The rectal microbiome in HIV infected Kenyan men who have sex with men (MSM): associations with sexual behaviour and inflammation in the rectum and blood

by

Henok Gebrebrhan

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

Master of Science

Department of Medical Microbiology and Infectious Diseases

University of Manitoba

Winnipeg

Copyright © 2022 by Henok Gebrebrhan

Abstract

HIV infection is responsible for massive destruction of T cells in the gut associated lypmphoid tissue. The damage is irreversible and affects the host immunity allowing the gut homeostasis to be disturbed. This disruption can translate to gut microbiota alterations, microbial translocation and chronic immune activation. Interest in studying the rectal microbiome has since accelerated as alteration to the microbiota has been implicated in diseases and found to be influenced by both the environment and genetics. One of the interesting factors that affect gut microbiota is sexual orientation as it was shown that MSM status had a distinct microbiota than non-MSM. There have been though few studies accounting for sexual behaviour practises and its effect on the gut microbiota and there is also limited research from sub Saharan Africa.

In this thesis, we investigated whether HIV and different sexual behaviours are linked to gut microbiota composition in an MSM cohort from Nairobi, Kenya. We found alpha diversity was consistently lower in Kenyan HIV-infected MSM, including those on antiretroviral therapy (ART) compared with HIV-uninfected MSM. Receptive anal sex with several types of sexual partners (paying, regular, casual) was associated with lower Chao1 and Simpson diversity, independent of HIV status.

We also hypothesized that MSM engaging in receptive anal sex might have microbial translocation and associated with immune activation due to local trauma to the rectal mucosa. We found limited evidence of markers of MT in HIV infection but that MSM who engage in receptive anal sex had elevated markers of MT that was correlated with systemic CD4+ T cell activation. We also compared differences between adherent bacteria and lumen bacteria by collecting rectal samples before and after enema to resemble fecal and mucosal samples respectively. We found the fecal samples to have less diversity than mucosal samples and that the mucosal samples were associated with systemic CD4+ T cell subsets.

These results demonstrate that both HIV infection and receptive anal sex affect the gut microbiome in MSM. We have also shown the feasibility of carrying out rectal sampling that captures mucosal adherent bacteria without the need for biopsies. This thesis adds to the body of knowledge on MSM behavioural effects on the gut microbiota and can be used for strategies to aimed at reducing HIV susceptibility in this high-risk population.

Dedication

This thesis is dedicated to my parents. Thank you for instilling in me your love, strength, courage and perseverance. I would not have been here now succeeding in my aspirations had it not been for the traits you taught me which gave me the drive to work hard to achieve my goals in life.

Acknowledgements

I would like to thank all the wonderful people I have come across in the department that has allowed me to make this thesis possible.

I would like to first thank my supervisor Dr. Lyle McKinnon. You took a chance on me as one of your first co-op students while the lab you were setting up your lab. I did not know what to expect but I quickly realized that I was in the right place once I got to know your personality, drive and aspirations. I always looked forward to coming to your office and I thank you for always extending your hand to guide me throughout my journey. All those moments in the hallway, cafeteria and public outings helped me in knowing that pursuit of graduate studies in your lab was an easy decision. Thank you for your guidance in choosing my career path and for making work feel like another day at the soccer field.

I thank members of my committee; Dr. Blake Ball, Dr. Paul McLaren and Dr. Jean-Eric Ghia. I appreciated your mentorship, guidance, feedback and for all the suggestions you gave me to successfully complete my thesis.

Thank you to all the support staff I had the pleasure to meet in the department. Thank you Angie and Jude for all the help in finding the right documents, submitting them on time, sending me reminders, arranging flights and for directing me to the right places for all my everyday needs.

I would also like to thank all the technicians in the department for their tireless work contribution to my development as a well rounded researcher. Thank you Adriana, Steve, Sue and Tash.

Thank you to all McKinnon lab members past and present who I have had the pleasure to meet. Thank you Andrew, Aida, Heather, Domenico, Ruth, Naima,Tosin, Sandra, Cheli, Faisal, Frank and Aloysious for being wonderful lab mates and great friends who I could always count on for support. I would like to give thanks to Andrew for instructing me in general lab use, flow cytometry and help me settle down when I joined the lab. I would also like to give thanks to Paul McLaren lab personnel and others I worked with at the JC Wilt for making me feel at home. It was a pleasure working with all of you and thanks for all the great memories I will cherish from the lunchroom to the outings where we all gathered together.

Thank you to all of my graduate and postdoctoral peers for the wonderful discussions, lunch breaks, picnics, and travel experiences. Thanks for all the good time and friendship along the way and I am sure our careers will cross path in the near future.

I would finally like to show gratitude for my family for all the love they showered on me during the good and tough times. I know you always have the best interest in me and appreciate all the courage you instilled in me. Habits start at home and I have been blessed to be in a loving and caring family that always drives me to reach my goals.

Table of Contents

Abstract ..... **i**

Dedication ..... **ii**

Acknowledgements ..... **iii**

List of Figures ..... **viii**

List of Tables: ..... **x**

Publications arising from this thesis ..... **xi**

Contributions of Authors ..... **xii**

1. Introduction **1**

1.1 The HIV epidemic 1

1.2 HIV virology 2

1.2.1 HIV classification and structure 2

1.2.2 HIV Life cycle 3

1.3 HIV transmission 4

1.4 HIV treatment and prevention 4

1.4.1 HIV treatment 4

1.4.2 HIV prevention 5

1.5 HIV pathogenesis 8

1.5.1 Overview of the human immune system 8

1.5.2 Innate immunity 8

1.5.3 Adaptive immunity 10

1.5.4 Host immune response against HIV 12

1.6 Gut microbiome overview 14

1.6.1 Factors affecting gut microbiome 16

1.6.2 Microbiome and health implications 16

1.6.3 Gut microbiome alterations associated with HIV 18

1.6.4 Factors influencing gut microbiota in MSM 20

1.6.5 Epidemiology of Kenyan MSM 21

Study Rationale, Hypothesis, and Objectives **23**

Study Rationale 23

Hypothesis 23

Objectives 24

2. Materials and Methods 25
2.1 General Reagents 25
2.1.1 Solutions 25
2.1.2 Stimulation reagents 25
2.2 Biological Samples 26
2.2.1 Ethics 26
2.2.2 Study participants 26
2.2.3 Demographic data collection 26
2.3 Sample collection and preparation 27
2.3.1 Blood collection and processing 27
2.3.2 Rectal sample collection and processing 27
2.4 General methods 28
2.4.1 16S rRNA sequencing 28
2.4.2 Flow Cytometry 31
2.5 HIV testing 33
2.6 HIV viral load analysis 34
2.7 STI testing 34
2.8 Elisa assay 35
2.8.1 Biorad assay 35
2.8.2 MSD assay 35
2.8.3 Microbial translocation assays 36
2.8.4 Elisa assay conditions and analysis 37
2.9 Statistics 37
3 Impact of HIV infection on gut microbiome 39
3.1 Rationale 39
3.2 Hypothesis 40
3.3 Objective 40
3.4 Section specific methods 40
3.4.1 Study design 40
3.4.2 Data collection 41
3.4.3 Specimen collection 41
3.4.4 16S rRNA targeted metagenomic sequencing of rectal microbiota 41
3.4.5 Sexually transmitted infection testing and HIV viral load 42

3.4.6 Statistical analysis 42
3.5 Results 43
3.5.1 Study participants 43
3.5.2 HIV infection and microbiome composition 45
3.5.3 Differential abundance of genera by HIV infection 48
3.5.4 Associations of sexual behaviour and microbiome diversity 50
3.6 Discussion 53
4. Microbial translocation in HIV infected MSM 56
4.1 Rationale 56
4.2 Hypotheses 57
4.3 Objective 58
4.4 Chapter specific methods 58
4.4 Results 58
4.4.1 Study participants and experimental conditions 58
4.4.2 Microbial translocation markers expressed at higher concentrations in HIV infected and uninfected MSM. 59
4.4.3 Microbial translocation markers elevated in MSM who engage in receptive anal sex 60
4.4.4 Microbial translocation markers were elevated in MSM compared to other low risk groups . 62
4.4.5 Association between markers of MT, microbiome diversity, systemic immune activation and rectal secretion immune activation 64
4.4.6 Microbial translocation markers not associated with gut microbiome clusters 66
4.5 Discussion 67
5. Mucosal and fecal sampling comparison 72
5.1 Rationale 72
5.2 Hypothesis 73
5.3 Objective 73
5.4 Section specific methods 74
5.5 Results 74
5.5.1 Gut microbiome composition between sites 74
5.5.2 Differential abundance between sample sites 76
5.5.3 Relationship between rectal inflammatory markers and microbiome clusters 77
5.5.4 Relationship between systemic T cells and microbiome composition 81
5.6 Discussion 86
6. Discussion 89

6.1 Major findings 89
6.2 Hypothesis and conclusion 94
6.3 Future directions 95
6.4 Concluding remarks 96
7. References 97
8. Appendix 109
8.1 Abbreviations 109
8.2 Supplementary data 112

List of Figures

Figure 1. Associations between HIV-defined groups and microbiome structure and diversity...47
Figure 2. Specific genera associated with HIV status using forward selection analysis generated in the *selbal* R package.....49
Figure 3. Report of different types of sexual activity with regular, casual and paid partners (denoted by 'reg', 'cas', 'pay'), and their association with alpha diversity in HIV-negative men…….....51
Figure 4. Microbial translocation in HIV infected MSM……..…….....60
Figure 5. Elevated microbial translocation in HIV negative MSM engaging in receptive anal sex……..61
Figure 6. Elevated microbial translocation in MSM cohort compared to healthy donor samples.....63
Figure 7. Heat map of spearman correlation analysis…….....65
Figure 8. Microbial translocation is not associated with gut microbiome cluster 67
Figure 9. Comparison of sampling method by microbiome structure and diversity 75
Figure 10. Volcano plot showing genera significantly different by sampling method…….....76
Figure 11. Microbiome cluster association with rectal secretion soluble markers of inflammation.....78
Figure 12. Multiple factor analysis of microbial clusters with rectal secretion soluble markers of inflammation…….....79
Figure 13. Effect size of association between microbial cluster and rectal inflammatory biomarkers.....80
Figure 14. Microbiome cluster association with systemic CD4+ T cell subsets.....82
Figure 15. Multiple factor analysis of microbial clusters with systemic CD+ T cell subsets.....84
Figure 16. Bar plot of log fold changes between microbial cluster and systemic CD+ T cell subsets in both sample types.....85
Figure 17. Specific genera associated with HIV status using forward selection analysis generated in the selbal R package.....117
Figure 18. Representative staining of PBMCs for panel 2 following four hour stimulation with PMA and Io..... 118

Figure 19. Representative staining of PBMCs from panel 1 for T cell activation.....119
Figure 20. Representative staining of PBMCs from panel 1 for memory T cell differentiation.....120

List of Tables:

Table 1. Clinical and demographic characteristics of the study population..... 44
Table 2. Linear regression analysis of clinical and behavioural variables associated with rectal microbiome alpha diversity 52
Table 3. Relative abundance of genera among participants..... 112
Table 4. Associations with microbiome cluster in HIV negative MSM 114
Table 5. Flow cytometry panels description 115

Publications arising from this thesis

Gebrebrhan H, Kambaran C, Sivro A, Adhiambo W, Siele N, Becker MG, Li J, Choi S, Mwatelah RS, Reyes NV, Akolo M, Njogu P, Cholette F, Ho J, Kim J, Peterson SW, Martin I, Sandstrom P, Mehta SD, Lorway RR, Ball TB, Kimani J, McLaren PJ, Ji H, McKinnon LR. Rectal microbiota diversity in Kenyan MSM is inversely associated with frequency of receptive anal sex, independent of HIV status. AIDS. 2021 Jun 1;35(7):1091-1101. doi: 10.1097/QAD.0000000000002829. PMID: 33534201.

Contributions of Authors

Contributions of authors from the manuscript titled “Rectal microbiota diversity in Kenyan MSM is inversely associated with frequency of receptive anal sex, independent of HIV status.” at chapter three.

Designed the study: AS TBB J. Kimani PJM HJ LRM

Performed the experiments: HG CK AS WA SC RSM NVR JL FC JH SP

Analysed the data: HG CK AS MGB JH SP SDM PJM LRM

Wrote the paper: HG CK AS WA NS MGB SC RSM NVR JL MA PN FC JH J. Kim SP IM PS SDM TBB J. Kimani PJM HJ LRM

Supervised clinical and/or experimental aspects of the study: NS MA PN J. Kim IM PS TBB J. Kimani HJ LRM

All authors listed have made a significant and direct contribution to the study and approved it for publication.

1. Introduction

1.1 The HIV epidemic

The HIV/AIDS pandemic has contributed to an enormous health toll on society over the last 4-5 decades. The first description of HIV infection was reported in 1981 in homosexual men from San Francisco and New York who had presented *Pneumocystis pneumonia* and rare malignancy Kaposi's sarcoma, which indicated compromised immunity [1]. According to recent numbers from UNAIDS in 2016, there are approximately 1.8 million new HIV infections, 1 million deaths attributed to AIDS annually and approximately 37 million people currently living with HIV/AIDS [2]. HIV prevalence varies considerably by geographic regions, with lower- and middle-income countries (LMIC) bearing the majority of the burden. The estimate from 2017 was that Sub-Saharan Africa accounted for 71% of people living with HIV, 75% of deaths and 65% of new infections globally [3]. HIV continues to disproportionally affect different "key" populations, with men who have sex with men (MSM), injection drug users and sex workers having higher prevalence than those who do not identify as belonging to these populations[4,5].

The severity of the HIV/AIDS pandemic made it an urgent priority for global health and reducing the HIV burden was included in the Millennium Development Goals (MDG) by the United Nations General Assembly in 2000. MDG number 6 specifically aimed to stop the spread of HIV/AIDS incidence and ensure accessible and affordable therapy for those who seek it by the year 2015. Although major gains have been made to achieve this goal, HIV/AIDS is still a major public health concern. The new goal that has been proposed by the Joint United Nations Programme on HIV/AIDS (UNAIDS) is called the 90-90-90 strategy which aimed to have 90% of those with HIV to know their status, 90% of those identified as positive to be on sustained therapy and 90% of those on therapy to achieve viral suppression by 2020 [2]. If achieved, it would mean 73% (90% x 90% x 90%) of all HIV positive people will be virally supressed although these goals have proven to be unrealistic [6]. Although the goals have not been met in the specified timeline, significant progress continues to be made towards reducing the burden of HIV/AIDS as a public health threat.

1.2 HIV virology

1.2.1 HIV classification and structure

HIV is a member of the family *Retroviridae*, sub-family *Lentrivirinae* and genus *Lentivirus*. The prefix "lenti" is used for slow viruses that have long incubation periods that can range from months to years and can induce a variety of pathologies. HIV is of zoonotic origin after cross-species transmission of simian immunodeficiency virus (SIV) from primates to humans [7]. Two sub-types of HIV have been classified; HIV-1 and HIV-2. The predominant type accounting for 95% of infections around the world is HIV-1, demonstrating increased virulence and pathogenicity compared to HIV-2. HIV-2 is mostly found in western Africa and is less virulent, reaching a lower viral set point, which decreases transmission and damage to the immune system than HIV-1 [8]. HIV-2 more commonly results in long term non-progressive (LTNP) infection. This could be due to the high similarity between SIV and HIV-2. The SIV precursor to the HIV-1 strain comes from apes such as chimpanzees and gorillas and is likely adapted to its hosts. HIV-1 can be further split into 4 subgroups (M, N, O, P) based on their lineages. Out of these groups, M is the major group and accounts for the majority of HIV infections. This M group can also be sub-divided into at least nine clades (A, B, C, D, F, G, H, J and K). Clades A and C, which are primarily present in Africa and some regions of southeast Asia, account for greater than 50% of new infections worldwide, while clade B dominates in western nations. There is additional genetic diversity in HIV due to recombination between various subtypes resulting in increasing prevalence of circulating recombinant forms (CRF) and unique recombinant forms (URF) [9].

The HIV virion encompasses two copies of positive-sense single stranded ribonucleic acid (RNA) found within the center of the virus. There are nine viral genes which can be grouped functionally as structural, regulatory and accessory genes. The *gag* genes encode the structural proteins of the nucleocapsid, capsid and matrix. The *env* gene encodes the gp160 glycoprotein that is cleaved into envelope proteins gp120 and gp41 and used for virus entry to the cell [10]. The *pol* gene encodes for reverse transcriptase (RT), integrase and protease which aid in viral replication. These enzymes are responsible for converting retroviral RNA to proviral DNA,

integration into the host genome and post translation modifications. The accessory and regulatory genes play a number of key roles in modulating virus replication.

1.2.2 HIV Life cycle

The HIV life cycle can be summarized into six distinct steps; 1) attachment; 2) capsid uncoating and reverse transcription; 3) integration ; 4) virus DNA transcription and translation; 5) assembly 6) budding. HIV entry into the target cells begins when its surface trimer gp120-gp41 heterodimer binds to the CD4 receptor expressed on CD4+ T cells and macrophages. Binding to the CD4 receptor results in the virus envelope having a conformational change allowing its chemokine binding domain to bind the co-receptor C-C motif chemokine receptor type 5 (CCR5) or C-X-C motif chemokine receptor type 4 (CXCR4) present on the target cell membrane. Most HIV strains sampled during acute infection and the most transmitted are CCR5 tropic which are most abundant in memory CD4+ T cells [11]. The result of gp120 binding to CD4 and a coreceptor (CCR5 or CXCR4) is a stable complex that subsequently promotes fusion of viral and target cell membrane [12]. Subsequent structural changes by gp41 enable fusion of membranes and entry of the virus capsid. The viral core is transported to the cytoplasm for viral RNA uncoating from the capsid proteins. Reverse transcription in the cytoplasm by the RT enzyme converts the RNA to DNA. The enzyme integrase facilitates integration of the cDNA into the host chromosome. The major target cell population of HIV infection are activated CD4+ T cells as they allow transcription of proviral DNA into RNA and abundantly express the co-receptor CCR5 [13]. Viral messenger RNA is spliced and moved to the cytoplasm for translation. New virions are formed when viral RNA strands associate and assemble with core proteins as the particle migrates to the cell membrane. Cleavage by HIV protease of viral precursor proteins assists in maturation of infectious viral particles which buds off from the host cell to search for target cells to infect.

1.3 HIV transmission

HIV transmission can occur through exposure to infected fluids during sexual contact, blood transfusion, sharing of needles and during birth from mother to baby. HIV is categorized as a sexually transmitted blood borne infection (STBBI) as the most predominant route of transmission is through sexual contact, especially in the hardest hit region of sub-Saharan Africa [14]. The risk of sexual HIV transmission, without use of condom or ARV, varies by type of sexual act. The estimated HIV transmission risk per 10,000 exposure ranged from low for oral sex (95% CI 0-4), followed by insertive penile-vaginal sex 8 (95% CI 6-11) and receptive anal intercourse 138 (95% CI 102-186)[15]. Anal intercourse has the highest transmission risk that could be due to the rectal mucosa surface being much thinner and more brittle than the vaginal mucosa. These relatively low per-coital rates can vary widely and increase in presence of co-factors of HIV transmission such as high viral load, genital ulcer disease, STIs, HIV infection stage and inflammation [15–17]. Circumcision significantly reduces HIV incidence as it removes access to receptive cells in the foreskin [18].

1.4 HIV treatment and prevention

1.4.1 HIV treatment

Since its discovery, much attention has been given to the development of HIV treatments and/or cures. The first drug approved for HIV treatment was zidovudine (AZT) in 1987 and has ever since grown substantially to include more than 40 drugs that can are readily available for patients. These drugs belong to six mechanistic classes that target various parts of the HIV life cycle. The most effective treatment was established in the late 1990s as combination antiretroviral therapy (cART) or commonly known as ART. This initial successful regimen was composed of two nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs) and a protease inhibitor. Advances in HIV care have resulted in the introduction of other classes such as integrase inhibitors that can be used in combination with NRTIs. Integrase inhibitors have gained a leading role in HIV treatment due to their favourable safety profile and clinical characteristics [19]. ART is highly effective in supressing plasma viremia, HIV associated morbidity and prolongs survival [20–23]. The viral suppression allows the host immune system to partially recover its

functionality by stopping further damage to immune cells although reconstitution of CD4+ T cells is incomplete. There was a major debate on when to start ART in the early days after it became available. Given limited resources, ART-related toxicity, and concerns on treatment adherence, a cautious approach was taken in the early 2000s to start ART for persons with CD4+ cell count less than 200 cells/mm[24]. The CD4+ cell count threshold was subsequently raised over the years to 500 cells/mm as it was shown conclusively that early initiation reduced mortality and disease progression [25]. The World Health Organization (WHO) now recommends immediate initiation of therapy after HIV diagnosis regardless of CD4+ cell count. Unfortunately, ART cannot serve as cure as it fails to clear latent reservoir of infected CD4+ T cells formed during acute infection. HIV latency establishment starts as early as Fiebig I stage when the first nucleic acid particles are detected in the first two weeks following infection before seroconversion. ART initiation during Fiebig I does not prevent HIV latency establishment as treatment interruption leads to HIV RNA rebound in approximately 3-4 weeks [26]. The drawback of ART is that it is currently a lifelong daily therapy. There have been recent advances in long-acting injectable ART that have been shown to be effective and tolerable. Long – acting injectable cabotegravir (CAB-LA) was the first long acting HIV prevention medication to be approved by the US Food and Drug Administration (FDA) in 2021. It is injected once every two months and is approved for HIV patients who are virologically suppressed and have no history of treatment failure [27].

1.4.2 HIV prevention

Despite significant progress in HIV treatment and prevention, HIV incidence has remained elevated although it continues to decrease after peaking in 1999 with 3.16 million cases a year to 1.94 million cases a year in 2017 [28]. Public health efforts that have aimed to modify behavioural risk factors (promotion of monogamy, abstinence and reduction in the number of sexual partners) have enjoyed limited success, leading to development of numerous biomedical interventions. Effective interventions such as circumcision can reduce HIV acquisition by 60% in men while proper condom use trims HIV acquisition by up to 80% in both male and females [15,29,30]. It has long been debated whether ART could have a dual benefit of both treating and preventing HIV transmission. It has been shown that decreased plasma HIV RNA after ART initiation is associated with low to undetectable viremia in genital secretions [31–33]. Plasma viral load of

<1500 copies/mL has been shown to result in low HIV transmission rate [32,34]. It was thought that because HIV sexual transmission is highly correlated to viral concentration in the blood [34] and genital tract secretions [31], ART may be able to reduce sexual transmission of HIV. The HPTN 052 study conclusively showed for the first time the benefits of using ART as a prevention tool against HIV infection [35]. HPTN052 enrolled 1763 discordant couples in which one person in the couple was HIV positive and had a CD4+ cell count of between 350 and 550 CD4 cells per milliliter (e.g. had not yet progressed to AIDS). Couples were randomized to immediate ART or standard of care until appearance of symptoms or CD4+ cell count below 300 cells/ml. Early ART reduced the number of genetically linked HIV-1 infections by 96% in comparison to the delayed ART group. There were also a 41% reduction in HIV-1 symptoms in the early group compared to the delayed therapy group. The INSIGHT START study group trial which was better powered with 4685 patients followed for an average of three years showed a 72% and 39% relative decrease in AIDS related events and non-AIDS related events respectively in the immediate ART initiation group [36]. The results of these studies led to guidelines for prompt ART therapy without meeting a CD4+ cell count threshold.

The concept of pre-exposure prophylaxis (PrEP) was developed in parallel, with the aim of providing HIV negative individuals at high risk of acquisition with some of the same drugs that are included in ART therapy to prevent new HIV infections. Animal models showed that intermittent PrEP given 1, 3, or 7 days before exposure and 2 hours after exposure was highly effective in preventing rectal SHIV infection in macaques [37]. One of the first double blind randomized controlled trials in humans was the CAPRISA 004 trial with the aim of assessing the effectiveness of a topical gel for prevention of HIV transmission in South African women. The results found a 39% decrease in HIV incidence in the tenofovir compared to placebo group, while a greater decrease of 54% was found if adherence was high [38]. At around the same time, a multinational study of men and transgender women called the Preexposure Prophylaxis Initiative (iPrEx) trial tested the efficacy of a combination of antivirals taken orally once per day. The antivirals (oral emtricitabine (FTC) and tenofovir disoproxil fumarate (TDF) assigned group had a 44% decrease in HIV incidence compared to the placebo group [39]. Drug levels that were detected in the blood had a relative decrease in HIV risk of 92% compared to those who had undetectable drug levels in the active arm which was similar to the CAPRISA 004 trial [40]. The

Partners PrEP trial in Uganda and Kenya randomized heterosexual discordant couples to receive daily oral TDF/TC, TDF alone, or placebo [41]. This study found a relative reduction in HIV incidence of 67% and 75% in the TDF and TDF-FTC assigned groups, respectively. Six clinical trials have since showed PrEP efficacy ranging from 0-75% with the differences likely due to low adherence [42].

The results of these trials resulted in the first approval of TDF and FTC as single tablet fixed dose PrEP by the FDA for use in HIV defined high risk groups. PrEP strategies have diversified over time and can be administered as oral pills [43–45], drug eluting rings [46] or long acting injectables [47,48] which are designed to improve adherence. Use of vaginal rings containing 25mg dapivirine reduced HIV incidence by approximately 30% compared to placebo in women in two phase 3 clinical trials [49]. The HPTN083 phase II /III trial was stopped ahead of schedule after results showed CAB-LA to be three times more effective in reducing HIV infection incidence than oral pills in cisgender men and transgender women [50].

Another HIV infection prevention tool is post-Exposure prophylaxis (PEP). The purpose of PEP is to prevent HIV infection by stopping viral replication of the initial inoculation of the virus. The evidence for use of PEP is sparse and is based on animal studies and observational studies. When treatment was started 24 hours after intravenous virus injection, antivirals have been shown to protect against SIV infection in macaques with initiation delays reducing efficacy [51]. A case controlled study on healthcare workers receiving zidovudine (AZT) for PEP resulted in the odds of HIV infection being reduced by approximately 81%. It is difficult to conduct a randomized controlled trial to assess efficacy of zidovudine due to ethical challenges [52]. PEP is recommended after a potential high risk exposure to HIV. The onus is on the person to perceive the risk associated with a given exposure, in order to start therapy within 72 hours. These challenges limit the widespread adoption of PEP to prevent HIV infection.

While ART has been a great tool to combat HIV infection, an effective vaccine can lead to the elimination of HIV/AIDS. Although different approaches have been tried, design of an effective vaccine has so far been an insurmountable challenge. There has been great difficulty due to biological challenges associated with HIV, including sequence diversity and viral latency

establishment. The fact that it is extremely rare to be spontaneously cured means the vaccine has to elicit stronger immune response that are more robust than natural infection [53]. Different vaccine approaches have been tried to overcome previous failures and use the lessons to guide vaccine development. The first study to show modest efficacy was the RV144 trial from Thailand conducted in 2009. The vaccine consisted of a combination of live recombinant canarypox vector vaccine (ALVAC-HIV) and a recombinant glycoprotein 120 subunit vaccine (AIDSVAX B/E) and had an estimated 31% vaccine efficacy [54]. In the HVTN 702 trial conducted in South Africa, an analogous regimen as the ALVAC-HIV/AIDSVAX B/E for HIV-1 subtype C virus had disappointing results and stopped for meeting the prespecified criteria for nonefficacy [55]. One of the exciting developments has been the elicitation of broadly neutralizing monoclonal antibodies (bNAbs) as a form of passive vaccination. Their main advantage is that they have strong antiviral activity against most of the circulating virus strains from all clades and have been shown to be protective after SHIV challenges [56,57]. Trials are ongoing in humans to find out the amount of bNAbs needed to provide sufficient protection.

1.5 HIV pathogenesis

1.5.1 Overview of the human immune system

The human immune system is tasked with defending the body from harmful microorganisms found in the surrounding environment. The immune system has to respond to potential pathogens by being able to identify them and mount an effective response to neutralize them without causing harm to the host itself. The immune system can be broken down into two arms; the innate and adaptive arms.

1.5.2 Innate immunity

The innate system is the first line of defence for the body against potential pathogens in the environment. It is made up of different cells, enzymes and body parts that form interfaces between the body and the outside environment. The other important function of innate immunity is to differentiate between self and non-self (foreign) particles.

The first line of defence starts with the human skin which prevents entry of pathogens by forming a barrier made up of epithelial cells. Mucosal surfaces are coated with viscous mucus that limit entry of microbes and serve as a trap. Enzymes such as lysosome in tears and digestive enzymes also have antimicrobial properties to protect against microbes.

The other component of the innate system are the innate cells which are the primary agents that target and destroy invading pathogens. They get activated by using their pattern recognition receptors (PRR's) to recognize pathogen- associated molecular patterns (PAMP's) found in the invading pathogens. Some common PAMPs include bacterial lipopolysaccharides (LPS), bacterial flagellin, peptidoglycan and endotoxins found on the cell membranes of Gram-negative bacteria. Some viruses are also recognized as PAMPs by their nucleic acid structure such as double stranded RNA viruses. The four major families of PRRs are toll like receptors (TLR), nod like receptors (NLR), RIG-I-like receptors (RLR) and C-type like receptors (CLR). TLRs are transmembrane proteins found on the surface of innate cells, adaptive (B and T cells) immunity cells and non-immune (epithelial) cells. There have been more than ten different TLRs characterized to date that are non-redundant, bind intracellular and cell surface components of pathogens [58]. TLRs play a crucial role in cell activation by inducing cytokine production and type I interferons from immune cells and initiate inflammation.

The innate immune system is also composed of immune cells such as granulocytes, natural killer (NK) cells and antigen presenting cells (APC) such as dendritic cells (DCs). NK cells are effector lymphocytes that have both cytotoxic and cytokine producing effector function. NK cells possess receptors on their surface that allows them to distinguish between diseased cells and healthy self cells. NK cells have major histocompatibility complex (MHC) class I specific inhibitory receptors that bind MHC class I which ensure tolerance of most healthy cells while allowing toxicity towards stressed cells that loose the inhibitory receptors on them. NK cells get stimulated when their activation receptors bind their ligand on target cells that have an absence of MHC class I inhibition. NK cells can also be stimulated by IL-12, IL-18 and other soluble factors released by activated immune cells which can lead to cytokine production that can activate T cells, DCs and macrophages [59]. NK cells are crucial in antiviral defence by using their CD16 receptor to bind specific host Ig subclasses bound to viral antigens in infected cells. They

proceed with their cytotoxic activity by killing the cells using perforin, which forms pores on the cell surface and granzymes which activate cell death by apoptosis [60]. The other main innate response against viral infection is production of type I interferon by plasmolytic DC (pDCs). TLRs bind common viral features such as unmethylated CpG-rich DNA which activates pDCs. Type I interferons suppress viral replication by inducing genes involved in RNA degradation, arresting cell cycle progression and apoptosis of target cells [61]. Professional APCs capture, process, and present antigens to other immune cells such as T cells and include DCs, macrophages and B cells. Mature DCs are the most effective at priming naive T cells and controlling their functional response. APCs are localized in peripheral tissues and monitor their environment for pathogens.

1.5.3 Adaptive immunity

The adaptive immune arm produces highly targeted responses against specific antigens. It is composed of lymphocytes that include B cells and T cells. These lymphocytes have a unique ability to have differential expression of receptors on their surface after passing thorough several developmental stages. After a stringent selection process, lymphocytes express functional receptors which can recognize virtually any antigen while also being able to differentiate between self and non-self antigens. Lymphocytes can differentiate into memory cells that are sustained after an infection has been cleared and produce a quicker immune response the next time the antigen is encountered. The adaptive immune arm has two components; the humoral response and cell-mediated response.

The humoral immunity involves the production of antibodies secreted by B cells. The response starts when B cell receptors bind an antigen, and digest and process protein into peptides bound to MHC molecules that are identified by T helper cells. Cytokines release by T helper cells allows B cells to mature and differentiate into antibody producing plasma cells. The released antibodies can neutralize pathogens by binding them, flagging pathogens for phagocytosis by other immune cells. Antibodies can also activate complement proteins that can kill the pathogen directly.

Cell mediated immune responses do not involve antibodies but are instead mediated by T cells. The major T cell classes include CD8+ cytotoxic T cell (CTL) and CD4+ T helper (Th) cells. APCs activate T cells by providing three important signals; recognition of MHC-bound antigen by the T cell receptor (TCR), signaling through co-stimulatory receptors and cytokine secretion. CTLs are identified by their expression of the CD8 receptor and can directly lyse pathogen infected cells after activation. Th cells facilitate B cell activation and are crucial in meditating the cell mediated response. DCs are crucial for the differentiation and proliferation of T cells by presenting antigens to both CD4+ and CD8+ T cells. DCs direct T cell differentiation in a way that depends on the nature of the antigen and by the cytokines they produce. They direct naïve CD4+ T cells to differentiate into a mixed CD4+ T cell phenotype made up of Th1, Th2, Th17 and Treg. Each of these T helper cell subsets have different effector functions and express differential receptors on their surface. IL-12 secretion by DCs leads to Th1 differentiation, which are important for clearance of intracellular pathogens. IL-4 secretion by DC leads to Th2 differentiation which helps clear extracellular pathogens and parasites. Production of TGF-β, IL-23 and IL-6 by DCs after phagocytosis of bacteria-infected cells promotes Th17 differentiation [62,63]. Th17 cells are enriched in epithelial barrier sites giving them ideal location to respond to extracellular bacteria and fungal pathogens and have been implicated in autoimmune diseases during uncontrolled activation[64]. Treg differentiation can be induced by release of TGF-β by DCs and these cells play an important role in inhibiting the immune response. The balance of Th17 and Treg subsets maintains homeostasis in the gut microbiota. Segmented filamentous bacteria (SFB), a member of the gram positive bacteria has been shown to impact Th17 production in mice and mediating resistance to intestinal pathogens [65]. Th17 responses are also important against invading commensal microbes when the integrity of the gut barrier is compromised and microbial products are disseminated. It has also been observed that commensal bacteria such as those that produce butyrate have also been implicated in inducing Treg, contributing to mucosal tolerance [66]. The dynamics of the Th17/Treg axis thus have important implications for maintaining gut homeostasis and preventing disease in this important organ.

1.5.4 Host immune response against HIV

Most HIV transmission occurs during sexual exposure through the genital or rectal mucosae. Crossing the epithelial cell barrier is a low probability event but can be increased if the mucosa is physically impaired or prevalent sexually transmitted infections (STIs). During acute infection, there are various clinical phases during acute infection that can be distinguished by the presence of HIV antigen and HIV-specific antibodies. The eclipse phase refers to the time between HIV acquisition and detection of viral RNA in the plasma and approximately lasts for 10 days. The virus replicates locally in the mucosa before reaching the draining lymph node near the end of eclipse phase; in this tissue the virus encounters a higher concentration of activated CD4+ CCR5+ T cells and replicates rapidly while spreading throughout the body. The virus does particularly profound damage in the gut-associated lymphoid tissue (GALT), which contains most of the activated CD4+ CCR5+ T cells [67]. During acute HIV infection, there is a significant depletion of the intestinal CD4+ T cells (80%) and viral load increases exponentially, reaching peak viremia at 21-28 days after initial HIV infection [67,68]. While CD4+ T cell numbers decrease during peak viremia and recover almost fully in the blood, this is not observed in the GALT [67]. Individuals can become symptomatic during peak viremia, displaying a typical viral syndrome that can include fever, cough, fatigue and myalgia [69,70]. The 'window of opportunity' for a HIV vaccine to be effective in stopping viral replication is believed to be the time between exposure and peak viremia, prior to severe CD4+ T cell depletion and viral reservoir formation. Viral set point is established as the virus levels decrease from peak levels and stabilize after acute infection, typically 3 to 12 months following infection, though this varies between individuals and is determined by balance between virus turnover and immune response [71,72].

The early innate response to HIV begins with increase in acute phase proteins and pro-inflammatory cytokines that coincides with peak viral load during primary HIV infection. As viral load increases, levels of cytokines and chemokines such as IL-15, IL-18, TNF, IL-22 and IFNγ increase in the plasma in a response described as the cytokine storm [73]. This intense cytokine response reflects systemic activation of innate cells such as DCs, NK cells and macrophages that is intended to be protective against HIV, but can also mediate immunopathology. pDCs numbers decrease during acute HIV infection, in part due to activation

induced cell death. Activated DCs enhance antiviral innate response but also facilitate the spread of the virus to CD4+ T cells [74,75]. Immune activation due to production of pro-inflammatory cytokines and/or induced by innate cells can promote HIV replication by increasing HIV target cell numbers.

