Hematopoietic stem and progenitor cell heterogeneity and susceptibility to HIV-1

Candidate

Name: Ms Juanita Mellet

Student Number: 27054374

Submitted in fulfilment of the requirements for the degree

Doctor of Philosophy (Medical Immunology)

In the Faculty of Health Sciences

Department of Immunology

University of Pretoria

Date: September 2019

Supervisor: Prof Michael S Pepper

DECLARATION

I, Juanita Mellet hereby declare that this thesis submitted to the University of Pretoria for the degree Doctor of Philosophy in Medical Immunology, is my own original work and has never been submitted for any academic award to any other tertiary institution for any degree.

_____

Juanita Mellet

ACKNOWLEDGEMENTS

This thesis becomes a reality with the support and help of many individuals and I would like to extend my sincere thanks to all of them.

Foremost, I would like to thank my *Heavenly Father*, for everything throughout my life and without whom nothing is possible.

My supervisor, *Prof Michael Pepper*, for this incredible opportunity and for letting me be part of the ICMM team/family. Your continuous scientific support, encouragement and guidance throughout the years are much appreciated.

The National Research Foundation (NRF), Poliomyelitis Research Foundation (PRF) and the Institute for Cellular and Molecular Medicine (ICMM) for financial assistance throughout this study period. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF and/or PRF.

Prof Fourie Joubert, thank you for your bioinformatics assistance and support.

Special thanks to all the cord blood donors and personnel at the private hospital where the blood was collected.

My colleagues at the ICMM. *Martie Madgwick* and *Candice Murdoch* for keeping the ICMM boat afloat and making sure that everything runs smoothly. *Dr Chrisna Durandt*, for your endless assistance and support in the lab and the long hours and weekends spent in front of the flow cytometers. *Dr Melvin Ambele*, for your expertise, help and support with the Affymetrix gene expression experiments and data analysis. The HIV team, *Dr Chrisna Durandt*, *Candice Herd* and *Catherine Wickham*, for each of your ideas and contributions to try and solve our HIV problems and for being the most amazing team players. The single-cell team*, Dr Chrisna Durandt* and *Elize Wolmarans*, the long and stressful days of single-cell capture and experiments certainly improved our knowledge in this rapidly-growing field. *Elize Wolmarans*,

thank you for your support with the bioinformatics analysis, it was good to have a partner in crime throughout this whole experience. *Candice Herd*, thank you for the inspiring scientist and incredible friend that you are and for always being willing to help where possible. To my office colleagues, thank you for many discussions, laughs, good times and for creating such an inspiring and positive working atmosphere in the office even on days when I did not feel all that positive. Being part of the ICMM family has been incredible and it is truly inspiring to be part of a team with so many talented individuals.

My close family and friends, thank you for your encouraging words and enduring friendship.

Finally, I would like to thank my parents. Your support, encouragement and love extends far beyond the years of my PhD. Thank you for your enthusiasm, positivity and interest in what I do. I could not have done this without you.

Above all, my loving husband and beautiful daughter, thank you for the support, patience, love and happy moments throughout this time. Thank you for encouraging me and believing in me. I love you both so much!!!

LIST OF PUBLICATIONS AND CONFERENCE OUTPUTS

I. Manuscripts for publication:

Mellet J, Durandt C, Ambele M, and Pepper MS. *The effect of StemRegenin-1 on the gene expression of umbilical cord blood derived hematopoietic stem and progenitor cells*. (To be submitted in 2020).

Mellet J, Herd C, Durandt C, and Pepper MS. *Ex vivo expansion of hematopoietic stem and progenitor cells*. (Additional work is required for this manuscript to be completed).

Herd C, Mellet J, Durandt C, and Pepper MS. *Hematopoietic stem and progenitor cells and HIV.* (To be submitted in 2020)

Wolmarans E, Nel S, Durandt C, Mellet J, Pepper MS. *Side population: Its use in the study of cellular heterogeneity and as a potential enrichment tool for rare cell populations.* Stem Cell International. 2018; 108: 2472137.

Wolmarans E, Mellet J, Ambele MA, Durandt C, Pepper MS. *Heterogeneity of cell therapy products.* South African Medical Journal 2019; 109(8b): 24-28.

Durandt C, Potgieter JC, Khoosal R, Nel JG, Herd CL, Mellet J, Rossouw T, Pepper MS. *HIV and haematopoiesis.* South African Medical Journal 2019; 109(8b): 40-45.

II. Presentations at conferences:

Mellet J, Wolmarans E, Durandt C, Alessandrini M, and Pepper MS. *Hematopoietic stem and progenitor cell transcriptomic heterogeneity: analysis at the single-cell level*. South African Tissue Bank Association (SATiBA). October 2017. Centurion, South Africa.

Mellet J, Wolmarans E, Durandt C, Alessandrini M, and Pepper MS. *Hematopoietic stem and progenitor cell transcriptomic heterogeneity: analysis at the single-cell level*. Flagship workshop. October 2017. Pretoria, South Africa.

Mellet J, Durandt C, Ambele M, and Pepper MS. *The effect of StemRegenin-1 on the gene expression of umbilical cord blood derived hematopoietic stem and progenitor cells.* 5th National Cell and Gene Therapy Conference. September 2019. Pretoria, South Africa.

III. Posters at conferences:

Mellet J, Durandt C, and Pepper MS. *Culture of HSCs in vitro according to GMP-standards*. CYTO Conference. June 2016. Washington, USA.

ABSTRACT

Umbilical cord blood (UCB) is a rich source of hematopoietic stem and progenitor cells (HSPCs). There are however limitations to using UCB as a regular source for hematopoietic stem cell transplantation (HSCT). The number of $CD34^+$ HSPCs is limited, while a minimum number of $CD34^+$ HSPCs is required for HSCT, which cannot always be achieved. New developments in HSCT are currently underway to expand current applications and improve safety and efficacy. This necessitates efficient *ex vivo* expansion of these cells to therapeutic numbers. HSCT is being investigated in therapies for non-hematopoietic disorders with the goal of replacing diseased cells or tissue with healthy cells. HSPC-based gene therapy strategies are becoming attractive applications of corrective *ex vivo* gene transfer given the reconstitutive potential of HSCT. The success of these strategies for the treatment of monogenic disorders resulted in the application of HSPC gene therapy being considered for other diseases such as the human immunodeficiency virus (HIV).

The optimal isolation method was determined for increased HSPC purity and viability by testing two different methods, magnetic activated cell sorting (MACS) and fluorescent activated cell sorting (FACS). FACS was considered optimal for our purposes and was used to isolate $CD34^+$ HSPCs for subsequent experiments. Different commercially available serum-free media were tested and compared to standard medium supplemented with foetal bovine serum (FBS). All commercial serum-free media outcompeted the standard medium based on viability and proliferation. Building on the previous work, StemSpan ACF was used to test combinations of cytokines for their expansion potential. The combination containing FLT3L, SCF, TPO, IL-3 and G-CSF resulted in the greatest expansion of HSPCs. The effect of StemRegenin-1 (SR1) on the expansion of HSPCs was explored by adding SR1 to the above-mentioned cytokine combinations. This resulted in minor effects on HSPC expansion based on viability and immunophenotype. Similarly, it resulted in only two significantly downregulated genes, cytochrome P450, family 1, subfamily B, polypeptide 1 (*CYP1B1)* and erythrocyte membrane protein band 4.1-like 3 *(EPB41L3),* in both $CD34^+$ and $CD34^-$ cells compared to non-treated controls. The use of $CD34^+$ HSPCs exclusively expanded with SR1 would be beneficial in cases where the HSPC cell dose of the initial harvested cell therapy product is suboptimal and

therefore not a feasible option for HSCT on its own. Single-cell RNA sequencing was performed on $CD34^+$ HSPCs and four populations were identified, which is in line with previous publications. HSPC gene therapy is a promising approach to treat HIV. However, this type of approach would require the presence of significant numbers of long-term repopulating HSPCs to enable successful long-term engraftment of gene-modified cells. One aspect that could result in this approach not succeeding is the presence of proviral DNA in HSPCs. It would therefore be important to identify a population of HSPCs that is resistant to HIV infection. It was therefore investigated whether HSPCs from leukapheresis products are susceptible to infection with HIV and whether a subset of HSPCs exists that is resistant to infection to use in HIV gene therapies. Unfortunately, this could not be achieved due to loss of viability of HSPCs from leukapheresis products.

Key words: Umbilical cord blood (UCB); hematopoietic stem and progenitor cells (HSPCs); expansion; StemRegenin-1 (SR1); gene expression; immunophenotyping; HIV-1C; co-receptor

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS, SYMBOLS AND UNITS

Abbreviation	Description
7AAD	7-Amino-Actinomycin-D
ACF	Animal component free
ACT	Alberts Cellular Therapy
AhR	Aryl hydrocarbon receptor
AIDS	Acquired immunodeficiency syndrome
ANGPTL	Angiopoietin-like proteins
ANOVA	Analysis of variance
APC	Allophycocyanin
APC-Cy7	Allophycocyanin-Cy7 conjugate
APRIL	A proliferation-inducing ligand
ARC	Agricultural Research Council
ARNT	AhR nuclear translocator
BD	Becton Dickinson, Inc.
BMP	Bone morphogenic protein
BV	Brilliant violet
CAL Factor	Calibration factor
CAR	CXCL12-abundant reticular
cART	Combined antiretroviral treatment
CC	Canonical correlation vectors
CCA	Canonical correlation analysis
CCR5	C-C chemokine receptor type 5
cDNA	Complementary DNA
CD	Cluster of differentiation
CFC	Colony forming cell
CLOUD	Continuum of low-primed undifferentiated
CLP	Common lymphoid progenitor
CMP	Common myeloid progenitor
CO_2	Carbon dioxide

CPD	Citrate phosphate dextrose
CRF	Circulating recombinant forms
cRNA	Complementary RNA
CRU	Clinical research unit
CSIR	Council for Scientific and Industrial Research
CXCL12	C-X-C motif chemokine ligand 12
CXCR4	C-X-C chemokine receptor type 4
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
ds	Double-stranded
EC	Endothelial cells
EDTA	Ethylene-diaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ESC	Embryonic stem cell
FACS	Fluorescent activated cell sorting
FBS	Foetal bovine serum
FC	Fold change
FDA	Food and Drug Administration
FITC	Fluorescein isothiocyanate
FLT3L	FMS-like tyrosine kinase 3 ligand
FS	Forward scatter
GCCP	Good cell culture practice
G-CSF	Granulocyte colony stimulating factor
gDNA	Genomic DNA
rDNA	Recombinant DNA
GFP	Green fluorescent protein
GLP	Good laboratory practice
GM-CSF	Granulocyte-macrophage colony stimulating factor
GMP	Good manufacturing practice

GMP	Granulocyte-macrophage progenitor
GO	Gene Ontology
GVHD	Graft-versus-host disease
Haplo-HSCT	Haploidentical hematopoietic stem cell transplant
HCl	Hydrochloric acid
HDAC	Histone deacetylase
HIV	Human immunodeficiency virus
HIV-1B	HIV-1 subtype B
HIV-1C	HIV-1 subtype C
HLA	Human leukocyte antigen
HOS	Human osteosarcoma
HOXB4	Homeobox B4
HPC	Hematopoietic progenitor cell
HSC	Hematopoietic stem cell
HSCT	Hematopoietic stem cell transplant
HSPC	Hematopoietic stem and progenitor cell
ICMM	Institute for Cellular and Molecular Medicine
IFC	Integrated fluidics circuit
IFN-γ	Interferon-γ
IGF-2	Insulin-like growth factor 2
IGFBP	Insulin-like growth factor 2 binding protein
IL	Interleukin
IMDM	Iscove's modified dulbecco's medium
iPSC	Induced pluripotent stem cells
KCl	Potassium chloride
KH_2PO_4	Monopotassium phosphate
Lin	Lineage marker cocktail (CD3, CD14, CD16, CD19, CD20, CD45)
LMPP	Lymphoid-primed multipotent progenitors
LogFC	Log fold change
LT-HSC	Long-term HSC
LTR	Long terminal repeats

MACS	Magnetic activated cell sorting
MAPK	Mitogen-activated protein kinases
MEP	Megakaryocyte-erythrocyte progenitor
MHC	Major histocompatibility complex
MIP-1α	Macrophage inflammatory protein-1 alpha
MIP-1β	Macrophage inflammatory protein-1 beta
MNC	Mononuclear cell
M-MLV	Moloney murine leukemia virus
MPP	Multipotent progenitor
mRNA	Messenger RNA
MSC	Mesenchymal stromal cell
N1-ICD	Notch1-intracellular domain
NaCl	Sodium chloride
$NaHCO_3$	Sodium bicarbonate
Na_2HPO_4	Disodium phosphate
NAM	Nicotinamide
NaOH	Sodium hydroxide (NaOH)
NF	Nuclear factor
NG2	Neuron-glial antigen 2
NH_4Cl	Ammonium chloride
NICD	National institute for communicable diseases
NIH	National Institutes of Health
NK	Natural killer
NOD	Non-obese diabetic
NSG	NOD/SCID gamma
NUP98	Nucleospin 98
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffered saline
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction

PE	Phycoerythrin
PE-Cy7	Phycoerythrin-Cy7 conjugate
PTN	Pleiotrophin
PGE_2	Prostaglandin E_2
PHA-P	Phytohemagglutinin-P
qPCR	Quantitative PCR
R5	CCR5-tropic
RAM	Random access memory
RANTES	Regulated on activation, normal T cell expressed and secreted
RIN	RNA integrity number
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RPMI	Roswell Park Memorial Institute
RT	Reverse transcription
SABMR	South African Bone Marrow Registry
SCF	Stem cell factor
SCID	Severe combined immunodeficiency
SD	Standard deviation
SDF1	Stromal-derived factor 1
SIV	Simian immunodeficiency viruses
SMART	Switching mechanism at 5′ end of RNA template
SNN	Shared nearest neighbour
SNP	Single nucleotide polymorphisms
SP	Side population
SR1	StemRegenin 1
SS	Side scatter
ss	Single-stranded
TAC	Transcriptome Analysis Console
TCDD	Tetrachlorodibenzo-p-dioxin
TE	Tris-EDTA
TGF-β	Transforming growth factor beta

TNC	Total nucleated cells
TNF-α	Tumour necrosis factor-α
TPM	Transcripts per kilobase million
TPO	Thrombopoietin
t-SNE	T-distributed stochastic neighbour embedding
UCB	Umbilical cord blood
UCBT	Umbilical cord blood transplant
URF	Unique recombinant forms
UV	Ultraviolet
VCAM-1	Vascular cell adhesion molecule 1
VDC	Vybrant Dye Cycle
VPA	Valproic acid
X4	CXCR4-tropic

Unit	Description
^{o}C	Degree Celsius
bp	Base pairs
cells/mL	Cells/millilitre
cells/ µL	Cells/microliter
g	Grams
kg	Kilograms
IU	Infectious units
IU/mL	Infectious units/millilitre
L	Litres
M	Molar
mM	Millimolar
mg	Milligrams
mg/mL	Milligrams/millilitre
min	Minutes
mL	Millilitres
ng	Nanograms

ng/mL	Nanograms/millilitre
ng/ µL	Nanograms/microlitre
nm	Nanometres
nM	Nanomolar
pg/mL	Picograms/millilitre
rpm	Rotations per minute
U	Units
V	Volts
µg	Micrograms
µg/mL	Micrograms/millilitre
µL	Microlitre
µM	Micrometres

Symbol	Description
%	Percentage
=	Equal to
<	Less than
≤	Less than or equal to
>	Greater than
≥	Greater than or equal to
α	Alpha
β	Beta
γ	Gamma
κ	Kappa
rcf (x g)	Relative centrifugal force

CHAPTER 1. INTRODUCTION

Cells are the 'building blocks of life' with the earliest cell being a stem cell, which has the ability to self-renew and differentiate into different cell types that work collectively to ensure a functional system. The concept of a 'stem cell' was originally proposed in 1963 by Till and McCullock after the formation of blood cell colonies post-bone marrow transplantation in the spleen of recipient mice (1,2). This occurrence established two defining characteristics of stem cells, multipotency and self-renewal. Multipotency is the ability of a cell to differentiate and give rise to multiple lineages, while self-renewal is the ability to produce an identical daughter cell which does not differentiate (3).

Stem cells can be broadly divided into three categories based on their differentiation potential, namely totipotent, pluripotent and multipotent. Totipotent stem cells have the highest differentiation potential and are able to form the embryo and extra-embryonic tissue such as the placenta. Examples include zygotes and early blastomeres (4,5). Pluripotent stem cells, which include embryonic stem cells (ESCs), are derived from the inner cell mass of the blastocyst and are able to differentiate into cells of the three germ layers (ectoderm, endoderm and mesoderm) upon appropriate stimulation.

Researchers have identified four essential transcriptions factors for pluripotency, *OCT4*, *SOX2*, *KLF4*, and *C-MYC* (6) which control genes required for cells to remain undifferentiated. Somatic cells can be reprogrammed by artificial stimulation with these four transcription factors, resulting in a reversion to pluripotency; these reprogrammed cells are known as induced pluripotent stem cells (iPSCs). Although iPSCs are equivalent to ESCs in their protein expression, these cells differ in their epigenetic structures. Upon differentiation, cells acquire specific methylation changes which influence induced pluripotency (7). The use of ESCs for treatment is controversial due to potential tumour progression following treatment and ethical concerns relating to the use of these cells. However, iPSCs offer great potential in regenerative and personalised medicine to better understand disease mechanisms and for the treatment of various diseases. Scientists are making progress with iPSCs-derived cells in Parkinson's disease, platelet deficiency disorders and spinal cord injury (8).

Multipotent stem cells, also referred to as adult stem cells or tissue-specific stem cells, exist in an undifferentiated state and have an important function in local tissue repair and maintenance. Hematopoietic stem cells (HSCs) are a prime example of multipotent stem cells. Even though HSCs have been studied for over 50 years and are the best-characterised stem cells, no standard criteria exist to classify HSCs or distinguish them from early hematopoietic progenitor cells (HPCs), which are already primed to a particular lineage. Both HSCs and HPCs reside in the bone marrow and both cell types have the ability to self-renew and differentiate into the different blood lineages. This thesis will therefore collectively refer to these cells as hematopoietic stem/progenitor cells (HSPCs), which encompasses both HSCs and HPCs.

In 1988, murine HSPCs were isolated for the first time from bone marrow using a combination of fluorescent activated cell sorting (FACS) and monoclonal antibodies. Lineage-specific markers are expressed once cells start to differentiate into the various blood lineages. Expression of lineage markers by mature blood cells assists in excluding lineage-committed and differentiated cells during characterisation of HSPCs. A Thy-$1^{low}$$Lin^{-}$$Sca\text{-}1^{+}$ population represented approximately 0.05% of the bone marrow cells and was able to reconstitute the hematopoietic system following transplantation into sub-lethally irradiated mice (9). A study performed in 1996 revealed that one in three $CD34^{-/low}$c-$Kit^{+}$$Sca\text{-}1^{+}$$Lin^{-}$ cells taken from murine bone marrow ensured long-term reconstitution of myeloid and lymphoid lineages following single-cell transplantation into sub-lethally irradiated mice (10).

1.1. THE HEMATOPOIETIC STEM CELL NICHE

The importance of the niche was first postulated in 1978 by Ray Schofield (11) and refers to a microenvironment in which HSPCs primarily reside in a non-cycling state. The location of the niche depends on the developmental stage of the organism and can shift from the placenta and foetal liver in fetuses, to the bone marrow in adults (12). Migration of HSPCs into and out of the niche is crucial during development, as well as later in life for dynamic hematopoietic reconstitution.

The behaviour of HSPCs is controlled by a balance of intrinsic (gene expression changes) and extrinsic factors (surrounding cell signals) that regulate self-renewal, quiescence, proliferation, differentiation and apoptosis. A fine balance between HSPC self-renewal and differentiation is achieved during development to maintain tissue homeostasis and produce cellular diversity of the HSPC population capable of rapid response to stimuli. This balance is achieved through asymmetric cell division, where an identical copy of a cell is produced that retains stem cell properties, while a daughter cell enters the path of differentiation. Disruption of this balance could lead to HSPC depletion or cancer development (13).

The presence of multiple niches within the bone marrow has been proposed (14). Phenotypically distinct HSPCs were found in close proximity to non-hematopoietic cell types, providing evidence that different niches within the bone marrow support different subsets of HSPCs (15). The bone marrow niche consists of bone matrix and various non-hematopoietic cells, including endothelial cells, stromal cells, neuronal cells and adipocytes (14). These specialised cell types and extracellular elements allow integration of signals from the periphery into the bone marrow niches, to which HSPCs respond.

Osteolineage cells (osteoblasts and osteoclasts) were some of the first cells shown to interact with HSPCs and were suggested to play an important role in HSPC fate. Osteoblast precursors are mesenchymal stromal cells (MSCs) that occupy the endosteal surface of flat and trabecular bones between the bone and the bone marrow. These cells are embedded within the bone matrix upon differentiation into osteoblasts. Transplanted $CD34^{+}$ HSPCs have been shown to preferentially localise in the endosteal region of the bone marrow in close proximity to osteolineage cells (16). The osteoblasts that support HSPCs have a distinct phenotype of N-cadherin^{+}CD45^{-} and are regulated by bone morphogenic protein (BMP) (17). Mice deficient in BMP receptor type IA showed increased numbers of N-cadherin^{+}CD45^{-} osteoblasts which correlated with increased numbers of HSPCs (17). In 2003, a study by Calvi *et al.* (18) showed the role of osteoblasts in hematopoiesis through a genetically modified mouse model producing osteoblast cells, resulting in increased osteoblastic cells in bone marrow. This increase was associated with an increase in lineage-negative, long-term repopulating HSPCs in the region between the bone and marrow. However, HSPC maintenance does not solely rely

on the presence of osteoblasts, since impaired osteoblastic function has been shown not to affect long-term HSPCs (19) and expansion of specific modalities of osteolineage cells has been shown to decrease HSPC numbers (20).

It is proposed that multiple subsets of osteolineage cells at different stages of differentiation play an important role in HSPC maintenance (21,22). Immature osteoprogenitor cells are thought to be more important in maintaining the HSPC niche (23), while mature cells are dispensable since they have an inhibitory effect on HSPCs (24). Osteolineage cells and MSCs are believed to regulate quiescence through the secretion of C-X-C motif chemokine ligand 12 (CXCL12), angiopoietin 1 and thrombopoietin (TPO) (25). It has been demonstrated that deletion of CXCL12, a major chemoattractant and retention factor in the bone marrow, in osteoprogenitors and MSCs resulted in increased mobilisation of HSPCs (26,27).

Evidence further suggests the presence of another niche component referred to as the vascular niche, involving perivascular cells. A variety of perivascular cells have been identified to date, including CXCL12-abundant reticular (CAR) cells, nestin-green fluorescence protein $(GFP)^{dim}$/leptinR^{+} MSCs and nestin-GFP^{bright}/neuron-glial antigen 2 $(NG2)^{+}$ pericytes. Increased numbers of HSPCs are present when located adjacent to CAR cells, (28), while ablation of CAR cells show reduction in osteogenic and adipogenic lineage differentiation of non-hematopoietic cells in the bone marrow. Deletion of CAR cells has further resulted in the depletion of mature lineages such as lymphoid lineages, suggesting a wider role of these cells in the niche (14).

Nestin-GFP^{low} cells express genes associated with HSPC retention, such as *CXCL12* and vascular cell adhesion molecule 1 *(VCAM-1)*. These genes are generally downregulated upon granulocyte colony stimulating factor (G-CSF) mobilisation. Depletion of Nestin-expressing cells in mice resulted in a 50% decrease in HSPCs (29). Nestin-GFP^{bright} cells are close to the endosteum and denote a quiescent niche, while Nestin-GFP^{low} cells surround the sinusoidal vessels and serve as a proliferative niche (22,28) from which HSPCs are mobilised for differentiation.

The presence of multiple cell types and their respective roles in the niche highlights the heterogeneity and complexity of these supportive cells required for HSPC maintenance. A recent study tried to decipher the bone marrow composition at a single-cell level, revealing the presence of 17 transcriptionally distinct sub-populations within the bone marrow of mice. These sub-populations expressed genes relating to MSCs, osteolineage cells, chondrocytes, fibroblasts, endothelial cells and pericytes. Some populations showed considerable heterogeneity within these subsets (30). This complex network of interactions has not been fully elucidated and additional studies are required to identify the different components, cell types, and interactions between the various cells in the niche.

1.2. HEMATOPOIESIS

The mammalian blood system contains more than 10 distinct mature blood cell types that all differentiate from HSPCs indicating the remarkable differentiation potential of these cells. Haematopoiesis involves the progression of HSPCs through different stages of differentiation for continuous maintenance of mature blood cell production. This process is well-studied and involves many cytokines, chemokines, cell-to-cell interactions and extracellular matrix interactions. However, the *in vivo* functionality, frequency and longevity of HSPCs in humans are not clearly defined.

The classical model of haematopoiesis suggests a hierarchical structure with long-term HSPCs at the apex of the hierarchy (Figure 1.1). Long-term HSPCs have self-renewal capabilities throughout life and give rise to short-term HSPCs with limited self-renewal capabilities. Short-term HSPCs differentiate to form multipotent progenitors (MPP), which are precursors of common lymphoid and myeloid progenitors (CLP/CMP). MPPs have no self-renewal abilities but are capable of full lineage differentiation. CLP progeny differentiate into lymphoid and natural killer (NK) cells. CMP progeny differentiate into granulocyte–macrophage progenitors (GMP) and megakaryocyte–erythrocyte progenitors (MEP), which will differentiate into granulocytes and macrophages, and erythrocytes and megakaryocytes, respectively (31).

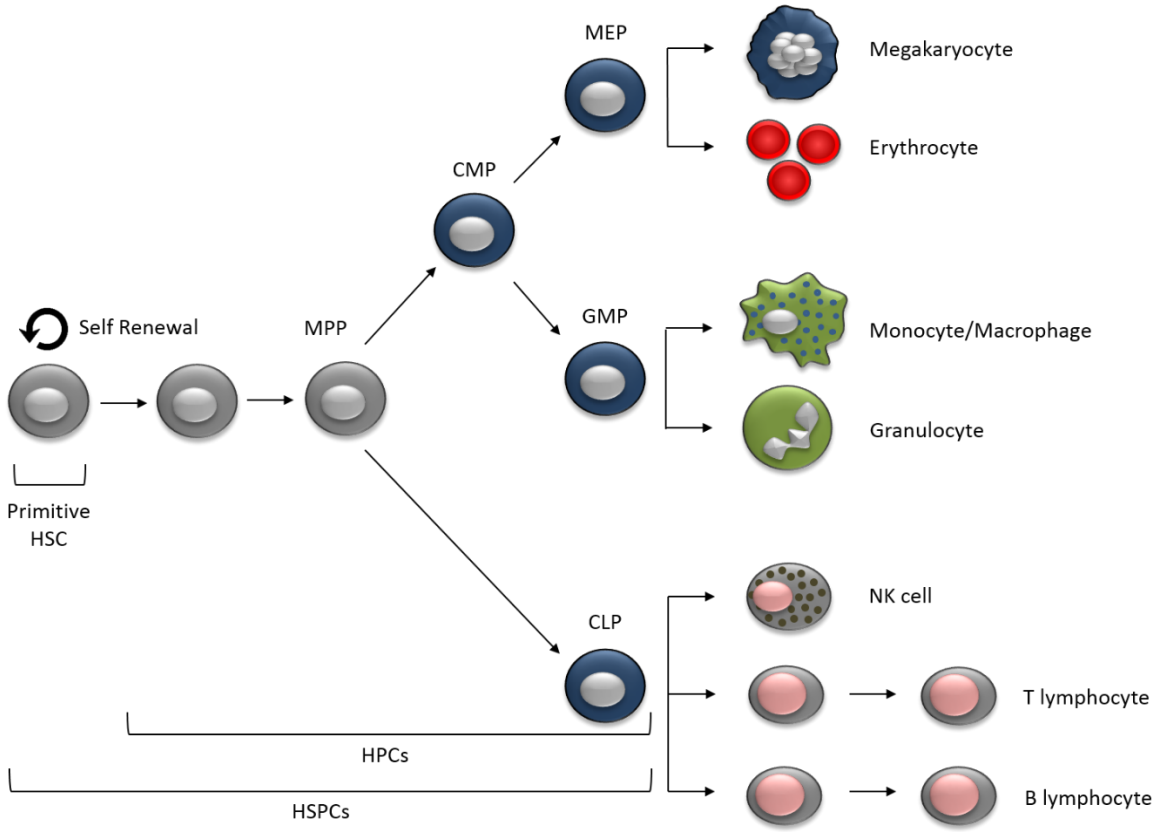


Figure 1.1. The classical model of hematopoiesis. Schematic illustration of the differentiation of hematopoietic stem/progenitor cells (HSPCs) into mature blood cell types. CLP, common lymphoid progenitor; CMP, common myeloid progenitor; GMP, granulocyte/monocyte progenitor; HPC, hematopoietic progenitor cell; HSC, hematopoietic stem cell; HSPC, hematopoietic stem and progenitor cell; MEP, megakaryocyte-erythrocyte progenitor; MK, megakaryocytes; MPP, multipotent progenitor; NK, natural killer cell. (Figure was created by Juanita Mellet and Candice Herd).

Even though hematopoiesis has been studied for many years, the HSPC hierarchy is more complex than previously assumed. Recent studies provide data that cannot be explained by the classical model. This includes myeloid-restricted progenitors with long-term repopulating potential (32) and HSPCs expressing platelet-biased genes while having the ability to self-renew (33). These two examples are only a fraction of the data demonstrating non-classical differentiation potential of HSPCs. Technological advances in the past decade resulted in a revised model of hematopoiesis, illustrated in Figure 1.2.

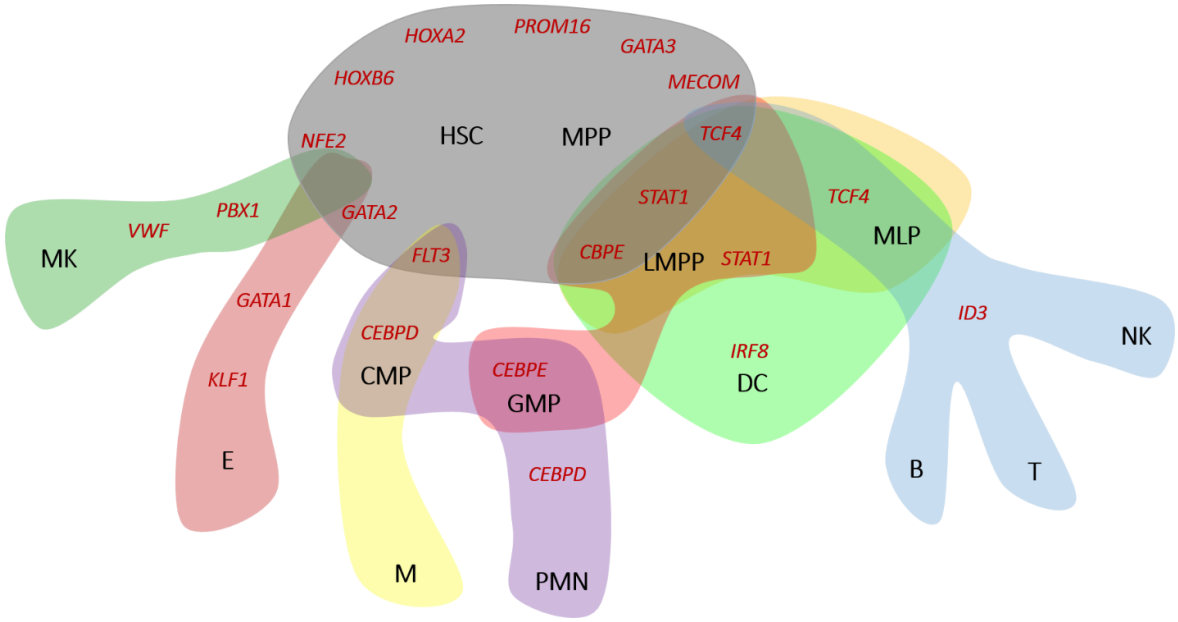


Figure 1.2. The revised model of hematopoiesis. A revised model of the hematopoietic hierarchy based on current literature. The hematopoietic populations are indicated in black, whereas differentiation-specific genes are indicated in red. B, B cell, CMP, common myeloid progenitor; DC, dendritic cell; E, erythrocyte; GMP, granulocyte/monocyte progenitor; HSC, hematopoietic stem cell; LMPP, lymphoid-primed multipotent progenitor; M, monocyte; MK, megakaryocytes; MLP, multipotent lymphoid progenitor; MPP, multipotent progenitor; NK, natural killer cell; PMN, polymorphonucleated cell; T, T cell. (Figure was created by Juanita Mellet, adapted from (34)).

Based on the revised model, HSCs do not pass through subsequent stages of differentiation and are not subjected to cell fate decisions at various discrete branching points. This model rather proposes a Continuum of LOw-primed UnDifferentiated ('CLOUD')-HSPCs, where HSCs and MPPs share similar gene expression patterns (*HOXB6/HOXA2/PRDM16*) despite different metabolic states. It is further suggested that HSCs and MPPs start to show signs of lineage priming while still in the CLOUD (35). Lineage-priming exists at low levels in the CLOUD and is reinforced upon differentiation into the various lineages. For many years, it was thought that erythrocytes and megakaryocytes originate from a common progenitor, MEP. However, recent evidence suggests multiple MEPs are already primed for differentiation into either erythrocytes (E-MEP) or megakaryocytes (MK-MEP) and only a small fraction exist that have true bipotent differentiation potential (pre-MEP) (36). Additionally, a population of lymphoid progenitors was identified containing partial myeloid potential, referred to as lymphoid-primed multipotent progenitors (LMPP) expressing *CBPE/STAT1/TCF4*, suggesting that granulocytes

can arise from both lymphoid and myeloid primed cells (37). With the advent of new technologies the hematopoietic process is continually being re-evaluated at increasing resolution.

1.3. PHENOTYPE AND HETEROGENEITY OF HSPCs

The cluster of differentiation (CD)34 marker was identified on HSPCs in 1984 (38) and is still routinely used to identify, enumerate and enrich for these cells. Although CD34 is primarily used to identify HSPCs, not all HSPCs express this marker. A rare population of HSPCs is CD34-negative ($CD34^-$) and becomes CD34-positive ($CD34^+$) prior to cell division (39). This population was first identified in murine bone marrow studies, in which a single $CD34^-$ HSPC was able to repopulate multiple lineages (10). Studying the $CD34^-$ population is difficult, since a positive marker for this population remains to be identified. These cells are highly quiescent and demonstrate low-level engraftment capabilities compared to $CD34^+$ cells (40). However, the $CD34^-$ cells are competent in long-term engraftment even though they remain undifferentiated for up to six weeks post-transplantation *in vivo* (41). Expression of CD34 is therefore associated with cell cycle entry, metabolic activation, mobilisation and homing of HSPCs (42).

The number of markers present on the cell surface is extremely vast. Different cell types are often phenotypically similar with regard to a small number of cell surface markers, while being genotypically, transcriptionally and functionally distinct. Absence of lineage markers (Lin^-), CD38, CD133, CD90, CD49f, CD117 (c-Kit) and C-X-C chemokine receptor type 4 (CXCR4) are among many cell surface markers used in combination with CD34 to identify and isolate different sub-populations of HSPCs.

CD133 is also commonly used to identify HSPCs, and enriches for a primitive population of HSPCs (43). The primitive cells are a small subset of cells that have the capacity to engraft and reconstitute the entire bone marrow. A recent study by Takahashi and colleagues identified a sub-population of $CD133^+$ cells as being $CD34^-$ HSPCs (44). Despite CD34 being used as the

marker to identify HSPCs, CD133 may be a better marker to identify both CD34-positive and -negative fractions of HSPCs. A variety of genes associated with HSPC maintenance and pluripotency (43) support the primitive state of these cells.

Furthermore, the CD38 surface marker is a glycoprotein present on several immune cells, including $CD4^+$ and $CD8^+$ lymphocytes, B lymphocytes and NK cells. CD38 functions in cell adhesion, signal transduction and cell signalling. This marker is generally absent on early HSPCs and is acquired upon differentiation (45,46).

Cell-surface expression of CD90 (THY1) is associated with long-term engraftment of HSPCs, but is generally expressed on a small fraction of cells (47). It was proposed that CD90 is expressed on more primitive HSCs, while MPPs lack CD90 expression and that this marker could therefore be used to separate these two populations. However, more than a third of recipients transplanted with $CD90^-$ cells engrafted in primary and secondary transplants (48). This raises uncertainty about whether CD90 is indicative of functional characteristics of HSPCs. This might indicate that both HSCs and MPPs have engraftment potential.

Integrin α6, also known as CD49f, is a stem cell marker expressed on more than 30 different stem cell populations, including HSPCs, pluripotent- and multipotent stem cells. This is the only marker commonly expressed on all stem cell populations and speaks to its importance in stem cell biology (49). Single HSPCs expressing CD49f in combination with additional HSPC markers ($CD34^+CD38^-CD45RA^-Thy1^+$ and $CD34^+CD38^-CD45RA^-Thy1^-$) engrafted in the bone marrow of transplant recipients and gave rise to multi-lineage grafts (45), demonstrating the utility of this marker in stem cell research.

CD117 (c-KIT) is the receptor for stem cell factor (SCF), an important HSPC growth factor *in vitro* and *in vivo*. CD117 is mainly expressed on multipotent and oligopotent HSPCs, but is also observed on more mature hematopoietic cells (50). Similar to the expression of CD90, CD117 is expressed on a small proportion of $CD34^+$ HSPCs (47).

CXCR4 is expressed on HSPCs and together with its ligand, stromal derived factor 1 (SDF1), also referred to as CXCL12, plays an important role in retaining HSPCs in the bone marrow and directing migration from the circulation to the niche (51). Blocking CXCR4 results in the impaired ability of $CD34^{+}CD38^{-}$ HSPCs to home and engraft in non-obese diabetic/severe combined immunodeficiency (NOD/SCID) mice (52). *Ex vivo* pre-treatment of $CD34^{+}$ HSPCs with SDF1α increased engraftment in NOD/SCID mice at a lower doses (53), while inhibition of the CXCR4-SDF1 pathway leads to proliferation of HSPCs (54).

Enhanced homing of human $CD34^{+}$ cells from mobilised peripheral blood and umbilical cord blood (UCB) to the bone marrow was observed following intravenous administration of SDF1 in mice (55). Drugs that block CXCR4, such as Mozobil (AMD3100) and Filgrastim, cause degradation of SDF1, leading to HSPC mobilisation from the niche and proliferation (56).

Historically, studies have focused on purified populations expressing specific cell surface markers, while neglecting transcriptional heterogeneity. The $CD34^{+}$ HSPC population represents a heterogeneous population of cells at different developmental stages (35), which is not restricted to $CD34^{+}$ cells as a whole, but expands to the populations within the $CD34^{+}$ population (36). Advances in the field of single-cell technology allow researchers to study multiple "-omes" (genome, transcriptome and proteome) of single cells simultaneously to further understand heterogeneity. Such studies are revealing the complexity within phenotypically defined populations, cellular differentiation and overall regulatory networks. Furthermore, remarkable functional heterogeneity exist within phenotypic and transcriptionally defined HSPC populations (57). True HSPCs can only be identified based on their ability to reconstitute the blood lineages in sub-lethally irradiated mice to assess the "stemness" of these cells, while self-renewal and the ability to form progeny are assessed upon serial transplantation.

1.4. HEMATOPOIETIC STEM CELL TRANSPLANTATION

The first successful hematopoietic stem/progenitor cell transplantation (HSCT) was performed in 1956 by Dr James Thomas, which involved transplantation between identical twins and progressed to HSCT between related siblings in 1968 (58). HSPCs are also used clinically to treat

the adverse effects of high-dose chemotherapy in cancer patients. Chemotherapy is a potent drug or combination of drugs used to kill rapidly-dividing cells, a characteristic of cancer cells. One of the adverse effects of chemotherapy is myelosuppression, where the stem cells within the bone marrow are destroyed. HSCT is utilised to reconstitute the entire hematopoietic system following chemotherapy. Transplanted stem cells should ideally migrate to the bone marrow, a process known as stem cell homing, for complete reconstitution of the damaged/destroyed cells.

Autologous transplantation, where the patient's own cells are infused, have been shown to be effective for specific disorders, such as lymphoma and myeloma. Autologous HSCT eliminates the need for immunosuppressive agents since donor and recipient cells are genetically identical and there is no risk of developing graft-versus-host disease (GVHD), resulting in reduced morbidity and mortality. The disadvantages include potential tumour cell contamination in the graft, which could result in relapse, and absence of a graft-vs-tumour effect. In contrast, allogeneic HSCT involves the procurement of cells from a human leukocyte antigen (HLA)-matched donor. Jean Dausset showed that grafts between siblings are more successful than grafts between unrelated individuals (59). The preferred donor is therefore an HLA-identical sibling, which is not always possible to achieve, whereby a partially matched sibling or family member donates instead (matched-related donor). HLA-matched unrelated donors also exist in the general population when a matched family member is not available. Unrelated donors are usually found through a National Bone Marrow Registry. In 1990, the South African Bone Marrow Registry (SABMR) was established, and assists in the identification of an adequately HLA-matched donor for those individuals not fortunate enough to possess an HLA-matched relative.

Haploidentical HSCT (haplo-HSCT) involves transplantation of T cell-depleted cells usually from a matched-related donor. This is a favourable approach since the majority of donors would have a partially matched family member; however, this approach has a high relapse risk, risk of developing GVHD and immune reconstitution is slow. A new method was developed for haplo-HSCT using cyclophosphamide to lower the risk of GVHD following HSCT (60). There are several advantages associated with allogeneic HSCT that include the absence of malignant cells

and the potential graft-versus-tumour effect, while the disadvantages include difficulty in finding a matched-unrelated donor and development of GVHD, leading to morbidity and mortality.

Quantification of $CD34^+$ HSPCs is widely used for conventional transplantation purposes. Although the optimal number of $CD34^+$ HSPCs to ensure hematopoietic recovery is a subject for debate, a minimum $CD34^+$ cell dose of $2 – 5 \times 10^6$/kg body weight is required to achieve consistent engraftment (61). However, for the purpose of gene therapy where cells undergo manipulation, increased numbers of $CD34^+$ cells need to be harvested due to cell loss during manufacturing procedures, such as mononuclear cell (MNC) purification, positive selection and transduction of $CD34^+$ cells. Harvesting increased volumes of bone marrow for gene therapy is well-tolerated and will allow for an adequate number $CD34^+$ HSPCs for gene-modification and subsequent *in vivo* administration (62).

The most widely applied HSPC-based therapy is conventional HSCT for the treatment of a variety of acquired and inherited malignant and non-malignant disorders. These include haematological malignancies (leukemia, lymphoma and myeloma) and non-malignant disorders (thalassemia, sickle cell anaemia and SCID). HSCT is also being investigated in therapies for non-hematopoietic disorders with the goal to replace diseased cells or tissue with healthy cells. Multiple clinical studies employing HSCT are in progress, including transplantation to treat neurological disorders (63) and autoimmune diseases (64).

Given the reconstitutive potential of HSCT, HSPC gene therapy is an attractive approach integrating *ex vivo* gene transfer strategies and well-established HSCT practice to treat several diseases. Gene therapy approaches for the treatment of various diseases aim to correct non-functional genes with various genome-editing technologies, for endogenous expression of corrected/functional genes. Gene therapy strategies targeting HSPCs followed by HSCT were performed in the 1990s for the treatment of X-linked SCID and were successful for 9 of 10 patients. However, four of nine patients developed leukemia 31 – 68 months post-treatment, questioning the future of gene therapy. This was mainly the result of insertional mutagenesis relating to the use of gammaretroviral vectors inserting near *LMO2* and *BMI1* proto-oncogenes

(65), resulting in leukemia. Safer lentiviral vectors currently available with a low risk of insertional mutagenesis resulted in the re-emergence of this treatment strategy which has already shown great potential in clinical trials for the treatment of primary immunodeficiencies (66). The success of these strategies for the treatment of monogenic disorders resulted in them being considered for other diseases, such as the human immunodeficiency virus (HIV). The first clinical trial for the use of autologous HSCT to treat HIV was approved by the Food and Drug Administration (FDA) in 2016 (NCT02500849@clinicaltrials.gov) (67).

1.5. SOURCES OF HSPCs

Adult HSPCs reside primarily within the bone marrow. The first HSCT, 50 years ago, made use of bone marrow derived HSPCs. However, the use of bone marrow for HSCT is decreasing as other, less invasive procedures are increasingly available, such as obtaining HSPCs from UCB and mobilised peripheral blood.

1.5.1. Mobilised peripheral blood HSPCs

Under normal conditions, very few HSPCs circulate in the peripheral blood. Treatment with mobilising agents (such as Mozobil and Filgrastim) disrupts the interaction of HSPCs with supporting niche cells in the bone marrow leading to mobilisation of HSPCs into the peripheral circulation. Mobilised HSPCs can then be harvested in the peripheral blood, a less invasive procedure when compared to acquiring HSPCs directly from the bone marrow. Granulocyte colony stimulating factor (G-CSF) is commonly used to mobilise HSPCs; however, several alternative mobilisation regimens are available. Collection of mobilised peripheral blood is performed for > 90% of autologous HSCT and > 70% of allogeneic HSCT indications (61). The advantages of using mobilised peripheral blood over bone marrow include a relatively easy collection process, faster hematopoietic recovery and better immunological reconstitution (61). Monitoring the number of $CD34^+$ cells in the peripheral blood by flow cytometry is reliably used to predict harvest time. The final yield of $CD34^+$ HSPC in the harvested product is often variable; however, a $CD34^+$ count ≥ $10 – 20 \times 10^3$/mL is considered reasonable for a successful harvest. Despite multiple mobilisation regimens available, many patients do not achieve mobilisation of the minimum required number of $CD34^+$ cells for transplantation (61).

When bone marrow and/or mobilised peripheral blood HSPCs are donated for allogeneic HSCT, HLA typing is performed for 10 HLA loci (HLA-A, -B, -C, -DRB1, and -DQB1), and a 9/10 or 10/10 match between donors and recipients is imperative to ensure engraftment success (68). As a result of the high degree of matching necessary for the use of bone marrow- and mobilised peripheral blood HSPCs, UCB has become an alternative source of HSPCs for transplantation. Advantages associated with the use of UCB include less stringent HLA matching, requiring three or four HLA loci to match between donors and recipients. In addition, there is a lower risk of developing GVHD following UCB transplantation (UCBT). UCB is also easily accessible and poses no risk to either mother or child during harvesting.

1.5.2. Umbilical cord blood HSPCs

UCB is an alternative, rich source of HSPCs (69) obtained from the foetal circulation in the placenta through the umbilical vein. This source of stem cells is particularly useful for allogeneic transplantation due to the immaturity of the immune cells in UCB as a result of minimal exposure to antigens (70). UCB has been used successfully for HSCT since 1989 (71).

A disadvantage of UCB is the limited number of $CD34^+$ cells in a single UCB unit due to the volume of cord blood available for harvest. Since $CD34^+$ cells are transplanted based on recipient body weight, the low number of HSPCs present may require two or more UCB units for transplantation. Although UCB requires less stringent donor-recipient matching, partial HLA matching between the units and recipient and also between the different units is essential (72). The safety and efficacy of using double UCBT have been documented in a retrospective study and were comparable to single UCBT. Double UCBT is therefore an alternative when an adequately dosed single unit is not available (73). Interestingly, following double UCBT only one unit is ultimately responsible for long-term engraftment in the majority of cases (74,75).

Another disadvantage associated with the use of UCB is delayed platelet and neutrophil recovery following UCBT, which tend to be slower when compared to bone marrow and mobilised peripheral blood HSCT, corresponding to an increased risk of infection and hospitalisation (76). These limitations can be overcome through the *ex vivo* expansion of UCB units to achieve clinically relevant $CD34^+$ numbers in single units and faster hematopoietic

reconstitution. Concerns regarding the effect of expansion on cells have led to the use of non-expanded (non-manipulated) units in combination with *ex vivo* expanded (manipulated) UCB units for HSCT (77–79) (Figure 1.3).

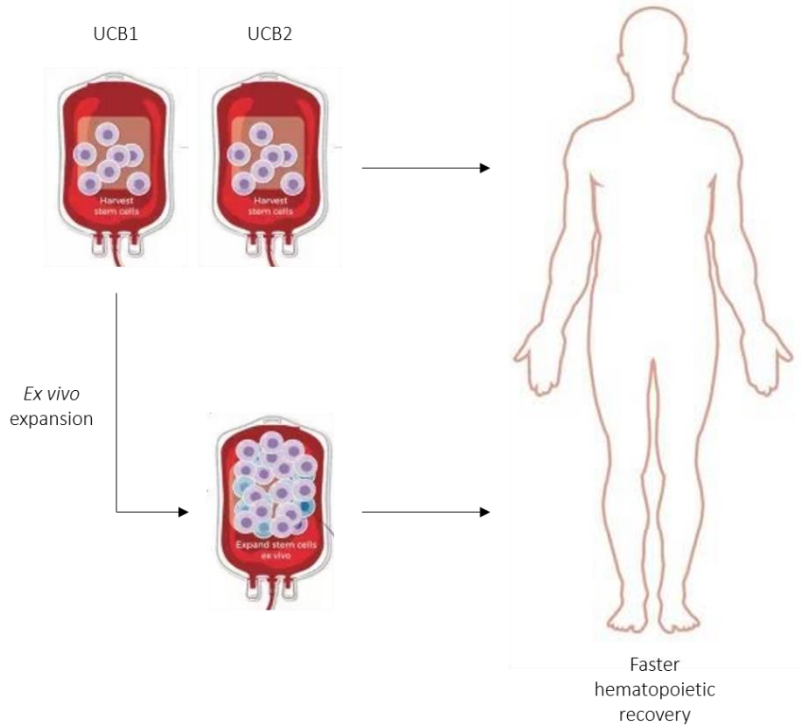


Figure 1.3. UCBT combining non-expanded and expanded UCB units. The use of a non-expanded (non-manipulated) UCB unit in combination with *ex vivo* expanded (manipulated) UCB HSPCs for HSCT (Figure was created by Juanita Mellet, adapted from (78)).

The use of HSPCs from each source has numerous advantages and disadvantages and the choice of HSPC source ultimately depends on donor availability, patient age, disease status and transplant centre preferences.

1.6. *EX VIVO* EXPANSION OF HSPCs

Despite the availability of multiple HSPC sources for life-saving transplantation, the low frequency of $CD34^+$ and $CD34^-$ HSPCs in transplant products means that sufficient HSPC numbers for successful transplantation is not always achieved. This, in conjunction with the restrictions of HLA-matching even in UCBT, necessitates novel approaches to increase HSPC numbers for transplantation, and resultant short- and long-term healthy hematopoietic

reconstitution. *Ex vivo* expansion of these cells to therapeutic numbers is one such approach that is currently employed in research and clinical studies. However, expansion of HSPCs outside their natural environment is challenging, since *in vitro* culture conditions induce spontaneous differentiation resulting in reduced stem cell characteristics. Numerous efforts have been made to expand these cells and simulate their natural environment *in vitro*. These efforts aimed to identify *ex vivo* conditions that could promote self-renewal and proliferation of HSPCs, while at the same time restricting their differentiation. Progress in the field of HSPC expansion has identified intrinsic regulators, biochemical pathways and small molecules that support expansion of these cells.

1.6.1. Growth factor-mediated expansion

Bone marrow stromal cells express numerous cytokines involved in maintaining HSPCs in the bone marrow including FMS-like tyrosine kinase 3 ligand (FLT3L), SCF, TPO, interleukin (IL)-3 and IL-6, granulocyte-macrophage colony stimulating factor (GM-CSF) and G-CSF (80–82) to name a few. Initial attempts to expand HSPCs *ex vivo* involved supplementing standard hematopoietic cell culture conditions with different combinations of the above-mentioned cytokines (83). No standard culture condition exists for the culturing of HSPCs and the variability between laboratories makes it difficult to directly compare results.

The cytokine combination, SCF, FLT3L, and TPO has been extensively studied and is essential for the *ex vivo* expansion of HSPCs, collectively referred to as 'early acting' cytokines. These cytokines are used in several clinical and pre-clinical HSPC expansion protocols (77,78,84) and have been shown to promote quiescence and self-renewal (83). The presence of FLT3L in cultures has shown enhanced HSPC survival and self-renewal (85). FLT3L in combination with IL-6 has been shown to improve the expansion of early HSPCs (86). *Ex vivo* expansion of HSPCs with SCF improves engraftment and expansion, while maintaining self-renewal and maintenance *in vivo* (87). The *ex vivo* use of SCF supplemented with various cytokines such as FLT3L, IL-3/6, and TPO has been shown to effectively expand HSPCs (88). HSPCs express c-mpl, the TPO receptor, which means that TPO can act directly on these cells (89). Thrombopoietin (TPO) suppresses differentiation and promotes survival of HSPCs (90) and assists in self-replication of these cells *ex vivo* (91). The use of IL-3 and IL-6 is critical for the expansion of

$CD133^+$ HSPCs, and IL-6 is important in maintaining an immature HSPC phenotype following expansion (92). Expansion of HSPCs exclusively with cytokines has not significantly improved engraftment parameters (93). The first HSCT that used cytokine-expanded and non-expanded HSPCs showed no improvement in engraftment (94), stressing the need for improved expansion methods.

Gene expression profiling of supportive liver stromal cells in mice revealed that these cells express insulin-like growth factor 2 (IGF-2), its binding protein (IGFBP-2) and several angiopoietin-like proteins (ANGPTLs). Supplementing hematopoietic culture medium with one or combinations of these growth factors increased expansion of both mouse and human HSPCs with long-term multilineage repopulating ability in mice (95,96). *In vivo* studies suggest that ANGPTLs form part of the foetal liver and adult microenvironment of HSPCs (97,98). Pleiotrophin (PTN) is another secreted growth factor that recently started attracting attention as a neuromodulator with several functions in neuronal development (99). PTN is also secreted by sinusoidal endothelial cells in the vascular HSPC niche regulating self-renewal, homing and retention of HSPCs inside the bone marrow (100). Supplementing hematopoietic culture conditions with growth factors alone deprives these cells of the supporting niche cell influences, which assists in retaining HSPCs in an undifferentiated state.

1.6.2. Stromal cell co-cultures

Dexter *et al.* (101) were the first to achieve long-term (> 6 months) *in vitro* culture of bone marrow-derived HSPCs using a stromal cell feeder layer. Since then, many co-culturing systems have been used with various sources of feeder cells including bone marrow stromal cells, immortalised stromal cells and endothelial cells. Co-culturing of HSPCs with different cell types exhibiting feeder properties is useful for the *ex vivo* expansion of these cells, but also allows researchers to establish the relationship between these cells within the bone marrow niche. Co-culture with stromal cells leads to increased phenotypic markers associated with primitive HSPCs. MSCs play an important role in the bone marrow niche and maintain HSPC quiescence. MSCs differentiated into osteoblasts have also been used in co-culture experiments representing the endosteal niche (102). MSCs secrete a plethora of cytokines, such as SCF, SDF1/CXCL12, transforming growth factor beta (TGF-β) and FLT3L, all of which support self-

renewal and quiescence of HSPCs (103). Exogenous cytokine supplementation has been shown to synergise with stromal cell cytokine secretion (104,105) in conserving primitive HSPCs. Even though bone marrow-derived MSCs are most commonly used as a feeder layer, MSCs from other sources have also been used. While MSCs from different sources might differ in proliferation and differentiation potential, they usually share similar characteristics (106,107). MSCs derived from human adipose tissue have been shown to support hematopoiesis *in vitro* and *in vivo* (108) and are a good alternative to bone marrow MSCs. These MSCs represent an easily accessible source of feeder cells for expansion of HSPCs for clinical use (109).

