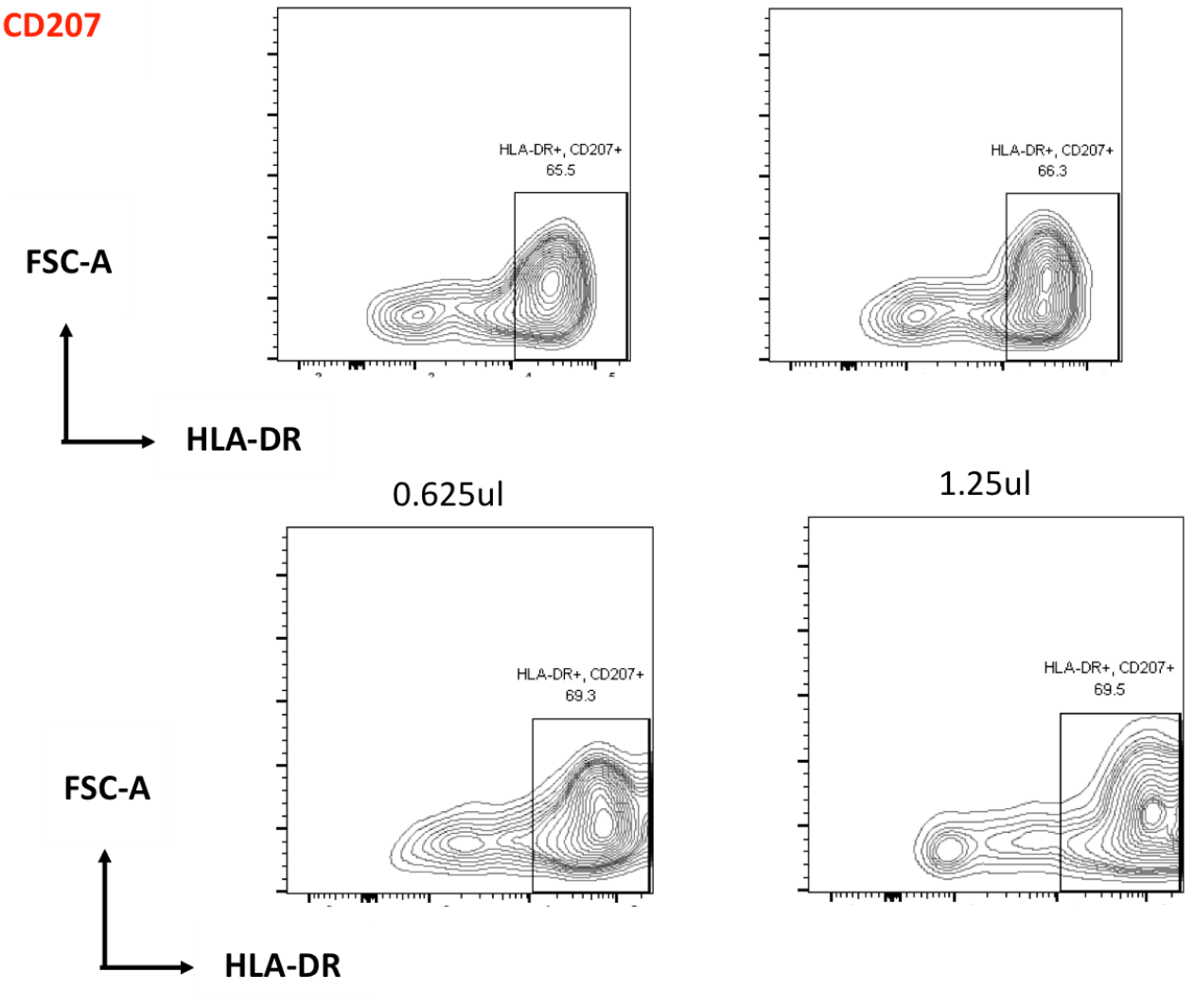