Early T cell responses during HIV infection is described by increased CD8+ T cell activation and potent cytotoxic activity. This effector function is evidenced by emergence of HIV specific CD8+ T cells correlated with decreased viral load, increase in viral load during SIV following CD8+ T cell depletion and viral escape mutations to avoid HIV specific CD8+ T cells [76,77]. This functional capacity is diminished during HIV infection progression through immune exhaustion and disruption of long-term memory CD8+ T cells formation [78,79]. Acute HIV infection severely harms the mucosal B cell generative microenvironments, resulting in widespread depletion of B cells and germinal centres [80]. This contributes to weak or short lived virus-specific antibody responses during the early period of infection. Despite the robust CD8+ T cell response, it is rarely sufficient to prevent HIV replication and disease progression occurs in a majority of cases (>95%) without ART. There is a short window of opportunity after transmission to eradicate viral infected cells before the virus is widely disseminated and forms long lasting latent reservoirs. The natural immune response is insufficient in breadth and too late to effectively control HIV replication and prevent immune dysregulation. Despite early therapy, the damage done to the immune system renders its efforts insufficient and persistent immune activation is evident during chronic HIV infection.

A key question in HIV infection relates to the cause of ongoing immune activation. While the virus is one obvious culprit, viral load is supressed during therapy yet increased activation persists. This activation is characterized by increased target cells, dysregulation of NK cell and pDC antiviral response, defective CD4+ T cell and CD8+ T cell function and risk of organ damage [61,79]. Immune activation during therapy in HIV-infected people is likely to be caused by several intertwined factors, including reactivation of the HIV reservoir, compromised gut barrier, increased levels of pro-inflammatory cytokines, destruction of mucosal surfaces and co-infections among other factors [81]. Persistent immune activation despite HIV therapy predicts mortality and is associated with comorbidities [82]. HIV infection has been linked with microbial

translocation (MT) and intestinal epithelial cell damage [83,84] that likely stems from major destruction to the gut and lymphoid memory CD4+ T cell compartment during acute HIV infection. MT occurs when microbial products cross through a leaky gut mucosa into systemic circulation. LPS may be derived from microbes translocated across the gut after destruction of HIV-infected CCR5+ T helper 17 cells [85,86]. The effects of this damage have been associated with inflammation as the body mounts a response to MT products that are in circulation. Binding of toll like receptors (TLR) on innate immune cells such as monocyte and DC induces these cells to release pro-inflammatory cytokines and lead to immune activation [84]. MT has been shown to predict disease progression in untreated HIV infected people [87].

1.6 Gut microbiome overview

The study of the gut microbiota was initially limited to culturable organisms [88]. Progress in molecular techniques have allowed for better understanding of the structure and function of the human gut microbiome. High-throughput sequencing of target genes, deep shotgun sequencing and other "omic" technologies have allowed study of complex microbial ecosystems in a culture independent fashion [89].

These methods have revealed that the gastrointestinal (GI) tract is home to trillions of microbes and the symbiotic relationship with the host is crucial to maintain gut immune homeostasis. Anaerobic bacteria make up the majority of the gut microbiota and dominate aerobic bacteria by 3-4 orders of magnitude. Although there are approximately 500-1000 different bacterial species that reside in the human gut intestine, most of this diverse array of species belong to a four dominant phyla: *Firmicutes, Bacteroidetes, Actinobacteria,* and *Proteobacteria* [90,91]. The composition of the gut microbiota also has diverse gradients along the tract. The microbial density and diversity increase from the proximal GI tract to the distal GI tract; the density increases from 10^1 to 10^{12} cells per gram from the stomach to the colon respectively. There is also variation in latitudinal composition in microbes present in the lumen and those adhering to the mucus layer [92]. Gut microbiota colonization begins immediately after birth and mode of delivery has been shown to influence the composition of the microbiota with reduced diversity for those delivered through caesarean section [93]. In early life, breast milk is the driving factor of

microbial structure and selection of gut bacteria [92]. With age, environment and lifestyle factors such as diet, urban versus rural setting, antibiotic use and other factors shape our gut microbiota structure [94]. It has also been shown that the gut microbiota of adolescents differs from adults, supporting the notion that changes continue to occur with age [95].

The gut microbiota is an important contributor to host physiology via a complex web of interactions. Gut bacteria help in digestion of insoluble fibers, in nutrient breakdown and production of essential vitamins [96]. Fermentative bacteria also produce short chain fatty acids (SCFA), which are crucial because they influence epithelial cell growth and facilitate metabolic effects [97]. The gut microbiota is also important for regulating immune homeostasis. Germ free animal models have shown deficits in lymphoid structures, reduced IgA-producing plasma cells, systemic T cell deficiencies and imbalances and irregular cytokine profiles in these animals [98–100]. Recolonization with specific commensal bacteria in germ free mice is adequate to restore some immune system functions [101,102].

Gut microbiota also protect the host by serving as a physical barrier to potential incoming pathogens, and prevent overgrowth of commensal bacteria. The host benefits from its gut microbiota through competitive exclusion, consumption of nutrients, and production of antimicrobial substances. Lactic acid produced by members of the *Lactobacillus* genus lowers pH and provides an inhibitory microbe growth environment. Lactic acid also permeabilizes the bacterial outer membrane of Gram negative bacteria which increases the susceptibility to host antimicrobials such as lysozyme for the pathogens [103].

1.6.1 Factors affecting gut microbiome

The gut is one of the most diverse body sites with regards to the bacterial communities it harbors. The gut microbiome varies widely due to plenty of factors such as age, diet, geography, disease, antibiotic use, genetics and other environmental exposures [104]. Diet has been shown to be important as at the phyla level. For example, children in Burkina Faso who have diets rich in fibre have a microbiome that is enriched with *Bacteroides* when compared to European children [105]. It has also been established at the genera level that plant and carbohydrate rich diets of people in LMIC tend to have a low *Bacteroides*:*Prevotella* ratio. In contrast, the typical Western diet, rich in animal products including high amounts of protein and saturated fat has been associated with a high *Bacteroides*:*Prevotella* ratio [94]. The environment we live in is also crucial in shaping the makeup of our microbiome. Households are more likely to have similar microbiomes as they might share similar dietary preferences, while those who move out from their homes will have their microbiome adjust to the new surroundings [106]. Antibiotic usage also can have dramatic effect on the gut microbiome, resulting in less diverse bacterial communities. Host genetics also contributes to shaping the gut microbiota as abundances of taxa are more highly correlated within monozygotic than dizygotic twins [107]. Although both host genetics and environment influence the gut microbiota composition, it's the environment exposure which has thus far been found to have a greater contribution in explaining microbiome variability compared to genetic ancestry [108].

1.6.2 Microbiome and health implications

Given the importance of the gut microbiota to immune modulation, it is not surprising that it has been associated to pathologies of the immune system. This makes it important that the immune response be tightly regulated to distinguish commensal from pathogenic bacteria. The nuclear factor (NF)-κB signalling which plays a pivotal role host inflammatory response to pathogenic bacteria is not activated by majority of the normal microflora [109,110].

There has been an increasing focus in understanding the relationship between the immune system and the gut microbiome in order to elucidate the crosstalk between the two components. Commensal microbes play a major role in maintaining intestinal homeostasis. This homeostasis

involves immune regulation to maintain a healthy gut while preventing an uncontrolled immune response again commensal microbes. A breakdown in this relationship as a result of alterations in the gut microbiota composition has been termed ‘dysbiosis’, which has been associated with various disease states including inflammatory bowel disease, obesity, diabetes, and cardiovascular diseases [111–114].

Inflammatory bowel disease (IBD) manifestations comprise of chronic inflammation along the digestive tract. Two types of IBD include ulcerative colitis, which affects the lining in the colon and rectum, and Crohn’s disease, which affects the deeper layers of walls of the digestive tract. Although the exact cause of IBD is unknown, some factors that play a role include genetics and defects in the immune response. Immune defects include aberrant responses to luminal antigens that lead to failure of tolerance to commensal flora [97]. The gut microbiome has been implicated in autoimmune arthritis, with studies showing colonization by a single commensal bacteria reinstated Th17 functionality, production of autoantibodies and development of arthritis [101].

Obesity has also been associated with dramatic changes in the structure of the gut microbiota. Lean mice had a 50% reduction in *Bacteroidetes* abundance and a corresponding rise in *Firmicutes* compared to obese mice. The obesity-associated microbiota was also linked to increase in genes involved in degradation of indigestible polysaccharides and elevated fat deposition using a microbiota transplantation model. There was low SCFA levels in the feces of the obese mice highlighting the increased ability to extract energy from diet in obese associated microbiota[111,115].

A link between microbial metabolism of phosphatidylcholine into the pro-atherosclerotic metabolite trimethylamine-N-oxide (TMAO) and cardiovascular disease has been reported [116]. Plasma TMAO levels were supressed after antibiotic treatment and in those who had a vegan diet after dietary phosphatidylcholine challenge. This has important health implications as plasma TMAO levels are associated with elevated risk of cardiovascular disease risk factors [113,117].

1.6.3 Gut microbiome alterations associated with HIV

HIV infection has also been associated with dysbiosis in a positive feedback loop. HIV infection leads to a severe depletion of intestinal CD4+ T cells, dramatically altering the structure and immunological property of this important organ. It is thought that HIV infection is accompanied by impairment of the gut barrier and MT, which contributes to persistent inflammation, further fueling HIV replication in a perpetual cycle.

Outcomes of interest in gut microbiome studies includes alpha diversity, beta diversity, and relative abundance of bacterial communities. Alpha diversity measures the amount of diversity and evenness of microbes within a sample. In adults, higher alpha diversity has been associated with good health outcomes, while lower alpha diversity has been associated to multiple gut diseases[118,119]. Beta diversity describes the variation in diversity of species between different environments. Individual changes in taxa can be analyzed across the classification rankings from phyla level to the species level, but a limitation in comparing studies is based on the resolution offered by the method used to generate the data.

There has been a proliferation of studies utilizing different molecular techniques, such as marker gene amplification and shotgun metagenomics, to find consistent patterns of association between the gut microbiota and HIV infection. The results of these studies have been varied, although one of the consistent patterns has been a decrease in alpha diversity in HIV infected patients. Interesting, this finding it is more consistent in heterosexuals compared to MSM, after stratification for sexual behaviour [120]. There are plenty of sources that contribute to the heterogeneity of HIV-microbiome outcomes, such as study designs, participants (untreated vs treated), small cohorts, MSM status, ART regimen, diet, sample type (stool and biopsy) and technical aspects of sample collection and analysis. The first studies that examined HIV infection and the gut microbiome environment found increases in the *Prevotella* genera in HIV positive patients compared to HIV uninfected persons. A metagenomic study on ART treated HIV infected people and healthy controls on stool samples found increased *Prevotella* and decreased *Bacteroides* in HIV infected individuals, in conjunction with a lower alpha diversity. They also found strong correlation of gut microbiota metabolism activity and T cell activation in HIV

infection linking gut microbiota composition to inflammation [121]. Several other studies that have utilized samples collected along various regions of the GI tract also found increased *Prevotella* and decreased alpha diversity [122–124].

Potential confounders to these studies were not fully appreciated until a large study showed convincingly in two independent cohorts in Europe that *Prevotella* dominance in the gut microbiota was linked with MSM than HIV status [125]. They also found MSM to have significantly richer diversity than non-MSM. They did not find HIV-specific microbiota changes but HIV progression was associated with lower bacterial diversity. Subsequent studies carried out in different cohorts similarly highlighted the importance of accounting for MSM status when studying gut microbiota composition[126,127]. Gut microbiota from MSM were much more pro-inflammatory when transferred to gnobiotic mice compared to those from non-MSM [128] . A large study also clearly showed that microbiota from HIV uninfected individuals clearly clustered by MSM status in terms of their beta diversity [127]. MSM status has thus been established to contribute to the diversity of the gut microbiota structure.

Although ART supresses viral production and reconstitutes immune cells, elevated chronic immune activation is observed in patients on long term ART. MT is thought to contribute to the immune activation but the influence of ART on gut microbiota structure is not well understood. One small study showed that the effect of HIV infection and therapy was minimal with regards to changes in microbiota composition in the MSM population; however, HIV infected men not on therapy had higher MT and CD8+ T cell activation compared to MSM on HIV therapy [126]. This study also found that fecal bacterial communities from MSM regardless of HIV status had higher activated monocytes response compared to HIV negative men who have sex with women (MSW) following stimulation. A study in Nigerian MSM also found subtle differences in microbiome composition and similar alpha diversity based on HIV status [129]. A longitudinal study of anal swabs carried out in the United States of America (USA) showed that HIV positive men had lower diversity and altered composition with enrichment in *Fusobacterium* and fewer *Firmicutes* [130]. It was shown that *Fusobacterium* was significantly elevated in those HIV infected people and was correlated with CD4+ T cell activation and poor CD4+ T cell recovery[131]. It seems that CD4+ T cell count is important to predict differences in gut microbiota composition

rather than HIV status alone. Low CD4+ nadir was associated with low microbial gene counts (LGC) in a dose-dependent fashion. LGC were elevated in Gram negative *Bacteroides* and enriched for bacterial virulence factors [132]. Another study also showed those with the lowest alpha diversity were those that had progressed HIV infection as measured by CD4+ T cell count [125]. In summary, these studies suggest that HIV infection might influence the rectal microbiome but that changes might not be visible if treatment is initiated early and viremia is controlled. Furthermore, it is critically important to carefully design studies that capture known gut microbiome modifiers such as medication, sexual activity, geography, diet and others that could have bigger effect than HIV itself.

1.6.4 Factors influencing gut microbiota in MSM

Although MSM status has been linked with the gut microbiota, it is still unclear what the potential drivers that shape the gut microbiota in this population are. A study of HIV negative MSM in the USA measured the impact of condomless receptive anal intercourse (CRAI) on the gut microbiome. Men who engaged in CRAI had higher levels of CD8+ activation and pro-inflammatory cytokines in their gut biopsies. The microbiota of those engaged in CRAI was also enriched in *Prevotellaceae* family members, while the controls were enriched in the *Bacteroidaceae* family. Alpha diversity was reported to be lower in those engaged in recent CRAI [133]. Studies from USA and Europe have found increased alpha diversity in MSM compared to non-MSM [125,127,134]. One of the consistent findings of these studies is that the MSM are enriched for *Prevotella* in comparison to non-MSM. A lot of these studies have found parallel decreases in the abundance of the *Bacteroides* genera [125,127,135]

Most of these MSM microbiome studies were generated in higher income countries and there is a scarcity of data from other locations. *Prevotella* and *Bacteroides* genera in particular have been identified as major clusters in the gut microbiome that vary based on different environmental exposures [94]. Studies in Africa suggest that the *Prevotella* genus is the most prevalent in different populations [136]. *Prevotella* was the most dominant in up to 70% of individuals in rural populations that engage in agro pastoralism in Tanzania and Botswana. In contrast, the dominant genus in the US is *Bacteroides,* which can be greater than 30% dominant in relative abundance

[136]. Another study showed distinct differences between urban and rural settings in children from Burkina Faso in comparison to children from Italy. Rural children had a mean abundance of 64% *Prevotella*, while the urban children had a mean of 10% and urban Italian children had less than 1% [137]. In light of these geographic influences on gut microbiome diversity, it is important to consider the role that MSM status has in this context. One Nigerian study in MSM found that *Prevotella* was the dominant genus at 40% in relative abundance, although they did not include a non-MSM group for comparison [129].There is a scarcity of studies from Africa, the region with the greatest HIV burden, that can help to answer whether identifying as MSM is associated with shaping the gut microbiome in the context of other influences.

1.6.5 Epidemiology of Kenyan MSM

MSM in sub-Saharan Africa face numerous challenges; at the forefront of these is the fact that homosexuality is still illegal in most African countries. MSM face a lot of stigma and hostility from society, contributing to many hardships that force men to keep their sexual orientation hidden in order to survive. Along with other key populations, MSM generally have a higher prevalence of HIV than the general population [4,138,139]. Until recently, MSM were a neglected population for HIV research in sub Saharan Africa. Subsequent research has highlighted the need for attention for prevention measures, treatment and care of this vulnerable population to stop the harmful effects of HIV borne by not only MSM but also by the rest of the society [140]. Early work in Kenya was done in Mombasa in an MSM cohort consisting mostly of sex workers. The HIV prevalence was 43% for men who exclusively had sex with men and 12.5% for men who had sex with both men and women [141]. A cohort of MSM sex workers in Nairobi, Kenya had a high HIV prevalence of 40%. The study also identified HIV risk factors to include history of urethral discharge, infrequent condom use during receptive anal sex and frequency of sex with male partners [142]. MSM sex workers are more likely to engage in risky sexual behaviour and have a higher chance to be infected with HIV or STIs. A study in Nairobi comparing male sex workers and other MSM found sex workers engaged in significantly more receptive anal sex and unprotected receptive anal intercourse than non sex workers. They also had a higher HIV prevalence 26.3% versus 12.2% compared to MSM who did not identify as sex workers [143]. Given this disease burden, it is crucial that studies are conducted to inform better HIV prevention

approaches in this population. This includes conducting gut microbiome research in this key population to determine their role in HIV acquisition and disease, and to replicate findings in more developed settings.

Study Rationale, Hypothesis, and Objectives

Study Rationale

Upon HIV infection, massive damage is incurred in the GALT that contributes to immune activation. HIV infection has been associated with gut microbiota alterations including elevated abundance of the *Prevotella* genera, lower alpha diversity and increased MT. ART initiation results in viral suppression and helps with CD4+ T cell reconstitution but immune responses remain impaired. Numerous factors can alter gut microbiota studies and cause variation between them. MSM status has been found to have a distinct gut microbiota profile independent of HIV status. These findings have also been strengthened by finding that MSM have a higher immune activation than MSW thus making them potentially more susceptible to HIV acquisition when exposed to the virus. It is therefore crucial to better understand the relationship between immune activation, MT and gut microbiome structure after HIV infection.

Potential drivers of gut microbiota structure in MSM have not been elucidated and there is also a scarcity of data from LMIC. One of the factors that could potentially affect the gut microbiota structure is different types of sexual behaviour. It is also not known how collecting gut microbiota samples from different sites affects their relationship with the immune system. There are variations in the bacterial communities that reside in the GI tract. There are adherent bacterial communities that reside in the mucosal layer while some reside in the lumen and they could have different interactions with the local immunity. Finally, it is of crucial importance to collect data from developing countries to assess consistency of results given the majority of studies are done in developed countries. Geographical and environmental exposure differences account for variation in gut microbiota difference and it is crucial to have that data in terms of HIV prevention efforts with targeted interventions.

Hypothesis

1. HIV infected men will have altered microbiome structure that results in reduced alpha diversity, but ART therapy will be associated with increased diversity. Men who engage in

different sexual behaviours will have, independent of HIV status, a distinct microbiome structure.

2. HIV infected men will have a higher level of MT than HIV negative men which will be associated with elevated systemic inflammation. MSM will have a higher background of MT compared to non-MSM populations.

3. Mucosal adherent samples will have less bacterial community diversity than fecal samples. Adherent bacteria communities in the mucosal samples will have stronger associations with local immunity than lumen bacterial communities.

Objectives

1. To determine if HIV infection and/or sexual behaviour are associated with an altered gut microbiota structure.
2. To determine the levels of MT in HIV infected men and links with immune activation. Assess differences in levels of MT in MSM and non-MSM populations.
3. To assess the difference in luminal vs mucosal sample with regards to gut microbiota structure and local and systemic inflammation.

2. Materials and Methods

2.1 General Reagents

2.1.1 Solutions

Phosphate-buffered saline (PBS)- Gibco-Invitrogen Thermo Scientific, USA

R-10 Cell Culture Media- RPMI-1640 (Hyclone) complemented with 10% FBS + 1% Penicillin/Streptomycin/Fungizone (PSF, Gibco-Invitrogen Thermo Scientific, USA)

Freezing Media- FBS + 10% dimethyl sulfide (DMSO, tissue culture grade; Sigma-Aldrich, St Louis-USA)

FACS Wash- PBS + 2% fetal bovine serum (FBS); (Gibco-Life Technologies-Thermo Scientific, Ontario, Canada). It's used to wash cells during surface staining.

BD Cytofix™ Fixation (BD Biosciences, California-USA). Used to fix cells for staining.

Perm/WashTM Buffer (BD Biosciences, California-USA). Used for permeabilizing cells. The stock solution is diluted to 1x using FACS Wash.

2.1.2 Stimulation reagents

Phorbol 12-myristate 13-acetate (PMA; Sigma-Aldrich, St Louis-USA)- Stimulate cytokine release by T cells.

Ionomycin (Io; Sigma-Aldrich, St Louis-USA)- calcium ionophore used to activate T cells

2.2 Biological Samples

2.2.1 Ethics

All study participants gave informed consent and Ethical approval was granted by the Kenyatta National Hospital Ethics Review Committee and the Institutional Review Board at the University of Manitoba.

2.2.2 Study participants

The study participants were enrolled from a well established MSM cohort in Nairobi, Kenya. The samples were collected from sexual workers outreach program (SWOP) and Health Options for Young Men on HIV/AIDS/STI (HOYMAS) clinics which are managed by a University of Manitoba/ University of Nairobi collaboration under the guidance of the ministry of health of Kenya. The clinics provide free services to participants that include provision of condoms, health education, counseling, STI screening and routine HIV rapid tests with same day results. The clinics offer treatment services on diagnosed infections and if necessary refer clients to appropriate service providers. The men are provided with HIV treatment (ART) and prevention services (PREP). During informed consent, participants completed a questionnaire that collected baseline information on demographics, sexual history and medical history. Physical and genital examination was performed before sample collection. Inclusion criteria included being male, ≥18 years of age and having a history of sex with men. Local donor samples without any gender bias, who are assumed to be relatively healthy and in low risk of infection from Winnipeg, Manitoba were used to optimize assays.

2.2.3 Demographic data collection

A standardized questionnaire was designed to collect demographics, medical history and sexual behaviour information. Data collected included age, marital status, education level, sex worker status, sex work period, practise of douching, number of partners in past month and medical

history. Sexual behavioural frequency in the past month with different partner types (regular, casual, paid) was also collected.

2.3 Sample collection and preparation

2.3.1 Blood collection and processing

For immunological testing, whole blood samples were taken in heparin vacutainers; for CD4+ cell count and HIV-1 viral load, they were taken in EDTA tubes. Through density gradient centrifugation, peripheral blood mononuclear cells (PBMC) were isolated from the blood samples. For 10 minutes, the whole blood sample was centrifuged at 550xg. Plasma from the top layer was taken out and kept at -80°C for storage. PBS + 2% FBS was used to dilute the remaining blood in the tube all the way to the top. In a 50ml conical tube, the tube diluted blood was gently combined with 10ml of Ficoll (Lymphoprep; MJS BioLynx Inc).With the brakes off, the tube was centrifuged at 500xg for 25 minutes. Without disrupting the red blood cells (RBC), the PBMC layer at the ficoll/plasma interface was extracted. The PBMC's were centrifuged at 500xg for 10 minutes after being rinsed with 40ml of R-10 cell culture medium. PBMC's were counted using the trypan blue dye exclusion method and washed for a second time at 500xg for 10 minutes. In a cryofreezing container, PBMC were resuspended in 2 ml of freezing media. Before being transferred for long-term storage into a liquid nitrogen tank, the cryofreezing container was incubated at a -80°C freezer for an entire night. For immunological tests, samples were sent to Winnipeg, Manitoba.

2.3.2 Rectal sample collection and processing

Rectal sponge samples were collected from participants by clinician assisted anoscopy. Two swabs were collected by inserting them 10cm inside rectum and turned twice. The swabs were placed in a cryovial containing 500 µl of RNA later. The patients conducted a self administered micro enema. The mucosa was allowed to reconstitute for half an hour before a two swabs were

collected by inserting 10cm inside the rectum, held in place for 2 minutes in which after sides were changed. The samples were transported to the lab for storage in a -80°C freezer. Individual sponges were taken out of their plastic stems and placed into 0.5 ml tubes that had already had sterile 18-gauge needles inserted into their distal ends. These individual tubes were then put into 2-ml tubes. Liquid was eluted by centrifuging for 5 minutes at maximum speed and both the pellet and supernatant were frozen in a −80° freezer until further processing. Supernatant was used for cytokine analysis and pellets were used for 16S rRNA sequencing of the gut microbiota were frozen in a −80° freezer until further processing.

2.4 General methods

2.4.1 16S rRNA sequencing

2.4.1.2 DNA extraction

DNA was extracted using QIAmp Fast DNA Stool Mini kit from Qiagen (Hilden, Germany). 180mg sample pellets from rectal sponges was weighed in a 2ml microcentrifuge tube and thawed on ice. 1ml of InhibitEX Buffer was added and the tube vortexed until the sample was thoroughly homogenized. The tube was centrifuged for 1 minute to form pellet. The supernatant was collected and added to 15 µl proteinase K. 200 µl Buffer AL was added to the tube and vortexed for 15s to create a homogenous solution. The tube was incubated at 70°C for 10 minutes, 600 µl of 95% ethanol was added to the lysate and vortexed briefly. The lysate was transferred to the QIAamp spin column and spanned for 1 minute. The filtrate was disposed and the QIAamp spin column was put into a new 2 ml collection tube. The QIAamp spion column was washed twice; first with 500 µl Buffer AW1 centrifuged for 1 minute and second with 500 µl Buffer AW2 centrifuged for 3 minutes with the filtrate discarded after each wash. The QIAamp spin column was put in a new collection tube and centrifuged for 3 minutes, moved into a new 1.5ml micro centrifuge tube. For DNA elution, 200 µl Buffer ATE was added. The tube was incubated for 1 minute at room temperature and centrifuged for 1 minute to elute DNA. All centrifugation steps was performed at room temperature at 20,000xg.

2.4.1.3 DNA quantification

DNA quantification was done by using qubit quantification. A working solution was made by mixing 199 µL QuBit buffer and 1 µL QuBit dye per sample/standard to be quantified. 199 µL QuBit buffer and 1 µL QuBit dye were added per sample to be quantified. 190 µL QuBit buffer and 10 µL QuBit dye were added per standard to be quantified. Tubes were vortexed for 5 seconds, were pulse spined and incubated at room temperature for 2 minutes. Tubes were inserted in the QuBit Fluorometer and readings were recorded. An aliquot of approximately 5 µL was removed for PCR.

2.4.1.4 PCR and Library preparation

The PCR workflow was done in a clean, template and amplicon free environment to avoid contamination. Primers were diluted in double distilled water to a 10 µM concentration. Samples were diluted to less than 125 ng/ µL prior to PCR using nuclease free water. A master PCR mix was prepared for each reaction by mixing 12.5µl of Invitrogen platinum II hot start master mix (2x) (Fisher Scientific), 0.4 µL of each 10 µM Reverse primer and 7.1 µL of nuclease free water. 20µL of the PCR master mix was added to each well (total of 96 wells). 1 µL of 5 µM of the barcoded Forward primer and 4 µL of each extracted DNA sample was added to reaction mixtures in each well separately. The reverse primer sequence ordered from Integrated DNA technologies read 5'GGACTACNVGGGTWTCTAAT 3'. The forward primer sequence ordered from Integrated DNA technologies read 5' GTGYCAGCMGCCGCGGTAA 3'. The PCR was done on a 96 well plate with each well having a final volume of 25 µl. Each extracted DNA sample is amplified in triplicate, 32 samples per 96-well plate. Sample DNA concentration was diluted by ultrapure water to <125 ng/microlitre for PCR library preparation. PCR was set up with the following settings: initial denaturation of 94 °C for 2 minutes; followed by 35 cycles of denaturation at 94 °C for 15 minutes, annealing at 50 °C for 15 minutes, and extension at 68 °C for 15 min; and a final extension at 68 °C for 5 min before a hold of 4 minutes.

2.4.1.5 Library quantification and cleaning

After PCR completion, each triplicate PCR reaction was combined into a single volume tube (60 µL per sample). Quality of PCR amplicons was assessed on a Screentape and quantified on qubit fluorometer. DNA primers and non-specific amplification products had to be cleaned up using SPRI based size selection (Beckman Coulter, Inc.). Amplicons that ranged in size between 150bp to 800bp were purified using double sized selection. A left side size selection process was followed by a right side size selection. The process involved adding a sample to SPRIselect of 0.85 and 0.56 for the left side and right side size collection respectively. Briefly, samples and SPRI beads were mixed on a tube, separated by magnetic stand, followed by ethanol wash and addition of elution buffer to extract the amplicons. Using quantifiable data from the tapestation and qubit, libraries were pooled in equimolar concentrations (5 ng of each library) into a single tube. The pooled library was quantified with qubit and tapestation for molarity determination. The libraries were diluted in 0.1X TE or ultrapure water to 4nM prior to loading to Miseq. PCR amplicons were assessed for size, quantity and purity by a tapestation. The ladder was prepared by adding 10 µL Genomic DNA Sample Buffer and 1 µL Genomic DNA. Samples were prepared by adding 10 µL Genomic DNA Sample Buffer and 1 µL DNA sample. The solutions were thoroughly mixed by a IKA MS3 vortexer at 2000 rpm for 1 minute. They were then spun down for 1 minute and centrifuged briefly before loading into the TapeStation Instrument for recordings.

2.4.1.6 Sequencing

The pooled library was diluted to a desired concentration of 8pM according to the sequencing machine protocol. Sequencing was carried out using the Illumina MiSeqTM platform, generating paired-end reads of nearly 300 base-length. A 20% PhiX spike- in was added to increase the complexity of the library. The forward sequencing primer read 5’

AGTCAGCCAGCCGGACTACNVGGGTWTCTAAT 3’, the reverse sequencing primer read 5’ TATGGTAATTGTGTGYCAGCMGCCGCGGTAA 3’ and the index primer read 5’

AATGATACGGCGACCACCGAGATCTACACGCT 3'. The Miseq was run and fastq files were obtained according to the manufactures instructions.

2.4.1.7 Data processing

The overlapping paired end FAST-Q files were processed in a data curation pipeline using QIIME version 1.0 [144]. Sample demultiplex and trimming is described in detail in Chapter 3.4. Chimeras were identified and removed using the consensus method. The naive Bayes classifier UCLUST was used for taxonomic classification, trained on the 2013 99% identity Greengenes database. Sequencing depth was visualized in rarefaction graphs and three samples having less than 26, 000 reads were removed.

2.4.2 Flow Cytometry

2.4.2.1 Surface Staining

Frozen PBMC vials were put in a hot bath of 37°C until nearly completely thawed before being resuspended in 10ml of R-10 cell culture medium. For DMSO removal, the sample was spanned at 500xg for eight minutes. Prior to surface labelling, cell viability and count was done by trypan dye exclusion. One million cells were transferred to 5 ml FAC tubes and centrifuged for 10 min at 500xg. The resuspended cells were then transferred to 96 well V bottom plates for staining after the supernatant was removed. The cells were washed in 150 µL of FACS wash and centrifuged for 10 minutes at 500xg. Cells were resuspended in 50 µL of surface staining antibody master mix after the supernatant was removed. The cells were incubated at 4ºC in the dark for 30 minutes. The cells were centrifuged at 500xg for 10 minutes for two washes with 150 µL of FACS wash. The supernatant was removed, cells were resuspended, fixed in 300 µL of cytofix buffer and moved to 5ml FACS tubes for acquisition.

2.4.2.2 Stimulation and Intracellular cytokine staining

Cells were resuspended and stimulated with phorbol myristate acetate (PMA) and ionomycin at a concentration of 100 ng/ml and 1 µg/ml respectively. Fetal calf serum (10%) and penicillin/streptomycin (5000 U/ml /5000 µg/ml) were added to the medium as supplements. A Golgi-block cocktail of brefeldin-A (1 mg/ml, Sigma-Aldrich, Seelze, Germany) and monensin (1 mg/ml, BioLegend, London, UK) was added. The samples were incubated at 37°C with 5% CO_2 for 4 hours. Cells were washed with FACS wash and transferred to 96 V -shaped plate for surface staining with a mixture of fluorochrome conjugated antibodies. Samples were incubated for a total of twenty minutes at room temperature in the dark and then washed with FACS wash. Samples were permibilized with 200 µl cytofix/cytoperm, with 0.25% cytofix/cytoperm, incubated at (4°C) for twenty minutes and washed with 1x perm Wash. Cells were centrifuged at 500xg for 10 minutes and subsequently stained with intracellular cytokine antibody and incubated at (4°C) for thirty minutes, washed twice with 1x perm Wash, resuspended in 1x PBS and transferred to FAC tubes for acquisition.

2.4.2.3 Compensation

Compensation is necessary in flow cytometry when multiple fluorochromes are used in a panel. Spectral overlap occurs when there are multiple fluorochromes that have similar energy emission characteristics. Compensation is the process of reducing signal in the channel of interest from other fluorochromes. Compensation was carried out during each flow cytometry run for consistency. Compensation for the surface and intracellular antibody fluorochromes was carried out using single stained positive BD CompBeads (anti rat or anti mouse antibodies). A drop of each positive and negative control beads was mixed with 400 µL of FACS wash for each batch of four single colour bead controls. The mixture was aliquoted in 100µL aliquots to four FACS tubes, 1 µL of each single stained antibody from the flow cytometry panel was added and incubated for fifteen minutes at room temperature. The tube was diluted to 300 µL with FACS wash before acquisition. Compensation for the live dead stain followed a similar protocol with some modifications and use of ArC (amine reactive compensation) beads. One drop of beads is

added to 1 µL of fixable dead cell stain and incubated for 30 minutes at room temperature. The beads were washed twice after incubation with 4ml FACS wash and diluted in 300 µL FACS wash. One drop of negative control bead was applied to the tube before flow cytometer analysis.

2.4.2.4 Data acquisition

The LSRII flow cytometer (BD Bioscience) at the lowest preset low flow rate was used for data acquisition. Fluorescence Minus One (FMO) tubes were used to modify the gates in order to define positive and negative gates. Each fluorochrome is given its own FMO tube, which contains all of the fluorochromes in the panel with the exception of one. After gathering 100,000 lymphocyte events, the acquisition was terminated. BD FACSDiva software (v6.1.3) was used for data acquisition and analysis. The data was exported to the FlowJo software for additional processing (v10.8.1; Tree Star Inc.).

2.5 HIV testing

HIV serology was confirmed via the fourth generation GS HIV Combo Ag/Ab EIA (Bio-Rad, Hercules, California, USA) and Avioq HIV-1 Microelisa System (Avioq Inc., Rockville, Maryland, USA) at the JC Wilt lab in Winnipeg. The GS HIV Combo Ag/Ab EIA test was performed on the Automated Microplate System (Bio-Rad Laboratories). The 96 well plate was coated with HIV p24 antigen and antibodies to HIV Type 1 and Type 2. Frozen plasma samples were initially thawed in an ice bath. 25µL of Anti-HIV biotinylated polyclonal antibodies was added to the plate before loading of 25µL of control/samples. The mixture was incubated for 1 hour at 37ºC and washed to remove excess sample. 100µL of streptavidin peroxidase was loaded into each well, incubated for 30 minutes at room temperature and unbound streptavidin peroxidase was removed by washing. The sample was then washed with 80µL of Working TMB Solution and incubated for 30 minutes at room temperature. 100 µL of the Stop Solution was added and the plate was read at the dual wavelength 450/615-630nm on a microelisa plate reader.

The Avioq HIV-1 Microelisa System is also an enzyme-linked immunosorbent assay (ELISA) for the detection of antibodies to HIV-1. The plate is coated with HIV-1 antigens and samples/controls are added. Following incubation, the plate is washed and anti-human enzyme conjugate is added to the plate. Following a wash, the plate is incubated with ABTS (2,2'-azino-di-(3-ethylbenzthiazoline-6-sulfonate)) substrate and the solution turns green. The reaction is inhibited by the addition of a fluoride solution. The amount of antibodies to HIV was quantified on a microelisa plate reader.