1.6.3. Small soluble factors

Small soluble molecules have been used for several years to induce *ex vivo* expansion of HSPCs since these molecules are relatively cheap and easy to manufacture. Histone deacetylase (HDAC) inhibitors were some of the first small molecules identified to have an effect of HSPC expansion. Treatment of HSPCs with an HDAC inhibitor and valproic acid (VPA) has been shown to significantly increase the number of $CD34^{+}CD90^{+}$ HSPCs up to approximately 60%. Even though homing of VPA-treated cells was inhibited due to decreased expression of CXCR4, these cells maintained their differentiation and long-term engraftment potential *in vivo* (110).

Supplementing *ex vivo* HSPC cultures with nicotinamide (NAM), a form of vitamin B3, resulted in delayed differentiation and increased engraftment of UCB-derived $CD34^{+}$ HSPCs. Expansion in the presence of NAM increases the number of HSPCs ($CD34^{+}CD38^{-}$) and decreases the number of lineage-restricted cells (111).

StemRegenin-1 (SR1), an aryl hydrocarbon receptor (AhR) antagonist, has been used for the *ex vivo* expansion of $CD34^{+}$ HSPCs with long-term engraftment potential (112–114). AhR regulates toxic environmental effects and plays a role in modulating hematopoiesis and the immune system (113). AhR undergoes several conformational changes to become a DNA-binding transcription factor able to upregulate certain genes, while suppressing others. A phase I/II clinical trial indicated improved neutrophil and platelet engraftment following UCBT with a single *ex vivo* SR1-expanded UCB unit.

A recent study showed that a pyrimidoindole derivative, UM171, identified by screening a library of 5280 low molecular weight compounds, likewise induces human HSPC self-renewal and *ex vivo* expansion. Culture with UM171 resulted in improved expansion of primitive human $CD34^+$ cells from mobilised peripheral blood compared to controls. Although UM171 was able to successfully expand human and macaque $CD34^+$ cells, it was unable to expand mouse HSPCs. Successful expansion with UM171 requires constant presence in the culture medium, which suggests that its effect is reversible (115). Transplantation studies revealed that HSPC division rate is not affected by UM171; even so, expansion of the absolute number of primitive cells was observed by culturing with UM171, which further increased when used in combination with SR1. Lymphoid-deficient differentiation and augmentation of myeloid cell differentiation at 30 weeks post-transplantation were observed following engraftment with UM171-cultured HSPCs, suggesting that UM171 targets more primitive cells than SR1. Lymphoid-deficient differentiation and augmentation of myeloid cell differentiation at 30 weeks post-transplantation were observed following engraftment with UM171-cultured HSPCs. RNA sequencing (RNA-seq) revealed fewer transcripts associated with erythrocyte and megakaryocyte differentiation and upregulation of *PROCR* (CD201), a cell surface marker expressed on long-term HSPCs in UM171-expanded HSPCs (115).

Reactive oxygen species have been shown to limit the life span of HSPCs through activation of the p38 mitogen-activated protein kinases (MAPK) pathway as a stress response. A recently identified novel structural analog of SB203580 (p38-MAPK inhibitor) expanded HSPCs from non-enriched UCB MNCs. The expanded cells were enriched for CD34, CD90 and CD49f and showed enhanced colony formation and long-term repopulating ability after transplantation into primary and secondary NOD/SCID gamma (NSG) mice (116).

1.6.4. Developmental pathways

Several key developmental genes have been identified that regulate hematopoiesis, and have been investigated as agonists in HSPC expansion. These include: Homeobox B4 *(HOXB4)*, Nucleosporin 98 *(NUP98)*, Notch ligands and prostaglandin E2 (PGE_2). Four Notch receptors (Notch 1 – 4) and five ligands (Delta-like1, 3, and 4 and Jagged1 and 2) have been identified in mammals and most are expressed on human HSPCs (117). Although Notch-signaling is not

crucial for *in vivo* expansion of HSPCs, it has been shown to play a role *in vitro.* The importance of Notch-signaling in hematopoiesis is a subject of debate due to contradictory results from several studies. Early reports indicated that when HSPCs were transduced to express the active form of Notch1-intracellular domain (N1-ICD) or the Notch target *HES1*, increased numbers of HSPCs with self-renewal capabilities were observed (118,119). Furthermore, *in vitro* exposure of primitive HSPCs to Notch ligands such as Jagged1 and Delta-like1 ligand promoted self-renewal and inhibited differentiation (18,120). Despite these positive results *in vitro*, the role of Notch-signaling remains to be established since loss of function assays showed no involvement of the Notch pathway in hematopoiesis (41,121).

Homeobox transcription factors are highly active during early development and also play an important role in HSPC development and especially primitive growth (122). *HOXB4* promotes *in vitro* and *in vivo* self-renewal and regulates proliferation and differentiation of murine HSPCs, and to a lesser extent human HSPCs (123). Overexpression of HOXB4 through exogenous supplementation with a recombinant fusion protein Tat-HOXB4, resulted in expansion of primitive HSPCs (124). In addition, short-term expression of HOXB4 through integration-deficient lentiviral vectors in HSPCs led to improved multi-organ engraftment of murine HSPCs after transplantation (125). The clinical translation of this approach has been discouraged due to the instability of the protein, its short half-life (approximately one hour) and production difficulty. NUP98-HOX fusion also promoted expansion of HSPCs suppressing differentiation (126). However, in one study the differentiation block led to leukemogenesis (127).

Prostaglandin E2 (PGE_2) is a primary mediator of inflammation and is critical for the activation, migration and cytokine secretion by various immune cells, especially immune cells related to innate immunity (128). PGE_2 increases CXCR4 expression on HSPCs (129) which explains the enhanced and preferential engraftment observed when HSPCs are pre-treated with PGE_2 prior to double UCBT in a phase I clinical trial. In this trial, multilineage hematopoiesis was sustained for up to 27 months (130).

1.7. HIV AND HEMATOPOIESIS

The human immunodeficiency virus has several concomitant effects once an individual becomes infected, including various haematological disorders such as cytopenias and myelodysplasia. These HIV-associated haematological abnormalities result in reduced numbers of erythrocytes, leukocytes and platelets in infected individuals.

In 2007, Timothy Brown, also known as the “Berlin patient”, was cured of HIV after receiving HSCT for acute myeloid leukaemia from a C-C chemokine receptor type 5 (CCR5)-null stem cell donor. A search was initiated to identify this naturally occurring mutation in the CCR5 gene (delta-32 deletion) in the donor cells; lack of CCR5 renders all reconstituted cells post-transplantation resistant to CCR5-tropic (R5) HIV-1 infection (131). Similar to the Berlin patient, a second HIV-infected patient underwent allogeneic HSCT for Hodgkin's lymphoma using cells from a CCR5-null donor. Combination antiretroviral therapy (cART) was interrupted 16 months post-transplant and the patient is still in HIV remission with undetectable viral nucleic acids at 18 months (132). Although it is premature to conclude that a cure was achieved, this is promising for future research.

Gene therapy strategies to develop a possible cure for HIV through genetic modification of autologous HSPCs by disrupting the CCR5 chemokine receptor are ongoing. The capabilities of HSPCs to self-renew and differentiate into all the different blood lineages make these cells uniquely positioned for the treatment of HIV, as immune cells are the primary target of HIV. HSCT of genetically modified cells would ultimately result in an HIV-resistant hematopoietic system, achieving a functional (and possibly sterile) cure.

1.7.1. The human immunodeficiency virus

The first case of HIV was reported in 1981 (133,134) and HIV/acquired immunodeficiency syndrome (AIDS) has become a serious health concern ever since, with an estimated 38 million infected individuals and 770 000 HIV-related deaths worldwide. South Africa has a high prevalence of HIV (20.4%), with more than seven million HIV-infected individuals and an estimated 240 000 AIDS-related deaths (135). Since the advent of cART, the number of new HIV infections and morbidity and mortality of AIDS-related diseases have decreased

significantly (136). HIV is a chronic infection for which there is no cure, and lifelong therapy is required to maintain viral suppression. The ultimate goal of cART is to sufficiently suppress viral load, allowing for the restoration of depleted immune cells. However, cART alone cannot eliminate the virus, and maintaining therapy for prolonged periods presents major challenges. Potential immune evasion, reservoir establishment and drug resistance remain problematic and contribute to HIV persistence.

1.7.1.1. HIV groups and subtypes

HIV belongs to the family *Retroviridae* and the genus *Lentivirus*. This genus consists of HIV-1, HIV-2 and various simian immunodeficiency viruses (SIV) that infect primate species in Africa. HIV-1 and -2 have similar properties with regard to gene arrangement, transmission, intracellular replication pathways and clinical consequence. However, HIV-2 is less common, has lower transmission, and is less likely to progress to AIDS. HIV-2 is limited to West Africa and accounts for a small proportion of HIV infections (137,138). Many differences exist between HIV-1 and -2 which are not yet completely understood. HIV-1 is responsible for the majority of infections worldwide and is divided into four groups, M (main), N, O, and P, which have different geographic distributions, but produce similar clinical symptoms (139). HIV-1 group M is the main cause of the worldwide pandemic, whereas groups N and O are restricted to West Africa. Group P, the result of a zoonotic event, was very recently identified in Cameroon (140). Based on divergent sequences, the M group can be classified into different subtypes (A – J), circulating recombinant forms (CFR) and unique recombinant forms (URF).

HIV-1 subtype B (HIV-1B) and C (HIV-1C) are the most prevalent subtypes, accounting for 11% and 48% of infections worldwide, respectively (141). Most HIV-1 infections in India, Brazil and sub-Saharan Africa are due to HIV-1C, while HIV-1B is confined to developed countries such as North America, Europe and Australia (142–144). However, HIV-1B is the most studied subtype and most of our understanding of HIV-1 is based on this subtype. However, subtypes differ with respect to phenotypic properties and co-receptor usage (145–149), replication rates (150,151), rate of disease progression (152–154), biology of transmission (155,156) and mutational patterns (157–159).

The widespread success of HIV-1C viruses relative to other HIV-1 subtypes suggests that there might be HIV-1C-specific factors affecting the transmission and/or replication of this subtype (160). The genomic construction is similar between HIV-1 subtypes, with sequence diversity ranging between 5 and 35% (161). Distinct characteristics with regard to viral entry and pathogenesis have been reported for HIV-1C, such as the use of either CCR5/CXCR4 or both co-receptors, during early and late infection (162). HIV-1C viruses also have high transmission fitness, which could increase the frequency of sexual and mother-to-child transmission (150,163). Counter-intuitively, HIV-1C has been reported to be less cytopathogenic compared to other subtypes, which may result in a superior ability to latently persist for long periods (164). At a genomic level, the long terminal repeats (LTR) have three potential nuclear factor (NF)-κB binding sites and a truncated Rev protein (165), which might enhance gene expression and thereby also enhance viral replication. An amino acid insertion in the Vpu polypeptide could likewise affect HIV-1C virulence, through mechanisms such as increased CD4 degradation or enhanced virion release from host cells (166). Despite these molecular characteristics that could lead to enhanced viral replication, HIV-1C viruses have shown to have lower replication fitness in primary $CD4^+$ T cells and peripheral blood mononuclear cells (PBMCs) when compared to other subtypes (167,168).

1.7.1.2. HIV tropisms

The primary HIV receptor, CD4, and two well described co-receptors, CCR5, and CXCR4, play a critical role in receptor-mediated entry of HIV into host cells. Although the general perception for many years has been that HIV-1 requires CD4 in combination with one of its co-receptors for receptor-mediated infection, there is increasing evidence that infection can occur independently of CD4 (169–172). CCR5 and CXCR4 function as co-receptors of R5- and X4-tropic strains of HIV, respectively. HIV-1 tropism (co-receptor usage) is determined by the variability of the V3 region of the viral gp120 glycoprotein. The R5-tropic virus dominates in initial to early infection (173), while X4-tropic virus emerges later during infection, leading to more advanced disease and AIDS (162). X4-tropic strains are present in approximately 50% of individuals chronically infected with HIV-1B (174).

Although different theories have been proposed to explain tropism switching, the exact mechanism involved is still poorly understood. X4-tropic variants might arise over time from R5-tropic strains as a result of evolution, which might emerge as minor X4-tropic strains during early infection. Second, X4-tropic strains may be transmitted with R5-tropic strains; however, effective immune responses during primary infection might prevent their infection, whereas late emergence is associated with a compromised immune system. Third, R5- and X4-tropic strains have different target cells due to the varying expression profiles of CCR5 and CXCR4 on somatic cells. Change in target cell populations during the course of infection may therefore also be responsible for co-receptor switching (175).

1.7.1.3. HIV life cycle

HIV is transmitted as an enveloped single-stranded (ss) RNA virus that has the ability to infect both dividing and non-dividing cells. Receptor-mediated HIV entry into host cells is facilitated by the binding of gp120 to CD4 and a co-receptor, allowing fusion of the virion with the host cell membrane and insertion of the viral RNA into the cell. Upon viral entry, viral RNA is reverse transcribed into double-stranded (ds) proviral DNA, imported into the nucleus and integrated into the host genome, where it can either actively replicate or remain inactive, known as latent infection. In actively replicating cells, the viral genome is transcribed and translated through host molecular mechanisms. Viral proteins are produced and assembled as viral particles (Figure 1.4) that will bud from the cell. Upon the release of a critical mass of viral particles, the cell breaks leading to cell death. This entire process is referred to as the lytic cycle of HIV infection.

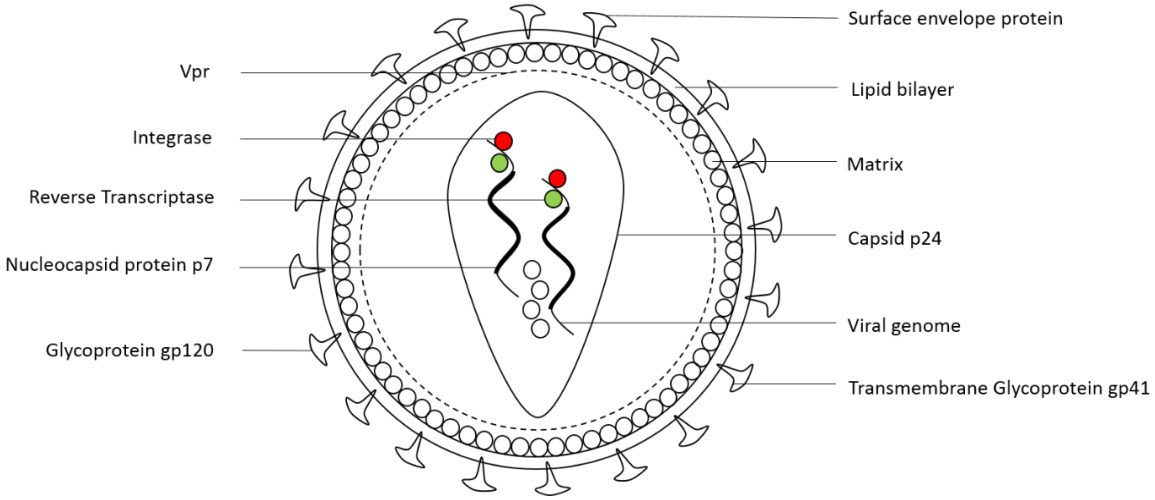


Figure 1.4. HIV virion. An illustration of an HIV virion and the presence of viral proteins and enzymes important for entry, reverse transcription and integration into the host cell genome. (Figure was created by Juanita Mellet).

1.7.1.4. Acute and latent infection

Acute HIV infection, resulting in continued replication of HIV and production of virions cause a progressive decrease in the number of $CD4^+$ T cells, ultimately leading to AIDS. AIDS is defined as the progressive decline of the immune system, which may result in opportunistic infections and cancer. $CD4^+$ T cells are the preferential targets of HIV and have the ability to interchange between an active and resting (inactive) state. Infection generally occurs when these cells are active, and latency is established once these cells revert to an inactive state allowing the virus to escape immune recognition, contributing to viral persistence (176,177). Latently infected cells are rare and have therefore not been accurately quantified to date (178). Memory $CD4^+$ T cells are a well-established HIV reservoir. Macrophages are another known target of HIV infection; however, their contribution to the reservoir pool is still largely unknown due to their extreme diversity and tissue specificity (179).

Latent reservoirs pose a major obstacle for the eradication of HIV, since the integrated provirus can be activated upon stimulation to produce viral particles. The viremia rebound from undetectable viral load observed in two separate cases, the Mississippi child (180) and the Boston patients (181), treated with allogeneic CCR5-null HSCT, provides a cautionary tale for the importance of the reservoir. A recent report describes two participants who started cART

10 – 12 days post-infection and continued treatment for 32 weeks and in whom HIV-1 was undetectable in blood and several organs; HIV relapse occurred more than seven months after treatment interruption (182).

1.7.2. HIV and HSPCs

Although $CD4^+$ T cells are the primary targets of HIV infection, hematopoiesis is also negatively impacted by HIV infection, leading to a wide range of HIV-associated haematological abnormalities. The most plausible mechanisms for this include: direct infection of HSPCs with HIV, the inability of stromal cells in the bone marrow to function normally, toxic effects of HIV-1 proteins and changes in the cytokine milieu.

1.7.2.1. Direct infection of HSPCs

HSPCs have been shown to express low levels of CD4 and CCR5, while CXCR4 is more abundantly expressed (183,184). Although HSPCs were initially thought to be resistant to HIV infection (185–188), expression of these receptors theoretically makes these cells susceptible. Since HSPCs tend to express increased CXCR4, they tend to be more susceptible to X4-tropic virus (183,189,190). This would explain the rapid disease progression observed upon viral transition from R5- to X4-tropic virus (191). However, due to conflicting reports in literature, it is still not clear whether HSPCs are directly infected with HIV. Several studies proposed that HIV infects HSPCs in the bone marrow, contributing to a latent reservoir pool (183,189,192), while others have opposed this notion, since they were unable to detect HIV in HSPCs (186,188).

The negative effects of HIV on HSPCs results in decreased proliferation and differentiation into mature blood cell lineages *in vitro* and *in vivo* (193–196). Whether this is due to direct or indirect effects of HIV remain to be established. The studies that attempted to determine whether HSPCs are susceptible to HIV infection were laboratory-based due to ethical and logistical challenges associated with obtaining bone marrow from HIV-infected individuals. A single study by Redd *et al*. (196) reported that HIV-1C, but not HIV-1B, has the ability to invade the bone marrow niche and infect HSPCs.

Although more primitive HSPCs tend to express higher levels of CXCR4, several studies indicate that progenitors (such as CMP and CLP) are more susceptible to infection and are directly associated with various HIV-associated cytopenias. This has consequences for autologous HSCT-based HIV gene therapy, in that the $CD34^+$ fraction used for genetic modification and HSCTs possibly harbour HIV. Determining whether primitive HSPCs are susceptible to infection is crucial for autologous HSCT of HIV-infected patients, since infected HSPCs could lead to viral rebound and expansion of the reservoir.

1.7.2.2. Impaired stromal cell and signaling network in the bone marrow niche

In vitro studies suggest that bone marrow MSCs and endothelial cells are directly infected with HIV (197), resulting in altered cytokine signaling and cell death. Endothelial cells contribute to the stromal impairment observed in individuals infected with HIV by exhibiting a decreased capacity to respond to regulatory signals to enhance blood cell production (197). Several studies reported infection of macrophages *in vivo* since they express receptors and co-receptors and can become infected with both T cell-tropic and macrophage-tropic HIV strains (198). Alterations in the bone marrow stromal composition and cell signaling result in a sub-optimal microenvironment for HSPCs and consequently defective haematopoiesis (199). It has been shown that infected stroma is unable to support uninfected HSPC expansion and differentiation (200,201).

Toxic effects of HIV-associated proteins have been reported to dysregulate the bone marrow niche and affect hematopoiesis. HIV-1 gp120 impairs the clonogenic potential of HSPCs and induces apoptosis as a result of endogenous upregulation of TGF-β through a Fas-dependent mechanism (202). The HIV-1 Gag protein has been shown to suppress colony-forming capacity of HSPCs (203), while Vpr has been shown to induce phagocytosis of bone marrow cells through mononuclear phagocytes (204). Additionally, Tat-exposure stimulates macrophages to produce TGF-β leading to myelosuppression *in vitro*. Blocking TGF-β in purified $CD34^+$ HSPCs from peripheral blood previously exposed to HIV-1 either *in vivo* or *in vitro*, improved their growth and survival (202). In agreement with these results, gp120 leads to TGF-β upregulation in $CD34^+$ HSPCs and simultaneous downregulation of a proliferation-inducing ligand (APRIL), a

cytokine known to induce proliferation (205). Results from studies investigating production of TGF-β1 are conflicting, since other studies failed to identify differences in HSPCs in the presence and absence of HIV-1.

Cytokines and hematopoietic factors, such as tumour necrosis factor-α (TNFα), TGF-β, interferon-γ (IFN-γ), IL-1, IL-6, IL-10, IL-18 and many others are frequently affected in HIV-infected individuals. Clinical studies have confirmed that differences in cell signaling profiles exist between HIV-infected and non-infected individuals, with higher levels of IL-1, IL-18, TNFα and IL-6 in plasma of HIV-positive patients. A recent study showed higher expression of TNF-α and IFN-γ in HIV-infected individuals when compared to non-infected individuals (206), in accordance with HIV-associated chronic inflammation. Furthermore, pro-inflammatory cytokines TNFα, IL-1, and IL-6 and the chemokines macrophage inflammatory protein-1 alpha and beta (MIP-1α and MIP-1β) and regulated on activation, normal T cell expressed and secreted (RANTES) were also found to be upregulated in the bone marrow of HIV-infected individuals. HIV infection also causes a decrease in endogenous G-CSF (207,208). Inhibition of G-CSF production by Vpr has also been described, which was restored when Vpr was deleted from the virus (209). Decreased endogenous levels of G-CSF results in impaired proliferation and differentiation of GMPs.

The derangement of growth factor and cytokine profiles in the bone marrow microenvironment related to HIV infection is clearly unfavourable for normal hematopoiesis, further supporting the need for a cure for HIV.

1.8. STUDY AIMS AND OBJECTIVES

Stem cells have sparked a great deal of interest in recent years due to their potential application in regenerative medicine. HSCT has provided the foundation for these technologies through problems encountered and knowledge gained throughout the years. New developments in HSCT are currently underway to expand current applications and improve safety and efficacy. The power of HSCT lies in the ability of HSPCs to self-renew and differentiate into various progenitors and mature blood cell lineages.

Aim 1. Determine the optimal isolation method and good manufacturing practice (GMP)-compliant culturing of HSPCs

The *in vitro* culturing and expansion of HSPCs is an essential part of research and initial stages of HSCT product development. However, there is currently no standard for culturing HSPCs *in vitro*. The overall aim of this project was to make the research performed in our laboratory not only compliant with good cell culture practice (GCCP), but also make it as clinically relevant as possible by conforming to good manufacturing practice (GMP) guidelines for the processing and culturing of HSPCs. A significant part of this study therefore aimed to standardise the culturing of HSPCs in our laboratory (Chapter 2 and 3). The optimal isolation method to obtain increased HSPC purity and viability was determined by testing different methods of isolation, including magnetic activated cell sorting (MACS) and FACS (Chapter 2). Initial culturing of HSPCs in our laboratory included the use Dulbecco's modified eagle's medium (DMEM) supplemented with foetal bovine serum (FBS). The ethical concerns and high degree of unknown variables involved with the use of FBS has led to the use of serum-free and animal-component free culture conditions. *Ex vivo* GMP-compliant culturing of HSPCs was also explored by testing different commercial serum-free media (Chapter 2).

Aim 2. Determine whether G-CSF leads to improved *ex vivo* expansion of HSPCs

Different laboratories make use of different cytokine combinations for the *ex vivo* expansion of HSPCs. Two commonly used cytokine combinations include: FLT3L, SCF, TPO, IL-3 and FLT3L, SCF, TPO, IL-6. G-CSF is a cytokine that induces granulocyte differentiation and proliferation. G-CSF is used clinically in mobilising agents to enrich for $CD34^+$ HSPCs in the peripheral circulation and to improve and/or prevent chemotherapy-induced neutropenia. Part of this

study aimed to determine whether the addition of G-CSF to the above-mentioned cytokine combinations would lead to improved *ex vivo* expansion of HSPCs (Chapter 3).

Aim 3. Determine the effect of SR1 on the transcriptome of seven-day expanded HSPCs

The AhR has recently been found to play a role in cell differentiation, pluripotency, and stemness, and could be involved in the balance between differentiation and pluripotency. Treatment of HSPCs with SR1 has been shown to increase the number of $CD34^+$ cells able to engraft in the bone marrow of recipient mice. However, the exact role of AhR in HSPC biology is still unknown, and the effects of SR1-induced expansion on HSPCs still needs to be fully determined. The aim was therefore to determine the effect of SR1 on the transcriptome of seven-day expanded $CD34^+$ and $CD34^-$ cells from UCB, and to establish to what extent the transcriptome of SR1-expanded $CD34^+$ HSPCs differ from non-expanded $CD34^+$ HSPC (Chapter 4). Chapter 4 will describe the optimisation techniques that were performed to determine the optimal concentration of SR1 required for expansion and subsequent gene expression analysis. The effect of SR1 on expansion and gene expression of HSPCs will also be discussed in this chapter.

Aim 4. Characterise the heterogeneity of $CD34^+$ HSPCs from UCB using single-cell transcriptome analysis

UCB-derived $CD34^+$ cells have been used successfully for UCBT across the globe to treat various diseases. A great deal of effort is being made to understand heterogeneity within $CD34^+$ HSPC populations and its contribution to the successful clinical use of these cells. Gene expression profiles provide a vast amount of information regarding the $CD34^+$ HSPC population phenotype, which underlies its molecular function. Until recently, gene expression studies have been limited to bulk RNA. Single-cell RNA-seq enables researchers to uncover previously unknown cellular heterogeneity in cell populations. This provides meaningful information into the behaviour of cells, enabling a deeper understanding of complex biological systems and processes that individual cells are involved in. Single-cell gene expression profiling has already contributed to improved understanding of cellular heterogeneity in a variety of cell types. The aim was to characterise the heterogeneity of $CD34^+$ HSPCs from UCB using single-cell transcriptome analysis (Chapter 5).

Aim 5. Determine whether HSPCs are susceptible to HIV-1 infection

HIV infection has a negative effect on hematopoiesis. Whether this is due to direct infection of HSPCs or impaired stromal cell signaling in the bone marrow is not entirely known. This could have implications for autologous HSCT-based HIV gene therapy. Therefore, to better understand the interactions between HSPCs and HIV, we aimed to determine whether HSPCs are susceptible to HIV-1 infection and whether a subset of HSPCs exists that is resistant to infection to use for HIV gene therapies (Chapter 6). Before this could be achieved, protocols for the culture of primary HIV-1C strains had to be established and optimised in order to produce high-quality stocks of virus. Methods for the successful detection of infected cells also had to be optimised.

The thesis is written in article format, where each chapter has its own introduction, relevant literature, materials and methods, results, discussion and conclusion. This thesis concludes with Chapter 7, summarising the final conclusions drawn from results of all the chapters and the implications this might have for future clinical applications of HSPCs.

CHAPTER 2. THE CULTURE OF HEMATOPOIETIC STEM AND PROGENITOR CELLS *IN VITRO* ACCORDING TO GMP STANDARDS

2.1. INTRODUCTION

Hematopoietic stem and progenitor cells (HSPCs) have self-renewal capabilities and are able to differentiate into all the blood lineages, including the myeloid (erythrocytes, monocytes, megakaryocytes, and neutrophils) and lymphoid (T cells, B cells, and NK cells) lineages (48,210). UCB is a rich source of HSPCs and is present at a frequency of less than 2% of total nucleated cells (TNC) (211). Numerous transplantations have been performed using UCB-derived HSPCs; however, the use of a single unit has been limited to pediatric use due to the relatively small number of HSPCs present in UCB units. The use of multiple UCB units has been successful, but require that the units match both the donor and the other units used for the transplant (73).

For several years, research has focused on expanding the $CD34^+$ cell population, with the objective being to maintain the 'primitive population' of stem cells. The *ex vivo* culturing of cells often requires the supplementation of basal growth medium with serum to supply nutrients for maintenance and proliferation. The most commonly used animal-derived serum is FBS. Previous HSPC culture conditions in our laboratory included the use of DMEM supplemented with FBS. Several concerns have been raised with regard to the use of FBS (212): (i) FBS contains many growth factors, adhesion molecules and other components which are not well defined; (ii) the addition of FBS to culture medium could potentially elicit an immune response when used for transplantation purposes, since it contains xenogeneic components that are not normally present in the human body (213); (iii) the use of FBS may expose cells to endotoxins, mycoplasma and viral contaminants (214); (4) the use of FBS for expansion and differentiation of primary cultures could result in non-reproducible results, largely due to the differences between batches used in different laboratories (215,216); and (iv) there are several ethical concerns, such as the welfare of the bovine fetuses during harvesting and collection of FBS (217,218).

The ethical concerns and high degree of unknown variables involved with the use of FBS has lead researchers to turn to serum-free and animal component free culture conditions. The culturing of cells under serum-free conditions is widely recognised as an alternative to using FBS (212,219). Several chemically defined serum-free media specific for the *ex vivo* culturing of HSPCs are commercially available. Serum-free culture conditions for HSPCs would still require the addition of growth factors. GMP-compliant growth factors are available commercially, and are less likely to elicit an immune response.

Research on HSPCs has been conducted for many years without uniform standards for culturing these cells. Cell culture procedures are becoming more controlled and the proposed GCCP guidelines for the *ex vivo* culturing of cells ensure reliable and reproducible results that will allow for international harmonisation between research groups (220,221). In the past few years, the use of terms such as quality assurance, quality control, good laboratory practice (GLP) and GMP have increased in the context of blood cells and blood products. The clinical use of stem cells requires that the isolation, preparation and manipulation of these cells conforms to GMP guidelines. GMP requires that products are consistently manufactured and controlled according to well-defined quality standards (222). Since *in vitro* work is an essential part of research and initial stages of product development, the aim is to make the research performed in our laboratory not only compliant with GCCP, but also make it as clinically relevant as possible by conforming to GMP guidelines for the processing and culturing of HSPCs. A GCCP and GMP compliant culture method therefore needs to be established for the *ex vivo* culturing and expansion of HSPCs. The purpose of this study was to determine the optimal serum-free medium for the culturing of HSPCs *in vitro*.

This chapter describes the results obtained when MACS was compared to FACS for the isolation of UCB-derived HSPCs. This chapter further describes the results obtained when $CD34^+$ HSPCs are expanded using different media, using the following end-points: viability, proliferation (TNCs and $CD34^+$ HSPC numbers and percentages) and HSPC immunophenotype.

2.2. MATERIALS AND METHODS

2.2.1. Ethics statement

This study was conducted at the Institute for Cellular and Molecular Medicine (ICMM), Department of Immunology, University of Pretoria. Approval was granted in 2014 by the Faculty of Health Sciences Research Ethics Committee (*Protocol Number 410/2014*) (Appendix G) and Netcare Research Operations Committee (*Approval Number UNIV-2014-0058*) (Appendix H).

2.2.2. Sample collections

Informed consent was obtained from healthy HIV-negative mothers scheduled for caesarian section (C-section) at 38 to 40 weeks gestation at a private hospital in Pretoria. An unsigned example of the informed consent document that was completed by mothers prior to UCB collection is included as Appendix E. The HIV status of patients from whom UCB was collected was obtained from patient files. UCB was collected post-delivery from the placenta via the umbilical vein into bags containing citrate phosphate dextrose (CPD) anticoagulant (Tianhe Pharmaceutical, China), using a 14-gauge needle. UCB units were codified (anonymised) to protect patient information and ensure patient anonymity. Samples were stored at 4°C and used within 24 hours.

2.2.3. Enrichment of $CD34^+$ HSPCs

The collection bag content was poured into sterile 50 mL Falcon tubes. UCB units were plasma depleted by centrifugation at 300 relative centrifugal force (rcf, x g) for 20 minutes (min). The plasma layer was aspirated and discarded. The buffy coat layer (white layer on top of the red blood cells) was aspirated and placed into a new 50 mL Falcon tube (pooled for different tubes from the same donor). Red blood cells were lysed by adding 20 mL of ammonium chloride (NH_4Cl; 0.155 M NH_4Cl (Sigma-Aldrich, USA), 0.0119 M sodium bicarbonate ($NaHCO_3$; Sigma-Aldrich, USA), 0.25 mM ethylenediaminetetraacetic acid (EDTA), pH 7.4) for 20 min at 4°C. The sample was centrifuged at 300 x g for 10 min at 4°C. Ammonium chloride (NH_4Cl) was aspirated and each sample was washed twice by filling the tube to 30 mL with TP buffer (phosphate buffered saline (PBS, pH7.4), 10 µg/mL human albumin (Sigma-Aldrich, USA) and 2 mM EDTA

(Sigma-Aldrich, USA)) and centrifugation at 300 x g for 10 min at 4°C. Supernatant was aspirated and cells were resuspended in 5 – 10 mL TP buffer. At this point a 50 µL aliquot was taken for counting.

2.2.3.1. Flow cytometer: Counting protocol

Stem-Kit™ reagents from Beckman Coulter (Miami, Florida, USA) were used to determine viability and the absolute number of $CD34^+$ cells in each sample using the Gallios flow cytometer (Beckman Coulter, Miami, Florida, USA). Stem-Kit reagents include a vial containing CD45 FITC and CD34 PE monoclonal antibodies, a vial containing CD45 FITC and IsoClonic™ control PE, and a vial containing the viability dye, 7-Amino-Actinomycin-D (7AAD). This specific CD34 antibody recognises a Class III epitope of the CD34 molecule.

The IsoClonic control is for possible non-specific antibody binding and is used to set the negative/positive boundaries of CD34 protein expression. The nucleic acid intercalating viability dye binds between cytosine and guanine bases of the DNA in cells with compromised cell membranes. Non-viable cells were identified and excluded from flow cytometry post-acquisition analysis based on 7AAD-positive staining. Antibodies were added to samples and incubated for 15 min prior to counting. Flow-Count™ fluorospheres (Beckman Coulter, Miami, Florida, USA) are fluorescent microbeads used to determine cell counts using a flow cytometer. Flow-Count fluorospheres is a calibrated suspension of microbeads with a known concentration that have a fluorescence emission range of 525 nm to 700 nm, when excited at 488 nm. On the Gallios flow cytometer, these Flow-Count fluorospheres can be detected in channels FL1, FL2, FL3 or FL4. A 1:1 ratio of sample to Flow-Count fluorospheres is required for accurate counts. Flow-Count fluorospheres were added to each sample immediately prior to analysis on the flow cytometer.

Viable cells were measured, using 7AAD viability dye [Excitation: 488 nm; Emission: 635/75 nm], in the FL4 channel (Figure 2.1a), which was gated on region [B]. Region [B] is a NOT Flow-Count fluorospheres Boolean gate applied post-acquisition and used to exclude the Flow-Count fluorospheres from downstream flow cytometric plots during data analysis. Flow-Count fluorospheres were detected in the FL3 channel (Figure 2.1d) for this particular protocol.

$CD34^{+}$ and $CD45^{dim}$ cells were measured in the FL2 channel [CD34 PE; Excitation: 488 nm; Emission: 568/90 nm] (Figure 2.1b) and in the FL1 channel [CD45 FITC; Excitation: 488 nm; Emission: 504/41 nm] (Figure 2.1c), respectively. Both the side scatter (SS) Lin vs. CD34 PE (Figure 2.1b) and SS Lin vs. CD45 FITC (Figure 2.1c) were gated using the "Viable" region. The Gallios flow cytometer filter configurations are shown in Table 2.1. All flow cytometer analyses were performed using Beckman Coulter Kaluza Analysis Software (version 2.1) (Beckman Coulter, California, USA).

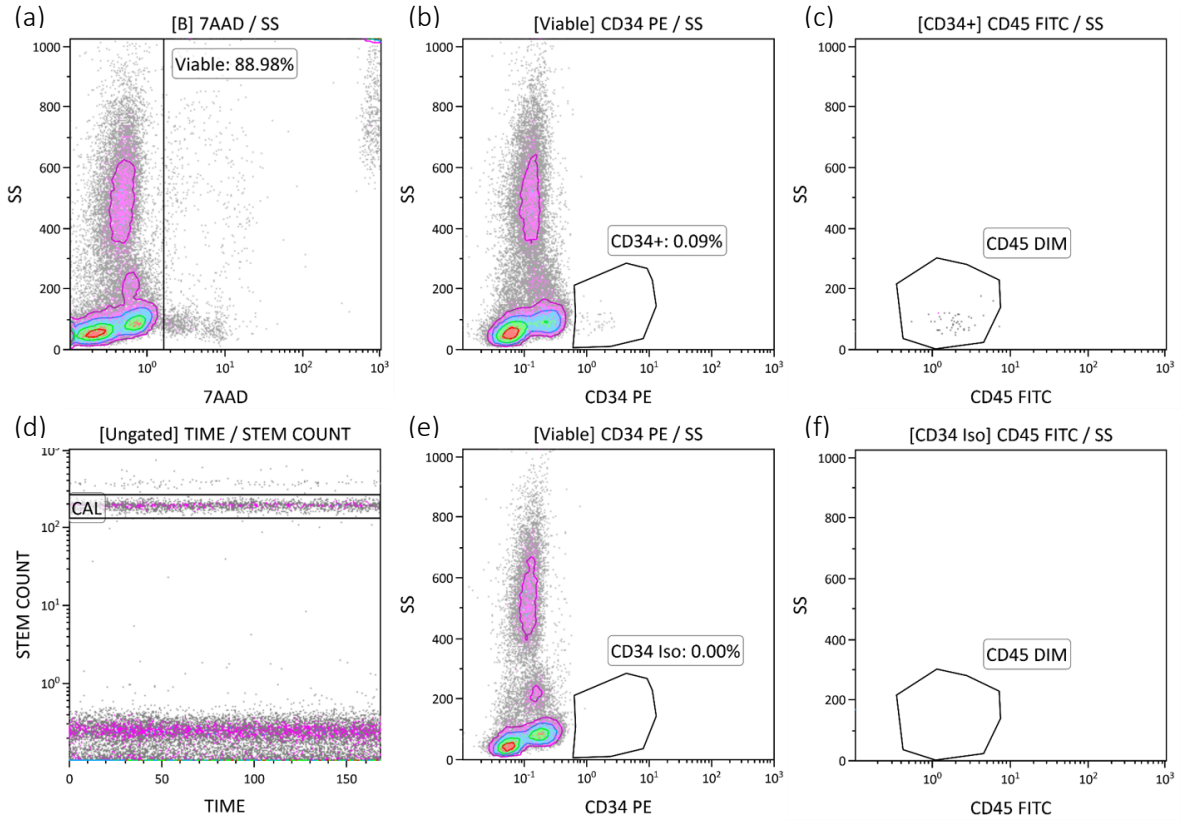


Figure 2.1. Day 0 counting protocol. Flow cytometry protocol and representative analysis of UCB-derived HSPCs (buffy coat) on Day 0. (a) Density plot showing viable, 7AAD-negative cells. (b) and (c) Density plots showing cells stained with Stem-Kit CD34 PE, CD45 FITC monoclonal antibodies. (e) and (f) Density plots showing cells stained with Stem-Kit IsoClonic control PE, CD45 FITC. (d) Density plot showing the Flow-Count fluorospheres in "CAL" region.

The number of Flow-Count fluorosphere events together with the number of cell events in the region of interest ("CD34+" region) were used to determine cell concentration. In general, > 1000 Flow-Count fluorosphere events were acquired to ensure accurate cell counts. The following formula was used to calculate the number of cells/µL:

$$Absolute\ cell\ count\ \left(\frac{cells}{\mu L}\right) = \frac{Number\ of\ events\ in\ region\ of\ interest,i,e.CD34+\ region}{Number\ of\ \text{Flow-Count fluorospheres}\ in\ \text{CAL}\ region} \times Flow-Count\ fluorospheres\ concentration$$

Flow-Count fluorospheres concentration refers to the calibration factor indicated on the product insert that is supplied with each Flow-Count fluorospheres vial.

Table 2.1. Gallios flow cytometer filter configuration.

Laser	Filter	FL	Fluorochrome/Dye	Clone
488nm, 22mW	525/40	1	CD45 FITC	J33
	575/30	2	CD34 PE	581
	620/30	3	Flow-Count fluorospheres	
	695/30	4	7AAD	

2.2.3.2. Magnetic activated cell sorting

The human CD34 MicroBead kit (Myltenyi Biotec, Germany) is a positive selection kit that we used together with LS columns (Myltenyi Biotec, Germany) and the QuadroMACS separator (Myltenyi Biotec, Germany), to magnetically enrich for $CD34^+$ cells.

2.2.3.2.1. Sample preparation

Once cell number had been determined, using the counting protocol above, the sample was centrifuged at 300 x g for 10 min at 4°C. The cell pellet was resuspended in TP buffer (300 µL for 10^8 total cells). FC receptor (FcR) Blocking Reagent (supplied with the kit) was added to the sample (100 µL for 10^8 total cells) and incubated for 10 min at 4°C. Thereafter, the CD34 Microbeads were added to the sample (100 µL for 10^8 total cells), mixed and incubated for

30 min at 4°C. After incubation, cells were washed by centrifugation with 5 mL TP buffer (5 – 10 mL for 10^8 total cells) at 300 x g for 10 min at 4°C. The cell pellet was resuspended in 500 µL TP buffer.

2.2.3.2.2. Magnetic separation

An LS column was placed in the magnetic field of the QuadroMACS separator with a 30 µM pre-separation filter (Myltenyi Biotec, Germany) at the top. The column was prepared by rinsing with 3 mL PBS (pH 7.4). Thereafter, the cell suspension was added onto the column. The column was washed three times with 3 mL PBS after which the column was removed from the separator and placed in a 15 mL collection tube. 5 mL TP buffer was added onto the column to flush out the labelled cells by firmly pushing the plunger into the column. A 100 µL aliquot was taken to determine the purity and number of $CD34^+$ cells present after magnetic isolation.

2.2.3.3. Fluorescent activated cell sorting

Fluorescent activated cell sorting is a specialised type of flow cytometry, where a cell population of interest is isolated/purified from a heterogeneous population of cells. Stem-Kit reagents (CD45 FITC, CD34 PE and 7AAD) were added to the sample based on the absolute number of cells present (approximately 5 µL per 10 x 10^6 viable cells). Once the antibodies had been added, the sample was incubated for 30 min at room temperature. Stained cells were washed once with TP buffer and the supernatant was aspirated. The pellet was resuspended in the desired volume based on the absolute cell number (final concentration of approximately 7 x 10^6 cells/mL). Cells were sorted by FACS based on CD34 expression using a BD FACSAria™ Fusion (BD Biosciences, USA). The 'Purity' sorting mode was used for sorting together with the 70 µM nozzle. The gating strategy was performed as described earlier in the counting protocol. A purity check was performed prior to actual sorting and > 90% purity was obtained for all samples (Table 2.3). A graphical illustration of the gating strategy used for sorting to verify the purity post-sort (purity check) is shown in Figure 2.2.

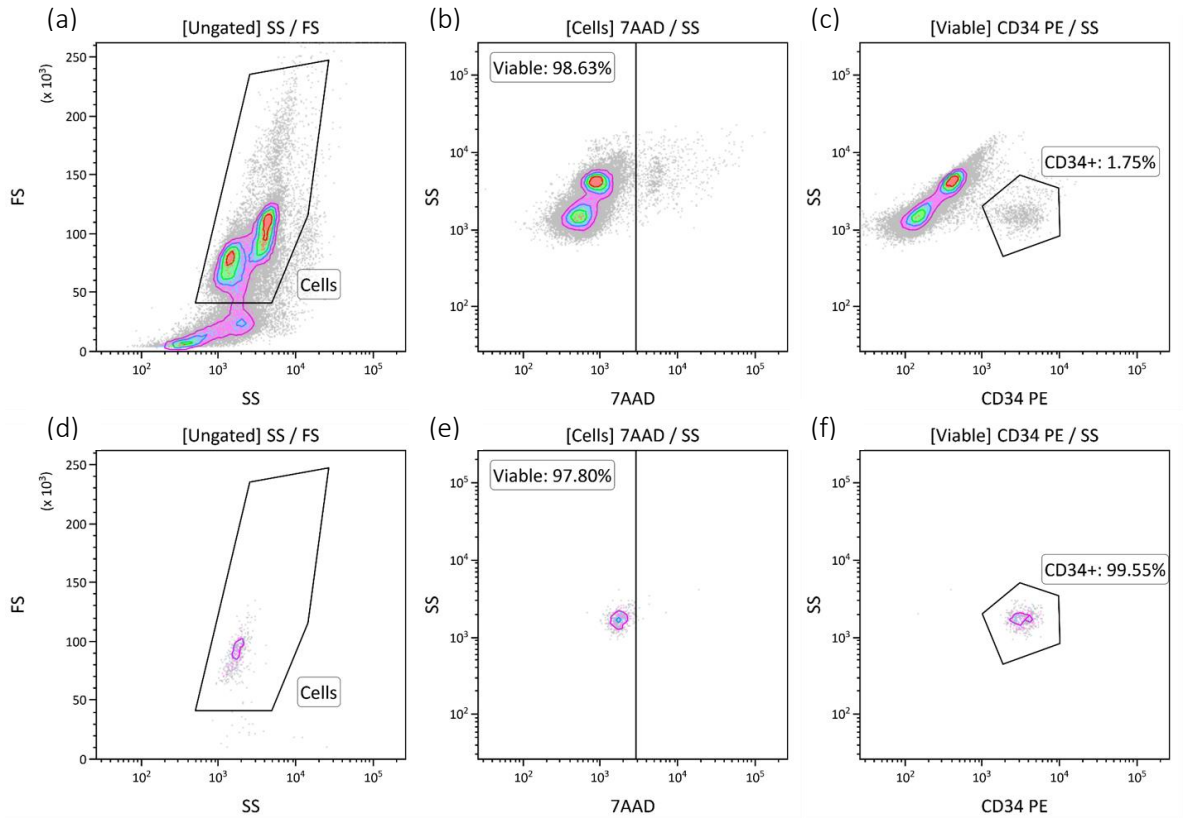


Figure 2.2. Sorting protocol for UCB-derived HSPCs. (a) Forward scatter (FS) vs. SS plot used to identify the cell population before sorting. (b) Dot plot showing cell viability (7AAD-negative) before sorting. (c) Dot plot showing cells stained with CD34 PE conjugated antibody before sorting. (d) A FS vs. SS plot used to identify the sorted cell population (post-sort). (e) Dot plot showing cell viability (7AAD-negative) of sorted cells. (f) Dot plot showing cells stained and sorted with CD34 PE conjugated antibody.

2.2.4. Culture conditions

Viable FACS-purified $CD34^+$ cells were sorted into wells of a 48-well plate (Thermo Fisher Scientific, USA) containing four different medium conditions, in triplicate. $CD34^+$ cells were cultured in 1 mL medium at 4 – 5 x 10^3 cells/well. Medium conditions included: DMEM (Gibco, Thermo Fisher Scientific, USA) supplemented with FBS (Gibco) (as previously used in our laboratory), Stemline II (Sigma-Aldrich, USA), StemPro-34 Plus (Thermo Fisher Scientific, USA) and StemSpan animal component free (ACF) (Stem Cell Technologies, Canada). The medium was supplemented with 2% Penicillin-Streptomycin (pen/strep, 10 000 units/mL Penicillin and 10 000 µg/mL Streptomycin, Gibco™, Life Technologies, Thermo Fisher Scientific, USA). The following recombinant human growth factors were added: G-CSF (100 ng/mL), SCF (100 ng/mL), TPO (25 ng/mL), and IL-3 (25 ng/mL) (Gibco by Life Technologies, Thermo Fisher Scientific, USA). The concentrations indicate the final concentration of the respective growth

factors present in a well. Sterile PBS was added to unused wells to maintain humidity. The cultures were maintained at 37°C and 5% carbon dioxide (CO_2) in a humidified atmosphere for seven days.

2.2.5. Viability, proliferation and $CD34^+$ cell numbers/percentages

HSPC expansion was examined after seven days by measuring viability, proliferation and $CD34^+$ cell numbers/percentages. A 50 µL aliquot was taken from each of the triplicate wells and transferred into flow tubes. As on Day 0, 3 µL of Stem-Kit reagents (CD45 FITC / CD34 PE and 7AAD) were added to the tubes. A separate flow tube containing CD45 FITC / IsoClonic control PE and 7AAD was also prepared. Again, as mentioned earlier in this chapter, Flow-Count fluorospheres were added immediately before the samples were analysed on the Gallios flow cytometer.

2.2.6. HSPC-associated immunophenotype

The analysis of HSPC immunophenotypic markers was performed after seven days using a Gallios flow cytometer. A 100 µL aliquot was taken from each of the triplicate wells and put into flow tubes with 5 µL of each of the following mouse anti-human monoclonal antibodies: Lin FITC, CD34 PE-Cy7, CD38 APC-Cy7, CD45 BV421 (Biolegend, USA) (Table 2.2). Two gating strategies were used for the HSPC immunophenotypic analysis. Figure 2.3 and 2.4 show the first and second gating strategies employed, respectively. For the first gating strategy, "Cells" were analysed using a FS vs. SS plot (Figure 2.3a). Lineage-negative (Lin^-) cells were identified using the lineage cocktail FITC [Excitation: 488 nm; Emission: 520 nm] (Figure 2.3b), which was gated on the "Cells" region. The $CD38^-$ cells were identified using CD38 APC-Cy7 [Excitation: 633 nm; Emission: 776 nm] (Figure 2.3c) and gated on the "Lin^-" region. The majority of cells were $CD45^+$ on day seven, which were identified (gated on $CD38^-$ region) using CD45 BV421 [Excitation: 405 nm; Emission: 421 nm] (Figure 2.3d). $CD34^+$ and $CD34^-$ cells were identified (gated on $CD45^+$ region) using CD34 PE-Cy7 [Excitation: 488/532/561 nm; Emission: 575/780 nm] (Figure 2.3e). For the second gating strategy, $CD34^+$ and $CD34^-$ cells were gated on "Cells" (Figure 2.4b). $CD34^+CD38^+$, $CD34^+CD38^-$ (Figure 2.4c), $CD34^-CD38^+$ and $CD34^-CD38^+$ (Figure 2.4d) cells were analysed by gating CD38 on $CD34^+$ and $CD34^-$ regions, respectively. Negative

and positive regions for each phenotypic marker were created based on positive and negative populations present in a sample (if applicable) or using a whole blood control sample when the distinction between the negative and positive populations was not clear.

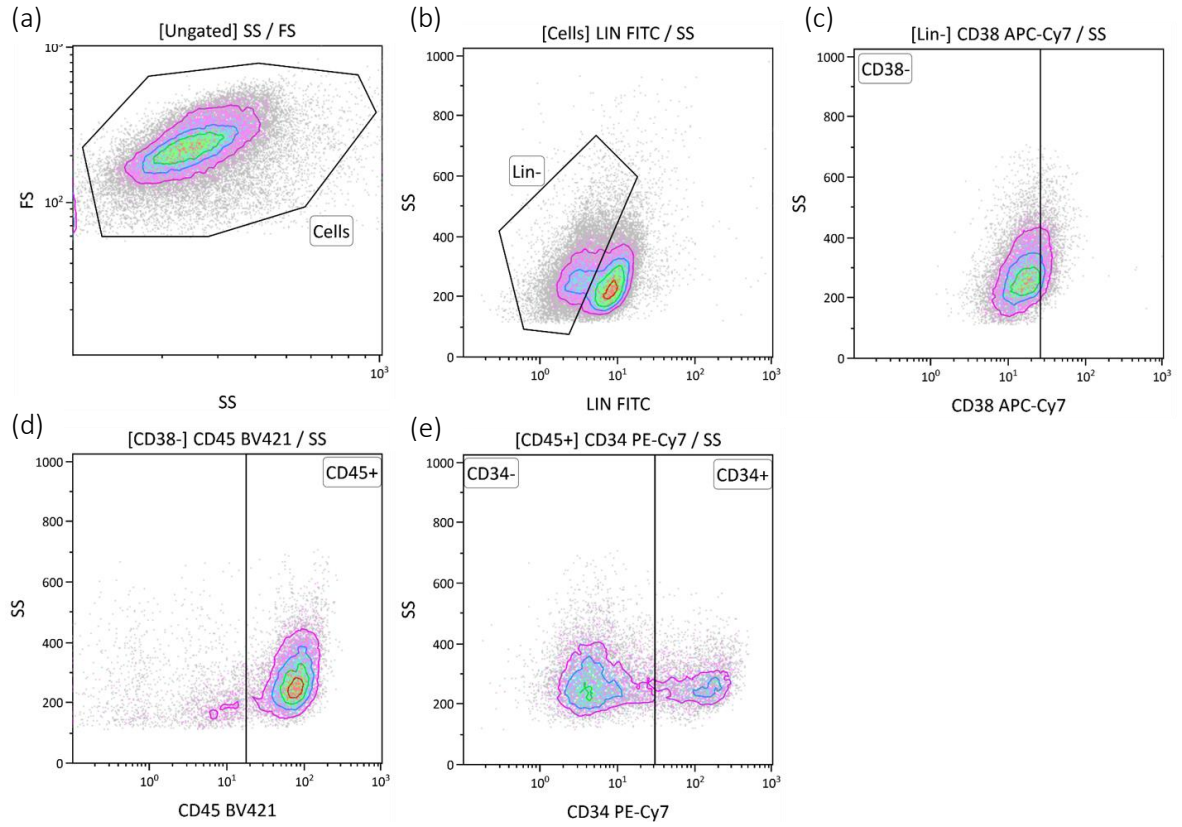


Figure 2.3. Immunophenotypic gating strategy 1. Sequential gating strategy for HSPC-associated immunophenotypic markers. Density plots showing (a) FS vs. SS and cells stained with (b) Lin FITC, (c) CD38 APC-Cy7 (d) CD45 BV421 and (e) CD34 PE-Cy7 monoclonal antibodies.

