IDENTIFYING THE CELLULAR HIV-1 RESERVOIR IN LYMPH NODES OF ANTIRETROVIRAL THERAPY SUPPRESSED INDIVIDUALS

Isabella Ferreira

Submitted in fulfilment of the requirements for the degree of Master of Medical Sciences in the School of Virology, University of KwaZulu-Natal.

May 2019

Dr. Alex Sigal

Isabella Ferreira

DECLARATION

I, Ms Isabella Ferreira declare as follows:

1. That the work described in this thesis has not been submitted to UKZN or other tertiary institution for purposes of obtaining an academic qualification, whether by myself or any other party.

2. That my contribution to the project was as follows:

   I was involved in selecting which lymph node samples to use for the study. The cohort had already been established prior to my entry into the lab but I chose the final lymph nodes that appear in the manuscript. I set up some of the initial PBMC in vitro infections with NL4-3 virus and optimized the broadly neutralizing anti-Env J3 antibody for use in identifying HIV-1 infected cells from lymph nodes. I cloned full-length HIV-1 sequences from the lymph node samples and sequenced them. And I was involved in optimizing the Seq-Well protocol for the lymph node samples using in vitro infections as well as the flow cytometry and the running of both of these on the samples. I conducted the flow cytometry analysis and sequence analysis from the full-length clones.

3. That the contributions of others to the project were as follows:

   Dr. Alex Sigal conceived of the idea for the project and started a collaboration with Prof. Alex Shalek, who also advised on the project, with the idea of running single-cell RNA-Sequencing on lymph node samples. Dr. Alex Sigal established the cross-sectional lymph node cohort in collaboration with clinicians in the Department of Cardiothoracic Surgery at Inkosi Albert Luthuli Central Hospital. Jessica Hunter and Laurelle Jackson ran the initial proof of concept and pilot experiments with Carly Ziegler. These included in vitro infections of Rev-CEM cells with two different virus strains and a pilot Seq-Well experiment on one lymph node sample. Dr. Kiera Clayton and Heather Stuart optimized the broadly neutralizing J3 antibody using in vitro infections of $CD4^+$ T cells. Carly Ziegler did the single-cell data analysis and optimized the Seq-Well protocol and flow cytometry protocol as well as the running of the experiments on the lymph node samples with myself. Dr. Alejandro Balazs and Dr. Martin Deymier advised and assisted with both the cloning as well as the overall data generated from the project.

4. Signed  Date: 29/05/2019

ACKNOWLEDGEMENTS

I would like to acknowledge:

Dr. Alex Sigal for his supervision and guidance on the project and for giving me the opportunity to pursue this research. Alex has really believed in me and I want to thank him for the incredible opportunities for training and expanding my knowledge that he made possible. I hope to continue collaborating with him in the future.

Professor Alex Shalek for inviting me into his lab as a visiting student at the Massachusetts Institute of Technology and for his advice and input during this time, which have been incredibly useful to the progress of the project.

Dr. Alejandro Balazs for inviting me into his lab as a visiting student at the Ragon Institute of MGH, MIT and Harvard. His advice has been incredibly valuable to the progression of the research.

Carly Ziegler for being an incredible collaborator on this project. It has been wonderful working alongside her doing this research and I am very appreciative for this opportunity as well as her expertise, particularly in the single-cell RNA-sequencing field.

Jessica Hunter for also being an incredible collaborator and lab mate who really helped me to find my feet in the lab, particularly with *in vitro* HIV-1 experiments, and has continually offered her support and advice on the project.

Dr. Martin Deymier for his help with cloning of the full-length HIV-1 sequences from the lymph node samples and his advice on the project.

The members of the Sigal lab at Africa Health Research Institute, the Shalek lab at the Massachusetts Institute of Technology, and the Balazs lab at the Ragon Institute of MGH, MIT and Harvard for their kindness and support of my research.

SANTHE and all its staff for funding the majority of my research and for their continual support as well as my SANTHE thesis committee: Dr. Henrik Kløverpris, Dr. Derseree Archery, and Dr. David Russell.

The Poliomyelitis Research Foundation for funding part of my research and the National Institute of Health for funding the research component.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER 1

CHAPTER 2

ACRONYMS

HIV-1	Human immunodeficiency virus 1
ARV	Antiretroviral
ART	Antiretroviral therapy
CNS	Central nervous system
PD-1	Programmed cell death 1
PBMCs	Peripheral blood mononuclear cells
EFV	Efavirenz
FTC	Emtricitabine
TFV	Tenofovir
LPV/r	Lopinavir and ritonavir
3TC	Lamivudine
NVP	Nevirapine
LC-MS-MS	Chromatography-mass tandem spectrometry
ddPCR	Digital droplet PCR
qVOA	Quantitative viral outgrowth assay
PHA	Phytohemagglutinin
TB	Tuberculosis
GFP	Green fluorescent protein
BFP	Blue fluorescent protein
UMIs	Unique molecular identifiers
t-SNE	*t*-distributed stochastic neighbor embedding
PID	Participant identification number
MIF	Macrophage migration inhibitory factor
MX2	Myxovirus resistance-2
MACS	Magnetic activated cell sorting
SNN	Shared nearest neighbour
HIV-1	Human immunodeficiency virus 1
ART	Antiretroviral therapy
CNS	Central nervous system
PD-1	Programmed cell death 1
PBMCs	Peripheral blood mononuclear cells
LC-MS-MS	Chromatography-mass tandem spectrometry

CHAPTER 1: INTRODUCTION

The Human Immunodeficiency Virus 1 (HIV-1) presents a major challenge, both locally and globally. KwaZulu-Natal is the epicenter of the pandemic with infection rates as high as 25% in areas of the population (Kharsany and Karim, 2016). Whilst antiretroviral (ARV) drug treatments are readily available to those infected with HIV-1 and are able to successfully suppress the virus/suppress viremia, they are not curative (Sigal and Baltimore, 2012). It is unsustainable for individuals living with HIV-1 to take daily medication for the rest of their lives, especially in the developing world, where access to medication may be limited. Therefore, it is critical that the HIV-1 reservoir is identified in order to pave a way for a functional cure.

Should an individual stop antiretroviral therapy (ART), viral levels in the plasma will rebound, even when their plasma viremia has been suppressed to undetectable levels for several years (Chun et al., 2015, Eriksson et al., 2013, Licht and Alter, 2016). This reservoir may be maintained by a population of latently infected cells (Siliciano and Greene, 2011), or by ongoing replication (Lorenzo-Redondo et al., 2016, Sigal et al., 2011). This is as a result of the HIV-1 reservoir: sanctuary sites throughout the body that are thought to include the central nervous system (CNS) (Fois and Brew, 2015, Hellmuth et al., 2015), gut (McElrath et al., 2013, Yukl et al., 2013), genital tract (Marcelin et al., 2008, Politch et al., 2012), and lymph nodes (Horiike et al., 2012, Spiegel et al., 1992) as well as latently infected T cells in the blood (Kulpa and Chomont, 2015, Svicher et al., 2014). These sanctuary sites may harbor replication-competent virus (Soriano-Sarabia et al., 2014, Wang et al., 2018a, Wang et al., 2018b) in specific cell subsets. Identifying which cells harbor replication-competent virus is crucial in the eradication of the reservoir without depleting cells that make up a large immune system compartment.

The latent HIV-1 reservoir persists in $CD4^+$ T cells, where the cells are infected but not actively producing virus. This is established during acute infection, likely within the first days of infection (Chun et al., 1998, Ruelas and Greene, 2013), and is transcriptionally silent with a long half-life (Finzi et al., 1999, Siliciano and Greene, 2011). Research has shown that cells with single integrations of HIV-1 decreased over time on therapy and that the cells that remained infected had integrated HIV-1 in the silent regions of their genome (Cohn et al., 2015). However, latency reversal agents are being studied, to reactivate transcription of HIV-1 in these cells, which could cause the immune system to either clear the cells or for them to spontaneously die (Ruelas and Greene, 2013). The HIV-1 reservoir in lymph nodes presents an obstacle to the eradication of the infection, as it allows for the virus to persist in the presence of ART (Kulpa and Chomont, 2015,

Kutscher et al., 2016, Reeves et al., 2018, Lorenzo-Redondo et al., 2016). Identifying which cell types in the lymph node harbour competent virus is key to developing a possible cure for HIV-1 (Chun et al., 2015, Siliciano and Siliciano, 2015, Churchill et al., 2015, Sigal et al., 2011).

Ongoing replication is another proposed mechanism whereby HIV-1 persists in the presence of ART (Lorenzo-Redondo et al., 2016, Sigal et al., 2011). Due to low drug penetrance in lymphoid tissue, HIV-1 is able to replicate and infect new cells, including through cell-to-cell spread of the virus (Sigal et al., 2011). This can allow for evasion of the immune system, a further contributing factor to HIV-1 persistence (Sattentau, 2008). Research in support of ongoing replication found an increase in episomal DNA, by increasing the intensity of a suppressive ART regimen through the addition of raltegravir, which is an indicator of ongoing replication (Buzon et al., 2010). However, ongoing replication is a controversial school of thought and the model of latency is generally preferred.

Several recent studies have proposed cellular locations for the HIV-1 reservoir in the face of ART. CD32a was identified as a marker of a potential $CD4^+$ T cell reservoir in peripheral blood mononuclear cells (PBMCs) (Descours et al., 2017). When cells from ART suppressed individuals were assessed for infection frequency, $CD32a^+CD4^+$ T cells had a 1000-fold higher level of HIV-1 DNA than $CD32a^-CD4^+$ T cells (Descours et al., 2017). A quantitative viral outgrowth assay for inducible replication-competent provirus found a 3000-fold enrichment in $CD32a^+$ cells in comparison to total $CD4^+$ T cells (Descours et al., 2017).

A second putative cellular HIV-1 reservoir in the T cell compartment is the programmed cell death protein 1 (PD-1) cell subset, which is mainly composed of T follicular helper cells (Banga et al., 2016, Perreau et al., 2013). Populations of these cells were found to harbour more replication-competent virus than other memory $CD4^+$ T cell subsets in both blood and lymph nodes from ART suppressed individuals (Banga et al., 2016). PD-1 lymph node cells from individuals who had been on long term ART had higher levels of HIV-1 RNA present than any other memory $CD4^+$ T cell subset, indicating that $PD\text{-}1^+$ cells could be a major source of replication-competent provirus in both the blood and lymph nodes and a site of active viral replication during long term ART (Banga et al., 2016).

The majority of research focused on the HIV-1 reservoir has been conducted in the Western World, using *in vitro* models with lab strains of HIV-1 or *in vivo* infected samples from individuals with Clade B HIV-1 (Omondi et al., 2019, Siliciano and Greene, 2011). It is currently unclear what the cellular source of the HIV-1 reservoir in sub-Saharan Africa, and specifically in South Africa, is. Very few studies have been conducted on lymph nodes (Cohen, 2011) and none

in lymph nodes from individuals infected with Clade C HIV-1 from South Africa. We investigated the cellular source of HIV-1 in the presence of suppressive ART in lymph nodes from surgical resections. These samples were unique, since they allowed access to the HIV-1 reservoir in large, non-superficial lymph nodes, which provided a rich cellular infection environment for the virus.

We hypothesized that the HIV-1 reservoir in the face of ART was contained in a specific T cell subset that could be identified by surface markers. The aims of the study were to:

1) Assemble a cohort of lung lymph nodes from ART suppressed individuals
2) Determine the cell subtype which could express inducible HIV-1 through large-number/high-throughput single-cell RNA sequencing using Seq-Well technology and flow cytometry, including anti-HIV-1 p24 Gag staining
3) Generate full length HIV-1 genome sequences from RNA and DNA extracted from the lymph node cells and functionalize these to create infectious molecular clones

We assembled a cross-sectional cohort of study participants in collaboration with clinicians at the Department of Cardiothoracic Surgery at Inkosi Albert Luthuli Central Hospital in Durban, South Africa. We selected lymph nodes from seven HIV-1 positive study participants and three HIV-1 negative study participants, based on their viral load, number of years on ART, and TB status (Table 1). Liquid chromatography-mass tandem spectrometry (LC-MS-MS) was run on the matched PBMCs from each study participant to detect ART levels and whether the study participant was adhering to their regimen (Table 2). We quantified the percentage of infection in the cells using anti-HIV-1 p24 Gag staining as well as staining with the J3 anti-Env broadly neutralizing nanobody. We found that the percentage of infected cells varied between study participants from 0.034% to 0.16%, indicated by $p24^+$ cells (Figure 1).

In collaboration with Professor Alex Shalek and his graduate student, Carly Ziegler, at the Massachusetts Institute of Technology, we optimized methods related to the single-cell RNA sequencing technology Seq-Well, that has been developed by the Shalek laboratory (Gierahn et al., 2017). We initially ran several pilot studies to determine whether it was possible to detect HIV-1 transcripts as well as the depth of detection down to single base-pair mutations. Rev-CEM-E7 cells were infected with NL4-3, NL4-3 containing the L100I mutation, or both virus strains. Single cells were sorted and sequenced and the HIV-1 transcripts were detected down to the base-pair mutation as well as the abundance of viral transcripts per cell (Figure 3). To detect the resolution of HIV-1 infection using the Seq-Well platform, lymph node cells from HIV-1 negative study participants were infected with blue fluorescent protein (BFP) tagged NL4-3 or JR-CSF

and BFP positive cells were sorted. A threshold for the level of background was determined by comparing BFP positive cells to mock infected cells (Figure 4).

Table 1: Clinical parameters of cross-sectional lymph node cohort

PID	HIV Status	Active TB	Sex	Age	Nadir CD4	Diagnosis date	Years untreated	ART start date	Years on ART	MDR/XDR episode	LN location
024-09-0198	Positive	Yes	Male	52	N/A	N/A	N/A	N/A	N/A	XDR	N/A
024-09-0207	Positive	No	Female	53	N/A	2012	2	2014	2	-	Lateral bronchus
024-09-0227	Positive	No	Male	40	227	2011	5	2016	<1	-	Left anterior medial basal
024-09-0276	Positive	No	Male	32	563	2016	<1	2016	1	-	Mediastinal
024-06-0312	Positive	No	Female	47	N/A	2003	2	2005	13	-	Peribronchial
024-06-0320	Positive	No	Female	33	85	2015	<1	2015	3	MDR	Peri-bronchus intermedius
024-06-0338	Positive	No	Female	37	387	N/A	N/A	2018	<1	-	Hilar
024-09-0274	Negative	No	Male	22	N/A	N/A	N/A	N/A	N/A	-	Parabronchial
024-09-0219	Negative	No	Female	30	N/A	N/A	N/A	N/A	N/A	-	Left lower lobectomy
024-09-0236	Negative	No	Male	40	N/A	N/A	N/A	N/A	N/A	-	Right upper lobectomy

N/A , not available
PID, participant identification number

Table 2: Measured parameters of cross-sectional lymph node cohort

PID	HIV Status	Viral load	TDF	FTC	EFV	LPV	RTV	ABC	3TC	Regimen by LC/MS/MS
024-09-0198	Positive	<40	51,8	380	1340	BQL	BQL	BQL	BQL	EFV/TDF/FTC
024-09-0207	Positive	<40	76,9	301	1860	BQL	BQL	BQL	BQL	EFV/TDF/FTC
024-09-0227	Positive	<40	50,8	327	1470	BQL	BQL	BQL	BQL	EFV/TDF/FTC
024-09-0276	Positive	<40	66,8	255	1310	BQL	BQL	BQL	BQL	EFV/TDF/FTC
024-06-0312	Positive	<40	BQL	44	1970	BQL	BQL	BQL	BQL	EFV/FTC
024-06-0320	Positive	<40	BQL	BQL	BQL	247	7	BQL	14	LPV/r/3TC
024-06-0338	Positive	58	-	-	-	-	-	-	-	-

All ARV in ng/ml
PID, participant identification number

Seq-Well allows for the high-throughput sequencing of thousands of single cells in parallel from low-input samples. A pilot was run on lymph node cells from PID 024-09-0198, and HIV-1 transcripts were detected in various T cell subsets as well as in myeloid lineage cells (Figure 6). However, due to the rarity of HIV-1 infected cells in the lymph node compartment from an individual who is suppressed on ART, we had to optimize methods to enrich for HIV-1 infected cells. We optimized the anti-Env broadly neutralizing camelid nanobody J3 (McCoy et al., 2012) to select for HIV-1 infected cells, firstly using flow cytometry and then using magnetic-activated cell sorting columns (MACS), before loading the J3 positive cells onto Seq-Well arrays (Figure 2 & 7). The MACS columns were also used to enrich for $CD4^+$ T cells, as these were most likely the cells that would harbor HIV-1, and to deplete $CD19^+$ B cells, as these made up a large portion of the cells found in the lymph nodes. We loaded seven arrays in total per study participant, both HIV-1 negative and HIV-1 positive: two unsorted, two $CD4^+$ enriched, two $CD19^+$ depleted, and one $J3^+$ enriched. We found HIV-1 transcripts in T cell and myeloid lineage cell types (Figure 7).

To validate the infected cell types, flow cytometry was run on the lymph node cells from PID 024-09-0276, using anti-HIV-1 p24 Gag and anti-Env J3 in combination with cell type markers for major cell lineages. We found cells doubly positive for p24 in $CD3^+CD4^+$ T cells as well as in $CD33^+$ myeloid cells. We looked further into different myeloid lineage cells, including $CD14^+$ cells and $CD68^+$ cells and found cells positive for both p24 and J3. Further characterization of the T cell subsets is still needed (Figure 7).