2.6 HIV viral load analysis

HIV viral load testing using the RealTime HIV-1 viral load assay (Abbott, Chicago, Illinois, USA) on an automated m2000 RealTime system (Promega, Madison, Wisconsin, USA). The assay combines automated RNA extraction, real time PCR amplification of the integrase conserved region of pol gene fragment and fluorescence detection. The m2000sp automated extractor was used to get purified RNA from the plasma sample while the automated m2000rt instrument was used for qRT-PCR amplification and fluorescence detection. An internal control is used as a positive control to make sure the reaction is proceeding properly. The sample input is 0.6ml of plasma and the detection range is from 40 to 10,000,000 copies/ml.

2.7 STI testing

Genomic DNA was extracted using QIAmp Fast DNA Stool Mini kit from Qiagen (Hilden, Germany) from preenema rectal swabs pellets according to the manufacture's instructions. STI testing was done using the Seegene Anyplex 6-plex kit. The kit was used to test for Chlamydia trachomatis, Neisseria gonorrhoea, Mycoplasma genitalium, Mycoplasma hominis and Ureaplasma urealyticum. PCR amplification was performed following the manufacturer's instructions. An internal control was present in the PCR mix for making sure the PCR went well and acted as a positive control. Sterile deionized water was used as the negative control. Amplified DNA products were detected and quantified using the tapestation.

2.8 Elisa assay

2.8.1 Biorad assay

A multiplex immunoassay (Bio-source 37-plex Human Cytokine Assay; Invitrogen) was used according to the manufacturer's instructions except where noted to measure levels of APRIL / TNFSF13,BAFF / TNFSF13B,sCD30 / TNFRSF8,sCD163,Chitinase-3-like 1,gp130 / sIL-6Rβ,IFN-α2,IFN-β,IFN-γ,IL-2,sIL-6Rα,IL-8,IL-10,IL-11,IL-12 (p40)IL-12 (p70),IL-19,IL-20,IL-22,IL-26,IL-27 (p28),IL-28A / IFN-λ2IL-29 / IFN-λ1,IL-32,IL-34,IL-35,LIGHT / TNFSF14,MMP-1,MMP-2,MMP-3,Osteocalcin,Osteopontin,Pentraxin-3,sTNF-R1,sTNF-R2,TSLP,TWEAK / TNFSF12. Briefly, post enema rectal sponge supernatant samples were incubated with antibody-coupled beads, washed to remove unbound protein, biotinylated detection antibody was added and subsequently streptavidin-phycoerythrin was added after a series of washes between each addition. Standard curves ran in duplicate, samples were ran as singlets and analyte concentrations were calculated using Bioplex Manager Software. The assays showed good reproducibility with the average calculated concentration %CV of 8.13 % for the standards. Most of the samples fell in the detection range of the assays and 30/37 analytes had greater than 60% detectability.

2.8.2 MSD assay

Chemokines and inflammatory cytokines were measured using two panels; the proinflammatory panel (IFN-γ, IL-1β, IL-2, IL-4, IL-6, IL-8, IL-10, IL-12p70, IL-13, TNF-α) and the chemokine kit (Eotaxin, Eotaxin-3, IL-8, IL-8 (HA), IP-10, MCP-1, MCP-4, MDC, MIP-1α, MIP-1β, TARC) respectively. The assay was used on post enema rectal sponge supernatant and the protocols for both kits were similar. Briefly, the 96 well plate was washed 3 times with 150 µL of PBS+ 5% tween, addition of 50 µL calibrator, addition of 25 µL of samples and standards and incubation of plate for 2 hours with shaking at room temperature. Plate was washed 3 times with PBS+ 5% tween, 25 µL of detection antibody solution was added and incubated for two hours with shaking at room temperature. The plate wash washed 3 times with PBS+ 5% tween, 150 µL

of read buffer T was added and the plate was read on the SI2400 Imager. The assays showed good reproducibility with the average calculated concentration %CV of 8.2% for the standards. Most of the samples fell below the detection range of the assays and 8/20 analytes had greater than 60% detectability.

2.8.3 Microbial translocation assays

Microbial translocation was assessed by measuring LPS-binding protein (LBP), IFABP (intestinal fatty acid binding protein), sCD14 using elisa assays. IFABP was measure using the Human FABP2/I-FABP Quantikine ELISA Kit (R&D Systems), LBP was measured using human LBP DuoSet ELISA Kit (R&D Systems) and sCD14 was measured using Human CD14 Quantikine ELISA Kit (R&D Systems). The assays were done on plasma samples according to the manufacturer's instructions and samples were diluted accordingly. The protocols were similar for both IFABP and sCD14; briefly assay diluent was added to the 96 well plate, sample and standards were loaded and plate incubated for two hours at room temperature. The plate was washed, conjugate antibody was added, incubated and washed again. The plate was incubated for 30 minutes at room temperature with substrate solution. Stop solution was added to stop the reaction before readings were taken on microplate reader set to 450nm. The LBP assay was different in that plate was prepared overnight with addition of capture antibody to coat the 96 well plate and addition of reagent diluent for blocking purpose. Samples and standards were loaded the next day, plate was washed, detection antibody was added and plate washed. Streptavidin-horseradish peroxidase (HRP) was added, incubated for 20 minutes at room temperature and plate was washed. The plate was incubated for 30 minutes at room temperature with substrate solution. Stop solution was added to stop the reaction before readings were taken on microplate reader set to 450nm.

2.8.4 Elisa assay conditions and analysis

The steps that were not in protocol was when we didn't dilute our post enema rectal sponge supernatant samples because pilot data had showed that most of the analytes were already at the bottom of the standard curve without dilution. We also thawed our samples overnight for a period of 12-14 hours in a cold room that was at 4^0C rather than thawing them out at the start of assay as suggested in the protocol. This was done to minimize degradation of the analytes by allowing them to thaw at a slower pace. To act as extra control for intra-variability, we ran two plasma samples, a local donor plasma and control vial from the manufacturer in duplicate. The local donors were subdivided into low risk samples and high risk samples for the MT kits. The low risk samples were obtained from Department of Medical Microbiology and Infectious diseases staff and the high risk samples were obtained from patients with IBD visiting Health Science Hospital in Winnipeg, Canada. The median coefficient of variation (CV%) for all the wells that were ran in duplicate was 8.13 with the interquartile range(IQR) being (6.5, 11.5). This was within the recommended range by the kit manufacturer of less than 15 CV%. The standard curve generated in the assays had an upper limit of quantification (ULOQ) and lower limit of quantification (LLOQ) which were defined by the standard recovery within a range of (80-120) % as recommended by the kit manufacturer. Concentrations less than the LLOQ were assigned half the concentration of the LLOQ and analytes that were greater than 60% detectable were treated as continuous variables

2.9 Statistics

Microbiome data analysis was performed in R and is described in detail in chapter 3.4. Multiple factor analysis (MFA) was used to overlay immune phenotype data to microbiome defined clusters with a priori labelling of variables. The "FactoMineR" package was used for the Principal component analysis (PCA) while the graphing was done using the "factoextra" package [145,146]. Data was mean centered and unit scaled prior to PCA. The Non-parametric Wilcoxon Rank Sum test was used to compare analyte and cell surface marker differences among two different groups. A two way ANOVA was used to compare differences in analytes between three

different groups and Tukey's test was conducted for pairwise comparisons. Spearman's test was used to measure correlations between variables and visualized on a heatmap generated from Metaboanalyst version 4.0 [147]. A P-value of less than 0.05 was considered significant.

3 Impact of HIV infection on gut microbiome

3.1 Rationale

The gut microbiome plays an important role in maintaining intestinal immune homeostasis, regulating mucosal immunity and intestinal barrier integrity [148]. HIV infection disrupts this barrier, leading to alterations in the composition of the microbiome, microbial translocation and inflammation [84,149]. Although ART reduces non-AIDS mortality, especially with early treatment, chronic HIV-associated inflammation is only partially reversed by viral suppression [150]. Therefore, understanding factors that influence rectal microbiome composition and associated inflammation remains of intense interest. In particular, how rectal microbiota might affect HIV susceptibility has only recently been considered more widely in higher risk individuals [151].

The most reproducible association between HIV and rectal microbiota is a reduction in alpha or absolute bacterial diversity [152]. However, confounding variables such as age, geography and diet may obscure differences attributable to HIV alone [94,105]. Although some studies suggest a shift from *Bacteroides* to *Prevotella* spp.-dominated communities following HIV infection [124], others argue this is confounded by sexual orientation (oversampling HIV-infected MSM compared with controls). This was supported by a large study of European men that found a high *Prevotella*-to-*Bacteroides* ratio was strongly associated with being MSM, independent of HIV infection [125]. In studies that controlled for being MSM, relatively subtle microbiome changes were observed in HIV infected individuals with higher $CD4^+$ T cell counts [153]. In contrast, low $CD4^+$ T cell count nadir during HIV infection was associated with decreased species richness and functional changes in the gut microbiome [132]. Men who recently practised condomless receptive anal intercourse (CRAI) were enriched for *Prevotellaceae* spp., and lower alpha diversity compared with individuals who never engaged in anal intercourse [133]. This study highlighted the possibility that sexual activity itself could lead to rectal microbiome changes, but this has not been investigated in follow-up studies. If anal sex itself influences the rectal microbiota and increases risk of HIV acquisition, then rectal microbiota changes may precede HIV infection. Rectal sexually transmitted infections (STIs) have also been associated with microbiome structure; this could be due to preexisting inflammation associated with sexual behaviour, STI-induced or

related to other factors [154]. Resolving how sexual activity and infections influence rectal microbiota remains an important question for the health of this key population [154].

Microbiome composition varies based on geography [94], necessitating replication of microbiome study findings in multiple contexts. One prominent example of population-specific differences comes from studies showing that *Prevotella* spp. are dominant genera in healthy individuals from non-Western countries, suggesting that this genus might not carry the same inflammatory effects that have been attributed to HIV and/or being MSM [155]. Furthermore, few microbiome studies have been conducted in sub-Saharan Africa, in particular in MSM. With that in mind, we used 16S rRNA sequencing to profile the composition of the rectal microbiome in Nairobi MSM at a high risk for HIV acquisition (~12 incident HIV cases per 100 person-years) [142]. Although this study does not compare MSM with non-MSM, it does compare microbiome variables in multivariable models that incorporate both the frequency of sexual behaviour with different types of clients in both HIV-infected and uninfected men. The findings of this study have important implications for elucidating rectal HIV transmission mechanisms among MSM in sub-Saharan Africa, a region with very high HIV incidence.

3.2 Hypothesis

HIV infected men will have altered microbiome structure that results in reduced alpha diversity, but ART therapy will be associated with increased diversity. Men who engage in different sexual behaviours will have, independent of HIV status, a distinct microbiome structure.

3.3 Objective

To determine if HIV infection and/or sexual behaviour are associated with an altered the gut microbiota structure.

3.4 Section specific methods

3.4.1 Study design

We enrolled MSM from a programme that provides HIV/STI care and treatment in Nairobi, Kenya. Inclusion criteria were being biologically male, at least 18 years of age and reporting sex with men. Participants were stratified into HIV-negative, HIV positive on ART ('treated') and HIV-positive ART-naive ('untreated') groups. All participants gave informed consent and ethical

approvals were obtained from the Kenyatta National Hospital Ethics Review Committee and the Institutional Review Board at the University of Manitoba.

3.4.2 Data collection

Clinical and demographic data were captured by a validated questionnaire. Key variables included age, education, HIV and ART status, type and frequency of sexual behaviour with different types of partners, condom use, number of partners, anal douching, participation in sex work and recent medications.

3.4.3 Specimen collection

Rectal mucosal sampling included two clinician-assisted rectal swabs followed by a self-administered enema using a commercial solution. After approximately 25–30 minutes, participants underwent clinician-assisted anoscopy to collect additional mucosal samples. Detailed sampling collection and storage conditions can be found in chapter 2.3.2

3.4.4 16S rRNA targeted metagenomic sequencing of rectal microbiota

Postenema rectal sponges were thawed and centrifuged, with pellets used for microbiome profiling. DNA was extracted using QIAmp Fast DNA Stool Mini kit from Qiagen (Hilden, Germany). Amplification of two variable regions (V3-V4) from the 16S rRNA gene (~460 bp) was carried out using primer pairs described in the Illumina MiSeq rRNA Amplicon Sequencing protocol. PCR reactions followed manufacturer's instructions. Sequencing was carried out using the Illumina MiSeqTM platform, generating paired-end reads of nearly 300 base-length. Individual samples were indexed using dual barcoding and sequences were de-multiplexed by binning barcodes in QIIME [144]. Sequences were trimmed with Trimmommatic using a cut-off value of Q30 for both ends [156]. A minimum mean threshold of Q20 for 30 base pairs-sliding window across sequences and a minimum read length of 250 base pairs was established. The naive Bayes classifier UCLUST was used for these analyses.

3.4.5 Sexually transmitted infection testing and HIV viral load

HIV-1 infection was assessed via rapid screening in the clinic, as per Kenyan Ministry of Health guidelines. HIV serology was confirmation via GS HIV Combo Ag/Ab EIA (Bio-Rad, Hercules, California, USA) and Avioq HIV-1 Microelisa System (Avioq Inc., Rockville, Maryland, USA) at the JC Wilt lab in Winnipeg. HIV viral load testing using the RealTime HIV-1 viral load assay (Abbott, Chicago, Illinois, USA) on an automated m2000 RealTime system (Promega, Madison, Wisconsin, USA). Infection with *Chlamydia trachomatis*, *Neisseria gonorrhoea*, *Mycoplasma genitalium, Mycoplasma hominis* and *Ureaplasma urealyticum* was tested using the Seegene Anyplex 6-plex kit from preenema rectal swabs. Detailed protocol can be found in chapter 2.5, 2.6 and 2.7.

3.4.6 Statistical analysis

Alpha diversity was calculated using the *vegan* package in R, including species richness [Observed, Chao1 and abundance-based coverage estimator (ACE)] and richness/evenness indices (Shannon and Simpson). Operational taxonomical units (OTUs) observed in only one sample and/or genera with less than 25% prevalence were filtered out to reduce the number of taxa for analyses. For each sample, a random subset of 26000 counts was selected for library size normalization prior to calculating alpha diversity. Beta diversity was visualized on nonmetric multidimensional scaling (NMDS) based on ecological distance matrices calculated by Bray-Curtis dissimilarities. Permutational analysis of variance (PERMANOVA) using Bray-Curtis distances was performed to assess the impact of covariates on the microbiome composition using the *vegan* package [157]. The partition around medoids (PAM) method was used to identify clusters using Bray-Curtis ecological distance between samples using the *cluster* package. Silhouette coefficients were used to assess the randomness of clustering and the optimal number of clusters using the *factoextra* package [145]. Genus abundance was visualized by consecutive bar plots with samples ranked according to NMDS1 coordinates using the *ggplot2* package. Differential abundance testing was performed using the *selbal* package [158], a forward selection method that accounts for compositionality. This method finds the best of ratio of genera that distinguish two

groups, with a 10-fold cross-validation performed to assess outcome validity. Multivariable linear regression was used to model receptive anal intercourse and HIV infection as predictors of alpha diversity using the *mixed* procedure in SPSS version 25, adjusting for various covariates including insertive anal intercourse. Composite behavioural variables were computed combining report of sex with one or more paid, casual or regular male partners in the past week. Outcome variables included two measures of alpha diversity (Chao1 and Simpson). Age and all variables associated with alpha diversity at *P* value less than 0.2 and/or HIV grouping were included in multivariable models, with beta coefficients, 95% confidence intervals and *P* values reported.

3.5 Results

3.5.1 Study participants

The study enrolled 80 participants; 46 were HIV negative, 10 were HIV-positive untreated and 24 were HIV-positive treated. Questionnaire data were available for 70 participants (Table 1). HIV-negative men were younger than both HIV groups (median 23, IQR: 21–29; versus median 28, IQR: 25–32 for HIV-positive treated; and median 27, IQR: 23–29 for HIV-positive untreated, respectively; $P=0.036$). HIV-negative men had lower rectal STI prevalence than HIV-positive untreated and HIV-positive treated groups (23.5 versus 46.7 and 60.0%, respectively, $P=0.062$). The study groups were similar with respect to participation in sex work, sexual positioning, condom use, douching and number of regular, casual or paid male sex partners. Insertive and receptive anal sex with all partner types was frequently reported, with receptive anal sex more common in HIV-positive men. $CD4^+$ T-cell counts were lower in both HIV-positive groups than the HIV-negative group ($P=0.0001$). A relatively small proportion (16%) of HIV-positive men had $CD4^+$ T-cell count less than 300 cells/μl, indicating that most participants were not at an advanced stage of HIV infection. Although most HIV-positive treated participants (13/18) had a plasma HIV viral load below the limit of detection, HIV-positive untreated men all had detectable viral load (median 11 334 copies/ml, IQR 4334–42 354).

Table 1. Clinical and demographic characteristics of the study population

Variable	HIV-negative (n=39)	HIV-positive treated (n=21)	HIV-positive untreated (n=10)
Age, median (IQR)	23 (21–29)*	28 (25–32)*	27 (23.3–29.8)*
Marital status, single	33 (84.6)	17 (81.0)	5 (50)
Education level			
Primary	5 (12.8)	3 (14.3)	1 (10)
Secondary	16 (41)	9 (42.9)	5 (50)
Tertiary	18 (46.2)	9 (42.9)	4 (40)
Age at first sex, median (IQR)	16 (13–18)	14 (12–16)	15.5 (14.3–17.8)
Works as a sex worker	20/39 (51.3)	14/21 (66.7)	6/10 (60)
Regular partner (≥1) past 1 month	24/34 (70.6)	14/18 (77.8)	4/8 (50)
Casual partner (≥1) past 1 month	23/33 (69.7)	13/18 (72.2)	7/9 (77.8)
Paying partner (≥1) past 1 month	18/34 (52.9)	11/17 (64.7)	6/7 (85.7)
Oral sex (sometimes or always)			
Regular[a]	13/39 (33.3)	10/21 (47.6)	4/9 (44.4)
Casual	5/30 (16.7)	6/14 (42.9)	2/8 (25)
Paid	5/23 (22.7)	7/13 (53.8)	2/6 (33.3)
Insertive sex (always)			
Regular	21/39 (53.8)	6/20 (30.0)	5/9 (55.6)
Casual	15/30 (50.0)	4/14 (28.6)	4/8 (50)
Paid	10/22 (45.5)	4/13 (28.6)	4/8 (50)
Receptive sex (sometimes or always)			

Variable	HIV-negative (n=39)	HIV-positive treated (n=21)	HIV-positive untreated (n=10)
Regular	18/39 (46.2)	15/20 (75.0)	6/9 (66.7)
Casual	11/29 (37.9)	11/14 (78.6)	5/8 (62.5)
Paid	12/22 (54.5)	11/13 (84.6)	4/6 (66.7)
Always uses condoms			
Regular	28/39 (71.8)	16/20 (80.0)	7/9 (77.8)
Casual	22/29 (75.9)	11/12 (91.7)	7/8 (87.5)
Paid	17/23 (77.3)	10/12 (83.3)	3/6 (50)
Practices rectal douching			
Soap and water	13/25 (52)	9/13 (69.2)	2/7 (28.5)
Water only	10/25 (40)	3/13 (23)	4/7 (57)
Prep use, ever	7/39 (17.9)	0	0
STI positive[b]	4/34 (11.7)	4/15 (26.7)	2/10 (20)

Differences between subgroups were assessed by chi-square analysis for categorical variables and Mann–Whitney *U* test for continuous variables.

a All sexual behavioural variables may differ in their denominator, depending on whether participants report sex with that type of sex partner.

b STI positive is a positive test result for *Chlamydia trachomatis* or *Neisseria gonorrhoea.*

$*P<0.05$.

3.5.2 HIV infection and microbiome composition

The unsupervised PAM algorithm and silhouette quality index showed moderate support for two clusters of *Bacteroides*-spp or *Prevotella*-spp dominant microbiota (silhouette index of 0.4). Although there was no significant clustering based on study group in two-dimensional NMDS, the *Prevotella*-spp cluster was less frequent in HIV-positive treated compared with HIV-negative

and HIV-positive untreated men, respectively (38.1 versus 59.0 and 60.0%, $P=0.27$, Fig. 1a). We did not find any association between alpha diversity and $CD4^+$ cell count, although very few participants had low $CD4^+$ cell counts. The ability of HIV/ART status to explain the variation in microbiome composition was not significant (PERMANOVA $R^2=0.033$, $P=0.13$, Fig. 1b). Across all alpha diversity metrics, HIV-positive treated men had lower diversity than HIV-negative men: these were statistically significant for Observed ($P=0.01$) and ACE ($P=0.013$, Fig. 1c), which measure overall richness. There were no statistical differences between the HIV-positive untreated and HIV-negative groups.

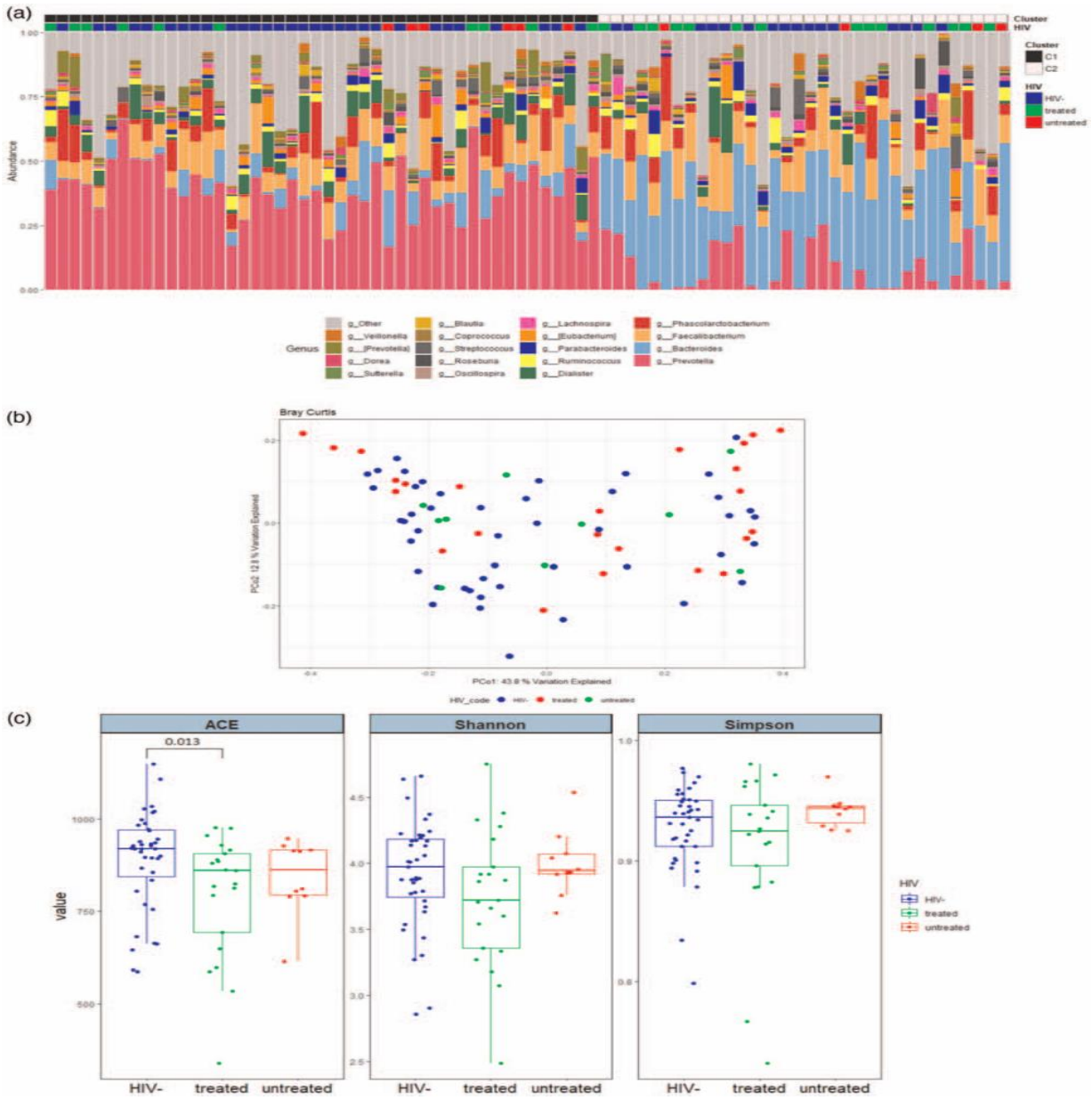
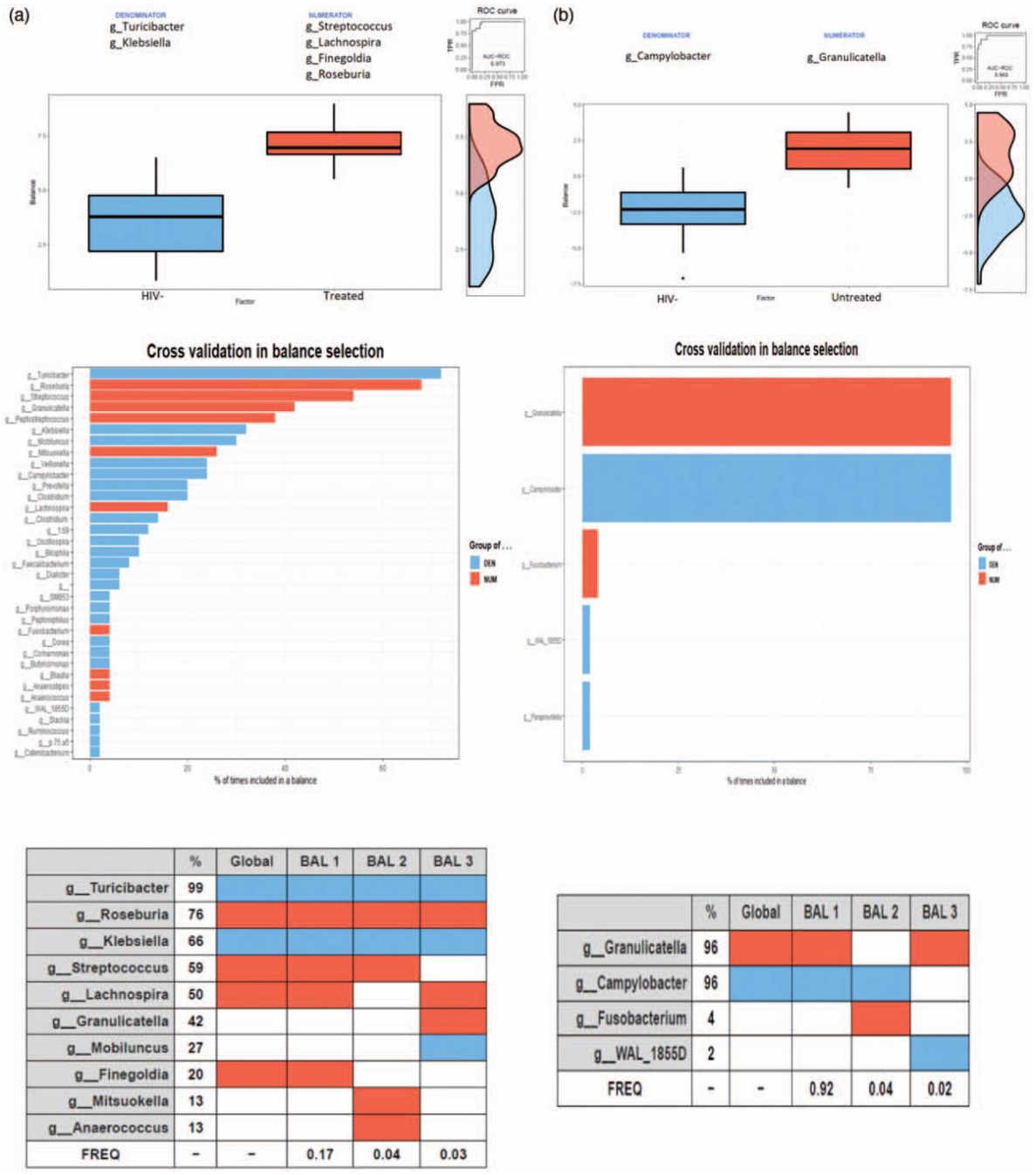


Figure 1. Associations between HIV-defined groups and microbiome structure and diversity. (a) Genus-level taxonomic summary plot. Genera with mean relative less than 0.1% abundance are binned together into the 'other' category. HIV status and cluster classification are indicated for each sample (top rows). Each column corresponds to an individual, with the relative abundance of each genera colour-coded and indicated below. Samples are ordered by the NMDS1coordinate on the Bray Curtis similarity matrix. (b) Plot of NMDS1 versus NMDS2 on Bray Curtis similarity matrix with each individual (circle) coloured by HIV study group. (c) Alpha diversity boxplots by study group. As per part B), HIV groups are coloured as HIV-negative (blue), HIV-positive treated (red) and HIV-positive untreated (green). Statistical details are listed in the text.

3.5.3 Differential abundance of genera by HIV infection

We next determined the ratios of rectal microbiota genera that best distinguished study groups. Compared with HIV-negative men, *Roseburia, Lachnospira* and *Streptococcus* genera were enriched in HIV-positive treated men, while *Turicibacter* and *Klebsiella* were enriched in HIV-negative men. Each of these genera were detected in more than 50% of the 10-fold cross-validation tests (Fig. 2a). When compared with HIV-negative men, *Granulicatella* was enriched, while *Campylobacter* was depleted in the HIV-positive untreated men; again, both appeared in more than 50% of 10-fold cross-validations (Fig. 2b). More differences were noted when comparing HIV-positive ART and HIV-negative groups, in line with the reduction in alpha diversity observed earlier for this group (Fig. 1c). Most differentiating genera were prevalent in greater than 50% of participants (Supplemental data, Table 3). We also performed the analysis on rare genera with prevalence of 5% or greater and found the results were mostly similar with the exception of elevated *Sneathia* in the HIV-treated men in comparison to the HIV-negative men appearing as a hit as shown in (Supplemental data, Figure 17)



	%	Global	BAL 1	BAL 2	BAL 3
g__Turicibacter	99				
g__Roseburia	76				
g__Klebsiella	66				
g__Streptococcus	59				
g__Lachnospira	50				
g__Granulicatella	42				
g__Mobiluncus	27				
g__Finegoldia	20				
g__Mitsuokella	13				
g__Anaerococcus	13				
FREQ	-	-	0.17	0.04	0.03

	%	Global	BAL 1	BAL 2	BAL 3
g__Granulicatella	96				
g__Campylobacter	96				
g__Fusobacterium	4				
g__WAL_1855D	2				
FREQ	-	-	0.92	0.04	0.02

Figure 2. Specific genera associated with HIV status using forward selection analysis generated in the *selbal* R package. (a) Box plots showing the balance value distribution (boxplots) in the HIV-negative (blue) and HIV-positive treated groups (red). (b) Box plots showing the balance value distribution (boxplots) in the HIV-negative and HIV-positive untreated groups (red). The genera contributing to the balance are listed in rank order of importance (bottom panels), with genera over-represented in HIV-negative (in blue) and HIV-positive (in red), sorted by frequency of appearance in multiple cross-validations.

3.5.4 Associations of sexual behaviour and microbiome diversity

It has been hypothesized that MSM-associated rectal microbiota may influence HIV susceptibility at the site of virus exposure. We stratified our sample by HIV status to investigate behavioural covariates of microbiome diversity that might impact HIV susceptibility (Fig. 3). Reporting receptive anal sex one or more times in the previous week with a paying partner was associated with lower alpha diversity (P=0.003). Similar lower alpha diversity trends were observed in men who reported receptive anal sex with regular and casual partners, although these were not significant (chi-square P=0.12 versus P=0.41, respectively). Reporting to not 'always' use condoms with regular partners was also associated with lower alpha diversity (P=0.05), with similar trends observed with other types of partners (P=0.21 for both casual and paid). The opposite trend was observed for insertive anal sex, whereby those who always practiced this type of sex had higher diversity. There was no significant difference in alpha diversity based on STI status. We did not find any significant association between the behaviour reported and the microbiome clusters we classified as *Bacteroides*-spp. Cluster and *Prevotella*-spp (Supplemental data,Table 4)

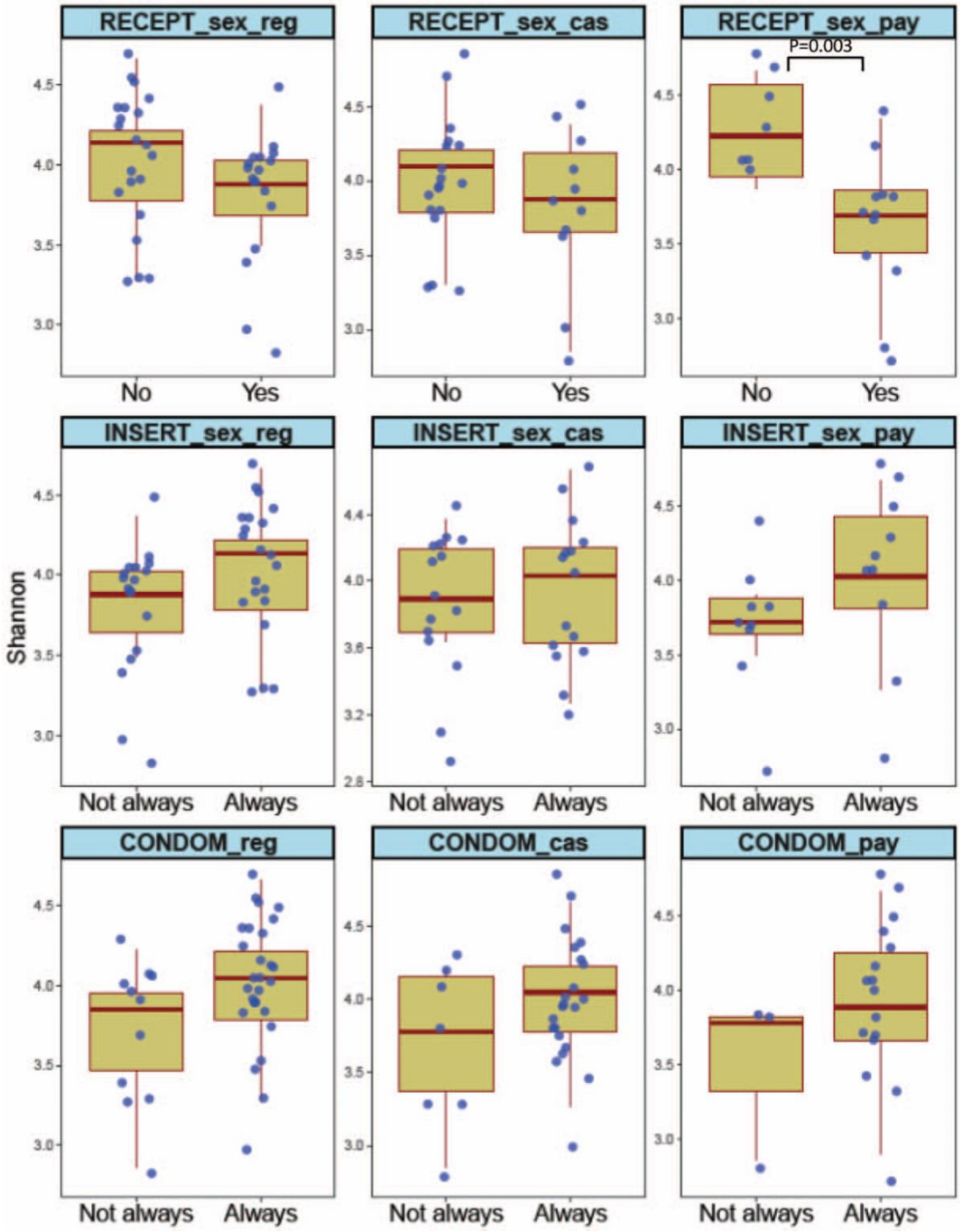


Figure 3. Report of different types of sexual activity with regular, casual and paid partners (denoted by 'reg', 'cas', 'pay'), and their association with alpha diversity in HIV-negative men. Responses were either coded as 'ever' (Y/N) or 'always/not always'. Boxes represent IQR, with lines denoting the median. Comparisons were made using the Mann–Whitney *U* test (only $P<0.05$ are indicated).