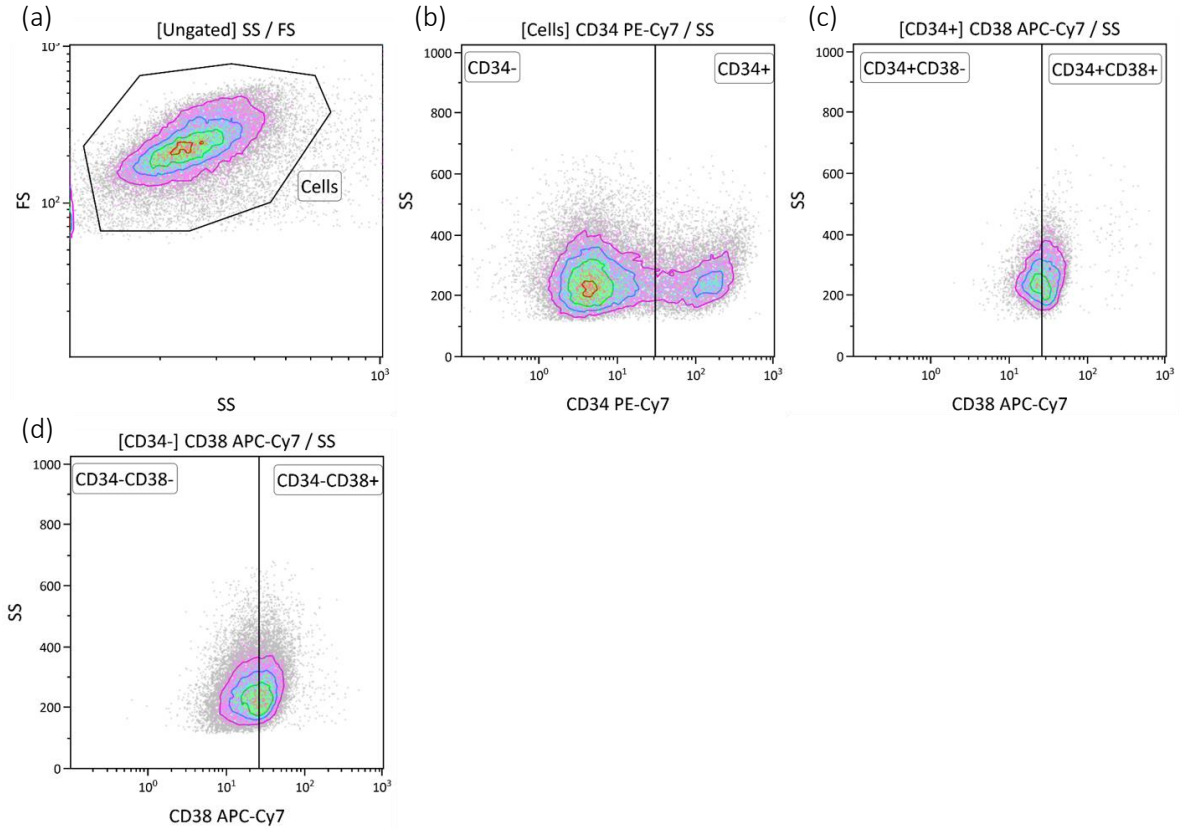


Figure 2.4. Immunophenotypic gating strategy 2. The sequential gating strategy for HSPC-associated immunophenotypic markers. Density plots showing (a) FS vs. SS and cells stained with (b) CD34 PE-Cy7, (c – d) CD38 APC-Cy7 monoclonal antibodies gated on (c) $CD34^{+}$ and (d) $CD34^{-}$ regions, respectively.

Table 2.2 shows the Gallios flow cytometer filter configurations and information regarding antibodies that were used, including their respective fluorochromes, clones and the composition of the Lineage cocktail. Spectral overlap is a common phenomenon when multi-colour antibody panels are used. Spectral overlap refers to signal spill over into adjacent fluorescent detection channels due to overlapping emission spectra. Spectral overlap may lead to erroneous, especially false positive results, if not corrected. An electronic process, known as colour compensation, is used to correct for this overlap and subtract the spill over signal from the true signal. Single-colour antibodies were used to set up colour compensation.

Table 2.2. Gallios flow cytometer filters configuration.

Laser	Filter	FL	Fluorochrome/Dye	Clone
488nm, 22mW	252/40	1	Lineage cocktail FITC	
			CD3	UCHT1
			CD14	HCD14
			CD16	3G8
			CD19	HIB19
			CD20	2H7
			CD56	HCD56
	575/30	2	-	-
	620/30	3	-	-
	695/30	4	-	-
	755LP	5	CD34 PE-Cy7	581
638nM, 25mW	660/20	6	-	-
	725/20	7	-	-
	755LP	8	CD38 APC-Cy7	HIT2
405nM, 40mW	450/40	9	CD45 BV421	30-F11
	550/40	10	-	-

2.2.7. Statistical analysis

Experiments were performed in triplicate (three technical repeats per condition) using three separate UCB samples (donors). Experimental data are represented as mean ± standard deviation (SD). All statistical analyses were performed using GraphPad Prism 7. A non-parametric one-way analysis of variance (ANOVA), a Kruskal Wallis, and a multiple comparisons Dunn's test were applied to determine if the differences were statistically significant, and a value of $P < 0.05$ was considered as significant.

2.3. RESULTS

2.3.1. MACS versus FACS

Researchers generally make use of MACS to enrich for $CD34^+$ cells prior to FACS. An experiment was performed to determine whether it is essential to perform an enrichment step (MACS) before FACS during the isolation of UCB-derived HSPCs. Sample viability, percentage $CD34^+CD45^{dim}$ cells and absolute $CD34^+CD45^{dim}$ cell numbers were used as endpoints. As expected, the absolute number of $CD34^+CD45^{dim}$ cells decreased with every processing step (Figure 2.5). The mean number of $CD34^+CD45^{dim}$ cells was highest in whole blood, 1.79 x 10^6 (± 1.35) and decreased after NH_4Cl treatment, 3.5 x 10^5 (± 0.25). Increased $CD34^+CD45^{dim}$ cell numbers were observed after MACS, 4.4 x 10^5 (± 0.51) compared to FACS, 2.1 x 10^5 (± 0.15). However, a higher level of variability was observed for MACS compared to FACS. The sorting algorithms used by FACS are generally quite strict and could be the reason for the lower cell numbers observed after FACS compared to MACS.

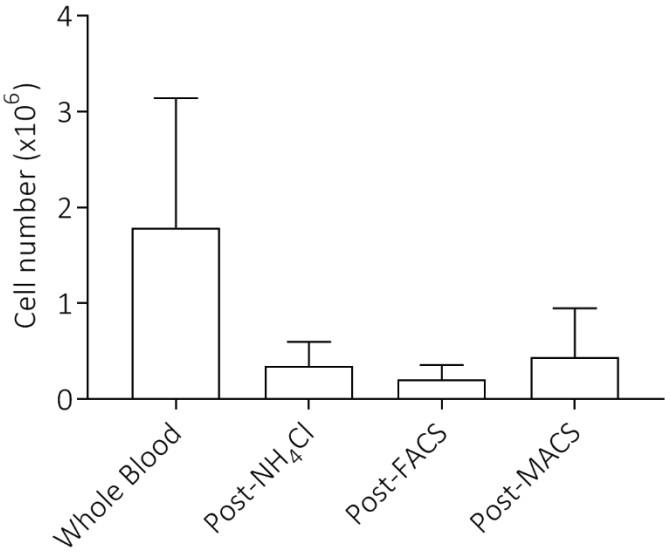


Figure 2.5. Absolute $CD34^+/CD45^{dim}$ HSPC numbers. The mean (SD) total number of $CD34^+CD45^{dim}$ HSPCs present before (whole blood), during (post-NH_4Cl) and post-isolation of HSPCs (post-FACS and post-MACS). Bars are representative of four independent donors (n = 4).

The proportion of $CD34^+CD45^{dim}$ cells in whole UCB on average is very low. The percentage $CD34^+CD45^{dim}$ cells present post-sort was used as an indication of purity (Figure 2.6). As expected, the CD34 percentages in whole blood were low, due to the presence of other cells.

The percentage $CD34^{+}CD45^{dim}$ increased from 0.21% (± 0.1%) in whole blood to 1.15% (± 0.7%) after the first enrichment step (post-NH_4Cl). The percentage/purity of $CD34^{+}CD45^{dim}$ was 83.8% (± 11.2%) and 67.6% (± 8.6%) after FACS and MACS isolation, respectively.

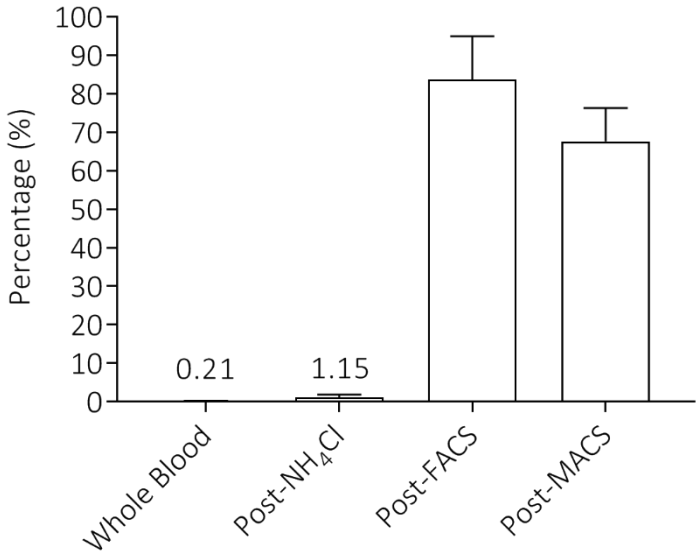


Figure 2.6. Percentage $CD34^{+}/CD45^{dim}$ HSPCs. The mean (SD) percentage $CD34^{+}CD45^{dim}$ HSPCs present before (whole blood), during (post-NH_4Cl) and post-isolation of HSPCs (post-FACS and post-MACS). Bars are representative of four independent donors (n = 4).

To ensure that cell viability was not compromised during the isolation steps, the viability of FACS- and MACS-isolated cells was investigated. Viability was high in whole blood, 96.9% (± 1.9%) and decreased slightly upon NH_4Cl treatment, 86% (± 11.3%) (Figure 2.7). Viability after FACS and MACS isolation was 95.6% (± 2.8%) and 92.7% (± 0.9%), respectively.

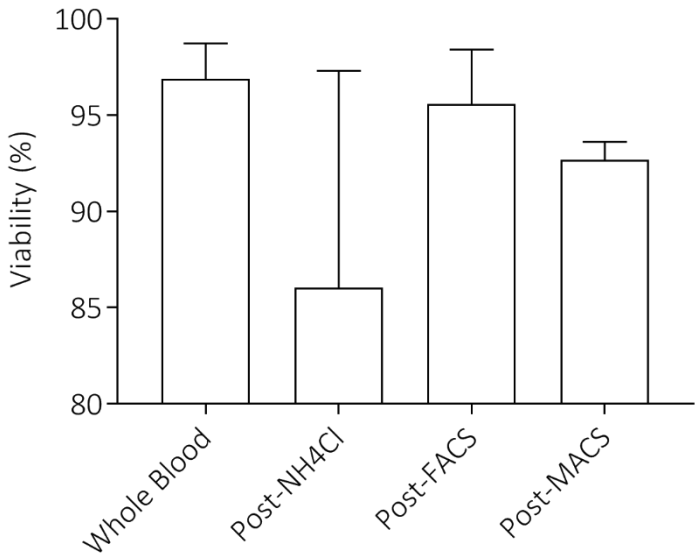


Figure 2.7. The percentage viability. The mean (SD) percentage viability of total cells before (whole blood), during (post-NH_4Cl) and post-isolation of HSPCs (post-FACS and post-MACS). Bars are representative of four independent donors (n = 4).

2.3.2. Comparison of different media conditions

A literature search was performed to determine the most frequently used GMP-compliant culture media for the culturing of HSPCs. Based on the search results, we decided to test three different serum-free media from three different companies: Stemline II (223,224), StemPro-34 Plus (225) and StemSpan ACF (116,226). These three serum-free media were compared to DMEM supplemented with FBS (Gibco™ by Life Technologies, Thermo Fisher Scientific, USA). DMEM supplemented with FBS was previously used in our laboratory for the *ex vivo* culturing of HSPCs.

Three UCB units were used in this series of experiments. Table 2.3 shows the UCB units collected together with their respective sample identities, blood volumes, viability percentages, absolute leukocyte and $CD34^+$ cell counts (post-NH_4Cl), CD34 percentages and the post-sort purity. Several studies have shown that UCB volume as well as the TNC count correlate with the number of $CD34^+$ cells in a single UCB unit (227,228).

Table 2.3. UCB units collected to determine the optimal medium for the culturing of HSPCs. UCB units collected with sample identities, blood volumes, viability percentages, absolute leukocyte and $CD34^+$ numbers (post-NH_4Cl), CD34 percentages (post-NH_4Cl) and the purity of the $CD34^+$ population post-sorting.

Sample	Volume (mL)	Viability (%)	Total leukocytes	$CD34^+$ cells	$CD34^+$ (%)	Sort purity (%)
CB291015	70	93.99	331.7×10^6	2.5×10^6	0.66	91.7
CB031115	42	96.67	136.6×10^6	9.8×10^5	0.64	90.8
CB041115	32	97.45	43.7×10^6	3.0×10^5	0.47	98.6

2.3.2.1. Viability, proliferation and CD34 cell numbers/percentages

Figure 2.8 shows the flow cytometry protocol used to determine sample viability, proliferation and CD34 cell numbers and percentages after seven days in culture.

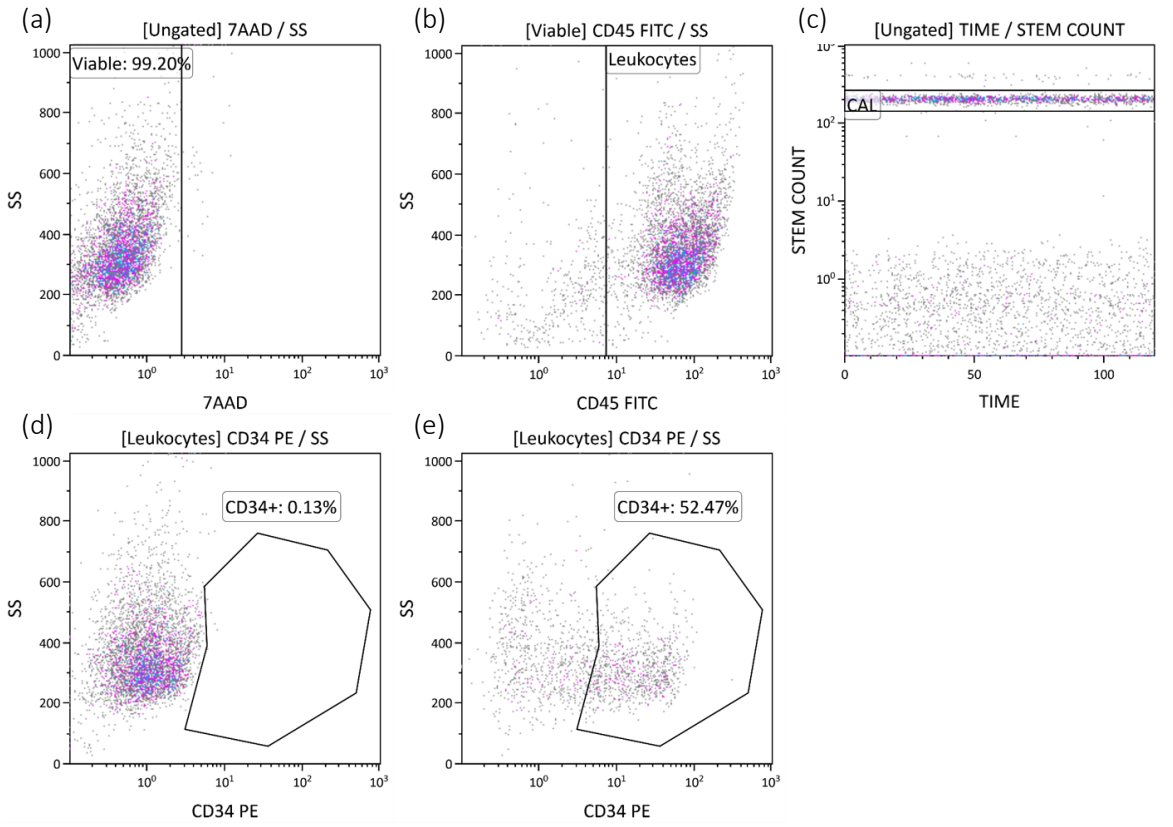


Figure 2.8. Counting protocol. Flow cytometry protocol and analysis of expanded UCB-derived HSPCs on Day 7. (a) Density plot showing viable (7AAD-negative) cells. (b) Density plot (gated on "Viable") showing $CD45^+$ cells. (c) Density plot showing Flow-Count fluorospheres in the "CAL" region. Density plot (gated on "Leukocytes") showing cells stained with (d) Stem-Kit IsoClonic control PE and (e) CD34 PE conjugated antibody. The $CD34^+$ region was set according to non-specific staining observed in (d) and kept the same to quantify the number of $CD34^+$ cells in (e).

Sample viability did not differ significantly between the different culture media used: Stemline II, 93.8% (± 4.6%), StemPro-34 Plus, 90% (± 0.9%) and StemSpan ACF, 87.4% (± 3.3%), while lower viability was observed in DMEM cultures supplemented with FBS, 57.1% (± 5.6%) (Figure 2.9).

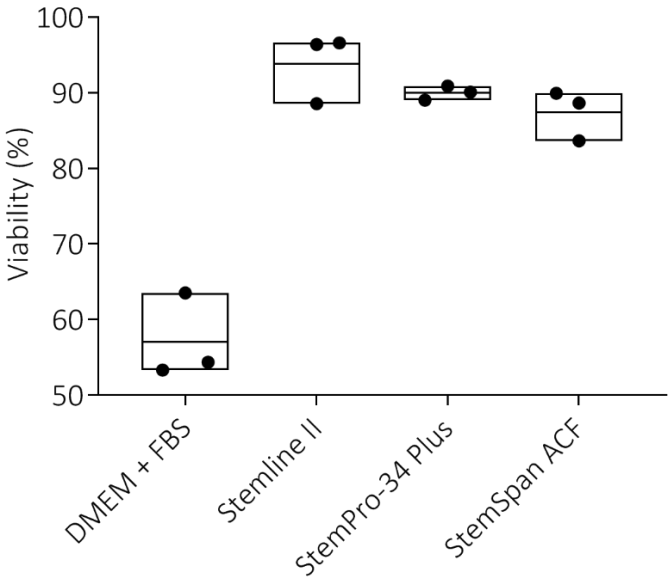


Figure 2.9. Viability after seven-day expansion. The mean percentage viable cells after seven-day expansion in different media. Bars are representative of three technical repeats of three independent donors (n = 3).

The effect of different cell culture media on cell proliferation was demonstrated as fold increase, which is described as the increase in the number of cells observed on Day 7 compared to the number of cells seeded on Day 0. In general, the use of serum-free medium resulted in higher cell numbers compared to DMEM supplemented with FBS (Figure 2.10). The highest overall proliferation (mean fold increase of 174.8 (± 29.1)) was observed in cells cultured in StemSpan ACF; the difference was statistically significant (P = 0.02) when compared to DMEM supplemented with FBS. The use of DMEM supplemented with FBS resulted in the lowest total proliferation (a mean fold increase of 24.3 (± 6.83)) (Figure 2.10).

On Day 0, the CD34$^+$ population was quite homogeneous in size and CD34 expression. The CD34$^+$ population was not as distinct on Day 7 and for that reason an IsoClonic control was included to determine the negative/positive boundary of CD34 expression. The same counting protocol was used on Day 0 and Day 7, even though the profile of the CD34$^+$ cells changed slightly upon expansion.

StemSpan ACF (fold increase 32.0 (± 8.3)) and StemPro-34 Plus (fold increase 30.8 (± 8.4)) showed the highest overall proliferation of CD34$^+$ cells (Figure 2.10). Interestingly, although StemPro-34 Plus showed intermediate total proliferation when compared to Stemline II and StemSpan ACF, it resulted in the highest proportion, 39.4% (± 1.4%) of CD34$^+$ cells (Figure 2.11), significantly (P = 0.02) higher than Stemline II, 10.6% (± 1.8%).

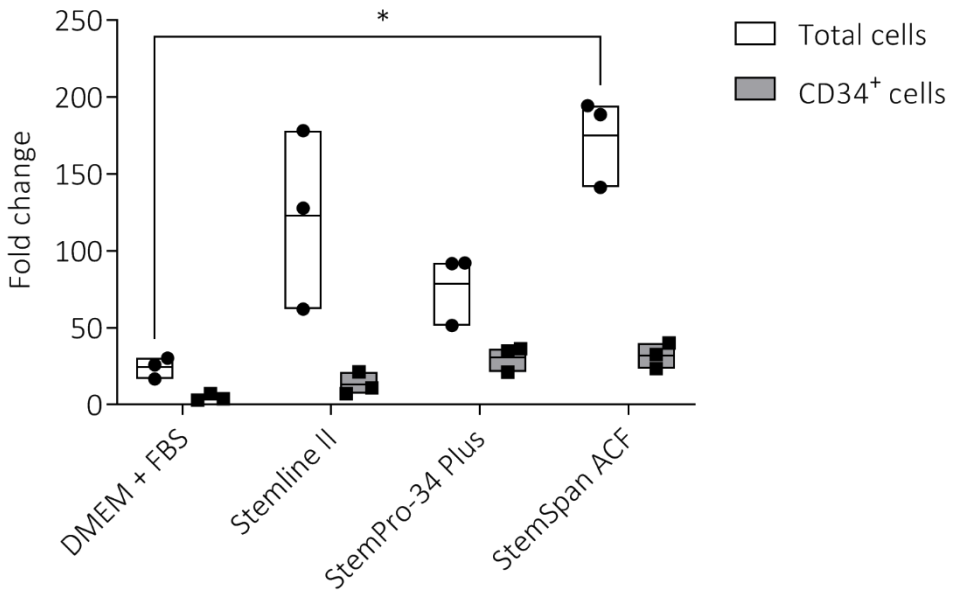


Figure 2.10. Proliferation of total and CD34$^+$ cells after seven-day expansion. The mean fold increase of viable absolute and CD34$^+$ cell numbers after seven-day expansion in different media. Bars are representative of three technical repeats of three independent donors (n = 3). * P < 0.05.

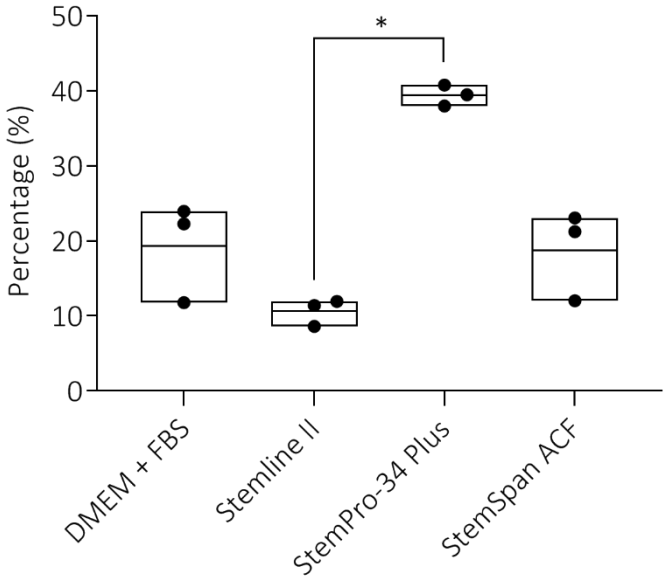


Figure 2.11. Percentage CD34$^+$ cells after seven-day expansion. The mean percentage of CD34$^+$ cells after seven-day expansion in the different media. Bars are representative of three technical repeats of three independent donors (n = 3). * $P < 0.05$.

2.3.2.2. HSPC-associated immunophenotype

There were not enough cells in DMEM supplemented with FBS cultures after seven days to perform HSPC-associated immunophenotypic analysis. Although none of the differences observed were statistically significant, StemSpan ACF showed the highest number of CD34$^+$CD38$^-$, 12.3 (± 13.6) and CD34$^+$CD38$^+$ cells, 29.7 (± 5.5) compared to Stemline II, 3.7 (± 3.8) and 6.7 (± 2.3), and StemPro-34 Plus, 6.0 (± 5.3) and 12.3 (± 2.3) (Figure 2.12). Immunophenotypic profiles for HSPCs cultured in StemPro-34 Plus (9.0 ± 8.2) and StemSpan ACF (8.0 ± 8.7) showed an increase in the Lin$^-$CD38$^-$CD45$^+$CD34$^+$ cell population (Figure 2.13), which is believed to include more primitive HSPCs. It was observed that all CD34$^+$CD38$^-$ cells were also Lin$^-$ on day seven, while the CD34$^+$CD38$^+$ cells were Lin$^+$. Numerous studies, however, have shown that the Lin$^-$CD38$^-$CD45$^+$CD34$^-$ population precedes the Lin$^-$CD38$^-$CD45$^+$CD34$^+$ population in the hematopoietic hierarchy (41,229), and for that reason, the results for the analysis of the CD34$^-$ fractions are also displayed in Figure 2.13 below. Between the three serum-free media tested, the lowest and highest numbers of Lin$^-$CD38$^-$CD45$^+$CD34$^+$ and Lin$^-$CD38$^-$CD45$^+$CD34$^-$ populations were observed in Stemline II, 1.7 (± 2.1) and 26.0 (± 11.0), and StemSpan ACF, 8.0 (± 8.7) and 42.7 (± 14.6), respectively.

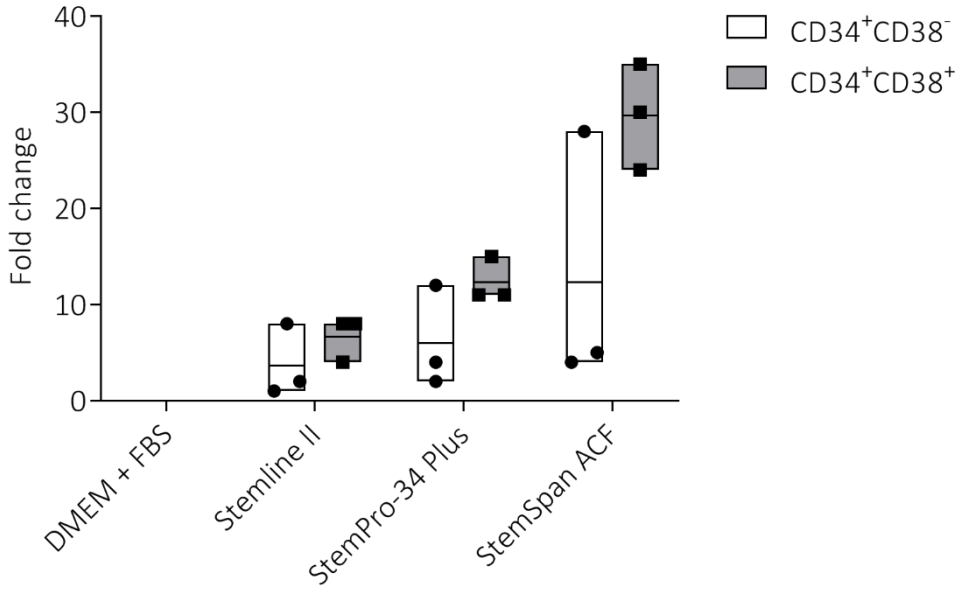


Figure 2.12. Expansion of $CD34^+CD38^-$ and $CD34^+CD38^+$ HSPCs. The mean total number of $CD34^+CD38^-$ and $CD34^+CD38^+$ HSPCs present after seven-day expansion in different media. Bars are representative of three technical repeats of three independent donors (n = 3).

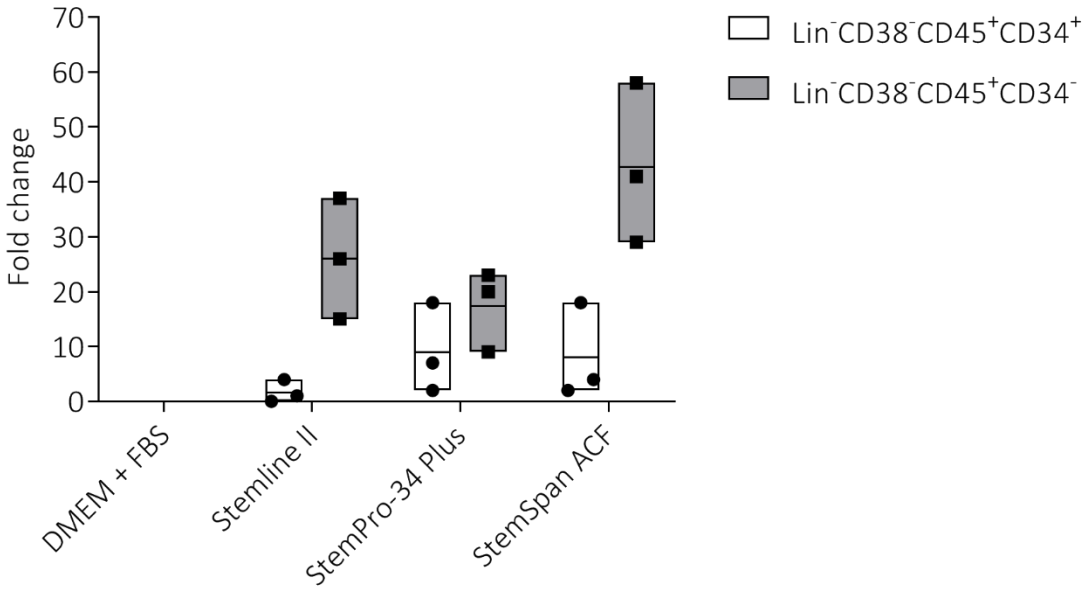


Figure 2.13. Expansion of $Lin^-CD38^-CD45^+CD34^-$ and $Lin^-CD38^-CD45^+CD34^+$ HSPCs. The mean total number of $Lin^-CD38^-CD45^+CD34^-$ and $Lin^-CD38^-CD45^+CD34^+$ HSPCs present after seven-day expansion in different media. Bars are representative of three technical repeats of three independent donors (n = 3).

2.4. DISCUSSION AND CONCLUSION

The results from the first part of this study, which compared MACS vs. FACS for the isolation of HSPCs, indicated that it is not essential to perform an enrichment step (such as MACS) before FACS sorting of rare cell populations, such as HSPCs. One advantage of MACS is that it yields greater HSPC cell numbers with acceptable sample viability. Therefore, for experiments that require high cell numbers, MACS isolation should be the preferred method of choice. FACS results in pure, viable HSPC isolations, but with greater cell loss. Therefore, if increased purity is desired, FACS isolation should be the preferred method for HSPC isolation. Cell loss is one of the great disadvantages regardless the isolation technique, and cell loss is therefore unfortunately inevitable.

The purpose of the second part of this study was to determine the optimal serum-free medium for the *ex vivo* culturing and expansion of HSPCs. In this study we compared DMEM supplemented with FBS (previously used in our laboratory), with three commercially available serum-free media, Stemline II, StemPro-34 Plus and StemSpan ACF. Although studies have reported the successful culture of HSPCs using DMEM supplemented with FBS (230), improved GMP-compliant standards and procedures are required to ensure reliable and reproducible results. Moving towards serum-free HSPC culturing conditions will cease the use of FBS and eliminate the associated safety risks and ethical concerns. This would ultimately lead to reproducible results between laboratories across the globe and ensure ease of adapting protocols for the preparation of clinical cell therapy products.

Jackson *et al.* (230) successfully cultured HSPCs using DMEM supplemented with FBS and showed good overall viability and expansion of HSPCs after seven days in culture. The current study showed that HSPCs cultured in DMEM supplemented with FBS had decreased total cell viability, decreased total and $CD34^+$ cell expansion compared to the serum-free media tested. The lower expansion achieved using DMEM resulted in insufficient cell numbers after seven-day expansion to perform HSPC-associated immunophenotypic analysis. A concentration of approximately 1000 – 2000 cells/µL was required for the immunophenotypic analysis. Although the results from our study showed smaller numbers after seven-day expansion compared to the study by Jackson and colleagues, it is important to note that the

number of HSPCs that were seeded on Day 0 by Jackson *et al.* (230) was 20 times higher. The fold increase reported by Jackson *et al.* (230) for the total cell numbers was less than was observed in this study for DMEM supplemented with FBS. The lower extent of expansion observed in DMEM supplemented with FBS could possibly be explained in part by the choice of medium. DMEM is generally used in most laboratories to culture and expand adherent cells, while cells in suspension, including HSPCs, are generally cultured in Iscove's Modified Dulbecco's Medium (IMDM) or Rosswell Park Memorial Institute (RPMI). HSPCs cultured in DMEM supplemented with FBS is not optimal and might therefore take longer to expand to sufficient numbers, which might also be a possible reason for the cell death observed.

Stemline II showed a greater extent of total cell expansion compared to StemPro-34 Plus, however, less expansion was observed when compared to StemSpan ACF. A higher fold increase of over 250 was observed by Leugers *et al.* (231) when $CD34^+$ HSPCs were cultured in Stemline II for 10 days. However, the $CD34^+$ HSPCs in the above-mentioned study were in culture for three additional days. Similar to this study, Spanholtz *et al.* (224) observed approximately 100-fold total expansion when UCB-derived $CD34^+$ cells were cultured in Stemline II for seven days. StemPro-34 Plus showed an increased proportion of $CD34^+$ cells compared to the other serum-free media; however, the total cell numbers were noticeably lower. StemSpan ACF showed greater expansion of both total and $CD34^+$ cells compared to all other serum-free media tested.

$CD34^+$ HSPCs are a heterogeneous population containing varying subsets of stem and progenitor cells at different stages of lineage commitment. Only a small proportion of HSPCs have repopulating abilities. Heterogeneity within the HSPC population is somewhat reflected in the variable CD38 expression observed. CD38 expression shows a gradient with increasing density of this surface antigen rather than clear positive and negative populations, and increased CD38 expression is associated with HSPC differentiation and maturation. Studies have shown that the $CD34^+CD38^+$ population encompass the hematopoietic progenitors, while the $CD34^+CD38^-$ population includes more primitive cells. Lineage-restricted ($CD34^+CD38^+$) progenitor cells are mostly responsible for short-term engraftment, while $CD34^+CD38^-$ cells are responsible for long-term engraftment (232). Thus, HSPCs with repopulating ability, known as

long-term HSCs, are enriched in the $Lin^{-}CD38^{-}CD34^{+}$ population (233). Studies have suggested that the $Lin^{-}CD38^{-}CD34^{-}$ population is a more primitive population of HSPCs (41,229). However, this population has slower repopulating ability compared to $Lin^{-}CD38^{-}CD34^{+}$ cells (234).

This study further revealed similar proportions of $CD34^{+}CD38^{-}$ and $CD34^{+}CD38^{+}$ cells in all serum-free cultures after expansion, measured at a phenotypic level. These results indicate that both primitive and committed $CD34^{+}$ populations are present in HSPC cultures expanded in serum-free medium. Similarly, $Lin^{-}CD38^{-}CD34^{+}$ and $Lin^{-}CD38^{-}CD34^{-}$ cells were present, with the $Lin^{-}CD38^{-}CD34^{-}$ population being much greater than the $Lin^{-}CD38^{-}CD34^{+}$ population. Overall, cultures expanded in StemSpan ACF resulted in increasing cell numbers when compared to Stemline II and StemPro-34 Plus. The use of serum-free medium for the culturing of HSPCs for clinical applications has raised several concerns, including that certain media formulations might favour the increase of certain populations of cells. A very recent study by Goncalves *et al.* (235) revealed that UM171 and HDAC inhibitors increased HSPC-associated cell surface markers, even though the HSPCs lacked engraftment potential.

Although flow cytometry is a fast and effective method to determine the nature of the cells present based on specific cell surface markers, one of its limitations is that it does not provide any functional data. Colony forming cell (CFC) assays constitute a functional *in vitro* method often used to prove that sub-populations of cells, identified using flow cytometry, are functional. It assumes that functional cells are able to proliferate and differentiate into different hematopoietic cell types. However, the most effective method to determine whether cells are long-term HSCs is through serial *in vivo* transplantations. This will determine whether the transplanted cells are able to repopulate recipients who have received myeloablative therapy, using experimental animal models (236). *In vivo* transplantation studies are however costly, and since the sole purpose of this study was to determine the optimal serum-free medium for the *ex vivo* culturing of HSPCs, flow cytometry alone was considered to be sufficient. For the purpose of expanding HSPCs for clinical applications, *in vivo* transplantation

studies would need to be performed to confirm the presence of long-term HSCs after *ex vivo* culturing of HSPCs in order to determine the effect of the different media (if any) on the success of engraftment.

In conclusion, this comparative study demonstrated improved viability, proliferation and HSPC-associated phenotype of HSPCs when cultured for seven days in three different serum-free media compared to DMEM supplemented with FBS. Based on the endpoints measured during this study, StemSpan ACF appears to be more suitable for culturing HSPCs. Also, the use of StemSpan ACF allows for GMP-compliant culturing of HSPCs *ex vivo*. We do however acknowledge that this study did not include all serum-free media commercially available. This study also revealed that heterogeneity exists in the response to different media conditions, which suggests that it will be important to move towards global standardisation of *ex vivo* HSPC culture conditions in the future. Functional *in vitro* colony forming assays and *in vivo* transplantation studies would be essential to ultimately determine the optimal serum-free medium for clinical applications.

This chapter ties in with Chapter 3, where HSPCs are cultured in serum-free expansion medium (StemSpan ACF) with different cytokine combinations. The work in Chapters 2 and 3 was mainly performed with the aim to standardise the culturing of HSPCs in our laboratory and to make this as clinically relevant as possible to ensure that we use products that conform to GMP standards.

CHAPTER 3. *EX VIVO* EXPANSION OF HEMATOPOIETIC STEM AND PROGENITOR CELLS

3.1. INTRODUCTION

Umbilical cord blood transplantation is performed for hematological malignancies and other disorders (237). However, the number of TNC and $CD34^+$ HSPCs is less than what is obtained from bone marrow and mobilised peripheral blood. Therefore to ensure engraftment success and hematopoietic reconstitution, several UCB units may be co-transplanted (73). Delayed hematopoietic recovery of neutrophils and platelets has been reported following UCBT, as a result of the immature nature of the cells (238). The lower rate of hematopoietic recovery increases the susceptibility of recipients to microbial and viral infections, which has contributed to the mortality rate of above 30% following UCBT (239).

Ex vivo expansion of HSPCs to increase cell numbers and shorten time to engraftment has been an area of intense interest for several years. *Ex vivo* expansion increases the number of long- and short-term HSPCs, which contribute to successful engraftment. One way in which HSPCs have been expanded *ex vivo* is with the use of exogenous cytokines (240). Cytokines are small secreted proteins that are involved in cell signaling and regulate many aspects of hematopoiesis (241). Different combinations of cytokines have been used for the *ex vivo* expansion of HSPCs, and several clinical trials have been undertaken to determine whether delayed engraftment following UCBT could be overcome with the combined infusion of *ex vivo* expanded (manipulated) and non-expanded (non-manipulated) HSPCs (Table 3.1).

Table 3.1. Cytokine combinations used for the *ex vivo* expansion of UCB-derived HSPCs in clinical trials.

Cells	Expansion time	Media	Growth factors	Additional components	Neutrophil engraftment	Platelet engraftment	References
$CD34^+$	16d	Serum-free medium	SCF, TPO, FLT3L, IL-3 and IL-6	Notch ligand Delta1	16	-	(84)
$CD133^+$	3w	α-MEM + 10% FBS	SCF, TPO, FLT3L and IL-6	Nicotinamide	13	33	(77)
$CD133^+$	3w	α-MEM + 10% FCS	SCF, TPO, FLT3L and IL-6	Copper chelator	30	48	(242)
UCB cells	12d	IMDM + 10% FBS + 10% horse serum	PIXY321, FLT3L and EPO	-	22	106	(243)
$CD34^+$	10d	Commercial medium	SCF, G-CSF and MGDF	-	28	94	(94)
$CD34^+$	15d	Commercial medium	SCF, TPO, FLT3L and IL-6	StemRegenin 1	15	49	(78)
MNCs	2w	α-MEM + 20% FBS	SCF, TPO, FLT3L and G-CSF	Mesenchymal stromal cell layer	15	42	(244)

d = days; w = weeks; Cells = cell types purified through various purification methods prior to expansion; Commercial medium = commercially available medium for the expansion of HSPCs containing proprietary ingredients.

G-CSF is a cytokine that induces granulocyte differentiation and proliferation (245). G-CSF is used clinically as a mobilising agent for the enrichment of $CD34^+$ HSPCs in the peripheral circulation (246) and also to improve and/or prevent chemotherapy-induced neutropenias (247).

The purpose of this study was to determine whether the addition of G-CSF to well-established cytokine combinations, FLT3L, SCF, TPO and IL-3 or IL-6 would lead to improved *ex vivo* expansion of HSPCs. This chapter describes the results obtained from expanding $CD34^+$ HSPCs with different cytokine combinations, using the following end-points: viability, proliferation (TNC and $CD34^+$ HSPCs numbers and percentages), HSPC-associated immunophenotype and side population (SP) analysis.

3.2. MATERIALS AND METHODS

3.2.1. Ethics statement

An amendment to the informed consent document, shown in Appendix E, was made in May 2018 to include HIV-1 testing of all collected UCB samples. A copy of the amended informed consent document that was completed by mothers prior to collection is included in Appendix F.

3.2.2. Sample collections

Informed consent and sample collections were performed as described in Chapter 2. Six independent UCB units were collected for this part of the study.

3.2.3. HIV testing

The HIV status of patients was obtained from their files at a private hospital in Pretoria, Gauteng. As confirmation, a 3 mL aliquot of the collected UCB samples was used to screen for HIV-1 using the GeneXpert 1 System (Cepheid, USA) and the Xpert® HIV-1 Qual cartridge (Cepheid, USA). The Xpert HIV-1 Qual is a qualitative molecular test to measure total nucleic acids of HIV-1 RNA and proviral DNA. An internal control is included in the Xpert HIV-1 Qual cartridge to ensure sample processing and to monitor the presence of inhibitors in reverse

transcription (RT) and PCR reactions. Samples with a status 'HIV-1 not detected' were used for these experiments.

3.2.4. Enrichment of $CD34^+$ HSPCs

The collection bag content was carefully layered onto Histopaque®-1077 (Sigma-Aldrich, USA) in sterile 50 mL Falcon tubes without mixing the two phases in a 2:1 volume ratio (30 mL of UCB onto 15 mL Histopaque-1077) and centrifuged for 30 min at 1700 rpm. Between three to four 50 mL Falcon tubes were processed for a single UCB unit. The plasma fraction was aspirated and the mononuclear cell (MNC) layer, from different tubes for the same UCB unit, was collected and pooled into a single 50 mL Falcon tube. A wash step was included (centrifugation at 300 x g for 10 minutes), by filling up the tube to 50 mL with TP buffer, to remove any excess Histopaque. TP buffer was aspirated and red blood cells were lysed by adding 20 mL of NH_4Cl for 20 min at 4°C. The sample was centrifuged at 300 x g for 10 min at 4°C. NH_4Cl was aspirated and each sample was washed twice by filling the tube to 30 mL with TP buffer and centrifugation at 300 x g for 10 min at 4°C. Supernatant was aspirated and cells were resuspended in 5 – 10 mL TP buffer. At this point a 50 µL aliquot was taken for to determine absolute cell count.

3.2.5. Flow cytometer: Counting protocol

Stem-Kit™ reagents from Beckman Coulter (Miami, Florida, USA) were used to determine viability and absolute number of $CD34^+$ cells in each sample using the Gallios flow cytometer, as discussed in Chapter 2, Section 2.2.3.1. Antibody staining and determination of total number of cells in each sample were also carried out as discussed in Chapter 2. The same protocol as in Chapter 2 was used to determine cell count. All flow cytometer analyses were performed using Beckman Coulter Kaluza Analysis Software (version 2.1).

3.2.6. FACS sorting

FACS sorting was performed as discussed in Chapter 2, Section 2.2.3.3. $CD34^+$ cells were stained and sorted from the MNC fraction. Viable $CD34^+$ cells were sorted into wells (1×10^4 cells/well) of a 24-well plate (Thermo Fisher Scientific, USA) containing four different media conditions in triplicate.

3.2.7. Culture conditions

FACS-isolated $CD34^+$ HSPCs (1 x 10^4 cells/well) were cultured in 24-well plates using 1 mL serum-free StemSpan ACF medium with 2% pen/strep. Four different combinations of growth factors (Life Technologies, Thermo Fisher Scientific, USA), each at a 100 ng/mL, were used in this study (Table 3.2). All experimental conditions were performed in triplicate. Cells were maintained at 37°C and 5% CO_2 in a humidified atmosphere for seven days.

Table 3.2. Cytokine combinations. The four different cytokine combinations used during this study.

Combinations	Cytokines
Combination 1	FLT3L, SCF, TPO and IL-3
Combination 2	FLT3L, SCF, TPO, IL-3 and G-CSF
Combination 3	FLT3L, SCF, TPO and IL-6
Combination 4	FLT3L, SCF, TPO, IL-6 and G-CSF

3.2.8. Viability, proliferation and $CD34^+$ absolute cell numbers and percentages

The number of cells was determined on Day 7 using the counting protocol as described in Chapter 2. Cells were stained with Stem-Kit reagents, either CD34 PE / CD45 FITC monoclonal antibodies and 7AAD or CD34 IsoClonic control PE / CD45 FITC and 7AAD viability dye.

3.2.9. HSPC-associated immunophenotype

HSPC-associated immunophenotype analysis was performed on Day 7 using the Gallios flow cytometer. The following monoclonal antibodies, specific for human antigens, were used: Lineage cocktail FITC, CD34 PE-Cy7, CD38 APC-Cy7, CD45 BV421, CD133/2 PE and Mouse IgG2b PE isotype (Myltenyi Biotec, Germany) (Table 3.3). A viability dye, 7AAD, was also included to exclude non-viable cells. Single-colour antibodies were used to set up colour compensation. A sequential gating strategy (Figure 3.1) was used for the HSPC-associated immunophenotypic analysis for both $CD34^+$ and $CD34^-$ HSPCs. Viable (7AAD-negative) and intact cells were analysed using a FS vs. SS plot. $CD34^+$ and $CD34^-$ cells were identified using CD34 PE-Cy7 [Excitation: 488/532/561 nm; Emission: 575/780 nm] gated on 'intact' cells. Lin^- and Lin^+ cells were identified using the lineage cocktail FITC [Excitation: 488 nm; Emission: 520 nm], which were gated on either the "$CD34^+$" or "$CD34^-$" regions. The $CD38^-$ and $CD38^+$ cells were

identified using CD38 APC-Cy7 [Excitation: 633 nm; Emission: 776 nm] and were gated on the "Lin^{-} " or "Lin^{+}" regions, respectively. $CD133^{+}$ cells were identified using the CD133 FITC [Excitation: 496 nm; Emission: 578 nm], which were gated on either the "$CD38^{+}$" or "$CD38^{-}$" regions, respectively. Data was also analysed using a second gating strategy to identify $CD34^{+}CD38^{+}$ and $CD34^{+}CD38^{-}$ cells. $CD34^{+}$ and $CD34^{-}$ cells were gated on "Intact" cells. $CD34^{+}CD38^{+}$, $CD34^{+}CD38^{-}$, $CD34^{-}CD38^{+}$ and $CD34^{-}CD38^{+}$ cells were identified by gating CD38 APC-Cy7 on "$CD34^{+}$" and "$CD34^{-}$ " regions, respectively. The negative and positive regions for each phenotypic marker were created based on positive and negative populations present in a sample (if applicable) or whole blood when the distinction between the negative and positive populations was not clear. The CD133 region was created using an isotypic control.

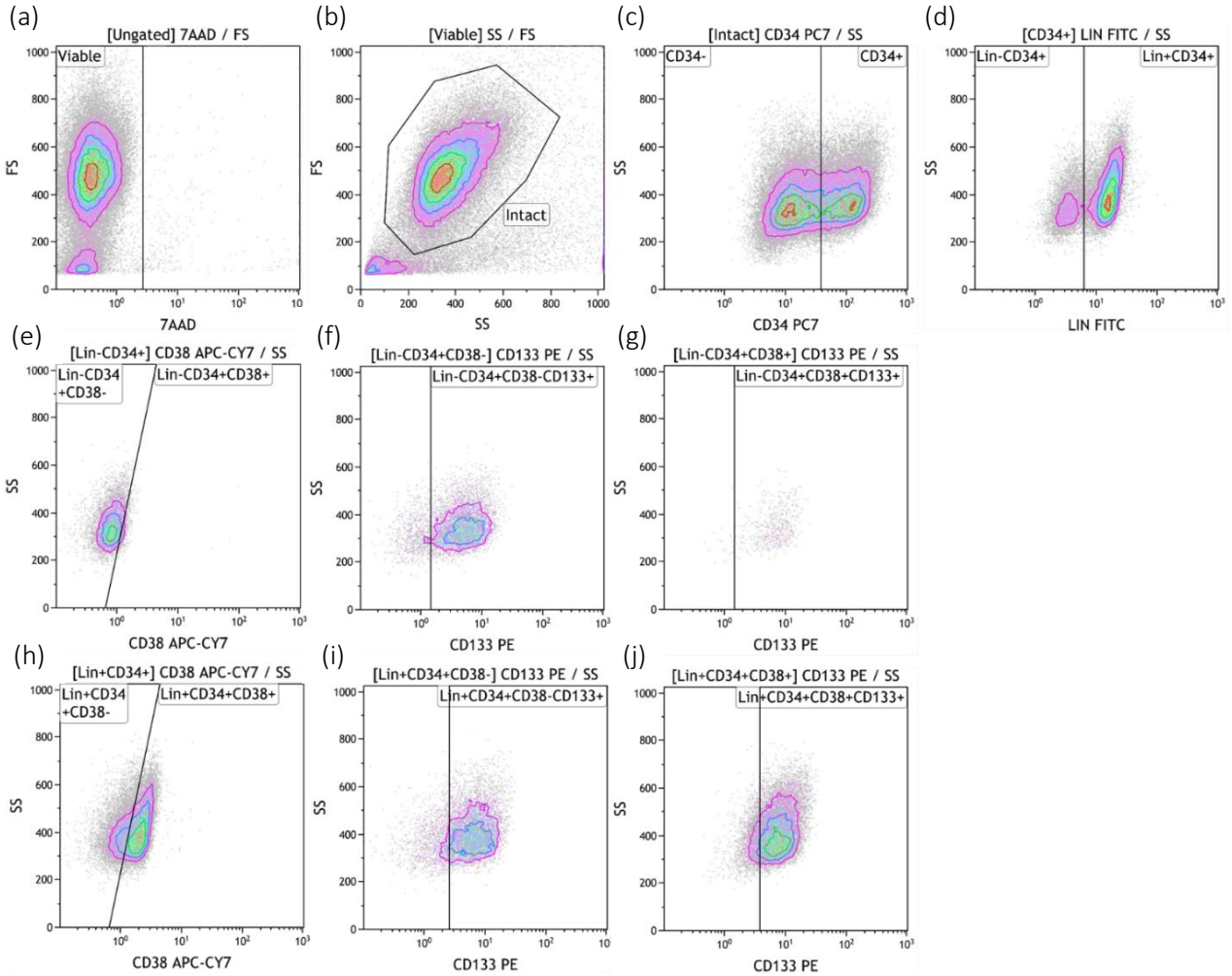


Figure 3.1. HSPC-associated immunophenotypic analysis. A schematic illustration of the sequential gating strategy used for the HSPC-associated immunophenotypic analysis. Density plot showing (a) viable (7AAD-negative) and (b) intact cells. (c) Density plot showing cells stained with (c) CD34 PE-Cy7, (d) lineage cocktail FITC, (e and h) CD38 APC-Cy7, (f – g and i – j) CD133 PE monoclonal antibodies.

Table 3.3. Gallios flow cytometer filter configuration. Gallios flow cytometer filters with respective fluorochromes/dyes for each monoclonal antibody used during HSPC-associated immunophenotypic analysis.

Laser	Filter	FL	Fluorochrome/Dye	Clone
488nm, 22mW	252/40	1	Lineage cocktail FITC	
			CD3	UCHT1
			CD14	HCD14
			CD16	3G8
			CD19	HIB19
			CD20	2H7
			CD56	HCD56
	575/30	2	CD133/2 PE	293C3
			Mouse IgG2b PE Isotype	IS6-11E5.11
	620/30	3	-	-
	695/30	4	7AAD	-
	755LP	5	CD34 PE-Cy7	581
638nM, 25mW	660/20	6	-	-
	725/20	7	-	-
	755LP	8	CD38 APC-Cy7	HIT2
405nM, 40mW	450/40	9	CD45 BV421	30-F11
	550/40	10	-	

3.2.10. Side population analysis

Analysis of the SP was performed after eight days in culture. The four different cytokine combinations were added in triplicate as mentioned above. On Day 8, all the cells (from triplicate wells) for a given cytokine condition, that remained after counting and immunophenotyping on Day 7, were pooled for SP analysis. A 50 µL sample aliquot was taken for counting (as previously described) and stained with 5 µL Stem-Kit reagents (CD34 PE / CD45 FITC monoclonal antibodies and 7AAD or CD34 IsoClonic control PE / CD45 FITC and 7AAD) to determine the number of viable and $CD34^+$ HSPCs present in the pooled sample. Prior to running the samples on the Gallios flow cytometer, 50 µL of Flow-Count Fluorospheres were added to each sample. Once the number of cells in a sample had been determined (previously described in Chapter 2, Section 2.2.3.1), cells were resuspended at a concentration of 1×10^3 viable cells/µl in StemSpan ACF medium. A control tube was included for each condition,

containing a calcium channel blocker, Verapamil (100 µM; Sigma-Aldrich, USA). Verapamil was added to appropriate tubes and incubated for 10 min prior to the addition of Vybrant® DyeCycle™ (VDC) Violet (Life Technologies, Thermo Fisher Scientific, USA). After the 10 min incubation, 5 µg/mL VDC Violet was added to all the tubes and incubated at 37°C for 120 min, protected from light. Cells were stained with CD34 PE-Cy7 and CD38 APC-Cy7 (Biolegend, USA) for 10 min prior to analysis. Annexin V FITC (Beckman Coulter, USA) and 7AAD were added to detect apoptotic and necrotic cells, respectively. After 120 min incubation, cells were placed on ice and analysed on the BD FACSAria Fusion.

A sequential gating strategy was applied during data analysis. For the sequential gating strategy, cells were first analysed using a FS vs. SS plot to capture intact cells in the "Cells" region (Figure 3.2). Necrotic and apoptotic cells were excluded using a 7AAD vs. Annexin V plots and viable cells were captured in the "Viable" region. Cells positive for 7AAD only and/or cells positive for both 7AAD and Annexin V were considered to be non-viable. Cells positive for only Annexin V generally include early apoptotic cells. The cell membranes of these cells are still intact and therefore the cells could still recover (248). Thus, cells only positive for Annexin V were included in the "Viable" region. The $CD34^{+}CD38^{-}$ and $CD34^{-}CD38^{-}$ cells (Figure 3.2c) were identified using CD38 APC-Cy7 [Excitation: 633 nm; Emission: 776 nm] and CD34-PE-Cy7 [Excitation: 488/532/561 nm; Emission: 575/780 nm] and were gated on the "Viable" region. The quadrants in Figure 3.2c were determined based on $CD34^{+}/CD34^{-}$ and $CD38^{+}/CD38^{-}$ populations present in the sample. $CD34^{+}$ SP and $CD34^{-}$ SP cells were identified using VDC Violet [Excitation: 369 nm; Emission: 437 nm] and were gated on "$CD34^{+}CD38^{-}$" and "$CD34^{-}CD38^{-}$" quadrants, respectively. The "SP" region in $CD34^{+}CD38^{-}$ (Figure 3.2d) and $CD34^{-}CD38^{-}$ (Figure 3.2e) cells was determined based on the Verapamil control.