In conjunction to the single-cell RNA-sequencing, full length genomes were cloned from the lymph node cells using extracted DNA and RNA. The genomes were cloned in two parts, sequenced, and then combined to form HIV-1 genomes specific to each study participant. The single-cell RNA-sequencing data from each study participant was aligned to the genome from the same study participant as well as to the HIV-1 Clade C consensus sequence. This was done in collaboration with Dr. Alejandro Balazs, and his post-doctoral fellow Dr. Martin Deymier, at the Ragon Institute of MGH, MIT and Harvard (Figure 5)

Using a variety of techniques optimized to specifically identify HIV-1 infected cells, in lymph nodes from ART suppressed individuals, allowed us to identify the infected cells and build on the research that exists in this field. Further characterization of the exact T cell subsets is still required however, the existing data presents T cells and myeloid lineage cells as being the main infected cell types in the presence of ART. Understanding how the reservoir presents itself in these individuals is critical for creating a functional cure that could reduce or eradicate the HIV-1 reservoir. This would allow individuals infected with HIV-1 to be in a state of monitored remission in future studies

CHAPTER 2: IDENTIFYING THE CELLULAR HIV-1 RESERVOIR IN LYMPH NODES OF ANTIRETROVIRAL THERAPY SUPPRESSED INDIVIDUALS

Abstract

HIV-1 infection is suppressed but not cured in the face of antiretroviral therapy (ART). Pinpointing the cellular HIV-1 reservoir, which allows HIV-1 to persist, is key to the eradication of the virus. Lymph nodes are known to be a reservoir site for HIV-1 persistence, and we have assembled lymph nodes from a cross-sectional cohort of participants on suppressive ART to better understand the cellular HIV-1 reservoir. We developed a novel single-cell RNA-Seq methodology to identify the cellular HIV-1 reservoir in the lymph node compartment in ART suppressed individuals. HIV-1 positive cells from these lymph nodes were stained with anti-HIV-1 antibodies and selected using flow cytometric sorting. Seq-Well, a high throughput single-cell RNA-Seq approach, was then performed to detect *gag* and *env* HIV-1 transcripts in individual cells, as well as the infected subtype using the cellular transcriptome. In parallel, the consensus near full length viral clone from the lymph node was sequenced and used for alignment. Using our methods for identifying HIV-1 infected cells from lymph nodes from chronically infected individuals, we have identified both known and novel putative host markers that are associated with persistent infection. These included co-expression of *APOBEC3G*, *NFAT5*, and *NFKB2* in cells that contained HIV-1 mRNA. Our results show that cells with transcriptomes consistent with a T cell origin are the main infected population, and we are in the process of deeply characterizing the cell subtypes involved that also express markers of HIV-1 infection.

Introduction

HIV-1 presents a major global public health burden, particularly in sub-Saharan Africa. South Africa, and in particular the province of KwaZulu-Natal, is one of the epicentres of the HIV-1 pandemic (Kharsany and Karim, 2016). Most individuals who take antiretroviral therapy (ART) are not viraemic but lifelong suppression is challenging, both economically and logistically. There are multiple mechanisms whereby HIV-1 infection may persist in the presence of ART. These include latency versus ongoing replication which can occur in several reservoir sites throughout the body including the lymph nodes, gut, brain, and genital tract (Fois and Brew, 2015, Horiike et al., 2012, Kulpa and Chomont, 2015, McElrath et al., 2013, Svicher et al., 2014).

In recent years, varying cell types including T follicular helper cells, programmed cell death 1 (PD-1) cells, and T cells expressing CD32a (Descours et al., 2017), have been described as the cellular HIV-1 reservoir (Banga et al., 2016, Noto et al., 2018, Fromentin et al., 2016). Cell types such as central memory T cells have been found to be preferentially infected over other T cell subsets, such as effector memory T cells and transitional memory T cells (Fromentin et al., 2019). These cell types have been shown to harbour replication-competent virus (Lee and Lichterfeld, 2016, Siliciano and Siliciano, 2015).

It has also been reported that macrophages are able to be infected by HIV-1 and lead to the persistence of the HIV-1 reservoir. T cells maintain the HIV-1 reservoir through homeostatic proliferation of latent T cells, (Chomont et al., 2009) whereas macrophages are long-lived and have been shown to be resistant to the cytopathic effects of HIV-1 replication in *in vitro* studies (Castellano et al., 2017, Clayton et al., 2018). Macrophages have also been shown to be infected *in vivo* and form HIV-1 reservoirs in an i*n vivo* mouse model (Honeycutt et al., 2016, Mlcochova et al., 2014) as well as in humans (Castellano et al., 2017). The primary cell types infected by HIV-1 in the brain are myeloid lineage cells such as microglia or perivascular macrophages (Gorry et al., 2001, Schnell et al., 2009, Schnell et al., 2011, Williams et al., 2001), and in some patients HIV-1 replication persists in the brain in the face of ART (Kugathasan et al., 2017).

There are several ways in which the HIV-1 reservoir is typically measured and characterised. These include measuring total proviral HIV-1 DNA through qPCR and digital droplet PCR (ddPCR). qPCR has traditionally been used to measure proviral HIV-1 DNA, with a limit of detection of 50 proviral copies of DNA per $CD4^+$ T cell (Kabamba-Mukadi et al., 2005). ddPCR has an advantage over qPCR, as it is able to quantitatively and discreetly measure the number of HIV-1 integrations, by measuring the actual amount of target DNA. Each molecule is contained in one droplet, which fluoresces if the droplet contains the molecule of interest. This allows for measurement of the initial template DNA (Rutsaert et al., 2018, Strain et al., 2013, Trypsteen et al., 2016). *Alu*-PCR is used to measure integrated HIV-1 DNA by using a non-kinetic preamplification step that uses primers to bind genomic *Alu* elements and HIV-1 *gag* sequences followed by a kinetic PCR which quantitates HIV-1 long terminal repeats (Liszewski et al., 2009, O'Doherty et al., 2002). The quantitative viral outgrowth assay (qVOA) is often described as the gold standard in terms of measuring replication-competent virus (Bruner et al., 2015). Typically, $CD4^+$ T cells are isolated from blood using a negative selection method. The cells are plated in limiting dilution, to allow for quantification, and stimulated with phytohemagglutinin (PHA) and irradiated, allogeneic PBMCs from a non-infected donor. $CD8^+$ T cells are depleted from the

PBMCs so as not to suppress viral replication. After several weeks replication competent HIV-1 is detected by ELISA for the HIV-1 p24 antigen (Barton and Palmer, 2016, Siliciano and Siliciano, 2015).

We have developed a new method approach whereby cells harbouring replication-competent virus across a reservoir site can be characterized. This utilizes single-cell RNA-Seq (Bradley et al., 2018, Rato et al., 2017), in particular the Seq-Well method (Gierahn et al., 2017). Seq-Well is a high throughput and easily portable method where many cells can be sequenced in parallel from low input samples. Here we apply this method to lymph node cells from ART suppressed individuals to identify various subsets of cells that are infected with HIV-1. We used the anti-Env broadly neutralizing nanobody J3 (McCoy et al., 2012, McCoy et al., 2014) to enrich for HIV-1 infected cells using magnetic beads. Our data shows that there is diversity in the cell types that harbour HIV-1 transcripts in the lymph node cells from ART suppressed individuals. $CD4^+$ T cells and macrophages make up the majority of cells that contain these transcripts, with different subsets of T cells containing transcripts including memory T cells and circulating T cells. This goes against the hypothesis that the HIV-1 reservoir is restricted to a particular cellular subset or can be identified with a singular cell marker.

Results

Quantifying HIV-1 infection in lymph node cells

In order to quantify the size of the cellular HIV-1 reservoir in lymph nodes from ART suppressed study participants, a cross-sectional cohort was established in collaboration with clinicians at the Department of Cardiothoracic Surgery at Inkosi Albert Luthuli Central Hospital in Durban, South Africa. We selected lymph nodes from seven HIV-1 positive study participant and three HIV-1 negative study participants based on their viral load, number of years on ART, and TB status (Table 1). Liquid chromatography-mass tandem spectrometry (LC-MS-MS) was run on the matched PBMCs from each study participant to detect ART levels and whether the study participant was adhering to their regimen (Table 2).

Lymph node cells were fixed and stained with anti-HIV-1 p24 Gag antibody to quantify the number of HIV-1 infected cells. This was done for each study participant. The percentage of $p24^+$ cells varied between study participants with a range between 0.034% and 0.16% (Figure 1A). We also determined the percentage of cells expressing Env by staining with the anti-Env broadly neutralizing antibodies 10-1074 and J3 and found the percentage of anti-Env positive cells to be slightly higher than those expressing anti-HIV-1 p24 Gag. This slight difference could be the

result of background staining of the J3 nanobody (Figure 1B). We optimized anti-Env antibody detection of infected cells using *in vitro* infected CD4$^+$ T cells that were infected with 89.6, a dual tropic virus. CD4$^+$ T cells were either infected with the original strain or a blue fluorescent protein (BFP) strain. The cells were stained with anti-HIV-1 p24 Gag and anti-Env J3 and fluorescence of J3 versus BFP measured and p24 versus J3 (Figure 1C). Env expression on the cell surface of cells infected with BFP-tagged virus was found to be decreased compared to cells infected with non-BFP tagged virus. This could be as a result of the BFP gene causing a disruption of transcription and therefore attenuating Env expression.

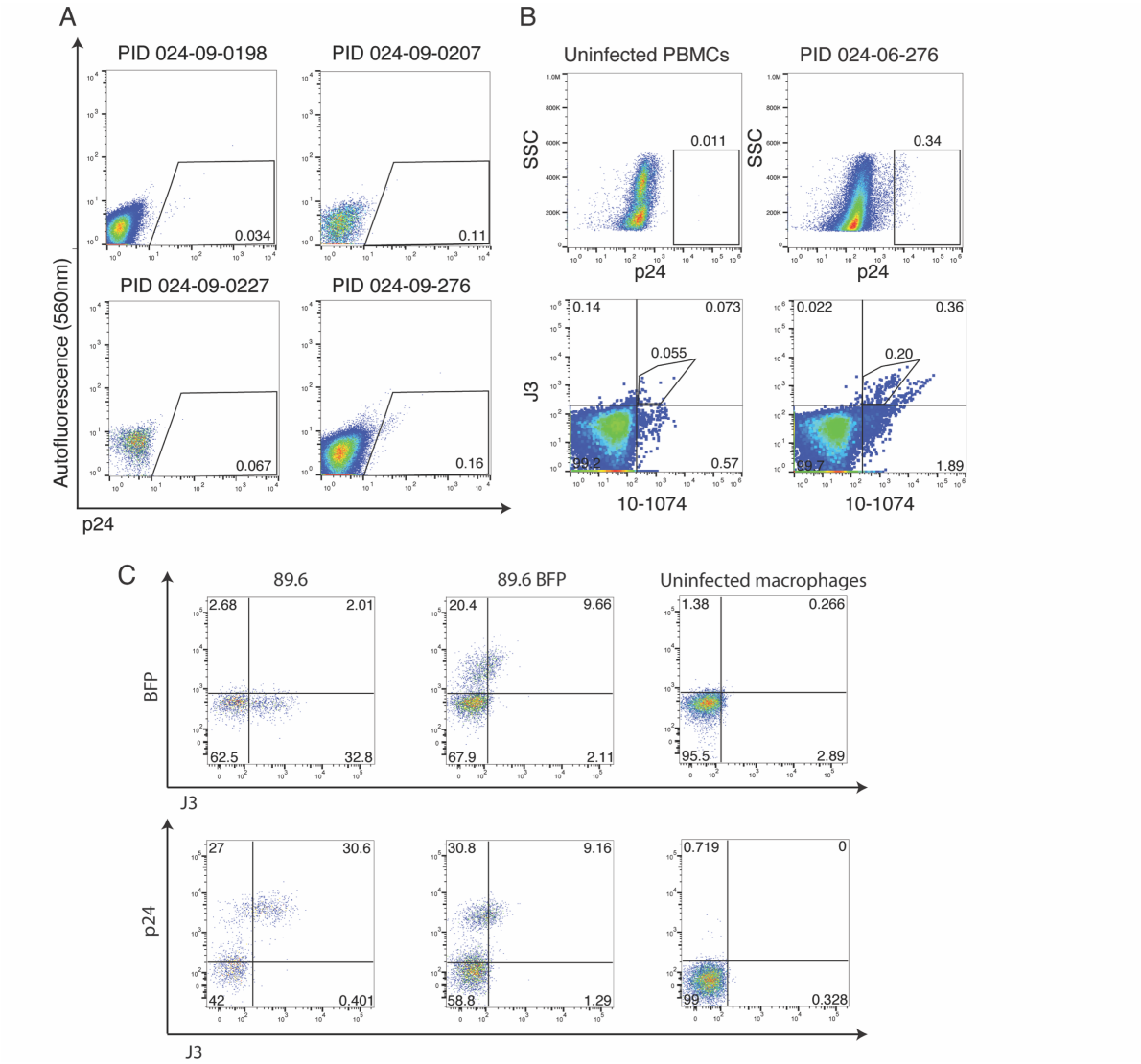

Figure 1. Quantifying HIV-1 infection in lymph node cells. (A) Anti-HIV-1 p24 Gag staining of lymph node cells from HIV-1 positive study participants to quantify the size of the HIV-1 reservoir in each individual. (B) Anti-HIV-1 p24 Gag staining of lymph node cells from PID 024-

09-0276 compared with staining using the broadly neutralizing anti-Env antibodies 10-1074 and J3. (C) Macrophages infected with either 89.6 or BFP tagged 89.6 compared to uninfected macrophages in terms of BFP fluorescence or anti-HIV-1 p24 Gag staining versus anti-Env J3 staining.

Optimization of anti-Env J3

The anti-Env nanobody J3 has typically been used as a broadly neutralizing antibody. Here, we optimized it for cellular staining with applications for both flow cytometry and magnetic-activated cell sorting (MACS) to select for HIV-1 infected cells. *In vitro* infected lymph node cells expressing surface HIV-1 envelope protein were stained with the anti-Env broadly neutralizing antibodies 10-1074 and J3, that were conjugated to fluorophores. The cells were infected with NL4-3 with a BFP tag. Enrichment was observed in comparison to staining of PBMC cells from an HIV-1 negative study participant when cells were acquired (Figure 2A). To verify that the anti-Env antibodies were selecting for HIV-1 infected cells, we sorted both the negative and positive fractions from the HIV-1 infected cells and the HIV-1 negative cells as a true negative control. A nested PCR was run on the cDNA from these cells to amplify a ~200bp region of the HIV-1 reverse transcriptase gene and gel electrophoresis to visualize the results. We found that the combination of J3 and 10-1074 enriched for HIV-1 infected cells. We also sorted BFP^+ and BFP^- cells to serve as controls (Figure 2B).

We then tested the staining on lymph node cells from the *in vivo* infected study participant PID 024-09-0276 (Figure 2C). We found that the combination of 10-1074 and J3 enriched for HIV-1 infected cells. However, we found that the population of cells, when sorted, was very fragile and that the data generated by running Seq-Well on these cells was of low quality. Therefore, due to cell loss we optimized staining with J3 for magnetic-activated cell sorting using a biotinylated version of the J3 nanobody and enriched for HIV-1 infected cells. We ran Seq-Well on the enriched fraction of cells.

We compared the number of unique molecular identifiers (UMIs), which are short sequences that are added to next generation sequencing libraries to identify and quantify mRNA, across all the samples that were sequenced and found comparable depth between them indicating sufficient sequencing. However, the total number of cells that were sequenced across all samples was variable and sample dependent. We then compared the number of UMIs between the J3 enriched fractions from each sample and found variability, which could be a direct result of the number of HIV-1 infected cells that were expressing *Env* on their surfaces as well as non-specific binding of uninfected cells to the J3 magnetic beads. The cell count from the J3 enriched fractions varied

greatly across the samples. The HIV-1 negative samples had some background indicating the level of non-specific binding that occurs. However, one sample which was antibody enriched for HIV-1 infected cells (PID 024-06-0312) had fewer cells with HIV-1 transcripts than the HIV-1 negative sample. This could be due to poor quality of the cells recovered rather than the lack of HIV-1 infected cells (Figure 2D). The *t*-distributed stochastic neighbor embedding (*t*-SNE) plot, colored by PID, showing the clustering of cells from the J3 enriched Seq-Well arrays (Figure 2E).