Overlap between types of sex was common, with most participants engaging in both insertive and receptive anal intercourse with multiple partner types, making it difficult to identify independent effects of a given sexual behaviour. To address this, we carried out multivariable linear regression models with alpha diversity defined by Simpson's index (capturing richness and evenness) as an outcome. Men who had receptive anal sex with casual, regular and paid partners in the previous week had lower alpha diversity than men who had receptive anal intercourse with 0, 1 or 2 of those partner types (P=0.017, P=0.089 and P=0.003, respectively, Table 2). Similar results were obtained for receptive anal sex in an adjusted model using Chao1 diversity as the outcome, which measures overall richness. Insertive anal sex was not associated with alpha diversity in either model. Interestingly, and in line with Fig. 1c, HIV was associated with alpha diversity in the Chao1 (P=0.03) but not the Simpson (P=0.49) alpha diversity model.

Table 2. Linear regression analysis of clinical and behavioural variables associated with rectal microbiome alpha diversity

	Simpson		**Chao1**	
Variable	**P**	**Adjusted[a] beta coefficient (95% CI)**	**P**	**Adjusted[a] beta coefficient (95% CI)**
HIV-positive	0.49	0.008 (−0.015 to 0.032)	0.03	−96.7 (−183.6 to −9.9)
Receptive anal sex with all three partner types[b]	Ref	ref	Ref	Ref
Receptive anal sex with all two of three partner types[b]	0.003	0.058 (0.02–0.096)	0.039	161.1 (8.7–313.5)
Receptive anal sex with all one of three partner types[b]	0.089	0.030 (−0.005 to 0.065)	0.134	94.4 (−30.1 to 218.8)
No receptive anal sex last week	0.017	0.044 (0.008–0.080)	0.092	117.6 (−20.1 to 255.4)

a Models included HIV (pos/neg), receptive anal sex (four categories), insertive anal sex (four categories), age, education level, marital status and STI (pos/neg).

b Combined variable capturing report of receptive anal sex with one or more regular, casual and/or paid partners in the past week.

3.6 Discussion

The gut microbiome has been an important area of interest in HIV, given implications for its role in pathogenesis and adverse health outcomes [159]. However, almost all studies to date have been carried out in Western cohorts, and in particular, studies of HIV in African MSM have only recently emerged, with biological studies still lacking [160]. It is critical that microbiome findings are replicated in genetically and geographically diverse populations to understand their generalizability. We addressed this in a population of Kenyan MSM at a high risk of HIV infection [142]. In agreement with a recent meta-analysis [152], alpha diversity was lower in HIV-positive than in HIV-negative MSM for indices that focused on richness, but not richness and evenness. These findings imply that HIV infection may decrease the number of taxa present in the mucosa, but not the frequency within those taxa. Alpha diversity was also associated with frequency of receptive anal sex and lower condom use, including in models adjusted for HIV. In contrast, neither HIV nor sexual activity was associated with microbial cluster, despite similar proportions of *Prevotella* and *Bacteroides*-dominant communities in this population as compared to other studies. Although the underlying explanation for HIV-alpha diversity association remains unclear, it is notable that lower alpha diversity in the gut carries negative consequences for numerous health conditions. Interestingly, our data confirm that receptive anal intercourse, a major HIV risk factor in MSM [161], may have a larger effect on rectal microbiota than HIV itself. These results underscore the importance of carrying out longitudinal studies to discern whether decreased alpha diversity precedes HIV infection and increases susceptibility [162].

Our results are partially in line with a previous study of HIV-negative MSM who recently engaged in CRAI [133]. Although we find similar associations with receptive anal sex and alpha diversity, no associations with *Prevotella* cluster were observed in our study. Nevertheless, data from both studies support that receptive anal sex itself, rather than another MSM-associated variable, drives microbiome changes in MSM that are consistent across different geographic contexts. Importantly, our results suggest that different partner types such as paid partners, which may be associated with differences in sexual intercourse, may drive this effect.

HIV associations in our study are similar to those in western nations that have accounted for being MSM [124–126]. Although we did not include non-MSM in our study, our results are generally

in agreement with other studies conducted in non-MSM populations in urban African settings [136,137,153]. The genus *Prevotella* was the most dominant genera in our setting with an average mean abundance of 22%, which is within the range of what is observed in other studies. Alpha diversity was particularly lower in HIV-positive treated men, and lower in immunological nonresponders than those whose $CD4^+$ cell counts increase on ART [126,127]. $CD4^+$ cell count nadir is a strong predictor of low alpha diversity and microbiome changes in MSM [132]. A Ugandan study also found alpha diversity was reduced in individuals with $CD4^+$ cell count less than 200 cells/μl [153]. A Nigerian MSM study, in contrast, did not find a reduction in alpha diversity, although a reduction was evident within the *Bacteroidetes* phylum [129]. The most likely explanation is that decreased alpha diversity is driven by duration of HIV infection, and HIV-positive untreated men in our study had more recently acquired HIV, as suggested by their relatively high $CD4^+$ cell counts. The HIV-positive treated group may also not experience microbiome restoration despite $CD4^+$ cell count gains, particularly if they initiated ART at a low $CD4^+$ cell count nadir, as has been suggested by other studies [123]. Although we did not find any associations between alpha diversity and $CD4^+$ cell count, the HIV-positive untreated and treated groups had similar $CD4^+$ cell counts, likely representing lack of decline in the former and partial recovery in the latter. Longitudinal studies that include $CD4^+$ cell count nadir will be critical to further unravel $CD4^+$-microbiome associations.

Compositional changes in rectal microbiota related to HIV were evident beyond the *Prevotella*/*Bacteroides* clusters. HIV-positive treated men were enriched for *Streptococcus*, *Roseburia* and *Lachnospira* and depleted of *Turicibacter* and *Klebsiella*, in comparison to HIV-negative men. These genera are Gram-positive Firmicutes, except *Klebsiella*, which is a Proteobacteria. *Lachnospira* and *Roseburia* are thought to be nonpathogenic, with *Roseburia* potentially beneficial due to butyrate production, which is the fermentable product of insoluble dietary fibre. Butyrate promotes anti-inflammatory effects by inducing colonic Treg differentiation, serves as an energy source and is depleted in inflammatory bowel disease and HIV [112,163–165]. *Klebsiella* is a proinflammatory pathogen and its overgrowth has been implicated in gut dysbiosis in colitis, HIV and hypertension [166–168]. *Streptococcus* dominates the oral microbiome, but is found throughout the body and is elevated in IBD and liver cirrhosis [169,170]. Our data agree with a previous MSM study in Nigeria, which observed a shift in HIV-positive treated men towards pathogenic bacteria not observed in

the untreated group [129]. Although similar, the genera that distinguished our HIV-negative and HIV-positive treated groups differed; these differences could be due to technical and/or population differences.

Our study had some limitations. We did not enrol a non-MSM control group, as it has been established in previous studies that MSM have a distinctive microbiome structure; instead, we focused on how heterogeneity in sexual behaviour within MSM might affect microbiome attributes. We did not collect data on diet and BMI, which could be confounders. That said, all men were recruited from the same clinic, with a similar age, education level and sexual behaviour, suggesting our study is reasonably representative of Nairobi MSM. Sequencing of V3-V4 regions of 16s rRNA provides taxonomic identification to the genus level and precludes species-level resolution. Follow-up studies using shot-gun sequencing are needed to better understand how bacterial gene expression is associated with HIV and sexual behaviour. Our study was cross-sectional and had a relatively small sample size; larger, longitudinal studies are needed to better understand the temporality of relationships, including whether rectal microbiomes predict the risk of HIV acquisition, including whether rectal microbiota predict the risk of HIV acquisition. Longitudinal studies can also evaluate whether very early ART prevents HIV-associated rectal microbiome changes that have been observed in HIV-positive individuals who started ART at very low $CD4^+$ cell counts.

In summary, we demonstrated the feasibility of carrying out rectal microbiome studies in Kenyan MSM, and replicated several findings related to HIV and being MSM in a non-western context. In addition to decreased alpha diversity in men with HIV infection, we also established a dose-dependent relationship between receptive anal sex with multiple partner types, controlling for HIV, and reduced rectal microbiome alpha diversity. These findings suggest that MSM-associated rectal microbiomes are likely linked directly to specific types of sexual activity, rather than confounding variables. These data provide an important perspective on the microbiome in a unique population of men at high risk of HIV acquisition, and could be included in formulation of strategies used to augment preexposure prophylaxis (PrEP) to reduce HIV transmission in this population.

4. Microbial translocation in HIV infected MSM

4.1 Rationale

MT is defined by the passage of viable gut microflora from the GI tract, after breaching the intestinal epithelial barrier, into normally sterile sites such as mesenteric lymph nodes, liver and the bloodstream [171]. The major factors that promote MT are bacterial overgrowth, impaired host immune defenses and physical damage to intestinal mucosa [171]. The gut barrier is composed of a thin epithelial layer that separates our body from the gut microbiota. The gut barrier functions selectively permitting nutrient absorption, hindering entrance of bacteria, training of immune cells and waste secretion. It has antimicrobial peptides, IgA and thick musical layer to serve its barrier function. Under normal conditions, the translocated products are phagocytosed within the lamina propria by epithelial cells acting as non-specific phagocytes [89,172]. However, bacteria can survive and egress when the host immune system is weakened. Alterations in composition of the commensal bacteria and immune system deficiencies can lead to breaches and potential disease state.

HIV infection depletion of CD4+ T cells in the gut compromises gut barrier integrity and impairs mucosal immune function [84]. Immune activation is a robust predictor of HIV disease progression with higher frequency of CD38 expression on CD8+ T cells predicting faster decline of CD4+ T cells [173,174]. The first landmark study showed elevated levels of LPS in the bloodstream of HIV infected patients and SIV-infected animals [84]. LPS was used as an indicator of MT and was correlated with immune activation [84]. LPS is part of the cell wall of Gram-negative bacteria and was found to be elevated in HIV infected individuals and linked to impaired CD4 gut homing and CD4+ T cell activation [175]. However, because assays to measure LPS are not ideal, many groups have measured LBP instead as an indirect measure of LPS. LBP produced by intestinal epithelial cells binds LPS and presents the molecule to CD14 [176]. MT can also be quantified indirectly using IFABP and sCD14. sCD14 is a co-receptor for LPS that is released from activated monocytes and the soluble form that appears after shedding is a strong predictor of mortality during HIV treatment. sCD14 correlates with other MT markers [177]. IFABP is a marker of enterocyte damage and was also shown to be elevated in HIV infected patients with lower nadir CD4+ cell count and correlates positively with sCD14 [173,174,178]. Although markers of MT

has been found to be elevated in HIV infected individuals in a majority of studies, there have been some studies from sub Saharan Africa that did not replicate those findings and raise questions on how environment exposure affects markers of MT [179,180].

Given the relationship between markers of MT and immune activation, it is possible that differences in gut microbiome structure between groups could be associated with markers of MT. It has been shown that HIV-associated gut microbiome dysbiosis is linked to inflammation, but is not known if the types of microbiome community in different populations might influence this association[159]. MSM have distinctive gut microbiome structure that can be characterized as pro-inflammatory [133,151]. Previous reports have shown that trauma or physical perturbations that disturb the gut mucosa can cause MT and this has been shown in patients after undergoing surgery [181,182]. It was also shown that MSM who engage in recent high risk exposure had elevated endotoxemia that was associated with increased level of inflammation [183]. It is possible that the pro inflammatory gut environment in MSM could be due to sexual behaviour preference and any disruption of the mucosa lining could translate into MT. We aimed to investigate this relationship by characterizing the gut microbiome and linking it to markers of MT and systemic immune activation from plasma samples in our MSM cohort. IBD patients were also included as positive controls, as IBD is associated with elevated MT [184]. We also compared markers of MT in HIV infected patients from Kenyan MSM to see if we can replicate the findings from the majority of studies conducted in western countries. The focus of this chapter is to understand whether alterations in the gut microbiota structure also correspond to systemic immune activation and markers of MT.

4.2 Hypotheses

IBD patients and HIV infected patients will have a higher amount of MT when compared to healthy controls. MSM who engage in receptive anal sex will also have elevated levels compared to MSM who don't practise receptive anal sex. Gut microbiome changes between groups will be linked with markers of MT and systemic markers of inflammation.

4.3 Objective

To compare the expression of markers of MT in HIV patients, IBD patients and healthy controls, and to compare markers of MT levels in MSM by their sexual behaviour preference stratified by their HIV status. To correlate markers of MT to systemic T cell activation markers, rectal inflammatory cytokines, and gut microbiome communities.

4.4 Chapter specific methods

Rectal analytes were run on post-enema rectal sponge supernatant using the biorad multiplex assay as described in chapter 2. Flow cytometry was run on PBMC from three panels used for gating T cells (Supplemental data, Table 5). Representative plot for each population of interest is shown in Supplemental data, Figure 18,19,20)

4.4 Results

4.4.1 Study participants and experimental conditions

We enrolled three study groups; MSM, IBD patients, and healthy controls. MSM participants were further subdivided by HIV status; 46 were HIV negative, 10 were HIV-positive untreated and 24 were HIV-positive treated. Demographics and clinical data were available on 70 samples and summarized in previous chapter (Chapter 3, Table 1). Receptive anal sex with at least one partner was practised by (18/30) MSM who tested HIV negative. CD4+ T-cell counts were lower in HIV patients while viral load was mostly undetectable (13/18) in those that initiated ART. The IBD patients were selected based on the severity of their symptoms and ten samples were chosen that had active IBD. The samples were taken before initiation of additional IBD treatment. Healthy controls included local donor samples from students and staff from the healthy donor program at the Department of Medical Microbiology and Infectious Diseases, University of Manitoba and HIV negative Kenyan women enrolled in a community clinic in Nairobi, Kenya. The healthy controls are not known to be diagnosed with a chronic disease, had never engaged in receptive anal intercourse nor were they presenting any disease symptoms during sample collection.

Plasma samples were assayed primarily as singlets, with a subset of samples in duplicates to calculate correlation of variation (CV). The intraplate CV for the all group runs for sCD14, IFABP and LBP plate was 7.31%, 7.87% and 10.42%, respectively. This is within the manufacturer recommended range of less than 15%.

4.4.2 Microbial translocation markers expressed at higher concentrations in HIV infected and uninfected MSM.

We first compared markers of MT on the basis of HIV status. When compared to HIV negative, HIV positive samples had elevated levels of sCD14 (median 2,191, IQR: 1,687–2,614 pg/ml; versus median 1,745, IQR: 1,491-2,364 pg/ml for HIV-negative men, respectively; p=0.013). A comparison of HIV patients showed no difference in sCD14 levels between HIV negative samples and HIV infected treated and untreated patients. We found no difference in LBP levels in HIV negative and HIV positive samples (median 19.7, IQR: 18.6-22.7 ng/ml; versus median 20.8, IQR: 19.7-23.4 ng/ml for HIV negative men, respectively; p= 0.24) as visualized in **Figure 4**. There were similar IFABP levels in HIV negative and HIV positive samples (median 1,199, IQR: 945–1825 pg/ml; versus median 1315, IQR: 757-1790 pg/ml for HIV-negative men, respectively; p=0.53).

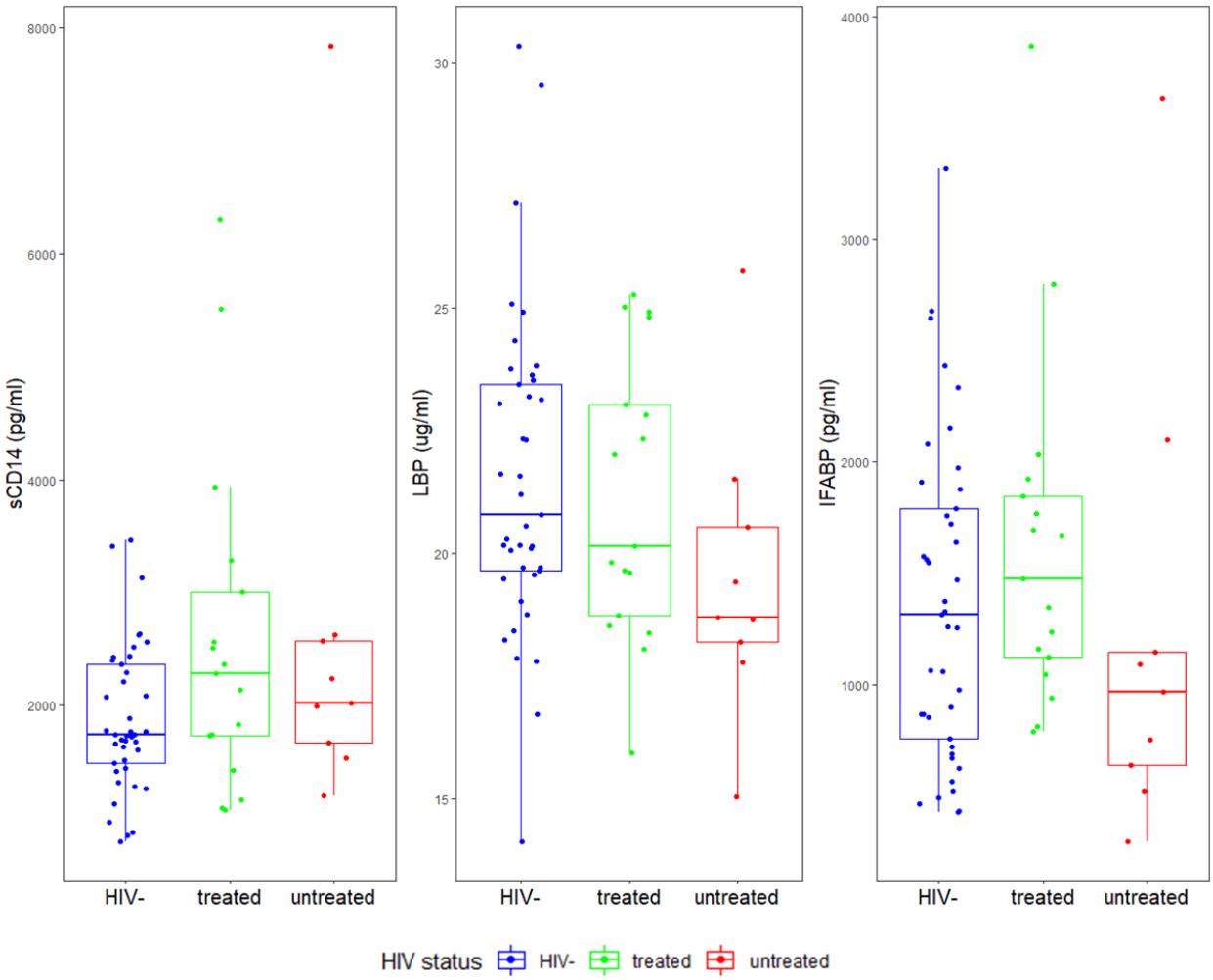


Figure 4. Microbial translocation in HIV infected MSM. Color of boxplot indicate HIV status; red indicates HIV uninfected men, blue indicates HIV infected men on ART and green indicate ART naïve HIV infected men. Boxes represent IQR, with lines denoting the median. Comparisons were done using a Kruskalis-Wallis test including turkey's post-hoc pairwise analyses.

4.4.3 Microbial translocation markers elevated in MSM who engage in receptive anal sex

Markers of MT were evaluated in men stratified by recent anal receptive sex with regular, casual or paid partners. In HIV negative men, a majority (18/30) engaged in receptive anal sex at least with one partner type over the past month. Men who practised receptive anal sex had elevated levels of LBP (median 22.0, IQR: 19.8-24.2 ng/ml; versus median 19.9, IQR: 18.4-20.4 ng/ml for men who didn't practise receptive anal sex, respectively; p=0.036, **Figure 5)**. IFABP levels

were elevated in men who practised receptive anal sex, although this difference was not significant (median 1312, IQR: 694-1880 pg/ml; versus median 1062, IQR: 796-1573 pg/ml for men who don't practise receptive anal sex, respectively; p=0.52). Similarly, sCD14 levels were higher in men practising receptive anal sex, but non-significantly (median 1,694, IQR: 1,485-2,011 pg/ml; versus median 1505, IQR: 1277-1800 pg/ml for men who don't practise receptive anal sex, respectively; p=0.42).

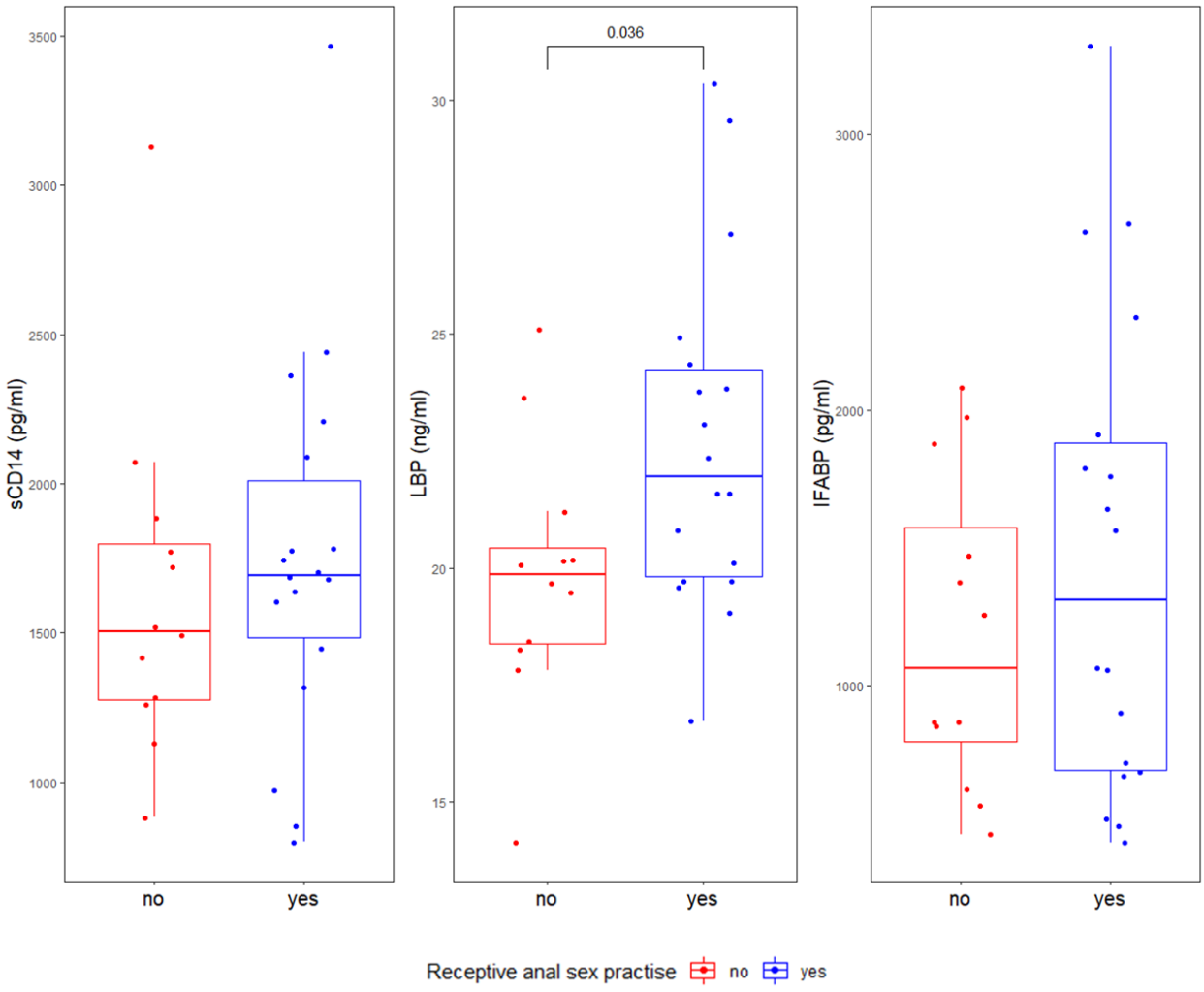


Figure 5. Elevated microbial translocation in HIV negative MSM engaging in receptive anal sex. Color of boxplot indicate risk group. Boxes represent IQR, with lines denoting the median. Comparisons were done using a Mann-Whitney U test. P- values less than 0.05 are shown on top of each analyte.

4.4.4 Microbial translocation markers were elevated in MSM compared to other low risk groups

We compared markers of MT levels in MSM who engage in receptive anal sex against different population samples to assess whether environmental differences also impact markers of MT. We included IBD patient samples who have high risk of markers of MT and two low risk groups which included donors with no known infection from Winnipeg and Nairobi to account for different environment exposure. The MSM population who had tested negative for HIV infection were stratified by receptive anal sex practices. When compared to healthy donors and IBD patients, MSM had elevated levels of IFABP (median 413, IQR: 356-525 pg/ml; versus median 188, IQR: 128-297 pg/ml; p=0.0031 for healthy donor samples; versus median 226, IQR: 181-328 pg/ml; p=0.027 for IBD patients respectively, **Figure 6).** Further analysis showed men who practised receptive anal sex had significantly higher IFABP levels than the low risk and IBD sample groups and MSM who did not practise receptive anal sex. We found no difference in LBP levels between MSM samples, low risk and high risk samples with a median and IQR of 12.64 (11.4-14.3) ng/ml; versus median 11.88, IQR: 10.6-13.5 ng/ml; versus median 14.92, IQR: 11.95-17.36 ng/ml for MSM, low risk and high risk samples respectively (**Figure 6)**. There was a trend in elevated LBP for IBD patients in comparison to MSM and low risk samples (p=0.09). MSM who practised receptive anal sex also trended towards elevated LBP in comparison to MSM who do not practise receptive anal sex (p=0.18).

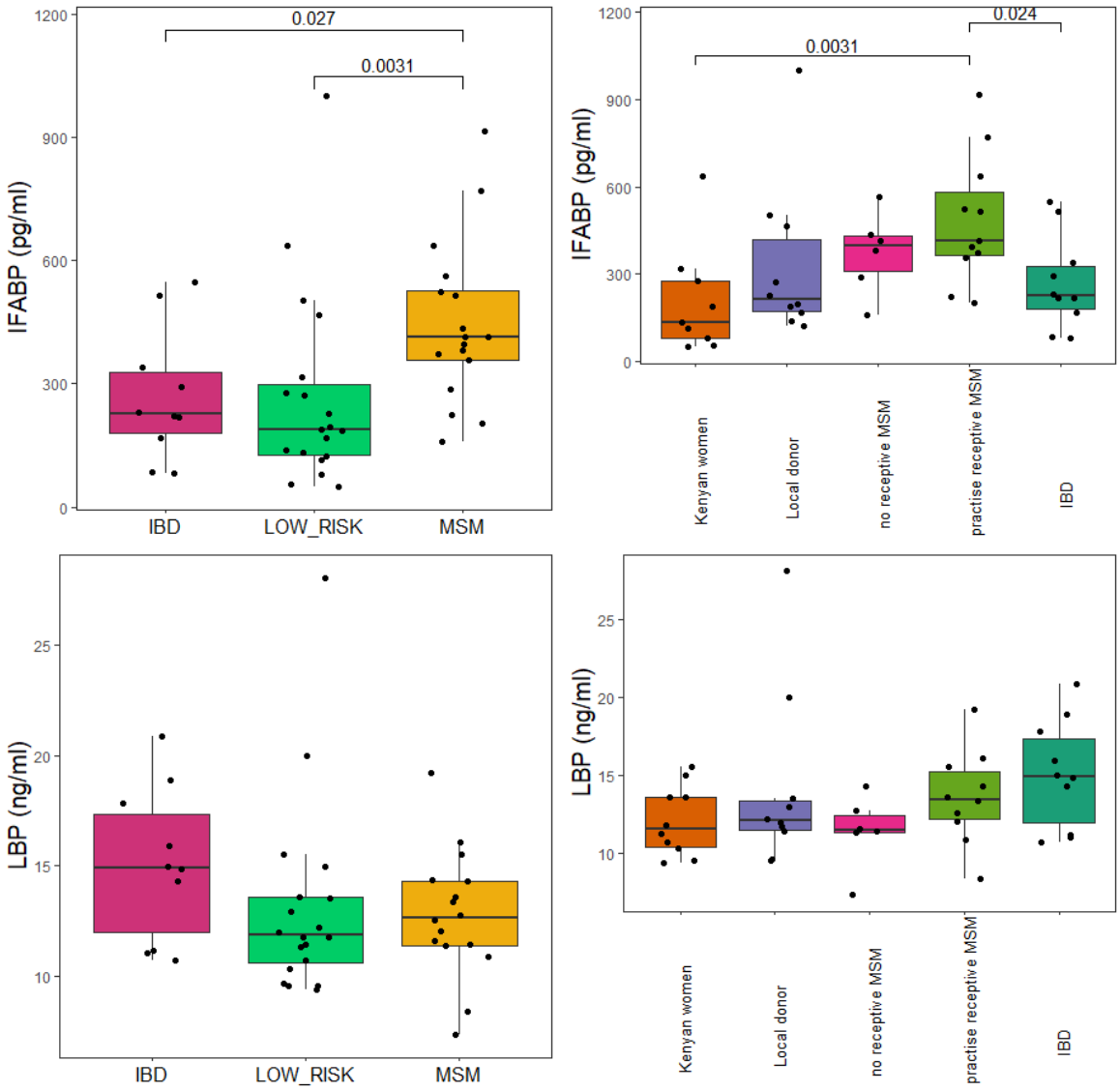


Figure 6. Elevated microbial translocation in MSM cohort compared to healthy donor samples. Color of boxplot indicate risk group. The low risk group include healthy donor samples from Kenyan women and local donors. Boxes represent IQR, with lines denoting the median. Comparisons were done using a Kruskal-Wallis test including turkey's post-hoc pairwise analyses. Post-hoc pairwise comparisons p- values are shown on top of each analyte.

4.4.5 Association between markers of MT, microbiome diversity, systemic immune activation and rectal secretion immune activation

We proceeded to conduct Spearman correlation analyses on a subset of systemic and rectal immune activation biomarkers after excluding those that were highly correlated. Correlation analysis on markers of MT markers, T cell activation and local rectal immune activation was visualized by heat map. The MT markers, IFABP and sCD14, were positively correlated with systemic T cell activation phenotypes such as %CCR5+CD4+, %Ki67+ CD4+, TEM CD4+ and %TTN CD4+ T cells in both HIV negative and HIV positive MSM. There was minimal correlation between markers of MT, alpha diversity and rectal secretion immune activation biomarkers as seen in **Figure 7A**. We also found positive correlations between markers of MT and systemic T cell activation in MSM stratified by their receptive anal sex practise status as shown in **Figure 7B**. The strongest correlations were found in MSM who did not engage in receptive anal sex including IFABP and sCD14 with CD38+HLA-DR+ CD4+ T cells, Ki67+ CD4+ T cells and CD8+ T cells. The patterns were similar in MSM who engage in receptive anal sex although the correlations were not significant. There was minimal correlation between markers of MT and rectal immune activation biomarkers as was the case with HIV status analysis. Of note, sCD14 was negatively correlated with IFN-B, Pentraxin-3 and TNFR-1 in MSM who practised receptive anal sex.

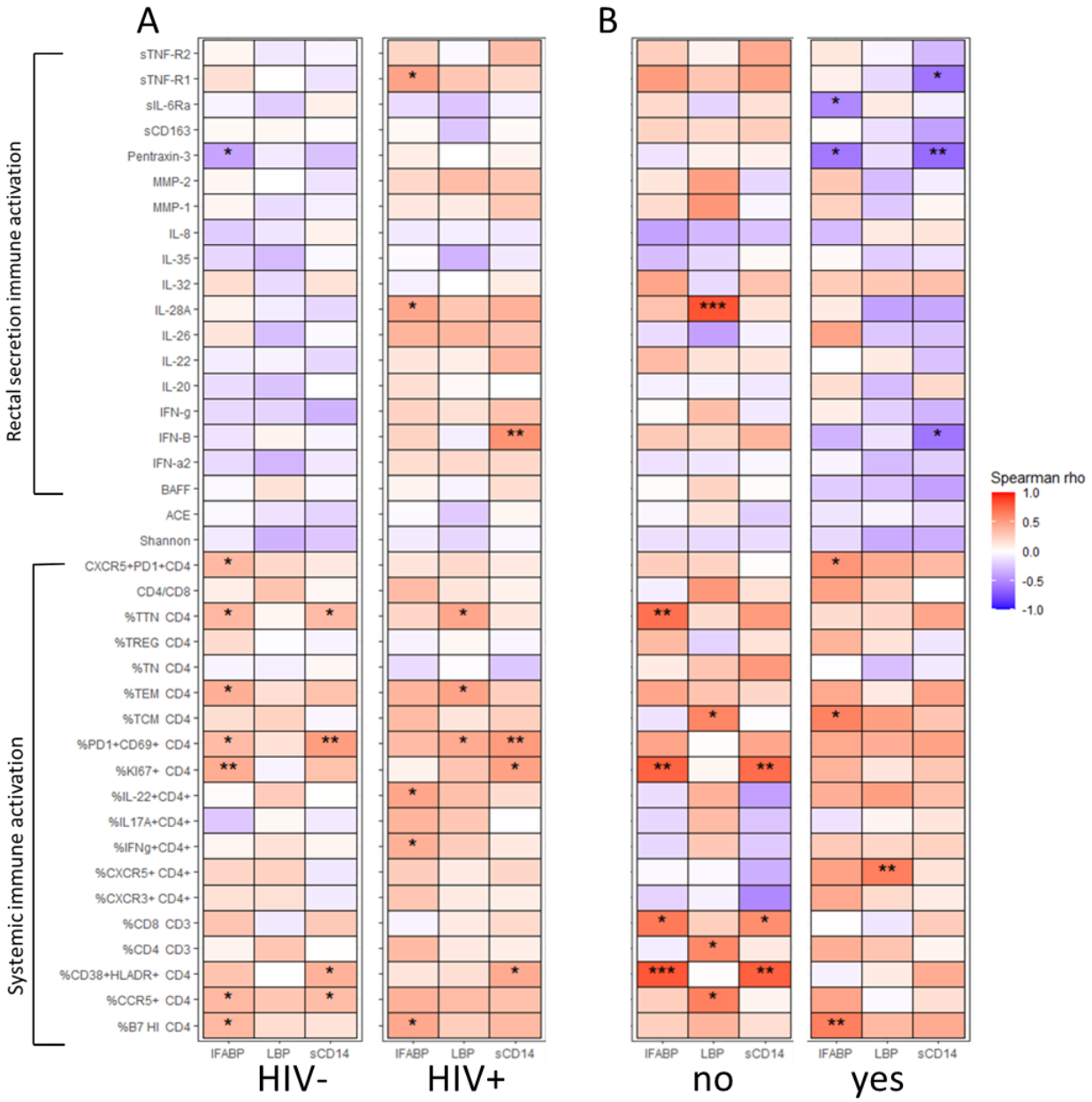


Figure 7. Heat map of spearman correlation analysis. The heat map represents correlation between markers of MT, systemic immune activation, alpha diversity and rectal secretion inflammatory biomarkers in HIV negative and HIV positive men (**A**). The data was also stratified by HIV status and we looked at MSM who never engage in receptive anal sex with those that practise receptive anal sex with at least one partner type (**B**) in HIV negative MSM. The red squares indicate positive correlations while blue squares indicate negative correlations. Stars indicating significance are given as *($p<0.05$), **($p<0.01$) and ***($p<0.001$).

4.4.6 Microbial translocation markers not associated with gut microbiome clusters

In addition to the alpha diversity metrics, we also compared markers of MT on two microbial clusters identified through the unsupervised PAM algorithm. We restricted the analysis to HIV negative men and had microbiome data available for a total of n=41 individuals, comprising of *Bacteroides*-spp (n=13) or *Prevotella* -spp (n=28) dominant identified individuals. There wasn't a significant difference in markers of MT between the *Bacteroides*-spp or *Prevotella* -spp dominant microbiota in HIV negative samples as seen in **Figure 8**. IFABP levels were elevated in the *Prevotella* -spp dominant cluster (median 1423.9, IQR: 864.3-1811.3 pg/ml; versus median 868.4, IQR: 719.3-1562.1 pg/ml for the *Bacteroides*-spp cluster, respectively; p=0.4). sCD14 levels were also elevated in the *Prevotella* -spp dominant cluster (median 1772.6, IQR: 1652.7-2373.3 pg/ml; versus median 1603.2, IQR: 1416.0-2208.2 pg/ml for the *Bacteroides*-spp cluster, respectively; p=0.31).