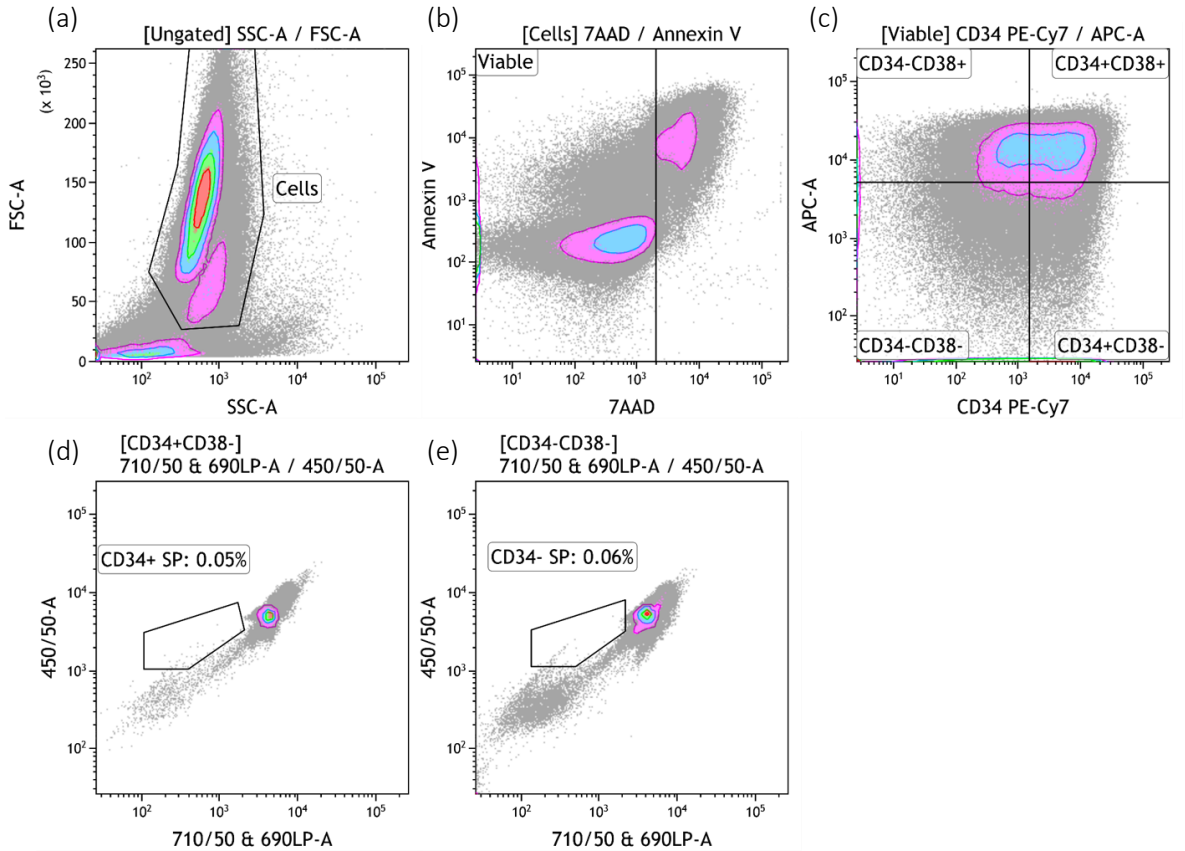


Figure 3.2. Side population analysis. A schematic illustration of the sequential gating strategy used to identify the SP in $CD34^{+}CD38^{-}$ and $CD34^{-}CD38^{-}$ cells. (a) Density plot showing intact cells (b) and viable (7AAD-negative) cells. (c) Density plot showing cells stained with CD38 APC-Cy7 and CD34 PE-Cy7 monoclonal antibodies. (d) and (e) Density plots showing cells stained with VDC Violet. Plots were used to detect SP in (d) $CD34^{+}CD38^{-}$ and (e) $CD34^{-}CD38^{-}$ cells.

3.2.11. Statistical analysis

Experiments were performed in triplicate (three technical repeats per cytokine combination) using three to five independent UCB samples. Statistical analysis was performed as described in Chapter 2.

3.3. RESULTS

Six independent UCB units were collected to determine the optimal cytokine combination for the *ex vivo* expansion of $CD34^{+}$ HSPCs. Table 3.4 shows the UCB units collected together with their respective sample identities, blood volumes, viability percentages, absolute leukocyte and $CD34^{+}$ numbers (post-NH_4Cl), CD34 percentages (post-NH_4Cl) and the purity of the $CD34^{+}$ population post-sorting. All of these units were HIV negative.

Table 3.4. UCB units collected to determine the optimal cytokine combination. UCB units collected with sample identities, blood volumes, viability percentages, absolute leukocyte and $CD34^+$ numbers (post-NH_4Cl), CD34 percentages (post-NH_4Cl) and the purity of the $CD34^+$ population post-sorting.

Sample	Volume (mL)	Viability (%)	Total leukocytes	$CD34^+CD45^{dim}$ cells	$CD34^+$ (%)	Sort purity (%)
CB080317	90	96.07	49.1×10^6	7.9×10^5	1.64	99.5
CB290317	65	97.90	35.0×10^6	5.2×10^5	1.53	96.7
CB110517	68	85.68	34.1×10^6	5.8×10^5	1.71	94.1
CB260517	45	85.19	26.8×10^6	9.9×10^5	3.68	98.5
CB310517	87	83.56	27.0×10^6	4.1×10^5	1.63	97.4
CB210617	120	97.00	30.0×10^6	1.7×10^6	6.66	93.2

3.3.1. Viability, proliferation and $CD34^+$ cell numbers/percentages

In this study, human $CD34^+$ HSPCs were cultured using four different cytokine combinations (Table 3.1). Proliferation was observed with all four cytokine combinations after seven days in culture (Figure 3.3). Proliferation is shown as fold-increase, which is described as the increase in the number of cells observed on Day 7 compared to the number of cells seeded on Day 0. 1×10^4 $CD34^+$ cells were seeded on the day of isolation (Day 0) for expansion. All cytokine combinations showed an overall increase in proliferation (total cell number) after seven days in culture (Figure 3.3a). Increased total proliferation (mean fold increase) was observed in cultures supplemented with Combination 2, 148.4 (± 34.7) and Combination 4, 90.2 (± 15.0) when compared to their respective controls, Combination 1, 113.4 (± 30.6) and Combination 3, 61.4 (± 15.2). The number of $CD34^+$ HSPCs in all cultures treated with different cytokine combinations also increased after seven days in culture. The mean fold increase in $CD34^+$ HSPCs was higher in cultures supplemented with G-CSF, Combination 2, 29 (± 5.9) and Combination 4, 25.2 (± 6.3), when compared to cultures without G-CSF, Combination 1, 21 (± 4.9) and Combination 3, 20 (± 6.3) (Figure 3.3b). The mean fold increase in $CD34^-$ cells was higher than the mean fold increase of $CD34^+$ cells in all cultures after seven days. Cultures supplemented with G-CSF showed increased numbers of $CD34^-$ cells, Combination 2, 107.2 (± 21.9) and Combination 4, 61.8 (± 11.8), when compared to their respective controls without G-CSF, Combination 1, 88.2 (± 17.9) and Combination 3, 35.2 (± 11.9).

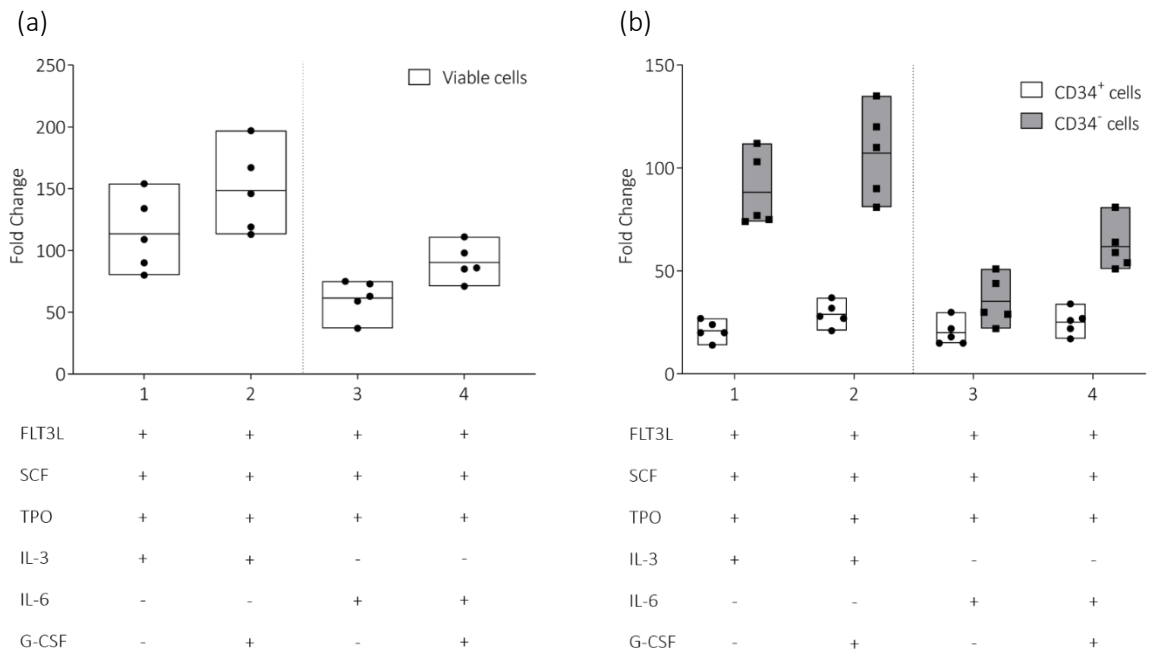


Figure 3.3. Proliferation after seven-day expansion. Fold increase of (a) viable and (b) $CD34^+$ and $CD34^-$ absolute cell numbers (mean fold increase) after seven-day expansion. Bars are representative of three technical repeats of five independent donors (n = 5).

Although the $CD34^+$ HSPC absolute numbers increased, the mean proportion of $CD34^+$ cells decreased in all cytokine combinations during the seven-day expansion period (Figure 3.4): Combination 1, 22.4% (± 4.3%); Combination 2, 25% (± 4.1%); Combination 3, 42.2% (± 3.3%); and Combination 4, 33.8% (± 1.5%). Combination 3 showed the highest proportion of $CD34^+$ cells, 42.2% (± 3.3%) after expansion, which decreased with the addition of G-CSF (Combination 4; 33.8% (± 1.5%)). Combinations 1 and 2, containing IL-3, had similar proportions of $CD34^+$ cells, namely 25% (± 4.1%) and 22.4% (± 4.3%), respectively.

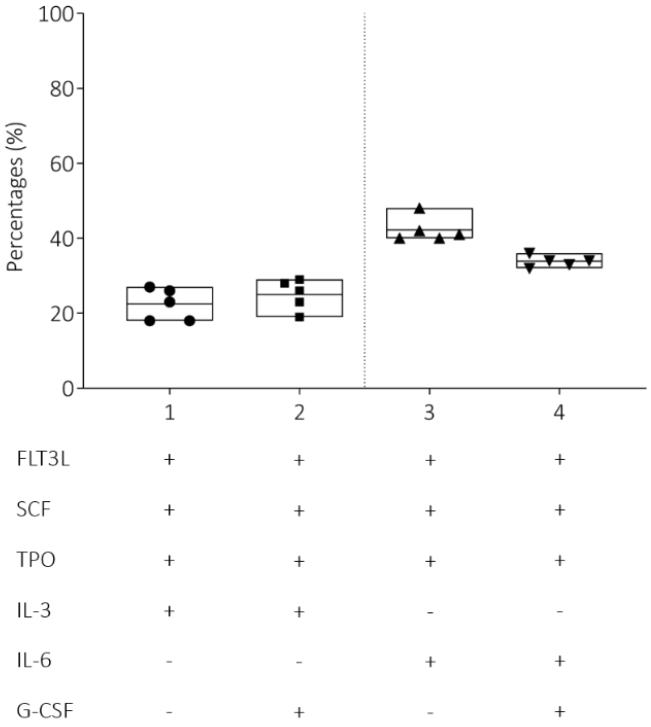


Figure 3.4. CD34⁺ proportions after seven-day expansion. The mean percentage of CD34⁺ cells after seven-day expansion in four different cytokine combinations. Bars are representative of three technical repeats of five independent donors (n = 5).

3.3.2. HSPC-associated immunophenotype

The CD34⁺ HSPC population is a heterogeneous population that consists of cells with short- and long-term repopulating ability. Combining the CD34 marker with the differentiation marker, CD38, assists in identifying HSPCs reported to have short- ($CD34^+CD38^+$) and long-term ($CD34^+CD38^-$) repopulating potential (232). Although not statistically significant, the HSPC-associated immunophenotypic results showed increased proliferation of the $CD34^+CD38^-$ population when compared to the $CD34^+CD38^+$ population after a seven-day expansion in all cytokine combinations (Figure 3.5). The mean fold increase for the $CD34^+CD38^-$ population was: Combination 1, 13.0 (± 3.1); Combination 2, 18.8 (± 2.9); Combination 3, 12.6 (± 3.6); and Combination 4, 16 (± 4.5). The mean fold increase for the $CD34^+CD38^+$ population was: Combination 1, 7.4 (± 4.3); Combination 2, 9.2 (± 4.6); Combination 3, 6.8 (± 4.3), and Combination 4, 8.4 (± 4.7).

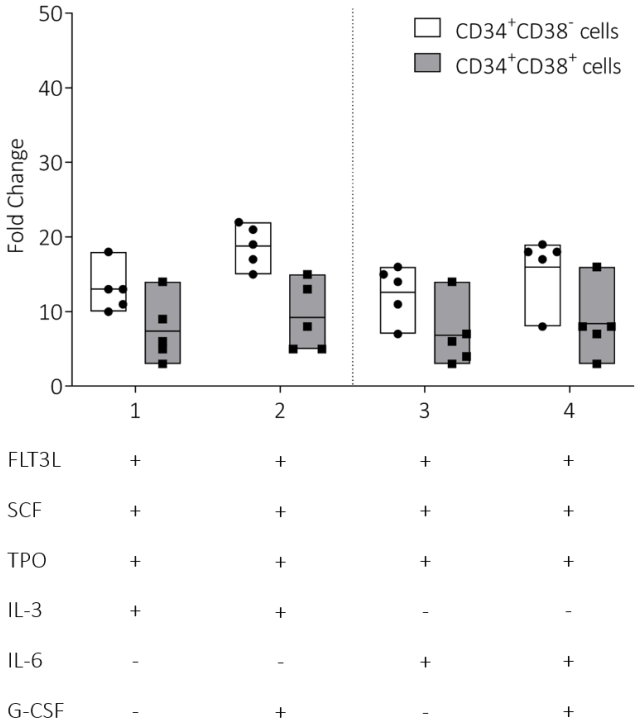


Figure 3.5. $CD34^+CD38^-$ and $CD34^+CD38^+$ HSPCs after seven-day expansion. The total number (fold change) of $CD34^+CD38^-$ and $CD34^+CD38^+$ HSPCs after seven-day expansion. Bars are representative of three technical repeats of five independent donors (n = 5).

The addition of G-CSF resulted in increased proliferation, although not statistically significant, of the $Lin^-CD34^-CD38^-$ population compared to the $Lin^-CD34^+CD38^-$ population in all four cytokine combinations after a seven-day expansion (Figure 3.6a). The most primitive stem cells are thought to be $CD34^-$ (41,229). $CD34^-$ stem cells have been shown to have slower repopulating ability compared to $CD34^+$ stem cells (234). $CD34^-CD133^+$ identifies a more primitive population of HSPCs (44), while the $CD34^+CD133^+$ population identifies HSPCs with multipotent potential (249). No difference was observed between the proliferation of the $Lin^-CD34^-CD38^-CD133^+$ population when compared to the $Lin^-CD34^+CD38^-CD133^+$ population in the four different cytokine combinations after a seven-day expansion (Figure 3.6b). The fold change (SD) of HSPC-associated immunophenotypic populations present in cultures after seven-day expansion with four different cytokine combinations is summarised in Table 3.5 and Table 3.6.

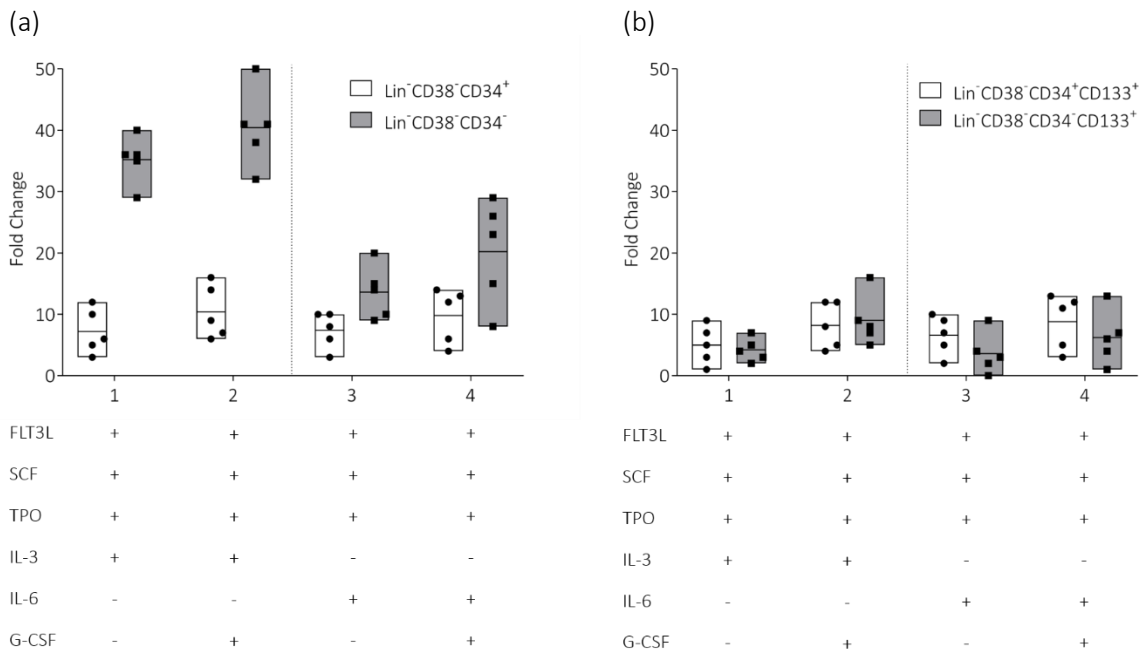


Figure 3.6. HSPC-associated immunophenophenotypic analysis after seven-day expansion. The total number (fold increase (SD)) of (a) $Lin^{-}CD38^{-}CD34^{+}$ and $Lin^{-}CD38^{-}CD34^{-}$ and (b) $Lin^{-}CD38^{-}CD34^{+}CD133^{+}$ and $Lin^{-}CD38^{-}CD34^{-}CD133^{+}$ HSPCs after seven-day expansion. Bars are representative of two technical repeats of five independent donors (n = 5).

Table 3.5. Fold increase (SD) of the total number of viable and $CD34^+$ cells after a seven-day expansion in different cytokine combinations.

	Condition			
	Combination 1	Combination 2	Combination 3	Combination 4
Viable cells	113 (± 31)	148 (± 35)	61 (± 15)	90 (± 15)
$CD34^+$	24 (± 6)	34 (± 7)	23 (± 7)	29 (± 7)
Lin^-CD34^+	8 (± 3)	11 (± 4)	8 (± 3)	10 (± 4)
$Lin^-CD34^+CD38^-$	7 (± 3)	11 (± 4)	7 (± 3)	10 (± 5)
$Lin^-CD34^+CD38^-CD133^+$	5 (± 3)	8 (± 4)	7 (± 3)	9 (± 5)
$Lin^-CD34^+CD38^+$	0	0	0	0
$Lin^-CD34^+CD38^+CD133^+$	0	0	0	0
Lin^+CD34^+	15 (± 6)	21 (± 9)	14 (± 7)	18 (± 14)
$Lin^+CD34^+CD38^-$	5 (± 2)	6 (± 1)	4 (± 3)	5 (± 5)
$Lin^+CD34^+CD38^-CD133^+$	2 (± 2)	4 (± 2)	4 (± 3)	4 (± 4)
$Lin^+CD34^+CD38^+$	10 (± 6)	14 (± 9)	9 (± 5)	12 (± 9)
$Lin^+CD34^+CD38^+CD133^+$	5 (± 3)	9 (± 5)	7 (± 4)	9 (± 7)

Data presented in this table is representative of two technical repeats of five independent donors (n = 5).

Table 3.6. Fold increase (SD) of the total number of viable $CD34^-$ cells after a seven-day expansion in different cytokine combinations.

	Condition			
	Combination 1	Combination 2	Combination 3	Combination 4
Viable cells	113 (± 31)	148 (± 35)	61 (± 15)	90 (± 15)
$CD34^-$	85 (± 17)	102 (± 21)	32 (± 12)	58 (± 11)
Lin^-CD34^-	36 (± 4)	42 (± 7)	14 (± 5)	24 (± 7)
$Lin^-CD34^-CD38^-$	35 (± 4)	40 (± 7)	14 (± 4)	20 (± 8)
$Lin^-CD34^-CD38^-CD133^+$	4 (± 2)	9 (± 4)	4 (± 3)	6 (± 5)
$Lin^-CD34^-CD38^+$	0	1 (± 1)	0	2 (± 2)
$Lin^-CD34^-CD38^+CD133^+$	0	0	0	0
Lin^+CD34^-	42 (± 12)	59 (± 25)	17 (± 9)	34 (± 14)
$Lin^+CD34^-CD38^-$	16 (± 2)	23 (± 11)	7 (± 3)	9 (± 1)
$Lin^+CD34^-CD38^-CD133^+$	2 (± 1)	4 (± 3)	1 (± 1)	1 (± 1)
$Lin^+CD34^-CD38^+$	25 (± 10)	35 (± 15)	9 (± 6)	24 (± 14)
$Lin^+CD34^-CD38^+CD133^+$	3 (± 3)	8 (± 5)	2 (± 3)	6 (± 5)

Data presented in this table is representative of two technical repeats of five independent donors (n = 5).

3.3.3. Side population analysis

Side population analysis has proven to be a valuable tool to identify an immature population of HSPCs (229). Side population analysis was performed on eight-day expanded cells and the results are representative of pooled cells from three independent experiments from three different donors. Stem cells preferentially exclude unbound DNA binding dyes resulting in a lower fluorescent intensity population of cells. Although Hoechst is commonly used to identify the SP, VDC Violet is another commonly used cell-permeable DNA binding dye similar to Hoechst that can be excited by a violet or an ultraviolet (UV) laser, where Hoechst can only be excited by a UV laser (250,251).

VDC Violet was used in this study to identify the SP. The SP fraction (displayed as percentage) was identified as cells with higher dye efflux ability and therefore cells with low/negative VDC Violet fluorescence compared to the rest of the population (Figure 3.7b). The SP disappears

when cells are treated with Verapamil (Figure 3.7a), which binds to the ABC transporters on the surfaces of the cells and blocks their dye efflux ability. The SP has previously been identified in $CD34^+$ (252,253) and $CD34^-$ HSPCs (229,253,254) and for that reason the presence of a SP was investigated in both $CD34^+CD38^-$ and $CD34^-CD38^-$ HSPCs. The average mean fluorescence intensity (MFI) for $CD34^+$ and $CD34^-$ cells was 4627.03 and 111.87, respectively, whereas, the average MFI for the $CD38^+$ and $CD38^-$ cells was 1394.67 and 658.43, respectively.

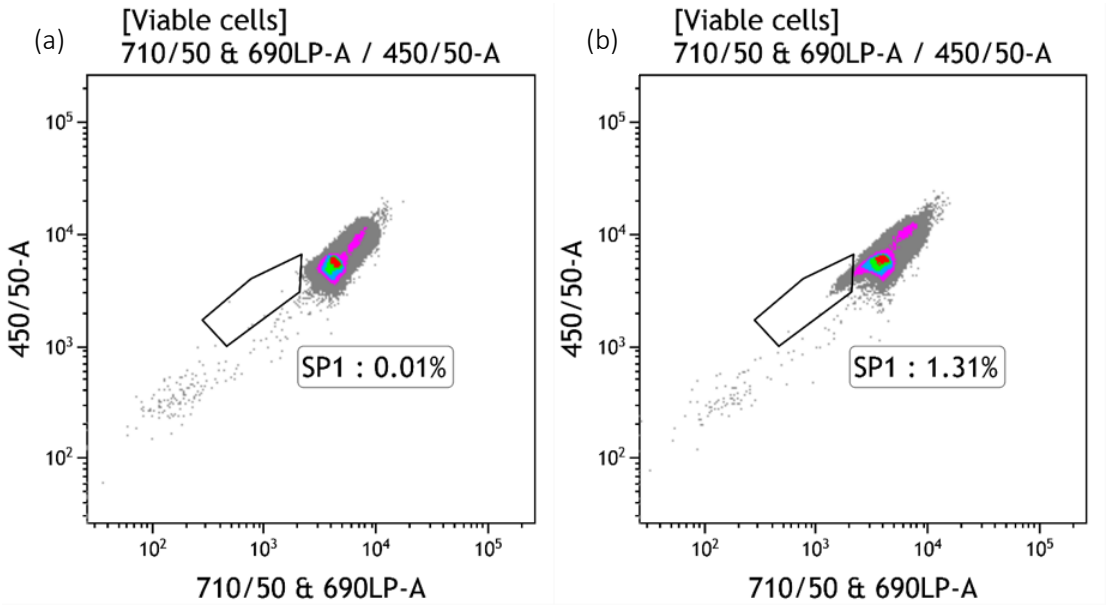


Figure 3.7. A schematic illustration of the SP. A representation of the (a) SP control, VDC Violet with verapamil and (b) VDC Violet without verapamil.

The mean percentages of SP cells are summarised in Figure 3.8 below. The differences observed between the different cytokine combinations were not statistically significant, but there seems to be a tendency of increased SP in Combinations 2 and 3 for both $CD34^+CD38^-$ and $CD34^-CD38^-$ cells. There also seems to be a tendency of increased SP in $CD34^-CD38^-$ cells in cultures without G-CSF, Combinations 1, 0.73 (± 0.53) and Combination 3, 1.13 (± 1.24), compared to the SP in $CD34^+CD38^-$ cells, Combination 1, 0.01 (± 0.24) and Combination 3, 0.80 (± 0.41). The lack of significance may be ascribed to the variability in the SP observed between different donors. Table 3.7 summarises the percentages SP observed in $CD34^+CD38^-$ and $CD34^-CD38^-$ populations after eight days.

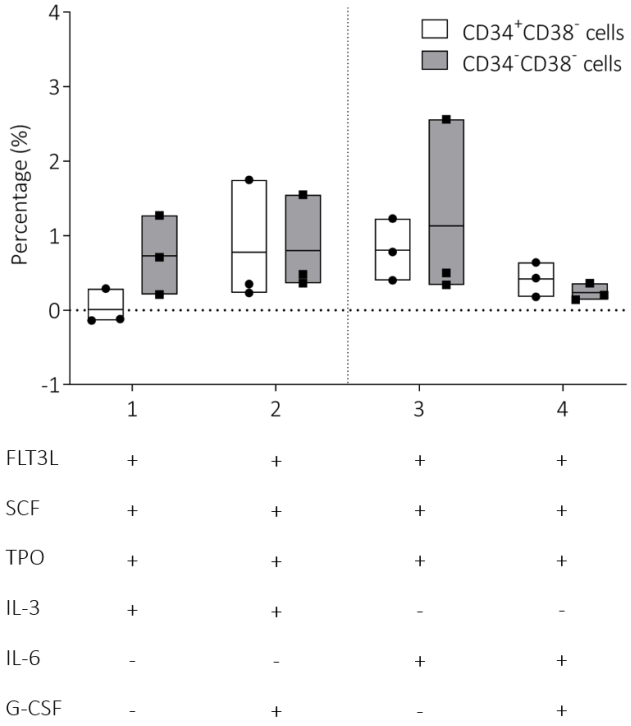


Figure 3.8. The effect of different cytokine combinations on the SP. The percentage side population observed in the $CD34^{+}CD38^{-}$ and $CD34^{-}CD38^{-}$ HSPC sub-populations after eight days in culture with different cytokine combinations. Bars are representative of three pooled technical repeats of three independent donors (n = 3).

Table 3.7. The mean percentage (SD) SP observed in $CD34^{+}CD38^{-}$ and $CD34^{-}CD38^{-}$ populations after eight days in culture with different cytokine combinations.

	VDC (%)	Verapamil control (%)	SP (%)
		$CD38^{-}CD34^{+}$ cells	
Combination 1	0.36 (± 0.19)	0.35 (± 0.10)	0.01 (± 0.24)
Combination 2	1.43 (± 0.93)	0.65 (± 0.31)	0.78 (± 0.85)
Combination 3	1.05 (± 0.47)	0.25 (± 0.13)	0.80 (± 0.42)
Combination 4	1.14 (± 0.38)	0.72 (± 0.53)	0.42 (± 0.23)
		$CD38^{-}CD34^{-}$ cells	
Combination 1	1.20 (± 0.71)	0.47 (± 0.20)	0.73 (± 0.53)
Combination 2	1.26 (± 0.70)	0.46 (± 0.14)	0.80 (± 0.66)
Combination 3	1.45 (± 1.05)	0.48 (± 0.26)	1.13 (± 1.24)
Combination 4	1.25 (± 0.53)	1.01 ± 0.52)	0.23 (± 0.11)

Data presented in this table is representative of three pooled technical repeats of three independent donors (n = 3).

3.4. DISCUSSION AND CONCLUSION

Cytokine signaling regulates proliferation, survival and differentiation of cells within the hematopoietic system. There is currently no standard cytokine combination used for the *ex vivo* expansion of HSPCs. The purpose of this study was to determine whether the addition of G-CSF to well-established cytokine combinations, FLT3L, SCF, TPO and IL-3/IL-6 would be beneficial for *ex vivo* culturing of HSPCs and whether these cytokine combinations have the potential to be used for *ex vivo* expansion of HSPCs for clinical applications in future.

FLT3L, SCF and TPO are the most potent and generally employed cytokines for the *ex vivo* expansion of HSPCs (255,256). A study by Bari *et al.* (257) has shown that the abovementioned cytokines are the most effective for HSPC maintenance *in vitro*, while IL-3, IL-6 and G-CSF promote rapid differentiation. It has previously been demonstrated that UCB-derived HSPCs express receptors for SCF (c-Kit) (258) and TPO (c-mpl) (259). The FLT3L cytokine has very little direct effect on HSPC proliferation, but rather acts synergistically when used in combination with other cytokines, such as interleukins and colony stimulating factors (260,261). FLT3L induces proliferation of quiescent human bone marrow and UCB $CD34^{+}CD38^{-}$ cells and maintains $CD34^{+}CD38^{+}$ progenitor cells *in vitro* (262). SCF and FLT3L have a synergistic effect on primitive HSCs, while TPO promotes survival and proliferation of HSPCs without bias to a particular lineage (263). Cells cultured in the presence of TPO have also shown to generate more colony forming cells compared to cells cultured without TPO (264).

The results from this study showed increased overall proliferation of both viable total cells and $CD34^{+}$ HSPCs expanded in FLT3L, SCF, TPO and IL-3 or IL-6 (Figure 3.2a). A further increase in proliferation of the above-mentioned cells was observed with the addition of G-CSF. Proliferation was highest in the presence of IL-3, with G-CSF (Figure 3.2a). The overall proliferation of cells expanded in the presence of IL-6 was noticeably lower than cells expanded with IL-3 (Figure 3.2a). Qui *et al.* (265) also observed an increase in the total number of nucleated and $CD34^{+}$ cells after seven-day expansion in FLT3L, SCF, TPO and IL-3 compared to cultures without IL-3. Likewise, Paek and Kim (266) showed increased expansion of $CD34^{+}$ cells in cultures supplemented with FLT3L, SCF, TPO, and IL-3 compared to controls (FLT3L, SCF and TPO). Increased expansion of cultures supplemented with IL-6 compared to the control was

also observed, but this was less than cultures supplemented with IL-3 (266). Increased $CD34^+$ HSPC numbers were also observed after an eight-day expansion, using G-CSF, in a study by Lam *et al.* (267).

The proportion of $CD34^+$ HSPCs decreases as cells expand and differentiate. The proportion of $CD34^+$ HSPCs was similar in cultures supplemented with FLT3L, SCF, TPO and IL-3, with and without G-CSF (Figure 3.3). Even though the proportion of $CD34^+$ cells was lower in cultures supplemented with IL-3 compared to IL-6, the increased number of total viable cells in cultures supplemented with IL-3, with and without G-CSF, resulted in increased numbers of $CD34^+$ cells after a seven-day expansion (Figure 3.2b). Qiu *et al.* (265) also observed lower CD34 percentages in cultures supplemented with IL-3 compared to cultures without IL-3. The number of $CD34^+$ HSPCs was lower in cultures supplemented with IL-6. However, the proportion of $CD34^+$ cells was higher, and the addition of G-CSF decreased the proportion slightly. Bordeaux-Rego *et al.* (92) also observed an increased proportion of $CD34^+$ and $CD133^+$ HSPCs after expansion with IL-6. The decreased proportion observed with the addition of G-CSF could suggest that G-CSF drives differentiation of $CD34^+$ HSPCs into mature lineages. Schuettpelz *et al.* (268) showed increased frequencies of murine HSPCs present after treatment with G-CSF. G-CSF exposed HSPCs also showed similar homing ability. However, secondary transplants revealed decreased repopulating ability of G-CSF exposed murine HSPCs.

The $CD34^-$ cell population was noticeably greater in cultures supplemented with IL-3, which further increased with the addition of G-CSF (Figure 3.2b). The greater number of $CD34^-$ cells present after a seven-day expansion could be the result of differentiation into mature lineages, since CD34 surface expression decreases as cells mature (269). Cultures supplemented with IL-6 had lower numbers of $CD34^-$ cells after seven days, which increased with the addition of G-CSF (Figure 3.2b).

An increase in the number of $CD34^+CD38^-$ cells was observed in cultures supplemented with G-CSF (Figure 3.4a). $CD34^+CD38^+$ cells remained relatively constant across cultures, with and without G-CSF. Qiu *et al.* (265) also reported an increase in long-term ($CD34^+CD38^-$) HSPCs

when cells were cultured with IL-3 compared to cultures without IL-3. The $Lin^{-}CD34^{-}CD38^{-}$ population and the $Lin^{-}CD34^{-}CD38^{-}CD133^{+}$ population were also increased in IL-3 cultures, with and without G-CSF (Figure 3.4b). The number of $Lin^{-}CD38^{-}CD34^{-}$ cells was much higher than $Lin^{-}CD38^{-}CD34^{+}$ cells in the presence of IL-3, with and without G-CSF. This suggests that the increased number of $CD34^{-}$ HSPCs ($Lin^{-}CD38^{-}CD34^{-}$ and $Lin^{-}CD34^{-}CD38^{-}CD133^{+}$) observed is due to the presence of IL-3. Cultures expanded with IL-6 showed a similar ratio of $Lin^{-}CD38^{-}CD34^{+}$ to $Lin^{-}CD38^{-}CD34^{-}$ HSPCs, with and without G-CSF. Schuettpelz *et al*. (268) found that treatment with G-CSF results in quiescent HSPCs, mainly in G0 of the cell cycle, which results in decreased repopulating ability of these cells. An increase in $CD34^{-}$ HSPCs was observed in cultures treated with G-CSF. G-CSF treatment usually results in increased $CD34^{+}$ HSPCs in mobilised peripheral blood (270); however, $CD34^{-}$ HSPCs are not enumerated prior to transplantation, since a positive marker for this population still needs to be identified (39).

FLT3L, SCF, TPO, IL-3, G-CSF and SCF, TPO, FLT3L, IL-6 supplemented cultures had a tendency to increase the proportion of SP cells. Side population cells were identified in both $CD34^{+}CD38^{-}$ and $CD34^{-}CD38^{-}$ cells, which is comparable to findings from Storms *et al.* (253), who identified SP cells in $CD34^{+}$ and $CD34^{-}$ HSPC isolated from UCB that had been lineage-depleted. The SP generally identifies an immature population of HSPCs enriched for long-term repopulating cells; however, the number of SP cells present does not predict time to engraftment (271).

FLT3L, SCF, TPO and IL-6 is a commonly used cytokine combination for the *ex vivo* expansion of HSPCs (77,112,242) and the IL-6 receptor (IL-6R) has been shown to be present on HSPCs from UCB (92). A study by Bordeaux-Rego *et al.* (92) demonstrated that cells cultured in IL-6 maintain an immature phenotype after expansion and also demonstrated through qPCR that cultures expanded in the presence of IL-6 maintained higher gene expression levels of *NANOG* and *SOX-2*. *NANOG* and *SOX-2* are generally expressed in pluripotent ESCs, but can also be present in somatic stem cells with increased expansion and differentiation potential (272). These two genes were also expressed at higher levels in $CD133^{+}$ cells on the day of isolation (92).

The role of IL-3 in the expansion of HSPCs has been controversial. There are some concerns that the addition of IL-3 may impair the ability of HSPCs to successfully engraft in transplant recipients. In the murine system, IL-3 has been indicated to be a negative regulator of HSPC function. A study by Yonemura *et al.* (273) showed that the presence of IL-3 reduced the number of colony forming cells and resulted in impaired engraftment of murine HSPCs. A similar study by Nitsche *et al.* (274) found that IL-3 treatment resulted in rapid total cell expansion *in vitro*, but the expanded cells possessed reduced *in vivo* hematopoietic regenerative potential. In primitive murine hematopoietic cells with lympho-myeloid potential, IL-3 impaired the self-renewal capacity of these cells and blocked the cytokine-stimulated generation of lymphoid precursors (275).

In contrast, Rossmanith *et al.* (276) found that including IL-3 into the FLT3L, SCF, and TPO cytokine cocktail, enhanced the *ex vivo* expansion of human UCB progenitor populations without impairing their engraftment potential. These authors showed that SCF, TPO, FLT3L and IL-3 enhanced the *ex vivo* expansion of primitive HSCs without abrogating the expression of the CXCR4 receptor, which is considered essential for the engraftment of transplanted cells (276). Mascarenhas *et al.* (263) found that IL-3 has the strongest effect on the most immature hematopoietic colonies. In combination, IL-3 and TPO have survival and proliferation effects on HSPCs and result in 90% repopulating ability of these cells into multiple lineages. SCF, TPO, FLT3L and IL-3 have the ability to initiate the proliferation of G0 $CD34^+$ human bone marrow cells, but only IL-3 sustained proliferation of SCF and FLT3L pre-stimulated G0 $CD34^+$ cells (256). IL-3 and G-CSF have been shown to act synergistically to increase cell proliferation (277).

A study by Bodine *et al.* (278) reported reduced repopulating ability of murine HSPCs after five days of G-CSF and SCF treatment. This was also observed by Schuettpelz and Link (279) who showed that treatment with G-CSF resulted in decreased repopulating ability of murine HSPCs in the bone marrow. Similarly, Winkler *et al*. (280) showed loss of repopulating ability after a five-day treatment with G-CSF. As HSPCs proliferate, they lose their regenerative capabilities, which might be one reason why cells treated with G-CSF have been shown to possess reduced repopulating abilities. Similarly, Li *et al.* (281) showed that G-CSF treatment increased the

number of $CD34^+$ cells during mobilisation, but inhibited the expression of endosteal cytokines, resulting consequently in major impairment of HSPC reconstitution potential in the mobilised bone marrow.

This study only shows data for *ex vivo* expansion of HSPCs with different cytokine combinations. One limitation of this study is that it did not include assays that predict *in vivo* outcomes. Numerous studies have shown that primitive phenotype after culture is not necessarily predictive of primitive function (235,264). Further studies will need to be conducted to assess the effect of each cytokine combination on the colony forming capacity of these cells as well as the effect on engraftment in an *in vivo* model.

Results for this study showed that G-CSF was able to increase HSPC populations after a seven-day expansion and that the greatest expansion was observed in cultures supplemented with FLT3L, SCF, TPO, IL-3 and G-CSF. Subsequent experiments for this study were therefore performed using this cytokine combination.

The cytokines and combinations of cytokines used would depend on individual applications. The expansion of HSPCs for HSCT aims to increase both short- and long-term repopulating cells to ultimately enable successful short- and long-term engraftment following HSCT. Expansion of HSPCs could additionally be applied to other non-stem cell-based therapies, such as cases of prolonged neutropenia, which increases the risk of infection. Some studies have been successful in expanding HSPCs into neutrophilic lineages to obtain functional and sufficient neutrophils for the treatment of neutropenias (282,283), since a large number of cells is required for this particular type of treatment. The process of expansion does not involve genetic manipulation and is GMP compliant with regard to manipulated products.

CHAPTER 4. THE EFFECT OF STEMREGENIN-1 ON THE GENE EXPRESSION OF UMBILICAL CORD BLOOD-DERIVED HEMATOPOIETIC STEM AND PROGENITOR CELLS

4.1. INTRODUCTION

Ex vivo expansion of UCB-derived HSPCs has long been an area of research interest. One of the main reasons for expanding HSPCs is the need to obtain a minimal number of cells required for transplantation success when UCB is used as a source of HSPC. The optimum $CD34^+$ cell dose has been defined as being between 2 and 5 x 10^6 cells/kg body weight (284). A single UCB unit seldom contains sufficient numbers of HSPCs required for transplantation into adults. Single unit transplantations are often associated with delayed hematopoietic reconstitution following HSCT, increasing the risk of infection and hospitalisation (76). These limitations can be overcome through the *ex vivo* expansion of UCB units. Concerns regarding the effect of expansion on cells have led to the use of non-expanded (non-manipulated) units in combination with *ex vivo* expanded (manipulated) UCB units for HSCT (77–79).

Ex vivo expansion has been achieved using different culture periods and culturing methods including the addition of different molecules such as cytokines (refer to Chapter 3) to promote cell proliferation. An unbiased screen searching for factors capable of expanding human HSPCs *ex vivo,* using a library of 100 000 small molecules, yielded a major breakthrough in expansion agent discovery (112). The expansion agent was the purine derivative SR1, which increased human UCB-derived $CD34^+$ HSPCs 50-fold during *ex vivo* expansion and resulted in a 17-fold increase in long-term engrafting HSPCs (112). A recent clinical study using SR1 demonstrated remarkable early neutrophil and platelet recovery and improved engraftment in patients who received UCB $CD34^+$ HSPCs expanded *ex vivo* with SR1, compared to recovery in recipients who received equal 'start' numbers of unexpanded $CD34^+$ cells from the same unit (78).

SR1 antagonises the AhR, a ligand-activated transcription factor localised in the cytosol, bound to various proteins when in a resting state (112). Upon ligand binding, AhR translocates into the nucleus where it binds to AhR nuclear translocator (ARNT). This dimerised complex activates transcription of genes such as *CYP1A1* and *CYP1B1* (285) (Figure 4.1).

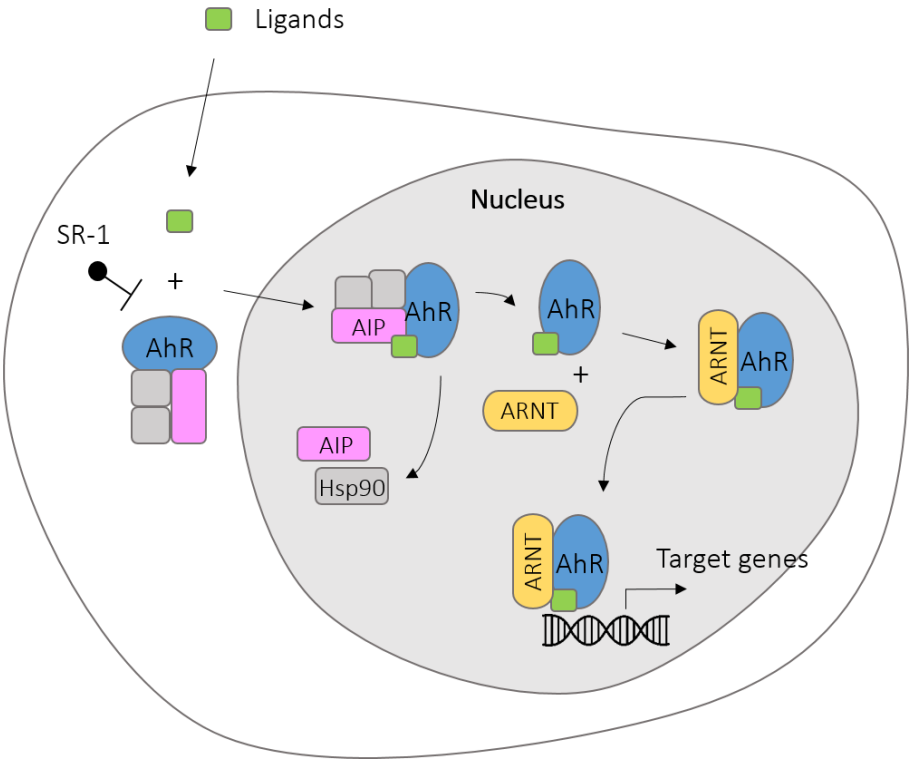


Figure 4.1. The AhR pathway. A schematic illustration of the mechanism of AhR activation and the antagonising effect of SR1. SR1, StemRegenin-1; AhR, Aryl hydrocarbon receptor; AIP, AhR interacting protein; Hsp90, Heat shock protein 90; ARNT, AhR nuclear translocator. (Figure was created by Juanita Mellet, adapted from (114)).

AhR has become an interesting area of research with a significant number of recent studies analysing its contribution in the functioning of the immune, hepatic, cardiovascular, vascular and reproductive systems (286). Recent reports have revealed two new exciting aspects in AhR biology; the findings imply that it plays a role in controlling cell differentiation, pluripotency and stemness. AhR has been shown to be involved in cellular differentiation *in vivo*, since activation of AhR with tetrachlorodibenzo-p-dioxin (TCDD) blocked the long-term self-renewal potential of murine HSPCs (286). In fact, it has been suggested that AhR could be involved in the balance between differentiation and pluripotency (286).

Messenger RNA (mRNA) and protein levels of AhR are abundant in HSPCs (287). It has been suggested that AhR is important for the long-term maintenance and function of HSPCs, and regulates gene expression of long-term HSPCs (288). HSPCs from mice that lack AhR were unable to repopulate the bone marrow of transplant recipients following secondary and tertiary transplantation, even though the number of progenitors was not directly affected

(288). Reports have indicated that AhR activation might lead to changes in gene expression that relate to HSPC migration and homing (288). The mRNA and protein levels of AhR in HSPCs are significantly reduced when these cells are cultured in the presence of SR1 (287). Treatment of HSPCs with SR1 has been shown to increase the number of $CD34^+$ cells able to engraft in the bone marrow of recipient mice (112). However, the exact role of AhR in HSPC biology is still unknown, and the effects of SR1-induced expansion on HSPCs still need to be fully determined.

The purpose of this study was to determine the effect of SR1 on the transcriptome of expanded $CD34^+$ and $CD34^-$ HSPCs from UCB, and to establish to what extent the transcriptome of the seven-day expanded cells differs from non-expanded $CD34^+$ cells on the day of isolation. This chapter describes the data obtained from expanding $CD34^+$ HSPCs in the presence of different concentrations of SR1 to determine the optimal concentration for expansion and subsequent gene expression analysis. Gene expression analysis performed on RNA isolated from non-expanded UCB $CD34^+$ HSPCs and seven-day expanded $CD34^+$ and $CD34^-$ cells (SR1-treated and non-treated) is also discussed in this chapter.

4.2. MATERIALS AND METHODS

4.2.1. Sample collections

Informed consent and sample collections were performed as discussed in Chapter 2, Section 2.2.2. Ten independent UCB units were collected for this part of the study; Appendix F shows the informed consent document that was used for this part of the study. Four UCB units were used to determine the optimal expansion concentration of SR1 on locally sourced HSPCs, while six UCB units (two units pooled at a time) were used to determine the effects of seven-day expansion on the transcriptome of SR1-treated and non-treated HSPCs. HIV testing was performed as described in Chapter 3, Section 3.2.3.

4.2.2. Enrichment of $CD34^+$ HSPCs

Enrichment of $CD34^+$ HSPCs was performed as previously described in Chapter 3, Section 3.2.4. This particular part of the study made use of both single and pooled UCB samples for the different experiments. Single (not pooled) UCB samples were used for the SR1 concentration

optimisation, while two pooled UCB units (repeated three times) were used for the gene expression expansion experiments. The reason for sample pooling was to be able to sort enough cells to obtain good quality RNA to proceed with the microarray gene expression experiments.

4.2.3. Flow cytometer: Counting protocol

Stem-Kit reagents from Beckman Coulter (Miami, Florida, USA) were used to determine viability and the absolute number of $CD34^+$ cells in each sample using the Gallios flow cytometer and counting protocol as described in Chapter 2. All post-acquisition analyses were performed using Beckman Coulter Kaluza Analysis Software (version 2.1) (Beckman Coulter, California, USA). The Gallios flow cytometer filter configurations are shown in Chapter 2, Table 2.2.

4.2.4. FACS sorting

Stem-Kit reagents (CD45 FITC, CD34 PE and 7AAD) were added to the pooled (Day 0) or cultured (seven-day expanded) samples based on the absolute number of cells present (approximately 5 µL per 10 x 10^6 viable cells). Once the antibodies had been added, samples were briefly mixed and incubated for 30 min at room temperature. After incubation, stained cells were washed using TP buffer and the supernatant was aspirated. The cell pellet was resuspended in 5 – 10 mL TP buffer and filtered using a 30 µM tube-top filter into a flow tube (Falcon, Corning, USA) for sorting. Sample sorting and purity check were performed as described in Chapter 3, Section 3.2.3.3. The $CD34^+$ region used to sort the cells was set using the Stem-Kit IsoClonic control. Viable $CD34^+$ cells were sorted into wells of a 24-well plate containing StemSpan ACF (1 x 10^4 cells/well) or directly into lysis buffer (Qiagen, Germany) (1 x 10^5 cells/well) for RNA extractions.

4.2.5. Culture conditions

FACS-purified $CD34^+$ cells (1 x 10^4 cells/well) were cultured in 1 mL serum-free StemSpan ACF medium (Stem Cell Technologies, Canada) and 2% pen/strep. Based on previous results (Chapter 3), the following growth factors were included, each at 100 ng/mL: SCF, TPO, FLT3L, G-CSF and IL-3. SR1 (Stem Cell Technologies, Canada) was dissolved in dimethyl sulfoxide

(DMSO, Sigma-Aldrich, USA) and further dilutions were made using StemSpan ACF medium. The following concentrations were tested to determine the optimal SR1 concentration for HSPC expansion: 0.25, 0.5, 0.75 and 1 µM. A vehicle control containing the highest percentage of DMSO (0.01%), used to dissolve SR1, was included. Based on the outcome, gene expression experiments were performed using 1 µM SR1. A vehicle control (SR1 untreated cells) was also included for this series of experiments. The concentrations indicated are the final concentrations of the respective growth factors and SR1 present in the culture wells. Sterile PBS was added to unused wells to maintain humidity. The cultures were maintained for seven days at 37°C and 5% CO_2 in a humidified atmosphere.

4.2.6. Viability, proliferation, $CD34^+$ cell numbers/percentages and HSPC-associated immunophenotype

4.2.6.1. SR1 concentration optimisation

HSPC expansion was examined after seven days by measuring viability, proliferation, $CD34^+$ cell numbers/percentages and HSPC phenotype. Phenotypic analysis of HSPCs was performed as described in Chapter 3, Section 3.2.9.

4.2.6.2. Immunophenotypic profiles of expanded and non-expanded cells

The immunophenotypic profiles of non-expanded HSPCs and seven-day expanded SR1-treated (1 µM) and non-treated cells were assessed on the day of RNA isolation. Immunophenotypic markers were assessed using the FACSAria Fusion. The following anti-human monoclonal antibodies were used: Lineage cocktail FITC, CD34 PE-Cy7, CD38 APC-Cy7, CD45 BV421, CD133/2 PE (Myltenyi Biotec, Germany). A viability dye, 7AAD, was also included and allowed for the exclusion of non-viable cells. Monoclonal antibodies (3 – 5 µL) were added to a flow tube containing 100 µL of the cell suspension. Non-expanded (on the day of isolation) and expanded (SR1-treated and non-treated) cells were phenotyped. Single-colour antibody staining controls were used to correct for spectral overlap between the fluorochromes (colour compensation).

Treated and non-treated HSPCs were analysed using the following gating strategy. First, the HSPC population ("Cells" region) was identified using a FS vs. SS plot. A 7AAD vs. FS plot, gated on "Cells" was used to identify viable cells ("Viable" region). All the following histogram plots were gated on the "Viable" region. Lin^+ and Lin^- cells were identified using a lineage cocktail FITC [Excitation: 488 nm; Emission: 520 nm]. $CD133^+$ cells were identified using CD133 PE [Excitation: 496 nm; Emission: 578 nm]. $CD34^+$ and $CD34^-$ cells were identified using CD34 PE-Cy7 [Excitation: 488/532/561 nm; Emission: 575/780 nm]. $CD38^+$ and $CD38^-$ cells were identified using CD38 APC-Cy7 [Excitation: 633 nm; Emission: 776 nm]. $CD45^+$ and $CD45^-$ cells were identified using CD45 BV421 [Excitation: 405 nm; Emission: 421 nm]. The negative/positive detection boundaries were established using peripheral blood-derived leukocytes that were stained with the same panel of monoclonal antibodies.

4.2.7. Side population analysis

Side population analysis was performed as described in Chapter 3, Section 3.2.10, after eight days in culture. The sequential gating strategy that was used to identify the SP cells are shown in Figure 3.2.

4.2.8. RNA extractions

The Qiagen RNeasy Micro Plus kit (Qiagen, Germany) was used to perform RNA extractions from non-expanded $CD34^+$ cells on the day of isolation (Day 0) and $CD34^+$ and $CD34^-$ cells after seven-day expansion (Day 7 SR1-treated and non-treated controls). In preparation for RNA extraction, β-Mercaptoethanol (Sigma-Aldrich, USA) was added to the supplied RLT buffer (10 µL per 1 mL RLT buffer). Cells (1×10^5 total cells) were sorted directly into 300 µL lysis buffer (β-Mercaptoethanol and RLT buffer). Cells were kept at 4°C throughout the sorting process. Once the cells had been sorted, the lysates were transferred to a genomic DNA (gDNA) Eliminator column in a 2 mL collection tube and centrifuged at 8000 x g for 30 seconds. This removed genomic DNA from the sample. The gDNA column was discarded after centrifugation. 300 µL of 70% ethanol was added to the flow-through and mixed by pipetting. The sample was transferred to an RNeasy MinElute spin column in a 2 mL collection tube and centrifuged at 8000 x g for 15 seconds. 700 µL RW1 buffer was added to the MinElute column and centrifuged at 8000 x g for 15 seconds. 500 µL of buffer RPE was added to the MinElute column and

centrifuged at 8000 x g for 15 seconds. 500 µL of 80% ethanol was added to the MinElute column and centrifuged at 8000 x g for 2 min to wash the membrane of the spin column. The RNeasy MinElute column was transferred to a new collection tube and centrifuged at maximum speed for 5 min to dry the membrane of the spin column. The flow through was discarded after each step. The MinElute column was placed in a new 1.5 mL collection tube and 11 – 14 µL of RNase-free water was added directly to the center of the membrane and centrifuged at maximum speed for 1 min to elute the RNA. The latter step was performed twice. RNA sample concentrations were measured using the Nanodrop ND 1000 spectrophotometer (Thermo Fisher Scientific, USA) and stored at -80^{o}C until further use.

Freshly isolated HSPCs generally yielded small amounts of RNA. Therefore, two identical RNA extractions from two aliquots of cells (1 x 10^5 cells/aliquot) were performed on Day 0 and concentrated by vacuum centrifugation. The two Day 0 RNA extractions were then pooled and vacuum centrifuged using a SpeedVac SVC-100 (Savant Instruments, USA) at high speed for 10 min until the pellet was dry. The pellet was resuspended in 7 µL of RNase-free water and allowed to dissolve for 10 min at room temperature.