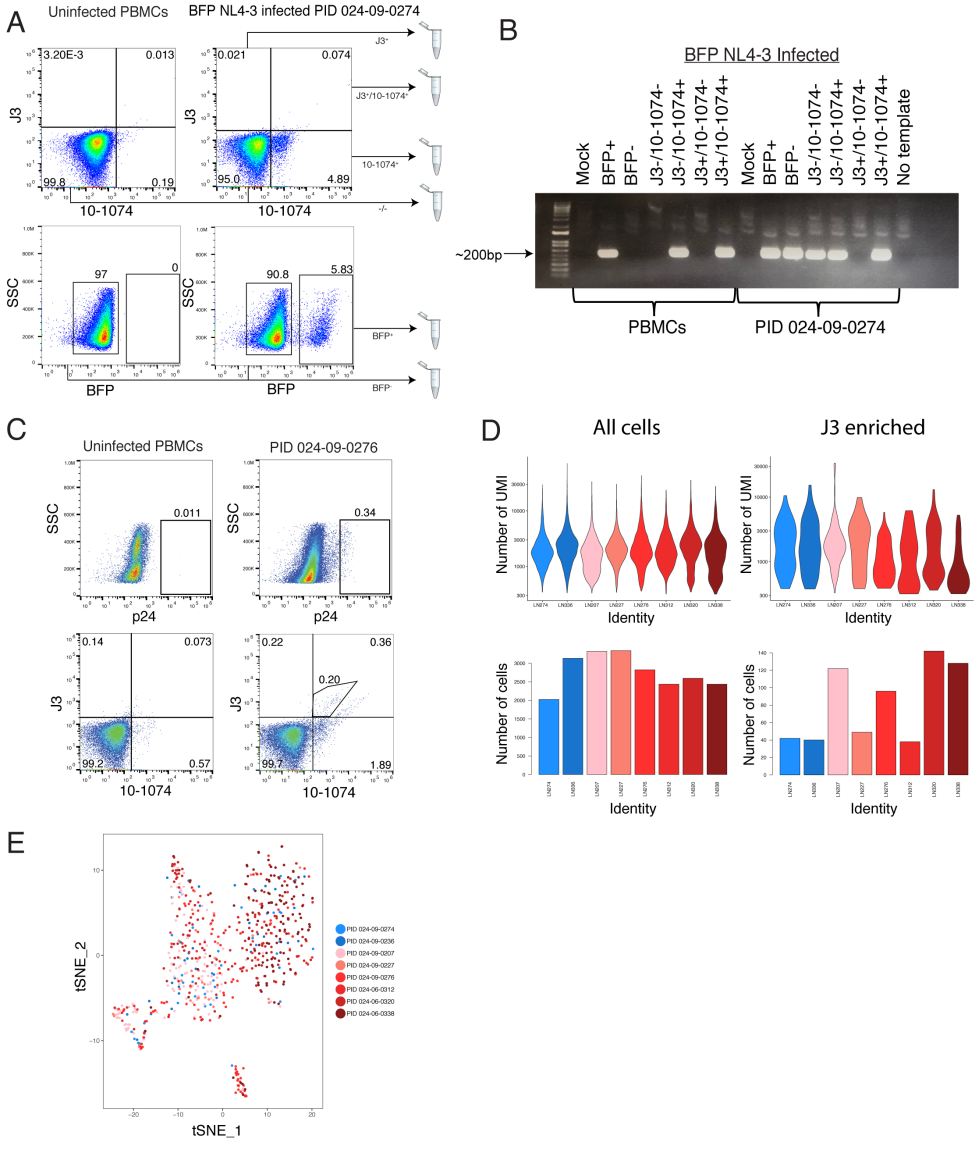

Figure 2. Optimization of anti-Env J3 detecting HIV-1 infected cells. (A) Enrichment of HIV-1 infection in lymph node cells, from PID 024-09-0274, infected with BFP-tagged NL4-3 and stained with anti-Env J3 and 10-1074. Cell populations were sorted (1000 cells per population), total RNA was extracted, and cDNA synthesized. (B) Gel electrophoresis of nested PCR

amplicons from the HIV-1 reverse transcriptase region (~200bp fragment) amplified from each sorted population. (C) Enrichment of HIV-1 infection in HIV-1 infected lymph nodes from PID 024-09-0276 using anti-Env 10-1074 and J3 and uninfected PBMCs as a negative control. (D) Violin plots of the number of UMIs compared between all samples as well as the number of UMIs per J3 enriched array across all samples. In parallel the number of cells between all samples and for the J3 enriched arrays was compared. (E) *t*-SNE plot of the cells from the J3 enriched arrays colored by study participant.

Single cell RNA-sequencing of in vitro infected cells to identify HIV-1 infected cells

The HIV-1 reservoir resides in tissues niche such as the lymph nodes, where it is able to persist in the presence of ART. Both intrinsic drivers, such as ongoing transcription of HIV-1 within cells, and extrinsic drivers, such as cell-to-cell spread, allow for the persistence of the virus (Figure 3A). Identifying gene signatures that indicate whether a cell is infected with HIV-1 or not is critical to determining *in vivo* what cells harbor HIV-1. We used *in vitro* infected Rev-CEM-E7 cells expressing GFP upon infection (Boulle et al., 2016). Cells were infected with wild-type NL4-3 virus, NL4-3 virus with the L100I drug resistant mutation, or a combination of both of these viruses. The cells from each infection were single-cell sorted into 96-well plates and single-cell RNA-sequencing run on the cells. Each cell expressed HIV-1 as indicated by GFP fluorescence (Figure 3B).

Upon sequencing of the single cells, we looked at the reverse transcriptase region of the genome where drug resistant mutations, such as L100I, occur. We were able to use HIV-1 transcripts sequenced by single-cell RNA-Seq to differentiate between the cells infected with wild type NL4-3, those infected with the NL4-3 L100I mutant, and those infected with a combination of both (Figure 3C and D). Analysis of the single-cell data comparing gene expression between GFP positive and GFP negative cells indicated distinct gene signatures between HIV-1 positive and negative cells (Figure 3E). This differential gene expression correlated with the abundance of HIV-1 infection in each cell. An example of two genes that play a role in HIV-1 infection that were looked at in terms of their expression were macrophage migration inhibitory factor (MIF) and myxovirus resistance-2 (MX2) (Figure 3F). MIF has been shown to be found at increased levels in the plasma from HIV-1 infected individuals and HIV-1 infected cells release a larger amount of MIF than uninfected cells (Regis et al., 2010). Our single-cell RNA-Seq data shows that MIF expression is increased in HIV-1 infected cells and therefore was used as a gene signature of HIV-1 infection. MX2 has been found to inhibit HIV-1 infection post entry and was also tagged as a gene signature of HIV-1 infection due to its low expression in the HIV-1 infected

cells, indicating that the virus may have been able to override the inhibition mechanism (Goujon et al., 2013) (Figure 3E).

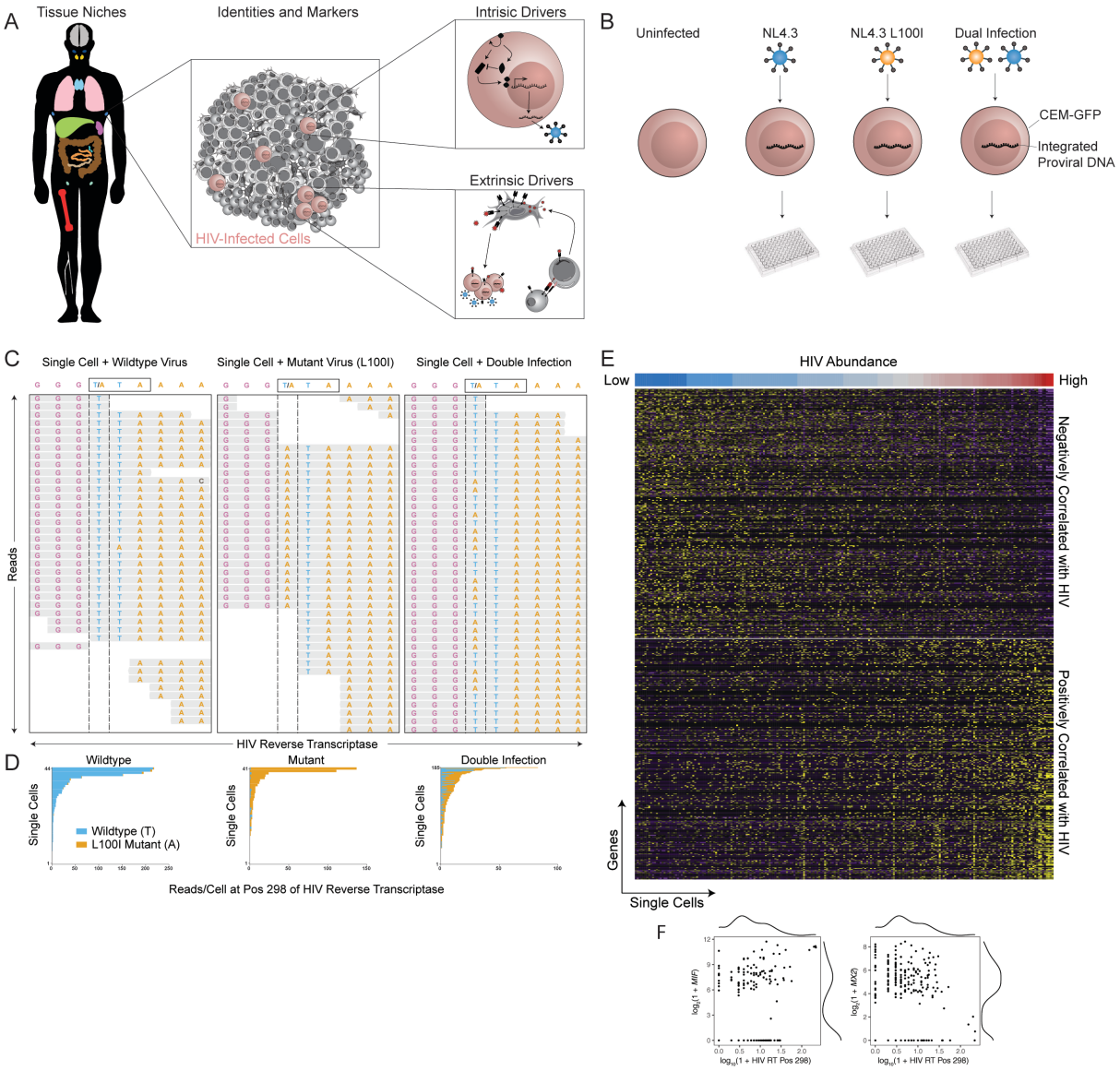

Figure 3: Proof of concept in *in vitro* infected Rev-CEM-E7s. (A) Schematic of the lymph node as the HIV-1 reservoir and what the drivers of reservoir persistence are. (B) Rev-CEM-E7 cells were infected with NL4-3, NL4-3 with the L100I drug resistant mutation, or a dual infection of both viruses. GFP cells were single cell sorted into plates. (C) The sequencing depth was able to recover the virus type and indicated that the depth was sufficient for detecting HIV-1 transcripts. (D) Number of reads per cell of the reserve transcriptase region of the HIV-1 genome indicating sufficient reads to identify HIV-1 infection. (E) Variance in genes that were correlated

with HIV-1 infection. (F) MIF and MX2 are two genes that are commonly associated with HIV-1 infection.

Detection of HIV-1 infection from in vitro infected HIV-1 negative lymph node cells using single-cell RNA-Seq

To optimize detection of HIV-1 transcripts from *in vivo* infected lymph node cells, we first optimized Seq-Well on *in vitro* infected lymph node cells from the HIV-1 negative study participants. The cells were stimulated with anti-CD3 and anti-CD28 antibodies for 5 days before spin infection with BFP tagged NL4-3 or BFP tagged JR-CSF. The cells were infected for 48 hours (Figure 4A). We validated the level of infection by measuring the number of $p24^+$ cells using anti-HIV-1 p24 Gag staining before sorting the BFP positive and negative fractions for both infections as well as the mock infected cells. We also saved an unsorted fraction of cells from each infection as well as the mock infected, for any downstream experiments that could be required for validation. Seq-Well was run on all of the conditions (Figure 4B). The transcripts were aligned to the Clade C consensus sequence and a cut-off determined by the number of transcripts from the mock infected lymph node cells that aligned to the sequence, to reduce the number of non-HIV-1 transcripts in the human genome with similarity to HIV-1 aligning to HIV-1. We obtained a clear population of cells with UMIs aligning to HIV-1 in the infected conditions (Figure 4C).

The UMIs that aligned to the Clade C consensus were log-normalized and 70% of BFP positive NL4-3 infected cells contained transcripts that aligned to the Clade C consensus sequence (Figure 4D). We then compared percentage of HIV-1 transcripts per condition with the percentage of anti-HIV-1 p24 Gag positive cells in each condition and found a 0.99 correlation (Figure 4E). Clustering of conditions by *t*-SNE showed clear clustering of sorted NL4-3 infected BFP positive cells compared with mock, NL4-3 negative cells, and sorted NL4-3 infected BFP negative cells (Figure 4F). Due to the percentage of infection by BFP (5.41%), the frequency of uninfected cells in the unsorted NL4-3 infected cells would be relatively high and therefore a likely reason as to these cells clustering more with the mock infected and NL4-3 infected BFP negative cells.

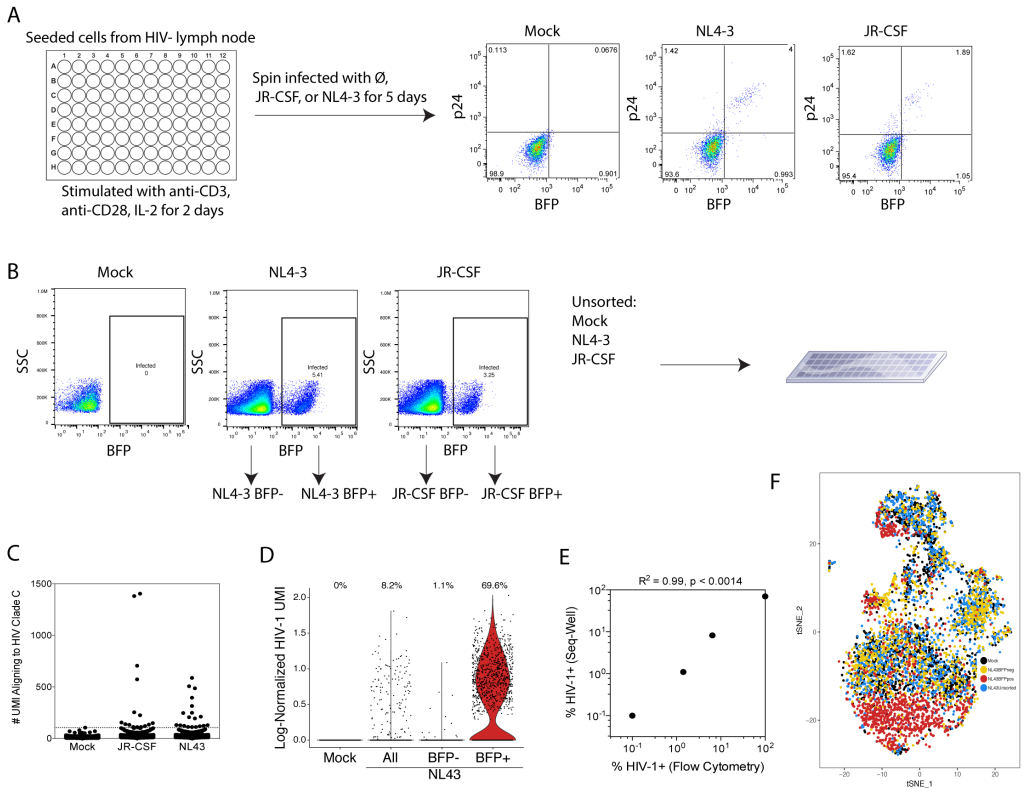

Figure 4: Resolution of HIV-1 infection from *in vitro* infected HIV-1 negative lymph node cells. (A) HIV-1 negative lymph node cells were stimulated with CD3 & CD28 antibodies for two days and spin infected with BFP NL4-3 or BFP JR-CSF. The infection was run for five days and validated with anti-HIV-1 p24 Gag staining. (B) BFP positive and negative fractions were sorted. Sorted and unsorted cells were loaded onto arrays for Seq-Well processing. (C & D) Resolution of HIV-1 alignment in infected cells compared to mock infected cells. (E) Correlation of HIV-1 infected cells recovered from Seq-Well compared to those recovered from flow cytometry based on p24 staining across all conditions (unsorted: mock, NL4-3 positive, JR-CSF positive; sorted: BFP NL4-3 positive, BFP NL4-3 negative, BFP JR-CSF positive, BFP JR-CSF negative). (F) *t*-SNE plot of each NL4-3 infection condition coloured by condition.

Amplification of HIV-1 virus from lymph node cells

To improve alignment of single cell RNA-sequencing data for HIV-1 reads for each study participant we cloned full length HIV-1 virus from the lymph node cells (Figure 5A). This was done by amplifying the genome in two parts, excluding the long terminal repeat (LTR) regions, from RNA and DNA that were extracted from the lymph node cells. We used the Phusion High-Fidelity DNA Polymerase (NEB) as opposed to previous methods (Deymier et al., 2014) that used Q-5 Hot-Start Polymerase as we found that the Phusion enzyme was better able to amplify HIV-1 from lymph node cells than Q-5, which was optimized for amplification from plasma. We were able to successfully amplify HIV-1 sequences from each study participant and combine the first and second half genome amplicons to form one genome which was then deep sequenced (Figure 5B). Each HIV-1 genome was used during alignment of the Seq-Well data generated from each lymph node. This improved alignment as opposed to aligning to the consensus HIV-1 Clade C genome (Figure 7D). We created a phylogenetic tree using each amplified HIV-1 genome in relation to the Clade C consensus sequence and NL4-3 using the neighbour-joining alignment type (Figure 5C). This was to ensure that there was no NL4-3 contamination as well as to look at the difference between each HIV-1 genome in relation to the Clade C consensus sequence, created from HIV-1 genomes amplified from study participants around Durban, South Africa.