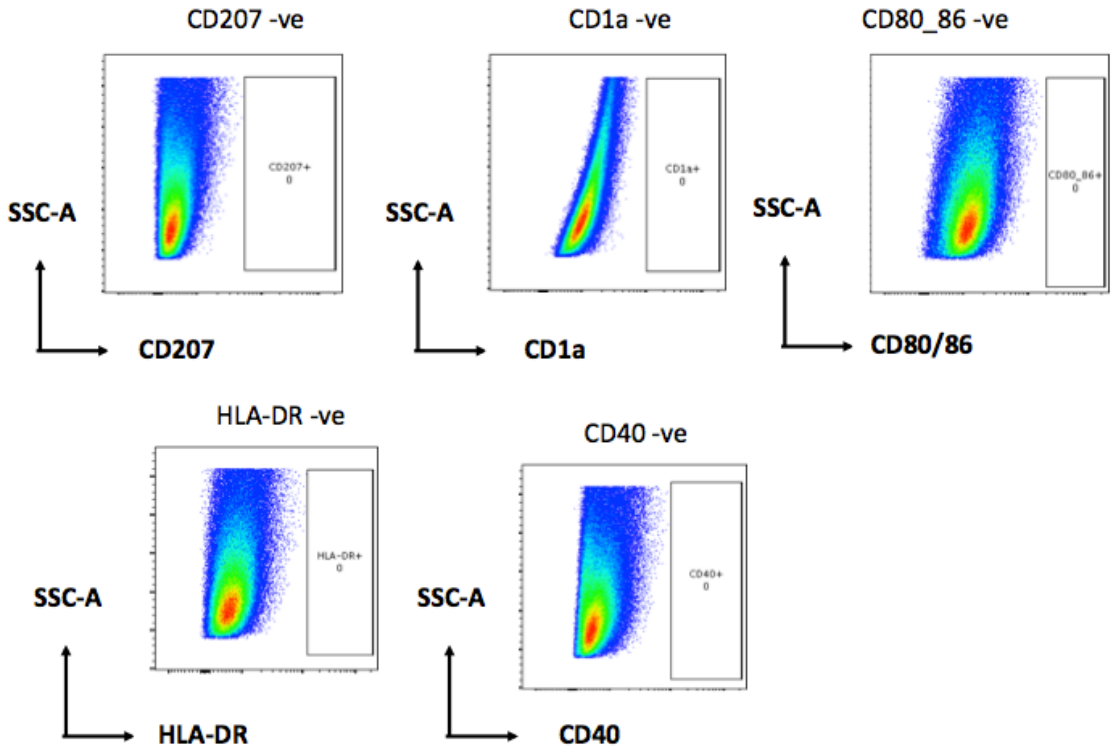