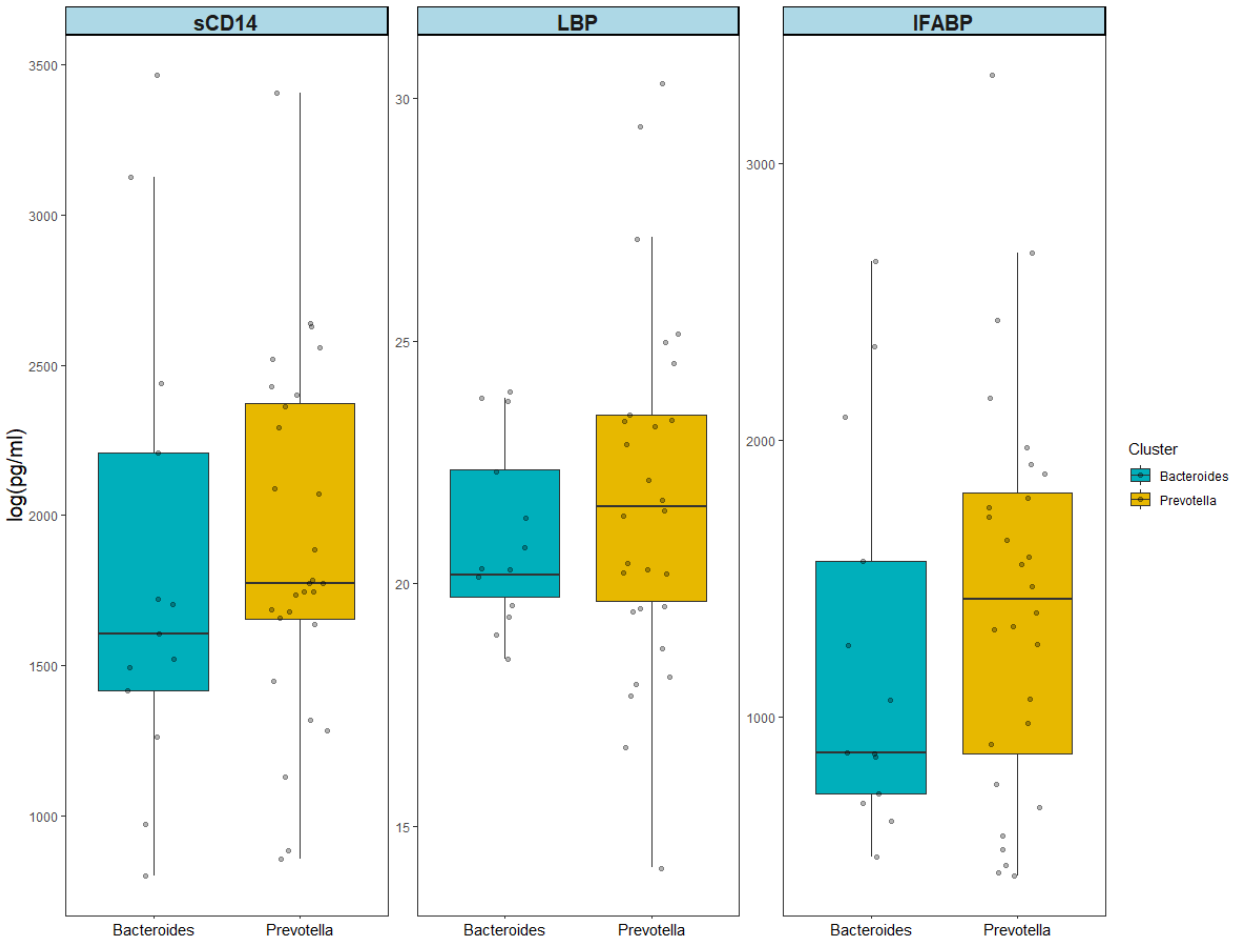


Figure 8. Microbial translocation is not associated with gut microbiome cluster. Color of boxplot indicate risk group. Boxes represent IQR, with lines denoting the median. Comparisons were done using a Mann-Whitney U test.

4.5 Discussion

By quantifying markers of MT indirectly from plasma samples, we were able to assess its relationship with systemic and rectal immune activation and gut microbiome structure on different study groups. Firstly, we found only sCD14 to be significantly elevated in HIV infected MSM in comparison to HIV negative MSM. Men who engaged in recent receptive anal sex had elevated levels of LBP. To understand where the distributions of markers of MT markers fit into a broader context, we also included external high and low risk samples and found increased markers of MT in IBD patients and the MSM samples compared to healthy donors. We found positive correlation between markers of MT and systemic T cell activation while correlations

with rectal secretion immune activation were minimal. Our results highlight that engaging in receptive anal sex independent of HIV status is associated with markers of MT and can potentially alter the host immune profile to be conductive for a pro-inflammatory environment.

Since the first study to link markers of MT to HIV infected patients and immune activation, subsequent studies have found similar findings and strengthened the association of HIV infection with elevated markers of MT [84,185]. Circulating levels of LPS in ART naïve patients independent of CD4+ cell count was found to be associated with increased risk of faster disease progression [87]. A study in SIV infected macaques also showed that treatment with an antimicrobial drug (sevelamer) that binds LPS to significantly reduce markers of MT, immune activation and reduction in viral replication [186]. Although the trial of sevelamer therapy on untreated HIV patients did not reduce markers of MT, there are ongoing attempts to gain insights into the dynamics of immune activation and markers of MT[187]. The contribution of markers of MT to inflammation in HIV disease progression requires more research as trials to reduce markers of MT have not translated to clinical improvements in untreated HIV infected patients[187]. Our finding of limited evidence of markers of MT in HIV infected men is contrary to most of the published literature. There have been inconsistent findings related to markers of MT in HIV infection that could be due to a variety of factors such as geography, length of HIV infection, clinical outcomes and difference in biomarker selection to measure markers of MT markers [177,188,189]. Reports from Africa have also looked at markers of MT in people living with HIV. A longitudinal study of chronically HIV infected patients from Uganda found no difference in amount of circulating LPS or sCD14 over time in different HIV-1 disease progression groups. They also showed baseline levels in Ugandan and USA uninfected subjects had similar level of LBP and LPS but that the sCD14 was elevated in the USA subjects. Other inflammatory cytokines and endotoxin core antibody (EndoCAb) levels were elevated in the Ugandan subjects [179]. Another study in Ugandan children infected with HIV found no link between markers of MT and immune activation. The researchers quantified bacterial DNA directly with broad range PCR and NGS as opposed to markers of markers of MT due to challenges with reproducibility associated with them. They found that although immune activation subsided following ART initiation and 96 weeks of follow up, there was no difference in bacterial DNA levels contrary to studies in USA children where markers of MT was persistent even after CD4 T cell

reconstitution over 96 week follow up [180,189]. These findings are similar to our results in that we did not see significant markers of MT in HIV infected men in comparison to HIV negative men, apart from sCD14.

A potential explanation for the different findings in African settings could be having a higher level of baseline markers of MT. We compared healthy donors and IBD patients to our MSM cohort to address this hypothesis. We found elevated levels of IFABP in the MSM cohort, compared to the others, but no difference in LBP levels. We also found the HIV negative men to have elevated levels of IFABP in comparison to the healthy donor samples. The IBD patients trended towards higher levels of LBP and IFABP, and although the results were not significant, the sample size is small and there may be variability in the severity of their IBD condition. Our findings show that background markers of MT markers could be higher in our cohort compared to healthy donor samples and this is not due to different geographical settings as the healthy Kenyan women samples are from the same region. Geographic explanations could include the prevalence of GI infections and environmental exposures endemic to the region which could alter immunological responses [190]; however, the healthy donor samples we collected from women in this region does not support this hypothesis. Another explanation could be high-risk behaviour in the MSM that could increase baseline markers of MT levels. Interestingly we found elevated markers of MT in MSM who engage in receptive anal sex compared to men who didn't engage in receptive anal sex over a similar time period. The data supports the thought that local trauma at the mucosal site due to engagement in receptive anal sex could contribute to markers of MT. There have been other studies also showing engaging in high risk behaviour is associated with a pro inflammatory environment characterized by endotoxemia, decreased CD4/CD8 T cell ratio and elevated levels of pro inflammatory cytokines [183].

We investigated whether markers of MT was associated with immune activation as has been found in previous studies [185]. We compared markers of MT to systemic T cell activation and rectal secretion immune activation levels. After stratification of samples by HIV status, we found associations between markers of MT and systemic CD4+ T cell activation regardless of HIV status. There was surprisingly very little correlation of the markers of MT markers and rectal secretion immune activation markers. We also looked at the relationship between markers of MT

and systemic immune activation based on engagement in receptive anal sex. We found the correlations to be similar to those we found based on HIV status as markers of MT was correlated with activated CD4+ T cell subsets. We also found the correlation with markers of MT and rectal immune activation to be minimal. Our results are consistent with plenty of studies showing markers of MT to be associated with inflammation including activated CD4+ and CD8+ T cells, activation of monocytes and pro inflammatory cytokines[159,172]. Our results of minimal correlation of markers of MT and rectal immune activation is in contrast to other studies showing in biopsy samples that markers of MT correlated with T cell activation [191] although we did not characterize mucosal T cells directly. Our results support the idea of MT being associated with T cell activation and contributing to inflammation. A previous study on MSM engaging in condomless receptive anal sex had found greater inflammation in the rectal mucosa compared to men who never engaged in rectal anal sex [133]. The elevated markers of MT in MSM who engage in receptive anal sex and the correlation with CD4+ T cell activation provide support for a pro inflammatory rectal environment.

We also found limited relationship between the rectal microbiome diversity and markers of MT. There was no difference in markers of MT between microbial clusters and no correlation with alpha diversity metrics. Previous studies have showed HIV induced gut microbiota such as increase in the genera *Prevotella* to be associated with mucosal and peripheral T cell activation and markers of MT by affecting metabolite pathways [149,167]. These findings share a common feature in linking immune activation with gut microbial changes and markers of MT in HIV infected subjects. The results could also be confounded by MSM status which has been found to independently also be linked to the rectal microbiota and markers of MT [125,183]. We did find the *Prevotella* microbiome cluster to be associated with CD4 T cell activation and pro-inflammatory cytokines in HIV negative individuals in Chapter 5.5. We did not find any significant differences between the gut microbiome clusters and MT markers. The markers of MT were also not correlated to CD4 T cell activation or alpha diversity metrics. Direct comparison with non-MSM group could be helpful in determining whether sexual orientation itself has a bigger influence in gut microbiota structure.

There are limitations in our study that are worth discussing. Although markers of MT are well established in the field, inconsistency in reproducibility has led to incorporation of direct bacterial DNA quantification in studies, although caution is warranted as there are possibilities for contamination. One of the ways to achieve this is through broad range PCR sequencing or next generation sequencing (NGS) sequencing. Due to our sample size, we also did not have enough power to assert a lack of association.

In conclusion, we found limited evidence for markers of MT in HIV infected MSM compared to uninfected MSM, although geographical context should be taken into consideration. We found MSM to generally have higher markers of MT compared to other low risk sample groups, with MSM engaging in receptive anal sex to have the highest markers of MT levels. We also found markers of MT to be associated with CD4+ T cell activation supporting the notion that MSM who engage in receptive anal sex can have a proinflammatory rectal environment that can affect their risk to infection due to local trauma.

5. Mucosal and fecal sampling comparison

5.1 Rationale

Research into the gut microbiota has grown extensively in recent years thanks to advances in next generation sequencing methods such as 16S rRNA gene sequence analysis, which is commonly used to profile the composition of bacterial genera. The gut microbiota has been found to have a crucial role in health and disease, implicated in numerous diseases such as IBD, cardiovascular disease, diabetes and colon cancer[192–194]. Host physiology, genetics, environment exposure and gut microbiota interactions have been shown to have a symbiotic relationship that maintains homeostasis. The gut microbiota has a complex set of microbes in the lumen and mucosa that allow it to have a variety of functions along the GI tract such as metabolism, immunologic development and gut epithelial health[192]. Since most studies use fecal samples, the role of mucosal associated bacteria has not been studied sufficiently and it is unclear whether feces samples reflect the gut microbiota host interactions along the GI tract.

There are differences in the bacterial communities that occupy different niches along the GI tract [195]. Previous studies suggest that not every species within bacterial communities will interact with the human mucosa. While certain bacterial taxa reside in the lumen, others attach to the mucosal layer that covers the epithelium. Due to anatomy and physiology of the GI tract, variations have been found in the function of the bacterial taxa found within these niches (fecal vs mucosal). The oxygen gradient differs along the intestinal mucosa, with the intestinal submucosa having the highest oxygen gradient and moving to near anoxia in the middle of the lumen [196]. This effect translates to mean that fecal (non-attached) microbes have a higher chance to be anaerobic than mucosal (attached) microbes and has indeed been observed [197]. It is thought that mucosal adherent microbiota directly interact with the host epithelium while the lumen associated microbiota indirectly interacts with the host through metabolic products and is enriched with phylotypes that enhance energy extraction from diets [198]. It has been observed that the mucosal adherent microbiota is also enriched for asaccharolytic bacteria which metabolize amino acids thought to be derived from the shedding of intestinal epithelial cells [197,199]. Another notable differences between the sites is that mucosal bacteria have been found to be less diverse than fecal bacterial communities, although they share some of the dominant genera [199,200].

Bacterial taxa along the digestive tract are more stable than fecal samples, possibly because they are more host-associated and less dependent on environmental exposure. The stability is translated into a similar diversity along the digestive tract from the ileum to the colon and rectum[201,202].

The gold standard for gut microbiome studies has relied on characterization of stool samples. The non-invasive nature of stool collection makes it convenient to carry out, although it mainly captures information about fecal bacterial communities which may be more transient than mucosal bacteria. The mucosal microbiota might also have a more important role in shaping the response against pathogens because it is closer to the epithelial and immune cells. Rectal swabs offer the potential of studying both the fecal and mucosal bacterial communities while being less invasive than biopsies. Rectal swab comparison with stool and mucosal biopsies comparison has shown that they capture similar bacteria communities as stool samples but that they might differ from mucosal biopsy samples [200,203,204]. Although the microbiome structure is similar in rectal swabs and stool, it still unknown whether there are host immunity responses that might differ according to the sample collected. In this study we aim to capture fecal and mucosal bacteria using 16S rRNA amplicon sequencing, comparing sample collection before and after enema use. We compared local cytokines and systemic CD4+ T cell subsets in relationship to the bacterial communities present by both of these sampling methods.

5.2 Hypothesis

There will be differences in the bacterial communities with the mucosal samples having less diversity than luminal samples. Adherent bacteria communities in the mucosal samples would have stronger associations with rectal and systemic immunity than fecal bacterial communities.

5.3 Objective

To characterize the gut microbiota using 16srRNA sequencing and compare bacterial diversity and composition between the mucosal and fecal samples. To correlate expression of soluble

markers of inflammation in the rectum and systemic CD4+ T cell subsets with bacteria present in the mucosal and fecal samples.

5.4 Section specific methods

The detailed methods can be found in detail in chapter 2.

5.5 Results

5.5.1 Gut microbiome composition between sites

We used the unsupervised PAM algorithm on the combined sample types and found moderate support for two clusters that were *Bacteroides-spp* or *Prevotella-spp* dominant microbiota (silhouette index of 0.4). We found that the *Prevotella- spp* dominant cluster was much more prevalent in feces compared to mucosal samples (84% versus 43% respectively; Chi square, p=0.01). Approximately one-quarter (22/84) of samples switched from *Prevotella- spp* dominant cluster to *Bacteroides- spp* dominant clusters from the fecal samples to the mucosal sample. There was no switching from the *Bacteroides - spp* to *Prevotella - spp* dominant cluster from the fecal samples to mucosal sample. The clusters were unbalanced in their distribution in the fecal samples (84% *Prevotella- spp* dominant) making it difficult to conduct stratified analysis (**Figure 9A**). The dataset was stratified by HIV status in order to account for HIV infection effect on gut microbiome structure. We found there was substantial overlap of common genera (prevalent in >50% of participants) present in both sample types. There was more differentiating genera present in the fecal samples than the mucosal samples as visualized in the Venn diagram (**Figure 9B**). In order to account for the distribution of the less common genera, we reduced the threshold of prevalence to 25% and found similar distribution of genera across both sample sites, which shows that most of the less common genera were present in both sample sites. The difference between the 50% and 25% prevalence thresholds point to uneven distribution of the genera with the fecal samples having more genera at higher prevalence than the mucosal samples. Across all alpha diversity metrics, we found that the alpha diversity was lower in the mucosal samples (**Figure 9C**). These were statistically significant for the Shannon (p=0.004) and Simpson (p=0.007) metrics that capture overall richness.

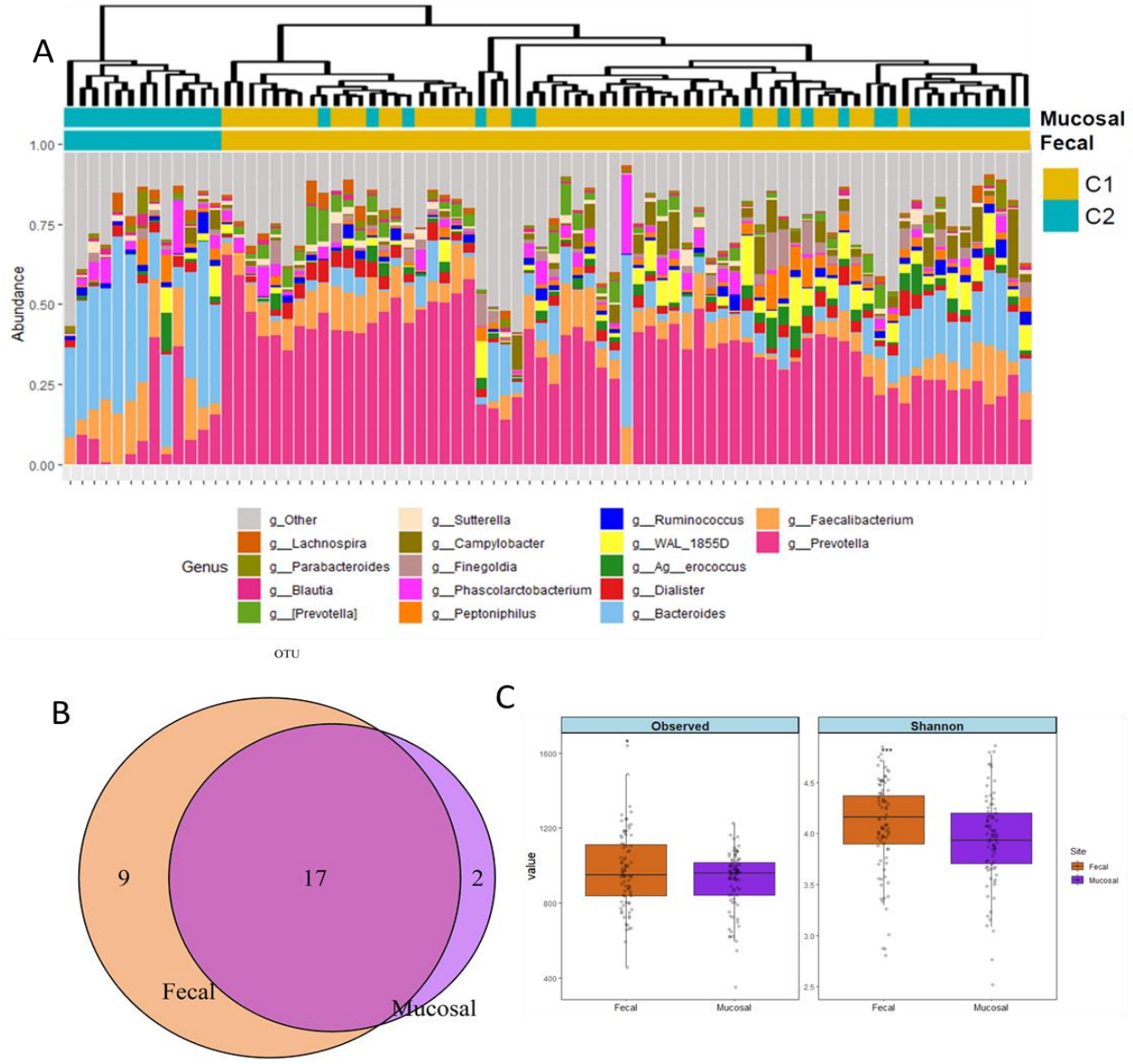


Figure 9. Comparison of sampling method by microbiome structure and diversity. (a) Genus-level taxonomic summary plot. Genera with mean relative less than 0.1% abundance are binned together into the 'other' category. Cluster classification by sampling method are indicated for each sample (top rows). Each column corresponds to an individual, with the relative abundance of each genera in the fecal sample colour-coded and indicated below. Samples are ordered by the NMDS1coordinate on the Bray Curtis similarity matrix. (b) Venn diagram depicting genera overlap between the sample methods. (c) Alpha diversity boxplots by sampling method. Comparisons were made using the Mann–Whitney U test. Stars indicating significance are given as *($p<0.05$), **($p<0.01$) and ***($p<0.001$).

5.5.2 Differential abundance between sample sites

We visualized differential abundance between the sample sites using a volcano plot. Of the 70 genera that were >25% prevalent, 26 genera were significantly different between the sample types, with 5 genera that passed the significance threshold after multiple test correction (Wilcoxon test with false discovery rate multiple test correction, **Figure 10**). These genera were more frequent in fecal samples, and included *Finegoldia, Corynebacterium, Anaerococcus, Petoniphilus* and *ph2*. The mean relative abundances of these genera collectively is less than 10% and are not the core representative of the overall microbiome composition. These five fecal-specific genera had low correlation between the sample types while the remaining genera showed a high correlation between the sample types. The only genera in higher abundance in the mucosal samples were *Ruminococcus* and *Methanobreviba*, although these differences were not significant.

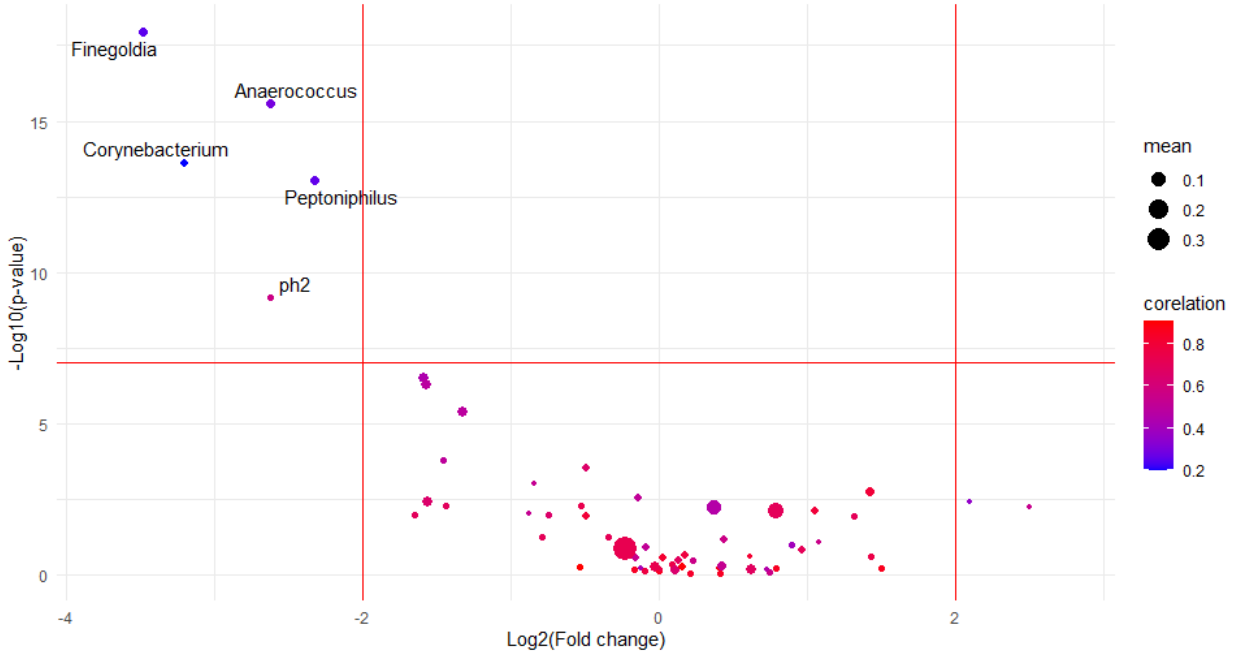


Figure 10. **Volcano plot showing genera significantly different by sampling method.** Log transformed fold change and p- value is plotted on the x-axis and y-axis respectively. The colour gradient represents the correlation between the sampling method for each genera with red and blue shading for positive and negative correlation respectively. The size of each dot represents the mean abundance of each genera.

5.5.3 Relationship between rectal inflammatory markers and microbiome clusters

To determine if rectal microbiome composition correlated with rectal inflammation, we quantified inflammatory cytokines in post enema rectal sponge secretions and correlated these with microbiome clusters (**Figure 11**). The *Prevotella*- spp. dominant cluster in feces was associated with elevated levels of MMP-1, IL-19, IL-20, IL-8, IL-35, GP130, IFN-a2 ($p<0.05$). The *Prevotella*- spp. dominant cluster in mucosa was associated with elevated levels of IL-26, IL-34 and IL-19 ($p<0.05$). Although HIV status was only minimally associated with differences in rectal inflammatory markers, we also stratified our mucosal cytokine-microbiome analysis by HIV status and analysed the HIV negative samples. The mucosal sample had significantly elevated levels of IL-26, IL-34 and MMP-2 ($p<0.05$) in the *Prevotella*- spp. dominant cluster. The fecal sample had significantly elevated levels of IL-26, MMP-1 ($p<0.05$) in the *Prevotella*- spp. dominant cluster. The associations seen in the fecal samples were not present in the stratified data due likely to the data set being underpowered as the sample size decreased.

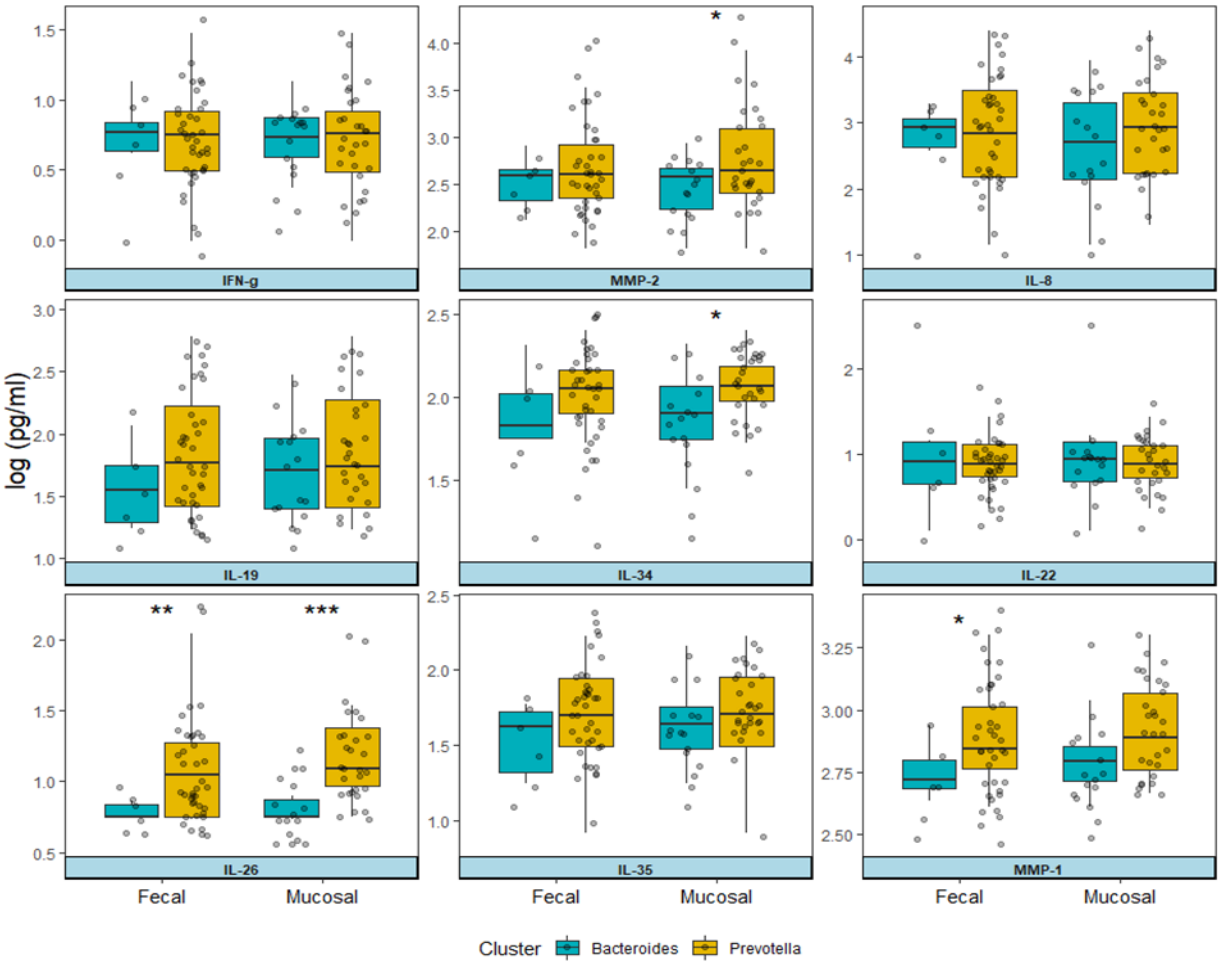


Figure 11. Microbiome cluster association with rectal secretion soluble markers of inflammation. Cluster classification was coloured as blue for *Bacteroides* cluster and orange for the *Prevotella* cluster in HIV negative samples. Boxes represent IQR, with lines denoting the median. Comparisons were made using student's t-test. P- values are shown on top of each analyte: *p < 0.1, **p < 0.05, ***p < 0.001.

MFA was performed with the cytokines labelled *a priori* according to their known functions (**Figure 12**). PC1 accounted for 27.1% of variation in cytokine expression, with 15.7% variation explained by PC2. Most pro-inflammatory cytokines (e.g. IL-6 and IL-10 families) clustered on the positive side of PC1 on the loadings plot, which was further away from the microbiome cluster centroids in both sample types. The centroids of the microbiome clusters were in the centre of the plot showing a weak association with the cytokines which were mostly on the positive side of PC1. There was a lot of overlap in the 95% confidence interval eclipses and the centroid distance separation between the clusters was not significant in both sample types (PERMANOVA; p=0.27 and p=0.21 for the fecal and mucosal sample type respectively).

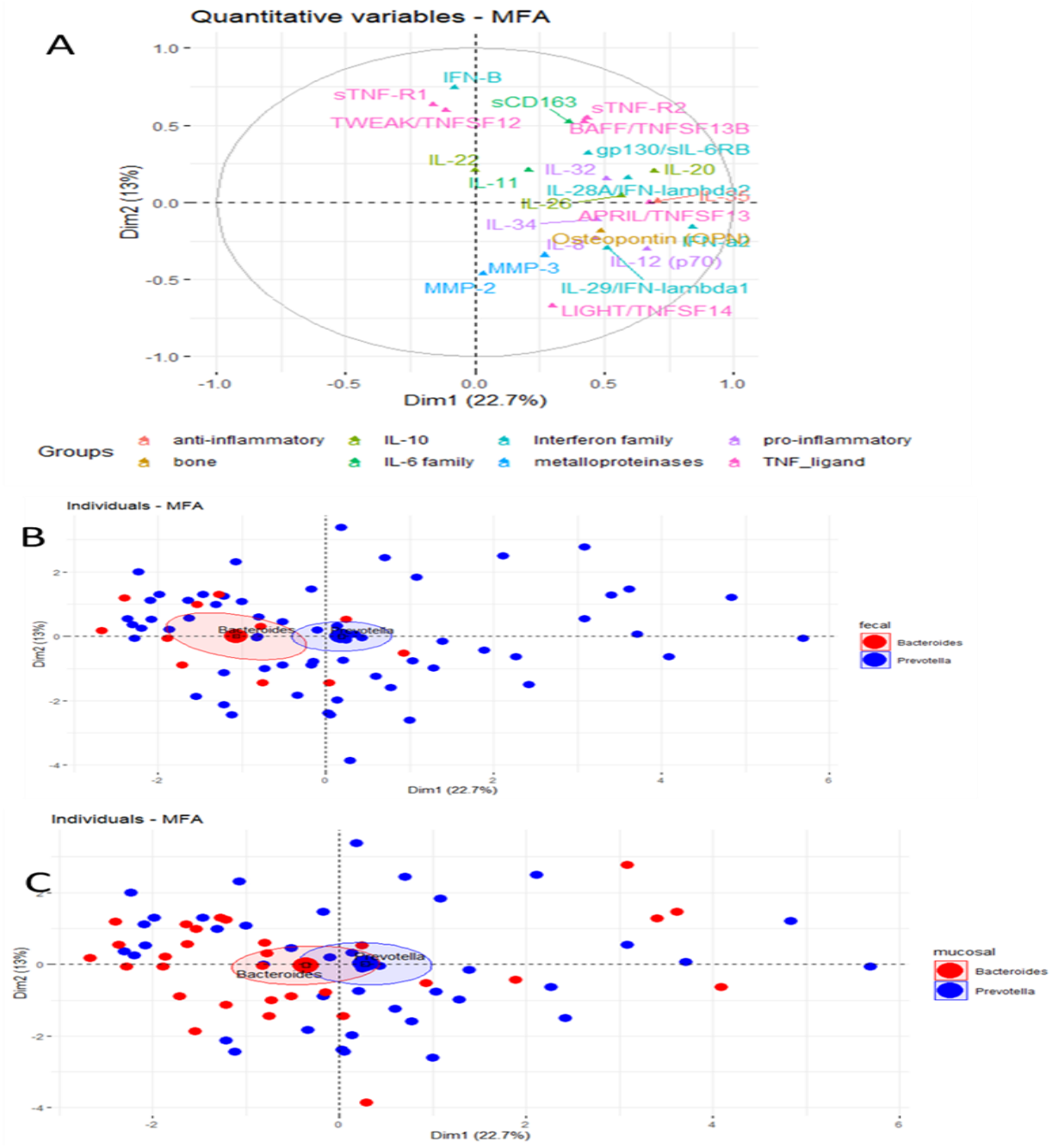


Figure 12. Multiple factor analysis of microbial clusters with rectal secretion soluble markers of inflammation. PCA was used to identify the first 2 principal components which explain 35.7% of the variation in the dataset. Part **A** is a loading plot which shows the markers which load on the respective principal components which are coloured according to pre-defined groups. Part **B** and **C** is a scores plot which shows the scores on each principal component for each of the observations in the fecal and mucosal samples respectively.