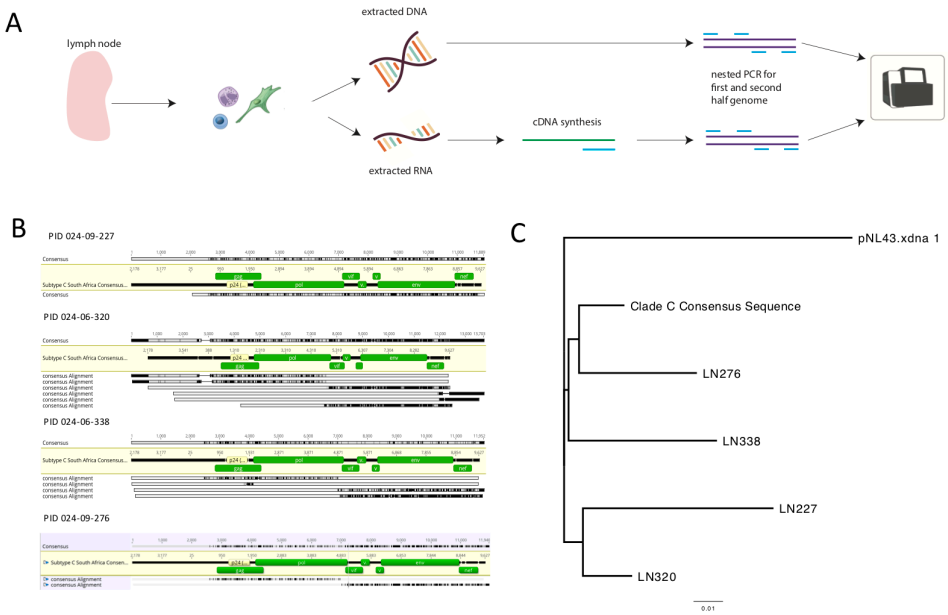

Figure 5: Amplification of HIV-1 virus from lymph node cells. (A) Lymph nodes were cellularized and total RNA and DNA extracted. cDNA was synthesized from the RNA and both DNA and cDNA underwent a nested PCR. To amplify the HIV-1 genome from lymph node cells, primers were designed for the first half and second half of the genome, therefore amplifying it in two parts. (B) Consensus sequences from the amplified HIV-1 genomes. Several amplicons were sequenced and aligned for some of the PIDs, indicated by the grey bars. The black lines within the grey bars indicate areas of alignment. (C) Phylogeny of the amplified genomes compared to the consensus HIV-1 Clade C genome and NL4-3.

Enrichment of cells most likely to harbor HIV-1

We found that using flow cytometry to sort cell populations for downstream use in Seq-Well was causing us to lose the population of cells most likely to harbour HIV-1 and possibly be the HIV-1 cellular reservoir in lymph nodes. Therefore, we used magnetic-activated cell sorting to enrich for $CD4^+$ cells, deplete $CD19^+$ cells, and enrich for HIV-1 infected cells using the anti-Env J3 nanobody (Figure 7). This also eliminated the majority of dead cells as determined by cell counting with Trypan Blue (ThermoFisher) pre and post magnetic sorting.

Cell phenotyping and identifying HIV-1 infected cells

We chose Seq-Well as the single-cell platform to run RNA sequencing on our lymph node cells as it is ideal for low abundance samples and many cells can be sequenced in parallel. We initially ran a pilot experiment with lymph node cells from PID 024-09-0198. Cells were loaded onto arrays, as mentioned above, and sequenced (Figure 6A). The analysis of the Seq-Well data showed different cell populations that are known to be found in lymph nodes including several different subtypes of T cells including memory-like T cells and activated T cells. Populations of macrophages and antigen presenting dendritic cells were observed as well as a substantial B cell population, as expected (Figure 6B). Using complete gene lists generated from the *t*-SNE analysis of the single-cell data, we were able to identify various lymphocytes typically found in the lymph node environment (Figure 6C) and the relative contribution of these cell types to the total number of cells and genes recovered from the single-cell data. The major cell types identified were highlighted with hallmark expressed genes such as TNFRSF4 in antigen presenting cells, PAX5 and BANK1 in B cells, and MAL and NKG7 for T cells (Figure 6D). Transcripts generated from

the single-cell data were also aligned to the HIV-1 Clade C consensus sequence so that we could identify which cells were infected (Figure 7D).

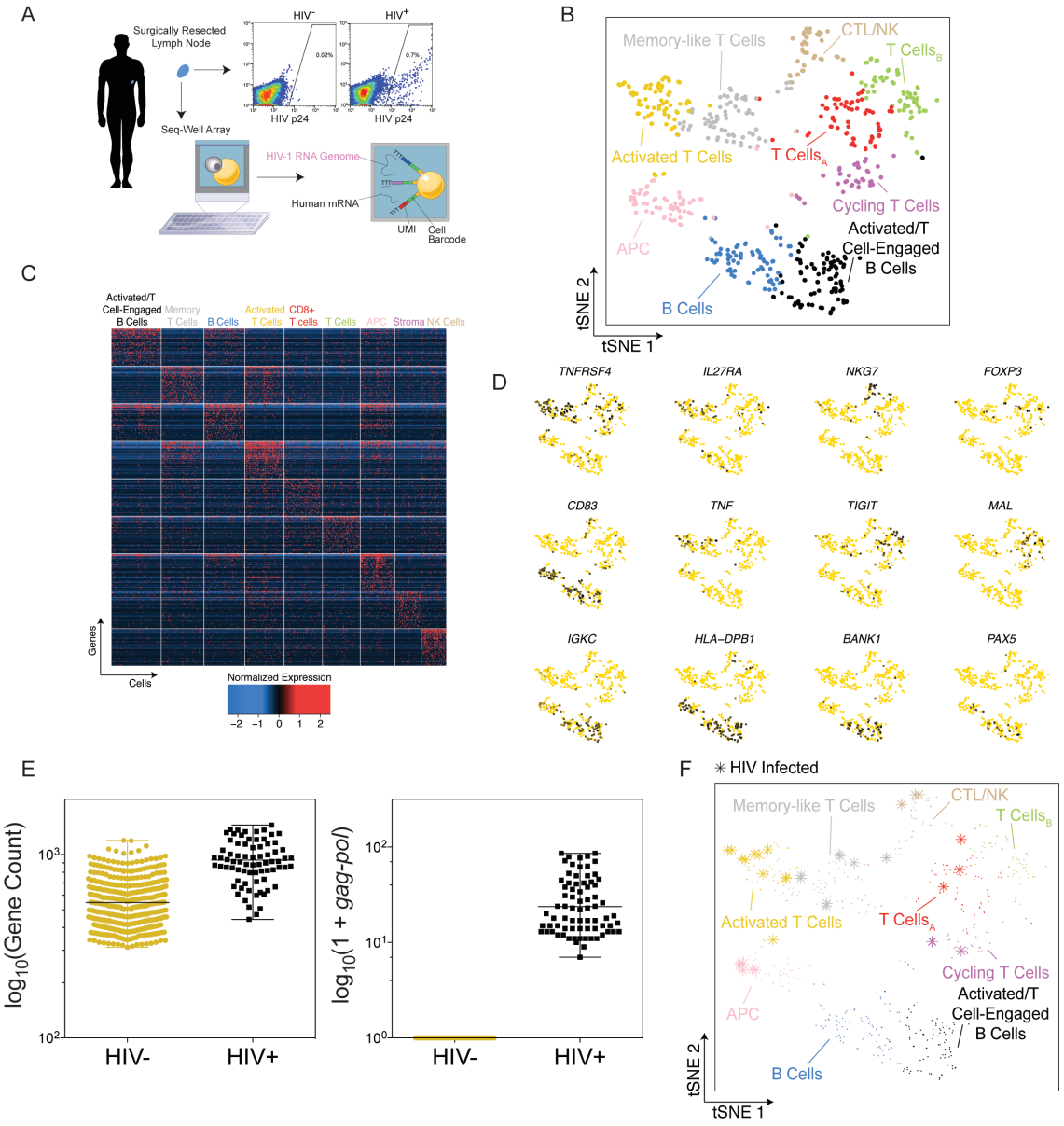

Figure 6: Cell phenotyping and identifying HIV-1 infected cells. (A) Schematic of workflow from lymph node to transcriptome. (B) *t*-SNE of cells from PID 024-09-0198 clustered and colored by cell type. (C) Heatmap of top 50 genes per identified cell cluster. (D) Highly expressed genes per cell type allowing for the identification of cell types. (E) Number of genes in HIV-1 negative compared with HIV-1 positive cells. We then compared the number of *gag-pol* transcripts in the HIV-1 negative cells to those in the HIV-1 positive cells to establish a threshold

for identifying truly HIV-1 infected cells as false positives may arise as a result of redundancy in the human genome. (F) We replotted the *t*-SNE from (B) and marked the cells that had HIV-1 transcripts using stars.

Total gene count between the HIV-1 negative cells and HIV-1 positive cells varied, with HIV-1 positive cells having a higher $\log_{10}$ normalized gene count. However, $\log_{10}$ normalization reads mapped to the *gag-pol* gene between HIV-1 positive and HIV-1 negative cells indicated a high gene count in the HIV-1 positive cells and the complete absence of *gag-pol* in the HIV-1 negative cells (Figure 6E). We were able to plot these cells in relation to the uninfected cell subtypes on a *t*-SNE plot (Figure 6F). This gave us an indication of where to the HIV-1 cellular reservoir in lymph nodes may be.

We then ran the optimized enrichment method for HIV-1 infected cells on the lymph node cells from six HIV-1 positive study participants and three HIV-1 negative study participants. A total of 22,114 cells were recovered from the Seq-Well data with 39,385 unique genes expressed (Figure 7A). The large number of genes recovered is as a result of alternative splicing. The lymph node cells that were enriched for HIV-1 infection using the anti-Env J3 magnetic column all clustered together from each lymph node (Figure 7B). Amongst the cell types that had transcripts that included the *env* and *gag-pol* genes were macrophages and $CD4^+$ T cells. The exact subsets of $CD4^+$ T cells are yet to be determined through further data analysis (Figure 7C).

Flow cytometric validation of HIV-1 infected cell types

To validate cell types that indicated the presence of viral transcripts in the transcriptomic data generated, we performed flow cytometry on lymph node cells from PID 024-09-0276 and included markers of HIV-1 infection including the anti-HIV-1 p24 Gag intracellular stain as well as the anti-Env broadly neutralizing antibodies 10-1074 and J3. We initially looked at the overall size of the reservoir as determined by these stains and the overlap between p24 positive cells and cells positive for the anti-Env antibodies. p24 positive cells constituted 0.069% of total cells, and cells doubly positive for anti-Env J3 and anti-HIV-1 p24 Gag constituted 0.052% of the total number of cells (Figure 7E). These frequencies were consistent with previously reported frequencies of HIV-1 infected cells in rectal biopsies in the face of ART as determined by HIV-1 DNA copies (Bruner et al., 2015, Bruner et al., 2019, Eriksson et al., 2013, Omondi et al., 2019).

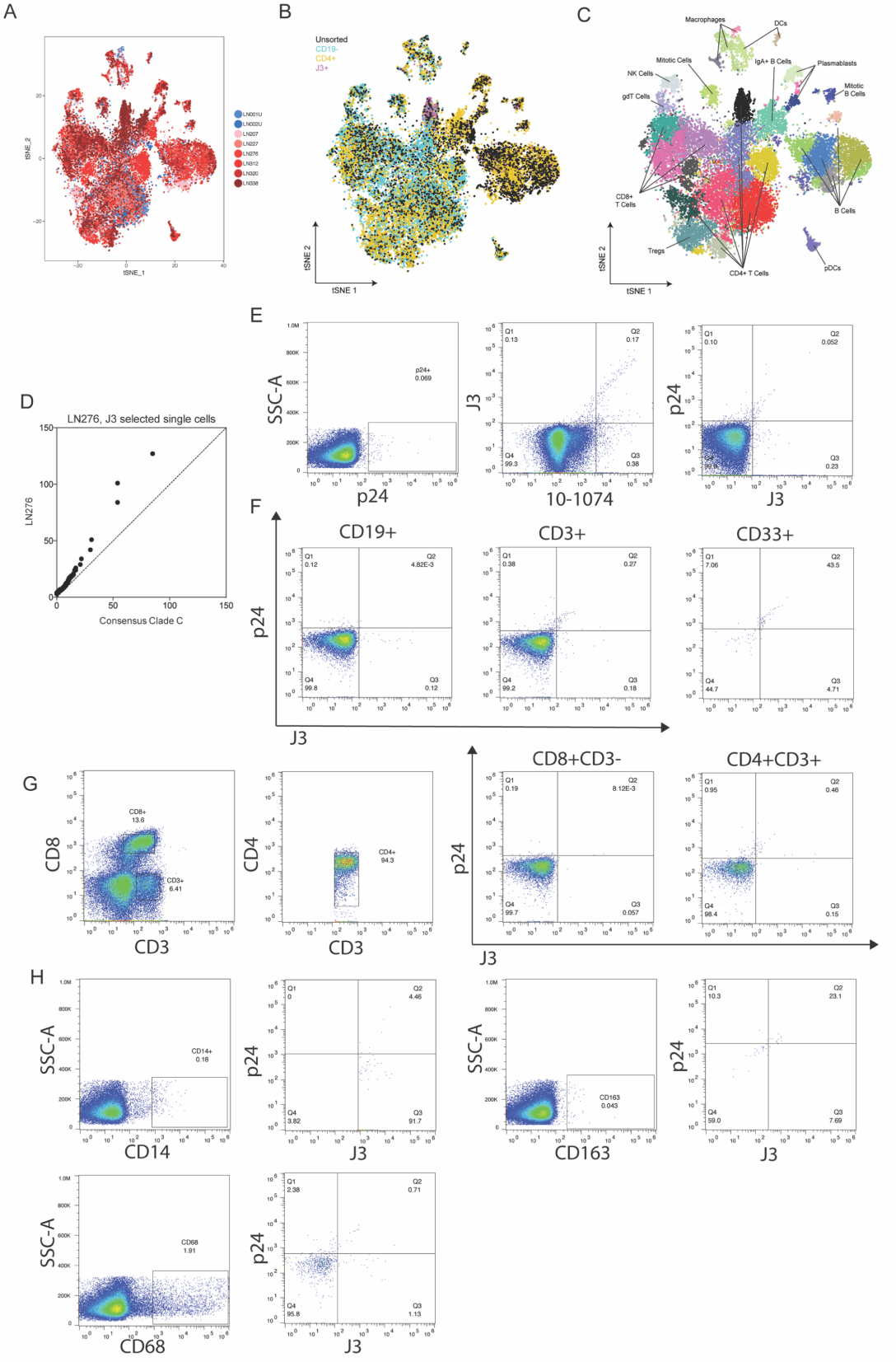

Figure 7: Identification of HIV-1 infected cells in the context of the HIV-1 reservoir in lymph nodes. (A) Our optimized Seq-Well workflow lead to a *t*-SNE plot displaying 22,114 cells

across all study participants (excluding PID 024-09-0219, as this had yet to be sequenced at the time of analysis). (B) The total number of cells coloured by experimental condition, with a defined J3 enriched population. (C) Cells clustered and coloured by cell type. (D) Increased alignment of J3 enriched single cells to the HIV-1 genome from PID 024-09-0276 than to the Clade C consensus sequence. (E-H) Flow cytometry validation of infected cell types using anti-HIV-1 p24 Gag and anti-Env J3.

We then looked at the difference in p24 and J3 double positive cells in the main cell lineages found in lymph nodes including $CD19^+$ cells, $CD3^+$ cells, and $CD33^+$ cells (Figure 7F). As to be expected, we found only background levels of p24 or J3 positive cells in the $CD19^+$ cell population. However, 0.27% of $CD3^+$ cells and 43.5% of $CD33^+$ (Mehraj et al., 2014, Rodrigues et al., 2017) cells were positive for both p24 and J3. T cells and myeloid lineage cells, such as macrophages, have been shown to form part of the HIV-1 reservoir (Castellano et al., 2017, Siliciano and Siliciano, 2015). Given the percentage of p24 and J3 positive cells in the $CD3^+$ T cell population, we identified infected cells within the $CD3^+$ population by gating out $CD8^+$ cells and examining $CD3^+CD4^+$ T cells (Figure 7G). We found that 0.46% of these cells were double positives. This indicates a significant $CD4^+$ T cell reservoir in the lymph node compartment. We note that while we have not tested transmission of HIV-1 from these cells, we observed both Gag and Env proteins, indicating the virus is functional.

The myeloid lineage cells showed a substantial fraction of HIV-1 infection (43.5%) by the CD33 marker. As the transcriptomic data also indicated HIV-1 infection in myeloid subsets, we looked further into different populations of myeloid cells, including $CD14^+$ cells, $CD68^+$ cells, and $CD163^+$ myeloid linage cells (Figure 7H). These populations of cells were rare (0.043 to 1.91% of total lymphocytes). We found that 4.46% of $CD14^+$ cells, 0.71% of $CD68^+$ cells, and 23.1% of $CD163^+$ cells were doubly positive for p24 and J3.

Discussion

Identifying the HIV-1 reservoir in the lymph node compartment from ART suppressed individuals required developing various assays to enrich for HIV-1 infected cells (Figure 1A). Previous research has identified cells expressing markers such as CD32a (Descours et al., 2017) and PD-1 (Banga et al., 2016) as composing the cellular HIV-1 reservoir. However, we hypothesized that the HIV-1 reservoir may also be present in other cell subsets or differ between individuals. We therefore we chose to employ the Seq-Well method of single-cell RNA

sequencing, which is unbiased in terms of selecting for cell types, to identify infected cell subtypes.

We assembled a cross-sectional cohort of lymph nodes from study participants, who were fully suppressed on ART (viral load of <40 viral copies per ml), except for PID 024-06-0338 who had a viral load of 58 viral copies per ml. We quantified ART regimen by LC/MS/MS and all individuals were in range for EFV/TFV/FTC or LPV/r/3TC suppression (Shen et al., 2008), except for PID 024-06-0312 who was below level of detection for TFV and had low levels of FTC. We initially quantified the size of the HIV-1 reservoir in the lymph node cells from each ART suppressed study participant by anti-HIV-1 p24 Gag staining and through the two anti-Env broadly neutralizing antibodies 10-1074 and J3. We established that the number of infected cells was less than 1% and therefore needed to be enriched for should we want to accurately identify the cellular subsets compromising the HIV-1 reservoir (Figure 1).