Figure A2 Fluorescence minus one: The antibodies did not have significant auto-fluorescence when combined with other antibodies. Combining all the antibodies used excluding the one in question was used to assess the effect of the other antibodies on auto-fluorescence.



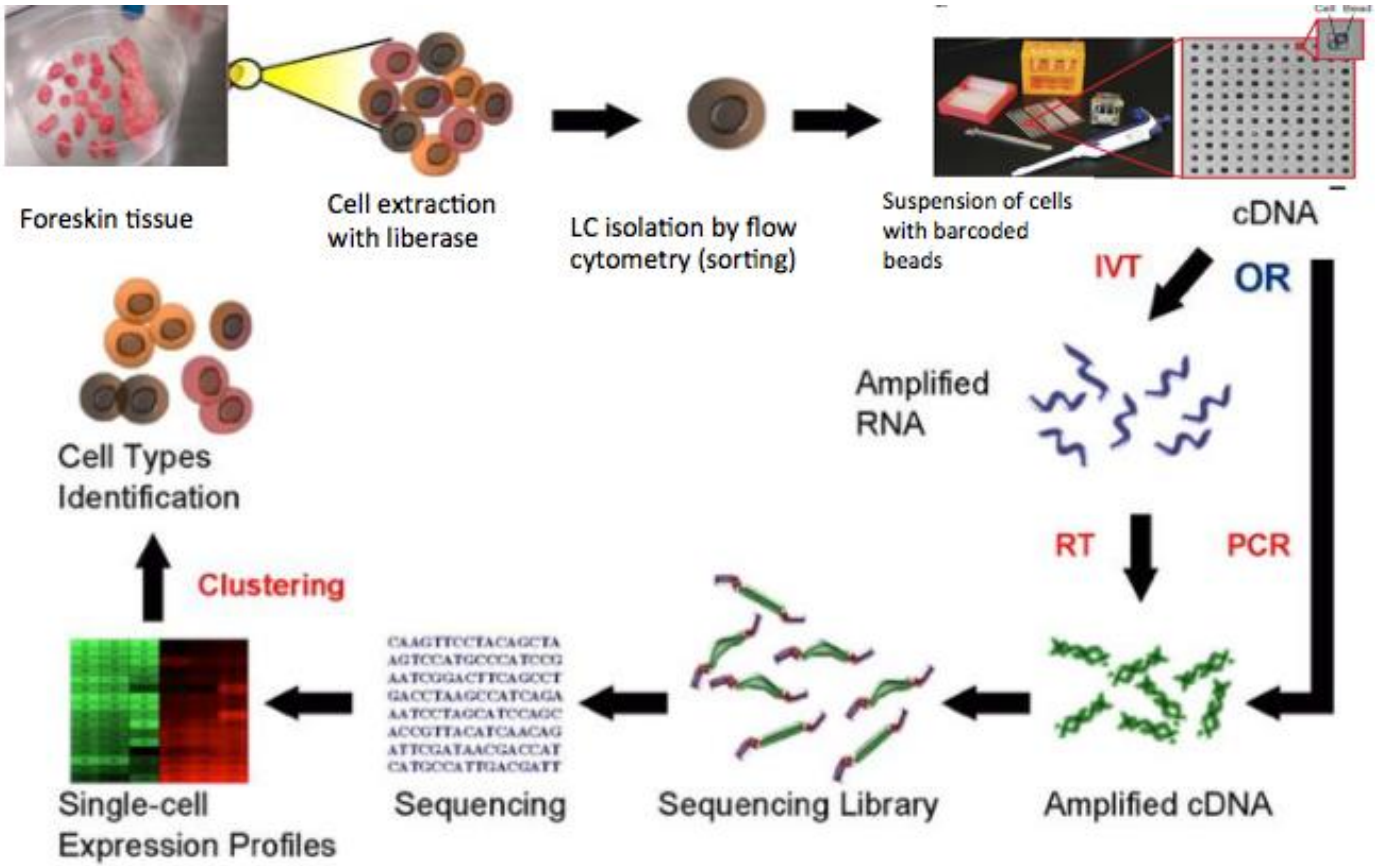
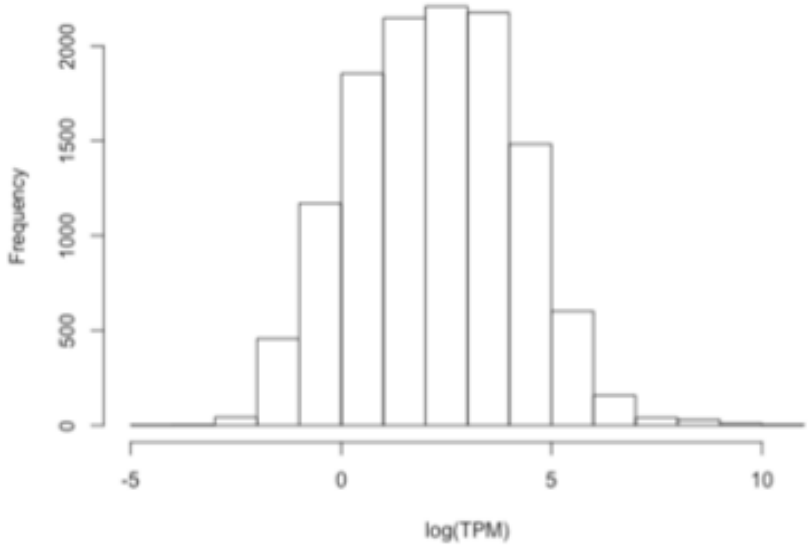