4.2.9. RNA integrity and quality

RNA integrity and quality were assessed using the TapeStation® 2200 (Agilent Tehnologies, USA), RNA ScreenTape® and Sample buffer kit (Agilent Technologies, USA). Samples were prepared by adding 5 µL of Sample buffer to 1 µL of the thawed RNA sample and vortexing for 1 min. Samples were then heated to 72°C for 3 min in a Thermal Cycler (Gene Amp®, PCR System 9700, Applied Biosystems, Life Technologies, Thermo Fisher Scientific, USA), and placed on ice for 2 min before being loaded into the TapeStation. RNA integrity (RIN) was analysed using the 2200 TapeStation software.

4.2.10. Microarray gene expression

RNA (50 ng) isolated from non-expanded (Day 0) and seven-day expanded (Day 7, SR1-treated and non-treated controls) cells was used for first- and second-strand complementary DNA (cDNA) synthesis. This was followed by the synthesis and amplification of complementary RNA (cRNA) by *in vitro* transcription using an Affymetrix GeneChip® WT Plus Reagent Kit (Affymetrix,

USA) according to the manufacturer's instructions. Magnetic purification beads supplied with the kit were used to purify the cRNA. The concentration of the purified cRNA was measured using the NanoDrop ND 1000 and 15 µg was used to synthesise the second cycle single-stranded cDNA (ss-cDNA). A second magnetic bead purification step was performed to purify the ss-cDNA and 5.5 µg of the ss-cDNA was enzymatically fragmented and labelled. The ss-cDNA was analysed using the TapeStation 2200 to ensure that the transcript size was correct. The fragmented and labelled ss-cDNA was added to the hybridisation cocktail as described in the Affymetrix GeneChip® WT PLUS Reagent Kit manual. Hybridisation was performed using the Hybridisation Wash and Stain Kit (Affymetrix, USA) following manufacturer's instructions. The hybridisation cocktail was hybridised to the Affymetrix GeneChip® Human Gene 2.0 ST arrays (Affymetrix, USA). The GeneChips were placed in the Affymetrix GeneChip® Hybridisation Oven-645 (Affymetrix, USA) rotating at 60 rpm at 45^{o}C for 17 hours. After incubation, the chips were washed and stained in an Affymetrix GeneChip® Fluidics Station-450Dx (Affymetrix, USA) before scanning using an Affymetrix GeneChip® Scanner-7G (Affymetrix, USA). The Scanner generates Affymetrix CEL files which contain intensity values for all probes present on the scanned chips. These files were used for further analysis.

4.2.11. Statistical analysis

Experiments were performed in triplicate (three technical repeats per condition) using three separate UCB samples. Experimental data are represented as mean (SD). All statistical analyses were performed as described in Chapter 2, Section 2.2.7.

4.2.12. Data analysis

The CEL files were imported into Affymetrix Transcriptome Analysis Console™ (TAC) Software 4.0 (Affymetrix, USA) to perform background correction, summarisation, normalisation and calculation of expression values with resulting CHP output files. The TAC software included a principal component analysis (PCA) tool to assess the relatedness of samples from the three biological replicates. Fold change was calculated using the Affymetrix TAC software. Genes with a fold change ≥ 2.5 and ≤ -2.5 and (p-value < 0.05) were considered to be differentially expressed. Fold change of a gene represents the change in gene expression based on the signal

measured between the SR1-treated and non-treated (control) cells, and between the seven-day expanded and non-expanded $CD34^+$ cells. The PANTHER Gene Ontology (GO) enrichment analysis tool (http://geneontology.org) was used to classify differentially expressed genes into biological processes/pathways (289).

4.3. RESULTS

4.3.1. SR1 concentration optimisation

Freshly isolated HSPCs were treated with four different concentrations of SR1 and a vehicle control for seven days: vehicle control (0.01% DMSO), 0.25 µM, 0.5 µM, 0.75 µM and 1 µM SR1. The particulars of each isolation are shown in Table 4.1. After the seven-day expansion, viability, proliferation, $CD34^+$ cell numbers/percentages and SP were assessed.

Table 4.1. UCB units collected to determine the optimal SR1 concentration for HSPC expansion.

Sample ID	Volume (mL)	Viability (%)	Total leukocytes	$CD34^+CD45^{dim}$ cells	$CD34^+$ (%)	Sort purity (%)
CB050917	70	85.4	16.7×10^6	5.3×10^5	3.23	97.4
CB060917	100	97.6	71.9×10^6	3.2×10^6	4.62	95.0
CB041017	90	97.1	28.3×10^6	4.9×10^5	1.79	94.3
CB070318	100	96.7	167.7×10^6	1.15×10^6	0.69	94.6

4.3.1.1. Viability, proliferation and $CD34^+$ HSPC numbers/percentages

The addition of reagents (especially those dissolved in DMSO) to standard media for the *ex vivo* expansion of HSPCs could influence the viability of the cells. It was therefore important to show that the highest percentage of DMSO added did not affect cell viability. The mean absolute viability was high in the vehicle control (95% (± 2.6%)) and all SR1 concentrations, 0.25 µM (95% (± 1.5%)), 0.5 µM (95% (± 2.1)) 0.75 µM (95% (± 2.6%)) and 1 µM (94% (± 2.6%)) (Figure 4.2).

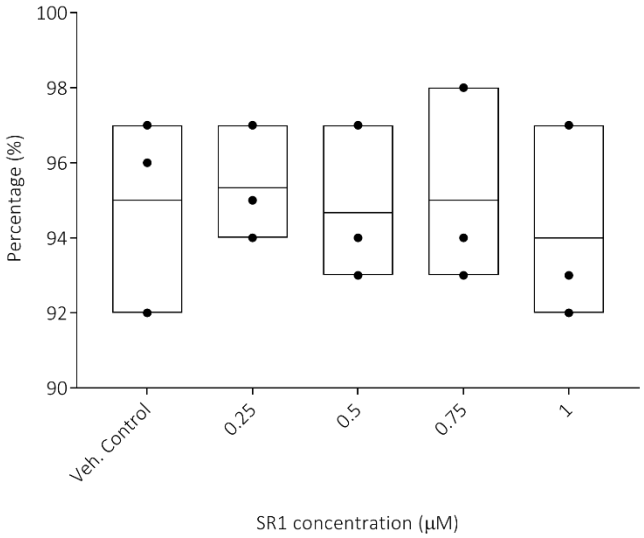


Figure 4.2. Absolute viability after seven-day expansion with SR1. The mean percentage (SD) viable cells at the various concentrations of SR1 after the seven-day expansion. Bars are representative of three technical repeats of three independent donors (n = 3).

Although the results were not statistically significant, the mean $CD34^+$ HSPC percentages increased with increasing concentrations of SR1 (Figure 4.3a and b). The vehicle control had the lowest percentage (21.5% (± 3.9%)) and total number of $CD34^+$ HSPCs (21.0 (± 7.4%)), while the highest concentration of SR1 (1 μM) had the highest percentage (49.3% (± 5.7%)) and number of $CD34^+$ HSPCs (47.2 (± 21.0%)) after seven-day expansion. The effect of different SR1 concentrations on cell proliferation was demonstrated as fold change, described as the increase in the number of cells observed on Day 7 relative to the number of cells seeded on Day 0. Percentages are sometimes misleading in cultures where the total number of cells is lower. However, since the total number of viable cells were similar between the different conditions, the average fold change of $CD34^+$ cells (Figure 4.4) was similar to the $CD34^+$ percentages (Figure 4.3).

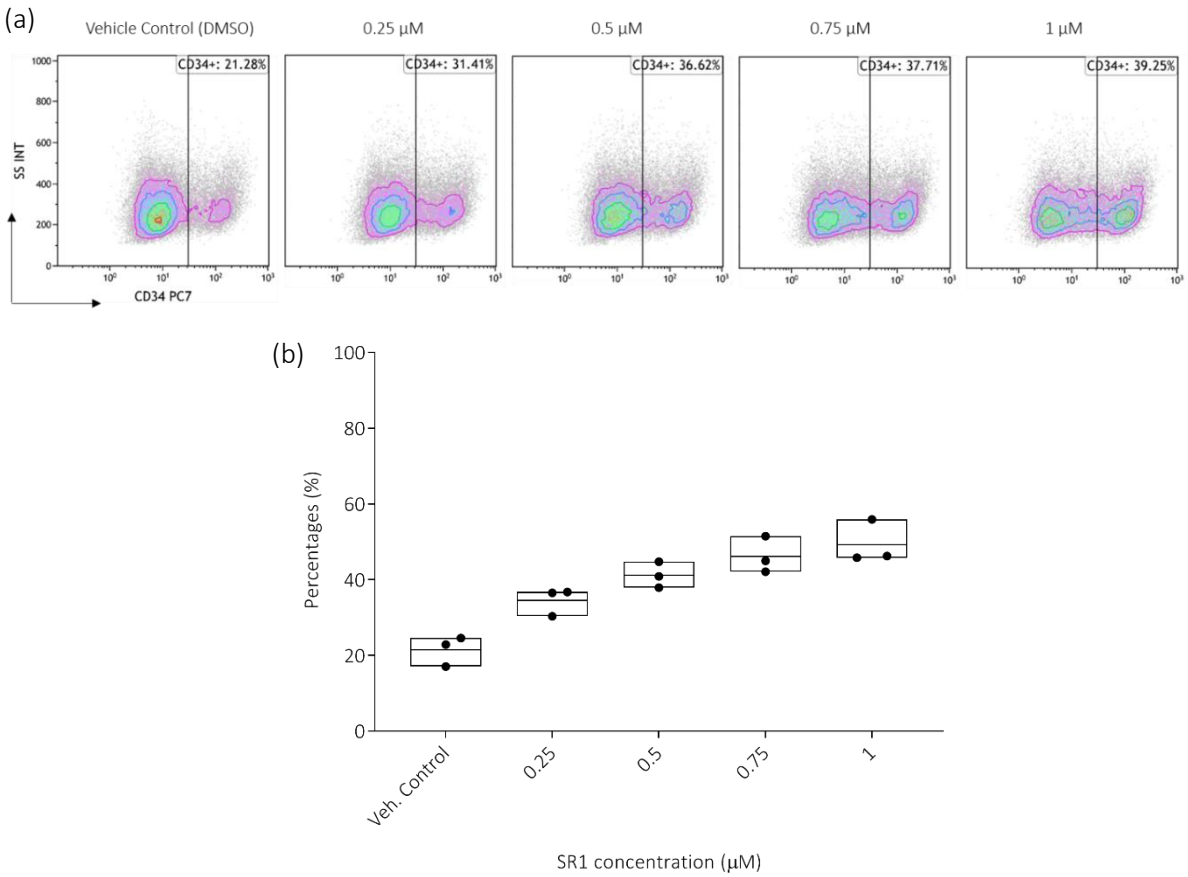


Figure 4.3. The effect of SR1 on CD34⁺ HSPC percentages. (a) A graphical illustration of flow cytometry density plots and (b) a bar graph indicating the increased percentage (mean ± SD) $CD34^+$ HSPCs as the concentration of SR1 increased. Bars are representative of three technical repeats of three independent donors (n = 3).

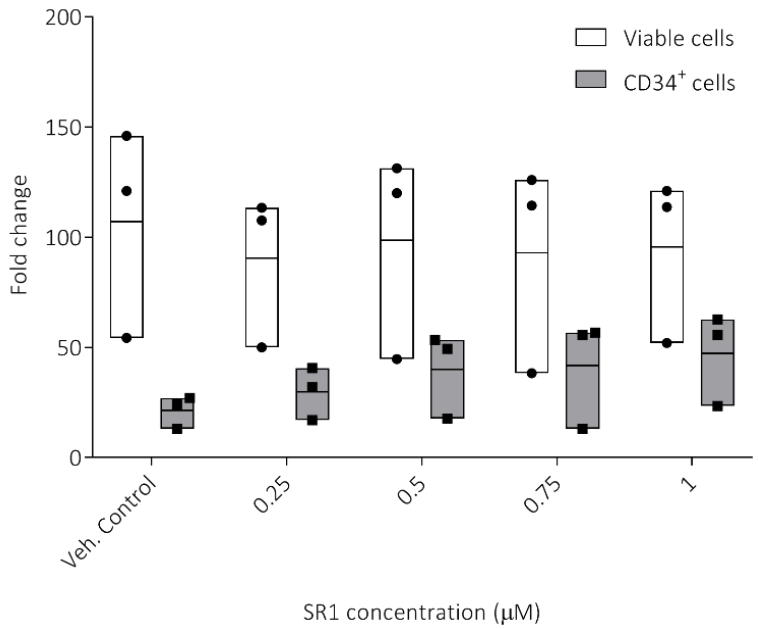


Figure 4.4. The effect of SR1 on cell proliferation. The mean fold increase (SD) of total number of viable cells and $CD34^+$ HSPCs after seven-day expansion in different concentrations of SR1. Bars are representative of three technical repeats of three independent donors (n = 3).

4.3.1.2. HSPC-associated immunophenotype

Due to the heterogeneity of the $CD34^+$ population, immunophenotyping was performed by combining CD34 with various other markers including the differentiation marker CD38, which allows the identification of cells enriched for short- ($CD34^+CD38^+$) and long-term ($CD34^+CD38^-$) repopulating potential (232). Both $CD34^+CD38^-$ and $CD34^+CD38^+$ cell numbers increased with increasing SR1 concentrations (Figure 4.5). The lowest total number of $CD34^+CD38^-$ (12.6 (± 4.9)) and $CD34^+CD38^+$ (9.0 (± 4.3) cells were present in the vehicle control, while the highest total number of $CD34^+CD38^-$ (25.0 (± 8.9)) and $CD34^+CD38^+$ (15.0 (± 8.5)) cells were present at the highest concentration of SR1 (1 μM). Thus, both the number of short- and long-term HSPCs increased with increased SR1 concentration. Overall, the $CD34^+CD38^-$ cell numbers were higher than $CD34^+CD38^+$ cell numbers in all cultures, which means that the fraction of cells containing the long-term HSPCs was greater than the fraction containing the short-term HSPCs when expanded with SR1.

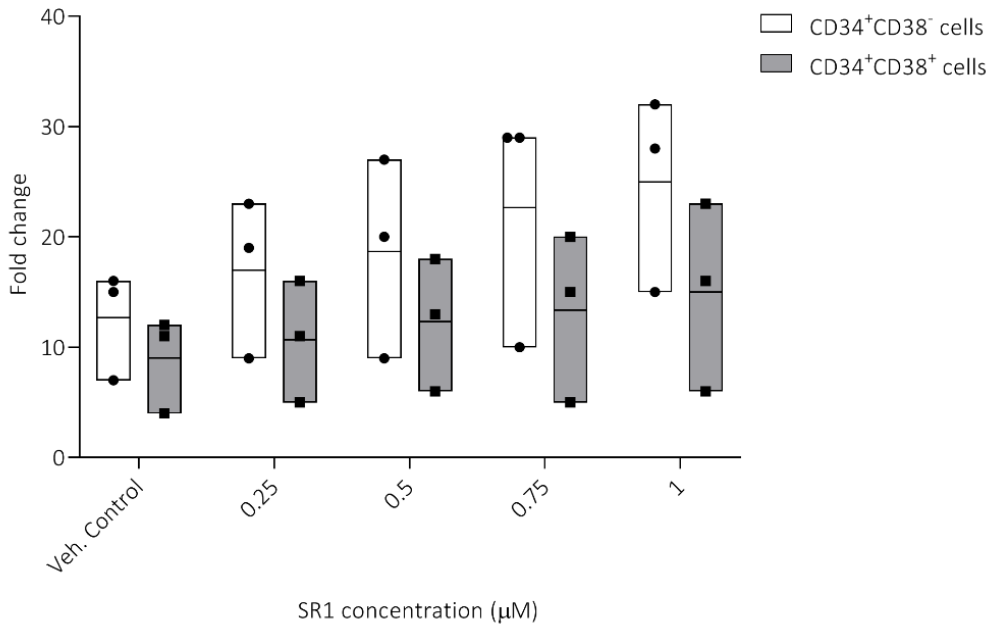


Figure 4.5. The effect of SR1 on the proliferation of $CD34^+CD38^-$ and $CD34^+CD38^+$ HSPCs. The mean total number (SD) of $CD34^+CD38^-$ and $CD34^+CD38^+$ HSPCs present after seven-day expansion in different concentrations of SR1. Bars are representative of two technical repeats of three independent donors (n = 3).

Although not statistically significant, the $Lin^-CD34^+CD38^-$ population increased, while the $Lin^-CD34^-CD38^-$ population decreased as the concentration of SR1 was increased (Figure 4.6a). The $Lin^-CD34^+CD38^-$ was highest, 21.3 (± 8.3) and the $Lin^-CD34^-CD38^-$ population was lowest, 28.0 (± 13.7) in 1 μM SR1 compared to the $Lin^-CD34^+CD38^-$, 10.3 (± 3.8) and $Lin^-CD34^-CD38^-$, 43.0 (± 17.3) populations in the vehicle control. The same trend was observed for the $Lin^-CD34^-CD38^-CD133^+$ population, which also decreased as the concentration of SR1 increased (Figure 4.6b). Again, the $Lin^-CD34^+CD38^-CD133^+$ population was highest, 16.7 (± 6.8) and the $Lin^-CD34^-CD38^-CD133^+$ population lowest, 2.7 (± 3.8) in 1 μM SR1 compared to the $Lin^-CD34^+CD38^-CD133^+$, 9.3 (± 2.9) and $Lin^-CD34^-CD38^-CD133^+$, 6.7 (± 4.7) populations in the control.

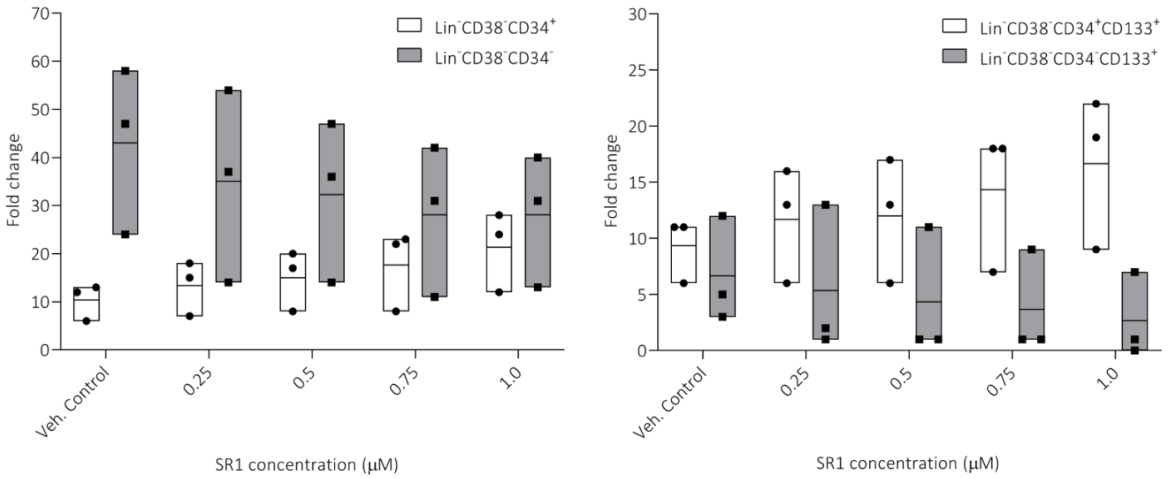


Figure 4.6. The effect of SR1 on the proliferation of CD34⁺ and CD34⁻ HSPC populations. The total number (mean fold increase) of (a) $Lin^-CD38^-CD34^+$ and $Lin^-CD38^-CD34^-$ and (b) $Lin^-CD38^-CD34^+CD133^+$ and $Lin^-CD38^-CD34^-CD133^+$ HSPCs after seven-day expansion with different concentrations of SR1. Bars are representative of two technical repeats of three independent donors (n = 3).

$CD34^+$ and $CD34^-$ HSPCs can be identified using different co-expression profiles of HSPC-associated markers. The average fold change of different $CD34^+$ HSPC populations, identified using different HSPC-associated marker permutations, was assessed after a seven-day expansion at different concentrations of SR1 (Table 4.2). The $CD34^+$ population of cells was subdivided into cells positive and negative for lineage markers after seven-day expansion. We noticed a 1:1 ratio of Lin^-CD34^+ and Lin^+CD34^+ cells present after seven-day expansion, with the cell numbers increasing slightly as the concentration of SR1 increased. The majority of the Lin^-CD34^+ cells were $CD38^-$, while the majority of the Lin^+CD34^+ cells tend to express CD38, which suggests that the latter cells are more mature HSPCs. Although previous studies have suggested that CD133 is expressed on more primitive cells (44), the results from this study show that CD133 is also expressed on later progenitors that have already acquired the expression of lineage markers and CD38 after a seven-day expansion. Although the number of cells after a seven-day expansion remained fairly constant with different concentrations of SR1, all populations positive for CD34 (irrespective of the presence of other phenotypic markers) increased as SR1 concentrations increased.

The average fold change of different $CD34^-$ HSPC populations, identified using different HSPC-associated marker permutations, were assessed after a seven-day expansion in different concentrations of SR1 (Table 4.3). The $CD34^-$ population of cells were also subdivided into cells positive and negative for lineage markers after a seven-day expansion. All populations negative for CD34 (irrespective of the presence of other phenotypic markers) decreased as the SR1 concentration increased. This gradual decrease with increasing SR1 concentrations was observed for a more primitive population of HSPCs ($Lin^-CD38^-CD34^-$) as well as populations considered to be more mature and already lineage committed ($Lin^+CD34^-CD38^+$).

Table 4.2. The effect of SR1 on the proliferation of $CD34^+$ HSPC populations. The mean fold increase (SD) of viable and total $CD34^+$ HSPC populations, identified using different HSPC-associated marker permutations, after a seven-day expansion in different concentrations of SR1.

	Condition				
	Vehicle control	0.25 µM	0.5 µM	0.75 µM	1 µM
Viable cells	112 (± 44)	99 (± 46)	99 (± 46)	97 (± 47)	104 (± 45)
$CD34^+$	23 (± 9)	28 (± 12)	32 (± 15)	37 (± 18)	45 (± 21)
Lin^-CD34^+	11 (± 3)	14 (± 5)	15 (± 6)	18 (± 8)	22 (± 9)
$Lin^-CD34^+CD38^-$	11 (± 4)	14 (± 6)	15 (± 6)	18 (± 8)	21 (± 9)
$Lin^-CD34^+CD38^-CD133^+$	9 (± 3)	12 (± 5)	12 (± 5)	14 (± 7)	17 (± 7)
$Lin^-CD34^+CD38^+$	0	0	0	0	0
$Lin^-CD34^+CD38^+CD133^+$	0	0	0	0	0
Lin^+CD34^+	11 (± 5)	13 (± 7)	16 (± 9)	18 (± 10)	22 (± 12)
$Lin^+CD34^+CD38^-$	1 (± 1)	2 (± 2)	3 (± 3)	4 (± 3)	6 (± 3)
$Lin^+CD34^+CD38^-CD133^+$	0	1 (± 1)	1 (± 2)	2 (± 2)	3 (± 2)
$Lin^+CD34^+CD38^+$	9 (± 4)	10 (± 5)	12 (± 6)	13 (± 8)	15 (± 9)
$Lin^+CD34^+CD38^+CD133^+$	8 (± 4)	9 (± 5)	10 (± 5)	11 (± 6)	12 (± 7)

Data presented in this table is representative of two technical repeats of three independent donors (n = 3).

Table 4.3. The effect of SR1 on proliferation of $CD34^-$ populations. The mean fold increase (SD) of viable and total number of $CD34^-$ HSPC populations, identified using different HSPC-associated marker permutations, after a seven-day expansion in different concentrations of SR1.

	Condition				
	Vehicle control	0.25 µM	0.5 µM	0.75 µM	1 µM
Viable cells	112 (± 44)	99 (± 46)	99 (± 46)	97 (± 47)	104 (± 45)
$CD34^-$	87 (± 36)	70 (± 38)	67 (± 34)	59 (± 31)	59 (± 27)
Lin^-CD34^-	44 (± 17)	36 (± 20)	33 (± 17)	29 (± 16)	29 (± 13)
$Lin^-CD34^-CD38^-$	43 (± 17)	35 (± 20)	32 (± 17)	28 (± 16)	28 (± 14)
$Lin^-CD34^-CD38^-CD133^+$	6 (± 5)	5 (± 7)	5 (± 6)	4 (± 5)	3 (± 4)
$Lin^-CD34^-CD38^+$	0	0	0	0	0
$Lin^-CD34^-CD38^+CD133^+$	0	0	0	0	0
Lin^+CD34^-	42 (± 18)	33 (± 17)	33 (± 17)	29 (± 15)	28 (± 14)
$Lin^+CD34^-CD38^-$	14 (± 8)	10 (± 5)	10 (± 5)	9 (± 5)	10 (± 5)
$Lin^+CD34^-CD38^-CD133^+$	2 (± 3)	0	0	0	0
$Lin^+CD34^-CD38^+$	28 (± 11)	22 (± 12)	22 (± 10)	19 (± 10)	18 (± 9)
$Lin^+CD34^-CD38^+CD133^+$	10 (± 12)	6 (± 7)	6 (± 6)	5 (± 5)	4 (± 4)

Data presented in this table is representative of two technical repeats for each of three independent donors (n = 3).

4.3.1.3. Side population analysis

The side population analysis was performed on eight-day expanded cells; the results are representative of pooled cells from three independent experiments from two different donors. As mentioned in Chapter 3, Section 3.3.3, the SP has previously been identified in $CD34^+$ (252,253) and $CD34^-$ HSPCs (229,253,254) and for this reason both $CD34^+CD38^-$ and $CD34^-CD38^-$ cells were included when assessing the SP.

The mean percentage SP in the vehicle control was similar for the $CD34^+CD38^-$, 1.03 (± 0.11) and $CD34^-CD38^-$, 1.04 (± 0.20) cells (Figure 4.7). Although not statistically significant when compared to the vehicle control, the highest mean percentage of SP cells was identified for both $CD34^+CD38^-$, 1.40 (± 1.34) and $CD34^-CD38^-$, 1.91 (± 0.95) cells using a 0.25 µM SR1 concentration. These SP percentages decreased as the concentration of SR1 increased.

SP percentages were higher in $CD34^{-}CD38^{-}$ cells compared to $CD34^{+}CD38^{-}$ cells for all SR1 concentrations. Table 4.4 shows the percentage SP (for two independent donors) observed in $CD34^{+}CD38^{-}$ and $CD34^{-}CD38^{-}$ populations after expansion with different concentrations of SR1 for eight days.

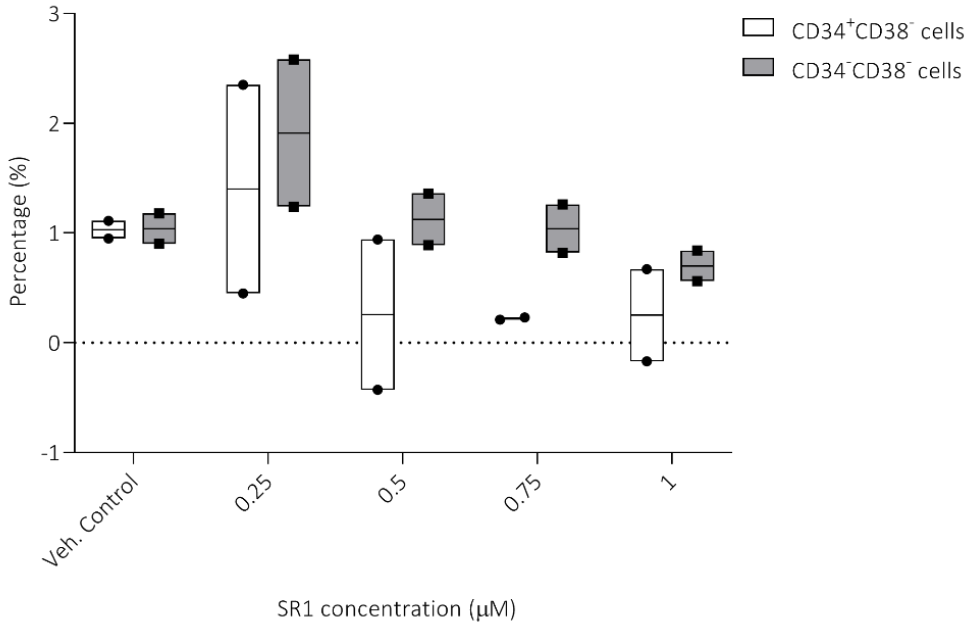


Figure 4.7. The effect of SR1 on SP phenotype. The percentage SP cells observed in the $CD34^{+}CD38^{-}$ and $CD34^{-}CD38^{-}$ populations after an eight-day expansion with different concentrations of SR1. Bars are representative of three pooled technical repeats of two independent donors (n = 2).

Table 4.4. Percentage SP cells in $CD34^+CD38^-$ and $CD34^-CD38^-$ cells. The percentage side population observed in $CD34^+CD38^-$ and $CD34^-CD38^-$ cells after expansion with different SR1 concentrations for eight days in culture.

	VDC (%)	Verapamil control (%)	SP (%)
		$CD38^-CD34^+$ cells	
Vehicle control	2.06 (± 0.33)	1.03 (± 0.44)	1.03 (± 0.11)
0.25 µM SR1	1.88 (± 1.23)	0.48 (± 0.11)	1.40 (± 1.34)
0.5 µM SR1	1.18 (± 0.89)	0.93 (± 0.08)	0.26 (± 0.97)
0.75 µM SR1	0.54 (± 0.02)	0.32 (± 0.01)	0.22 (± 0.01)
1 µM SR1	0.79 (± 0.43))	0.54 (± 0.16)	0.25 (± 0.59)
		$CD38^-CD34^-$ cells	
Vehicle control	1.71 (± 0.43)	0.67 (± 0.23)	1.04 (± 0.20)
0.25 µM SR1	2.49 (± 1.19)	0.58 (± 0.24)	1.91 (± 0.95)
0.5 µM SR1	1.85 (± 0.78)	0.72 (± 0.45)	1.13 (± 0.33)
0.75 µM SR1	1.51 (± 0.54)	0.47 (± 0.23)	1.04 (± 0.31)
1 µM SR1	1.72 (± 0.19)	1.02 (± 0.01)	0.70 (± 0.20)

Data presented in this table is representative of three technical repeats (pooled) for each of two independent donors (n = 2).

4.3.2. Gene expression analysis

Six independent UCB units were collected to determine the effect of SR1 (1 µM) on gene expression in seven-day expanded HSPCs. Table 4.5 contains the particulars of each UCB isolation. Two UCB units, collected on the same day from two different donors, were pooled to obtain sufficient $CD34^+$ HSPCs on the day of isolation, and represent a single biological replicate.

Table 4.5. UCB units collected for gene expression experiments.

Isolate	Volume (mL)	Viability (%)	Total leukocytes	$CD34^+CD45^{dim}$ cells	$CD34^+$ (%)	Sort purity (%)
CB300818-01	80	98.6	7.2×10^7	8.5×10^5	1.19	98.0
CB300818-02	50	98.2	7.9×10^6	1.7×10^5	2.29	
CB050918-01	30	93.4	2.0×10^6	7.0×10^4	3.36	95.0
CB050918-02	110	97.4	1.5×10^8	2.8×10^6	2.02	
CB041018-01	53	97.9	1.5×10^7	2.4×10^5	1.60	94.7
CB041018-02	83	98.9	9.8×10^7	8.1×10^5	0.83	

Several conditions were compared at the transcriptome level and are illustrated in Table 4.6. Bulk RNA concentrations from non-expanded $CD34^+$ cells on the day of isolation (Day 0) and $CD34^+$ and $CD34^-$ cells after a seven-day expansion (Day 7 SR1-treated and non-treated controls) and their respective 260/280 ratios and RNA integrity numbers (RIN), are summarised in Table 4.7.

Table 4.6. Transcriptome level comparisons.

Comparisons	
Expanded (D7) SR1-treated $CD34^+$	Expanded (D7) non-treated $CD34^+$
Expanded (D7) SR1-treated $CD34^-$	Expanded (D7) non-treated $CD34^-$
Expanded (D7) SR1-treated $CD34^+$	Non-expanded (D0) $CD34^+$

Table 4.7. RNA samples for gene expression experiments. RNA samples with their respective RNA concentrations (µM), 260/280 ratios and RIN values.

Isolate	Sample	Concentration (ng/µL)	260/280	RIN
CB300818	Day 0	20.4	1.59	8.8
	Day 7 SR-1-treated $CD34^+$ cells	47.7	1.54	10.0
	Day 7 SR-1-treated $CD34^-$ cells	27.3	1.95	9.9
	Day 7 Non-treated $CD34^+$ cells	44.1	1.77	10.0
	Day 7 Non-treated $CD34^-$ cells	36.3	1.48	10.0
CB050918	Day 0	27.7	1.79	7.8
	Day 7 SR-1-treated $CD34^+$ cells	48.5	1.91	9.7
	Day 7 SR-1-treated $CD34^-$ cells	39.1	1.85	9.5
	Day 7 Non-treated $CD34^+$ cells	47.1	1.59	9.8
	Day 7 Non-treated $CD34^-$ cells	35.7	1.86	9.8
CB041018	Day 0	30.2	1.54	9.2
	Day 7 SR-1-treated $CD34^+$ cells	37.1	1.63	10.0
	Day 7 SR-1-treated $CD34^-$ cells	47.6	1.62	9.9
	Day 7 Non-treated $CD34^+$ cells	40.0	1.67	10.0
	Day 7 Non-treated $CD34^-$ cells	53.5	2.03	8.6

4.3.2.1. SR1-treated vs. non-treated

4.3.2.1.1. Immunophenotypic profiles of SR1-treated vs. non-treated cells

Immunophenotypic profiles were assessed for SR1-treated and non-treated cells after a seven-day expansion by flow cytometric analysis. Figure 4.8 shows the differences in HSPC-associated immunophenotypic profiles for SR1-treated (purple) and non-treated (grey) cells. The three columns of histograms represent the three independent donors (CB300818, CB050918 and CB041018). Overlay histograms were gated on viable (7AAD-negative) cells. The positive regions for each marker were created based on inherent negative populations present in the sample (if applicable). For markers that did not have a definitive inherent negative population, whole blood samples were used to set the negative/positive boundaries. For certain markers, such as CD133, an isotypic control was used to set the negative/positive detection boundary. The majority of SR1-treated and non-treated cells expressed lineage markers after seven-day expansion (Figure 4.8a). In contrast, only a small fraction of the cells expressed the CD133 marker after seven-day expansion (Figure 4.8b). The proportion of CD133 cells was similar for SR1-treated and non-treated cells. There was a slight increase in $CD34^+$ numbers and percentages in SR1-treated cells (also shown earlier in this chapter) compared to non-treated cells (Figure 4.8c). There were fewer $CD34^-$ cells in SR1-treated cells compared to the non-treated cells. CD38 expression ranged from cells expressing low CD38 to cells expressing high CD38. Similar CD38 expression profiles were observed for SR1-treated and non-treated cells (Figure 4.8d). The majority of the cells were $CD45^+$ and only a small fraction was $CD45^-$ (Figure 4.8e). The $CD45^-$ cells were slightly fewer in number in SR1-treated cells compared to non-treated cells.

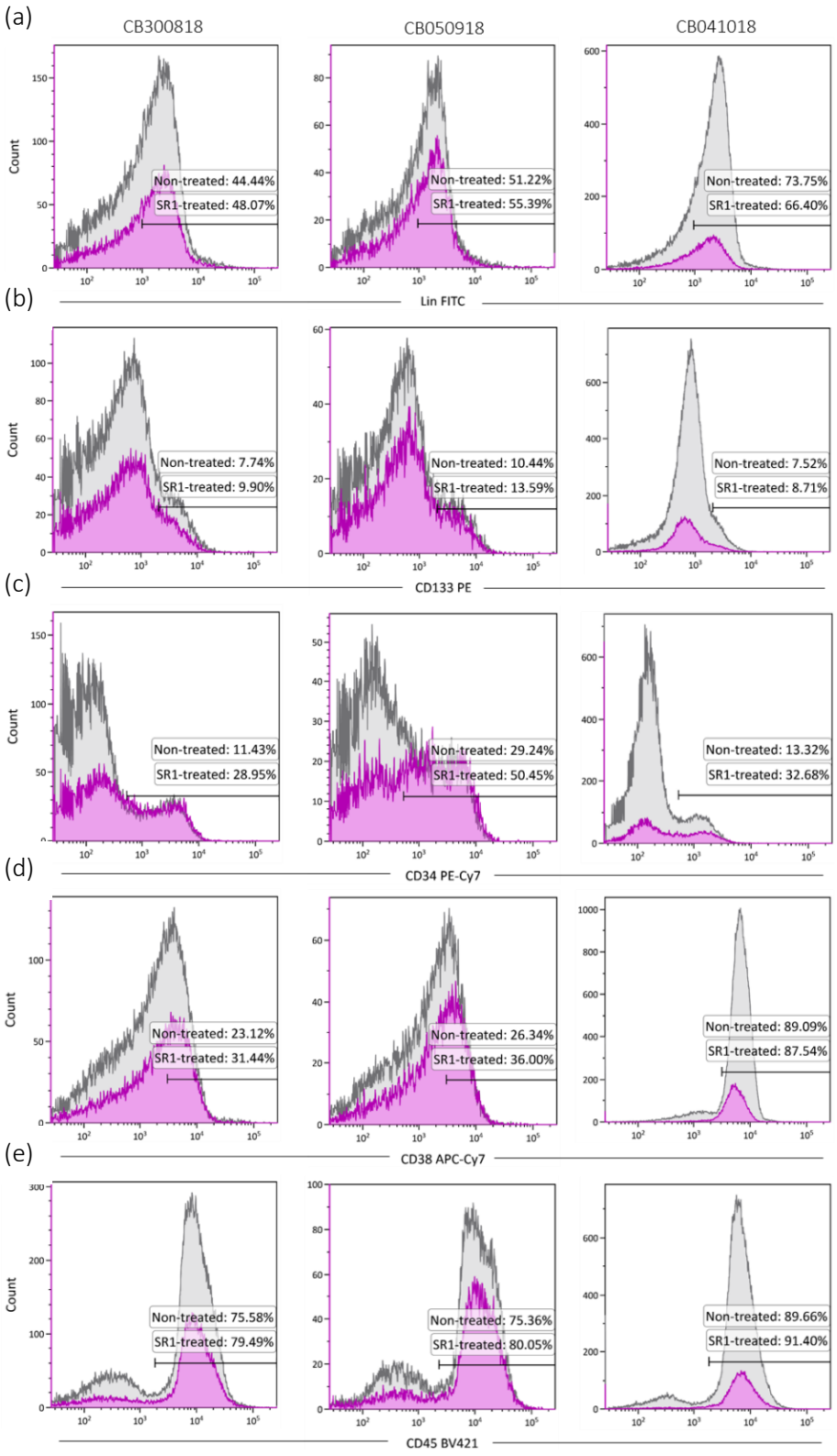


Figure 4.8. Expression of HSPC-associated phenotypic markers on SR1-treated and non-treated cells after seven-day expansion. Expression of HSPC-associated markers was assessed on SR1-treated (purple) and non-treated (grey) cells after a seven-day expansion. Overlay histograms were gated on viable (7AAD-negative) cells and include the following HSPC-associated markers: (a) Lin FITC, (b) CD133 PE, (c) CD34 PE-Cy7, (d) CD38 APC-Cy7, and (e) CD45 BV421.

Bulk gene expression analysis represents an average of gene expression patterns in a population of cells of interest. Cells with a similar function will have similar gene expression patterns, which makes it possible to group these cells into clusters, identified through PCA. Principal component analysis reduces high-dimensional data into fewer dimensions, called principal components (PCs) to simplify its complexity while retaining trends and patterns.

Figure 4.9 shows three distinct clusters representing different cells with different gene expression patterns. A clear distinction between cells on the day of isolation, Day 0 (D0, blue), and seven-day expanded cells (D7), was observed (far-right turquoise, purple, green and red clusters) (Figure 4.9). The $CD34^-$ (green and red) and $CD34^+$ (turquoise and purple) expanded cells also separated into two distinct clusters (Figure 4.9). Seven-day expanded $CD34^+$ and $CD34^-$ cells differ from one another, but had more similar gene expression patterns compared to D0 $CD34^+$ HSPCs. Interestingly, when comparing gene expression patterns between SR1-treated ($CD34^+$ and $CD34^-$) and non-treated control cells, very few differences were observed.

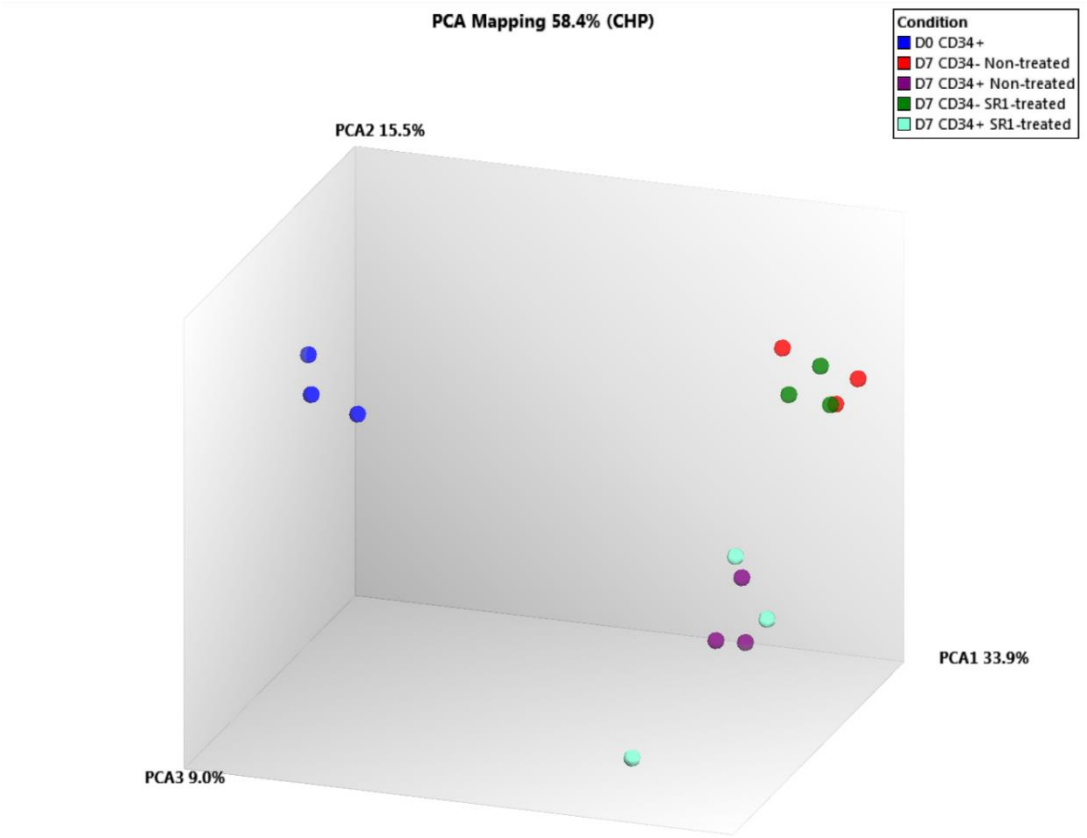


Figure 4.9. 3D PCA plot. A 3D plot of the first three PCs of non-expanded (D0) $CD34^+$ HSPCs (blue) and seven-day expanded (D7) $CD34^+$ (SR1-treated (turquoise) and non-treated (purple)) and $CD34^-$ (SR1-treated (green) and non-treated (red)) cells. Data is representative of three independent donors.

4.3.2.1.2. Gene expression profiles of SR1-treated vs. non-treated cells

Only two genes passed the filter criteria, both were downregulated in $CD34^+$ SR1-treated HSPCs compared to $CD34^+$ non-treated control HSPCs (Table 4.8). The effect of SR1 was more pronounced in $CD34^-$ cells, where eight genes passed the filter criteria, with four genes being significantly upregulated and four significantly downregulated in $CD34^-$ SR1-treated cells compared to $CD34^-$ non-treated control cells (Table 4.9). The most significantly downregulated genes in the SR1-treated $CD34^+$ and $CD34^-$ cells were cytochrome P450, family 1, subfamily B, polypeptide 1 (*CYP1B1)* and erythrocyte membrane protein band 4.1-like 3 *(EPB1L3)*. Gene ontology analysis did not reveal any significant processes associated with the up- and downregulated genes between SR1-treated and non-treated controls.

Table 4.8. Genes significantly expressed in $CD34^+$ cells expanded in SR1 for seven days.

Gene	Gene symbol	Fold change	P-value	FDR P-value
Cytochrome P450, family 1, subfamily B, polypeptide 1	CYP1B1	-9.38	1.98E-07	0.0096
Erythrocyte membrane protein band 4.1-like 3	EPB41L3	-2.57	2.45E-05	0.5916

FDR = False Detection Rate. Data presented in this table is representative of three independent donors (n = 3).

Table 4.9. Genes significantly expressed in $CD34^-$ cells expanded in SR1 for seven days.

Gene	Gene symbol	Fold change	P-value	FDR P-value
Cytochrome P450, family 1, subfamily B, polypeptide 1	CYP1B1	-4.54	8.15E-06	0.1718
Erythrocyte membrane protein band 4.1-like 3	EPB41L3	-2.59	1.44E-05	0.1735
Alpha hemoglobin stabilising protein	AHSP	4.73	0.0085	0.9999
5-Aminolevulinate synthase 2	ALAS2	2.69	0.0114	0.9999
Chemokine (C-C motif) ligand 2	CCL2	2.64	0.0129	0.9999
Dipeptidyl-peptidase 4	DPP4	-2.74	0.0168	0.9999
Glycophorin A (MNS blood group)	GYPA	2.7	0.0176	0.9999
C-type lectin domain family 10, member A	CLEC10A	-2.64	0.0303	0.9999

Data presented in this table is representative of three independent donors (n = 3).

4.3.2.2. Seven-day expanded CD34⁺ HSPCs vs. non-expanded CD34⁺ HSPCs

4.3.2.2.1. Immunophenotypic profiles of SR1-treated vs. non-treated cells

Immunophenotypic profiles of seven-day expanded and non-expanded CD34⁺ HSPCs were assessed by flow cytometric analysis. Figure 4.10 shows the differences in HSPC-associated immunophenotypic profiles for seven-day expanded (SR1-treated (purple), non-treated (grey)) and non-expanded (green) CD34⁺ HSPCs from which RNA was extracted. The three columns of histograms represent the three independent donors (CB300818, CB050918 and CB041018). Overlay histograms were gated on viable (7AAD-negative) and CD34⁺ HSPCs. A clear difference was observed with regard to lineage marker (Lin FITC) expression between expanded and non-expanded CD34⁺ HSPCs (Figure 4.10a). Two distinct populations (Lin⁻ and Lin⁺) were present in non-expanded CD34⁺ HSPCs, while expanded CD34⁺ HSPCs expressed lineage markers to a lesser extent (based on MFI) when compared to the Lin⁺ population in non-expanded CD34⁺ HSPCs. Expanded CD34⁺ HSPCs showed higher expression of CD133 compared to non-expanded CD34⁺ HSPCs (Figure 4.10b). The expression profiles for CD38 were similar in expanded and non-expanded CD34⁺ HSPCs (Figure 4.10c). Active cells generally express higher CD38 (290). Expanded and non-expanded CD34⁺ HSPCs generally included a combination of active and non-active cells, with no clear distinction between CD38⁻ and CD38⁺ cells. With regard to CD45, two distinct populations were present in non-expanded CD34⁺ HSPCs; one showed increased expression of CD45, while the other expressed intermediate levels of CD45 (Figure 4.10d). The majority of the non-expanded CD34⁺ HSPCs expressed intermediate levels of CD45, while only a small fraction showed increased expression of CD45. The expanded CD34⁺ HSPCs expressed higher levels of CD45.

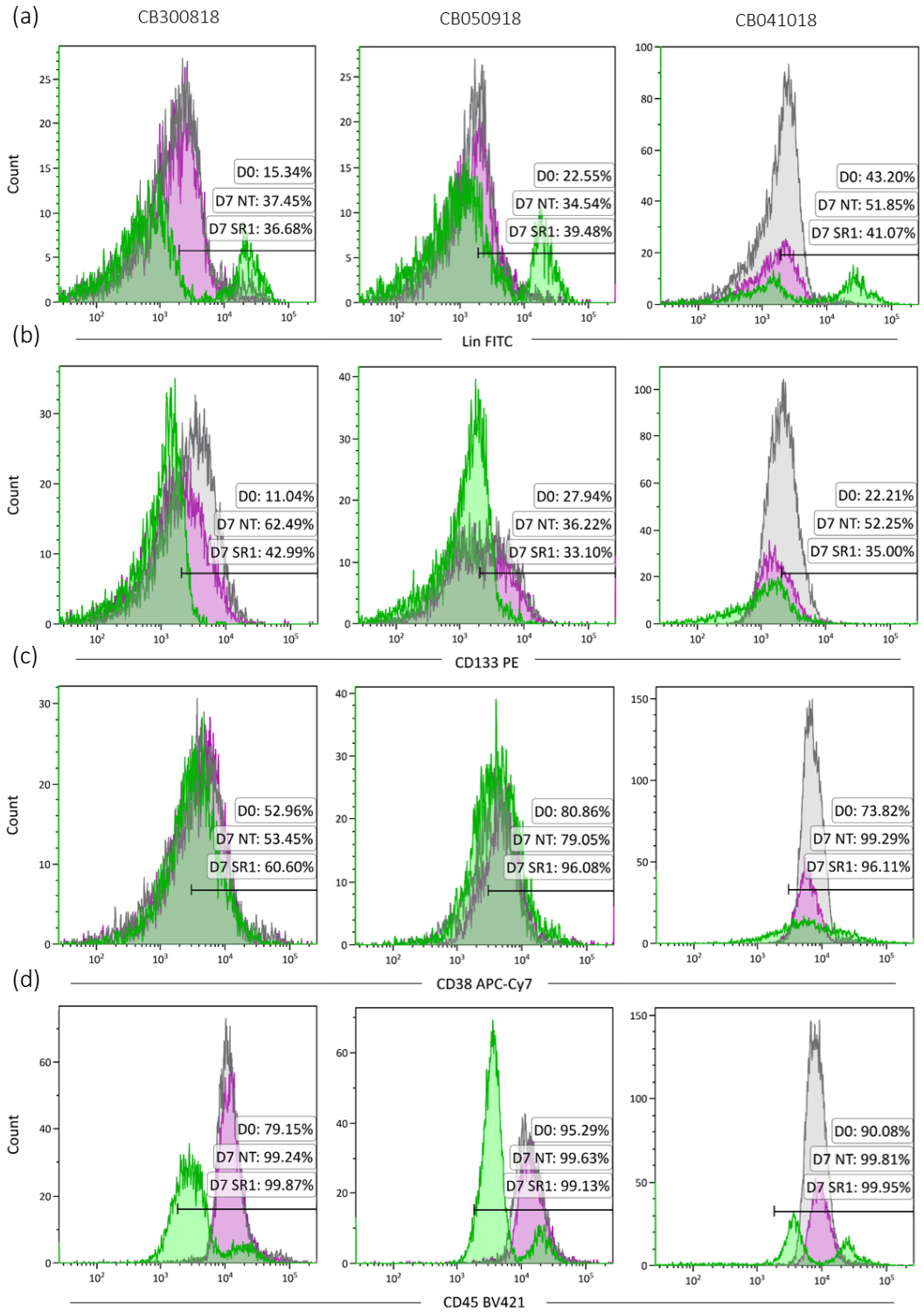


Figure 4.10. Expression of HSPC-associated phenotypic markers on expanded and non-expanded CD34⁺ HSPCs. Expression of HSPC-associated markers was assessed by flow cytometry on non-expanded HSPCs on Day 0 (D0, green) and seven-day expanded CD34⁺ HSPCs (Day 7 SR1-treated (D7 SR1, purple) and non-treated (D7 NT, grey)). Overlay histograms were gated on viable (7AAD-negative) and CD34⁺ cells.

4.3.2.2.2. Gene expression profiles of SR1-treated vs. non-treated cells

A total of 847 genes passed the filter criteria; 391 genes were significantly upregulated, while 456 genes were significantly downregulated in expanded vs. non-expanded CD34⁺ HSPCs (Figure 4.11). The heat map reveals different gene expression patterns for expanded and non-expanded CD34⁺ HSPCs (Figure 4.12). The gene expression data indicate distinct changes in the CD34⁺ HSPCs upon expansion. GO classification revealed several processes enriched for upregulated (Figure 4.13a) and downregulated (Figure 4.13b) genes.

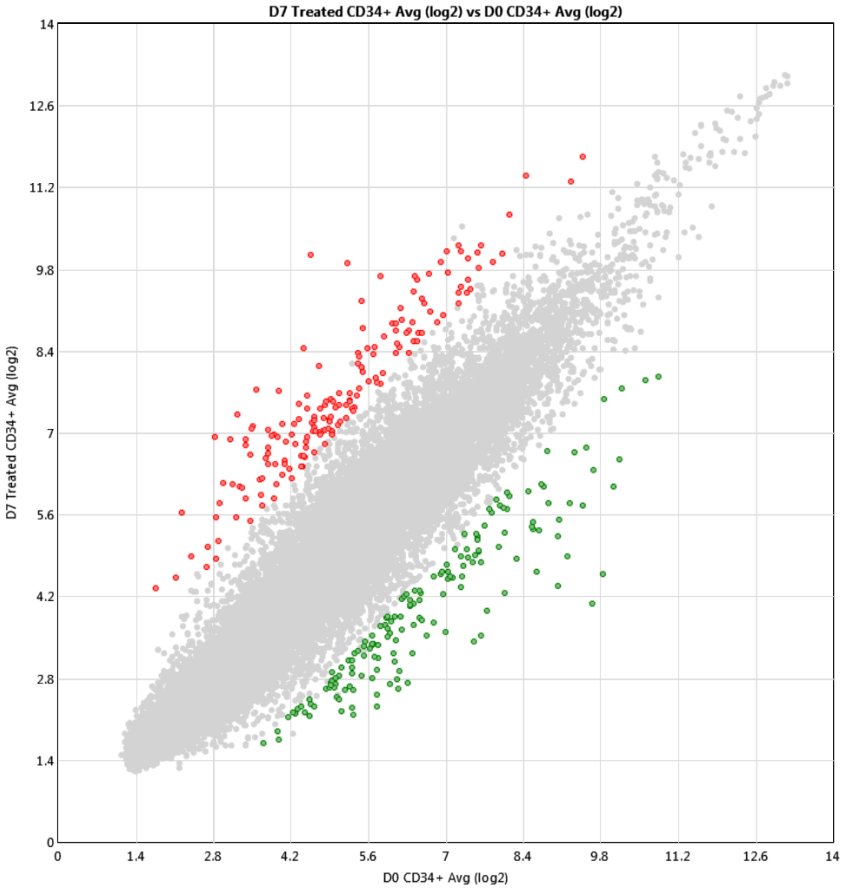


Figure 4.11. Scatter plot showing differentially expressed genes. A scatter plot showing significant up- (red) and downregulated (green) genes in expanded compared to non-expanded CD34⁺ cells.