A common marker of HIV-1 infection used during flow cytometry is anti-HIV-1 p24 Gag. However, as this involves paraformaldehyde fixing, it rapidly degrades the fragile RNA, and when de-crosslinked, very little RNA is recovered and the most fragile cells that are most likely infected with HIV-1, are lost. An alternative approach was to use the anti-Env broadly neutralizing nanobody J3 which we optimized in conjunction with the anti-Env broadly neutralizing nanobody 10-1074 (Caskey et al., 2017). Optimization of the anti-Env broadly neutralizing nanobody J3 for flow cytometry allowed us to enrich for HIV-1 positive cells when these cells were sorted and Seq-Well performed. However, we found that sensitive populations of cells were being lost in comparison to unsorted cells and therefore chose to optimize J3 enrichment using magnetic activated cell sorting. We optimized this through *in vitro* infections of lymph node cells with BFP-tagged NL4-3 and BFP-tagged JR-CSF virus and this yielded HIV-1 infected cells as quantified through PCR amplification of the reverse transcriptase region of the cells.

Before using *in vivo* infected lymph node samples, we performed a proof of concept experiment, to determine whether we could sequence single cells to sufficient depth to identify HIV-1 transcripts. Using two strains of NL4-3 virus, one with a drug resistant L100I mutation in the reverse transcriptase region, we infected Rev-CEM cells and were able to determine the virus with which each single cell was infected based on the presence or lack of the L100I mutation (Figure 3). We were also able to show gene expression in cells that correlated with HIV-1 infection and expression that negatively correlated with HIV-1 infection. Recent studies have shown this in a similar model (Bradley et al., 2018, Cohn et al., 2018), where they were able to

look at the regulation of HIV-1 transcription in infected $CD4^+$ T cells or reconstruct the viral sequence in each cell from the single-cell RNA sequencing data.

Once we were able to identify viral transcripts in single cells, we optimized our Seq-Well protocol with *in vitro* infected HIV-1 negative lymph node cells. We used a BFP fluorescent NL4-3 virus and BFP positive and negative cells as well as mock infected cells. Due to redundancy in the human genome, the short reads generated from Seq-Well may non-specifically align to the HIV-1 genome. This experiment allowed us to determine a cut-off based on comparing the number of mock infected cells that aligned to the HIV-1 genome in comparison to the number of BFP positive cells that aligned.

To improve HIV-1 alignment of Seq-Well reads, we successfully cloned full-length genomes from the lymph node cells from each HIV-1 positive study participant. We aligned the single-cell RNA sequencing data to both the HIV-1 Clade C consensus sequence and the cloned HIV-1 genome.

The pilot experiment on HIV-1 infected lymph node cells from PID 024-09-0198 indicated that various T cell subsets such as activated T cells, cycling T cells, and memory T cells, as well as antigen presenting cells (Figure 6), were infected with HIV-1. Previous research has shown some of these cell types to harbour the latent HIV-1 reservoir such as memory T cells (Chomont et al., 2009, Murray et al., 2016) and activated T cells (Murray et al., 2014). Macrophages have also been shown to form part of the HIV-1 reservoir (Baxter et al., 2014, Castellano et al., 2017). This indicates that they HIV-1 cellular reservoir in lymph nodes may not be specific to a singular cell type or cellular marker.

Our optimized protocol, which we ran on eight lymph node samples (two HIV-1 negative and six HIV-1 positive), yielded 22,14 cells and recovered 39,385 unique genes. We identified a population of cells that were enriched for infection using the J3 Env nanobody and also contained *gag-pol*, *nef*, and *vif* transcripts. These cells were characterized as $CD4^+$ T cells, but deeper characterization of the exact subset is still needed. We also observed HIV-1 transcripts in macrophages. To validate the transcriptomic data, we ran flow cytometry using cellular markers, anti-HIV-1 p24 Gag staining, and J3 anti-Env staining, to verify the infected cell subtypes. We found that there were cells in the $CD3^+$, and $CD4^+$ T cell populations and cells in the $CD33^+$, $CD14^+$, $CD68^+$, and $CD163^+$ myeloid lineage cell populations.

Cells of myeloid origin have been previously reported to harbour the HIV-1 reservoir; however, the mechanisms of infection and persistence remain controversial. Research has shown that macrophages infect $CD4^+$ T cells (Abbas et al., 2015) or conversely that $CD4^+$ T cells infected

macrophages (Groot et al., 2011). Or that macrophages are able to maintain HIV-1 infection independently of $CD4^+$ T cells (Honeycutt et al., 2016). Therefore, the finding that the cellular HIV-1 reservoir in lymph nodes is partly harboured in cells of myeloid origin is a unique and novel find that requires further in-depth studies to understand how this cellular reservoir is formed and what the implications in terms of persistence and eradication strategies are.

Our data showed the cellular HIV-1 reservoir in lymph nodes from ART suppressed individuals is harboured in $CD4^+$ T cells and macrophages. The exact $CD4^+$ T cell subtypes still need to be defined through further analysis of the transcriptomic data. It is clear that the HIV-1 reservoir is not in one cellular compartment as previously suggested (Banga et al., 2016, Descours et al., 2017). This will make a functional cure more difficult as a variety of cell types will need to be targeted in order to reduce or eradicate the HIV-1 reservoir.

Methods

Study participants

To investigate the HIV-1 LN reservoir, have assembled a cohort of LN from individuals on suppressive ART undergoing indicated lung lobe resection or diagnostic procedures in collaboration with clinicians at the Department of Cardiothoracic Surgery at Inkosi Albert Luthuli Central Hospital in Durban, South Africa. Lymph nodes were obtained from study participants undergoing surgery for diagnostic purposes or complications of inflammatory lung disease. Informed consent was obtained from each participant. The study protocol was approved by the University of KwaZulu-Natal Institutional Review Board (approval BE024/09). We obtained matched samples of peripheral blood from each individual. We used peripheral blood to measure viral load to determine if the ART regimen was effective at suppressing peripheral viremia below the level of detection (<40 HIV-1 copies/ml using Roche AMPLICOR). We also measured the concentrations of currently used antiretroviral drugs (ARVs) in South Africa in the peripheral blood as well as in cells from the lymph node compartment. Measured ARVs were efavirenz (EFV), emtricitabine (FTC) and its triphosphorylated intracellular form, tenofovir (TFV), lopinavir and ritonavir (LPV/r), lamivudine (3TC), and nevirapine (NVP). Levels of ARVs were an independent measure of the ART regimen and study participant adherence.

Primary cells

Peripheral blood mononuclear cells (PBMCs, from HIV-1 negative donors) were isolated by density gradient centrifugation using Histopaque 1077 (Sigma-Aldrich) and cultured at a concentration of 10^6 cells/ml in complete RPMI 1640 medium. The RPMI 1640 medium was supplemented with sodium pyruvate, L-Glutamine, HEPES, non-essential amino acids (Lonza), 10% heat-inactivated FBS (GE Healthcare Bio-Sciences, Pittsburgh, PA) as well as IL-2 at 5ng/ml (PeproTech, Rocky Hill, NJ). Phytohemagglutinin at 10μg/ml (Sigma-Aldrich) was added to activate the cells.

Cells from lymph node samples, isolated from HIV-1 positive and negative study participants, were obtained by mechanical separation of the lymph nodes. The cells were frozen at 5×10^6 cells/ml in a solution of 90% FBS, 10% DMSO, and 2.5 μg/ml Amphotericin B (Lonza). Cells were stored in liquid nitrogen until use, then thawed and resuspended at 10^6 cells/ml in RPMI 1640 medium supplemented with sodium pyruvate, L-Glutamine, HEPES, non-essential amino acids (Lonza), 10% heat-inactivated FBS (GE Healthcare Bio-Sciences, Pittsburgh, PA) as well as IL-2 at 5ng/ml (PeproTech, Rocky Hill, NJ).

Cell lines

Rev-CEM cells were obtained from the NIH AIDS Reagent Program. They were cultured in R10 media (500ml RPMI 1640 supplemented with 5ml L-glutamine, 5ml HEPES, and 50ml FBS). We have subcloned the Rev-CEM-E7 clonal line which has >70% GFP expression upon infection (Boulle et al., 2016).

Infection

Lymph node cells and $CD4^+$ T cells from the HIV-1 negative study participants were stimulated with anti-CD3 antibody. The plate was prepared the day before. Coating buffer (8.4g $NaHCO_3$, 3.56g Na_2CO_3, 1L ddH_2O, pH to 9.5, sterile filtered for TC) was made and 2μg/ml (20μl) of anti-CD3 antibody (Biolegend, Cat # 317315) was added to the coating buffer and vortexed to mix. 400μl of coating buffer was aliquoted into each well of a sterile, non-TC treated24 well plate (Corning, Cat #3473) and sealed. The plate was incubated overnight at 4C°. The plate was washed twice with phosphate buffered saline (PBS) and then 400μl of R10/50 added to each well. The

lymph node cells were then thawed in 10ml of R10, centrifuged at 300g for 7 minutes and resuspended in 1ml of R10. The cells were counted for viability and resuspended at 2 x 10^6/ml with R10/15. 4µg/ml of anti-CD28 antibody (Biolegend, Cat # 302923) was added to the cells and then 400µl of cell were added to each well of the 24 well plate and incubated in a 37C° incubator. After five days the cells were harvested, counted, centrifuged at 300g for 3 minutes, washed with 15ml of R10, and resuspended at 10^6 in R10/50. The cells were transferred to a 96-well V bottomed plate and centrifuged at 2400rpm at 4C° for 3 minutes. The cells were washed with 200µl of R10 and resuspended in 85µl of R10/50 (R10 media with 1ng/ml IL-2 – recombinant IL-2) and transferred to a 96-well flat-bottomed plate. 40µl of the appropriate virus (JR-CSF BFP, NL4-3 BFP, or 89.6 BFP) was added to each well. Two wells were left for a mock infection per sample. The plates were centrifuged at 800g with an acceleration of 1 for 10 minutes. At 9 minutes the temperature was changed to 37C° and the time to 59 minutes. The plate was then placed a 37C° incubator for 3 hours. The cells were then transferred to a 96 well V bottomed plate and centrifuged at 2400rpm for 3 minutes and then resuspended in 100µl of R10/50 and transferred to a round-bottomed plate. The plate was placed in a 37C° incubator for 48 hours.

Rev-CEMs were infected with 2 x 10^8 NL4-3 viral copies/ml as previously described (Jackson et al., 2018) for a cell-free infection.

Staining and flow cytometry for anti-HIV-1 p24 Gag and anti-Env 10-1074 and J3

The 10-1074 monoclonal antibody was conjugated to Alexa Fluor 488 (ThermoFisher Scientific) and the J3 nanobody was conjugated to Alexa Fluor 647. Cells were stained in phosphate buffered saline (PBS) with 10% FBS (GE Healthcare Bio-Sciences, Pittsburgh, PA) and 1mM ETDA with 1µl of 10-1074 (4mg/ml) and 1µl J3 (50nM). Cells were acquired and sorted on a SH800S Cell Sorter and data analyzed using FlowJo 10.0.8 software. 1000 cells per population were sorted into a solution of RLT buffer (Qiagen) and 1% β-mercaptoethanol. RNA was extracted using the Qiagen RNeasy Plus Micro Kit and converted to cDNA. cDNA synthesis was performed in 20µl reactions following the SuperScript IV manufacturer's instructions (ThermoFisher Scientific) with primer 1.3MOD2Malbec (5' – GTGAGTGATGGTTGAGGTGTGAGTGAGAGCACTCA AGGCAAGCTTTATTGAGGC – 3'). Nested PCRs of the HIV-1 reverse transcriptase region were performed as described previously (Jackson et al., 2018). The amplified PCR products were visualized using gel electrophoresis as well as on a 4200 TapeStation System (Agilent).

To determine the number of HIV-1 infected cells from both the *in vitro* infected PBMCs and *in vivo* infected lymph node cells by anti-HIV-1 p24 staining, cells were then fixed and permeabilized using the BD Cytofix/Cytoperm Fixation/Permeabilization kit (BD Biosciences). Cells were then stained with anti-p24 FITC or PE conjugated antibody (KC57, Beckman Coulter, Brea, CA). Cells were then acquired on a SH800S Cell Sorter (SONY) and data analyzed using FlowJo 10.0.8 software.

Amplification of full-length HIV-1 genomes from lymph node cells

RNA was extracted using the Qiagen RNeasy Plus Micro Kit and converted to cDNA. cDNA synthesis was performed in 20µl reactions following the SuperScript IV manufacturer's instructions (ThermoFisher Scientific) with primer 1.3MOD2Malbec (5' – GTGAGTGATGGTTGAGGTGTGA GTGAGAGCACTCAAGGCAAGCTTTATTGAGGC – 3'). DNA was extracted using the Qiagen DNeasy Blood & Tissue Kit. The cDNA product and the DNA were then amplified using a nested PCR in 50ul reactions following the Phusion High-Fidelity DNA Polymerase manufacturer's instructions (NEB). The HIV-1 genome was amplified in two 4.5kb fragments as opposed to one 9kb fragment. The first half genome first round primers were as follows: SubC_R_For1 (5' – GTCTCTCTAGGTAGA CCAG – 3') and SubC_Int_Rev1 (5' – CCTTTCCAAATAGGGTCTC – 3'). The second half genome second round primers were as follows: SubC_Int_For1 (5' – CGGGTTTATTACAGAGAGACAGC – 3') and SubC_R_Rev1 (5' – TARAGCACTCAAGGCAAGC – 3'). The first half genome second round primers were as follows: R_For_2 (5' – GTCTCTCTAGGTAGACCAGATCTGAGC – 3') and Int_Rev_2 (5' – GTCTCTGCTGTCTCTGTAATAAACCC – 3'). The second half genome second round primers were as follows: Int_For_2 (5' – AGAGACAGCAGAGACCCTATTTGGAAAGG – 3') and R_Rev_2 (5' – TARAGCAC TCAAGGCAAGCTTTATTGAGG – 3'). Amplified products were visualised using gel electrophoresis. The products were gel purified using the Wizard SV Gel and PCR Clean-Up System. The samples were then prepped for sequencing using the Nextera XT DNA Library Preparation Kit (Illumina) and sequenced on a MiSeq System (Illumina). Sequences and phylogenies were assembled using Geneious (v11.0.5). The phylogenetic tree was created using the neighbour-joining method with the HKY85 genetic distances model.

Staining, magnetic activated cell sorting, and Seq-Well

Lymph node cells were thawed in 10mls of R10 media and centrifuged at 300g for 7 minutes and resuspended in 1ml MACS buffer (phosphate buffered saline, 2% heat-inactivated FBS (GE Healthcare Bio-Sciences, Pittsburgh, PA), and 1mM EDTA). Cells were counted and resuspended at different concentrations: 40,000 cells in 800ul for unsorted Seq-Well arrays; 500,000 cells for CD4 enrichment; 500,000 cells for CD19 depletion; and the remaining cells (>3 x 10^6) for J3 enrichment. The CD4 enrichment, CD19 depletion, and J3 enrichment tubes were centrifuged at 300g for 3 minutes. The cells for CD4 enrichment and CD19 depletion were separately resuspended in 80µl MACS buffer and 2.5µl FcR blocking reagent for 10 minutes at 4C° after which 20µl of CD4 MACS beads and 20µl CD19 MACS beads (Miltenyi Biotec) were added to their respective tubes and kept at 4C° for a further 15 minutes. The cells for J3 enrichment were resuspended in 50µl of MACS buffer and 1µl of J3 conjugated to biotin and incubated for 30 minutes at 4C°. 5ml of MACS buffer was then added to the cells, before they were centrifuged at 300g for 3 minutes to wash. Cells were resuspended in 90µl of MACS buffer and 10µl of streptavidin beads were added and then incubated at 4C° for 15 minutes. 2ml of MACS buffer was added to each condition after incubation, and the cells were centrifuged at 300g for 3 minutes before being resuspended in 500µl of MACS buffer. LD columns (Miltenyi Biotec) were used for the CD19 depletion and CD4 enrichment and LS columns (Miltenyi Biotec) were used for the J3 enrichment. The columns were prepared with three washes of 3ml MACS buffer. The cells were then passed through the columns and the columns washed three times whilst the columns were on the MACS Manual Separators magnet (Miltyeni Biotec). The negative fraction was collected. The columns were then taken off the magnet and the positively selected cells were plunged into a 15ml conical, washed three times, centrifuged at 300g for 3 minutes, and resuspended in 500µl of R10. The cells were counted with Trypan Blue (ThermoFisher Scientific) and 15,000 live cells from each condition added to each array.

The Seq-Well protocol was then followed on the arrays (Gierahn et al., 2017) and the samples were either sequenced on a NextSeq 550 Series (Illumina) using an Illumina 75 cycle NextSeq500/550 v2 kit (Illumina FC-404-2005).