Figure A3: Workflow for seq-well single cell RNA sequencing. The work flow for scRNA seq is similar to bulk RNA seq except that it analyses a single cell. It involves cell extraction from a solid tissue, which was a FS in this study, isolation of single cell done using cell sorter in this case. Suspension of barcoded beads that bind to mRNA which is subsequently amplified as reverse transcribed RNA or second strand synthesis (cDNA) using PCR. The amplified cDNA is subject to sequencing library, sequenced then clustered according to expression profiles. Cell types are identified based on gene expression profiles.

QC2 Expression pattern

12,398 genes are expressed (total gene: 57820)



Expression of Langerhans cell marker genes
Langerin (CD207): 11.14
CD1a: 0
E-cadherin (CDH1): 16.35

Expression of Fibroblast marker genes
S100A4: 2.31
COL1A1:0

Figure A4: Gene expression pattern showing the number of cells that expressed Langerhans cells markers in percentages CD1a, CD207 and E-cadherin. The units of the expression are shown as %. CD1a was not detected to RNA errors or it was expressed in low levels since scRNA seq is not as sensitive as bulk RNA-seq. Gene mapping could be another reason CD1a was not detected. If the CD1a exon sequence is similar to other regions, it would've been impossible to align uniquely and therefore removed in this analysis.