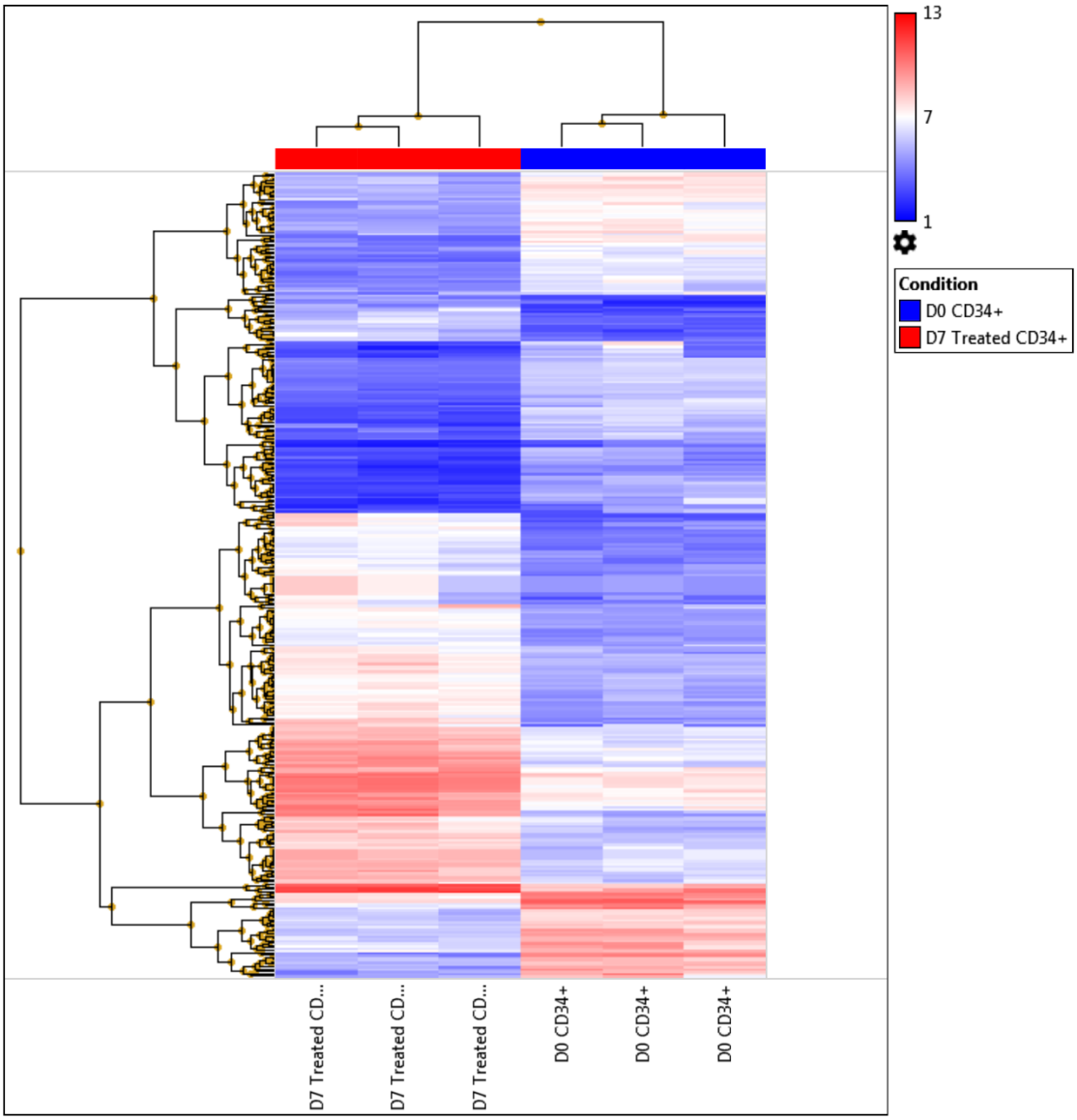


Figure 4.12. Heat map showing gene expression patterns. (a) Heat map showing the expression of genes in expanded (top red) and non-expanded (top blue) cells. Up- and downregulated genes are indicated in red and blue, respectively.

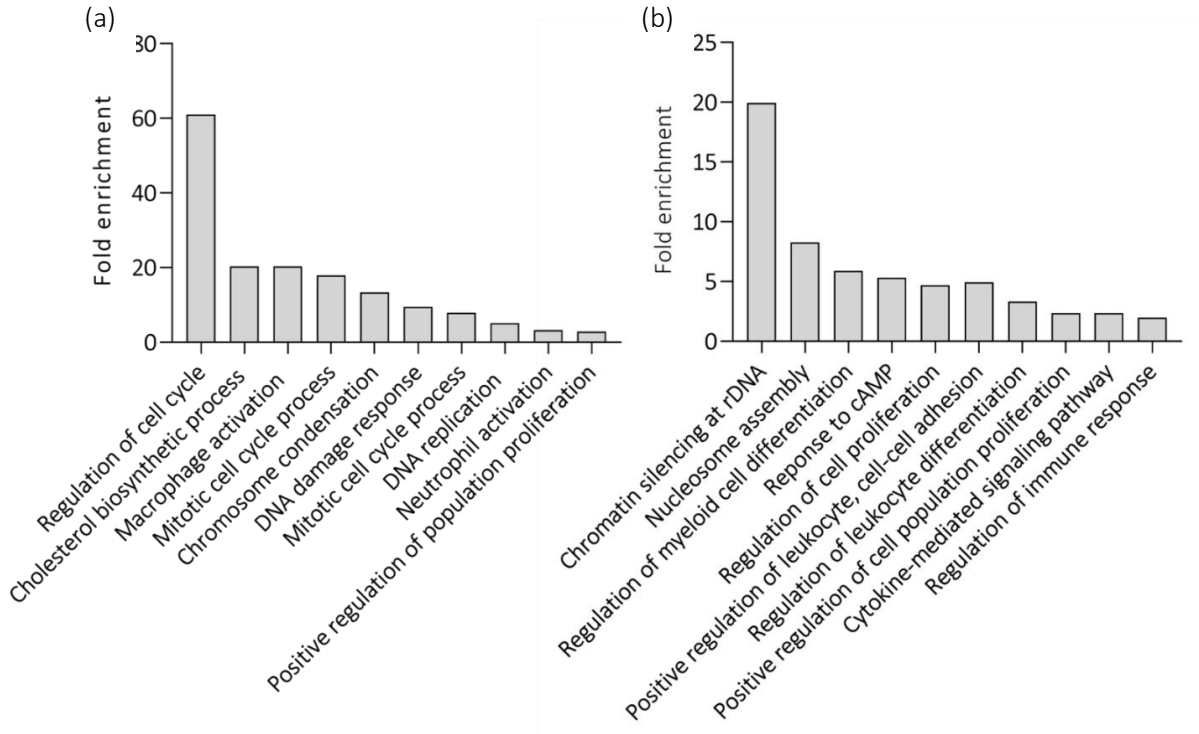


Figure 4.13. Biological processes using GO. GO classification revealed processes (FDR p-value < 0.05) enriched for significantly (a) up- and (b) downregulated genes in seven-day expanded $CD34^+$ HSPCs compared to non-expanded $CD34^+$ HSPCs.

4.4. DISCUSSION AND CONCLUSION

Several studies have used SR1 for the *ex vivo* expansion of HSPCs. In 2016, a Phase I/II clinical trial was published with SR1-expanded HSPCs transplanted in combination with a non-expanded UCB unit (78). The purpose of our study was to determine the effect of SR1 on the transcriptome of expanded $CD34^+$ and $CD34^-$ cells from UCB, and to establish to what extent the transcriptome differs between seven-day expanded and non-expanded $CD34^+$ cells.

4.4.1. SR1 concentration optimisation

SR1 has been used at various concentrations to expand HSPCs *ex vivo* (112,114,291,292). The first objective of our study was to determine the optimal concentration of SR1 needed for the expansion of locally sourced HSPCs with in-house culture conditions to ensure that the optimal SR1 concentration would be used in gene expression experiments. Freshly isolated UCB-derived HSPCs were treated with four different concentrations of SR1 namely 0.25 µM, 0.50 µM, 0.75 µM and 1 µM.

Viability remained high in all cultures (Figure 4.2), irrespective of the concentration of SR1, which is consistent with findings from Boitano *et al*. (112) and Tao *et al.* (114). The authors used SR1 concentrations ranging from 0.1 to 10 μM and showed high viability in all cultures supplemented with SR1, (112,114).

The $CD34^+$ percentages were higher in SR1-treated cultures compared to the untreated vehicle controls and increased as the concentration of SR1 increased (Figure 4.3). The absolute number of viable cells was similar across the different concentrations of SR1 (Figure 4.4), which indicated that SR1 did not increase the absolute number of cells, but rather increased the proportion of $CD34^+$ cells, which also resulted in an increased number of $CD34^+$ cells (Figure 4.4). It was also observed by Boitano *et al.* (112) and Tao *et al*. (114) that various concentrations of SR1 (0.1 – 1 μM) used for the expansion of UCB and mobilised peripheral blood, did not result in increased absolute cell numbers; however, the number of $CD34^+$ cells increased as the concentration of SR1 increased.

HSPC-associated immunophenotypic analysis revealed increased cell numbers for both $CD34^+CD38^-$ and $CD34^+CD38^+$ populations as the concentration of SR1 increased (Figure 4.5). Overall, $CD34^+CD38^-$ cell numbers were higher than $CD34^+CD38^+$ cell numbers in all cultures. As already mentioned in previous chapters, the $CD34^+CD38^-$ fraction of cells is enriched for cells with long-term repopulating potential, while $CD34^+CD38^+$ population is enriched for short-term repopulating cells (232). SR1 treatment resulted in an increase in both of these populations after a seven-day expansion. All populations that expressed CD34 ($CD34^+$), irrespective of the presence of other HSPC-associated phenotypic markers, increased in the presence of SR1 (Table 4.2), while the opposite was observed for $CD34^-$ cell populations. A more primitive and dormant population of HSPCs has been shown to be $CD34^-$ (41,229). Several studies suggest that the $CD34^-$ HSPC population only acquires CD34 expression prior to cell division (234,293,294). The $CD34^-CD133^+$ population is thought to be an even more primitive population of HSPCs than $CD34^+$ sub-populations (44). The $Lin^-CD34^-CD38^-$ population (Figure 4.6a) and the $Lin^-CD34^-CD38^-CD133^+$ population (Figure 4.6b) decreased as the concentration of SR1 increased. Studying the $CD34^-$ population of HSPCs has been

challenging, since a positive marker for this population remains to be identified (39). $CD34^-$ HSPCs are kinetically and functionally more dormant than $CD34^+$ cells. CD34 expression indicates entry into the cell cycle, metabolic activation and homing ability (39).

Our findings are in line with the findings of Boitano *et al.* (112), who also found that the various $CD34^+$ sub-populations increased when UCB-derived $CD34^+$ HSPCs were expanded *ex vivo* for five weeks in the presence of 1 µM SR1. Tao *et al.* (114) likewise showed that Lin^-CD34^+ cells expand in an SR1 concentration-dependent manner, and also observed a decrease in the $CD34^-$ population. It is, however, important to take into consideration that the latter study used higher concentrations of SR1 (1 – 5 µM). The highest SR1 concentration used in this study was 1 µM.

$CD34^+CD38^-$ and $CD34^-CD38^-$ cells displaying side population characteristics were highest at a concentration of 0.25 µM SR1 and decreased as the concentration of SR1 increased (Figure 4.7). Although 0.25 µM SR1 did not increase the $CD34^+$ cell numbers to the same extent as 1 µM SR1, it retained cells with dye efflux ability; these cells are suggested to be a more primitive population of HSPCs (271). The percentage SP was consistently higher in $CD34^-CD38^-$ cells than in other populations. The decreased SP observed as the concentration of SR1 was increased might be a result of the decreased number of $CD34^-CD38^-$ cells.

The data from this study suggest that SR1 expands primitive $CD34^+$ HSPCs ($CD34^+CD38^-$) along with progenitors ($CD34^+CD38^+$). However, proliferation of the more primitive $CD34^-$ HSPCs is reduced in the presence of SR1, especially at higher concentrations. The $CD34^+CD133^+$ population increased in the presence of SR1, which identified an HSPC population with multipotent potential (249). Several studies have shown that $CD34^-$ cells have less repopulating potential (234), and it has been suggested that engraftment success is a function of $CD34^+$ cell number (84,295).

Although no functional assays were performed during this study, Boitano *et al*. (112) and Jackson et al. (230) showed that UCB-derived HSPCs cultured in the presence of SR1 have colony forming potential using CFU assays. The above-mentioned studies also performed

secondary transplantation studies, where SR1-expanded HSPCs were transplanted into primary and secondary humanised mouse recipients, which resulted in improved engraftment compared to cells expanded without SR1 (112,291). A Phase I/II clinical trial has shown successful engraftment when SR1-expanded HSPCs were transplanted in combination with non-expanded UCB units (78). Based on our results, 1 µM SR1 was used for the expansion of $CD34^+$ HSPCs for gene expression experiments.

4.4.2. SR1-treated vs. non-treated cells

For gene expression analysis, HSPCs were cultured with 1 µM SR1 for seven days and a non-treated vehicle control was included. The effect of SR1 on gene expression of HSPCs has previously been evaluated for shorter culture periods (112,292). To the knowledge of the authors, this is the first time that the effect of SR1 on the transcriptome of seven-day expanded HSPCs has been reported. This is relevant since SR1 is currently being used to expand HSPCs *ex vivo* for extended periods rather than the shorter periods previously assessed (78). Boitano *et al.* (112) found AhR pathway-related genes, aryl hydrocarbon receptor repressor (*AHRR)* and heat-shock protein 90α (*HSP90AA1)* to be downregulated in SR1-treated cells after 48 hours. This study did not report differential expression among the AhR pathway-related genes, such as *AHRR and HSP90AA1*. This might suggest that SR1 initially acts on the above-mentioned genes, but that this particular effect is restored after seven days. Cytochrome P450 1B1 (*CYP1B1)* and erythrocyte membrane protein band 4.1-like 3 (*EPB41L3*) were most significantly reduced after seven-day expansion with SR1 in this study. *CYP1B1* is a target of AhR and its downregulation suggests that SR1 antagonises AhR in HSPCs. Downregulation of *CYP1B1* was also observed by Boitano *et al.* (112) after 48 hours of HSPC expansion and likewise by Koide *et al.* (292) in the expansion of NB4 cells. *CYP1B1* was further reduced in this study compared to the above-mentioned studies, which possibly results from using RNA from seven-day expanded cells in this study, compared to 48 hours (112) and four days (292). This suggests that the effect of SR1 on *CYP1B1* seems to be long lasting, since it remains significantly downregulated after seven days of expansion with SR1. Boitano *et al.* (112), however, have previously shown that the effect of SR1 is reversible. Therefore, *CYP1B1* expression might return to normal once SR1 is removed from the culture medium.

4.4.3. Seven-day expanded CD34$^+$ cells vs. non-expanded CD34$^+$ cells

This is the first time that SR1-expanded CD34$^+$ HSPCs have been compared to non-expanded CD34$^+$ HSPCs at the transcriptome level. The gene expression profiles between seven-day expanded CD34$^+$ HSPCs and non-expanded CD34$^+$ HSPCs were noticeably different (Figure 4.12). GO classification revealed up-regulated genes in expanded vs. non-expanded CD34$^+$ HSPCs enriched for processes including regulation of cell cycle, macrophage activation, DNA replication, neutrophil activation and positive regulation of population proliferation (Figure 4.13a). This indicates that *ex vivo* expansion of cells activates genes involved in proliferation and differentiation, which was expected since several cytokines were included in addition to SR1 to promote expansion of HSPCs.

GO classification revealed downregulated genes in expanded vs. non-expanded CD34$^+$ cells to be enriched for processes such as chromatin silencing at recombinant DNA (rDNA), nucleosome assembly and regulation of myeloid cell differentiation (Figure 4.13b). Although there are noticeable differences between expanded and non-expanded CD34$^+$ HSPCs, several studies have reported successful engraftment *in vivo* in murine models of SR1-expanded HSPCs (112,291). Also, the use of expanded HSPCs in combination with non-expanded UCB units for HSCT is already underway clinically (78). Gori *et al.* (296) cultured primate HSPCs in the presence of endothelial cells (EC), and even though the transcriptomic profiles of EC-expanded vs. non-expanded HSPCs were very different, long-term engraftment (assessed 800 days post-transplant) was still achieved.

Some of the limitations of the study include that these experiments were performed using bulk RNA. As mentioned earlier in this chapter, bulk RNA represents the average expression of the cells in the population of interest, while the effect on smaller populations is usually masked. This could be particularly problematic if SR1 has a greater effect on smaller sub-populations within the CD34$^+$ HSPC population. Another limitation is that an increased number of cells is generally required to obtain sufficient good quality RNA. To obtain sufficient cells and RNA from smaller sub-populations, such as the CD34$^-$ population of HSPCs, is challenging. Single-cell transcriptome analysis would therefore be beneficial to study the effect of SR1 on CD34$^+$ and CD34$^-$ HSPC populations. This would allow us to determine whether the effect of SR1 is more

pronounced in particular sub-populations within these populations. Single-cell analysis would further allow identification of sub-populations present following SR1-mediated expansion, and will aid in determining differences between expanded and non-expanded cells to increase applications of expanded cells in the future.

In conclusion, the findings from this study confirm the effect of SR1 on the proliferation of $CD34^+$ HSPCs and further reveal little effect of SR1 on the transcriptome of seven-day expanded $CD34^+$ and $CD34^-$ cells, with only two genes being significantly downregulated in SR1-treated cells ($CD34^+$ and $CD34^-$) compared to the non-treated controls. In contrast, the transcriptome of expanded $CD34^+$ cells differs significantly from the transcriptome of non-expanded $CD34^+$ cells. GO classification of differentially expressed genes suggests the gene expression changes are related to differentiation and proliferation processes. In addition, the use of $CD34^+$ HSPCs exclusively expanded with SR1 would be beneficial in cases where the HSPC cell dose in the initial harvested cell therapy products is suboptimal and therefore not a feasible option for HSCT when used alone.

CHAPTER 5. HUMAN CD34$^+$ HEMATOPOIETIC STEM AND PROGENITOR CELL TRANSCRIPTOME AT A SINGLE-CELL LEVEL

5.1. INTRODUCTION

Umbilical cord blood-derived CD34$^+$ HSPCs have been used successfully for UCBT across the globe to treat malignant and non-malignant disorders. The UCB-derived CD34$^+$ HSPC population is heterogeneous in nature. This population of cells, as the name suggests, encompasses primitive stem cells and early and late progenitors at different stages of differentiation. Cells belonging to different HSPC sub-populations are distinct in their self-renewal potential, *in vitro* colony-formation and *in vivo* behaviour (34). A great deal of effort is being made to understand the heterogeneity within CD34$^+$ HSPC populations and its contribution to the successful clinical use of these cells, and is consequently an active area of research.

Even though all cells within the human body are genetically almost identical, containing the same set of roughly 20 000 genes, gene regulation and expression are unique for individual cells. The gene expression profile of a single cell provides a vast amount of information regarding its phenotype, which underlies its molecular function. Until recently, gene expression studies have been limited to pooled populations of cells in order to obtain sufficient amounts of RNA, referred to as bulk RNA (297). This provides an average of levels of gene expression across the cell types present. Gene expression studies using RNA from pooled cell populations mask the uniqueness and heterogeneity of gene expression patterns present in individual cell types. This results mainly in abundant cell types being studied, while rare cell populations remain poorly characterised (298).

The first mammalian single-cell sequencing of DNA (299) and RNA (300) were performed in 2011 and 2009, respectively. Single-cell RNA sequencing (RNA-seq) enables researchers to uncover the uniqueness of each cell and investigate previously unknown cellular heterogeneity in cell populations (301,302). This provides meaningful information into the behaviour of cells, enabling a deeper understanding of complex biological systems and processes that individual cells are involved in. Single-cell gene expression profiling has already contributed to

improved understanding of cellular heterogeneity in a variety of cell types such as neurons (303,304), embryonic stem cells (305,306) and cancer cells (307,308). In addition, several single-cell studies have focused on the heterogeneity of murine and human HSPCs (309–311).

The purpose of this study was to characterise the heterogeneity of $CD34^+$ HSPCs from UCB using single-cell transcriptome analysis. This chapter describes the experimental set-up for the capture of individual cells using the Fluidigm C1™ system, and the lysis, reverse transcription, amplification and sequencing of the messenger RNA (mRNA) from individual cells. This chapter further describes the computational methods that were used to characterise the heterogeneity by means of clustering.

5.2. MATERIALS AND METHODS

This project made use of the Fluidigm C1™ Single-Cell Auto Prep System together with C1™ Single-Cell Auto Prep Array Integrated Fluidics Circuit (IFC) plates to capture single cells. The IFC plates capture a maximum of 96 cells per plate. The SMARTer® and SMART-Seq® V4 technologies were used for the lysis, reverse transcription and amplification of cDNA.

There are four critical steps in the single-cell RNA-seq workflow: (1) Isolation and capture of single cells, (2) lysis, reverse transcription and cDNA amplification, (3) library preparation and sequencing, and (4) computational analysis. These steps will be discussed in more detail in the sections to follow.

5.2.1. Isolation and capture of individual $CD34^+$ HSPCs

5.2.1.1. Sample preparation

$CD34^+$ cells isolated from UCB were used for these experiments. Informed consent and sample collections were performed as discussed in Chapter 2, Section 2.2.2. Samples were processed on the day of collection and processing started within an hour of collection. Six independent UCB units were collected for this part of the study (Table 5.7). Enrichment of $CD34^+$ HSPCs was performed as described in Chapter 2, Section 2.2.3. A cell count was performed on each sample after red blood cells had been lysed. After verifying that there were sufficient numbers of $CD34^+$ cells post-enrichment (a minimum of 1×10^5 cells), cells were prepared for sorting.

$CD34^+$ cells were sorted directly from the buffy coat (containing all the white blood cells). Stem-Kit reagents (CD45 FITC, CD34 PE and 7AAD) were added to the buffy coat fraction (approximately 5 µL per 10 x 10^6 viable cells). Once the antibodies had been added, samples were incubated for 30 min at room temperature. Stained cells were washed using TP buffer and the supernatant was aspirated. The cell pellet was resuspended in 5 – 10 mL TP buffer and filtered using a 30 µM tube-top filter into a flow tube (Falcon, Corning, USA) for sorting. The purity check was performed as previously described in Chapter 2, Section 2.2.3.3. The sort purities achieved are represented in Figure 2.2.

5.2.1.2. Priming of the C1 IFC plate and capture of the cells

Three IFC plates sizes, namely small- (5 – 10 µM), medium- (10 – 17 µM), and large-cell (17 – 25 µM) are available from Fluidigm for the C1 system. Since the average size of HSPCs is approximately 10 µM, both small (x3) and medium (x3) IFC plates were used to capture HSPCs for this study. The IFC plates needed to be primed before use with reagents provided in the C1™ reagent kits (Figure 5.1) (Fluidigm, USA). All reagents were thawed and mixed thoroughly before use. Reagents were thawed and added according to the user manual.

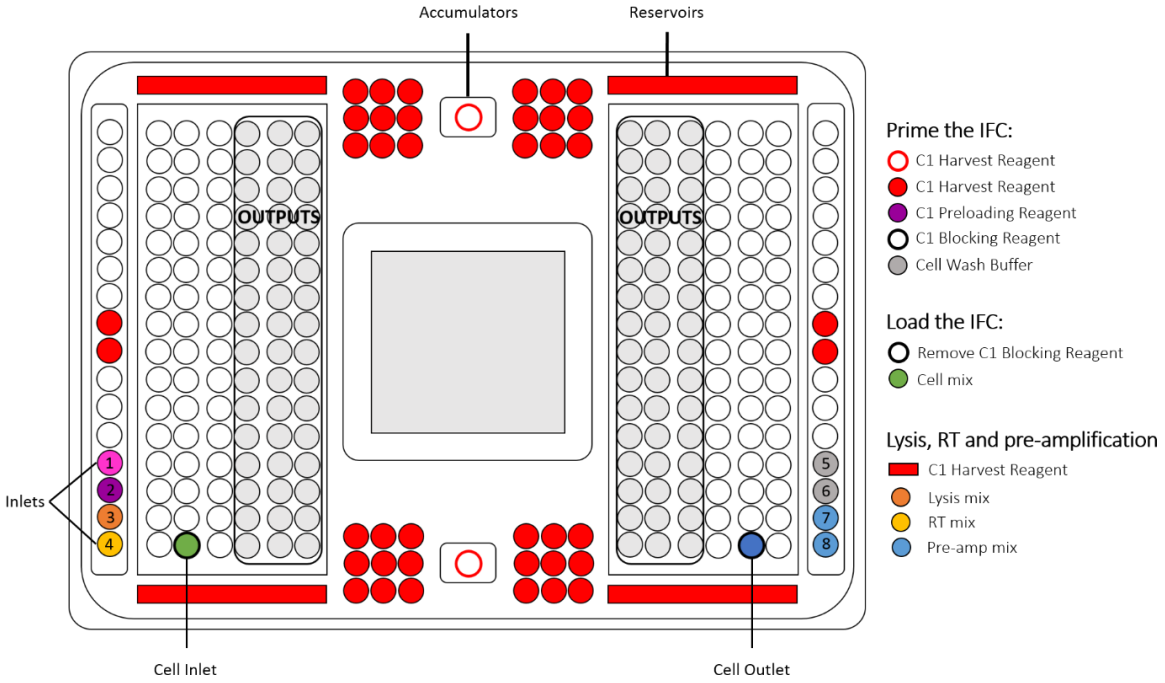


Figure 5.1. C1 IFC plate. A schematic illustration of the C1 IFC plate and the positions on the plate for the various reagents used to prime and load the IFC. The positions for lysis, reverse transcription (RT) and cDNA amplification mixes are also indicated. (Figure was created by Juanita Mellet, adapted from the SMART-Seq® V4 Ultra® Low Input RNA Kit for the Fluidigm® C1™ System, IFCs User Manual).

C1™ Harvest Reagent was added to the accumulators, the 36 inlets surrounding the accumulators and the two inlets in the middle of the outside columns on both sides of the IFC. C1™ Preloading Reagent was added to inlet 2 and C1™ Cell Wash Buffer was added to inlets 5 and 6. C1™ Blocking Reagent was added to the cell inlet and outlet wells. The IFC plate was placed inside the C1 and the "mRNA Seq: Prime" program was selected.

After the completion of the priming procedure, the Blocking reagent was removed from the cell inlet and outlet wells and the C1 Cell Wash Buffer was added to inlet 1. A total of 10 000 viable (7AAD-negative) $CD34^+$ cells (final cell concentration of 714 cells/µL) were sorted into the cell inlet of the IFC plate containing 3 µL of PBS. After the cells were sorted, 1 µL of Fluidigm's C1™ Suspension Reagent was added to the cell inlet, to achieve the correct buoyancy, which allows the cells to move into the microfluidic tubing. Optimal buoyancy is achieved from the correct ratio of cells (re-suspended in PBS) to C1 Suspension Reagent. Buoyancy was verified after addition of the C1™ Suspension Reagent using the Zeiss Axio Vert.A1 inverted microscope (Zeiss, Germany), after which the IFC plate was loaded back into the C1 platform.

The “mRNA Seq: Cell Load” program was selected. After the “mRNA Seq: Cell Load” program was completed, capture efficiency was confirmed using the Zeiss Axio Vert.A1 inverted microscope, noting capture sites containing multiple cells and/or debris. Once it was confirmed that sufficient cells had successfully been captured, the lysis, reverse transcription and PCR mixes were prepared.

5.2.2. Lysis, reverse transcription and cDNA amplification

The lysis, reverse transcription and cDNA amplification were performed using the SMARTer® Ultra® Low RNA kit for the Fluidigm® C1™ system (Clontech, Takara Bio Inc., USA) and the SMART-Seq® V4 Ultra® Low RNA kit for the Fluidigm® C1™ system (Clontech, Takara Bio Inc., USA). Reagent mixtures were prepared according to the relevant user manuals (SMARTer® Ultra® Low Input RNA Kit for the Fluidigm® C1™ System, IFCs User Manual and SMART-Seq® V4 Ultra® Low Input RNA Kit for the Fluidigm® C1™ System, IFCs User Manual).

RNA Spike-ins served as a positive control for thermal cycling of the C1 system independent of cell capture. A Spike-in mixture (1:40 000 dilution), consisting of ERCC RNA Spikes (Ambion, Life Technologies, Thermo Fisher Scientific, USA) and C1 loading reagent, was used per recommendation by Dr Christelle Borel, a senior associate at the University of Geneva working with HSPCs. A 1:2000 dilution was stored at -80°C and further diluted to 1:40 000 before each run. A small volume (1 µL) of the Spike-In mixture was used for each experiment. The lysis, reverse transcriptase and PCR mixtures were made up as per the manufacturer’s instructions. Tables 5.1 and 5.2 list the reagents present in the lysis mixture for the SMARTer and SMART-Seq V4 kits, respectively.

Table 5.1. SMARTer lysis mix.

Reagent	Volume
RNA Spike Mix	1 µL
RNase Inhibitor	0.5 µL
3’ SMART CDS Primer IIA	7 µL
Clontech Dilution Buffer	11.5 µL
Total volume	**20 µL**

Table 5.2. SMART-Seq V4 lysis mix.

Reagent	Volume
RNA Spike Mix	1 µL
3' SMART-Seq CDS Primer IIA	2.4 µL
Nuclease-Free Water	14 µL
10X Reaction Buffer	2.6 µL
Total volume	**20 µL**

Both the SMARTer and SMART-Seq V4 chemistries use oligo-dT primers for the reverse transcription of polyadenylated mRNA to prime first-strand cDNA synthesis (297,312). The total RNA in a human cell includes messenger RNA (mRNA), microRNA and ribosomal RNA (313). Ribosomal RNA accounts for over 95% of the total RNA in a cell (314). The SMART technology includes a poly(A)+ selection strategy to ensure capture of mRNA transcripts, while excluding the other RNA molecules in a given cell (297). The cDNA is then amplified via PCR using a single primer. Tables 5.3 and 5.4 list the reagents present in the reverse transcriptase mixture for the SMARTer and the SMART-Seq V4 kits, respectively.

Table 5.3. SMARTer reverse transcriptase mix.

Reagent	Volume
C1™ Loading Reagent	1.2 µL
5X First-Strand Buffer	11.2 µL
Dithiothreitol (DTT)	1.4 µL
dNTP Mix	5.6 µL
SMARTer IIA Oligonucelotide	5.6 µL
RNase Inhibitor	1.4 µL
SMARTScribe Reverse Transcriptase	5.6 µL
Total volume	**32 µL**

Table 5.4. SMART-Seq V4 reverse transcriptase mix.

Reagent	Volume
C1™ Loading Reagent	1.2 µL
5X First-Strand Buffer	11.2 µL
SMART-Seq IIA Oligonucelotide	2.8 µL
RNase Inhibitor	1.4 µL
Nuclease-Free Water	9.8 µL
SMARTScribe Reverse Transcriptase	5.6 µL
Total volume	**32 µL**

After converting RNA into first-strand cDNA, second-strand synthesis was achieved using the SMART technology (switching mechanism at 5′ end of RNA template). This chemistry exploits the transferase and strand-switch activity of Moloney Murine Leukemia Virus (M-MLV), which incorporates template-switching oligonucleotides as adaptors for downstream PCR amplification (315). Tables 5.5 and 5.6 list the reagents present in the PCR mixture for the SMARTer and the SMART-Seq V4 kits, respectively.

Table 5.5. SMARTer PCR mix.

Reagent	Volume
PCR-Grade Water	63.5 µL
10X Advantage 2 PCR Buffer	10 µL
50× dNTP Mix	4 µL
IS PCR Primer	4 µL
50X Advantage 2 Polymerase Mix	4 µL
C1 Loading Reagent	4.5 µL
Total volume	**90 µL**

Table 5.6. SMART-Seq V4 PCR mix.

Reagent	Volume
C1 Loading Reagent	4.5 µL
PCR primer IIA	4.4 µL
SMART-Seq IIA Oligonucelotide	3 µL
SeqAmp PCR Buffer	75.2 µL
SeqAmp DNA Plolymerase	2.9 µL
Total volume	**90 µL**

C1 Harvest Reagent was added to each of the four reservoirs of the IFC plate (Figure 5.1). Lysis mixture was added to inlet 3, while the reverse transcriptase mixture was added to inlet 4 and the PCR mixture was added to inlets 7 and 8. The IFC was placed inside the C1 system and the program "mRNA-Seq: RT and Amp" was selected. The "mRNA-Seq: RT and Amp" program takes approximately eight hours to run and was therefore left to run overnight. The resultant cDNA samples were harvested the following morning. Tube controls were included with each run (Appendix A) and were run in parallel to ensure that the kit chemistry was working outside of the C1 instrument.

C1™ DNA Dilution Reagent (10 µL) pre-warmed to room temperature was added to each well of a 96-well plate for harvesting. The harvested cDNA (volume of around 3.5 µL) products from the output wells (Figure 5.1) of the IFC plate were added into the same 96-well plate (containing the Dilution Reagent) using a pipetting map provided in the user manual. The 96-well plate containing the harvested single-cell cDNA products was stored at -20°C until further use.

5.2.3. Library preparation and sequencing of single HSPCs

Upon harvesting single-cell cDNA products, quality control checks were undertaken to ensure that only the best quality samples were sent for sequencing. Only cDNA from single cell capture sites that did not contain visible traces of debris was sent for sequencing.

5.2.3.1. Quality control check of the single-cell cDNA products

Quality control checks included fluorometric quantification of all samples and determining the size distribution of randomly selected cDNA products.

5.2.3.1.1. Fluorometric quantification of single-cell cDNA products

All samples were quantified using the Quantus™ Fluorometer (Promega, USA) together with the Quant-iT™ PicoGreen® dsDNA kit (Invitrogen, Thermo Fisher Scientific, USA). The Quant-iT PicoGreen dsDNA kit is an ultrasensitive fluorescent nucleic acid stain for quantitating double-stranded DNA.

Quant-iT PicoGreen dsDNA reagents were pre-warmed to room temperature before use. A 1X TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.5) working solution was prepared on the day of quantification by diluting the 20X TE buffer with nuclease-free water. The 1X TE buffer solution was used to dilute the Quant-iT PicoGreen dsDNA reagents and the cDNA samples. The Quant-iT PicoGreen reagent is sensitive to light and was therefore protected from light throughout the quantification process.

cDNA products were quantified using the Quantus Fluorometer, which is equipped with one red and one blue fluorescence channel. The red channel has an excitation wavelength of 640 nm and emission wavelength of 660 – 720 nm. The blue channel has an excitation wavelength of 495 nm and emission wavelength of 510 – 580 nm. The Quantus Fluorometer uses a single-point calibration process and was calibrated before each quantification.

Samples were prepared by adding 99 µL of the 1X TE working solution and 100 µL PicoGreen working solution to 1 µL single-cell cDNA. Samples were mixed gently by pipetting and incubated for 5 min at room temperature, protected from light. Each single-cell cDNA product was measured twice, and the concentrations were recorded. The average concentration calculated from the two respective measurements were used for the library preparations.

5.2.3.1.2. Size distribution of single-cell cDNA products

The TapeStation® 2200 (Agilent Technologies, USA), together with the High Sensitivity D5000 ScreenTape® and Sample buffer kit (Agilent Technologies, USA) were used to confirm cDNA size distribution and quality. Tube controls and several randomly selected samples for each biological replicate were analysed using the TapeStation 2200. The High Sensitivity D5000 ladder control was prepared by adding 2 µL of D5000 Sample buffer to 2 µL of the D5000 ladder. Samples were prepared by adding 2 µL of Sample buffer to 2 µL of the thawed cDNA sample and vortexed for 1 min. Samples were loaded into the TapeStation and analysed using the TapeStation 2200 Controller Software. A gel image and electropherograms for the visual assessment of the cDNA fragment size and distribution were generated for each sample (Figure 5.2).

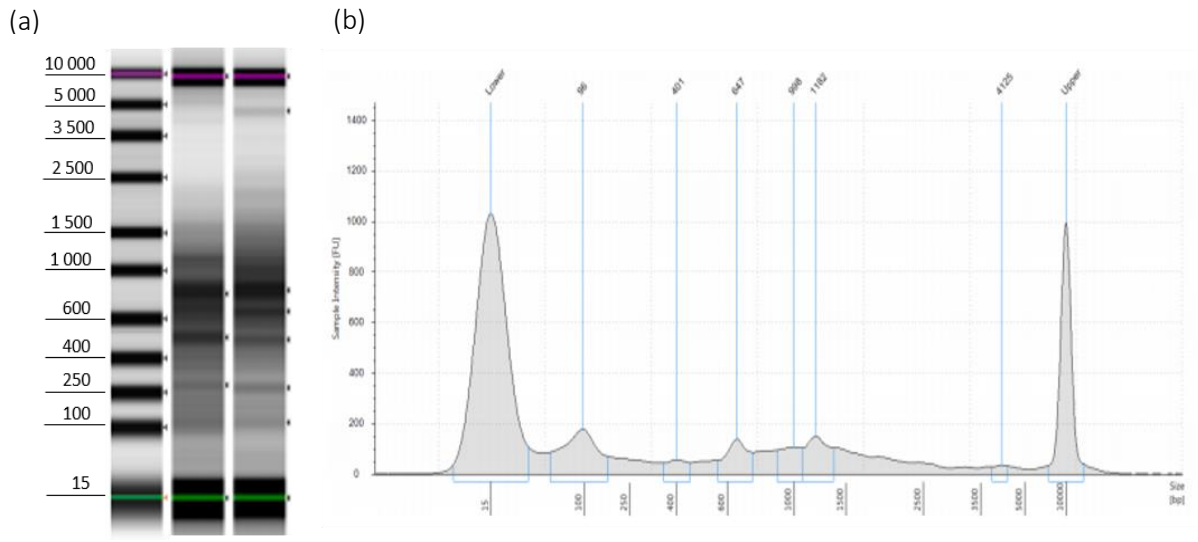


Figure 5.2. Gel image and electropherogram from the TapeStation 2200 system. Representative results generated from a single-cell cDNA sample with fragment sizes ranging from 600 – 2500 bp in length. (a) Gel image (ladder and two representative samples) and (b) electropherogram showing size distribution of the amplified cDNA fragments. The two peaks present on both ends of the electropherogram represent the upper (10 000 bp, right) and lower (15 bp, left) limit of detection of High Sensitivity D5000 ScreenTape.

5.2.3.2. Library preparation and sequencing

Several methods have been developed for single-cell RNA-seq and can be categorised into full-length and tag-based methods. For the purpose of cell-type discovery and characterisation of tissue composition, both full-length and tag-based methods are suitable. We opted for the full-length method in this study. Full-length methods aim to achieve uniform read coverage and increase the number of mappable reads, which enables isoform discovery, detection of splicing events and single nucleotide polymorphism (SNP) identification.

Library preparation and sequencing were performed by the Agricultural Research Council (ARC) at Onderstepoort Vetenary Research Campus, Onderstepoort, Gauteng, South Africa, using the Nextera XT DNA Library Preparation Kit (Illumina, USA) as per the manufacturer's instructions. Single-cell libraries were sequenced using the Illumina Hiseq 2500 sequencing system (Illumina, USA) in high output mode, obtaining paired-end reads of 125 bp in length. The average sequencing depth was 5.9×10^6 ($\pm 6.4 \times 10^6$) reads per single-cell sample.

5.2.4. Computational analysis of single-cell RNA-seq data

Several computational tools were used to analyse single-cell RNA-seq data for this study. These bioinformatics tools were used to remove sequencing adapters and poor-quality reads, assess the quality of the data, align reads to the human reference genome (GRCh38) and perform gene and transcript quantification. Further data analysis was performed in R (version 3.4.1) and RStudio (version 1.0.153) using the Seurat package (version 2.3.3). Seurat is an R package that enables exploration and analysis of single-cell data. Figure 5.3 illustrates the bioinformatics workflow used for this study. Each step of the workflow is discussed in more detail below.

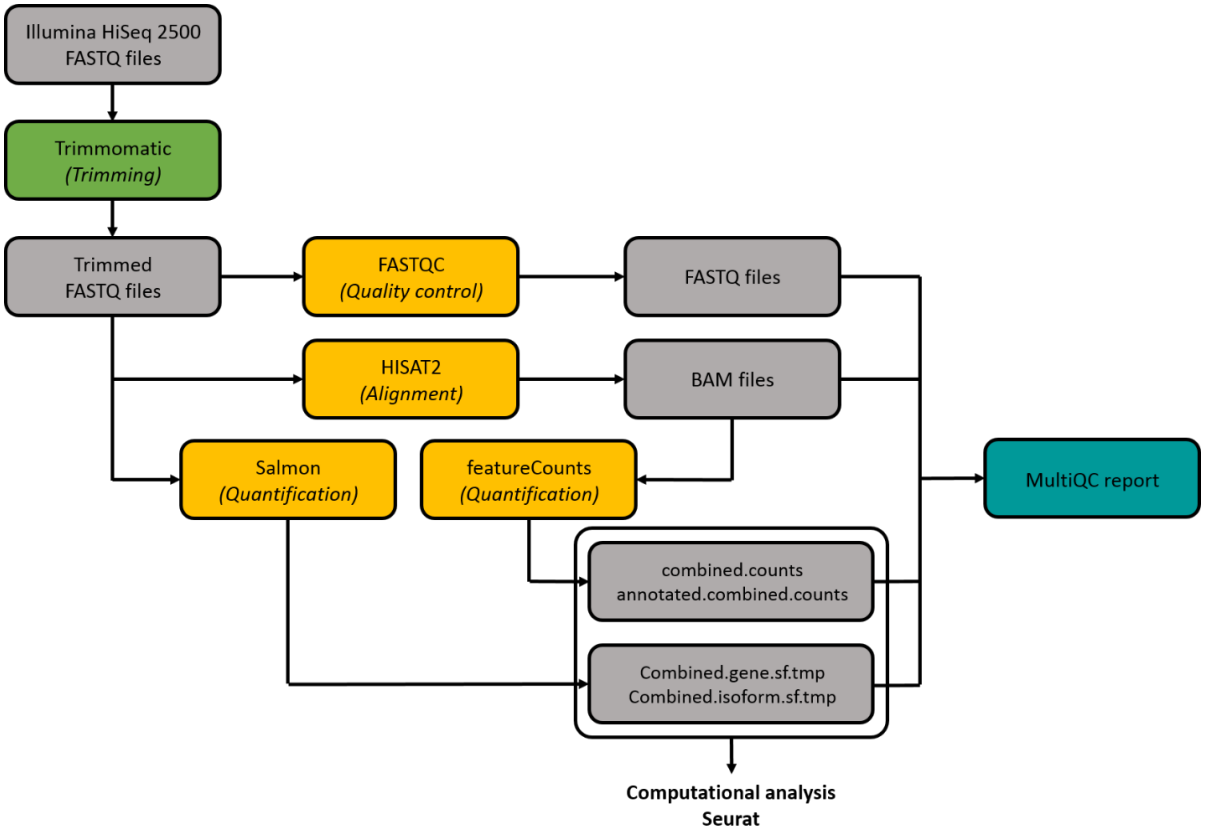


Figure 5.3. Bioinformatics workflow. A schematic illustration of the bioinformatics workflow that was used for the single-cell RNA-seq analysis.

5.2.4.1. Trimming of raw sequencing reads

The Illumina Hiseq 2500 sequencing system generated two FASTQ files (forward and reverse), containing genomic sequences of the cDNA fragments and the quality information for each base of the sequence. Poor-quality reads and the presence of sequence adapters could interfere with downstream analysis, therefore, trimming of raw sequencing data was required. Trimmomatic is a paired-end sensitive trimming tool that was used to remove Illumina Nextera sequencing adapters and poor-quality sequence reads (316). Reads with mean sequence quality (Phred) scores < 20 and length < 60 bp were removed from further analysis. Trimmed reads that passed the filter criteria were stored in two separate FASTQ files (forward and reverse) for further analysis.

5.2.4.2. The bcbio bioinformatics workflow

The bcbio-nextgen workflow was used to assess the quality of the data, align reads to the human reference genome (GRCh38) and perform gene and transcript quantification. This

community-developed bioinformatics pipeline was specifically designed for automated, high-throughput sequencing analysis (https://github.com/chapmanb/bcbio-nextgen). The already trimmed forward and reverse FASTQ files generated by Trimmomatic were used as input files for the bcbio workflow.

5.2.4.2.1. FastQC quality control

FastQC is a quality control tool for high-throughput sequencing data, providing information on the sequence counts, Phred scores, percentage GC content and percentage duplication for each sample. This information was viewed in the MultiQC report.

5.2.4.2.2. HISAT2 alignment to GRCh38

A human reference genome is a digital nucleic acid sequence database, assembled by scientists as a representative example of a species' set of genes. The latest human reference genome, released in December 2013, is GRCh38, and was used as a proxy for the human transcriptome. Single-cell RNA-seq reads were aligned to GRCh38, using the alignment tool HISAT2 (hierarchical indexing for spliced alignment of transcripts) (317). HISAT2 has increased speed, low use of random access memory (RAM) and an enhanced ability to detect spliced alignments when compared to other alignment tools available (317). Reads may map uniquely or could be multi-mapped reads. HISAT2 alignment returned binary alignment map (BAM) files, containing alignment data, and a BAMQC file containing information on the quality of the alignments and their position with regard to the reference genome. The BAMQC information was also compiled into the MultiQC report.

5.2.4.2.3. FeatureCounts for quantifying genes and transcripts

RNA-seq data allows for the quantification and comparison of gene and various gene-isoform (transcript) expression levels. Reads that aligned to the reference genome were quantified using featureCounts and Salmon. FeatureCounts is a read summarisation program that measures a particular gene's abundance by counting the number of reads per gene (318). The BAM files generated by HISAT2 were used as input to quantify the gene counts per samples. In parallel, transcript abundance was also calculated for each sample using Salmon. Several attempts have been made in recent years to develop tools that will allow for accurate

quantification of transcripts (319). However, accurate transcript reconstruction from short reads is difficult, and methods typically show substantial disagreement (320). Salmon is a fast, alignment-free transcript and gene quantification method (321). The transcript abundance level was indicated as the number of reads per length of all expressed transcripts per kilobase million (TPM values).

5.2.4.2.4. Output files

The bcbio workflow provided an output directory that contained both project- and sample-related files. The sample-related files included a BAM file that contained aligned and unaligned reads, the featureCounts output files and Salmon output files (including the TPM values) for each sample.

The project-related files contained the MultiQC report. The MultiQC program was used to gather all quality control metrics from the different tools (FASTQC, HISAT2, featureCounts and Salmon) used during the analysis to create an HTML MultiQC report. The project output files contained the gene count matrix (combined.gene.sf.tmp) and the transcript count matrix (combined.isoform.sf.tpm) generated by Salmon. Both the matrices contained counts normalised to the TPM values. FeatureCounts generated a combined.counts file and an annotated_combined.counts file. The combined.counts file containing the ENSEMBL IDs for the genes, and an annotated_combined.counts file containing both the ENSEMBL IDs and gene symbols.

There are currently two quantification approaches for RNA-seq data, gene- and transcript-level quantification. The bcbio workflow provided data for both gene- and transcript-level quantification. Gene-level quantification is more robust and actionable compared to transcript-level quantification (319). However, using gene-level quantification, can often mask the differences at the transcript-level. Soneson *et al.* (319) published a technical paper on gene- and transcript-level quantification. The authors showed that gene-level quantification data was more accurate and interpretable than transcript-level quantification data. However,

transcript-level data provided more in-depth information with regard to SNPs. This study only made use of gene-level quantification data. Thus the annotated_combined.counts file (mentioned above) was used for downstream analysis in R.

5.2.4.3. Pre-processing single-cell RNA-seq data

Quality control measures were applied in order to identify and remove poor-quality samples from downstream analysis. There is currently no consensus on the exact filter and quality control parameters that should be used for single-cell data. The filtering strategies and quality control parameters used are often based on the cells or tissues of interest (322,323). The most frequently used quality parameters for single-cell data include the number of reads per cell, the alignment percentage to the reference genome, and the percentage of reads aligning to mitochondrial genes or synthetic spike-in RNAs (322,323). Appendix B shows the scripts and functions that were used for pre-processing the data.

5.2.4.3.1. Sample quality

The MultiQC report was used to assess the quality of the data (Phred scores), the number of reads per sample and the percentage alignment to the human reference genome (GRCh38). Poor-quality samples with a Phred score < 20, the number of reads ≤ 100 000, the number of genes expressed < 200, and/or had < 60% alignment to the human reference genome, were removed from further analysis.

5.2.4.3.2. Genes and cells

Genes not expressed by three or more cells are unlikely to reach statistical significance between cell types (323) and were therefore removed from further analysis. If the cells expressed 10 000 genes or more, they were also removed. Although the Fluidigm system allows researchers to confirm capture of individual cells and only single cells were selected for downstream analysis, this parameter was only included to ensure that accidentally missed doublets were not included in downstream analysis.

5.2.4.3.3. Mitochondrial and pseudogenes

High expression of mitochondrial-derived genes is generally an indication of cellular stress (323), therefore percentage expression of mitochondrial genes was calculated for each sample. Samples with > 10% mitochondrial genes, were excluded from further analysis. The mitochondrial genes were also regressed out of the single-cell data since they can cause interference in differential gene expression analysis.

Pseudogenes are ubiquitous and abundant in the genome and are often transcribed into RNA. Although pseudogenes were once considered to be 'junk DNA', it has been recognised that these genes play essential roles in gene regulation of the parent genes (324). However, the role of pseudogenes in the genome is still poorly understood. Upon initial analysis of the data, several pseudogenes were detected in some of the clusters that slightly skewed the data, and they were therefore removed.

5.2.4.3.4. Cell cycle genes

The cell cycle is a highly regulated process during cell growth and proliferation. The four cell cycle phases are G1 phase, in which the cells are preparing for DNA synthesis; S phase, in which the DNA is duplicated; G2 phase, in which the cells are preparing for mitosis; and M phase, in which cell division occurs. Cell cycle genes introduce large within-cell-type heterogeneity that can compromise the interpretation of differential gene expression between cell types (325).

The cell cycle genes were identified for each sample and scores were calculated based on the relative expression of S and G2/M gene-sets (326). These scores were used to identify the phase of the cell cycle for each individual cell. Cell cycle genes were also removed from downstream analysis to ensure that these genes did not interfere with downstream differential gene expression data.

5.2.4.3.5. Normalisation

All cells were normalised using a global-scale normalisation approach, where the gene expression measurement for each cell was normalised to the total expression and then multiplied by a scale factor of 10 000 (default); the resulting values were then

log-transformed. Dividing the gene expression measurement of each cell by the total expression allowed a relative measure of counts for all expression values. Log-transformation is commonly used for gene expression data.

5.2.4.3.6. Batch effect

Initial clustering attempts showed clustering of cells based on the two different kits (SMARTer and SMART-Seq V4) used for the lysis, reverse transcription and cDNA amplification. Batch effect refers to technical variation or non-biological differences in gene expression measurements of different samples (327).

5.2.4.4. Clustering and differential gene expression

The batch effect observed was corrected by importing each dataset (data generated from the SMARTer and SMART-Seq kits) into R as individual datasets. A Seurat object was created for the two datasets mentioned above, using the Seurat package. Seurat objects are S4-type matrices which store all information associated with corresponding datasets. The Seurat object for each dataset was pre-processed as described earlier, after which variable genes for each dataset were identified. A multi-set canonical correlation analysis (CCA) was performed on the variable genes present in both datasets, identifying common sources of variation. Canonical correlation (CC) vectors, where cells were embedded into a shared low dimensional space to identify conserved cell types across the different conditions, were also identified.

The most significant dimensions (CCs) were identified, using the MetageneBicorPlot and Heat maps (CC1 – CC12), for downstream analysis. This allowed for the exploration of sources of heterogeneity in the dataset and was useful in identifying the significant CCs to be used for downstream analysis. The heat maps showed the top 20 differentially expressed genes identified by each CC for each cell. Yellow indicates highly expressed genes (positive values), while purple indicates genes expressed at low levels (negative values). In some cases, differentially expressed genes were not distinct on the heat maps. In such cases, the CC was considered non-significant and was excluded from downstream analysis. The MetageneBicorPlot (Figure 5.8) measured the correlation strength for each CC between the different datasets.

Cell clusters were identified using a shared nearest neighbour (SNN) modularity optimisation-based clustering algorithm (328). Non-linear dimensional reduction (t-distributed stochastic neighbour embedding; t-SNE) was used to run dimensional reduction on the aligned dataset, and t-SNE plots were used to visualise the clusters. The statistically significant differentially expressed genes were identified between different clusters.

5.2.4.5. Statistical analysis

Experimental data are represented as mean ± SD. A non-parametric Wilcoxon rank sum test with Bonferroni correction was applied and a P-value of ≤ 0.05 was considered significant.

5.3. RESULTS

5.3.1. Purity and capture efficiency

Six independent UCB samples were collected for this part of the study. Environmental changes, such as removing cells from a living organism, might affect gene expression patterns to accommodate the changing environment. The changes that occur are not well described and are therefore poorly understood. To ensure limited environmental and transcriptomic changes as much as is possible, samples were processed on the day of collection. A purity check was performed prior to sorting of the desired cell population (viable, $CD34^+$) to determine the sort purity and is shown in Table 5.7. The mean sort purity for the six UCB samples was 96.9% (± 1.58%).

Confirmation of capture was done using a light microscope. Each capture site was individually viewed to verify the capture of a single cell (Figure 5.4). Capture was considered unsuccessful if the capture site was empty, if more than one cell was present in the capture site or surrounding wings, or if the cells had an unusual morphology. The capture efficiency (Table 5.7) for each IFC plate was determined by dividing the number of captured cells by the number of capture sites (96). The mean capture efficiency for the six IFC plates used was 46% (± 12.1%). Using the Fluidigm C1 system together with three small and three medium IFC plates, a total of 266 single $CD34^+$ HSPCs were successfully captured from six individual donors.

(a) (b)

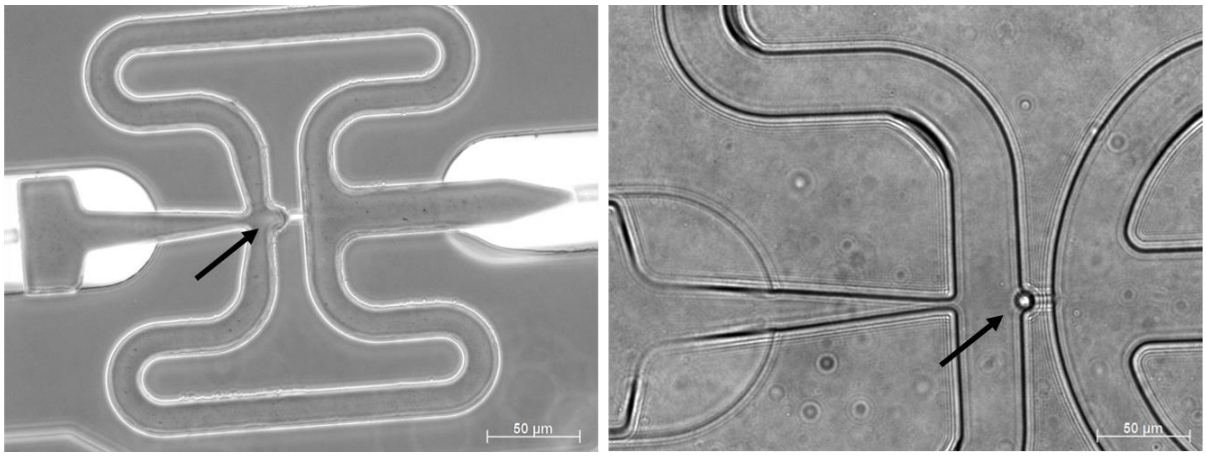


Figure 5.4. Single-cell capture. Representative cell capture in the IFC plate, where (a) is an empty capture site and (b) is a single captured cell.