A linear regression model was conducted to assess the effect size of the rectal inflammatory factors by cluster classification on both sample types. The model was adjusted for HIV status and rectal inflammatory cytokine concentration were normalized by a log10 transformation. The trend showed most of the rectal inflammatory biomarkers to be increased in the *Prevotella* cluster but the effect size on most of the biomarkers was greater in the fecal samples than the mucosal samples as shown in **Figure 13.**

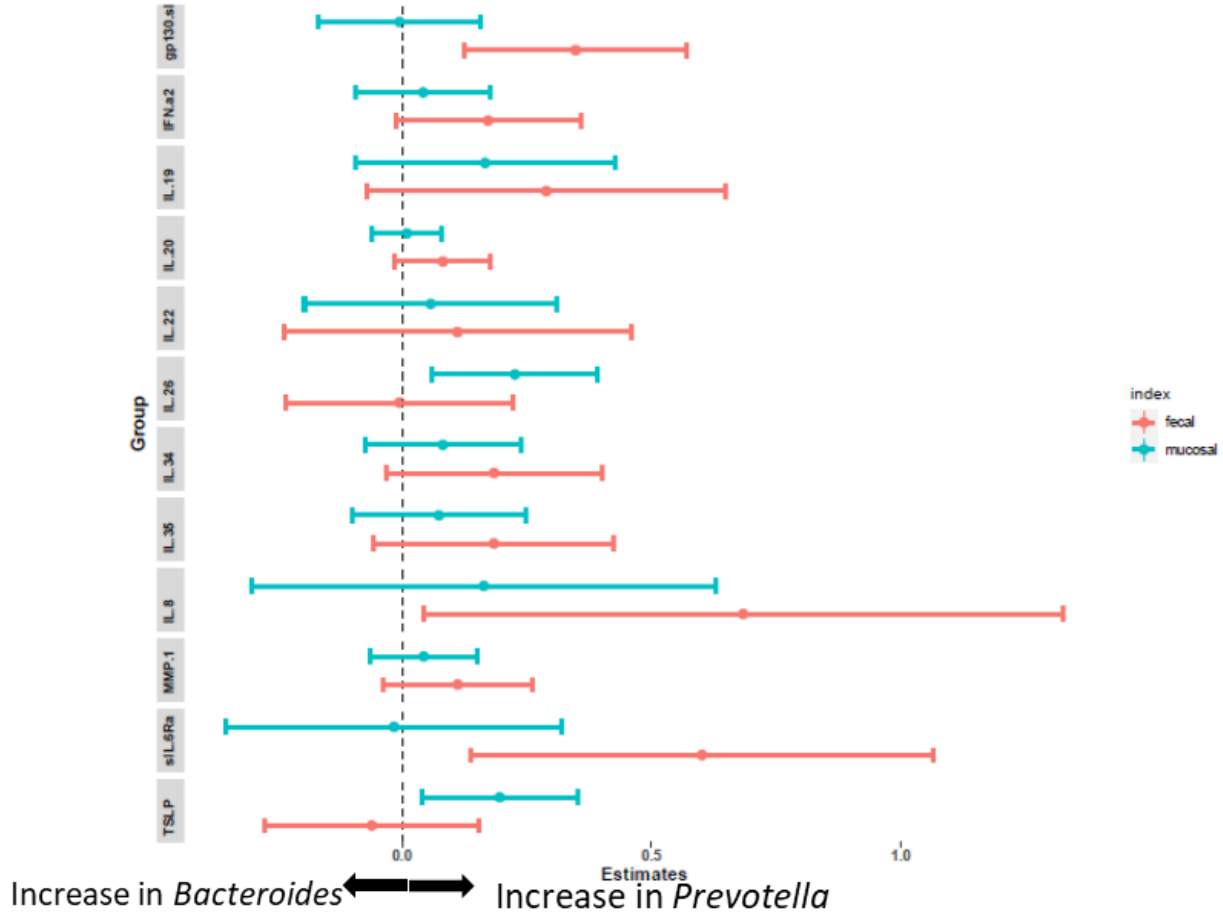


Figure 13. Effect size of association between microbial cluster and rectal inflammatory biomarkers. A linear regression model with microbiome cluster as predictor of rectal inflammatory biomarkers adjusted for HIV status was conducted. The effect size was considered significant if the 95% confidence interval did not cross the dashed line.

5.5.4 Relationship between systemic T cells and microbiome composition

We next compared gut microbiome composition and systemic CD4+ T cells, given the link between gut microbiota and chronic immune activation that has been observed in HIV infected individuals [167]. We analyzed immune cells in the blood using different flow cytometry panels (**Supplemental data, Table 5**) in HIV negative MSM. Various CD4+ T cell subsets were elevated in the *Prevotella- spp* cluster in both sites. *Prevotella*- spp dominant participants defined in mucosal samples had higher frequencies of IL-22 producing (p=0.01), effector memory (p=0.02) and Ki67+ CD4+ T cells (p=0.04, **Figure 14**) in their blood. The trend was similar for the same CD4+ T cells in the fecal samples but did not reach a significance level.

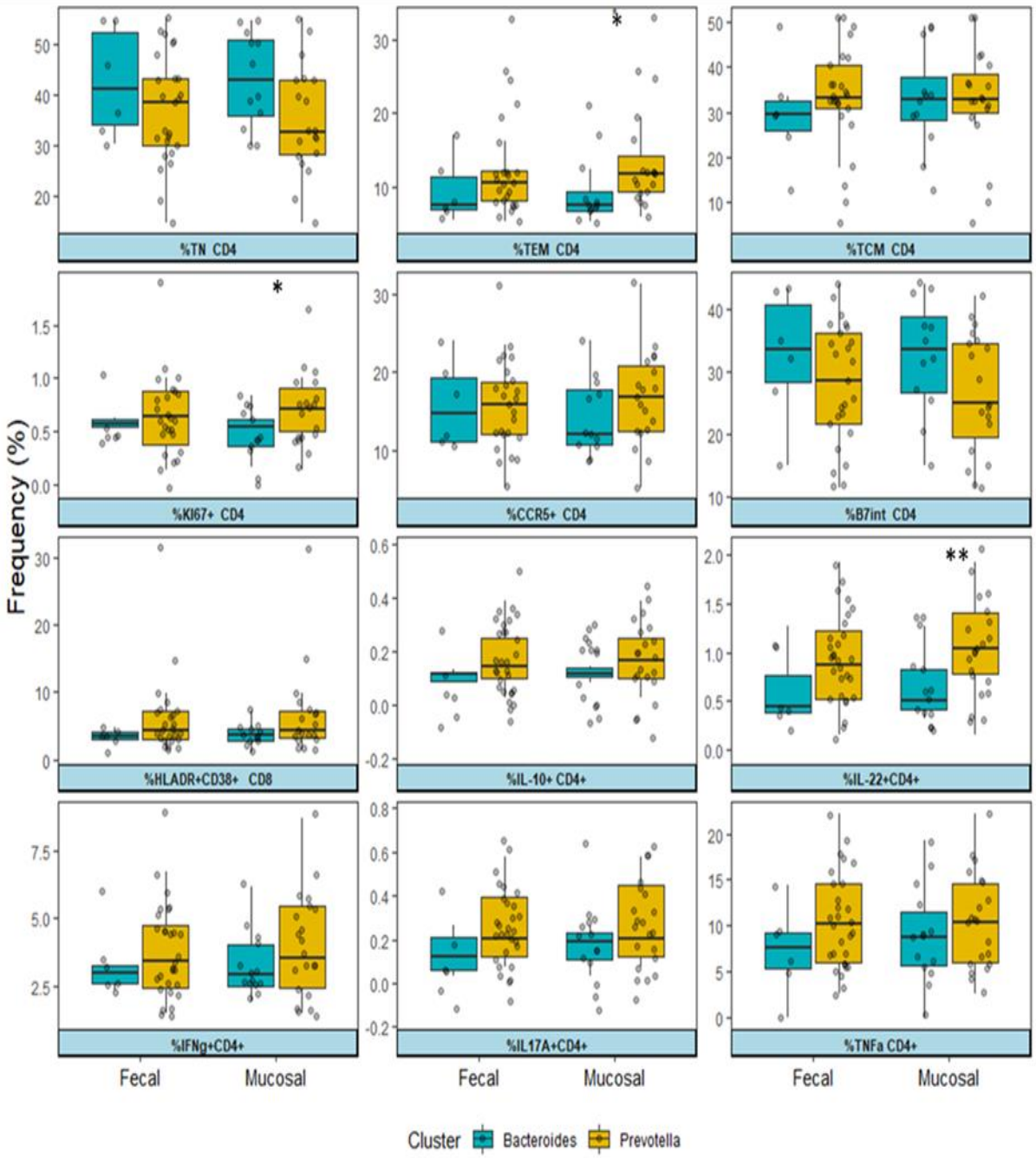


Figure 14. Microbiome cluster association with systemic CD4+ T cell subsets. Cluster classification was coloured as blue for *Bacteroides* cluster and orange for the *Prevotella* cluster in HIV negative samples. Boxes represent IQR, with lines denoting the median. Comparisons were made using the Mann–Whitney *U* test. P- values are shown on top of each analyte: *p < 0.1, **p < 0.05, ***p < 0.001.

MFA analysis of systemic CD4+ T cell subsets demonstrated that the top 2 principal components explained 40% of the variability in systemic CD4+ T cell frequencies (**Figure 15**). The *Prevotella*-spp dominant cluster was positively associated with CD4 T cell activation, Th1, Th17 and activated memory CD4+ T cell phenotypes in both sample types. The *Prevotella- spp* cluster had a more inflammatory systemic immune phenotype compared to the *Bacteroides- spp* cluster, similar to what was observed previously with the rectal inflammatory biomarkers. There was some level of overlap between the 95% confidence interval eclipses and the centroid distance separation between the clusters was not significant in both sample types (PERMANOVA; p=0.25 and p=0.07) for the fecal and mucosal sample type respectively. Although the *Prevotella*-spp dominant participants had elevated levels of systemic T cell activation in both sample types, classification was more robust in the mucosal samples.

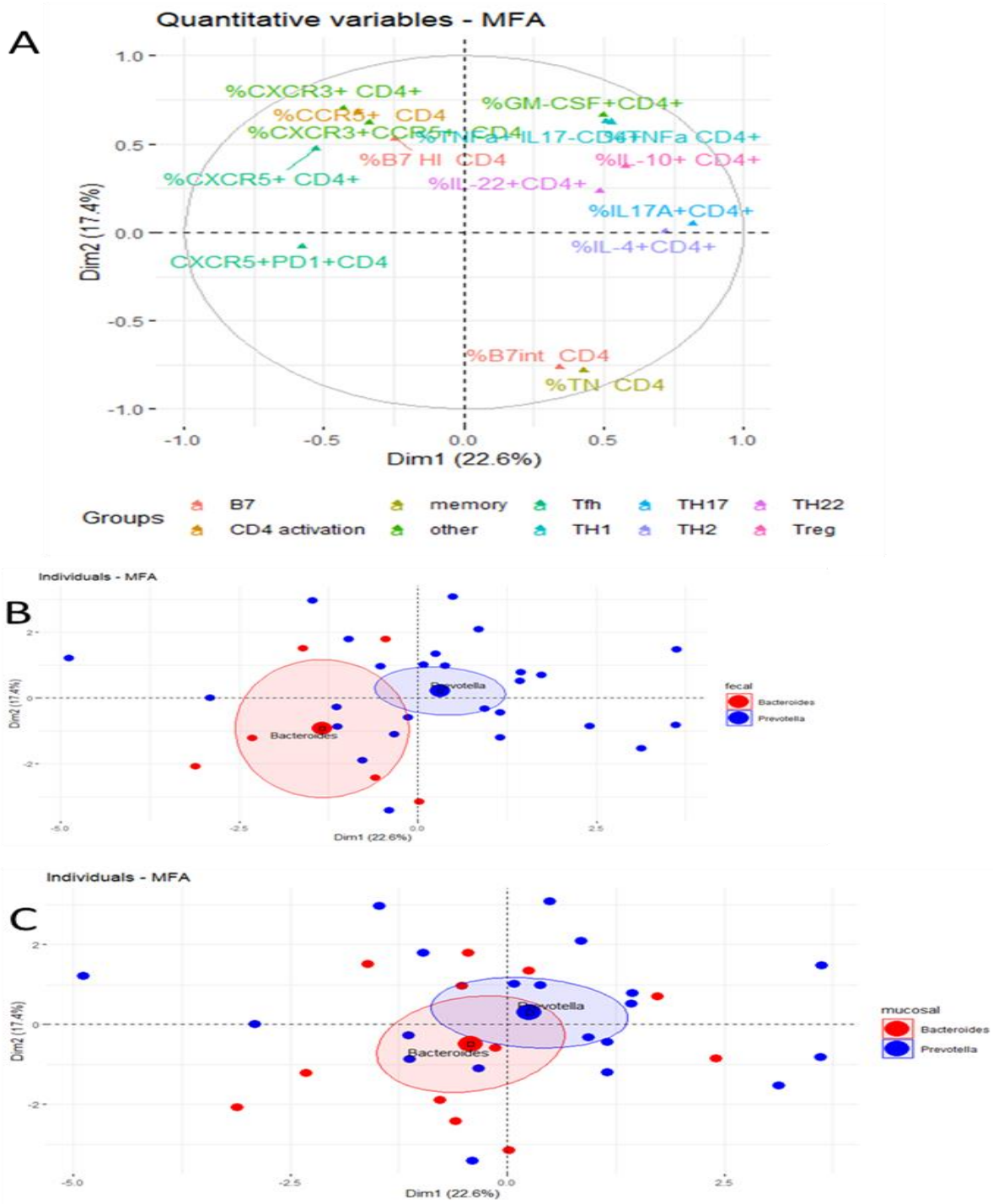


Figure 15. Multiple factor analysis of microbial clusters with systemic CD+ T cell subsets. PCA was used to identify the first 2 principal components which explain 35.7% of the variation in the dataset. Part **A** is a loading plot which shows the markers which load on the respective principal components which are coloured according to pre-defined groups. Part **B** and **C** is a scores plot which shows the scores on each principal component for each of the observations in the fecal and mucosal samples respectively.

A log fold change bar plot was used to visualize the difference in the systemic T cells by cluster classification on both sample types. The samples were stratified by HIV status and only included HIV negative samples. The trend showed most of the systemic CD4+ T cell subsets including those producing IFN-g, KI67 and IL-22 to be increased in the *Prevotella* cluster but the log fold change on most of the biomarkers was greater in the mucosal samples than the fecal samples as shown in **Figure 16**. The naïve CD4+ T cell were higher in the *Bacteroides* cluster with log fold change greater in the mucosal sample.

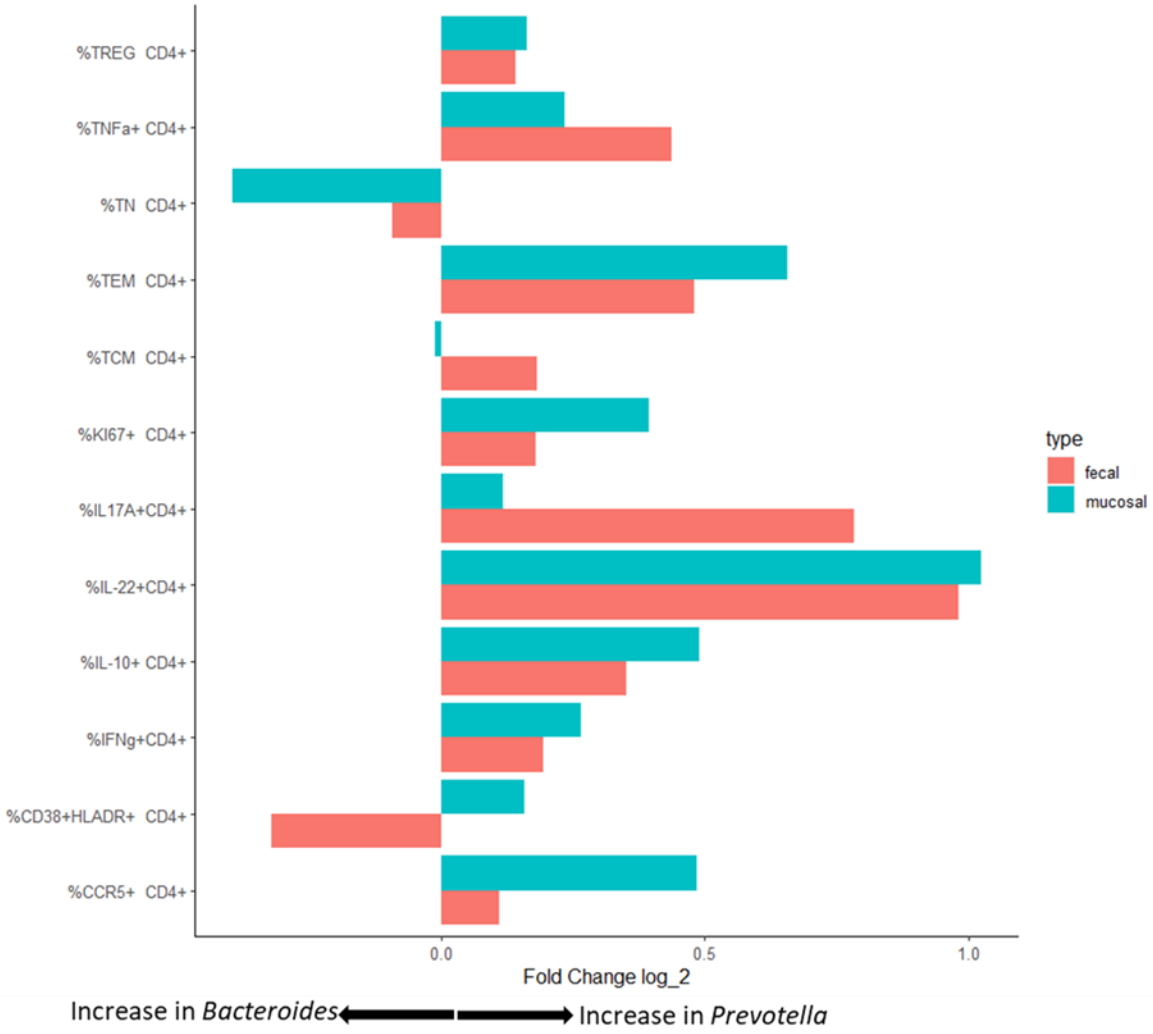


Figure 16. Bar plot of log fold changes between microbial cluster and systemic CD+ T cell subsets in both sample types. The plot shows log fold change in *Prevotella/bacteroides* clusters. The sample type was denoted as blue and red for mucosal and fecal samples respectively.

5.6 Discussion

Understanding the links between the gut microbiome and host immunity using reproducible and practical sampling has the potential to improve patient care and improve gut microbiome research outcomes. In this study we investigated the difference in microbiome structure between collecting rectal swab before and after enema representing fecal and mucosal samples respectively, and how well these explained host immunity. Both sample types shared a lot of similarity in terms of microbiome structure and host immunity associations although they were more pronounced in the mucosal samples for systemic T cell activation. In agreement with previous studies, fecal communities had a higher alpha diversity than mucosal communities. We also found that there was high correlation between most taxa, even though their distribution was uneven between sample types. We found there was a major difference in cluster classification as the *Prevotella-spp* dominant cluster was much more prevalent in the fecal samples, while some of these individuals were classified as *Bacteroides- spp* dominant in mucosal samples. The *Prevotella-spp* dominant cluster in both sample types was associated with a range of rectal and systemic immune correlates including Th17-related and soluble inflammatory factors, cytokine-producing CD4+ T cells, and a skewing of memory T cell populations.

Our results are in line with other studies which show that biopsy samples have a distinct community and less diversity than stool samples [199,200,205]. The fecal sample in our case will more likely resemble stool samples while the mucosal sample should have less diversity due to some of the bacteria being washed away during enema. At the genera level, there was a greater variability in the proportion of dominant genera in the fecal samples than in the mucosal samples. There was a high correlation between the genera, with a small proportion having a significant difference in their relative abundance between samples. These genera were higher in the fecal samples and all of these genera except *Corynebacterium* belong to the *Firmicutes* phylum and the *Clostridia* class which makes them anaerobic. In relative terms, the rectum was found to have a higher abundance of strict anaerobes compared to the ileum, due to various host factors such as lower concentration of antimicrobials and slower transit time and lower oxygen content [206]. The presence of oxygen gradient along the rectal lumen could also explain why some

of these anaerobes were in higher abundance in the fecal samples, as this site can accommodate mucosal-associated oxygen tolerant taxa [197].

There was also a profound effect on the beta microbiota diversity as samples clustered into two noticeable clusters identified as *Prevotella* dominant or *Bacteroides* dominant. The *Prevotella* and *Bacteroides* dominant cluster were characterized by the genera *Prevotella* and *Bacteroides* being the most prevalent in the majority of samples respectively. The prevalence of the *Prevotella* dominant cluster was higher in the fecal sample compared to the mucosal sample. The switch in cluster classification showed a trend towards the *Prevotella* dominant cluster switching to a *Bacteroides* dominant cluster from the fecal to mucosal sample classification. The profound effect on cluster classification agrees with studies that also showed fecal and rectal samples clustered separately on principal coordinates analysis (PCoA) plots [197]. Curiously, both the *Prevotella* genera and the *Bacteroides* genera are from the *Bacteroidetes* phylum and are also anaerobes. Although *Prevotella* and *Bacteroides* are the most prevalent genera in their respective clusters, other less prevalent genera in the cluster could have different forms of metabolism with the expectation that the *Prevotella* cluster would contain more anaerobic bacteria than the *Bacteroides* cluster. These results show that the fecal microbiome is not reflective of the mucosal microbiome and that they have sufficient distinctions in their overall diversity.

When looking at host immune associations, we found similar patterns in both rectal and systemic immune activation with the *Prevotella* dominant cluster being more inflammatory. The associations were more robust with systemic immune activation and in the mucosal samples. In contrast, there was a trend towards greater association of rectal inflammatory biomarkers with the fecal samples. This supports the hypothesis that mucosally-adherent bacteria better reflect the host gut microflora that interact directly with the host. Our data support the inflammatory effects of *Prevotella-spp*. dominant compared to *Bacteroides-spp* dominant communities, which likely underpin the associations between beta diversity and immunological factors. This may be due to a direct bacterial signalling effect in the mucosa leading to immune activation, as has been shown in several systems and diseases ranging from rheumatoid arthritis (RA) to bacterial vaginosis (BV) [207,208]. Separation in clustering analysis was observed for both rectal inflammatory and systemic T cell datasets, along with a theme of more inflammation and its

effects in those with *Prevotella* dominance. The decrease in naive memory T cells compared to effector memory cells in *Prevotella- spp*. dominant individuals is similar to what is observed in HIV infection and could decrease the host ability to initiate new immune responses. We also found the *Prevotella- spp* dominant cluster to have higher rectal concentrations of IL-26, IL-34 and MMP-1. The increase in MMPs (a trend also shown for MMP3) could imply a decrease in barrier function. Intriguingly, IL-26 released by activated Th17 cells mediates protective immunity by its direct antimicrobial activity and it has been found to be elevated in Crohn's disease and rheumatoid arthritis [209]. Circulating production of IL-22 (and to some extent, IL-17) by CD4+ T cells were also increased in those with *Prevotella- spp* dominance. These findings implicate the Th17 response against *Prevotella- spp* since IL-22 is released by Th17 and Th22cells. This is in line with studies in developed countries and also *in vitro* data that have shown *Prevotella- spp* to be associated with Th17 response and implicated in diseases such as RA, BV, IBD and also HIV [210].

Our study had some limitations. Sequencing of V3-V4 regions of 16srRNA only provides taxonomic resolution to the genus level. Use of shot-gun sequencing with greater resolution can help to better understand how the expression of bacterial genes is associated with the immune phenotypes. Our study was cross-sectional which hinder our ability to determine the temporal effects. In addition to pre- and post-enema differences between our sample types, the DNA recovery could have been different due to swab vs sponge use.

In conclusion, we have shown the feasibility of using rectal swabs pre- and post-enema to characterize the rectal microbiome and capture local host response. The protocol had high acceptability and establishes a platform for reproducible and accurate rectal microbiome sampling in Kenya. The results have shown that there is a high correlation between fecal and mucosal samples, with the fecal samples having a greater diversity and a higher proportion of men classified as *Prevotella-spp* dominant. The *Prevotella- spp* dominant cluster was associated with elevated local and systemic inflammation, regardless of sample type. The mucosal sampling after enema with rectal sponges capture bacterial adherent bacteria without the need for biopsies while providing vital information on gut microflora host interactions.

6. Discussion

6.1 Major findings

In this thesis, we investigated the relationship between the rectal microbiota and immune activation in both the rectum and the blood in a high risk MSM cohort from Nairobi. We found alpha diversity to be decreased in HIV infected MSM and subtle changes in the microbiome structure with regards to beta diversity and gut composition by HIV status. ART initiation likely reduced immune activation, although it did not restore rectal microbiota diversity. We also established that receptive anal sex with multiple partners was associated with lower alpha diversity independent of HIV status. We found limited evidence for elevated markers of MT in HIV infected MSM. We did find MSM to have higher markers of MT than other non-MSM populations and was especially elevated in MSM engaging in receptive anal sex. The feasibility of rectal sampling was well tolerated in our cohort and we show that both fecal and mucosal rectal samples have similar gut microbiome composition although fecal samples have increased diversity. We also found the mucosal sample to strongly associate with systemic immune activation while fecal samples associated strongly with rectal inflammatory markers.

Due to the inherent greater risk of transmission to anal intercourse, MSM have a high HIV burden, but they also experience stigma and have limited access to healthcare particularly in sub-Saharan Africa. HIV infection does major damage to the immune system but it was also found that it alters the gut microbiota until further studies highlighted the confounding effect of MSM status. The interest in gut microbiota studies of MSM started to get elevated once it became apparent that their gut microbiota was distinct from non-MSM. This increased interest though has not translated into an increase in diversity of the studied populations with studies from sub Saharan Africa severely lacking. Our study narrows the gap in study diversity and we were able to replicate findings from western nations. We were also able to show that engaging in receptive anal sex affected the gut diversity and was also associated with higher MT. Our findings give support to the idea that sexual orientation could alter the gut microbiota and this might influence risk to infection or response to therapy although longitudinal studies would be needed to confirm our findings.

The extent to which HIV infection is responsible for altering the rectal microbiota through affecting the host mucosal immune system still remains to be known. Massive damage to the GALT during acute HIV infection has been linked with compromised epithelial integrity, MT and systemic inflammation. Major correlates of progression to AIDS are chronic inflammation and MT which have been associated with increased risk of mortality and comorbidities. ART helps to restore immune function but the damage inflicted in early HIV infection is not completely reversible. There has been intense interest to seek whether there are pathobionts that exacerbate the gut dysbiosis and drive chronic inflammation. There has been a plurality of factors that affect the rectal microbiota, such as geography, diet, antibiotic use, ART regimen, sexual practises among others that can confound results and make it harder to compare studies. The concept of a universal healthy microbiome is not established nor is there a specific HIV induced gut microbiome dysbiosis. One of the major confounding factors in HIV microbiome studies was found to be MSM status, which was characterized with a gut microbiota composition that was distinct from non-MSM and women, regardless of HIV infection. This was demonstrated in two independent cohorts in Europe; MSM had a distinct gut microbiota that was enriched at the genera level for *Prevotella* [125]. Subsequent studies have shown that MSM status actually had a greater impact on the gut microbiota than HIV infection. In a study of 217 individuals from USA, samples clustered based on MSM status compared to non-MSM or women, more so than HIV infection status, based on PCoA analysis of beta diversity [127]. We found similar results in that alpha diversity was significantly reduced in MSM who engage in receptive anal sex with multiple partners to a greater extent than the reduction seen for HIV infection. We did not find a significant association with alpha diversity and insertive anal sex or condom use. We did not find associations between beta diversity, classified as *Prevotella* or *Bacteroides* dominant, with either HIV infection or receptive anal sex. At the genera level, we also found relatively subtle HIV induced gut microbiota alterations. HIV infection was enriched with Several taxa, including *Roseburia, Lachnospira, Streptococcus* and *Granulicatella*. The specific genera changes could be different from other MSM studies due to a myriad of factors that can affect the gut microbiome such as age, diet and genetics. The overall finding of minimal differences on the basis of HIV status has been observed in other MSM studies with some showing little to no effect on alpha diversity [125,127,129]. HIV-microbiome studies that control for MSM studies differ dramatically from previous studies that didn't do so, the latter reporting

significant gut microbiota alterations with a relative increase in the *Prevotella* genera and decrease in the *Bacteroides* genera [121,123]. The novelty of our study is in the ability to provide data from sub Saharan Africa which has sorely been missed in literature especially with context to HIV infection and gut microbiota relationship in the MSM population. Our findings are similar to those showing MSM have a gut microbiota enriched for *Prevotella*, with approximately 86% and 60% classified as belonging in the *Prevotella* cluster in our fecal and mucosal samples respectively which is comparable to the 79% found in the European MSM cohort [125]. Our study also had the important addition of capturing sexual frequency among the MSM population and assessing its impact on the gut microbiota. We found although engaging in receptive anal sex reduced alpha diversity, it did not affect beta diversity with the *Prevotella* dominance in our cohort being pronounced especially in the fecal samples, which is the more common sampling method used by the field. In our study [211] the overall effect of HIV on the gut microbiota was found to be relatively subtle, although we did observe a reduction in alpha diversity which is an indicator associated with poor health outcomes[118,212].

Since it was observed that MSM have a distinct gut microbiota compared to non-MSM, interest has increased with regards to how this could be related to mucosal immunity and subsequently HIV transmission or disease progression. MSM who are HIV negative had elevated endotoxemia, lower CD4/CD8 ratio, and increased plasma cytokine levels compared to age and ethnically matched low risk control male samples [183]. In that study there was a subset within the high risk MSM cohort who had detectable LPS in plasma which correlated with significantly elevated plasma pro-inflammatory biomarker levels [183]. One of potential factors that can explain the variability in immune activation within MSM is the heterogeneous sexual practises within the population. MSM who engaged in condomless receptive anal sex had a gut environment that had an inflammatory environment composed of increased levels of inflammatory cytokines, Th17 cells, and upregulation of genes involved in mucosal injury [133]. We also explored whether immune activation and markers of MT were associated with sexual behaviour, controlling for HIV infection. Although we found little evidence to support the consensus finding of HIV infection having increased markers of MT [84,185], our MSM cohort did have elevated markers of MT compared to other low risk groups and is similar to studies carried out from the region [180,189]. However, engaging in receptive anal sex was associated with increased markers of MT

independent of HIV infection. We did not find a significant association of markers of MT and alpha diversity metrics. Markers of MT were also associated with systemic inflammation markers but limited association with rectal inflammatory biomarkers. The increased markers of MT in those who engage in receptive anal sex could be due to physical trauma at the rectal mucosa and this could make them susceptible to HIV infection as receptive anal sex has the highest risk of HIV transmission compared to other sexual practises.

Although we did not find profound evidence of systemic or rectal mucosa inflammation in MSM who engaged in receptive anal sex, we did find reduction in alpha diversity and evidence of MT. This does support the hypothesis that MSM linked gut microbiota could contribute to an inflammatory immune phenotype that can potentially affect HIV susceptibility. It has been shown in SIV models that macaques with distinct gut microbiomes have different rates of infection that were linked to increase in HIV target cells in the susceptible group[162]. Studies have shown MSM to have elevated systemic immune activation and MSM-associated gut microbes *in vitro* elicited an elevated immune response and enhanced HIV infection compared to gut microbiota from non-MSM [126,128]. The reduction in gut alpha diversity and elevated markers of MT in MSM engaging in receptive anal sex could translate to affecting the mucosal immunity and potentially creating an inflamed rectal environment. Although markers of MT was not associated to alpha diversity and rectal inflammatory biomarkers, that could also show a higher baseline of markers of MT that is tolerated. The potential mechanism through which gut microbiota alterations could impact mucosal immunity and potentially affect HIV susceptibility, and any interventions that could result from these mechanisms, remain a focus of ongoing research.

This thesis also compared sampling methods for gut microbiome characterization, and how each method could explain host immune phenotypes. Although fecal sampling is a standard method in the field, rectal swabs also present a convenient alternative for certain types of study designs. Our focus was on evaluating the impact of an enema on rectal microbiome phenotypes. We hypothesized that pre-enema samples will resemble fecal samples while the post-enema samples will reflect mucosal adherent bacteria that were not washed away. The mucosal samples would thus more closely resemble the immune-interacting gut microbiota than fecal samples. We found

the fecal sample to have a *Prevotella* dominant gut microbiota and also richer diversity compared to mucosal samples (84% vs 43% respectively). We also noted that the trend to switch classification was only from *Prevotella* dominant to *Bacteroides* dominant. These results suggest that the *Prevotella* dominant cluster is affected by the enema and it potentially has less mucosal adherent genera than the *Bacteroides* dominant cluster. We also found relatively few gut microbiota compositional shifts between the fecal and mucosal samples with most of the enriched genera in the fecal samples being anaerobes. Although we did see the gut diversity affected by the sampling method, the associations with systemic and rectal immune markers of inflammation remained similar although there was difference in effect size by sampling method. The *Prevotella* cluster was in both sampling methods associated with elevated systemic and rectal immune activation. There was a trend for robust clustering of mucosal samples with systemic T cell immune activation in comparison to fecal samples. Previous research has shown the mucosal adherent gut microbiota to be more stable along the GI tract and also different from fecal samples [201,202]. This stability is explained by the similarity of the gut microbiota along the digestive tract and is thought to be more important in gut microbiota associations with the mucosal immunology and risk of disease. Our results support the idea that rectal sampling after enema could capture mucosal adherent bacteria without the need to extract biopsies. Although our results need to be confirmed in larger studies, this would provide researchers with a valuable tool that has plenty of advantages as it relates to collecting rectal swabs rather than invasive biopsies. Although it is highly unlikely that stool samples will stop being collected from gut microbiome research, it would be advisable to be cautious to draw assumptions from fecal samples as representative of mucosal samples.

There remain challenges in the microbiome field with regards to study design, data collection, and analysis. Due to cost associated with high throughput sequencing, the sample size of studies in the field tend to be small. Standardization of studies would significantly help in reproducibility and study comparisons. There are also different sequencing methods such as 16s rRNA, chaperonin 60 (CPN60) and shotgun sequencing which brings with it variability. Data analysis pipelines also can differ between studies with microbiome data being full of zeros, which add to the complexity of handling them. Different normalization methods are used in the field such as rarefaction, relative abundance, square root transformation and centre log

transformation. New analysis tools are also added constantly to the tool set, thus it's important to compare different tools and use that to measure weight of evidence when reporting results.

6.2 Hypothesis and conclusion

The major hypothesis of this thesis were;
1. HIV infected men will have altered microbiome structure that results in reduced alpha diversity, but ART therapy will be associated with increased diversity. Men who engage in different sexual behaviours will have, independent of HIV status, a distinct microbiome structure.
2. HIV infected men will have a higher level of MT than HIV negative men which will be associated with elevated systemic inflammation. MSM will have a higher background of MT compared to non-MSM populations.
3. Mucosal adherent samples will have less bacterial community diversity than fecal samples. Adherent bacteria communities in the mucosal samples will have stronger associations with local immunity than lumen bacterial communities.

Consistent with our first hypothesis, we found HIV infection MSM to have lower alpha diversity that was not restored by ART. We also found an even stronger effect on alpha diversity by MSM who engaged in receptive anal sex with multiple partner types. We did not find notable differences in beta diversity with regards to HIV infection or other high risk behaviours.
In contrast to our second thesis, we found limited evidence of markers of MT in HIV infected MSM. We did see though that MSM have higher amount of markers of MT compared to low risk samples. We also found engaging in receptive anal sex to be associated with markers of MT although limited link to gut diversity or immune activation.
Consistent with our third hypothesis, we found mucosal samples to have lower diversity than fecal samples. We also found the mucosal samples to have a more robust association with systemic immune activation than fecal samples although few associations with rectal inflammatory cytokines.

6.3 Future directions

Although many studies have been done on HIV throughout the infection phase, there is still a gap in studying pre infection predictors of HIV susceptibility. Potential future direction of this study is to conduct a longitudinal study with HIV as a biological endpoint to investigate immune correlates of protection. Correlates of HIV protection have been published in plenty of cohorts including in Kenyan female sex workers (FSW) in HIV exposed seronegative individuals (HESN) who are individuals highly exposed to HIV but are persistently seronegative [213–217]. Factors identified to be protective against HIV-1 in the Kenyan FSW cohort include reduced cytokine and chemokine production, low level of T cell activation, reduced IRF-1 expression, elevated Tregs and elevated antiprotease. These factors are thought to contribute to mediating innate immune response to reduce HIV replication but also maintain T cell immune quiescence to reduce target cells availability [218–220]. On the other hand increased levels of T cell activation means more target cells available for HIV infection. Early Microbicide trials were not effective in preventing HIV infection but were found to increase risk of HIV acquisition. The female genital tract's (FGT) tight junctions between epithelial cells were found to be impaired and inflammation enhanced by the microbicides cellulose sulphate and nonoxynol-9, [221]. It is thought that this inflammation increased target cell availability and increased risk of HIV acquisition [222]. The highest efficacious microcode to data was the 1% tenofovir gel with an efficacy of 39% [38]. The trial results showed that immune activation prior to exposure was independently linked to increase risk of HIV acquisition irrespective of prevention arm (microbicide or placebo) [223].
In theory, this could mean that there could be heterogeneity in infection probability among MSM who are at high risk of infection. A subset of individuals highly susceptible to HIV infection within the high risk MSM population could exhibit increased baseline immune activation, mucosal inflammation and gut microbiota disturbances. There has been only one study that was done to look at data from the early HIV pandemic retrospectively in MSM before HIV seroconversion. The study characterized the gut microbiota from fecal samples and systemic inflammation from plasma samples. They found significant gut microbiota composition differences at the genera level between HIV seroconverters and HIV negative controls prior to HIV infection. The seroconverters also had increased plasma inflammatory markers compared to HIV negative controls [224]. The study though was not able to link directly gut microbial

composition changes with inflammation but it is the first to show that variation in gut microbiota composition could affect HIV susceptibility. A well designed study looking at gut microbiota influence on HIV susceptibility would also quantify local inflammation and also control for known gut microbiome factors. The findings could have clinically implications in that targeted therapeutic interventions could reduce HIV acquisition.