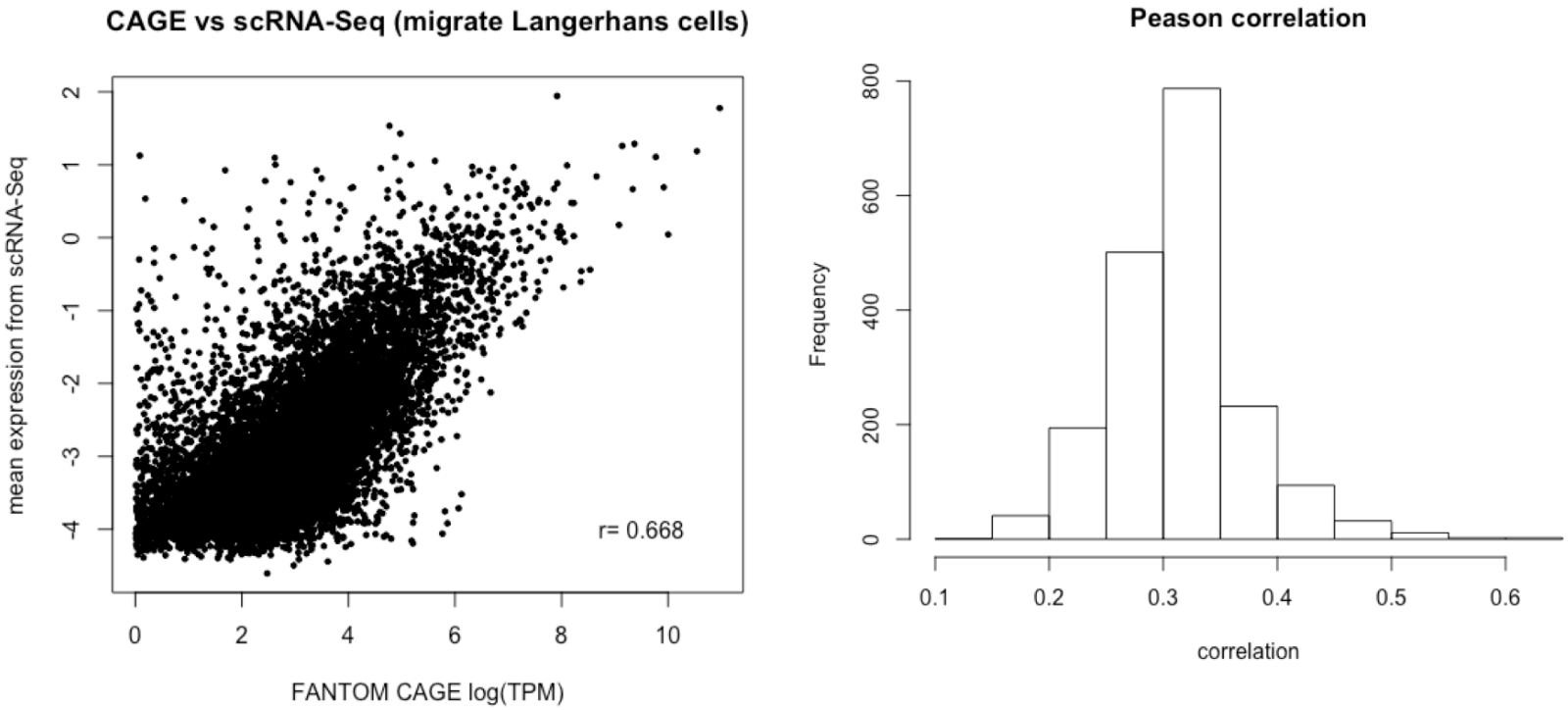


Figure A5: Comparison of sorted LCs with publicly available LCs data. The results showed that the mean expression values detected from 827 cells by Seq-Well showed the highest correlation (R=0.668) with spontaneously migrated Langerhans cells.

Table 3: List of marker genes of different populations found in sorted cells after RNA sequencing

gene	LogFC	p-value	FDR	cluster
RPS29	0,54004	9,34E-11	1,57E-06	0
RP11-318C24.1	0,6083726	7,25E-09	1,22E-04	0
RP11-123J14.2	0,432845	4,16E-08	7,02E-04	0
AC073271.1	0,4325626	1,82E-07	3,07E-03	0
AC016739.2	0,405062	4,27E-07	7,20E-03	0
RPS14	0,3520002	5,82E-07	9,81E-03	0
RPL35A	0,4001776	9,04E-07	1,53E-02	0
LOC124685	0,5562818	1,44E-06	2,44E-02	0
RP11-51O6.1	0,5608041	2,08E-06	3,50E-02	0
RPL13AP5	0,4603693	2,27E-06	3,82E-02	0
RP11-141J13.4	0,4503847	2,97E-06	5,01E-02	0
RPLP1	0,2949477	4,26E-06	7,18E-02	0
RP11-40C6.2	0,5743759	6,07E-06	1,02E-01	0
RP11-367G18.2	0,6171156	8,29E-06	1,40E-01	0
RPL13A	0,4442812	1,05E-05	1,77E-01	0
RP11-179H18.7	0,3743699	1,40E-05	2,37E-01	0
RPS27A	0,4804798	2,40E-05	4,05E-01	0
RPL21P116	0,5240464	2,80E-05	4,72E-01	0
FNBP11	0,5445104	2,80E-13	4,73E-09	1
MTSS11	0,5186553	1,69E-10	2,86E-06	1
KLF61	0,4444636	2,49E-10	4,21E-06	1
BIRC31	0,4265615	2,63E-10	4,44E-06	1
SLC7A111	0,5542338	7,02E-09	1,18E-04	1
MAP4K41	0,3975487	2,61E-08	4,40E-04	1
REV3L1	0,6435401	3,35E-08	5,65E-04	1
LCP1	0,3314161	4,54E-08	7,65E-04	1
NR4A31	0,4932466	7,04E-08	1,19E-03	1
DUSP51	0,4751913	1,35E-07	2,27E-03	1
PDIA31	0,5209469	3,74E-07	6,30E-03	1
ZFAND51	0,4180054	5,65E-07	9,53E-03	1
NUB1	0,3448468	5,82E-07	9,82E-03	1
LYST1	0,4868208	6,26E-07	1,06E-02	1
STK41	0,4472215	8,01E-07	1,35E-02	1
FAM107B1	0,4851795	8,01E-07	1,35E-02	1
TES	0,362302	1,85E-06	3,12E-02	1