Single-cell RNA-seq of infected Rev-CEM-E7 cells using Smart-Seq2

Single NL4-3 *in vitro* infected Rev-CEM cells were sorted into 96-well plates and processed through reverse transcription, whole transcriptome amplification (WTA), and library preparation

following the previously described protocol (Trombetta et al., 2014). The samples were sequenced on a NextSeq 550 Series (Illumina) using 30bp ends. Reads were aligned to the Hg19 genome and to the NL4-3 genome as previously described (Macosko et al., 2015) using STAR (Dobin et al., 2013) and Hisat2 (Kim et al., 2015). Low quality cells were trimmed (cells with <25,000 mapped reads or <1000 genes) and differential expression analysis of genes based on HIV-1 infection of the cells was conducted in Seurat (Satija et al., 2015). Gene expression based on HIV-1 infection was plotted in a heat map showing genes that were found to be expressed in HIV-1 infected cells versus genes that were associated with cells not infected with HIV-1.

Single-cell RNA-seq alignment and cell type identification from Seq-Well data

Read alignment was performed as in previously described (Macosko et al., 2015, Ordovas-Montanes et al., 2018) using the Hg19 genome and the HIV-1 Clade C consensus genome. Cell barcode discrimination and UMI/transcript collation were performed as previously described (Ordovas-Montanes et al., 2018). The pilot study run on the lymph node cells from PID 024-09-0198 was analysed separately from the rest of the cohort. Cells with less than 500 UMIs detected or more than 5000 UMIs were excluded as well as any gene expressed in less than 5 cells from any downstream analysis. Any cell with at least 300 unique cells was retained. This was done on all of the arrays that were processed per study participant, on an individual basis. These objects were then merged into one object for cell type identification across all study participants, except for PID 024-09-0198, which was analysed separately using the same parameters.

A combined Seurat (version 2.3.4) (Satija et al., 2015) object was created for all the study participants except PID 024-09-0198, and log normalized with a size factor of 10,000. The object was scaled with centering. A total number of 22,114 cells passed these parameters with 39,385 unique genes. We then ran principle component analysis (PCA) over this list of variable genes. For the *t*-SNE plots we used the first 36 principal components as indicated by the inflection on the elbow plot. We used FindClusters within Seurat (uses shared nearest neighbour (SNN) modularity optimization-based clustering algorithm to identify clusters) with a resolution of 2.4 to identify 31 clusters across the 8 input samples. We also plotted the cells on a *t*-SNE by study participant and by array type, to better understand the distribution of cells.

For PID 024-09-0198, we merged data from two arrays and found variable genes using 9 principal components to create a *t*-SNE plot. With FindClusters we found 9 clusters. We used the 'bimod' function in FindMarkers implemented in Seurat which uses a likelihood-ration test for identifying

differentially expressed genes in single-cell data. The FeaturesPlot implemented in Seurat, which colours cells on a dimensional reduction plot according to 'feature', was used to indicate genes that were highly expressed in each cluster and were used as markers to identify the cell type of each cluster. The normalized expressions of the top 100 genes per cluster were plotted on a heatmap. We then merged the datasets that had aligned to HIV-1 into the matrix and plotted the same *t*-SNE and including the total *gag-pol* transcripts across it. We then created vectors for each HIV-1 gene including *env*, *vif*, and *tat* to determine the sensitive range for each gene. Once the cut-off had been determined, the same *t*-SNE was plotted as above with shapes for the cells that had HIV-1 transcripts.

The *in vitro* infected cells were analysed separately. Four arrays were sequenced per infection: unsorted, NL4-3 infected BFP negative, NL4-3 infected BFP positive, and mock infected. A violin plot was created in Seurat to show the abundance of HIV-1 UMIs across the four conditions and a *t*-SNE plotted by infection condition. The correlation HIV-1 infection as determined by flow cytometry in comparison to HIV-1 transcripts from the Seq-Well data was plotted using Prism (version 7.0).

Antibodies

The following antibodies were used: PE conjugated CD3, APC conjugated CD33, PE-Cy7 conjugated CD19, BV711 conjugated CD4, APC conjugated CD8, PE conjugated CD14, BV711 conjugated CD68, PE conjugated CD163. Monoclonal antibodies were purchased from Beckton Dickinson; CA, USA). FITC conjugated and PE conjugated anti-HIV-1 p24 Gag (KC57) was from Beckman Coulter (CA, USA). Anti-Env broadly neutralizing antibody 10-1074 was obtained from the NIH AIDS Reagents Program and conjugated to Alex Fluor 488 (ThermoFisher Scientific). The anti-Env antibody J3 was acquired from Bruce Walker's laboratory conjugated to Alex Fluor 647 or to biotin.

Flow Cytometry and sorting of in vivo infected lymph node cells

Lymph node cells were thawed in 10mls of R10 media and centrifuged at 300g for 7 minutes and resuspended in 1ml MACS buffer (phosphate buffered saline, 2% heat-inactivated FBS (GE Healthcare Biosciences, Pittsburgh, PA), and 1mM EDTA). Cells were counted and resuspended at 1 x 10^6 cells/ml before staining. Cells were acquired and sorted on a Sony SH800S Cell Sorter and data analyzed using FlowJo 10.0.8 software. 1000 cells were sorted for each population of

cells positive for anti-HIV-1 p24 Gag into PKD buffer and RNA was extracted using the RNeasy FFPE Kit (Qiagen). The Smart-Seq2 protocol was followed for sequencing and libraries were prepared as previously described (Trombetta et al., 2014).

References

ABBAS, W., TARIQ, M., IQBAL, M., KUMAR, A. & HERBEIN, G. 2015. Eradication of HIV-1 from the macrophage reservoir: an uncertain goal? *Viruses,* 7**,** 1578-98.

BANGA, R., PROCOPIO, F. A., NOTO, A., POLLAKIS, G., CAVASSINI, M., OHMITI, K., CORPATAUX, J.-M., DE LEVAL, L., PANTALEO, G. & PERREAU, M. 2016. PD-1+ and follicular helper T cells are responsible for persistent HIV-1 transcription in treated aviremic individuals. *Nature Medicine,* 22**,** 754-761.

BARTON, K. M. & PALMER, S. E. 2016. How to Define the Latent Reservoir: Tools of the Trade. *Curr HIV/AIDS Rep,* 13**,** 77-84.

BAXTER, A. E., RUSSELL, R. A., DUNCAN, C. J., MOORE, M. D., WILLBERG, C. B., PABLOS, J. L., FINZI, A., KAUFMANN, D. E., OCHSENBAUER, C., KAPPES, J. C., GROOT, F. & SATTENTAU, Q. J. 2014. Macrophage infection via selective capture of HIV-1-infected CD4+ T cells. *Cell Host Microbe,* 16**,** 711-21.

BOULLE, M., MULLER, T. G., DAHLING, S., GANGA, Y., JACKSON, L., MAHAMED, D., OOM, L., LUSTIG, G., NEHER, R. A. & SIGAL, A. 2016. HIV Cell-to-Cell Spread Results in Earlier Onset of Viral Gene Expression by Multiple Infections per Cell. *PLoS Pathogens,* 12**,** e1005964.

BRADLEY, T., FERRARI, G., HAYNES, B. F., MARGOLIS, D. M. & BROWNE, E. P. 2018. Single-Cell Analysis of Quiescent HIV Infection Reveals Host Transcriptional Profiles that Regulate Proviral Latency. *Cell Rep,* 25**,** 107-117 e3.

BRUNER, K. M., HOSMANE, N. N. & SILICIANO, R. F. 2015. Towards an HIV-1 cure: measuring the latent reservoir. *Trends Microbiol,* 23**,** 192-203.

BRUNER, K. M., WANG, Z., SIMONETTI, F. R., BENDER, A. M., KWON, K. J., SENGUPTA, S., FRAY, E. J., BEG, S. A., ANTAR, A. A. R., JENIKE, K. M., BERTAGNOLLI, L. N., CAPOFERRI, A. A., KUFERA, J. T., TIMMONS, A., NOBLES, C., GREGG, J., WADA, N., HO, Y. C., ZHANG, H., MARGOLICK, J. B., BLANKSON, J. N., DEEKS, S. G., BUSHMAN, F. D., SILICIANO, J. D., LAIRD, G. M. & SILICIANO, R. F. 2019. A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. *Nature,* 566**,** 120-125.

CASKEY, M., SCHOOFS, T., GRUELL, H., SETTLER, A., KARAGOUNIS, T., KREIDER, E. F., MURRELL, B., PFEIFER, N., NOGUEIRA, L., OLIVEIRA, T. Y., LEARN, G. H., COHEN, Y. Z., LEHMANN, C., GILLOR, D., SHIMELIOVICH, I., UNSON-O'BRIEN, C., WEILAND, D., ROBLES, A., KÜMMERLE, T., WYEN, C., LEVIN, R., WITMER-

PACK, M., EREN, K., IGNACIO, C., KISS, S., WEST JR, A. P., MOUQUET, H., ZINGMAN, B. S., GULICK, R. M., KELER, T., BJORKMAN, P. J., SEAMAN, M. S., HAHN, B. H., FÄTKENHEUER, G., SCHLESINGER, S. J., NUSSENZWEIG, M. C. & KLEIN, F. 2017. Antibody 10-1074 suppresses viremia in HIV-1-infected individuals. *Nature Medicine,* 23**,** 185.

CASTELLANO, P., PREVEDEL, L. & EUGENIN, E. A. 2017. HIV-infected macrophages and microglia that survive acute infection become viral reservoirs by a mechanism involving Bim. *Sci Rep,* 7**,** 12866.

CHOMONT, N., EL-FAR, M., ANCUTA, P., TRAUTMANN, L., PROCOPIO, F. A., YASSINE-DIAB, B., BOUCHER, G., BOULASSEL, M. R., GHATTAS, G., BRENCHLEY, J. M., SCHACKER, T. W., HILL, B. J., DOUEK, D. C., ROUTY, J. P., HADDAD, E. K. & SEKALY, R. P. 2009. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med,* 15**,** 893-900.

CLAYTON, K. L., COLLINS, D. R., LENGIEZA, J., GHEBREMICHAEL, M., DOTIWALA, F., LIEBERMAN, J. & WALKER, B. D. 2018. Resistance of HIV-infected macrophages to CD8+ T lymphocyte–mediated killing drives activation of the immune system. *Nature Immunology,* 19**,** 475-486.

COHN, L. B., DA SILVA, I. T., VALIERIS, R., HUANG, A. S., LORENZI, J. C. C., COHEN, Y. Z., PAI, J. A., BUTLER, A. L., CASKEY, M., JANKOVIC, M. & NUSSENZWEIG, M. C. 2018. Clonal CD4(+) T cells in the HIV-1 latent reservoir display a distinct gene profile upon reactivation. *Nat Med,* 24**,** 604-609.

DESCOURS, B., PETITJEAN, G., LOPEZ-ZARAGOZA, J. L., BRUEL, T., RAFFEL, R., PSOMAS, C., REYNES, J., LACABARATZ, C., LEVY, Y., SCHWARTZ, O., LELIEVRE, J. D. & BENKIRANE, M. 2017. CD32a is a marker of a CD4 T-cell HIV reservoir harbouring replication-competent proviruses. *Nature,* 543**,** 564-567.

DEYMIER, M. J., CLAIBORNE, D. T., ENDE, Z., RATNER, H. K., KILEMBE, W., ALLEN, S. & HUNTER, E. 2014. Particle infectivity of HIV-1 full-length genome infectious molecular clones in a subtype C heterosexual transmission pair following high fidelity amplification and unbiased cloning. *Virology,* 468-470**,** 454-461.

DOBIN, A., DAVIS, C. A., SCHLESINGER, F., DRENKOW, J., ZALESKI, C., JHA, S., BATUT, P., CHAISSON, M. & GINGERAS, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* 29**,** 15-21.

ERIKSSON, S., GRAF, E. H., DAHL, V., STRAIN, M. C., YUKL, S. A., LYSENKO, E. S., BOSCH, R. J., LAI, J., CHIOMA, S., EMAD, F., ABDEL-MOHSEN, M., HOH, R., HECHT, F., HUNT, P., SOMSOUK, M., WONG, J., JOHNSTON, R., SILICIANO, R. F., RICHMAN, D. D., O'DOHERTY, U., PALMER, S., DEEKS, S. G. & SILICIANO, J. D. 2013. Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. *PLoS Pathogens,* 9**,** e1003174.

FOIS, A. F. & BREW, B. J. 2015. The Potential of the CNS as a Reservoir for HIV-1 Infection: Implications for HIV Eradication. *Current HIV/AIDS Reports,* 12**,** 299-303.

FROMENTIN, R., BAKEMAN, W., LAWANI, M. B., KHOURY, G., HARTOGENSIS, W., DAFONSECA, S., KILLIAN, M., EPLING, L., HOH, R., SINCLAIR, E., HECHT, F. M., BACCHETTI, P., DEEKS, S. G., LEWIN, S. R., SEKALY, R. P. & CHOMONT, N. 2016. CD4+ T Cells Expressing PD-1, TIGIT and LAG-3 Contribute to HIV Persistence during ART. *PLoS Pathog,* 12**,** e1005761.

FROMENTIN, R., DAFONSECA, S., COSTINIUK, C. T., EL-FAR, M., PROCOPIO, F. A., HECHT, F. M., HOH, R., DEEKS, S. G., HAZUDA, D. J., LEWIN, S. R., ROUTY, J. P., SEKALY, R. P. & CHOMONT, N. 2019. PD-1 blockade potentiates HIV latency reversal ex vivo in CD4(+) T cells from ART-suppressed individuals. *Nat Commun,* 10**,** 814.

GIERAHN, T. M., WADSWORTH, M. H., 2ND, HUGHES, T. K., BRYSON, B. D., BUTLER, A., SATIJA, R., FORTUNE, S., LOVE, J. C. & SHALEK, A. K. 2017. Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nat Methods,* 14**,** 395-398.

GORRY, P. R., BRISTOL, G., ZACK, J. A., RITOLA, K., SWANSTROM, R., BIRCH, C. J., BELL, J. E., BANNERT, N., CRAWFORD, K., WANG, H., SCHOLS, D., DE CLERCQ, E., KUNSTMAN, K., WOLINSKY, S. M. & GABUZDA, D. 2001. Macrophage tropism of human immunodeficiency virus type 1 isolates from brain and lymphoid tissues predicts neurotropism independent of coreceptor specificity. *J Virol,* 75**,** 10073-89.

GOUJON, C., MONCORGE, O., BAUBY, H., DOYLE, T., WARD, C. C., SCHALLER, T., HUE, S., BARCLAY, W. S., SCHULZ, R. & MALIM, M. H. 2013. Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature,* 502**,** 559-62.

GROOT, F., RUSSELL, R. A., BAXTER, A. E., WELSCH, S., DUNCAN, C. J. A., WILLBERG, C., OCHSENBAUER, C., KAPPES, J. C., SHAW, M. & SATTENTAU, Q. J. 2011. Efficient macrophage infection by phagocytosis of dying HIV-1 -infected CD4+T cells. *Retrovirology,* 8.

HONEYCUTT, J. B., WAHL, A., BAKER, C., SPAGNUOLO, R. A., FOSTER, J., ZAKHAROVA, O., WIETGREFE, S., CARO-VEGAS, C., MADDEN, V., SHARPE, G., HAASE, A. T., ERON, J. J. & GARCIA, J. V. 2016. Macrophages sustain HIV replication in vivo independently of T cells. *J Clin Invest,* 126**,** 1353-66.

HORIIKE, M., IWAMI, S., KODAMA, M., SATO, A., WATANABE, Y., YASUI, M., ISHIDA, Y., KOBAYASHI, T., MIURA, T. & IGARASHI, T. 2012. Lymph nodes harbor viral reservoirs that cause rebound of plasma viremia in SIV-infected macaques upon cessation of combined antiretroviral therapy. *Virology,* 423**,** 107-18.

JACKSON, L., HUNTER, J., CELE, S., FERREIRA, I. M., YOUNG, A. C., KARIM, F., MADANSEIN, R., DULLABH, K. J., CHEN, C. Y., BUCKELS, N. J., GANGA, Y., KHAN, K., BOULLE, M., LUSTIG, G., NEHER, R. A. & SIGAL, A. 2018. Incomplete inhibition of HIV infection results in more HIV infected lymph node cells by reducing cell death. *Elife,* 7.

KABAMBA-MUKADI, B., HENRIVAUX, P., RUELLE, J., DELFERRIERE, N., BODEUS, M. & GOUBAU, P. 2005. Human immunodeficiency virus type 1 (HIV-1) proviral DNA load in purified CD4+ cells by LightCycler real-time PCR. *BMC Infect Dis,* 5**,** 15.

KIM, D., LANGMEAD, B. & SALZBERG, S. L. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods,* 12**,** 357-60.

KUGATHASAN, R., COLLIER, D. A., HADDOW, L. J., EL BOUZIDI, K., EDWARDS, S. G., CARTLEDGE, J. D., MILLER, R. F. & GUPTA, R. K. 2017. Diffuse White Matter Signal Abnormalities on Magnetic Resonance Imaging Are Associated With Human Immunodeficiency Virus Type 1 Viral Escape in the Central Nervous System Among Patients With Neurological Symptoms. *Clinical Infectious Diseases,* 64**,** 1059-1065.

KULPA, D. A. & CHOMONT, N. 2015. HIV persistence in the setting of antiretroviral therapy: when, where and how does HIV hide? *Journal of Virus Eradication,* 1**,** 59-66.

LEE, G. Q. & LICHTERFELD, M. 2016. Diversity of HIV-1 reservoirs in CD4+ T-cell subpopulations. *Curr Opin HIV AIDS,* 11**,** 383-7.

LISZEWSKI, M. K., YU, J. J. & O'DOHERTY, U. 2009. Detecting HIV-1 integration by repetitive-sampling Alu-gag PCR. *Methods,* 47**,** 254-60.