Table 5.7. UCB units collected for the single-cell transcriptome analysis of CD34$^+$ HSPCs.

Sample	Sort purity (%)	Capture efficiency (%)	IFC plate	Kit
CB070515-01	95.0	26	Small	SMARTer
CB241115	95.5	60	Medium	SMARTer
CB221116	98.3	45	Small	SMARTer
CB021216	98.9	40	Small	SMARTer
CB080217	97.6	50	Medium	SMART-Seq v4
CB240317	96.2	50	Medium	SMART-Seq v4

5.3.2. cDNA concentration and size distribution

Picogreen quantification of the single-cell cDNA products revealed that a single HSPC on average yielded approximately 0.51 (± 0.18) ng/µL cDNA. Slight differences in the amount of cDNA were observed between cells captured using the small vs. the medium IFC plates. Interestingly, HSPCs captured using the medium IFC plates on average yielded lower concentrations of cDNA, 0.36 ng/µL (± 0.02) compared to HSPCs captured using the small IFC plates, 0.66 ng/µL (± 0.07). Only cells with cDNA concentrations ≥ 0.1 ng/µL were sent for sequencing. This may have introduced some bias in that HSPCs that were transcriptomically more active were sequenced, while the transcriptomically inactive HSPCs were excluded at this point.

The cDNA fragment size and distribution from only a few randomly selected cells from each IFC plate was determined using the TapeStation 2200. Figure 5.5 shows the electropherograms generated by the Tapestation Controller Software. Figure 5.5a represents the ladder, while Figures 5.5b and 5.5c represent the bulk RNA tube control, no template control (NTC) and tube control (cells), respectively. The tube controls were run in parallel with each IFC plate run (Appendix A). Neither the negative nor the positive control (Figure 5.5b) showed the presence of contaminants or cDNA product. The absence of cDNA product in the positive control might be the result of the small amounts of RNA present in $CD34^+$ HSPCs, which was difficult to extract manually. Figures 5.5d – h represent the amplified cDNA of a single cell from randomly selected samples. The single-cell sample represented in Figure 5.5f contained no cDNA product. Each sample was identified based on the biological replicate and the position of the sample in the 96-well harvest plate. All sample fragments (if present) showed a size distribution of between 400 – 2500 bp. The majority of fragments were distributed in the range of 600 – 1500 bp. Some samples showed the presence of shorter fragments (100 – 300 bp), which could indicate sample degradation.

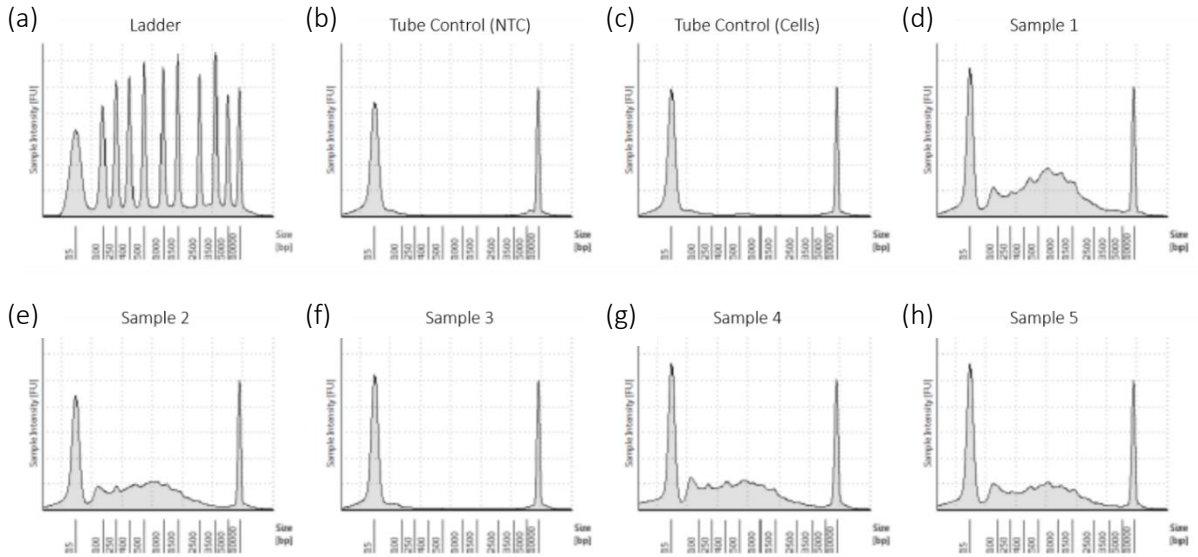


Figure 5.5. Electropherogram from the TapeStation 2200 system. Representative traces of (a) the D5000 ScreenTape Ladder, (b) Tube control (NTC), (c) Tube control (Cells); (d – h) each represent the amplified cDNA of a single $CD34^+$ HSPC identified by their position in the 96-well harvest plate. No cDNA product was present in Sample (f). The other samples (d, e, g and h) showed a normal distribution of fragment sizes, illustrating good quality single cell cDNA.

5.3.3. Quality control and pre-processing of the single-cell RNA-seq data

A total of 176 single cells provided good quality data after trimming. The quality control criteria set out in the methods section resulted in the removal of several samples either after trimming or based on the MultiQC report. The complete MultiQC report can be viewed at (http://wiki.bi.up.ac.za/j_multiqc). The Phred scores for all samples were above 25.

The average number of genes expressed in a single HSPC was between 5000 and 7000 (Figure 5.6a). None of the cells expressed less than 200 genes and only a few cells expressed more than 10 000 genes. The majority of the $CD34^+$ HSPCs expressed mitochondrial genes (Figure 5.6b), ranging from between 2.5 – 10%. Cells that expressed more than 10 000 genes and/or more than 10% mitochondrial genes were removed from further analysis.

The presence of cell cycle genes was also investigated since it has been shown that these genes can affect downstream differential gene expression (325). Figure 5.6c and 5.6d illustrate the S and G2/M phase scores for each cell, respectively. The scores were calculated based on the expression of S and G2/M phase gene markers. Cells with low scores in both were either not cycling or were in the G1 phase. The majority of $CD34^+$ HSPCs showed low S and G2/M phase scores. Bone marrow-derived LT-HSCs have been shown to be cell cycle dormant (329,330). A recently published study showed that 30 – 40% of UCB-derived $CD34^+$ HSPCs were in G0 phase, while less than 1% were in the S and G2M phases (309). The mitochondrial and cell cycle genes were regressed out of the datasets to prevent interference with differential gene expression.

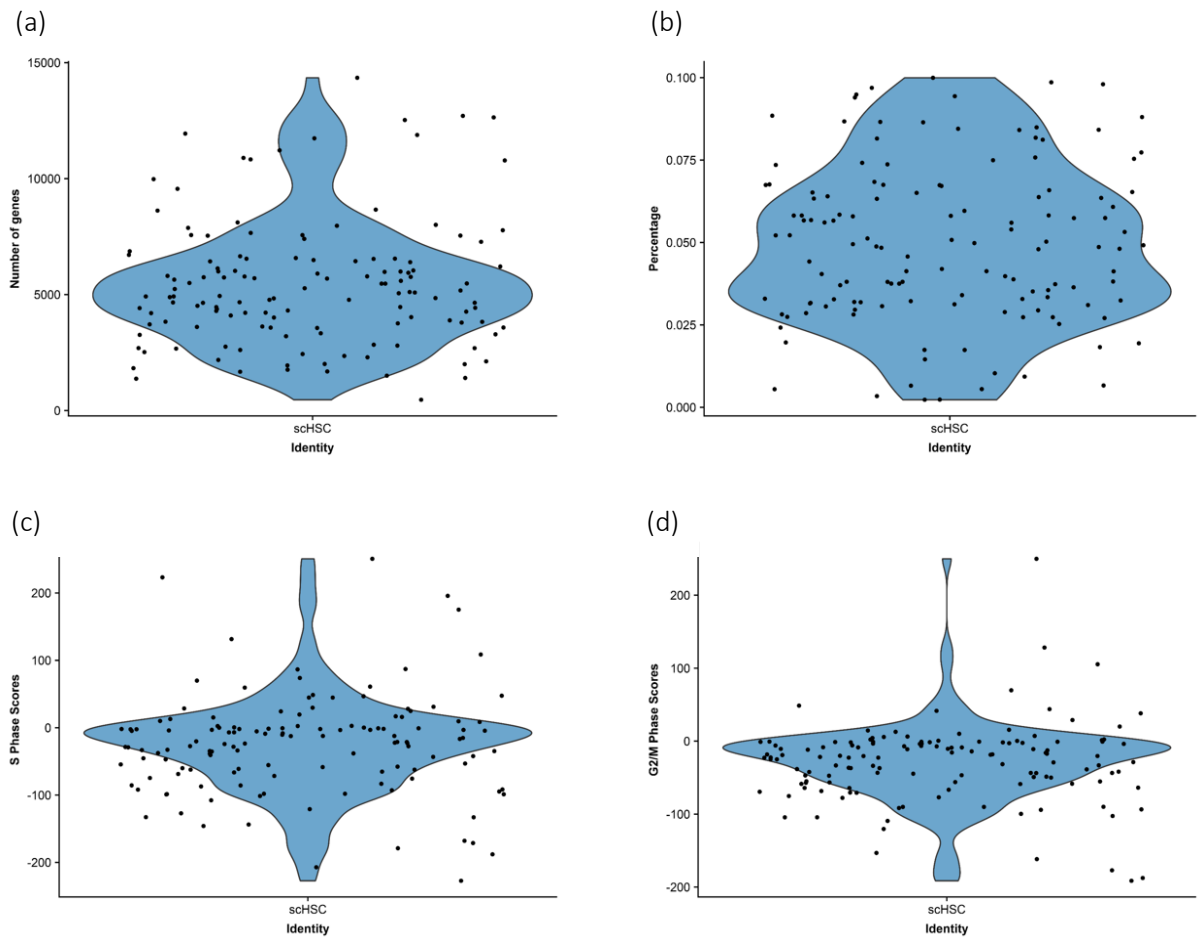


Figure 5.6. Quality control parameters assessed during pre-processing of the data. Violin plots showing (a) the number of genes, (b) the percentage mitochondrial genes, (c) the S and (d) G2/M phase scores per cell. Each dot represents a single cell. Phase scores ≤ 0 in (c) and (d) indicate that no S or G2/M genes were expressed.

5.3.4. Clustering and differential gene expression

Preliminary clustering of the CD34$^+$ HSPCs identified two clearly defined clusters (Figure 5.7). However, in-depth analysis of the individual clusters revealed that the cells were clustering according to the different kits that were used to obtain cDNA (SMARTer and SMART-Seq V4). We ascribed the clustering observed to batch effect, since batch effect refers to differences observed in gene expression measurements as a result of technical or non-biological variation (327).

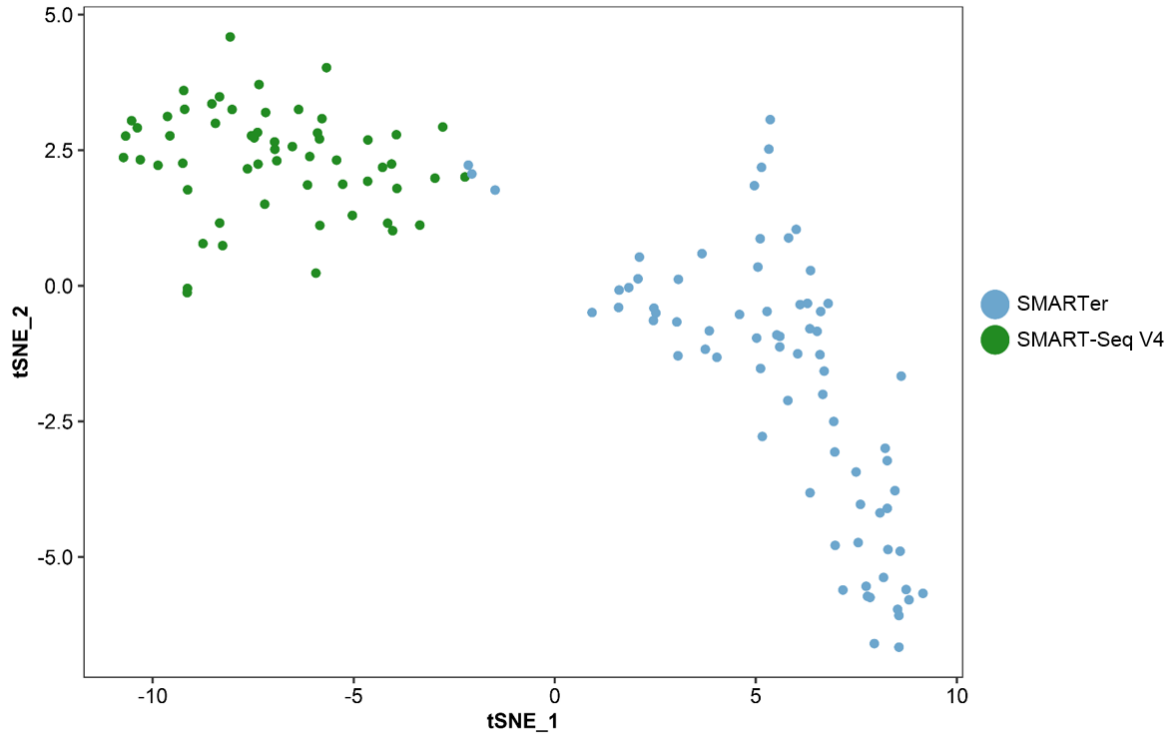


Figure 5.7. A t-SNE plot illustrating the batch effect. t-SNE clustering identified two clusters of CD34$^+$ HSPCs clustering according to the two different kits, SMARTer (blue) and SMART-Seq V4 (green), that were used for lysis, reverse transcription and cDNA amplification.

A computational strategy introduced by Butler *et al.* (331) was used to correct for the batch effect observed by importing each dataset (data generated from the SMARTer and SMART-Seq V4 kits) into R as individual datasets. A Seurat object was created separately for the individual datasets. The SMARTer Seurat object contained 58 735 genes across 107 samples, while the SMART-Seq V4 Seurat object contained 58 735 genes across 69 samples. Before the datasets could be aligned for further analysis, a set of variable genes was identified for each dataset. Seurat identified 13 406 variable genes for the SMARTer dataset and 12 234 variable genes for the SMART-Seq V4 dataset. The variable genes present in both datasets (12 146) were identified and used for downstream analysis.

The first step in the Seurat alignment was to perform CCA, to identify common sources of variation, after which the cells were combined into a shared low dimensional space. A MetageneBicorPlot and CC heatmaps were used to choose the CCs to use for downstream analysis. The MetageneBicorPlot showed the shared correlation strength versus the canonical

correlation vectors (CCs) to determine the useful dimensions. A drop-in correlation strength indicated the CCs to use for alignment. The two curves from each dataset reached a saturation point (flattened) at approximately five CCs (Figure 5.8).

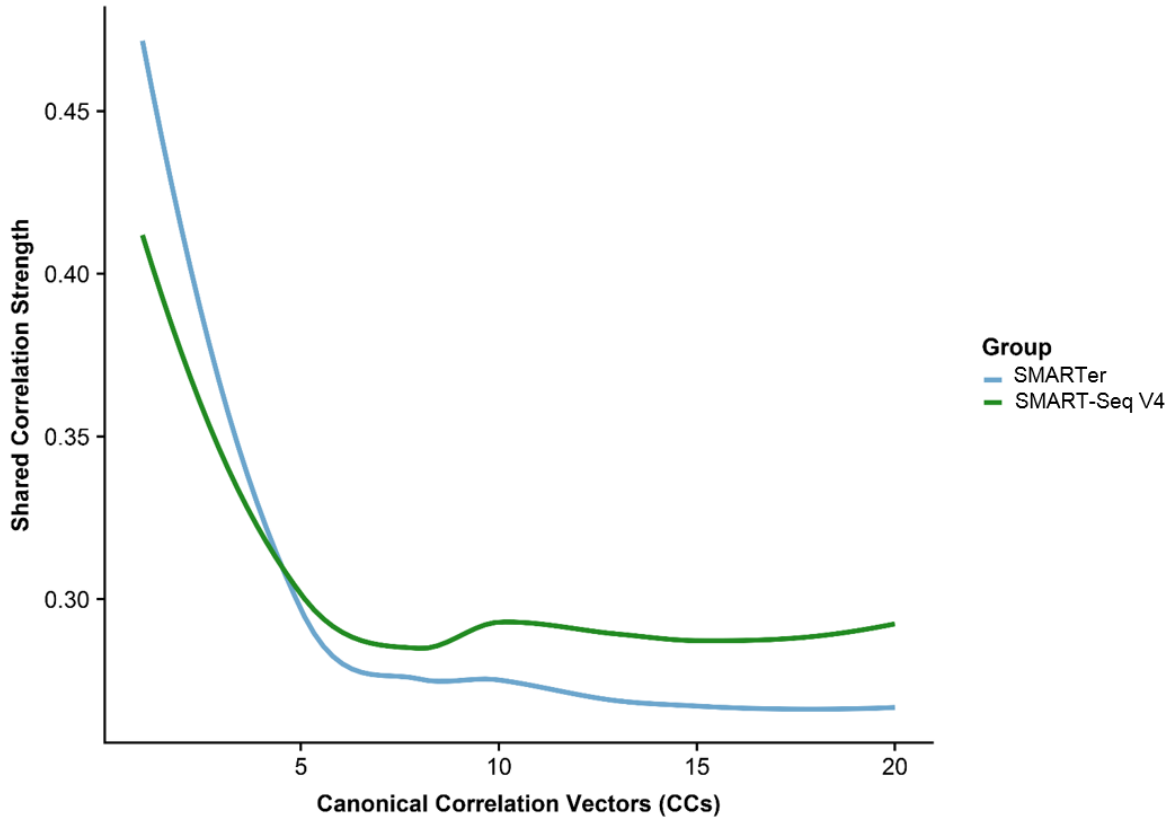


Figure 5.8. A MetageneBicorPlot. A CC Bicor saturation plot showing the correlation strength vs. canonical correlation vectors (CCs) between the SMARTer and SMART-Seq V4 kits. The green and blue coloured lines represent the SMARTer and SMART-Seq V4 datasets, respectively. This plot was used to determine the dimensions to proceed with for alignment of the datasets in Seurat.

The genes associated with each of the 12 CC heatmaps could also be used in combination with the MetageneBicorPlot to indicate the number of CCs to use. An output file containing 12 CC heatmaps (CC1 – CC12) (Figure 5.9) showed the top 20 differentially expressed genes (rows) identified for each cell (columns) in the individual CCs. Differentially expressed genes could be visualised for CC1 – CC5. There were no clear gene expression patterns observed from CC6 – CC12. Based on the results obtained from the MetageneBicorPlot (Figure 5.8) and the CC heatmaps (Figure 5.9), CC1 – CC5 were used to align the datasets.

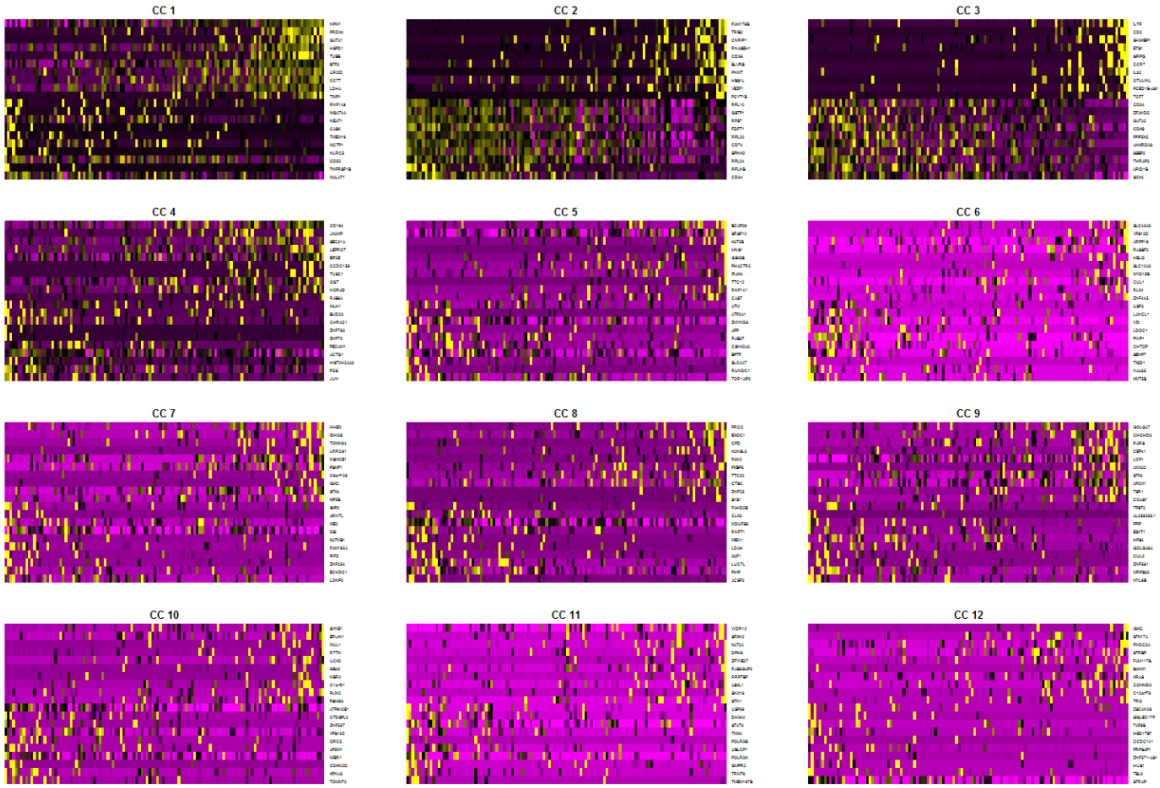


Figure 5.9. Heatmaps for CC1 – CC12. Heatmaps showing the correlation strength for the first 12 CCs of the different datasets. The top 20 differentially expressed genes (rows) identified for each cell (columns) in the individual CCs. Yellow indicates highly expressed genes (positive values), while purple indicates genes expressed at a low level (negative values).

The CCA subspaces for the two datasets were aligned using CC1 – CC5 to make the dimensions more comparable for clustering. t-SNE clustering was performed after alignment to confirm that the batch effect was indeed corrected. The t-SNE plot (Figure 5.10) showed no clear separation between the datasets as before.

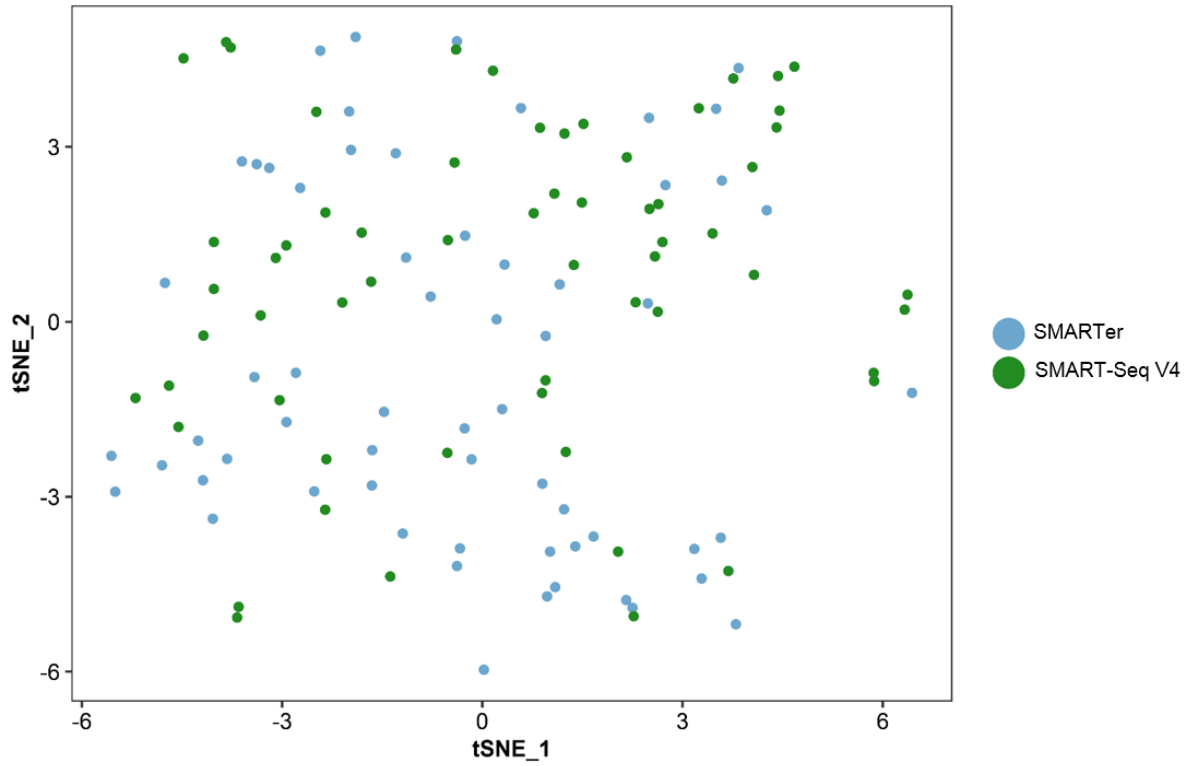


Figure 5.10. A t-SNE plot illustrating the removal of the batch effect. The two different coloured dots represent the two different datasets based on the two kits that were used. After adjusting for batch effect, no clear separation between the two datasets was observed.

The final integrated analysis to identify clusters was conducted on 122 cells from six independent donors. The results were visualised using a t-SNE plot. The function used to cluster the cells enabled adjustment of the resolution. An increase in the resolution, resulted in an increase in the number of clusters identified.

A resolution of 0.7 and 0.8 identified two transcriptionally distinct populations (Figure 5.11a). Differentially expressed genes were identified for each population. The differentially expressed genes reported had a logFC ≥ 0.5 with a p-value of ≤ 0.01. A positive logFC identified upregulated genes per cluster. Cells within each cluster were characterised using heat maps to illustrate the different gene expression patterns. The heatmaps showed distinct gene expression profiles, confirming the presence of two sub-populations present within the $CD34^{+}$ HSPC population (Figure 5.11b). Table 5.8 shows the top 20 statistically significant differentially expressed genes identified in the two clusters. Appendix C shows the full list of all the genes expressed in the two clusters identified at a resolution of 0.7 and 0.8.

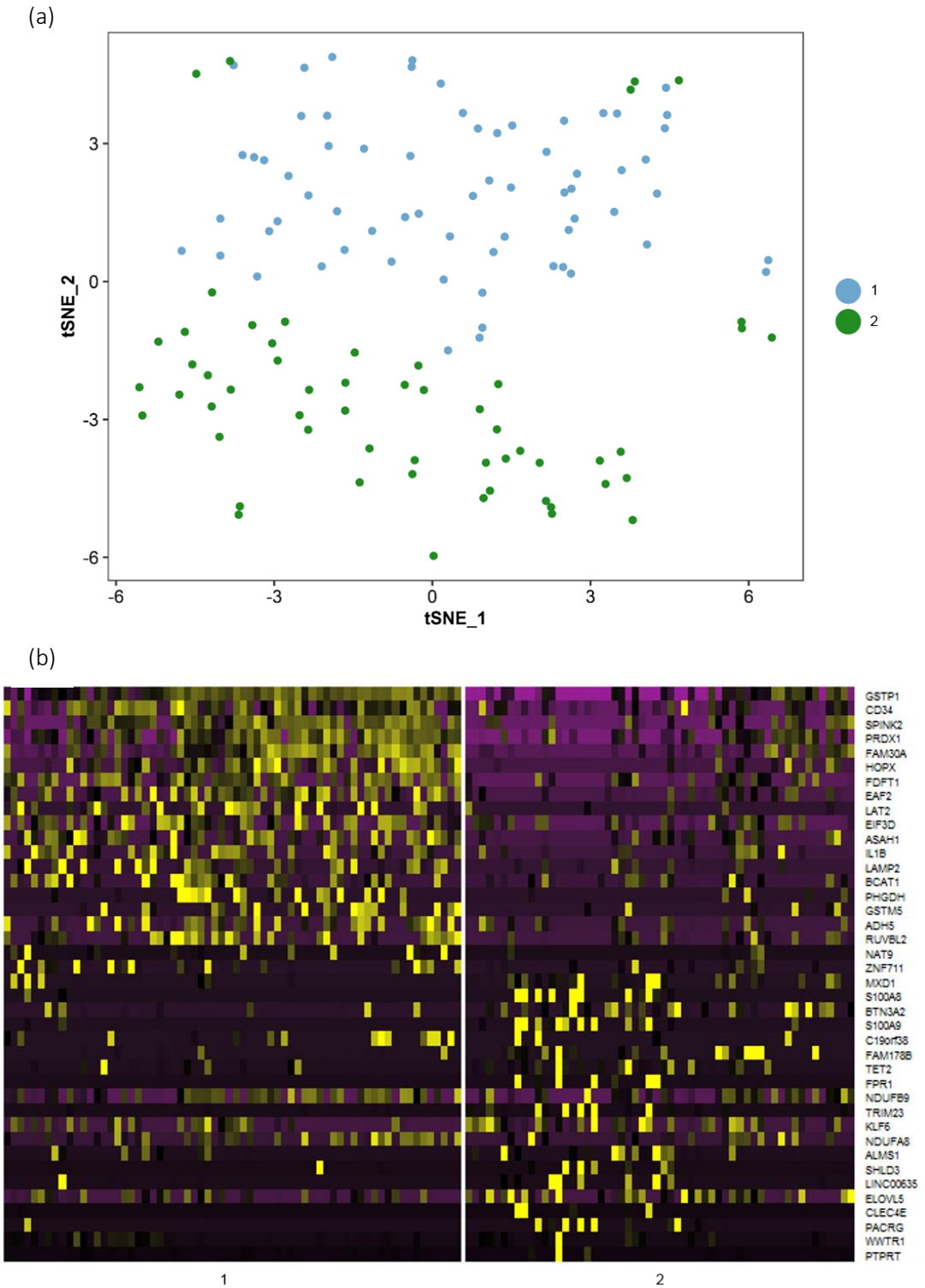


Figure 5.11. t-SNE plot and heatmap identifying three sub-populations. (a) t-SNE clustering identified three clusters at a resolution of 0.8. (b) A heatmap illustrating the top 20 differentially expressed genes (based on the average logFC) in each of the three clusters identified, showing distinct gene expression profiles. Yellow indicates highly expressed genes (positive values), while purple indicates lowly expressed genes (negative values).

Table 5.8. The top 20 significantly differentially expressed genes in the two clusters identified.

Cluster 1			Cluster 2		
Gene	LogFC	P-value	Gene	LogFC	P-value
GSTP1	1.24	3.60E-11	MXD1	2.26	2.38E-04
CD34	0.91	3.23E-08	S100A8	3.29	2.47E-03
SPINK2	1.00	5.39E-08	BTN3A2	1.31	3.91E-03
PRDX1	0.87	1.93E-07	S100A9	2.99	4.00E-03
FAM30A	1.06	2.53E-07	C19orf38	1.14	4.27E-03
HOPX	1.00	3.66E-07	FAM178B	1.07	7.12E-03
MDK	0.86	8.83E-07	TET2	1.01	1.03E-02
FDFT1	0.99	1.22E-06	FPR1	2.19	1.31E-02
EAF2	1.04	1.48E-06	NDUFB9	1.21	1.56E-02
LAT2	1.05	5.86E-06	TRIM23	1.68	1.58E-02
EIF3D	0.90	5.34E-05	KLF6	1.09	1.77E-02
ASAH1	0.89	6.83E-05	NDUFA8	1.59	2.30E-02
IL1B	0.89	2.46E-04	ALMS1	2.27	2.38E-02
LAMP2	1.16	2.62E-04	SHLD3	1.28	2.43E-02
BCAT1	0.86	1.01E-03	LINC00635	1.66	2.49E-02
PHGDH	0.87	1.47E-03	ELOVL5	1.12	2.55E-02
GSTM5	0.95	2.27E-03	CLEC4E	1.52	2.88E-02
ADH5	0.87	3.81E-03	PACRG	1.08	3.27E-02
RUVBL2	0.93	4.20E-03	WWTR1	1.01	4.59E-02
NAT9	1.87	1.44E-02	PRUNE1	1.01	4.59E-02

LogFC = Log fold change

Cluster 1 expressed *CD34*, *SPINK2*, *CD74* and MHC-related genes. *CD34* and *SPINK2* are generally expressed in HSCs (332), while the expression of *CD74* and MHCII-related genes, such as *HLA-DPA1*, is indicative of the presence of dendritic cell progenitors in Cluster 1 (333). Cluster 2 expressed genes, such as *HBB* and *S100A8*, which indicates the presence of erythroid and neutrophil progenitors in a single cluster (333). We recognise that cells belonging to different clusters might share gene expression patterns to some extent. However, the presence of genes frequently used to identify distinct populations of HSPCs were observed in single clusters using resolutions 0.7 and 0.8. Therefore, we decided to increase the resolution to identify additional clusters and determine whether the HSPCs cluster according to the predefined classification of populations previously identified within the $CD34^{+}$ HSPC population. The resolution was increased to 0.9 and 1.0. Both these resolutions identified four transcriptionally distinct populations (Figure 5.12a). Differentially expressed genes were identified for each of the four populations. As stated previously, the differentially expressed genes reported had a logFC ≥ 0.5 with a p-value ≤ 0.05. The heatmap showed distinct gene expression profiles, confirming the presence of four (Figure 5.12b) sub-populations. Table 5.9 shows the top 20 statistically significant differentially expressed genes identified in the four clusters. Appendix C shows the full list of all the genes expressed in the four clusters identified at a resolution of 0.9 and 1.0.

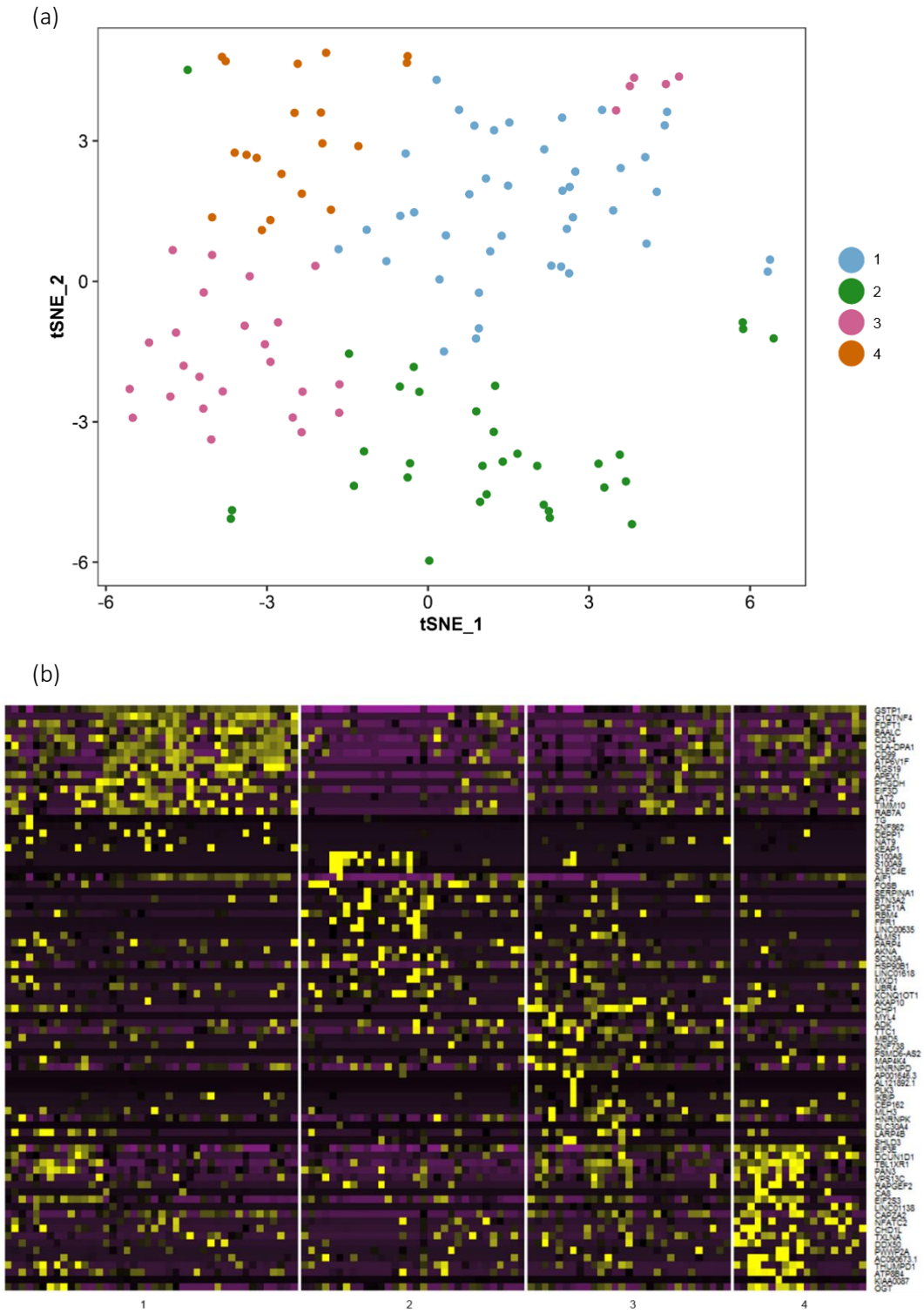


Figure 5.12. t-SNE plot and heatmap identifying four sub-populations. (a) t-SNE clustering identified four clusters at both resolutions 0.9 and 1.0. (b) A heatmap illustrating the top 20 differentially expressed genes in each of the four clusters identified, showing distinct gene expression profiles.

Table 5.9. The top 20 significantly differentially expressed genes in the four clusters identified.

Cluster 1			Cluster 2			Cluster 3			Cluster 4		
Gene	LogFC	P-value	Gene	LogFC	P-value	Gene	LogFC	P-value	Gene	LogFC	P-value
GSTP1	0.93	7.59E-08	S100A8	2.85	1.30E-06	AKAP10	1.23	8.99E-04	EIF3E	1.11	6.70E-05
C1QTNF4	0.90	2.63E-07	S100A9	3.28	3.00E-05	CHP1	1.06	4.48E-03	DCUN1D1	1.01	6.74E-05
FDFT1	0.89	4.40E-07	CLEC4E	1.98	2.25E-04	MYL4	2.89	5.04E-03	TBL1XR1	1.10	3.22E-04
BAALC	0.85	3.34E-06	MEFV	0.82	5.59E-03	ADK	1.26	5.29E-03	PAN3	1.07	6.26E-03
CD34	0.87	4.44E-06	FPR1	2.69	1.08E-02	TTC1	1.33	5.84E-03	VPS13C	1.06	8.51E-03
HLA-DPA1	0.82	5.62E-06	NRIP1	0.78	1.13E-02	MBNL2	1.01	6.16E-03	PRR14L	0.91	9.70E-03
CD99	0.82	1.24E-05	RAC2	0.82	1.18E-02	MBD5	2.52	8.61E-03	RAPGEF2	0.95	1.03E-02
ATP6V1F	0.79	1.89E-05	EXOC6B	0.79	1.33E-02	ZNF738	1.10	9.85E-03	CA8	0.98	1.27E-02
RGS19	0.79	3.77E-05	ANAPC16	1.08	1.34E-02	PSMD6-AS2	1.21	1.03E-02	EIF2S3	1.03	1.50E-02
APEX1	0.81	1.07E-04	LINC00635	2.16	1.51E-02	MAP4K4	1.24	1.23E-02	LINC01138	1.18	1.77E-02
PHGDH	0.92	3.21E-04	SLC43A2	0.87	1.64E-02	HNRNPD	1.61	1.32E-02	CAPZA2	0.97	2.00E-02
EIF3D	0.85	3.40E-04	ALMS1	2.58	2.02E-02	PLK3	1.81	1.67E-02	NFATC2	1.11	2.17E-02
LAT2	1.08	5.51E-04	NPAT	0.85	2.24E-02	IKBIP	1.20	2.00E-02	CHD1L	1.00	2.18E-02
TIMM10	0.91	1.60E-03	SCN3A	2.10	3.14E-02	CEP162	1.19	2.13E-02	TXLNA	1.01	2.50E-02
RAB7A	0.79	3.87E-03	IL6ST	0.78	3.38E-02	MLH3	1.03	2.43E-02	DDX50	1.42	2.61E-02
TG	1.42	5.26E-03	CSF1R	0.93	3.42E-02	HNRNPK	2.06	2.84E-02	PWWP2A	1.12	3.37E-02
ZNF862	1.02	6.31E-03	MXD1	2.58	3.82E-02	SLC30A4	1.14	3.81E-02	THUMPD1	1.04	3.97E-02
DEPP1	0.82	1.11E-02	NCF2	0.79	3.98E-02	LARP4B	1.14	3.91E-02	ATP8B4	1.10	3.98E-02
NAT9	2.10	1.29E-02	UBR4	1.92	4.22E-02	SHLD3	1.77	4.05E-02	KIAA0087	1.02	4.21E-02
KEAP1	0.81	3.20E-02	KCNQ1OT1	2.99	4.80E-02	MS4A2	1.00	4.72E-02	OGT	1.07	4.32E-02

Several cell surface markers are classically used to identify HSPCs. UCB-derived HSPCs are typically $CD34^+$ (*CD34*), $CD38^-$ (*CD38*), $CD45^{dim}$ (*PTPRC*), $CD133^+$ (*PROM1*), $CD117^+$ (*KIT*) and $CD90^+$ (*THY1*) at the protein level. Since the correlation between transcript and protein level expression is not well known for these markers, the transcript level expression of these markers was explored. The previously identified cell surface markers (Figure 5.13) were mapped onto the original t-SNE plot to display the clusters in which the markers were expressed. Blue indicates high expression, while grey indicates no expression. The CD34 surface marker expression was used to identify and isolate $CD34^+$ HSPCs from UCB. Despite the high sort purity obtained for each sample (> 95%, Table 5.7), the results indicated that CD34 expression was absent at the transcript level in some populations.

Figure 5.12a shows that *CD34* was expressed in a subset of the cells (Clusters 1 and 4), while Clusters 2 and 3 showed no expression of *CD34*. CD38 expression at the protein level is generally negative on HSCs and early progenitors, while being positive on later progenitors (232). Figure 5.13b shows expression of *CD38* on only a few cells, while most of the cells expressed low or no *CD38* at the transcript level. Previous reports have indicated a positive correlation between CD38 transcript presence and protein level expression (334,335). The CD45 antigen, encoded by the *PTPRC* gene, is a leukocyte marker typically intermediately expressed on freshly isolated $CD34^+$ cells. Figure 5.13c shows high expression of *PTPRC* in several random cells. The majority of the cells, however, expressed intermediate/low levels of *PTPRC*, which is similar to its protein expression. As mentioned in previous chapters, CD133 expression at the protein-level, together with CD34 expression, identifies progenitor cells with multipotent potential (249). The CD133 antigen is encoded by the *PROM1* gene. *PROM1* expression was more pronounced in Clusters 1 and 4, while being almost absent in Clusters 2 and 3, similar to CD34 expression (Figure 5.13d). The CD117 and CD90 antigens are encoded by the *KIT* and *THY1* genes, respectively, and play important roles in HSPC function (47). CD117 and CD90 proteins are typically expressed at low levels on $CD34^+$ HSPCs (47). Transcript-level expression of *KIT* and *THY1* was low or absent in the majority of the HSPCs analysed in this study (Figure 5.13e and Figure 5.13f).

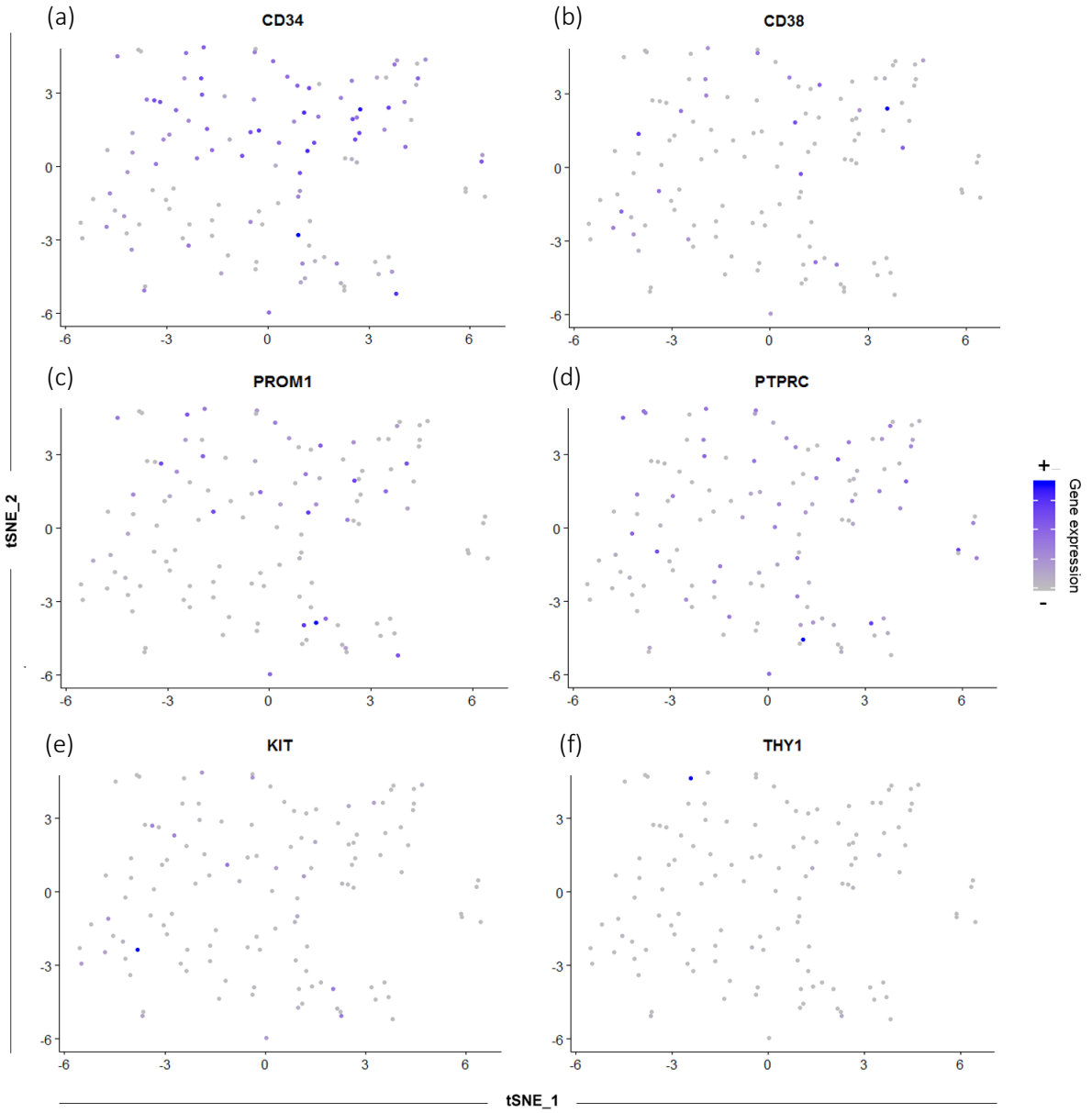


Figure 5.13. t-SNE plots showing expression of various HSPC-associated surface markers at the transcript level. Expression of (a) *CD34*, (b) *CD38*, (c) PTPRC (CD45), (d) *PROM1* (CD133), (e) *KIT* (CD117) and (f) *THY1* (CD90) is represented as blue dots while no expression is represented as grey dots. Purple dots indicate intermediate expression.

Several transcription factors have been identified to play an important role in hematopoiesis; these include *GATA1*, *GATA2*, *EGR1* and *SPI1*. The GATA-binding factor 1 and 2 (*GATA1* and *GATA2*) belong to the family of GATA transcription factors. In humans, the transcription factor PU.1 is encoded by the *SPI1* gene and has also been shown to be involved in hematopoietic development (336). A direct interaction between PU.1 and GATA1 has previously been described (337,338). Early growth response 1 (EGR1) is encoded by the *EGR1* gene and has been shown to play a role in proliferation and localisation of HSPCs in the bone marrow (339). The abovementioned transcription factors were also mapped onto the original t-SNE plot to display the clusters in which these genes were expressed. Figure 5.14 illustrates the HSPC-associated transcription factor expression in the various clusters identified.

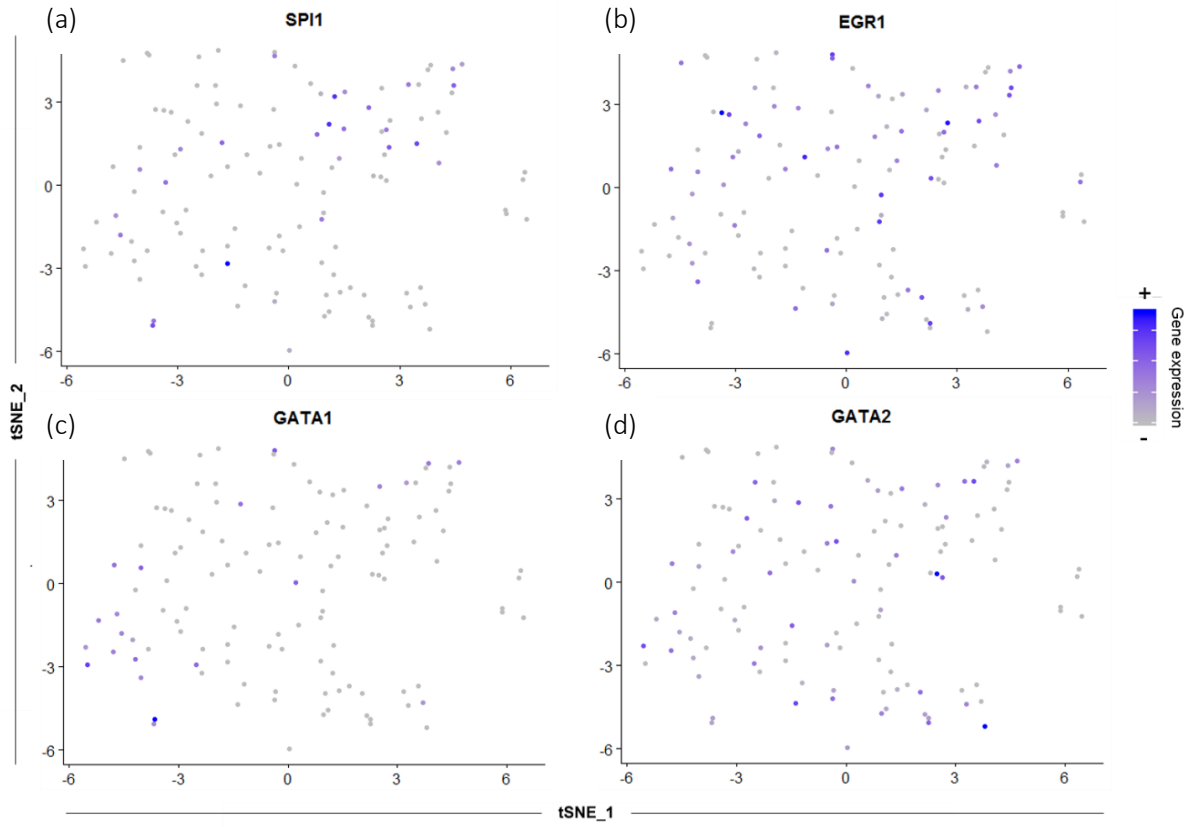


Figure 5.14. t-SNE plots showing expression of various HSPC-associated transcription factors. Expression of (a) *SPI1,* (b) *EGR1,* (c) *GATA1* and (d) *GATA2* is represented as blue dots while no expression is represented as grey dots.

5.4. DISCUSSION AND CONCLUSION

CD34 is a well-established marker for HSPCs and is used clinically to predict engraftment success. It is well accepted that $CD34^+$ HSPCs are extremely heterogeneous and encompass primitive HSCs as well as early and late progenitor cells. Single-cell transcriptome analysis is a powerful tool to study the heterogeneity within cell populations. The purpose of this study was therefore to characterise the heterogeneity (or lack thereof) of $CD34^+$ HSPCs isolated from UCB using single-cell transcriptome analysis.

The results from this study revealed the absence of CD34 expression in some populations at the transcript level (Clusters 2 and 3). The exact correlation between CD34 transcript and protein level expression is not well understood. Hittinger *et al.* (340) showed a positive correlation of 68.2% between CD34 transcript and protein level expression in $Sca^+/CD34^+$ murine cells. The phenomenon of stochastic variation in gene expression has been repeatedly reported and suggests that genes can transition from being transcriptionally active to being transcriptionally inactive (341). This would explain the variation observed between cell surface CD34 expression used to sort $CD34^+$ HSPCs and the lack of CD34 expression at the transcript level noticed during RNA-seq analysis.

Four sub-populations within the $CD34^+$ HSPC population were identified in this study. A recent study by Zheng *et al.* (311) identified 10 sub-populations present within the $CD34^+$ HSPCs from UCB. However Zheng *et al.* analysed > 19 000 single cells in their study. Twenty sub-populations were identified in mobilised peripheral blood $CD34^+$ cells when > 44 000 single cells were analysed (333). Even though fewer sub-populations were identified in the current study, we nevertheless wanted to determine whether the four sub-populations resemble previously identified sub-populations of $CD34^+$ HSPCs.

A sub-population of cells (Cluster 1, Figure 5.12) was identified that expressed *CD34*, *SPINK2* and *HOPX*. The same sub-population also expressed *SELL* (CD62L) and *CD99*. *SPINK2* has been shown to be expressed in UCB-derived $CD34^+CD133^+$ cells using microarray technology (332). Zheng *et al.* (311) showed that *SELL* and *CD99* identify HSCs transitioning into MPPs. A study by Lai *et al.* (333) identified *CD34* and *SOX4* to be expressed in the MPP population; however,

SOX4 was not differentially expressed in the HSC or MPP (Cluster 1) populations identified in this study. *HOPX* is a stem cell marker and has been identified to play a role in murine HSPC biology (342). High *HOPX* expression is associated with normal HSPC expression patterns in humans (343). *CD74* and the MHCII-related gene *HLA-DPA1* were also expressed in Cluster 1. These two markers are frequently used to identify dendritic progenitors (333,344). This observation might indicate the presence of $CD34^+$ dendritic progenitors in Cluster 1. Several studies that included thousands of $CD34^+$ cells observed a distinct dendritic progenitor population that expressed *CD74* and MHCII-related genes (310,333,344). Co-expression of *CD74* and MHCII-related genes with transcription factors *IRF8* and *ID2* is also seen in dendritic cell progenitors (344). However, these transcription factors were not significantly upregulated in the current study. It is possible that by increasing the number of cells in the current study, that the HSC and dendritic progenitor populations might split to form two distinct clusters.

Cluster 2 expressed *S100A8*, *S100A9* and *FOSB,* which are frequently observed in granulocyte-macrophage progenitors (GMPs). Cluster 2 likewise expressed *CFS1R* and *IRF8*, which are associated with monocyte development (345,346). Similarly, Zhu *et al.* (347) identified a population of cells that expressed high levels of GMP-associated genes, such as *EGR1*, *FOSB*, *JUN*, *GATA2* and *GATA1*. Several other genes have been identified to be critical for neutrophil development, which include *GFI1*, *CEBPA*, *CEBPE*, *PER3*, and *ETS1* (347). Not all genes previously identified as GMP and neutrophil-associated were detected in our dataset. Lai *et al.* (333) identified a sub-population expressing *S100A9*, *CSF1R* and *CD177*. Paul *et al.* (344) likewise identified a neutrophil progenitor population, in which *S100A8* and *S100A9* were upregulated.