CD831	0,4106462	2,03E-06	3,42E-02	1
MALT11	0,4027718	2,74E-06	4,62E-02	1
TYRP1	2,9339305	1,47E-56	2,47E-52	2
IGFBP7	1,7303058	2,44E-45	4,12E-41	2
DCT	2,6000468	2,80E-45	4,72E-41	2
TYR	1,4828524	1,89E-43	3,19E-39	2
PLP1	1,9644584	2,44E-41	4,12E-37	2
PMEL	2,121655	7,18E-40	1,21E-35	2
MLANA	2,1969188	9,58E-40	1,62E-35	2
EDNRB	1,8767352	3,16E-39	5,32E-35	2
GPNMB	2,0462907	9,22E-34	1,56E-29	2
CAPN3	1,3568065	7,73E-31	1,30E-26	2
DSTYK	1,4621893	9,07E-31	1,53E-26	2
KIT	1,8664234	3,80E-28	6,42E-24	2
CYB561A3	1,9278106	7,27E-26	1,23E-21	2
MITF	1,1982042	1,85E-25	3,12E-21	2
EMP1	2,0031385	4,74E-25	8,00E-21	2
TFAP2A	1,3610165	8,63E-22	1,46E-17	2
MYO10	1,417237	5,76E-18	9,72E-14	2
TRPM1	1,6900737	5,52E-15	9,32E-11	2
CCNG1	1,0620376	2,90E-14	4,89E-10	2
PHLDA1	1,3459827	3,33E-14	5,61E-10	2
TDRD3	1,4767432	4,15E-13	7,00E-09	2
CD44	1,4140601	4,29E-12	7,24E-08	2
SEMA3C	1,7806307	7,86E-12	1,33E-07	2
OSTM1	1,1801769	1,46E-11	2,47E-07	2
ZMAT3	1,582867	5,44E-10	9,18E-06	2
DNAJC13	1,2232069	8,12E-10	1,37E-05	2
KLF9	0,922333	2,42E-07	4,08E-03	2