MACOSKO, E. Z., BASU, A., SATIJA, R., NEMESH, J., SHEKHAR, K., GOLDMAN, M., TIROSH, I., BIALAS, A. R., KAMITAKI, N., MARTERSTECK, E. M., TROMBETTA, J. J., WEITZ, D. A., SANES, J. R., SHALEK, A. K., REGEV, A. & MCCARROLL, S. A. 2015. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell,* 161**,** 1202-1214.

MCCOY, L. E., GROPPELLI, E., BLANCHETOT, C., DE HAARD, H., VERRIPS, T., RUTTEN, L., WEISS, R. A. & JOLLY, C. 2014. Neutralisation of HIV-1 cell-cell spread by human and llama antibodies. *Retrovirology,* 11**,** 1-15.

MCCOY, L. E., QUIGLEY, A. F., STROKAPPE, N. M., BULMER-THOMAS, B., SEAMAN, M. S., MORTIER, D., RUTTEN, L., CHANDER, N., EDWARDS, C. J., KETTELER, R., DAVIS, D., VERRIPS, T. & WEISS, R. A. 2012. Potent and broad neutralization of HIV-1 by a llama antibody elicited by immunization. *J Exp Med,* 209**,** 1091-103.

MCELRATH, M. J., SMYTHE, K., RANDOLPH-HABECKER, J., MELTON, K. R., GOODPASTER, T. A., HUGHES, S. M., MACK, M., SATO, A., DIAZ, G., STEINBACH, G., NOVAK, R. M., CURLIN, M. E., LORD, J. D., MAENZA, J., DUERR, A., FRAHM, N., HLADIK, F. & NETWORK, N. H. V. T. 2013. Comprehensive assessment of HIV target cells in the distal human gut suggests increasing HIV susceptibility toward the anus. *J Acquir Immune Defic Syndr,* 63**,** 263-71.

MEHRAJ, V., JENABIAN, M. A., VYBOH, K. & ROUTY, J. P. 2014. Immune suppression by myeloid cells in HIV infection: New targets for immunotherapy. *The Open AIDS Journal,* 8.

MLCOCHOVA, P., WATTERS, S. A., TOWERS, G. J., NOURSADEGHI, M. & GUPTA, R. K. 2014. Vpx complementation of 'non-macrophage tropic' R5 viruses reveals robust entry of infectious HIV-1 cores into macrophages. *Retrovirology,* 11**,** 1-12.

MURRAY, A. J., KWON, K. J., FARBER, D. L. & SILICIANO, R. F. 2016. The Latent Reservoir for HIV-1: How Immunologic Memory and Clonal Expansion Contribute to HIV-1 Persistence. *J Immunol,* 197**,** 407-17.

MURRAY, J. M., ZAUNDERS, J. J., MCBRIDE, K. L., XU, Y., BAILEY, M., SUZUKI, K., COOPER, D. A., EMERY, S., KELLEHER, A. D., KOELSCH, K. K. & TEAM, P. S. 2014. HIV DNA subspecies persist in both activated and resting memory CD4+ T cells during antiretroviral therapy. *J Virol,* 88**,** 3516-26.

NOTO, A., PROCOPIO, F. A., BANGA, R., SUFFIOTTI, M., CORPATAUX, J.-M., CAVASSINI, M., FENWICK, C., GOTTARDO, R., PERREAU, M. & PANTALEO, G. 2018. CD32+ and PD-1+ Lymph Node CD4 T Cells Support Persistent HIV-1 Transcription in Treated Aviremic Individuals. *Journal of Virology*.

O'DOHERTY, U., SWIGGARD, W. J., JEYAKUMAR, D., MCGAIN, D. & MALIM, M. H. 2002. A sensitive, quantitative assay for human immunodeficiency virus type 1 integration. *J Virol,* 76**,** 10942-50.

OMONDI, F. H., CHANDRARATHNA, S., MUJIB, S., BRUMME, C. J., JIN, S. W., SUDDERUDDIN, H., MILLER, R., RAHIMI, A., LAEYENDECKER, O., BONNER, P., YUE, F. Y., BENKO, E., KOVACS, C. M., BROCKMAN, M. A., OSTROWSKI, M. & BRUMME, Z. L. 2019. HIV subtype and Nef-mediated immune evasion function correlate with viral reservoir size in early-treated individuals. *J Virol*.

ORDOVAS-MONTANES, J., DWYER, D. F., NYQUIST, S. K., BUCHHEIT, K. M., VUKOVIC, M., DEB, C., WADSWORTH, M. H., 2ND, HUGHES, T. K., KAZER, S. W., YOSHIMOTO, E., CAHILL, K. N., BHATTACHARYYA, N., KATZ, H. R., BERGER, B., LAIDLAW, T. M., BOYCE, J. A., BARRETT, N. A. & SHALEK, A. K. 2018. Allergic inflammatory memory in human respiratory epithelial progenitor cells. *Nature,* 560**,** 649-654.

RATO, S., RAUSELL, A., MUNOZ, M., TELENTI, A. & CIUFFI, A. 2017. Single-cell analysis identifies cellular markers of the HIV permissive cell. *PLoS Pathog,* 13**,** e1006678.

REGIS, E. G., BARRETO-DE-SOUZA, V., MORGADO, M. G., BOZZA, M. T., LENG, L., BUCALA, R. & BOU-HABIB, D. C. 2010. Elevated levels of macrophage migration inhibitory factor (MIF) in the plasma of HIV-1-infected patients and in HIV-1-infected cell cultures: a relevant role on viral replication. *Virology,* 399**,** 31-38.

RODRIGUES, V., RUFFIN, N., SAN-ROMAN, M. & BENAROCH, P. 2017. Myeloid Cell Interaction with HIV: A Complex Relationship. *Front Immunol,* 8**,** 1698.

RUTSAERT, S., BOSMAN, K., TRYPSTEEN, W., NIJHUIS, M. & VANDEKERCKHOVE, L. 2018. Digital PCR as a tool to measure HIV persistence. *Retrovirology,* 15**,** 16.

SATIJA, R., FARRELL, J. A., GENNERT, D., SCHIER, A. F. & REGEV, A. 2015. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol,* 33**,** 495-502.

SCHNELL, G., JOSEPH, S., SPUDICH, S., PRICE, R. W. & SWANSTROM, R. 2011. HIV-1 replication in the central nervous system occurs in two distinct cell types. *PLoS Pathog,* 7**,** e1002286.

SCHNELL, G., SPUDICH, S., HARRINGTON, P., PRICE, R. W. & SWANSTROM, R. 2009. Compartmentalized human immunodeficiency virus type 1 originates from long-lived cells in some subjects with HIV-1-associated dementia. *PLoS Pathog,* 5**,** e1000395.

SHEN, L., PETERSON, S., SEDAGHAT, A. R., MCMAHON, M. A., CALLENDER, M., ZHANG, H., ZHOU, Y., PITT, E., ANDERSON, K. S., ACOSTA, E. P. & SILICIANO, R. F. 2008. Dose-response curve slope sets class-specific limits on inhibitory potential of anti-HIV drugs. *Nat Med,* 14**,** 762-6.

SILICIANO, J. M. & SILICIANO, R. F. 2015. The Remarkable Stability of the Latent Reservoir for HIV-1 in Resting Memory CD4+ T Cells. *J Infect Dis,* 212**,** 1345-7.

STRAIN, M. C., LADA, S. M., LUONG, T., ROUGHT, S. E., GIANELLA, S., TERRY, V. H., SPINA, C. A., WOELK, C. H. & RICHMAN, D. D. 2013. Highly precise measurement of HIV DNA by droplet digital PCR. *PLoS One,* 8**,** e55943.

SVICHER, V., CECCHERINI-SILBERSTEIN, F., ANTINORI, A., AQUARO, S. & PERNO, C. F. 2014. Understanding HIV compartments and reservoirs. *Curr HIV/AIDS Rep,* 11**,** 186-94.

TROMBETTA, J. J., GENNERT, D., LU, D., SATIJA, R., SHALEK, A. K. & REGEV, A. 2014. Preparation of Single-Cell RNA-Seq Libraries for Next Generation Sequencing. *Curr Protoc Mol Biol,* 107**,** 4 22 1-17.

TRYPSTEEN, W., KISELINOVA, M., VANDEKERCKHOVE, L. & DE SPIEGELAERE, W. 2016. Diagnostic utility of droplet digital PCR for HIV reservoir quantification. *Journa of Virus Eradication,* 2**,** 162-169.

WILLIAMS, K. C., COREY, S., WESTMORELAND, S. V., PAULEY, D., KNIGHT, H., DEBAKKER, C., ALVAREZ, X. & LACKNER, A. A. 2001. Perivascular macrophages are the primary cell type productiviely infected by simian immunodeficiency virus in the brains of macaques: Implications for the neuropathogenesis of AIDS. *J. Exp. Med.,* 193**,** 905-915.

CHAPTER 3: DISCUSSION/SYNTHESIS

HIV-1 persistence in the face of ART necessitates lifelong adherence to treatment. Identifying the cellular HIV-1 reservoir in ART suppressed individuals is critical to creating a functional therapeutic cure. Previous studies have identified different cell subsets including CD32a expressing T cells (Descours et al., 2017), PD-1 expressing T cells (Banga et al., 2016, Fromentin et al., 2016), follicular T helper cells (Perreau et al., 2013, Vinuesa, 2012), or resting memory T cells as harbouring the HIV-1 reservoir (Siliciano and Siliciano, 2015). However, the establishment and size of the reservoir could depend on many factors and different cell types could be infected in different individuals. For example, CD32a expressing T cells have recently been shown to not harbour the HIV-1 reservoir (Badia et al., 2018, Perez et al., 2018). Rather, CD32a was postulated to be a marker of T cell activation.

Persistence results from reservoirs of HIV-1 infected cells at several possible anatomical reservoir sites (Fois and Brew, 2015, Horiike et al., 2012, Kulpa and Chomont, 2015, McElrath et al., 2013, Svicher et al., 2014). A critical reservoir site is lymphoid tissue, comprising lymph nodes (Horiike et al., 2012, Spiegel et al., 1992) and gut associated lymphoid tissue (Marcelin et al., 2008, Politch et al., 2012). However, most of the studies investigating the cellular HIV-1 reservoir have been done either using *in vitro* experiments or PBMCs isolated from ART suppressed individuals (Chargin et al., 2015, Cohn et al., 2018, Ho et al., 2013). Few studies have been performed in lymph nodes from ART suppressed study participants (Denton et al., 2012, Honeycutt et al., 2017, Marsden et al., 2012). In addition, there is very little known about the HIV-1 reservoir in southern Africa, the geographical area with by far the most HIV-1 infected individuals on ART (Granich et al., 2012, Vermund et al., 2015). Here, we aimed to identify the cellular HIV-1 reservoir in lymph nodes from ART suppressed participants from South Africa who were infected with Clade C viruses, hence investigating the HIV-1 reservoir at the epicentre of the HIV-1 epidemic.

We established a cohort of participants in collaboration with cardiothoracic clinicians at the Inkosi Albert Luthuli Central Hospital in Durban. All of study participants used in this study were suppressed on ART and had no detectable peripheral viremia (<40 HIV-1 copies/ml). Antiretroviral (ARV) drug levels were quantified in the matched blood samples from each study participant using liquid chromatography–tandem mass spectrometry (LC-MS/MS).

To investigate the HIV-1 infected cell type in the HIV-1 reservoir in the face of ART, we developed a novel single-cell RNA sequencing approach, which had not been previously used to identify HIV-1 infected cells from lymphoid tissues. We coupled HIV-1 transcript detection to

Seq-Well (Gierahn et al., 2017), a high throughput single-cell RNA-Seq methodology developed by our collaborating lab at the Massachusetts Institute of Technology and the Ragon Institute of MGH, MIT, and Harvard. Seq-Well allows for the transcriptomic profiling of thousands of cells in parallel. However, due to the low frequency of HIV-1 infected cells (<1%) we had to identify ways to enrich for these cells. A previous study (Cohn et al., 2018) had used biotinylated anti-Env broadly neutralizing antibodies to enrich for HIV-1 infected cells using magnetic activated cell sorting. We employed the same method using the J3 nanobody (McCoy et al., 2012) and ran the Seq-Well protocol on the total lymph node cell population and on cells enriched for HIV-1 Env expression.

Analysis of our single-cell data indicated that $CD4^+$ T cells and macrophages contained HIV-1 transcripts. In our pilot study, we identified various T cell subsets including memory T cells and cycling T cells as containing HIV-1 transcripts. These cell types have been identified to harbour the HIV-1 reservoir in previous studies (Baxter et al., 2014, Siliciano and Siliciano, 2015). However, our data shows that the HIV-1 reservoir is not isolated to single cell type or cellular marker as suggested by previous studies (Banga et al., 2016, Descours et al., 2017, Noto et al., 2018). We showed that there is a diversity in cell types which contain HIV-1 transcripts, although this does not necessarily mean that the virus is fully functional.

The use of single-cell RNA sequencing to identify cells which contain HIV-1 transcripts is a novel technique in the field of HIV-1 reservoirs. It allows us to identify these cells across all cell types found in the lymph node in an unbiased way without having to pre-select cell types which are most likely to contain virus or are generally thought to be the HIV-1 reservoir. However, due to the low frequency of cells containing transcripts, enrichment was necessary and therefore we used enrichment based on HIV-1 expression in parallel to the total cell population. To investigate if cell populations detected by our approach actually produce infection competent virus, quantitative viral outgrowth assays would need to be conducted. This a drawback of the methods we have used.

In addition to the development of a novel technique for HIV-1 reservoir detection, we amplified whole HIV-1 genome sequences from extracted genomic DNA from the lymph node cells from each study participant. We used these genomes in alignments of single-cell RNA-Seq transcripts. We found better alignment when the transcriptomic data was aligned to the HIV-1 genome sequence from the same study participant compared to the HIV-1 Clade C consensus sequence. Furthermore, we ran flow cytometry on lymph node cells to identify populations that were anti-HIV-1 p24 Gag positive. We bulk sorted the p24 positive populations and ran population RNA sequencing on these cells after decrosslinking the RNA. We are currently comparing the gene

signatures from these cells with those from the single-cell data to further validate our transcriptomic data (data not shown in this thesis).

A further application of single-cell RNA sequencing to this project would be to sort single cells into 96-well plates and sequence the HIV-1 virus in these cells to determine clonality. This is important since latency would predict clonal sequences whilst ongoing replication would predict sequences that diverge. Therefore, it is possible to use this technique to investigate the mechanisms of HIV-1 persistence in the context of the HIV-1 reservoir and whether this mechanism is latency or not. We have already begun doing so by sorting single cells positive for anti-Env broadly neutralizing antibodies 10-1074 and J3 from each HIV-1 infected study participant. However, we are still running downstream experimental components and therefore the data is not shown in this thesis.

In conclusion, our data shows that there is diversity in the cell types that harbour HIV-1 transcripts in the lymph node cells from ART suppressed individuals. $CD4^+$ T cells and macrophages make up the majority of cells that contain these transcripts, with different subsets of T cells containing transcripts including memory T cells and circulating T cells. This goes against the hypothesis that the HIV-1 reservoir is restricted to a particular cellular subset or can be identified with a singular cell marker. However, further data analysis of our data is required to determine exactly which T cell subsets are infected and the diversity of these subsets across all study participants is, as the above results are based on preliminary analysis. The cells which are infected have poor quality RNA due to being infected with HIV-1 and even with sufficient sequencing depth, it is difficult to correctly identify them as belonging to a specific cell subset.

This study used novel techniques and methods to identify HIV-1 infected cells in the lymph node compartment and these can be further optimized to be used on different reservoir sites as well as to better identify the infected cells. The next steps of this project, along with further data analysis, are to further optimize the pull-down of HIV-1 transcripts from the mRNA captured on the Seq-Well beads as well as and to complete the flow cytometry validation for all study participants.

REFERENCES

BADIA, R., BALLANA, E., CASTELLVI, M., GARCIA-VIDAL, E., PUJANTELL, M., CLOTET, B., PRADO, J. G., PUIG, J., MARTINEZ, M. A., RIVEIRA-MUNOZ, E. & ESTE, J. A. 2018. CD32 expression is associated to T-cell activation and is not a marker of the HIV-1 reservoir. *Nat Commun,* 9**,** 2739.

BANGA, R., PROCOPIO, F. A., NOTO, A., POLLAKIS, G., CAVASSINI, M., OHMITI, K., CORPATAUX, J.-M., DE LEVAL, L., PANTALEO, G. & PERREAU, M. 2016. PD-1+ and follicular helper T cells are responsible for persistent HIV-1 transcription in treated aviremic individuals. *Nature Medicine,* 22**,** 754-761.

BAXTER, A. E., RUSSELL, R. A., DUNCAN, C. J., MOORE, M. D., WILLBERG, C. B., PABLOS, J. L., FINZI, A., KAUFMANN, D. E., OCHSENBAUER, C., KAPPES, J. C., GROOT, F. & SATTENTAU, Q. J. 2014. Macrophage infection via selective capture of HIV-1-infected CD4+ T cells. *Cell Host Microbe,* 16**,** 711-21.

BUZON, M. J., MASSANELLA, M., LLIBRE, J. M., ESTEVE, A., DAHL, V., PUERTAS, M. C., GATELL, J. M., DOMINGO, P., PAREDES, R., SHARKEY, M., PALMER, S., STEVENSON, M., CLOTET, B., BLANCO, J. & MARTINEZ-PICADO, J. 2010. HIV-1 replication and immune dynamics are affected by raltegravir intensification of HAART-suppressed subjects. *Nat Med,* 16**,** 460-5.