Data integration with the components of the gut environment might also give additional insight into their interactions. The introduction of 'omics has given access to plenty of data that can be used to elucidate relationships between rectal microbiota composition and metabolic changes. Shotgun sequencing allows for identification of genes involved in metabolism pathways. The field is currently dominated by descriptive studies are limited in their ability to control various factors that can affect the microbiome and metabolome. Longitudinal studies would be able to asses the persistence of changes over time and also give more predictive value rather than associations.

6.4 Concluding remarks

This thesis sought to understand how sexual orientation and HIV infection affect the gut microbiota and its relationship to immune activation. Previous studies in western countries had shown MSM status to have distinct gut microbiota from non-MSM although is not known what drives these differences. We have shown that mucosal samples generated from rectal sponges after enema to be able to capture mucosal adherent bacteria and are distinct from fecal samples and important for capturing host physiology interactions. We found both HIV infection and engaging in receptive anal sex to be associated with reduced alpha diversity but modest effect on beta diversity. Although we found limited evidence of markers of MT in HIV infected MSM, we did find engaging in receptive anal sex to be also associated with elevated MSM that was associated with CD+ T cell activation. These findings suggest that gut alterations and markers of MT are linked to the type of sexual behaviour in MSM individuals. These findings should contribute to further research on how potentially distinct gut microbiota within this high risk cohort can affect HIV susceptibility.

7. References

1. Rutherford, G. W. *et al.* The Epidemiology of AIDS-Related Kaposi's Sarcoma in San Francisco. *J. Infect. Dis.* **159**, 569–572 (1989).
2. Green, D. *et al.* Demographic and risk group heterogeneity across the UNAIDS 90-90-90 targets: A systematic review and meta-analysis protocol. *Syst. Rev.* **8**, 1–7 (2019).
3. James, S. L. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1789–1858 (2018).
4. Beyrer, C. *et al.* The increase in global HIV epidemics in MSM. *AIDS* **27**, 2665–2678 (2013).
5. Makofane, K., Elst, E. M. van der, Walimbwa, J., Nemande, S. & Baral, S. D. From general to specific: moving past the general population in the HIV response across sub-Saharan Africa. *J. Int. AIDS Soc.* **23**, (2020).
6. Bain, L. E., Nkoke, C. & Noubiap, J. J. N. UNAIDS 90–90–90 targets to end the AIDS epidemic by 2020 are not realistic: comment on "Can the UNAIDS 90–90–90 target be achieved? A systematic analysis of national HIV treatment cascades". *BMJ Glob. Heal.* **2**, e000227 (2017).
7. Wolfe, N. D., Dunavan, C. P. & Diamond, J. Origins of major human infectious diseases. *Nat. 2007 4477142* **447**, 279–283 (2007).
8. Marlink, R. *et al.* Reduced rate of disease development after HIV-2 infection as compared to HIV-1. *Science (80-.).* **265**, 1587–1590 (1994).
9. AF, S. *et al.* Characterization of a new circulating recombinant form comprising HIV-1 subtypes C and B in southern Brazil. *AIDS* **20**, 2159–2164 (2006).
10. Sharma, G., Kaur, G. & Mehra, N. Genetic correlates influencing immunopathogenesis of HIV infection. *Indian J. Med. Res.* **134**, 749 (2011).
11. Bleul, C. C., Wu, L., Hoxie, J. A., Springer, T. A. & Mackay, C. R. The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 1925 (1997).
12. Freed, E. O. HIV-1 assembly, release and maturation. *Nat. Rev. Microbiol.* **13**, 484–497 (2015).
13. Okoye, A. A. & Picker, L. J. CD4+ T-cell depletion in HIV infection: mechanisms of immunological failure. *Immunol. Rev.* **254**, 54–64 (2013).
14. Buvé, A., Bishikwabo-Nsarhaza, K. & Mutangadura, G. The spread and effect of HIV-1 infection in sub-Saharan Africa. *Lancet* vol. 359 2011–2017 (2002).
15. Patel, P. *et al.* Estimating per-act HIV transmission risk: A systematic review. *AIDS* **28**, 1509–1519 (2014).
16. Galvin, S. R. & Cohen, M. S. The role of sexually transmitted diseases in HIV transmission. *Nat. Rev. Microbiol. 2004 21* **2**, 33–42 (2004).
17. Powers, K. A., Poole, C., Pettifor, A. E. & Cohen, M. S. Rethinking the heterosexual infectivity of HIV-1: a systematic review and meta-analysis. *Lancet Infect. Dis.* **8**, 553–563 (2008).
18. Gray, R. H. *et al.* The effectiveness of male circumcision for HIV prevention and effects on risk behaviors in a post-trial follow up study in Rakai, Uganda. *AIDS* **26**, 609 (2012).
19. Blanco, J. L., Whitlock, G., Milinkovic, A. & Moyle, G. HIV integrase inhibitors: a new

era in the treatment of HIV. *http://dx.doi.org.uml.idm.oclc.org/10.1517/14656566.2015.1044436* **16**, 1313–1324 (2015).
20. Lancet, A. T. C. C.-T. & 2008, undefined. Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies. *Elsevier*.
21. Vittinghoff, E. *et al.* Combination antiretroviral therapy and recent declines in AIDS incidence and mortality. *J. Infect. Dis.* **179**, 717–720 (1999).
22. Palella, F. J. *et al.* Declining Morbidity and Mortality among Patients with Advanced Human Immunodeficiency Virus Infection. *N. Engl. J. Med.* **338**, 853–860 (1998).
23. Mocroft, A. *et al.* Changing patterns of mortality across Europe in patients infected with HIV-1. *Lancet* **352**, 1725–1730 (1998).
24. Yeni, P. G. *et al.* Antiretroviral Treatment for Adult HIV Infection in 2002: Updated Recommendations of the International AIDS Society-USA Panel. *JAMA* **288**, 222–235 (2002).
25. Kitahata, M. M. *et al.* Effect of Early versus Deferred Antiretroviral Therapy for HIV on Survival. *N. Engl. J. Med.* **360**, 1815–1826 (2009).
26. Colby, D. *et al.* Rapid HIV RNA rebound after antiretroviral treatment interruption in persons durably suppressed in Fiebig I acute HIV infection. *Nat. Med.* **24**, 923 (2018).
27. SH, D. & EB, C. Cabotegravir-Rilpivirine: The First Complete Long-Acting Injectable Regimen for the Treatment of HIV-1 Infection. *Ann. Pharmacother.* **55**, (2021).
28. Frank, T. D. *et al.* Global, regional, and national incidence, prevalence, and mortality of HIV, 1980–2017, and forecasts to 2030, for 195 countries and territories: a systematic analysis for the Global Burden of Diseases, Injuries, and Risk Factors Study 2017. *Lancet HIV* **6**, e831–e859 (2019).
29. Auvert, B. *et al.* Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: The ANRS 1265 trial. *PLoS Med.* **2**, 1112–1122 (2005).
30. Weller, S. & Davis, K. Condom effectiveness in reducing heterosexual HIV transmission. *Cochrane database Syst. Rev.* CD003255 (2001) doi:10.1002/14651858.CD003255.
31. Baeten, J. M. *et al.* Genital HIV-1 RNA predicts risk of heterosexual HIV-1 transmission. *Sci. Transl. Med.* **3**, 77ra29-77ra29 (2011).
32. Attia, S., Egger, M., Müller, M., Zwahlen, M. & Low, N. Sexual transmission of HIV according to viral load and antiretroviral therapy: Systematic review and meta-analysis. *AIDS* **23**, 1397–1404 (2009).
33. Graham, S. M. *et al.* Initiation of antiretroviral therapy leads to a rapid decline in cervical and vaginal HIV-1 shedding. *AIDS* **21**, 501–507 (2007).
34. Quinn, T. C. *et al.* Viral Load and Heterosexual Transmission of Human Immunodeficiency Virus Type 1. *N. Engl. J. Med.* **342**, 921–929 (2000).
35. Cohen, M. S. *et al.* Antiretroviral Therapy for the Prevention of HIV-1 Transmission. *https://doi.org/10.1056/NEJMoa1600693* **375**, 830–839 (2016).
36. Group, T. I. S. S. Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection. *N. Engl. J. Med.* **373**, 795 (2015).
37. Cohen, M. S. *et al.* Prevention of HIV-1 Infection with Early Antiretroviral Therapy. *N. Engl. J. Med.* **365**, 493–505 (2011).
38. Karim, Q. A. *et al.* Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. *Science (80-.).* **329**, 1168–1174 (2010).

39. Grant, R. M. *et al.* Preexposure Chemoprophylaxis for HIV Prevention in Men Who Have Sex with Men. *N. Engl. J. Med.* **363**, 2587–2599 (2010).
40. Karim, S. S. A., Kashuba, A. D., Werner, L. & Karim, Q. A. Drug concentrations after topical and oral antiretroviral pre-exposure prophylaxis: implications for HIV prevention in women. *Lancet* **378**, 279–281 (2011).
41. Baeten, J. M. *et al.* Antiretroviral Prophylaxis for HIV Prevention in Heterosexual Men and Women. *N. Engl. J. Med.* **367**, 399–410 (2012).
42. Haberer, J. E. Current Concepts for PrEP Adherence: In The PrEP revolution; from clinical trials to routine practice. *Curr. Opin. HIV AIDS* **11**, 10 (2016).
43. Donnell, D. *et al.* HIV protective efficacy and correlates of tenofovir blood concentrations in a clinical trial of PrEP for HIV prevention. *J. Acquir. Immune Defic. Syndr.* **66**, 340–348 (2014).
44. Murnane, P. M. *et al.* Efficacy of preexposure prophylaxis for HIV-1 prevention among high-risk heterosexuals: Subgroup analyses from a randomized trial. *AIDS* **27**, 2155–2160 (2013).
45. Anderson, P. L. *et al.* Emtricitabine-tenofovir concentrations and pre-exposure prophylaxis efficacy in men who have sex with men. *Sci. Transl. Med.* **4**, (2012).
46. Nel, A. *et al.* A safety and pharmacokinetic trial assessing delivery of dapivirine from a vaginal ring in healthy women. *AIDS* **28**, 1479–1487 (2014).
47. Spreen, W. *et al.* Pharmacokinetics, safety, and tolerability with repeat doses of GSK1265744 and rilpivirine (TMC278) long-acting nanosuspensions in healthy adults. in *Journal of Acquired Immune Deficiency Syndromes* vol. 67 487–492 (Lippincott Williams and Wilkins, 2014).
48. Andrews, C. D. *et al.* Long-acting integrase inhibitor protects macaques from intrarectal simian/human immunodeficiency virus. *Science (80-.).* **343**, 1151–1154 (2014).
49. Ahmed, K. *et al.* HIV incidence among women using intramuscular depot medroxyprogesterone acetate, a copper intrauterine device, or a levonorgestrel implant for contraception: a randomised, multicentre, open-label trial. *Lancet* **394**, 303–313 (2019).
50. Smith, J. A., Garnett, G. P. & Hallett, T. B. The Potential Impact of Long-Acting Cabotegravir for HIV Prevention in South Africa: A Mathematical Modeling Study. *J. Infect. Dis.* (2020) doi:10.1093/INFDIS/JIAA296.
51. CC, T. *et al.* Effectiveness of postinoculation (R)-9-(2-phosphonylmethoxypropyl) adenine treatment for prevention of persistent simian immunodeficiency virus SIVmne infection depends critically on timing of initiation and duration of treatment. *J. Virol.* **72**, 4265–4273 (1998).
52. DM, C. *et al.* A case-control study of HIV seroconversion in health care workers after percutaneous exposure. Centers for Disease Control and Prevention Needlestick Surveillance Group. *N. Engl. J. Med.* **337**, 1485–1490 (1997).
53. Turk, G. *et al.* A Possible Sterilizing Cure of HIV-1 Infection Without Stem Cell Transplantation. *Ann. Intern. Med.* **175**, 95–100 (2022).
54. S, R.-N. *et al.* Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N. Engl. J. Med.* **361**, 2209–2220 (2009).
55. Gray, G. E. *et al.* Vaccine Efficacy of ALVAC-HIV and Bivalent Subtype C gp120–MF59 in Adults. *N. Engl. J. Med.* **384**, 1089–1100 (2021).
56. Moldt, B. *et al.* Highly potent HIV-specific antibody neutralization in vitro translates into effective protection against mucosal SHIV challenge in vivo. *Proc. Natl. Acad. Sci.* **109**,

18921–18925 (2012).
57. Sok, D. & Burton, D. R. Recent progress in broadly neutralizing antibodies to HIV. *Nat. Immunol.* **19**, 1179–1188 (2018).
58. Fitzgerald, K. A. & Kagan, J. C. Toll-like Receptors and the Control of Immunity. *Cell* **180**, 1044–1066 (2020).
59. Vivier, E., Tomasello, E., Baratin, M., Walzer, T. & Ugolini, S. Functions of natural killer cells. *Nat. Immunol. 2008 95* **9**, 503–510 (2008).
60. Barber, G. N. Host defense, viruses and apoptosis. *Cell Death Differ. 2001 82* **8**, 113–126 (2001).
61. Boasso, A., Shearer, G. M. & Chougnet, C. Immune dysregulation in human immunodeficiency virus infection: know it, fix it, prevent it? *J. Intern. Med.* **265**, 78–96 (2009).
62. Zygmunt, B. & Veldhoen, M. T Helper Cell Differentiation: More than Just Cytokines. *Adv. Immunol.* **109**, 159–196 (2011).
63. Torchinsky, M. B., Garaude, J., Martin, A. P. & Blander, J. M. Innate immune recognition of infected apoptotic cells directs TH17 cell differentiation. *Nat. 2009 4587234* **458**, 78–82 (2009).
64. Yasuda, K., Takeuchi, Y. & Hirota, K. The pathogenicity of Th17 cells in autoimmune diseases. *Semin. Immunopathol.* **41**, 283–297 (2019).
65. Ivanov, I. I. *et al.* Induction of Intestinal Th17 Cells by Segmented Filamentous Bacteria. *Cell* **139**, 485–498 (2009).
66. Omenetti, S. & Pizarro, T. T. The Treg/Th17 axis: A dynamic balance regulated by the gut microbiome. *Front. Immunol.* **6**, 639 (2015).
67. Brenchley, J. M. *et al.* CD4+ T Cell Depletion during all Stages of HIV Disease Occurs Predominantly in the Gastrointestinal Tract. *J. Exp. Med.* **200**, 749–759 (2004).
68. Veazey, R. S. *et al.* Gastrointestinal Tract as a Major Site of CD4+ T Cell Depletion and Viral Replication in SIV Infection. *Science (80-.).* **280**, 427–431 (1998).
69. JO, K. & BD, W. Acute human immunodeficiency virus type 1 infection. *N. Engl. J. Med.* **339**, 33–39 (1998).
70. Daar, E. S., Pilcher, C. D. & Hecht, F. M. Clinical presentation and diagnosis of primary HIV-1 infection. *Current Opinion in HIV and AIDS* vol. 3 10–15 (2008).
71. Schacker, T. W., Hughes, J. P., Shea, T., Coombs, R. W. & Corey, L. Biological and virologic characteristics of primary HIV infection. *Ann. Intern. Med.* **128**, 613–620 (1998).
72. Ho, D. D. *et al.* Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nat. 1995 3736510* **373**, 123–126 (1995).
73. Stacey, A. R. *et al.* Induction of a Striking Systemic Cytokine Cascade prior to Peak Viremia in Acute Human Immunodeficiency Virus Type 1 Infection, in Contrast to More Modest and Delayed Responses in Acute Hepatitis B and C Virus Infections. *J. Virol.* **83**, 3719–3733 (2009).
74. Killian, M. S., Fujimura, S. H., Hecht, F. M. & Levy, J. A. Similar changes in plasmacytoid dendritic cell and CD4 T-cell counts during primary HIV-1 infection and treatment. *AIDS* **20**, 1247–1252 (2006).
75. McMichael, A. J., Borrow, P., Tomaras, G. D., Goonetilleke, N. & Haynes, B. F. The immune response during acute HIV-1 infection: clues for vaccine development. *Nat. Rev. Immunol. 2009 101* **10**, 11–23 (2009).

76. Matano, T. *et al.* Administration of an Anti-CD8 Monoclonal Antibody Interferes with the Clearance of Chimeric Simian/Human Immunodeficiency Virus during Primary Infections of Rhesus Macaques. *J. Virol.* **72**, 164 (1998).
77. McBrien, J. B., Kumar, N. A. & Silvestri, G. Mechanisms of CD8+ T cell-mediated suppression of HIV/SIV replication. *Eur. J. Immunol.* **48**, 898 (2018).
78. Goonetilleke, N. *et al.* The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. *J. Exp. Med.* **206**, 1253–1272 (2009).
79. Perdomo-Celis, F., Taborda, N. A. & Rugeles, M. T. CD8+ T-cell response to HIV infection in the era of antiretroviral therapy. *Front. Immunol.* **10**, 1896 (2019).
80. Zhang, Z. Q. *et al.* Early depletion of proliferating B cells of germinal center in rapidly progressive simian immunodeficiency virus infection. *Virology* **361**, 455–464 (2007).
81. Klatt, N. R., Chomont, N., Douek, D. C. & Deeks, S. G. Immune activation and HIV persistence: implications for curative approaches to HIV infection. *Immunol. Rev.* **254**, 326–342 (2013).
82. Zicari, S. *et al.* Immune Activation, Inflammation, and Non-AIDS Co-Morbidities in HIV-Infected Patients under Long-Term ART. *Viruses* **11**, (2019).
83. Estes, J. D. *et al.* Damaged Intestinal Epithelial Integrity Linked to Microbial Translocation in Pathogenic Simian Immunodeficiency Virus Infections. *PLOS Pathog.* **6**, e1001052 (2010).
84. Brenchley, J. M. *et al.* Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat. Med.* **12**, 1365–1371 (2006).
85. Raffatellu, M. *et al.* Simian immunodeficiency virus–induced mucosal interleukin-17 deficiency promotes Salmonella dissemination from the gut. *Nat. Med. 2008 144* **14**, 421–428 (2008).
86. Brenchley, J. M. *et al.* Differential Th17 CD4 T-cell depletion in pathogenic and nonpathogenic lentiviral infections. *Blood* **112**, 2826–2835 (2008).
87. Marchetti, G. *et al.* Microbial translocation predicts disease progression of HIV-infected antiretroviral-naive patients with high CD4+ cell count. *AIDS* **25**, 1385–1394 (2011).
88. Moore, W. E. C. & Holdeman, L. V. Human Fecal Flora: The Normal Flora of 20 Japanese-Hawaiians. *Appl. Microbiol.* **27**, 961 (1974).
89. Frank, D. N. & Pace, N. R. Gastrointestinal microbiology enters the metagenomics era. *Current Opinion in Gastroenterology* vol. 24 4–10 (2008).
90. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
91. Huttenhower, C. *et al.* Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
92. Sekirov, I., Russell, S. L., Caetano M Antunes, L. & Finlay, B. B. Gut microbiota in health and disease. *Physiological Reviews* vol. 90 859–904 (2010).
93. Jakobsson, H. E. *et al.* Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by caesarean section. *Gut* **63**, 559–566 (2014).
94. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* vol. 486 222–227 (2012).
95. Agans, R. *et al.* Distal gut microbiota of adolescent children is different from that of adults. *FEMS Microbiol. Ecol.* **77**, 404–412 (2011).
96. Sommer, F. & Bäckhed, F. The gut microbiota-masters of host development and

physiology. *Nature Reviews Microbiology* vol. 11 227–238 (2013).
97. Shanahan, F. The host-microbe interface within the gut. *Bailliere's Best Pract. Res. Clin. Gastroenterol.* **16**, 915–931 (2002).
98. Mazmanian, S. K., Cui, H. L., Tzianabos, A. O. & Kasper, D. L. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **122**, 107–118 (2005).
99. O'Hara, A. M. & Shanahan, F. The gut flora as a forgotten organ. *EMBO Reports* vol. 7 688–693 (2006).
100. Macpherson, A. J. & Harris, N. L. Interactions between commensal intestinal bacteria and the immune system. *Nat. Rev. Immunol. 2004 46* **4**, 478–485 (2004).
101. Wu, H. J. *et al.* Gut-residing segmented filamentous bacteria drive autoimmune arthritis via T helper 17 cells. *Immunity* **32**, 815–827 (2010).
102. Ivanov, I. I. *et al.* Specific Microbiota Direct the Differentiation of IL-17-Producing T-Helper Cells in the Mucosa of the Small Intestine. *Cell Host Microbe* **4**, 337–349 (2008).
103. Alakomi, H. L. *et al.* Lactic acid permeabilizes gram-negative bacteria by disrupting the outer membrane. *Appl. Environ. Microbiol.* **66**, 2001–2005 (2000).
104. Gilbert, J. A. *et al.* Current understanding of the human microbiome. *Nat. Med.* **24**, 392 (2018).
105. De Filippo, C. *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 14691–14696 (2010).
106. Lax, S. *et al.* Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science (80-.).* **345**, 1048–1052 (2014).
107. Goodrich, J. K. *et al.* Human Genetics Shape the Gut Microbiome. *Cell* **159**, 789–799 (2014).
108. Rothschild, D. *et al.* Environment dominates over host genetics in shaping human gut microbiota. *Nature* **555**, 210–210 (2018).
109. Petrof, E. O. *et al.* Probiotics inhibit nuclear factor-κB and induce heat shock proteins in colonic epithelial cells through proteasome inhibition. *Gastroenterology* **127**, 1474–1487 (2004).
110. Kelly, D. *et al.* Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-γ and RelA. *Nat. Immunol. 2004 51* **5**, 104–112 (2003).
111. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nat. 2006 4447122* **444**, 1027–1031 (2006).
112. Morgan, X. C. *et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol.* **13**, 1–18 (2012).
113. Tang, W. H. W. *et al.* Intestinal Microbial Metabolism of Phosphatidylcholine and Cardiovascular Risk. *http://dx.doi.org/10.1056/NEJMoa1109400* **368**, 1575–1584 (2013).
114. Livanos, A. E. *et al.* Antibiotic-mediated gut microbiome perturbation accelerates development of type 1 diabetes in mice. *Nat. Microbiol. 2016 111* **1**, 1–13 (2016).
115. Ley, R. E. *et al.* Obesity alters gut microbial ecology. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 11070–11075 (2005).
116. Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nat. 2011 4727341* **472**, 57–63 (2011).
117. Koeth, R. A. *et al.* Intestinal microbiota metabolism of l -carnitine, a nutrient in red meat,

promotes atherosclerosis. *Nat. Med. 2013 195* **19**, 576–585 (2013).
118. Turnbaugh, P. J. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457**, 480 (2009).
119. Astbury, S. *et al.* Lower gut microbiome diversity and higher abundance of proinflammatory genus Collinsella are associated with biopsy-proven nonalcoholic steatohepatitis. *Gut Microbes* **11**, 569–580 (2020).
120. SA, T. *et al.* The Impact of Human Immunodeficiency Virus Infection on Gut Microbiota α-Diversity: An Individual-level Meta-analysis. *Clin. Infect. Dis.* **70**, 615–627 (2020).
121. Vázquez-Castellanos, J. F. *et al.* Altered metabolism of gut microbiota contributes to chronic immune activation in HIV-infected individuals. *Mucosal Immunol.* **8**, 760–772 (2015).
122. Lu, W. *et al.* Association between gut microbiota and CD4 recovery in HIV-1 infected patients. *Front. Microbiol.* **9**, (2018).
123. Mutlu, E. A. *et al.* A Compositional Look at the Human Gastrointestinal Microbiome and Immune Activation Parameters in HIV Infected Subjects. *PLoS Pathog.* **10**, (2014).
124. Lozupone, C. A. *et al.* Alterations in the gut microbiota associated with HIV-1 infection. *Cell Host Microbe* **14**, 329–339 (2013).
125. Noguera-Julian, M. *et al.* Gut Microbiota Linked to Sexual Preference and HIV Infection. *EBioMedicine* **5**, 135–146 (2016).
126. Neff, C. P. *et al.* Fecal Microbiota Composition Drives Immune Activation in HIV-infected Individuals. *EBioMedicine* **30**, 192–202 (2018).
127. Armstrong, A. J. S. *et al.* An exploration of Prevotella-rich microbiomes in HIV and men who have sex with men. *Microbiome* **6**, (2018).
128. Li, S. X. *et al.* Gut microbiota from high-risk men who have sex with men drive immune activation in gnotobiotic mice and in vitro HIV infection. *PLoS Pathog.* **15**, (2019).
129. Nowak, R. G. *et al.* Rectal microbiota among HIV-uninfected, untreated HIV, and treated HIV-infected men who have sex with men (MSM) in Nigeria. *AIDS* **31**, 857 (2017).
130. Yu, G., Fadrosh, D., Ma, B., Ravel, J. & Goedert, J. J. Anal microbiota profiles in HIV-positive and HIV-negative MSM. *AIDS* **28**, 753–760 (2014).
131. Lee, S. C. *et al.* Enrichment of gut-derived Fusobacterium is associated with suboptimal immune recovery in HIV-infected individuals. *Sci. Rep.* **8**, (2018).
132. Guillén, Y. *et al.* Low nadir CD4+ T-cell counts predict gut dysbiosis in HIV-1 infection. *Mucosal Immunol.* **12**, 232–246 (2019).
133. Kelley, C. F. *et al.* The rectal mucosa and condomless receptive anal intercourse in HIV-negative MSM: Implications for HIV transmission and prevention. *Mucosal Immunol.* **10**, 996–1007 (2017).
134. Kehrmann, J. *et al.* Gut Microbiota in Human Immunodeficiency Virus-Infected Individuals Linked to Coronary Heart Disease. *J. Infect. Dis.* **219**, 497–508 (2019).
135. Hensley-McBain, T. *et al.* Increased mucosal neutrophil survival is associated with altered microbiota in HIV infection. *PLoS Pathog.* **15**, (2019).
136. Hansen, M. E. B. *et al.* Population structure of human gut bacteria in a diverse cohort from rural Tanzania and Botswana. *Genome Biol.* **20**, (2019).
137. De Filippo, C. *et al.* Diet, environments, and gut microbiota. A preliminary investigation in children living in rural and Urban Burkina Faso and Italy. *Front. Microbiol.* **8**, (2017).
138. Sanders, E. J. *et al.* High HIV-1 incidence, correlates of HIV-1 acquisition, and high viral loads following seroconversion among MSM. *AIDS* **27**, 437–446 (2013).

139. Price, M. A. *et al.* Identifying at-risk populations in kenya and south africa: HIV incidence in cohorts of menwho report sex with men, sex workers, and youth. *J. Acquir. Immune Defic. Syndr.* **59**, 185–193 (2012).
140. Smith, A. D., Tapsoba, P., Peshu, N., Sanders, E. J. & Jaffe, H. W. Men who have sex with men and HIV/AIDS in sub-Saharan Africa. *Lancet* **374**, 416–422 (2009).
141. Sanders, E. J. *et al.* HIV-1 infection in high risk men who have sex with men in Mombasa, Kenya. *AIDS* **21**, 2513–2520 (2007).
142. McKinnon, L. R. *et al.* High HIV risk in a cohort of male sex workers from Nairobi, Kenya. *Sex. Transm. Infect.* **90**, 237–242 (2014).
143. Muraguri, N. *et al.* HIV and STI Prevalence and Risk Factors Among Male Sex Workers and Other Men Who Have Sex With Men in Nairobi, Kenya. *J. Acquir. Immune Defic. Syndr.* **68**, 91 (2015).
144. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* vol. 7 335–336 (2010).
145. Kassambara, A., The, F. M.-E. and visualize & 2017, U. *Package 'factoextra' Type Package Title Extract and Visualize the Results of Multivariate Data Analyses*. *Mran.Revolutionanalytics.Com* https://github.com/kassambara/factoextra/issues (2016).
146. Husson, F., Josse, J., Le, S. & Maintainer, J. M. *Package 'FactoMineR' Title Multivariate Exploratory Data Analysis and Data Mining*. http://factominer.free.fr (2020).
147. Chong, J. *et al.* MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* **46**, W486–W494 (2018).
148. Maynard, C. L., Elson, C. O., Hatton, R. D. & Weaver, C. T. Reciprocal interactions of the intestinal microbiota and immune system. *Nature* vol. 489 231–241 (2012).
149. Dillon, S. M. *et al.* An altered intestinal mucosal microbiome in HIV-1 infection is associated with mucosal and systemic immune activation and endotoxemia. *Mucosal Immunol.* **7**, 983–994 (2014).
150. Dieffenbach, C. W. & Fauci, A. S. Thirty years of HIV and AIDS: Future challenges and opportunities. *Ann. Intern. Med.* **154**, 766–771 (2011).
151. Coleman, S. L. *et al.* Can gut microbiota of men who have sex with men influence HIV transmission? *Gut Microbes* **11**, 610–619 (2020).
152. Tuddenham, S. A. *et al.* The Impact of Human Immunodeficiency Virus Infection on Gut Microbiota α-Diversity: An Individual-level Meta-analysis. *Clin. Infect. Dis.* **70**, 615–627 (2020).
153. Monaco, C. L. *et al.* Altered Virome and Bacterial Microbiome in Human Immunodeficiency Virus-Associated Acquired Immunodeficiency Syndrome. *Cell Host Microbe* **19**, 311–322 (2016).
154. Ceccarani, C. *et al.* Rectal Microbiota Associated With Chlamydia trachomatis and Neisseria gonorrhoeae Infections in Men Having Sex With Other Men. *Front. Cell. Infect. Microbiol.* **9**, (2019).
155. Albenberg, L. G. & Wu, G. D. Diet and the intestinal microbiome: Associations, functions, and implications for health and disease. *Gastroenterology* **146**, 1564–1572 (2014).
156. Bolger, A. M., Lohse, M. & Usadel, B. Genome analysis Trimmomatic: a flexible trimmer for Illumina sequence data. **30**, 2114–2120 (2014).
157. Oksanen, J. *et al. Package 'vegan' Title Community Ecology Package Version 2.5-7*. (2020).

158. Rivera-Pinto, J. *et al.* Balances: a New Perspective for Microbiome Analysis. *mSystems* **3**, (2018).
159. Zevin, A. S., McKinnon, L., Burgener, A. & Klatt, N. R. Microbial translocation and microbiome dsybiosis in HIV-associated immune activation. *Curr. Opin. HIV AIDS* **11**, 182 (2016).
160. Sanders, E. J., Jaffe, H., Musyoki, H., Muraguri, N. & Graham, S. M. Kenyan MSM: No longer a hidden population. *AIDS* vol. 29 S195–S199 (2015).
161. Buchbinder, S. P. *et al.* Who should be offered HIV pre-exposure prophylaxis (PrEP)?: A secondary analysis of a Phase 3 PrEP efficacy trial in men who have sex with men and transgender women. *Lancet Infect. Dis.* **14**, 468 (2014).
162. Sui, Y. *et al.* Influence of gut microbiome on mucosal immune activation and SHIV viral transmission in naive macaques article. *Mucosal Immunol.* **11**, 1219–1229 (2018).
163. Lewis, K. *et al.* Enhanced translocation of bacteria across metabolically stressed epithelia is reduced by butyrate†. *Inflamm. Bowel Dis.* **16**, 1138–1148 (2010).
164. Furusawa, Y. *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**, 446–450 (2013).
165. Dillon, S. M. *et al.* Low abundance of colonic butyrate-producing bacteria in HIV infection is associated with microbial translocation and immune activation. *AIDS* **31**, 511–521 (2017).
166. Garrett, W. S. *et al.* Enterobacteriaceae Act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. *Cell Host Microbe* **8**, 292–300 (2010).
167. Vujkovic-Cvijin, I. *et al.* Dysbiosis of the gut microbiota is associated with HIV disease progression and tryptophan catabolism. *Sci. Transl. Med.* **5**, 193ra91 (2013).
168. Yan, Q. *et al.* Alterations of the gut microbiome in hypertension. *Front. Cell. Infect. Microbiol.* **7**, 381 (2017).
169. Fyderek, K. *et al.* Mucosal bacterial microflora and mucus layer thickness in adolescents with inflammatory bowel disease. *World J. Gastroenterol.* **15**, 5287–5294 (2009).
170. Qin, N. *et al.* Alterations of the human gut microbiome in liver cirrhosis. *Nature* **513**, 59–64 (2014).
171. Berg, R. D. Bacterial translocation from the gastrointestinal tract. *Trends Microbiol.* **3**, 149–154 (1995).
172. Marchetti, G., Tincati, C. & Silvestri, G. Microbial translocation in the pathogenesis of HIV infection and AIDS. *Clin. Microbiol. Rev.* **26**, 2–18 (2013).
173. Hunt, P. W. *et al.* T cell activation is associated with lower CD4+ T cell gains in human immunodeficiency vires-infected patients with sustained viral suppression during antiretroviral therapy. *J. Infect. Dis.* **187**, 1534–1543 (2003).
174. Wilson, C. M., Ellenberg, J. H., Douglas, S. D., Moscicki, A. B. & Holland, C. A. CD8+CD88+ T cells but not HIV type 1 RNA viral load predict CD4+ T cell loss in a predominantly minority female HIV + adolescent population. *AIDS Res. Hum. Retroviruses* **20**, 263–269 (2004).
175. Mavigner, M. *et al.* Altered CD4 + T cell homing to the gut impairs mucosal immune reconstitution in treated HIV-infected individuals. *J. Clin. Invest.* **122**, 62–69 (2012).
176. Muta, T. & Takeshige, K. Essential roles of CD14 and lipopolysaccharide-binding protein for activation of toll-like receptor (TLR)2 as well as TLR4 Reconstitution of TLR2- and TLR4-activation by distinguishable ligands in LPS preparations. *Eur. J. Biochem.* **268**, 4580–4589 (2001).