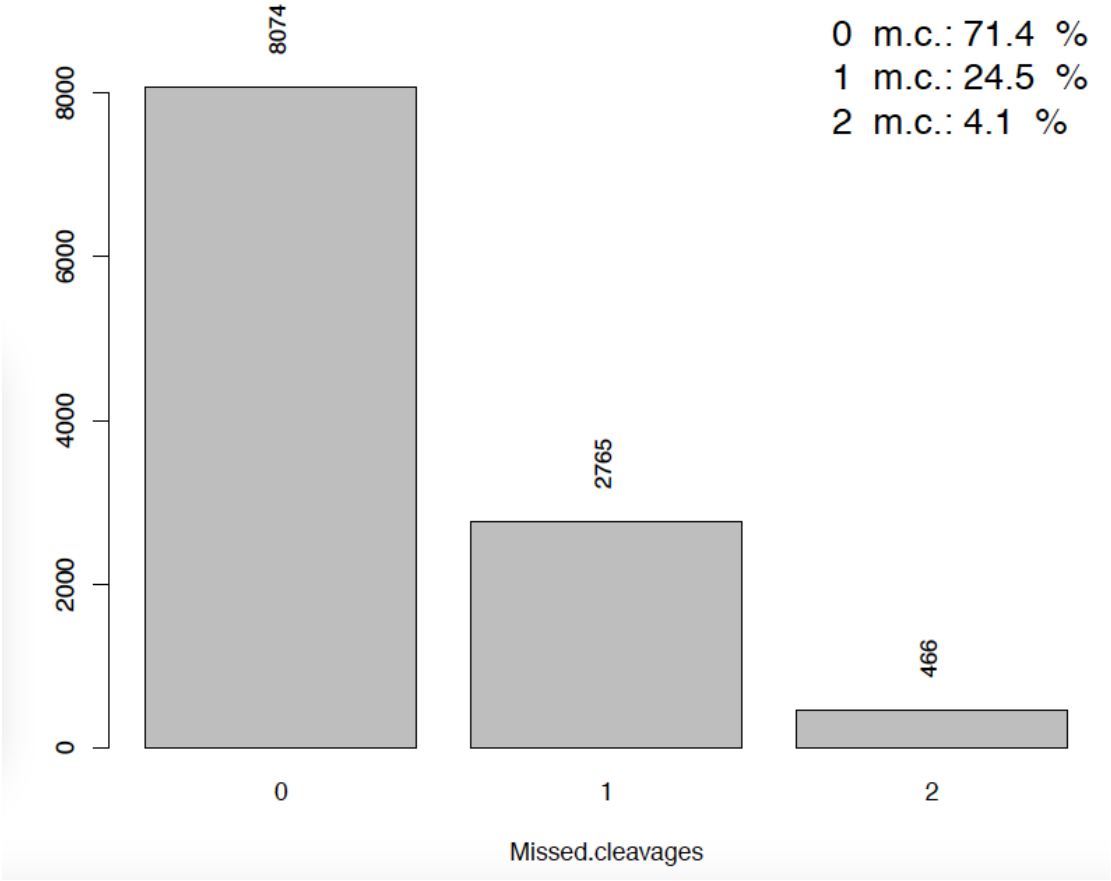


Figure A6: Number of missed cleavages, peptides that were undigested during trypsin digestion.

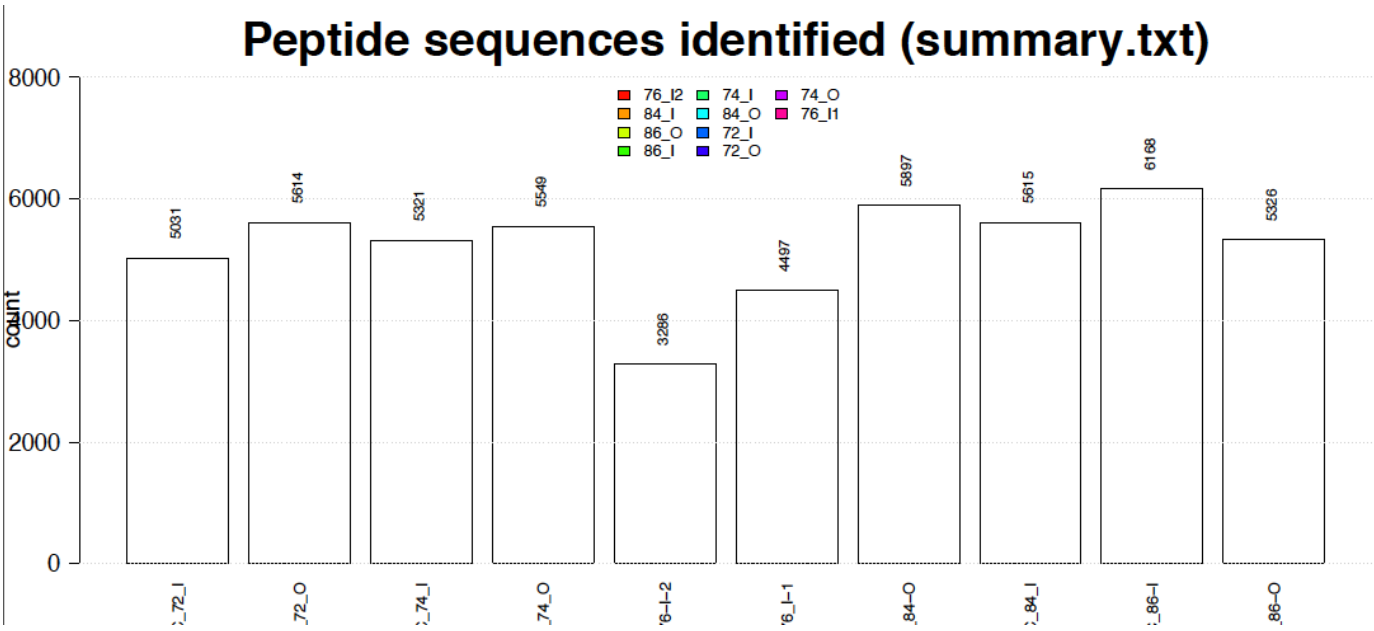


Figure A7: Number of peptide sequences identified per sample.