CHARGIN, A., YIN, F., SONG, M., SUBRAMANIAM, S., KNUTSON, G. & PATTERSON, B. K. 2015. Identification and characterization of HIV-1 latent viral reservoirs in peripheral blood. *J Clin Microbiol,* 53**,** 60-6.

CHUN, T. W., ENGEL, D., BERREY, M. M., SHEA, T., COREY, L. & FAUCI, A. S. 1998. Early establishment of a pool of latently infected, resting CD4+ T cells during primary HIV-1 infection *PNAS,* 95**,** 8869-8873.

CHUN, T. W., MOIR, S. & FAUCI, A. S. 2015. HIV reservoirs as obstacles and opportunities for an HIV cure. *Nature Immunology,* 16**,** 584-9.

CHURCHILL, M. J., DEEKS, S. G., MARGOLIS, D. M., SILICIANO, R. F. & SWANSTROM, R. 2015. HIV reservoirs: what, where and how to target them. *Nature Reviews Microbiology,* 14**,** 55-60.

COHN, L. B., DA SILVA, I. T., OLIVEIRA, T. Y., ROSALES, R. A., PARRISH, E. H., LEARN, G. H., HAHN, B. H., CZARTOSKI, J. L., MCELRATH, M. J., LEHMANN, C., KLEIN, F., CASKEY, M., WALKER, B. D., SILICIANO, J. D., SILICIANO, R. F., JANKOVIC, M. & NUSSENZWEIG, M. C. 2015. HIV-1 Integration Landscape during Latent and Active Infection. *Cell,* 160**,** 420-432.

COHN, L. B., DA SILVA, I. T., VALIERIS, R., HUANG, A. S., LORENZI, J. C. C., COHEN, Y. Z., PAI, J. A., BUTLER, A. L., CASKEY, M., JANKOVIC, M. & NUSSENZWEIG,

M. C. 2018. Clonal CD4(+) T cells in the HIV-1 latent reservoir display a distinct gene profile upon reactivation. *Nat Med,* 24**,** 604-609.

COHEN, J. 2011. Tissue Says Blood Is Misleading, Confusing HIV Cure Efforts. *Science,* 334**,** 1614-1614.

DENTON, P. W., OLESEN, R., CHOUDHARY, S. K., ARCHIN, N. M., WAHL, A., SWANSON, M. D., CHATEAU, M., NOCHI, T., KRISKO, J. F., SPAGNUOLO, R. A., MARGOLIS, D. M. & GARCIA, J. V. 2012. Generation of HIV latency in humanized BLT mice. *J Virol,* 86**,** 630-4.

DESCOURS, B., PETITJEAN, G., LOPEZ-ZARAGOZA, J. L., BRUEL, T., RAFFEL, R., PSOMAS, C., REYNES, J., LACABARATZ, C., LEVY, Y., SCHWARTZ, O., LELIEVRE, J. D. & BENKIRANE, M. 2017. CD32a is a marker of a CD4 T-cell HIV reservoir harbouring replication-competent proviruses. *Nature,* 543**,** 564-567.

ERIKSSON, S., GRAF, E. H., DAHL, V., STRAIN, M. C., YUKL, S. A., LYSENKO, E. S., BOSCH, R. J., LAI, J., CHIOMA, S., EMAD, F., ABDEL-MOHSEN, M., HOH, R., HECHT, F., HUNT, P., SOMSOUK, M., WONG, J., JOHNSTON, R., SILICIANO, R. F., RICHMAN, D. D., O'DOHERTY, U., PALMER, S., DEEKS, S. G. & SILICIANO, J. D. 2013. Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. *PLoS Pathogens,* 9**,** e1003174.

FINZI, D., BLANKSON, J., SILICIANO, J. D., MARGOLICK, J. B., CHADWICK, K., PIERSON, T., SMITH, K., LISZIEWICZ, J., LORI, F., FLEXNER, C., QUINN, T. C., CHAISSON, R. E., ROSENBERG, E., WALKER, B., GANGE, S., GALLANT, J. & SILICIANO, R. F. 1999. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nature Medicine,* 5**,** 512-517.

FOIS, A. F. & BREW, B. J. 2015. The Potential of the CNS as a Reservoir for HIV-1 Infection: Implications for HIV Eradication. *Current HIV/AIDS Reports,* 12**,** 299-303.

FROMENTIN, R., BAKEMAN, W., LAWANI, M. B., KHOURY, G., HARTOGENSIS, W., DAFONSECA, S., KILLIAN, M., EPLING, L., HOH, R., SINCLAIR, E., HECHT, F. M., BACCHETTI, P., DEEKS, S. G., LEWIN, S. R., SEKALY, R. P. & CHOMONT, N. 2016. CD4+ T Cells Expressing PD-1, TIGIT and LAG-3 Contribute to HIV Persistence during ART. *PLoS Pathog,* 12**,** e1005761.

GIERAHN, T. M., WADSWORTH, M. H., 2ND, HUGHES, T. K., BRYSON, B. D., BUTLER, A., SATIJA, R., FORTUNE, S., LOVE, J. C. & SHALEK, A. K. 2017. Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nat Methods,* 14**,** 395-398.

GRANICH, R., KAHN, J. G., BENNETT, R., HOLMES, C. B., GARG, N., SERENATA, C., SABIN, M. L., MAKHLOUF-OBERMEYER, C., DE FILIPPO MACK, C., WILLIAMS, P., JONES, L., SMYTH, C., KUTCH, K. A., YING-RU, L., VITORIA, M.,

SOUTEYRAND, Y., CROWLEY, S., KORENROMP, E. L. & WILLIAMS, B. G. 2012. Expanding ART for treatment and prevention of HIV in South Africa: estimated cost and cost-effectiveness 2011-2050. *PLoS One,* 7**,** e30216.

HELLMUTH, J., VALCOUR, V. & SPUDICH, S. 2015. CNS reservoirs for HIV: implications for eradication. *Journal of Virus Eradication,* 1**,** 67-71.

HO, Y. C., SHAN, L., HOSMANE, N. N., WANG, J., LASKEY, S. B., ROSENBLOOM, D. I., LAI, J., BLANKSON, J. N., SILICIANO, J. D. & SILICIANO, R. F. 2013. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell,* 155**,** 540-51.

HONEYCUTT, J. B., THAYER, W. O., BAKER, C. E., RIBEIRO, R. M., LADA, S. M., CAO, Y., CLEARY, R. A., HUDGENS, M. G., RICHMAN, D. D. & GARCIA, J. V. 2017. HIV persistence in tissue macrophages of humanized myeloid-only mice during antiretroviral therapy. *Nat Med,* 23**,** 638-643.

HORIIKE, M., IWAMI, S., KODAMA, M., SATO, A., WATANABE, Y., YASUI, M., ISHIDA, Y., KOBAYASHI, T., MIURA, T. & IGARASHI, T. 2012. Lymph nodes harbor viral reservoirs that cause rebound of plasma viremia in SIV-infected macaques upon cessation of combined antiretroviral therapy. *Virology,* 423**,** 107-18.

KHARSANY, A. B. & KARIM, Q. A. 2016. HIV Infection and AIDS in Sub-Saharan Africa: Current Status, Challenges and Opportunities. *Open AIDS J,* 10**,** 34-48.

KULPA, D. A. & CHOMONT, N. 2015. HIV persistence in the setting of antiretroviral therapy: when, where and how does HIV hide? *Journal of Virus Eradication,* 1**,** 59-66.

KUTSCHER, H. L., PRASAD, P. N., MORSE, G. D. & REYNOLDS, J. L. 2016. Emerging Nanomedicine Approaches to Targeting HIV-1 and Antiretroviral Therapy. *Future Virol,* 11**,** 101-104.

LICHT, A. & ALTER, G. 2016. A Drug-Free Zone--Lymph Nodes as a Safe Haven for HIV. *Cell Host Microbe,* 19**,** 275-6.

LORENZO-REDONDO, R., FRYER, H. R., BEDFORD, T., KIM, E. Y., ARCHER, J., POND, S. L. K., CHUNG, Y. S., PENUGONDA, S., CHIPMAN, J., FLETCHER, C. V., SCHACKER, T. W., MALIM, M. H., RAMBAUT, A., HAASE, A. T., MCLEAN, A. R. & WOLINSKY, S. M. 2016. Persistent HIV-1 replication maintains the tissue reservoir during therapy. *Nature,* 530**,** 51-56.

MARCELIN, A., TUBIANA, R., LAMBERT-NICLOT, S., LEFEBVRE, G., DOMINGUES, S., BONMARCHAND, M., VAUTHIER-BROUZES, N., MARGUET, F., MOUSSET-SIMEON, N., PEYTAVIN, G. & POIROT, C. 2008. Detection of HIV-1 RNA in seminal plasma samples from treated patients with undetectable HIV-1 RNA in blood plasma. *AIDS,* 22**,** 1673-1681.

MARSDEN, M. D., KOVOCHICH, M., SUREE, N., SHIMIZU, S., MEHTA, R., CORTADO, R., BRISTOL, G., AN, D. S. & ZACK, J. A. 2012. HIV latency in the humanized BLT mouse. *J Virol,* 86**,** 339-47.

MCCOY, L. E., QUIGLEY, A. F., STROKAPPE, N. M., BULMER-THOMAS, B., SEAMAN, M. S., MORTIER, D., RUTTEN, L., CHANDER, N., EDWARDS, C. J., KETTELER, R., DAVIS, D., VERRIPS, T. & WEISS, R. A. 2012. Potent and broad neutralization of HIV-1 by a llama antibody elicited by immunization. *J Exp Med,* 209**,** 1091-103.

MCELRATH, M. J., SMYTHE, K., RANDOLPH-HABECKER, J., MELTON, K. R., GOODPASTER, T. A., HUGHES, S. M., MACK, M., SATO, A., DIAZ, G., STEINBACH, G., NOVAK, R. M., CURLIN, M. E., LORD, J. D., MAENZA, J., DUERR, A., FRAHM, N., HLADIK, F. & NETWORK, N. H. V. T. 2013. Comprehensive assessment of HIV target cells in the distal human gut suggests increasing HIV susceptibility toward the anus. *J Acquir Immune Defic Syndr,* 63**,** 263-71.

NOTO, A., PROCOPIO, F. A., BANGA, R., SUFFIOTTI, M., CORPATAUX, J.-M., CAVASSINI, M., FENWICK, C., GOTTARDO, R., PERREAU, M. & PANTALEO, G. 2018. CD32+ and PD-1+ Lymph Node CD4 T Cells Support Persistent HIV-1 Transcription in Treated Aviremic Individuals. *Journal of Virology*.

OMONDI, F. H., CHANDRARATHNA, S., MUJIB, S., BRUMME, C. J., JIN, S. W., SUDDERUDDIN, H., MILLER, R., RAHIMI, A., LAEYENDECKER, O., BONNER, P., YUE, F. Y., BENKO, E., KOVACS, C. M., BROCKMAN, M. A., OSTROWSKI, M. & BRUMME, Z. L. 2019. HIV subtype and Nef-mediated immune evasion function correlate with viral reservoir size in early-treated individuals. *J Virol*.

PEREZ, L., ANDERSON, J., CHIPMAN, J., THORKELSON, A., CHUN, T. W., MOIR, S., HAASE, A. T., DOUEK, D. C., SCHACKER, T. W. & BORITZ, E. A. 2018. Conflicting evidence for HIV enrichment in CD32(+) CD4 T cells. *Nature,* 561**,** E9-E16.

PERREAU, M., SAVOYE, A. L., DE CRIGNIS, E., CORPATAUX, J. M., CUBAS, R., HADDAD, E. K., DE LEVAL, L., GRAZIOSI, C. & PANTALEO, G. 2013. Follicular helper T cells serve as the major CD4 T cell compartment for HIV-1 infection, replication, and production. *The Journal of Experimental Medicine,* 210**,** 143-56.

POLITCH, J. A., MAYER, K. H., WELLES, S. L., O'BRIEN, W. X., XU, C., BOWMAN, F. P. & ANDERSON, D. J. 2012. Highly active antiretroviral therapy does not completely suppress HIV in semen of sexually active HIV-infected men who have sex with men. *AIDS,* 26**,** 1535-43.

REEVES, D. B., DUKE, E. R., WAGNER, T. A., PALMER, S. E., SPIVAK, A. M. & SCHIFFER, J. T. 2018. A majority of HIV persistence during antiretroviral therapy is due to infected cell proliferation. *Nat Commun,* 9**,** 4811.

RUELAS, D. S. & GREENE, W. C. 2013. An integrated overview of HIV-1 latency. *Cell,* 155**,** 519-29.

SATTENTAU, Q. 2008. Avoiding the void: cell-to-cell spread of human viruses. *Nat Rev Microbiol,* 6**,** 815-26.

SIGAL, A. & BALTIMORE, D. 2012. As good as it gets? The problem of HIV persistence despite antiretroviral drugs. *Cell Host Microbe,* 12**,** 132-8.

SIGAL, A., KIM, J. T., BALAZS, A. B., DEKEL, E., MAYO, A., MILO, R. & BALTIMORE, D. 2011. Cell-to-cell spread of HIV permits ongoing replication despite antiretroviral therapy. *Nature,* 477**,** 95-8.

SILICIANO, J. M. & SILICIANO, R. F. 2015. The Remarkable Stability of the Latent Reservoir for HIV-1 in Resting Memory CD4+ T Cells. *J Infect Dis,* 212**,** 1345-7.

SILICIANO, R. F. & GREENE, W. C. 2011. HIV Latency. *Cold Spring Harbor Perspectives in Medicine,* 1**,** a007096-a007096.

SORIANO-SARABIA, N., BATESON, R. E., DAHL, N. P., CROOKS, A. M., KURUC, J. D., MARGOLIS, D. M. & ARCHIN, N. M. 2014. Quantitation of replication-competent HIV-1 in populations of resting CD4+ T cells. *J Virol,* 88**,** 14070-7.

SPIEGEL, H., HERBST, H., NIEDOBITEK, G., FOSS, H. & STEIN, H. 1992. Follicular denritic cells are a major reservoir for human immunodeficiency virus type 1 in lymphoid tissues facilitating infection of CD4+ T-helper cells. *American Journal of Pathology,* 140**,** 15-21.

SVICHER, V., CECCHERINI-SILBERSTEIN, F., ANTINORI, A., AQUARO, S. & PERNO, C. F. 2014. Understanding HIV compartments and reservoirs. *Curr HIV/AIDS Rep,* 11**,** 186-94.

VERMUND, S. H., SHELDON, E. K. & SIDAT, M. 2015. Southern Africa: the Highest Priority Region for HIV Prevention and Care Interventions. *Curr HIV/AIDS Rep,* 12**,** 191-5.

VINUESA, C. G. 2012. HIV and T follicular helper cells: a dangerous relationship. *J Clin Invest,* 122**,** 3059-62.

WANG, Z., GURULE, E. E., BRENNAN, T. P., GEROLD, J. M., KWON, K. J., HOSMANE, N. N., KUMAR, M. R., BEG, S. A., CAPOFERRI, A. A., RAY, S. C., HO, Y. C., HILL, A. L., SILICIANO, J. D. & SILICIANO, R. F. 2018a. Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. *Proc Natl Acad Sci U S A,* 115**,** E2575-E2584.

WANG, Z., SIMONETTI, F. R., SILICIANO, R. F. & LAIRD, G. M. 2018b. Measuring replication competent HIV-1: advances and challenges in defining the latent reservoir. *Retrovirology,* 15**,** 21.

YUKL, S. A., SHERGILL, A. K., HO, T., KILLIAN, M., GIRLING, V., EPLING, L., LI, P., WONG, L. K., CROUCH, P., DEEKS, S. G., HAVLIR, D. V., MCQUAID, K., SINCLAIR, E. & WONG, J. K. 2013. The distribution of HIV DNA and RNA in cell subsets differs in gut and blood of HIV-positive patients on ART: implications for viral persistence. *J Infect Dis,* 208**,** 1212-20.

ANNEX

ETHICS CERTIFICATE

UNIVERSITY OF
KWAZULU-NATAL
INYUVESI
YAKWAZULU-NATALI

RESEARCH OFFICE
BIOMEDICAL RESEARCH ETHICS ADMINISTRATION
Westville Campus
Govan Mbeki Building
Private Bag X 54001
Durban
4000
KwaZulu-Natal, SOUTH AFRICA
Tel: 27 31 2604769 - Fax: 27 31 260-4609
Email: BREC@ukzn.ac.za
Website: http://research.ukzn.ac.za/Research-Ethics/Biomedical-Research-Ethics.aspx

Amended letter
13 April 2017

Dr Adrie Steyn
KwaZulu-Natal Research Institute for TB and HIV (K-RITH),
Nelson R Mandela School of Medicine
University of KwaZulu-Natal

Dear Dr Steyn

PROTOCOL: Lung tissue and pleural effusion collection for the study on *Mycobacterium Tuberculosis* Immunology. REF: BE019/13.

RECERTIFICATION APPLICATION APPROVAL NOTICE

Approved: 30 July 2017
Expiration of Ethical Approval: 29 July 2018

I wish to advise you that your application for Recertification received on 07 April 2017 for the above protocol has been **noted and approved** by a sub-committee of the Biomedical Research Ethics Committee (BREC) for another approval period. The start and end dates of this period are indicated above.

If any modifications or adverse events occur in the project before your next scheduled review, you must submit them to BREC for review. Except in emergency situations, no change to the protocol may be implemented until you have received written BREC approval for the change.

The approval will be **ratified** by a full Committee at a meeting to be held on **09 May 2017.**

Yours sincerely

Mrs A Marimuthu
Senior Administrator: Biomedical Research Ethics