177. Sandler, N. G. *et al.* Plasma levels of soluble CD14 independently predict mortality in HIV infection. *J. Infect. Dis.* **203**, 780–790 (2011).
178. Pelsers, M. M. A. L. *et al.* Intestinal-type and liver-type fatty acid-binding protein in the intestine. Tissue distribution and clinical utility. *Clin. Biochem.* **36**, 529–535 (2003).
179. Redd, A. D. *et al.* Microbial translocation, the innate cytokine response, and HIV-1 disease progression in Africa. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 6718–6723 (2009).
180. Fitzgerald, F. C. *et al.* Microbial translocation does not drive immune activation in ugandan children infected with HIV. *J. Infect. Dis.* **219**, 89–100 (2019).
181. Reddy, B. S. *et al.* Commensal bacteria do translocate across the intestinal barrier in surgical patients. *Clin. Nutr.* **26**, 208–215 (2007).
182. Woodcock, N. P., Sudheer, V., El-Barghouti, N., Perry, E. P. & Macfie, J. Bacterial translocation in patients undergoing abdominal aortic aneurysm repair. *Br. J. Surg.* **87**, 439–442 (2000).
183. Palmer, C. D. *et al.* Enhanced Immune Activation Linked to Endotoxemia in HIV-1 Seronegative Men who have Sex with Men. *AIDS* **28**, 2162 (2014).
184. Pasternak, B. A. *et al.* Lipopolysaccharide exposure is linked to activation of the acute phase response and growth failure in pediatric Crohn's disease and murine colitis. *Inflamm. Bowel Dis.* **16**, 856–869 (2010).
185. Jiang, W. *et al.* Plasma Levels of Bacterial DNA Correlate with Immune Activation and the Magnitude of Immune Restoration in Persons with Antiretroviral-Treated HIV Infection. *J. Infect. Dis.* **199**, 1177–1185 (2009).
186. Kristoff, J. *et al.* Early microbial translocation blockade reduces SIV-mediated inflammation and viral replication. *J. Clin. Invest.* **124**, 2802–2806 (2014).
187. Sandler, N. G. *et al.* Sevelamer does not decrease lipopolysaccharide or soluble cd14 levels but decreases soluble tissue factor, low-density lipoprotein (ldl) cholesterol, and oxidized ldl cholesterol levels in individuals with untreated hiv infection. *J. Infect. Dis.* **210**, 1549–1554 (2014).
188. Chevalier, M. F. *et al.* The Th17/Treg Ratio, IL-1RA and sCD14 Levels in Primary HIV Infection Predict the T-cell Activation Set Point in the Absence of Systemic Microbial Translocation. *PLoS Pathog.* **9**, (2013).
189. Wallet, M. A. *et al.* Microbial translocation induces persistent macrophage activation unrelated to HIV-1 levels or T-cell activation following therapy. *AIDS* **24**, 1281–1290 (2010).
190. Glennie, S. J., Williams, N. A. & Heyderman, R. S. Mucosal immunity in resource-limited setting: is the battle ground different? *Trends Microbiol.* **18**, 487–493 (2010).
191. Ciccone, E. J. *et al.* Cycling of Gut Mucosal CD4+ T Cells Decreases after Prolonged Anti-Retroviral Therapy and is Associated with Plasma LPS Levels. *Mucosal Immunol.* **3**, 172 (2010).
192. Shreiner, A. B., Kao, J. Y. & Young, V. B. The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.* **31**, 69 (2015).
193. Leslie, J. L. & Young, V. B. The rest of the story: The microbiome and gastrointestinal infections. *Current Opinion in Microbiology* vol. 23 121–125 (2015).
194. Hold, G. L. Gastrointestinal Microbiota and Colon Cancer. *Dig. Dis.* **34**, 244–250 (2016).
195. Stearns, J. C. *et al.* Bacterial biogeography of the human digestive tract. *Sci. Rep.* **1**, 1–9 (2011).
196. Espey, M. G. Role of oxygen gradients in shaping redox relationships between the human

intestine and its microbiota. *Free Radical Biology and Medicine* vol. 55 130–140 (2013).
197. Albenberg, L. *et al.* Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology* **147**, 1055-1063.e8 (2014).
198. Chen, W., Liu, F., Ling, Z., Tong, X. & Xiang, C. Human Intestinal Lumen and Mucosa-Associated Microbiota in Patients with Colorectal Cancer. *PLoS One* **7**, e39743 (2012).
199. Ringel, Y. *et al.* High throughput sequencing reveals distinct microbial populations within the mucosal and luminal niches in healthy individuals. *Gut Microbes* **6**, 173–181 (2015).
200. Araújo-Pérez, F. *et al.* Differences in microbial signatures between rectal mucosal biopsies and rectal swabs. *Gut Microbes* **3**, 530–535 (2012).
201. Lepage, P. *et al.* Biodiversity of the mucosa-associated microbiota is stable along the distal digestive tract in healthy individuals and patients with IBD. *Inflamm. Bowel Dis.* **11**, 473–480 (2005).
202. Zoetendal, E. G. *et al.* Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Appl. Environ. Microbiol.* **68**, 3401–3407 (2002).
203. Budding, A. E. *et al.* Rectal swabs for analysis of the intestinal microbiota. *PLoS One* **9**, 101344 (2014).
204. Bassis, C. M. *et al.* Comparison of stool versus rectal swab samples and storage conditions on bacterial community profiles. *BMC Microbiol.* **17**, 1–7 (2017).
205. Carroll, I. M. *et al.* Molecular analysis of the luminal- and mucosal-associated intestinal microbiota in diarrhea-predominant irritable bowel syndrome. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **301**, G799 (2011).
206. Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota. *Nature Reviews Microbiology* vol. 14 20–32 (2015).
207. Maeda, Y. *et al.* Dysbiosis Contributes to Arthritis Development via Activation of Autoreactive T Cells in the Intestine. *Arthritis Rheumatol.* **68**, 2646–2661 (2016).
208. Si, J., You, H. J., Yu, J., Sung, J. & Ko, G. P. Prevotella as a Hub for Vaginal Microbiota under the Influence of Host Genetics and Their Association with Obesity. *Cell Host Microbe* **21**, 97–105 (2017).
209. Stephen-Victor, E., Fickenscher, H. & Bayry, J. IL-26: An Emerging Proinflammatory Member of the IL-10 Cytokine Family with Multifaceted Actions in Antiviral, Antimicrobial, and Autoimmune Responses. *PLoS Pathogens* vol. 12 (2016).
210. Larsen, J. M. The immune response to Prevotella bacteria in chronic inflammatory disease. *Immunology* vol. 151 363–374 (2017).
211. Gebrebrhan, H. *et al.* Rectal microbiota diversity in Kenyan MSM is inversely associated with frequency of receptive anal sex, independent of HIV status. *AIDS* **35**, 1091–1101 (2021).
212. Kriss, M., Hazleton, K. Z., Nusbacher, N. M., Martin, C. G. & Lozupone, C. A. Low Diversity Gut Microbiota Dysbiosis: Drivers, Functional Implications and Recovery. *Curr. Opin. Microbiol.* **44**, 34 (2018).
213. Hirbod, T. *et al.* HIV-neutralizing immunoglobulin A and HIV-specific proliferation are independently associated with reduced HIV acquisition in Kenyan sex workers. *AIDS* **22**, 727–735 (2008).
214. Burgener, A. *et al.* Identification of differentially expressed proteins in the cervical mucosa of HIV-1-resistant sex workers. *J. Proteome Res.* **7**, 4446–4454 (2008).
215. Mazzoli, S. *et al.* HIV-specific mucosal and cellular immunity in HIV-seronegative

partners of HIV-seropositive individuals. *Nat. Med.* **3**, 1250–1257 (1997).
216. Plummer, F. A., Ball, T. B., Kimani, J. & Fowke, K. R. Resistance to HIV-1 infection among highly exposed sex workers in Nairobi: What mediates protection and why does it develop? in *Immunology Letters* vol. 66 27–34 (Immunol Lett, 1999).
217. Horton, R. E., McLaren, P. J., Fowke, K., Kimani, J. & Ball, T. B. Cohorts for the study of HIV-1-exposed but uninfected individuals: Benefits and limitations. in *Journal of Infectious Diseases* vol. 202 (J Infect Dis, 2010).
218. Lajoie, J. *et al.* A distinct cytokine and chemokine profile at the genital mucosa is associated with HIV-1 protection among HIV-exposed seronegative commercial sex workers. *Mucosal Immunol.* **5**, 277–287 (2012).
219. Card, C. M., Ball, T. B. & Fowke, K. R. Immune Quiescence: A model of protection against HIV infection. *Retrovirology* vol. 10 1–8 (2013).
220. Tomescu, C., Abdulhaqq, S. & Montaner, L. J. Evidence for the innate immune response as a correlate of protection in human immunodeficiency virus (HIV)-1 highly exposed seronegative subjects (HESN). *Clinical and Experimental Immunology* vol. 164 158–169 (2011).
221. Mesquita, P. M. M. *et al.* Disruption of tight junctions by cellulose sulfate facilitates HIV infection: Model of microbicide safety. *J. Infect. Dis.* **200**, 599–608 (2009).
222. Fichorova, R. N. Guiding the vaginal microbicide trials with biomarkers of inflammation. in *Journal of Acquired Immune Deficiency Syndromes* vol. 37 184–193 (Lippincott Williams and Wilkins, 2004).
223. Naranbhai, V. *et al.* Innate immune activation enhances HIV acquisition in women, diminishing the effectiveness of tenofovir microbicide gel. in *Journal of Infectious Diseases* vol. 206 993–1001 (Oxford Academic, 2012).
224. Chen, Y. *et al.* Signature changes in gut microbiome are associated with increased susceptibility to HIV-1 infection in MSM. *Microbiome 2021 91* **9**, 1–18 (2021).

8. Appendix

8.1 Abbreviations

ACE: abundance-based coverage estimator
AIDS: acquired immunodeficiency syndrome
APC: antigen presenting cell
ART: antiretroviral therapy
ARV: Anti retroviral drugs
AZT: zidovudine
bNAbs: broadly neutralizing monoclonal antibodies
BV: bacterial vaginosis
cART: combination antiretroviral therapy
CAB-LA: long – acting injectable cabotegravir
CCR5: C-C motif chemokine receptor type 5
CLR: C-type like receptors
CPN60: chaperonin 60
CRAI: condomless receptive anal sex
CRF: circulating recombinant forms
CTL: cytotoxic T cell
CV: correlation of variation
CXCR4: C-X-C motif chemokine receptor type 4
DC: dendritic cells
DMSO: Dimethyl sulfoxide
ELISA: enzyme-linked immunosorbent assay
FDA: US Food and Drug Administration
FGT: Female genital tract
FSW: female sex workers
FTC: emtricitabine
FMO: Fluorescence Minus One
GALT: gut associated lymphoid tissue
GI: gastrointestinal
HESN: HIV exposed seronegative individuals
HIV: human immunodeficiency virus
HOYMAS: Health Options for Young Men on HIV/AIDS & STIs
HRP: horseradish peroxidase
IBD: Inflammatory bowel disease
IFABP: intestinal fatty acid binding protein
IFN: interferon
IL: interleukin
IQR: interquartile range
LBP: LPS-binding protein
LGC: Low microbial gene counts
LLOQ: Lower limit of quantification
LPS: lipopolysaccharide
LMIC: lower- and middle-income countries

LTNP: long term non-progressive
MDG: Millennium Development Goals
MFA: Multiple factor analysis
MHC: major histocompatibility complex
MSM: Men who have sex with men
MSW: Men who have sex with women
MT: Microbial translocation
NF: nuclear factor
NGS: next generation sequencing
NLR: nod like receptor
NK: natural killer
NMDS: nonmetric multidimensional scaling
NRTI: Nucleoside/nucleotide reverse transcriptase inhibitors
OTU: Operational taxonomical units
PAM : Partitioning around medoids
PAMP: pathogen- associated molecular patterns
PBMC: peripheral blood mononuclear cell
PCA: Principal component analysis
PCoA: Principal Coordinates Analysis
pDC: plasmolytic DC
PEP: post-Exposure prophylaxis
PERMANOVA: Permutational analysis of variance
PMA: phorbol 12-myristate 13-acetate
PrEP: pre-exposure prophylaxis
PRR: pattern recognition receptors
RA: rheumatoid arthritis
RBC: Red blood cell
RLRs: RIG-I-like receptors
RNA: ribonucleic acid
RT: reverse transcriptase
SCFA: short chain fatty acids
SFB: segmented filamentous bacteria
SIV: simian immunodeficiency virus
STBBI: Sexually transmitted blood borne infection
STI: sexually transmitted infections
SWOP: Sex Workers Outreach Programme
TCR: T cell receptor
TH: T helper
TDF: tenofovir disoproxil fumarate
TLR: toll like receptor
TMAO: trimethylamine-N-oxide
Treg: regulatory T cell
UNAIDS: United Nations Programme on HIV/AIDS
ULOQ: Upper limit of quantification
URF: unique recombinant forms
USA: United States of America

WHO: World Health Organization

8.2 Supplementary data

Table 3. Relative abundance of genera among participants

Genera	Prevalence(%)	HIV- (N=39)	Treated (N=21)	Untreated (N=10)	p value
		Median % (IQR)	Median % (IQR)	Median % (IQR)	
g_Bacteroides	100	6.7 (2.6-24.2)	15.0 (11.7-33.3)	11.6 (2.6-22.6)	0.2
g_Blautia	100	0.2 (0.1-0.7)	0.3 (0.2-0.6)	0.2 (0.1-0.4)	0.86
g_Dorea	100	0.2 (0.2-0.5)	0.4 (0.1-0.6)	0.5 (0.2-0.8)	0.01
g_Faecalibacterium	100	11.0 (6.9-13.5)	8.6 (5.9-11.8)	9.0 (6.5-11.4)	0.22
g_Lachnospira	100	0.4 (0.2-1.1)	0.8 (0.3-1.1)	0.2 (0.1-0.4)	0.33
g_Oscillospira	100	0.5 (0.3-0.8)	0.2 (0.1-0.5)	0.6 (0.4-0.8)	0.03
g_Prevotella	100	25.0 (8.0-35.5)	13.0 (1.0-37.0)	23.0 (11.0-29.5)	0.7
g_Roseburia	100	0.4 (0.1-0.9)	0.6 (0.2-1.4)	0.2 (0.1-0.3)	0.18
g_Coprococcus	98.78	0.3 (0.1-0.6)	0.2 (0.1-0.5)	0.2 (0.1-0.4)	0.53
g_Ruminococcus	98.78	1.0 (0.3-2.0)	0.7 (0.2-1.4)	0.9 (0.4-1.2)	0.92
g_Streptococcus	97.56	0.2 (0.1-0.7)	0.5 (0.1-2.4)	1.6 (0.3-2.0)	0.01
g_Dialister	95.12	2.6 (0.2-5.2)	0.2 (0.0-0.9)	1.4 (0.2-8.1)	0.05
g_Veillonella	92.68	0.2 (0.1-0.5)	0.1 (0.0-0.3)	0.1 (0.0-0.7)	0.69
g_Parabacteroides	91.46	0.8 (0.2-1.9)	0.4 (0.1-1.4)	1.7 (0.3-2.8)	0.46
g Phascolarctobacterium	90.24	1.8 (0.0-5.4)	2.4 (0.0-10.5)	2.0 (0.2-9.4)	0.74
g_.Eubacterium.	89.02	0.5 (0.0-1.5)	0.8 (0.0-1.7)	0.5 (0.2-1.7)	0.77
g_SMB53	85.37	0.1 (0.0-0.2)	0.0 (0.0-0.2)	0.1 (0.0-0.1)	0.44
g_Sutterella	84.15	0.2 (0.0-0.3)	0.2 (0.0-1.0)	0.1 (0.0-0.2)	0.23
g_Porphyromonas	82.93	0.2 (0.0-0.9)	0.1 (0.0-0.1)	0.0 (0.0-0.0)	0.28
g_Haemophilus	81.71	0.0 (0.0-0.1)	0.0 (0.0-0.1)	0.1 (0.0-0.3)	0.05
g_Peptoniphilus	81.71	0.0 (0.0-0.3)	0.0 (0.0-0.1)	0.0 (0.0-0.1)	0.81
g_WAL_1855D	81.71	0.1 (0.0-0.8)	0.0 (0.0-0.3)	0.0 (0.0-0.0)	0.55
g_.Ruminococcus.	80.49	0.0 (0.0-0.0)	0.0 (0.0-0.0)	0.1 (0.0-0.3)	0.04
g_Campylobacter	79.27	0.1 (0.0-0.4)	0.0 (0.0-0.1)	0.0 (0.0-0.0)	0.44
g_Bifidobacterium	78.05	0.1 (0.0-0.4)	0.1 (0.0-0.3)	0.0 (0.0-0.2)	0.93
g_Finegoldia	76.83	0.0 (0.0-0.1)	0.1 (0.0-0.3)	0.0 (0.0-0.2)	0.2
g_Anaerococcus	73.17	0.0 (0.0-0.2)	0.0 (0.0-0.3)	0.0 (0.0-0.1)	0.27
g_Bilophila	71.95	0.1 (0.0-0.1)	0.0 (0.0-0.0)	0.0 (0.0-0.1)	0.53
g_Clostridium	71.95	0.1 (0.0-0.2)	0.0 (0.0-0.0)	0.0 (0.0-0.1)	0.08
g_.Clostridium.	70.73	0.0 (0.0-0.1)	0.0 (0.0-0.0)	0.0 (0.0-0.0)	0.26

g_Catenibacterium	69.51	0.1 (0.0-0.4)	0.0 (0.0-0.1)	0.2 (0.0-0.4)	0.63
g_Odoribacter	68.29	0.1 (0.0-0.2)	0.0 (0.0-0.1)	0.0 (0.0-0.1)	0.53
g_Anaerostipes	67.07	0.0 (0.0-0.0)	0.0 (0.0-0.1)	0.0 (0.0-0.0)	0.05
g_Succinivibrio	63.41	0.0 (0.0-0.8)	0.0 (0.0-0.1)	0.0 (0.0-3.5)	0.19
g_Bulleidia	60.98	0.0 (0.0-0.3)	0.0 (0.0-0.1)	0.0 (0.0-0.1)	0.35
g_.Prevotella.	57.32	0.0 (0.0-0.6)	0.0 (0.0-0.1)	0.2 (0.0-0.8)	0.83
g_Collinsella	57.32	0.0 (0.0-0.5)	0.0 (0.0-0.3)	0.0 (0.0-0.6)	0.48

Differences between subgroups were assessed by Kruskal Wallis Test

Table 4. Associations with microbiome cluster in HIV negative MSM

Variable		Prevotella (n=23)	Bacteroides (n=16)	P-value
Age, Median (IQR)		23 (21-29)	23.5 (21-28.8)	0.91
Marital status, single		19 (82.6%)	14 (87.5%)	0.68
Education level	Primary	2 (8.7%)	3 (18.8%)	0.54
	Secondary	9 (39.1%)	7 (43.8%)	
	Tertiary	12 (52.2%)	6 (37.5%)	
Age at first sex, median (IQR)		15.0 (13.5-18)	16.5 (12.8-18)	0.85
Works as a sex worker		12 (52.2%)	8 (50.0%)	0.89
Regular partner (≥1) past 1 month		15 (65.2%)	11 (68.8%)	0.82
Casual partner (≥1) past 1 month		16 (80.0%)	11 (68.8%)	0.44
Paying partner (≥1) past 1 month		10 (43.5%)	8 (50.0%)	0.69
Oral sex (sometimes or always)	Regular	7 (30.4%)	6 (37.5%)	0.65
	Casual	3 (17.6%)	2 (18.2%)	0.97
	Paid	4 (36.4%)	1 (12.5%)	0.24
Insertive sex (always)	Regular	22 (95.7%)	14 (87.5%)	0.35
	Casual	14 (82.4%)	9 (81.8%)	0.97
	Paid	8 (72.7%)	8 (100.0%)	0.10
Receptive sex (sometimes or always)	Regular	10 (43.5%)	8 (50.0%)	0.69
	Casual	6 (37.5%)	5 (45.5%)	0.68
	Paid	8 (72.7%)	4 (50.0%)	0.31
Always uses condoms	Regular	17 (73.9%)	11 (68.8%)	0.73
	Casual	12 (75.0%)	9 (81.8%)	0.68
	Paid	10 (90.9%)	6 (75.0%)	0.35
Practices rectal douching		12 (52.2%)	13 (81.2%)	0.06
STI positive**		4/18 (22.2%)	4/16 (25%)	0.85

Differences between subgroups were assessed by Chi-square analysis for categorical variables and Mann Whitney U test for continuous variables.

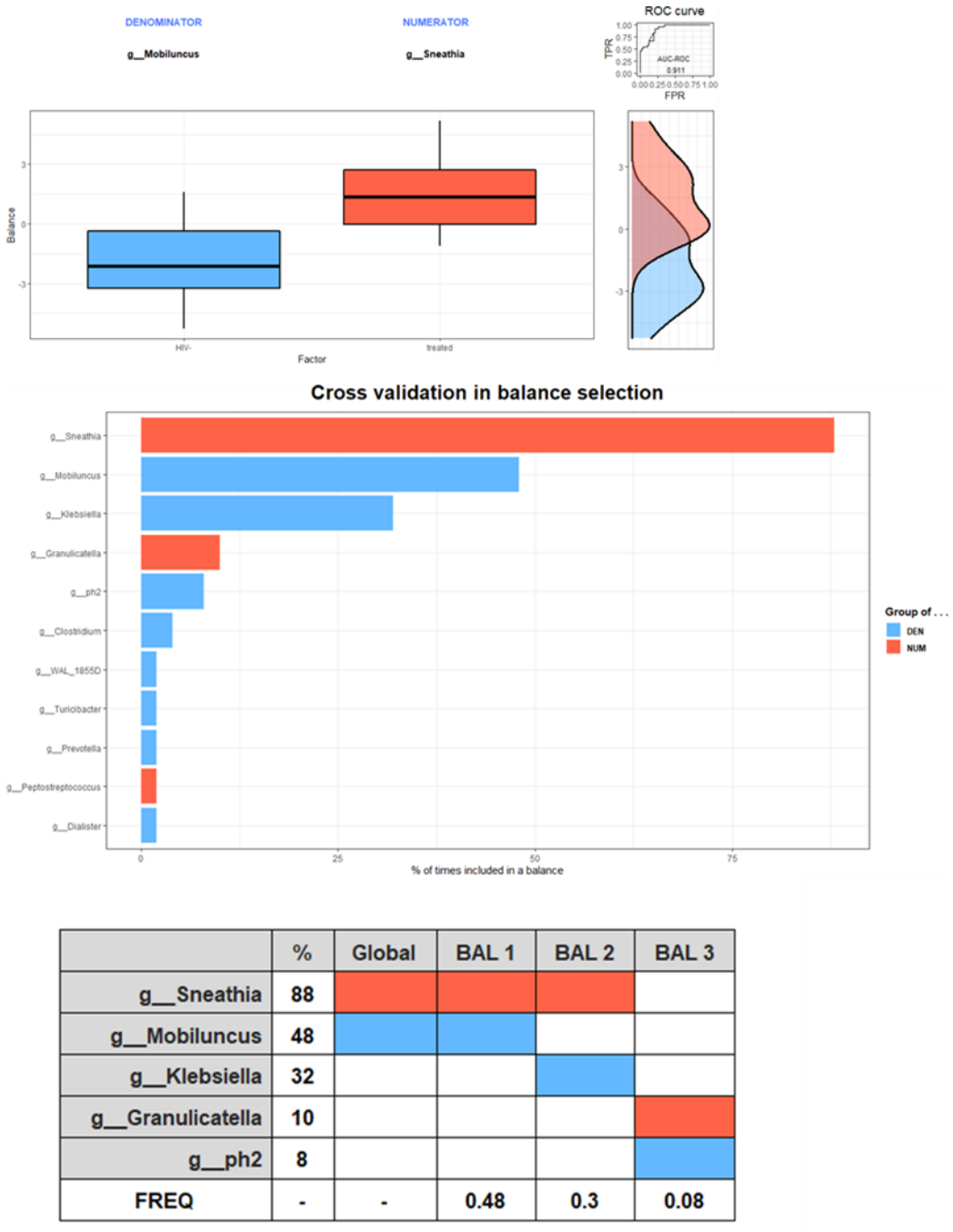
*p value < 0.05 was considered significant

** STI panel included *Chlamydia trachomatis, Neisseria gonorrhoea, Mycoplasma genitalium, Mycoplasma hominis and Ureaplasma urealyticum.*

Table 5. Flow cytometry panels description

Channel	Flourochrome	Clone	Company / Catalogue #
Panel 1 – T cell activation and memory			
FITC/BB515	CD38	HIT2	BD/ 564499
PerCP-Cy5-5			
PeCy7	CD27	M-T271	BD/ 560609
PeCy5			
Pe-Dazzle 594	CCR7	150503	BD/ 562381
PE	IntB7	FIB504	BD/ 555945
BV421	CCR5	2D7	BD/ 562576
V500	aqua fluorescent reactive dye		Thermo Fisher/ L34966
BV605			
BV650	HLA DR	L243	BD/ 564231
BV711			
BV786			
APC-H7/Fire 750	CD3	SK7	BD/ 560275
A700/APC-R700	CD69	FN50	BD/ 565154
APC	Ki67	B56	BD/ 558615
BUV 395	CD8	RPA-T8	BD/ 563796
BUV 496	CD4	SK3	BD/ 564652
BUV 737	CD45RA	HI100	BD/ 564442
Panel 2 – intracellular cytokine production			
FITC/BB515	TNFa	Mab11	Thermo Fisher/ 11-7349-41
PerCP-Cy5-5			
PeCy7	IFNg	B27	BD/ 557643
PeCy5			
Pe-Dazzle 594	GM-CSF	BVD2-21C11	BD/ 562857
PE	IL-22	22URTI	eBioscience/ 12-7229-42
BV421	IL-17A	N49-653	BD/ 562933
V500	aqua fluorescent reactive dye		Thermo Fisher/ L34966
BV605			
BV650	CD8	SK1	Biolegend/ 344730
BV711	IL-10	JES3-9D7	BD/ 564050
BV786	IL-4	MP4-25D2	BD/ 564113
APC-H7/Fire 750	CD3	SK7	BD/ 560275
A700/APC-R700			

APC			
BUV 395	CD45RO	UCHL1	BD/ 564291
BUV 496	CD4	SK3	BD/ 564652
BUV 737	IntB7	FIB504	BD/ 565604
Panel 3 – Regulatory and Tfh cells			
FITC/BB515	CD25	2A3	BD/ 564467
PerCP-Cy5-5			
PeCy7	Foxp3	236A	eBioscience/ 25-4777-41
PeCy5			
Pe-Dazzle 594	CXCR3	G025H7	Biolegend/ 353736
PE	IntB7	FIB504	BD/ 555945
BV421	CCR5	2D7	BD/ 562576
V500	aqua fluorescent reactive dye		Thermo Fisher/ L34966
BV605	CD3	SK7	BD/ 563219
BV650			
BV711	CD127	HIL-7R-M21	BD/ 563165
BV786	CTLA4	BM13	BD/ 563931
APC-H7/Fire 750	PD1	EH12.2H7	Biolegend/ 329954
A700/APC-R700	CD69	FN50	BD/ 565154
APC			
BUV 395	CXCR5	RF8132	BD/ 740266
BUV 496	CD4	SK3	BD/ 564652
BUV 737			



	%	Global	BAL 1	BAL 2	BAL 3
g__Sneathia	88				
g__Mobiluncus	48				
g__Klebsiella	32				
g__Granulicatella	10				
g__ph2	8				
FREQ	-	-	0.48	0.3	0.08

Figure 17. Specific genera associated with HIV status using forward selection analysis generated in the selbal R package.(a) Box plots showing the balance value distribution (boxplots) in the HIV-negative (blue) and HIV-positive treated groups (red). (b) Box plots showing the balance value distribution (boxplots) in the HIV-negative and HIV-positive untreated groups (red). The genera contributing to the balance are listed in rank order of importance (bottom panels), with genera over-represented in HIV-negative (in blue) and HIV-positive (in red), sorted by frequency of appearance in multiple cross-validations. The analysis included and genera that had a prevalence rate greater than 5%.

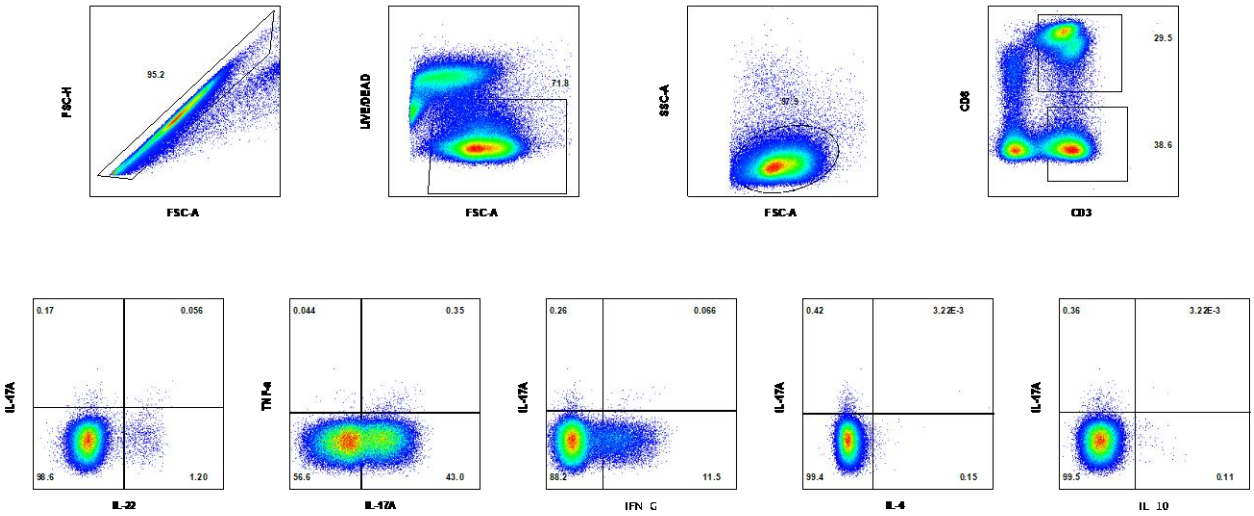


Figure 18. Representative staining of PBMCs for panel 2 following four hour stimulation with PMA and Io. Singlets, live cells and lymphocyte populations were selected on their respective plots. CD4+ T cells were identified as CD3+ and CD4+. CD4+ T cell subsets were differentiated by IL-22, IL-17A, TNF-a, IFN-g, IL-4 and IL-10.

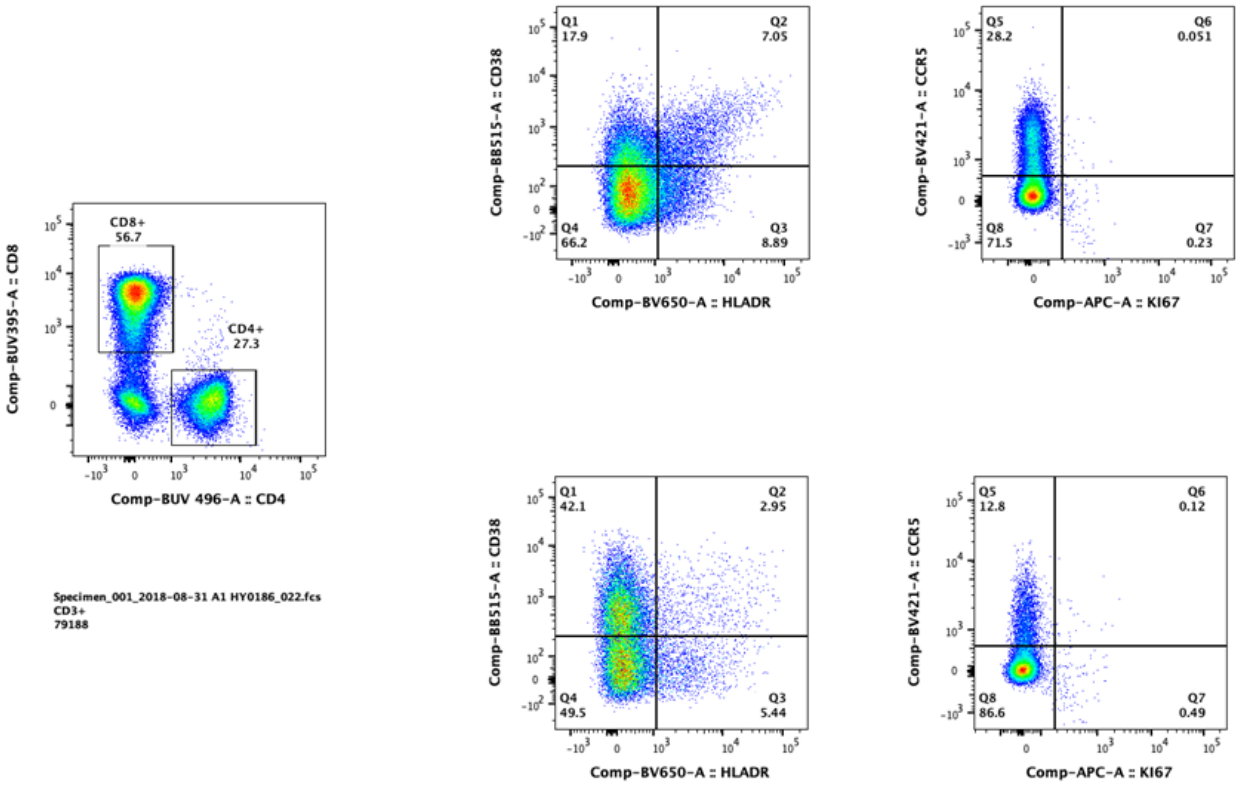


Figure 19. Representative staining of PBMCs from panel 1 for T cell activation. CD4+ T cells were identified as CD3+ and CD4+. CD4+ T cell activation were identified using CD38, HLA-DR, Ki67 and CCR5.

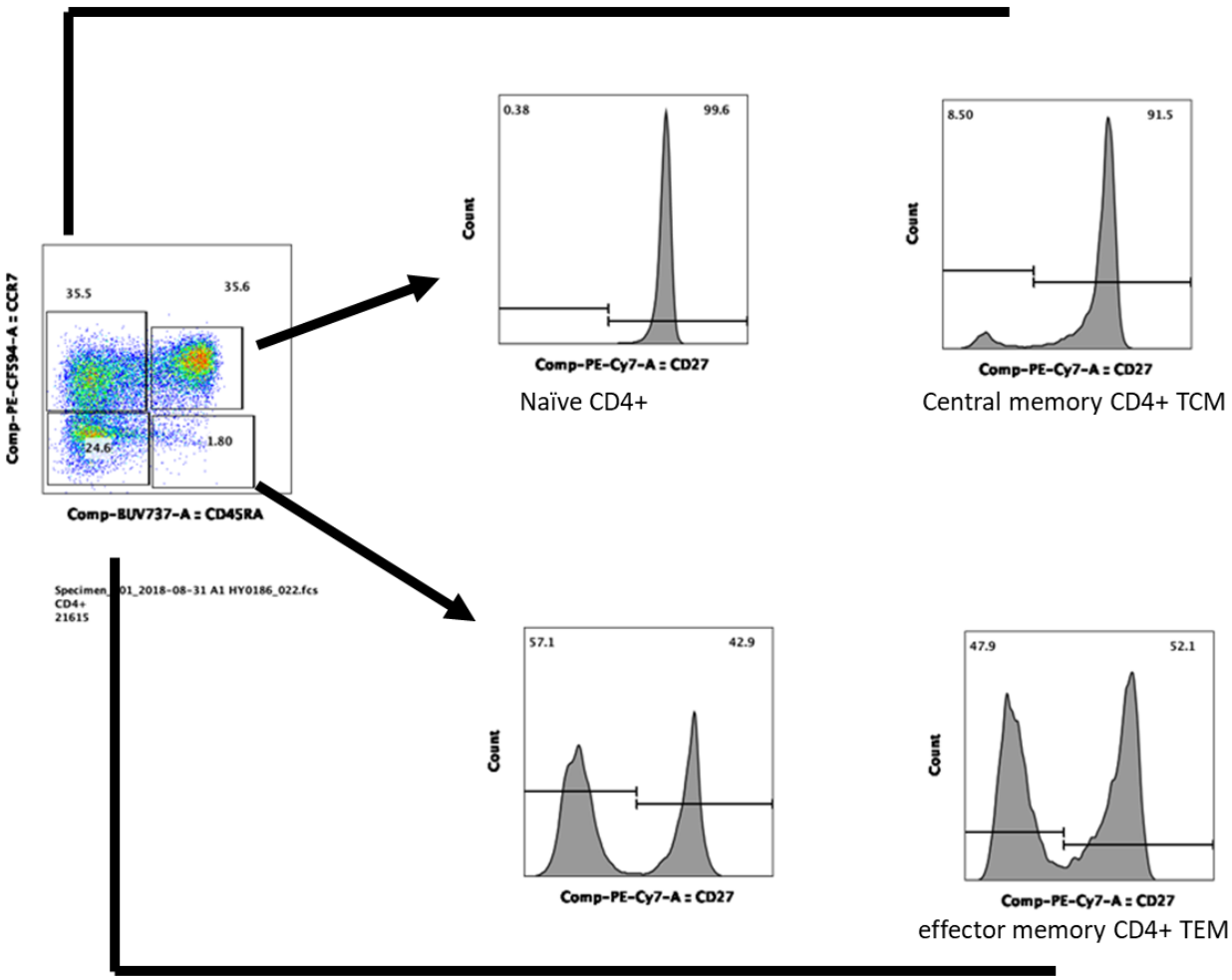


Figure 20. Representative staining of PBMCs from panel 1 for memory CD4+ T cell differentiation. CD4+ T cells were identified as CD3+ and CD4+. CD4+ T cell memory were identified as; naive (TN, CD45RA+CCR7+CD27+), central-memory (TCM, CD45RA-CCR7+CD27+), transitional-memory (TTM, CD45RA−CCR7−CD27+) and effector-memory cells (TEM, CD45RA−CCR7−CD27−).