Cluster 3 expressed *GATA1*, *HBD*, *HBB* and *TFRC*, which are genes frequently expressed in erythroid progenitors (344). *GATA1* plays a well-established role in erythrocyte development (344), where it regulates the transcription of various genes involved in the maturation of erythrocytes and platelets. *GATA1* expression is usually suppressed in HSPCs, since its activation leads to loss of self-renewal capacity (348,349). We also found that *GATA1* was not expressed in Clusters 1 and 4 (Figure 5.14), which are thought to be more primitive populations

of HSPCs. Although the number of cells in this cluster was limited, the presence of these four genes provides strong evidence for erythroid development.

The *FLI1* gene was expressed in Cluster 4, which usually directs megakaryocyte-erythroid fate (350). Several other genes, such as *HOXA9*, *MEIS1* and *MLLT3* were also expressed in Cluster 4. The HOX genes and *MEIS1* are known regulators of HSPCs (351), while *MLLT3* has been shown to be involved in the self-renewal of HSPCs (352). Paul *et al.* (344) found that early progenitors expressed transcription factors, such as *GATA2* and *MEIS1*; however, reduced expression of these transcription factors was observed in more differentiated cells. A study by Tan *et al.* (353) obtained long-term HSCs from induced pluripotent stem cells (iPSCs) by transducing several transcription factors, including *HOXA9*. *HOXA9*-deficient HSPCs lacked the ability to repopulate the bone marrow of irradiated recipients in competitive transplantation assays (354). In a study by He *et al*. (332), *MLLT3,* among several other genes, showed the highest differential expression in $CD34^{+}CD133^{+}$ HSPCs from UCB. An ABC transporter protein gene, *ABCD4* was also expressed in Cluster 4. This gene has been reported to distinguish HSCs from non-HSCs (355).

CD164, also expressed in Cluster 4, has been shown to identify a primitive population of HSPCs (356,357). It has been suggested that CD164 facilitates $CD34^{+}$ cell adhesion in the bone marrow niche and also functions in preventing HSPC proliferation (356). Pellin *et al.* (310) likewise identified a population of HSPCs that expressed *CD164*. Phenotypic and functional assays identified two populations of CD164-expressing cells ($CD164^{high}$ and $CD164^{low}$) in human $CD34^{+}$ bone marrow cells. By combining CD164 with several classical HSPC flow cytometric markers, the authors confirmed that the $CD164^{high}$ population is enriched for primitive progenitors, megakaryocyte-erythroid progenitors (MEPs) and early common myeloid progenitors (CMPs) (310). $CD34^{+}CD164^{+}$ HSPCs constitute a population with greater myeloid and megakaryocyte-differentiating potential compared to $CD34^{+}CD90^{+}$ HSPCs. Pellin *et al.* (310) further suggested that this $CD34^{+}CD164^{+}$ HSPC population could serve as a potential new population for transplantation purposes. Since $CD34^{+}CD90^{+}$ cells are typically rare in UCB HSPCs (47), the CD164 marker might be a better marker to use in identifying early HSPCs for future *in vitro* and *in vivo* experiments. The presence of CD164 as a cell surface marker on a sub-population of these cells allows us to identify and isolate these cells for future experiments.

The Fluidigm C1 system was used to capture single cells for this study. The main limitations include the limited range of cell sizes that one is able to capture using the various C1 IFC plates. We overcame this limitation by using small and medium sized IFC plates to capture the $CD34^+$ HSPCs. Also, the maximum number of cells that can be captured at a time (96) is limited, and with a capture efficiency of 46% on average for the HSPCs (Table 5.7) it will be challenging to increase the cell numbers in order to be competitive in the single-cell field. Another limitation is the high cost associated with the consumables for this system, especially if these experiments need to be repeated several times in order to reach the cell numbers required in single-cell publications. Even though SMART technologies drastically improved the coverage of the transcriptome and had superior sensitivity over other methods at the time of development, there were several limitations (358). Lower read coverage towards the 5' end, under-representation of transcripts with a higher GC-content and the effect of complex RNA structures insurmountable by the DNA polymerase were some limitations observed with the SMART technologies (359,360). A recent publications compared frequently used methods with regard to sensitivity, accuracy and cost efficiency (361). Smart-Seq showed increased sensitivity and reproducibility, as a result of increased mappable reads. It is important to keep in mind that sensitivity is also a function of sequence depth. At a sequencing depth of 1×10^6 reads/sample there is somewhat of a plateau in sensitivity and the sequenced mRNA complexity reaches a level of saturation (361).

In conclusion, a total of 122 cells (out of 234 sequenced) were analysed for this particular study and four definitive clusters were identified. The clusters identified and the genes expressed within the various clusters are in strong agreement with previously defined clusters. We identified two populations of earlier undifferentiated HSPCs (Cluster 1 and 4). Another population identified GMP gene signatures and genes associated with neutrophil development (Cluster 2) and lastly, an erythrocyte progenitor population (Cluster 3) was identified. Both the neutrophil and erythroid sub-populations form part of the myeloid lineage during HSPC differentiation. Our data therefore suggest that the transcriptomic profiles of progenitor cells, especially late myeloid progenitors, are sufficiently distinct from early, primitive HSPC populations to allow resolution of these populations with a limited number of single cells

analysed. However, a clear pitfall of this study is the limited number of cells analysed compared to other similar studies, which did not allow discovery of the total extent of heterogeneity within the $CD34^+$ HSPC population.

Previous knowledge regarding HSPCs was based on flow cytometry analysis and bulk transcriptome analysis. Although limitations remain to be overcome, single-cell technologies are improving rapidly and are starting to provide a detailed atlas of the gene expression patterns of all cell types in the human body. Although this study was unable to define the extent of heterogeneity within the $CD34^+$ HSPC population, several other studies have explored this heterogeneity. These cells are heterogeneous in their differentiation potential and their functional abilities, and the heterogeneity is not exclusive to the $CD34^+$ population but extends to the sub-populations within the already heterogeneous $CD34^+$ HSPC population. Even though the hierarchical model of hematopoiesis has been studied for many years, single-cell technologies are unravelling the classic HSPC hierarchy, which is more complex than previously assumed. A recent study on single-cell transplantation revealed a population of myeloid-restricted progenitors with long-term repopulating ability, a concept that was disregarded before the advent of single-cell technologies (32). Another study identified HSPCs expressing platelet-biased genes that have the ability to self-renew, which immediately put these cells at the top of the HSPC hierarchy (33). These two examples are only a fraction of the data and information emerging as a result of single-cell analysis. This fast-evolving technology provides a powerful tool to dissect cellular heterogeneity and uncover the uniqueness of each cell, and also to identify new markers and pathways that previously could not be identified. New emerging single-cell techniques enable the simultaneous study of multiple omics (genome, transcriptome and proteome) of the same cell. This will further contribute to and improve our understanding of the heterogeneity of cell populations in the future.

CHAPTER 6. HEMATOPOIETIC STEM AND PROGENITOR CELL SUSCEPTIBILITY TO HIV-1

6.1. INTRODUCTION

Hematopoietic stem cell therapies lend themselves to the development of gene therapy approaches, which integrate *ex vivo* gene transfer strategies into well-established HSCT practice to treat neoplastic, monogenic and infectious diseases. Clinical success in treating several monogenic disorders through genetic modification of autologous HSPCs support the feasibility, amongst other things, of using this approach to treat HIV.

HIV-1 depletes $CD4^+$ T-cells eventually leading to AIDS. AIDS refers to the advanced stages of HIV and is defined by the occurrence of HIV-associated opportunistic infections and/or malignancies. South Africa has a high prevalence of HIV, with more than 7×10^6 individuals living with HIV, which contributes significantly to the health burden in South Africa (135). HIV infects cells expressing CD4 in combination with the co-receptors CCR5 or CXCR4; however, the mechanism of infection is not limited to these receptors. Effective cART has reduced morbidity and mortality of AIDS-related diseases (136) and although cART can keep the virus from replicating, it cannot completely eliminate the virus and therefore needs to be maintained for life. This is expensive, requires lifelong compliance and has several associated side-effects, providing a rationale for the development of an HIV cure.

There has been increased interest in developing an HIV cure using HSPCs ever since Timothy Brown, also known as the 'Berlin patient', received an HSCT with cells from a CCR5 null donor and who has been HIV-free ever since (131,362). The CCR5-delta32 (Δ32) mutation results in the absence of CCR5 on the surface of $CD4^+$ cells, making these cells resistant to infection by an R5-tropic virus. CCR5 null donors are not readily available and the shortage in CCR5 null donors has resulted in research focusing on genetically modifying HSPCs by disrupting the CCR5 chemokine receptor. Reconstitution of the immune system after HSCT using genetically modified cells would result in an HIV-resistant hematopoietic system. The use of healthy

unrelated donors would be favourable for this approach; however, an HLA-matched donor is not always available. Autologous transplantation of genetically modified cells overcomes last-mentioned challenge and also reduces the risk of transplant-related complications.

HSPCs were thought to be resistant to HIV infection (185,186,188,363,364); however, several studies revealed that HSPCs express receptors and co-receptors required for HIV-1 entry, and are therefore susceptible to infection (183,189,190), at least theoretically. The susceptibility of HSPCs to HIV infection has implications for HSPC-based HIV gene therapy, as this could affect outcomes when using autologous cells. Direct infection of HSPCs could result in the genetic modification and transfusion of latently infected cells, which would produce latently infected progeny.

The purpose of this study was to determine whether HSPCs are susceptible to HIV-1 infection and/or whether a subset of HSPCs exists that is resistant to HIV-1 infection. Such a subset of HSPCs would be ideal for HSPC-based HIV gene therapies. Unfortunately, the aim of this study was not fully achieved due to several unforeseen circumstances and time constraints. This chapter will describe several HIV-related techniques and the problems encountered that resulted in this section not being fully completed.

6.2. MATERIALS AND METHODS

6.2.1. HIV-1 propagation using PBMCs

HIV is a single-stranded (ss) RNA virus that is reverse transcribed into double-stranded (ds) proviral DNA able to integrate into a host cell genome. The transcription and translation of HIV-related proteins is facilitated by host cell mechanisms, resulting in the production and packaging of HIV virions.

For experimental purposes, HIV was produced by exposing susceptible PBMCs from healthy uninfected donors to either HIV^+ PBMCs or HIV-containing supernatant. The production of virus depletes the number of viable cells as the virus replicates and buds from host cells, resulting in cell death. The depleted PBMC population was replaced with a fresh/new pool of PBMCs at

numerous timepoints to ensure constant viral production. Released virus was harvested as part of the cell culture supernatant. Viral production and titration were determined using a p24 enzyme-linked immunosorbent assay (ELISA) detection method (Section 6.2.2.1) and a functional GHOST cell assay (Section 6.2.2.3), respectively.

6.2.1.1. HIV-1 isolates

Primary HIV-1C isolates of three different tropisms (R5-tropic, R5X4-tropic and X4-tropic) were produced for this study (Table 6.1). These viral isolates were provided by Professor Lynn Morris from the National Institute for Communicable Disease (NICD) and were propagated using activated PBMCs according to an adapted Montefiori Laboratory protocol.

Table 6.1. Primary HIV-1C isolates used for this study.

HIV tropism	Isolate	Accession number	Reference
R5	CM1	AY505003	(162)
	Du422F	Ay043175.1	(365)
	Du156	Ay529660	(365)
	Du123	AF544007	(365)
	COT1	DQ235624	(147)
R5X4	CM9	AF411966	(162)
X4	SW7	AF411966	(162)

6.2.1.2. Processing of PBMCs

6.2.1.2.1. Peripheral blood collections

Peripheral blood was obtained from healthy donors at the Clinical Research Unit (CRU), Pathology Building, Prinshof Campus, University of Pretoria. Donors completed an informed consent document, after which 250 mL peripheral blood was collected into CPD collection bags (Tianhe Pharmaceuticals, China). HIV testing of the PBMC samples was performed as described in Chapter 3, Section 3.2.3. Ethical approval for collecting peripheral blood at the CRU and testing for HIV from consenting donors was granted by the Faculty of Health Sciences Research Ethics Committee at the University of Pretoria (*Protocol Number 204/2016*).

6.2.1.2.2. PBMC isolations from peripheral blood

PBMCs were isolated by density-gradient centrifugation using Histopaque-1077 as described in Chapter 3, Section 3.2.4. An additional step was included to remove platelets. Platelets serve as a possible binding site for infectious HIV particles which could compromise HIV production in PBMCs (366). Platelets were removed by centrifugation at 100 x g for 10 min without the brake. The supernatant was aspirated and the cells were washed by filling the tube to 30 mL with TP buffer and centrifugation at 300 x g for 10 min. Supernatant was aspirated and cells were resuspended in 10 mL TP buffer. A 100 µL aliquot was taken to determine absolute cell count.

6.2.1.2.3. PBMC counting by flow cytometry

Viability, absolute number of PBMCs and phenotype were determined by flow cytometry using the Gallios flow cytometer. A 100 µL aliquot was stained with 3 µL of 7AAD, CD4 FITC and CD45 KO and incubated for 15 min, protected from light. Flow-Count fluorospheres (100 µL) were added to each sample immediately prior to analysis on the flow cytometer. CD4 phenotyping was performed for parallel, but independent projects using the same samples. The Gallios flow cytometer filter configurations and channels used to detect the relevant fluorochromes are shown in Table 6.2. All flow cytometer analyses were performed using Beckman Coulter Kaluza Analysis Software (version 2.1) (Beckman Coulter, California, USA).

Table 6.2. Gallios flow cytometer laser and filter configurations for PBMC counting and phenotyping.

Laser	Filter	FL	Fluorochrome/Dye	Clone
488nm, 22mW	252/40	1	CD4 FITC	Okt4
	575/30	2	-	-
	620/30	3	Flow-Count fluorospheres	-
	695/30	4	7AAD	-
	755LP	5	-	-
638nM, 25mW	660/20	6	-	-
	725/20	7	-	-
	755LP	8	-	-
405nM, 40mW	450/40	9	-	-
	550/40	10	CD45 KO	J33

Viable cells were measured, using 7AAD viability dye [Excitation: 488 nm; Emission: 635/75 nm], which was gated on region [Not Beads] (Figure 6.1a). The [Not Beads] region is a NOT Flow-Count fluorospheres Boolean gate applied post-acquisition and used to exclude the Flow-Count fluorospheres from downstream flow cytometric plots during data analysis. "PBMCs" were measured using a FS vs. SS plot, which was gated on "Viable" (7AAD-negative) cells (Figure 6.1b). Leukocytes ($CD45^+$ cells) were measured in the FL10 channel [CD45 KO; Excitation: 398 nm; Emission: 528 nm] (Figure 6.1c), which was gated on "PBMCs". "Monocytes" ($CD45^+CD4^{dim}$) and "CD4 T-cells" ($CD45^+CD4^+$) were measured in the FL1 channel [CD4 FITC; Excitation: 488 nm; Emission: 504/41 nm] (Figure 6.1d), gated on "Leukocytes". Flow-Count fluorospheres were detected in the FL3 channel (Figure 6.1e) and used for PBMC counting. The absolute number of PBMCs present in each sample was calculated using the equation in Chapter 2, Section 2.2.3.1.

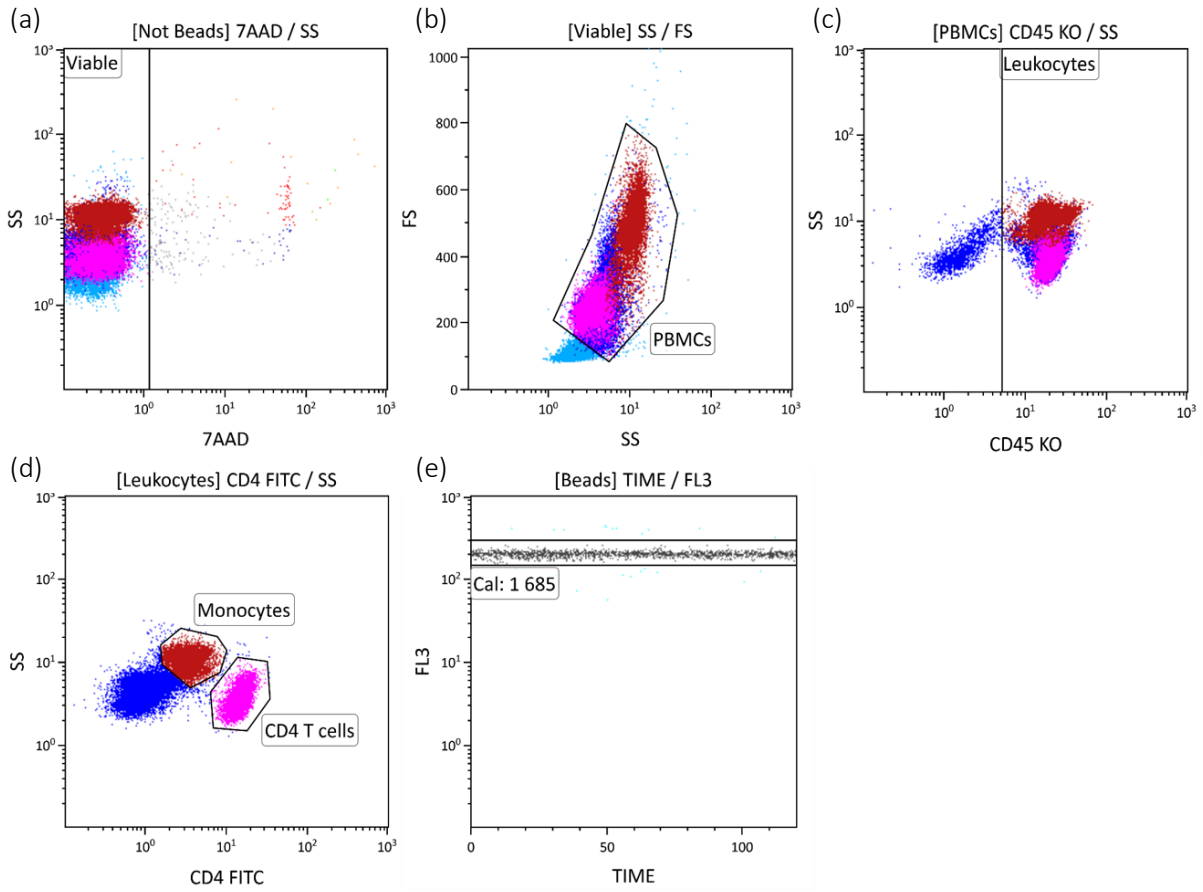


Figure 6.1. PBMC counting and phenotyping. A schematic illustration of the sequential gating strategy used for PBMC counting and phenotyping. Density plots showing (a) viable (7AAD-negative) cells (b) PBMCs, (c) cells stained with CD45 KO and (d) CD4 FITC monoclonal antibodies. Density plot showing the Flow-Count fluorospheres in "Cal" region.

6.2.1.2.4. PBMC cryopreservation

Once the number of PBMCs had been determined, the cells were centrifuged at 300 x g for 10 min and resuspended in 1 mL freezing medium (90% FBS and 10% DMSO) per 10 x 10^6 PBMCs. Aliquots were transferred to a 1.5 mL cryotubes, which were stored at -80°C in a Mr. Frosty™ freezing container for 24 hours, after which the cryotubes were transferred to liquid nitrogen until further use.

6.2.1.3. Propagation of HIV-1C isolates in PBMCs

Propagation of HIV-1C isolates was performed according to a modified Montefiori method for the production of HIV using activated PBMCs. The Montefiori protocol includes the various stages of HIV propagation: (i) activating PBMCs, (ii) infecting activated PBMCs with HIV-1C

isolates, (iii) feeding co-cultures with newly activated PBMCs and (iv) harvesting virus-containing supernatant and cells (Figure 6.2).

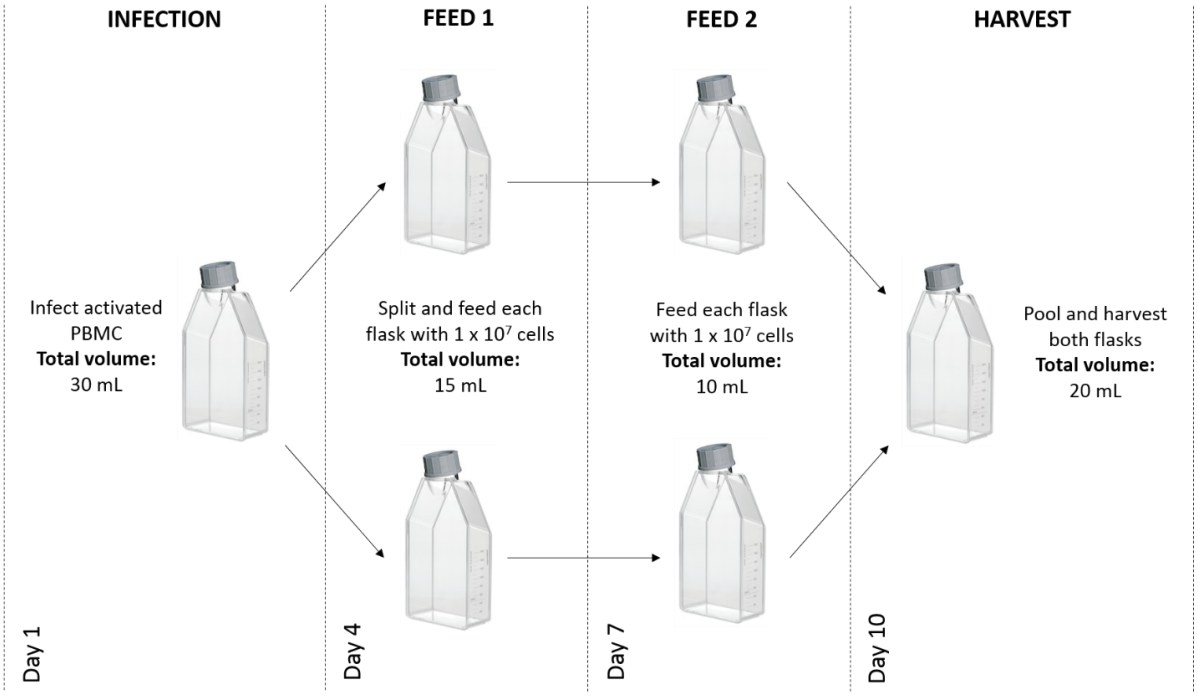


Figure 6.2. HIV-1 propagation in PBMCs. An illustration of our adapted method for propagating HIV-1C isolates in PBMCs. This protocol included activating PBMCs, infecting activated PBMCs with HIV-1C isolates (Day 1), feeding co-cultures with newly activated PBMCs (Day 4 and Day 7) and harvesting virus-containing supernatants and cells (Day 10). (Figure was created by Juanita Mellet).

6.2.1.3.1. PBMC activation

A combination of phytohemagglutinin (PHA-P, Sigma-Aldrich, USA) and T-cell growth factor, IL2 (Roche, Switzerland) was used to activate PBMCs. Mitogen stimulation of PBMCs with PHA-P upregulates the IL-2 receptor on the surfaces of PBMCs. Cryopreserved PBMCs were removed from liquid nitrogen and the tube content was diluted in 30 mL pre-warmed Roswell Park Memorial Institute-1640 medium (RPMI, Thermo Fisher Scientific, USA) and centrifuged at 300 x g for 10 min. The supernatant (containing DMSO) was aspirated and cells were resuspended in IL-2 growth medium (IL2GM, RPMI supplemented with 50 µL/mL from 10 units/mL IL-2 stock, 20% FBS and 2% pen/strep, 1 mL per 1 x 10^6 PBMCs) supplemented with 5 µg/mL PHA-P. PBMCs were activated three days prior to infection, since HIV requires activated cells for productive infection (367). PBMCs were incubated for 24 hours in culture

flasks that were kept in an upright position at 37°C and 5% CO_2. After a 24 hour incubation, the activation medium was removed by centrifugation at 300 x g for 10 min, and replaced with fresh IL2GM (1 mL per 1 x 10^6 PBMCs). Activated PBMCs were cultured for an additional 48 hours after medium change.

6.2.1.3.2. Infection of PBMCs with HIV-1

Activated PBMCs were centrifuged at 300 x g for 10 min and resuspended in 2 mL pre-warmed RPMI (2% pen/strep). Supernatant was aspirated and activated PBMCs were exposed to either HIV^+ PBMCs or HIV-containing supernatant for 2 hours in a small volume (usually 1 – 2 mL) at 37°C in 50 mL centrifuge tubes. The small volumes used for HIV infection are more likely to increase contact between the activated PBMCs and HIV. After incubation, the volume was increased to 30 mL with RPMI (2% pen/strep) for the HIV-exposed PBMCs and transferred to a T75 culture flask. Flasks were incubated in an upright position at 37°C and 5% CO_2 until the feeding cycle.

6.2.1.3.3. Feeding

Freshly activated PBMCs were added to the HIV/PBMC co-cultures on Day 4 and Day 7 of production to maintain a pool of viable cells for HIV to infect. PBMCs for each feeding were thawed and activated as described earlier in Section 7.2.1.3.1. On feeding days, the activated PBMCs were centrifuged at 300 x g for 10 min and cells were resuspended in 2 mL RPMI (2% pen/strep). The HIV/PBMC co-cultures from a single HIV-1C isolate were pooled in 50 mL centrifuge tubes and centrifuged at 300 x g for 10 min. The cell pellets were resuspended in 2 mL RPMI (2% pen/strep). The activated PBMCs and HIV/PBMC co-culture were pooled in a 50 mL tube, after which it was split in two and transferred to two T75 flasks. The volume in each T75 flask was increased according to the number of activated PBMCs present (1 mL per 1 x 10^6 PBMCs). The co-cultures were incubated in an upright position at 37°C and 5% CO_2.

6.2.1.3.4. Harvesting HIV-1

In preliminary experiments, we determined that the viral load is highest at Day 10 (Appendix D). Therefore, virus-containing supernatant was harvested on day 10. The HIV/PBMC co-cultures were transferred to 50 mL centrifuge tubes and centrifuged at 300 x g for 10 min.

The virus-containing supernatants from a single isolate were then collected and pooled in 50 mL centrifuge tubes. Cells and debris were removed from the supernatant by filtration using a 0.22 µm filter. Supernatant aliquots were stored at -80°C until further use. HIV^+ PBMCs were resuspended in 1 mL freezing medium and stored at -80°C until further use.

An appropriate control should be considered when working with virus-containing supernatant, since viral particles are harvested and frozen in supernatant containing an overabundance of growth factors. A control supernatant was therefore produced and used to distinguish between the effects of HIV and the effects of the supernatant. Production of PBMC control supernatant was performed in parallel with HIV-1 production, however, without exposure to HIV. Control supernatant was harvested every three days, filtered using a 0.22 µm filter and stored at -80°C.

6.2.2. HIV-1 detection methods

HIV can be detected through direct and indirect methods. Direct detection methods target the viral proteins, viral mRNA transcripts or viral RNA/DNA, while indirect detection methods target antibodies produced in response to HIV infection. Several HIV detection methods, such as PCR and p24 ELISA, were optimised to detect HIV. The GHOST cell assay was used to titrate the produced HIV.

6.2.2.1. p24 ELISA

A p24 ELISA using the Lenti-X™ Rapid Titre p24 ELISA kit (ClonTech, USA) was used to detect HIV and determine productive infection. The HIV *Gag* gene encodes core/capsid protein, p24, typically involved in virion assembly and often indicates productive infection (368). The p24 ELISA kit allows for the calculation of the number of viral particles present using a standard curve. It does, however, not provide information on whether the viral particles are infectious.

The p24 ELISA kit consists of 12 separable 8-well strips, which were used individually. Before each use, the 20X wash buffer (provided as part of the kit) was diluted to 1X wash buffer using distilled water. Positive p24 controls were prepared by diluting the p24 control (provided as part of the kit) to 12.5 pg/mL (low concentration control) and 200 pg/mL (high concentration

control) respectively, using RPMI. Figure 6.3 illustrates the 8-well strip layout, including the (1) negative control, (2) low (12.5 pg/mL) and (3) high (200 pg/mL) positive controls, (4, 6 and 8) undiluted and (5 and 7) 1/10 diluted samples. Undiluted samples can yield absorbance (OD) values outside the detection range of the spectrophotometer and therefore a 1/10 dilution of each sample was included.

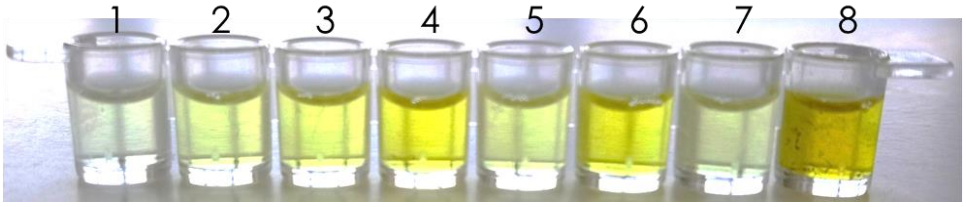


Figure 6.3. p24 ELISA sample layout. A schematic illustration of the sample layout of an 8-well Lenti-X Rapid p24 ELISA kit. Samples include: (1) negative control, (2) low (12.5 pg/mL) and (3) high (200 pg/mL) positive controls, (4, 6 and 8) undiluted and (5 and 7) 1/10 diluted samples.

A total of 200 µL of sample, followed by 20 µL of lysis buffer, was added to each well, and incubated for 60 min at 37°C. Upon lysis, p24 protein is released and binds to anti-p24 antibodies in the wells. After incubation, cell lysates were aspirated and wells were washed six times by adding 200 µL of wash buffer to each well. A 100 µL secondary biotinylated anti-p24 antibody was added to each well and incubated at 37°C for 60 min to tag the antibody-bound p24. After incubation, unbound antibody was aspirated and wells were washed again six times as described above. A streptavidin-HRP conjugate (100 µL) was added to each well and incubated at room temperature for 30 min. After incubation, wells were once again aspirated and washed six times using 200 µL of wash buffer. TMB substrate (100 µL) was added to each well, and incubated for 20 min, protected from light. After incubation, stop solution (100 µL) was added to each well. The absorbance was measured at 450 nm on a PowerWaveX spectrophotometer (Bio-Tek, USA). Absorbance values were normalised by subtracting the blank value from the measured values.

6.2.2.2. GHOST cells for HIV-1 titrations

Functional titrations of propagated HIV-1C isolates were performed using the GHOST cell infectivity assay, which makes use of the HIV indicator GHOST cell line.

6.2.2.2.1. GHOST/X4/R5 indicator cells

The GHOST/X4/R5 indicator cell line (GHOST cells) is derived from human osteosarcoma (HOS) cells expressing CD4 (Figure 6.4a), CCR5 (Figure 6.4b) and CXCR4 (Figure 6.4c). These cells contain a GFP under the control of an HIV-2 LTR promoter. Upon infection with HIV-1, the GFP gene is transactivated through Tat, which can be detected within 24 hours post-infection through a flow cytometric-based assay. These cells were used to determine the number of infectious viral particles per milliliter viral supernatant. The GHOST cell line was obtained through the National Institutes of Health (NIH) AIDS Reagent Program, Division of AIDS, NIAID, NIH from Dr V.N. KewalRamani and Dr D.R. Littman.

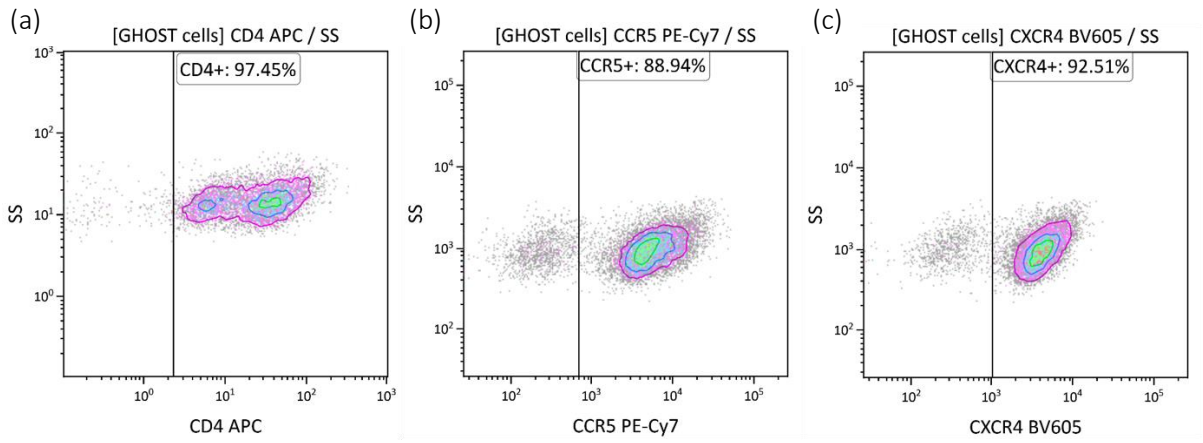


Figure 6.4. HIV-1 receptor and co-receptor expression on GHOST cells. A schematic illustration of the percentage HIV-1 receptor, (a) CD4 and co-receptor (b) CCR5 and (c) CXCR4 expression on GHOST cells stained with CD4 APC, CCR5 PE-Cy7 and CXCR4 BV605. The negative/positive region boundaries were created for each phenotypic marker using an isotypic controls.

6.2.2.2.1.1. Cell culture and maintenance of GHOST cells

GHOST cells were removed from liquid nitrogen and the tube content was diluted in 30 mL complete medium (DMEM, 10% FBS and 2% pen/strep). The sample was washed twice by centrifugation at 300 x g for 10 min. The supernatant was aspirated and the cells were resuspended in 8 mL complete medium supplemented with 1 µg/mL puromycin (Sigma-Aldrich, USA), 100 µg/mL hygromycin (Sigma-Aldrich, USA), and 500 µg/mL G418 (Roche, Switzerland), seeded in a T75 flask and incubated at 37°C and 5% CO_2 until confluent. These antibiotics ensure stable expression of CD4 and co-receptors, CCR5 and CXCR4. The

medium was replaced every three days by aspirating the medium and washing cells with 3 – 4 mL PBS (2% pen/strep) before adding fresh complete DMEM (8 mL) was added to the flask. The culture flasks were then incubated at 37°C and 5% CO_2 until confluent.

When the cells covered the majority of the surface area of the flask, the cells were considered to be confluent. Once the cells reached confluency, the medium was aspirated from the flask and the cells were rinsed twice with 3 – 4 mL PBS (2% pen/strep). Pre-warmed 0.25% Trypsin-EDTA (Thermo Fisher Scientific, USA) was added (4 mL) to the flask and incubated for 3 – 5 min, until cells were dissociated. For some experiments, GHOST cells were dissociated with a non-enzymatic dissociation buffer (Sigma-Aldrich, USA) instead of trypsin. To inactivate the trypsin or dissociation buffer, 4 mL complete DMEM was added to the flask. The dissociated cells were transferred to 15 mL centrifuge tubes. The flasks were rinsed twice with 2 – 3 mL of PBS (2% pen/strep) and the supernatant added to the cell suspension. The cells were centrifuged at 300 x g for 10 min after which the supernatant was aspirated. Cells were resuspended in 2 mL complete DMEM for each T75 flask used. The absolute number of GHOST cells was determined by flow cytometry. After counting, cells were either seeded again for culturing or were cryopreserved at 2 x 10^6 cells per cryovial in freezing medium (1 mL per 2 x 10^6 cells).

6.2.2.2.1.2. GHOST cell count and phenotype

GHOST cells were counted and phenotyped by flow cytometry using the Gallios flow cytometer. Table 6.3 indicates the laser and filter configurations of the Gallios flow cytometer. A 100 µL aliquot of GHOST cells was stained with 3 µL 7AAD. For phenotyping, two aliquots were stained with 3 µL of each antibody, one with CD4 APC (Beckman Coulter, USA), and the second with an appropriate isotype control, mouse IgG1 (Beckman Coulter, USA). Cells were incubated for 15 min, protected from light. Before analysis, 100 µL Flow-Count fluorospheres was added to the flow tube.

Table 6.3. Gallios flow cytometer laser and filter configurations for GHOST cell counting and phenotyping.

Laser	Filter	FL	Fluorochrome/Dye	Clone
488nm, 22mW	252/40	1	GFP	-
	575/30	2	-	-
	620/30	3	Flow-Count fluorospheres	-
	695/30	4	7AAD	-
	755LP	5	-	-
638nM, 25mW	660/20	6	CD4 APC	13B8.2
			CD4 APC isotype	mouse IgG_1

The sequential gating strategy used during data analysis is shown in Figure 6.5. The region "Not Beads" is a NOT Flow-Count fluorospheres Boolean gate applied post-acquisition and used to exclude the Flow-Count fluorospheres from downstream flow cytometric plots during data analysis. Viable (7AAD-negative) cells were measured, using 7AAD viability dye [Excitation: 488 nm; Emission: 635/75 nm], in the FL4 channel, which was gated on region "Not Beads" (Figure 6.5a). "GHOST cells" were measured using a FS vs. SS plot, which was gated on "Viable" (Figure 6.5b). "CD4+" GHOST cells were measured in the FL6 channel [Excitation: 650 nm; Emission: 660 nm], which was gated on "GHOST cells" (Figure 6.5d and Figure 6.5e). "GFP+" GHOST cells were measured in the FL1 channel [Excitation: 395 nm; Emission: 509 nm], which was gated on "GHOST cells" (Figure 6.5f and Figure 6.5g). For this protocol, Flow-Count fluorospheres were detected in the FL3 channel (Figure 6.5c). The number of GHOST cells in a sample was calculated using the equation in Chapter 2, Section 2.2.3.1.

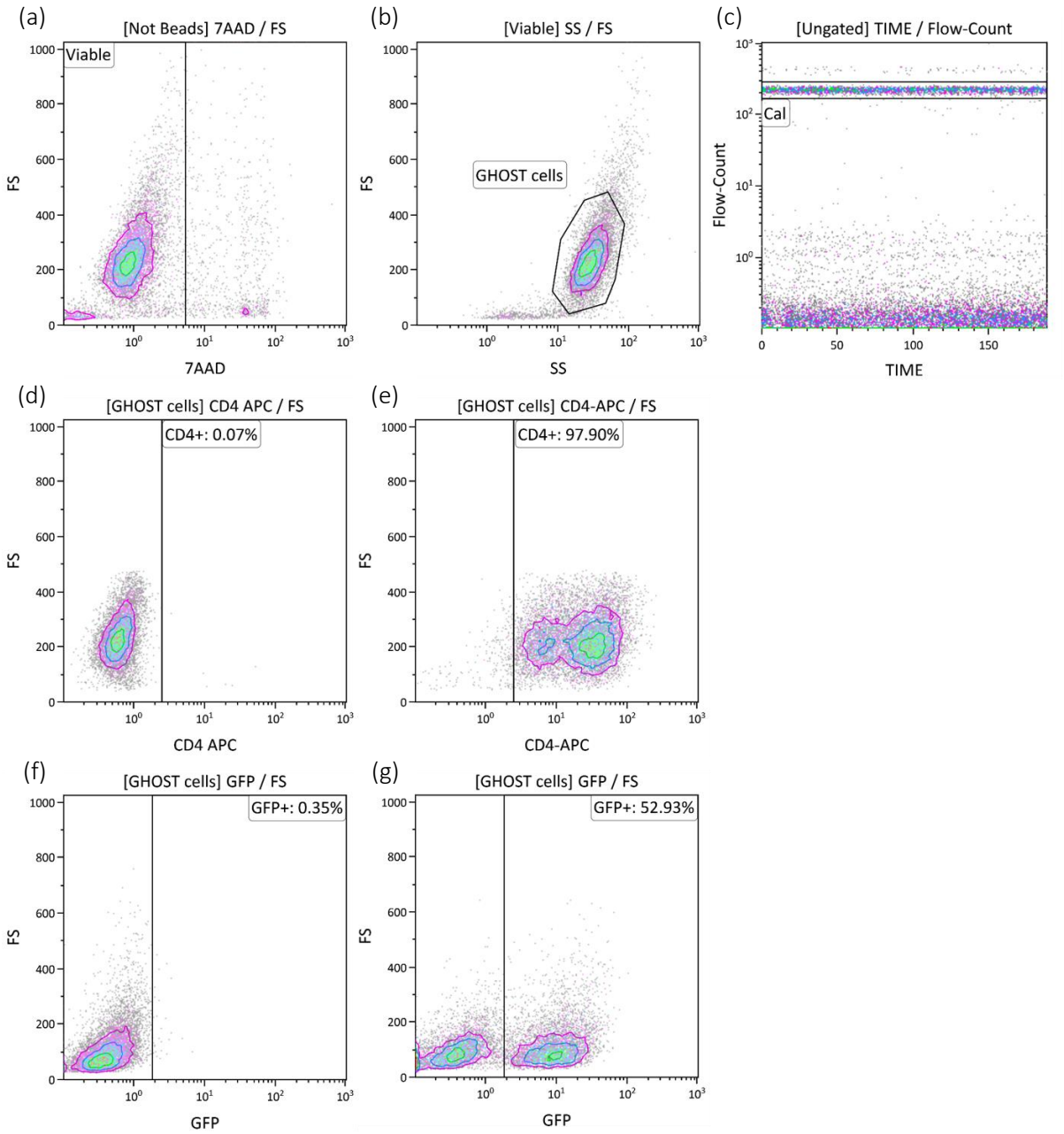


Figure 6.5. GHOST cell count and phenotype. A schematic illustration of the sequential gating strategy used for the GHOST cell count and phenotype. Density plot showing (a) viable (7AAD-negative) cells. A FS vs. SS density plot showing (b) GHOST cells. (c) Density plot indicating Flow-Count fluorospheres in "Cal" region. Density plots showing GHOST cells stained with (d) APC isotype and (e) CD4 APC monoclonal antibodies. Density plots showing (f) GFP^- (HIV-unexposed) and (g) GFP^+ (HIV-exposed) GHOST cells.

6.2.2.2.1.3. GHOST cell sorting

GHOST cells tend to lose CD4 expression over time. GHOST cells with high CD4, CCR5 and CXCR4 expression should be used to titrate HIV-1. Initial phenotyping indicated a sub-population of GHOST cells with lower CD4 expression. Therefore, it was decided to sort GHOST cells with higher CD4 expression to obtain a population with more uniform CD4 expression. GHOST cells were sorted by FACS using CD4 APC. Low-passage GHOST cells were removed from liquid nitrogen and cultured until confluent. GHOST cells were non-enzymatically dissociated as described earlier. Once dissociated, GHOST cells were stained with 7AAD and CD4 APC (approximately 2 µL antibody per 1 x 10^6 cells) and incubated for 15 min, protected from light. Cells were centrifuged at 300 x g for 10 min to remove unbound antibody, and were resuspended in PBS for sorting. Viable, high CD4-expressing ($CD4^{high}$) GHOST cells were sorted with > 95% post-sort purity into 15 mL centrifuge tubes (containing 2 mL DMEM). The flow cytometry protocol and sequential gating strategies used for sorting are indicated in Figure 6.6. $CD4^{high}$ GHOST cells were seeded at 5 x 10^5 cells per T75 flask in complete DMEM with puromycin, hygromycin, and G418 as described earlier for expansion. Once cells reached confluence, they were enzymatically dissociated and cryopreserved for future use.

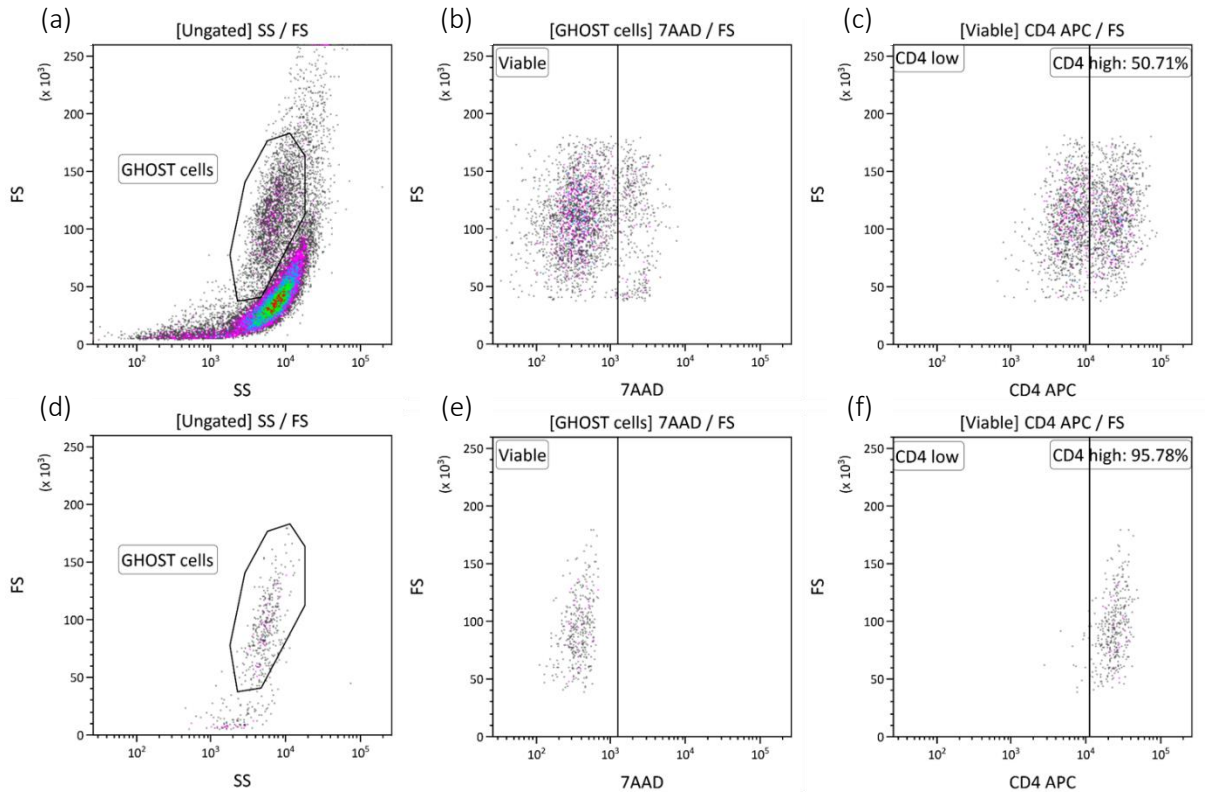


Figure 6.6. Sorting protocol for high CD4-expressing GHOST cells. (a) FS vs. SS plot used to identify the cell populations before sorting. (b) Density plots showing cell viability (7AAD-negative) before sorting. (c) Density plot showing cells stained with CD4 APC conjugated antibody before sorting. (d) A FS vs. SS plot used to identify the sorted cell population (post-sort). (e) Density plot showing cell viability (7AAD-negative) of sorted cells. (f) Density plot showing cells stained and sorted with CD4 APC conjugated antibody.

Primary HIV-1C isolates were titrated using the GHOST infectivity assay. As mentioned earlier, GHOST cells express GFP upon infection with HIV-1. HIV-exposed (GFP^+) and unexposed (GFP^-) GHOST cells were sorted for various optimisation procedures. Cells were prepared for sorting by centrifugation at 300 x g for 10 min and resuspended in PBS. GFP^+ and GFP^- cells were sorted into 1 mL PBS and a post-sort purity of 99.95% HIV-unexposed GHOST cells was achieved (Figure 6.7).

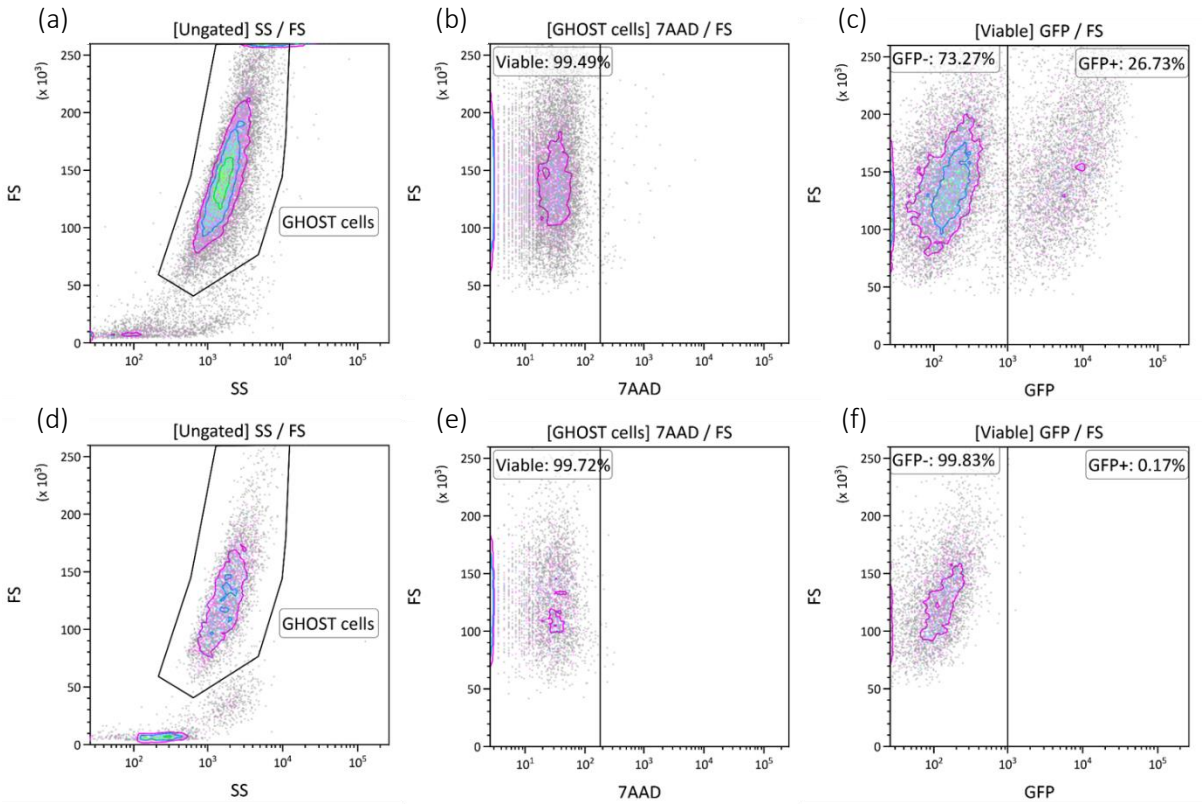


Figure 6.7. Sorting protocol for HIV-exposed and non-exposed GHOST cells. (a) FS vs. SS plot used to identify the cell populations before sorting. (b) Density plots showing cell viability (7AAD-negative) before sorting. (c) Density plot showing GFP^+ and GFP^- GHOST cells before sorting. (d) A FS vs. SS plot used to identify the sorted GHOST cell population (post-sort). (e) Density plot showing cell viability (7AAD-negative) of sorted cells. (f) Density plot showing cells GFP- sorted GHOST cells.

6.2.2.2.2. GHOST cell assay

Functional titration of HIV-1 using the GHOST cell assay was achieved by making serial dilutions of viral stocks and exposing HIV-susceptible GHOST cells to the various dilutions of the viral stocks (Figure 6.8). The GHOST cell assay was performed in 12-well plates (Greiner Bio-One, Austria) at a seeding density of 1.5 x 10^5 cells per well in 1 mL complete DMEM, and was performed in triplicate.

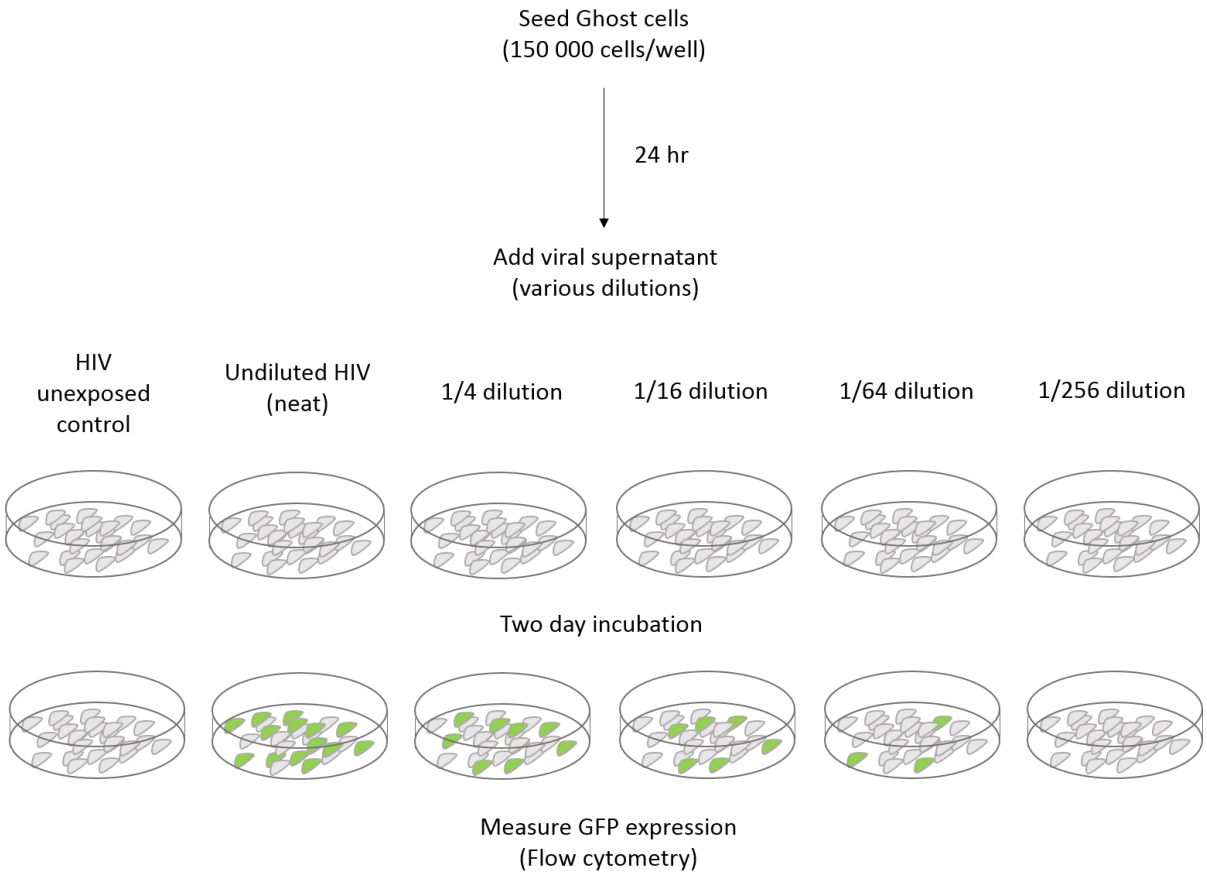


Figure 6.8. GHOST cell assay. A schematic representation of the GHOST cell assay using the GHOST reporter cell line. (Figure was created by Juanita Mellet).

6.2.2.2.2.1. Day 1 count

Seeded GHOST cells were allowed to attach for 24 hours prior to infection. The absolute counts of three triplicate wells seeded for the GHOST cell assay were determined 24 hours post-seeding of the GHOST cells. Cells were dissociated using 200 µL trypsin per well and centrifuged at 300 x g for 10 min. Cell pellets were resuspended in 100 µL PBS for counting.

6.2.2.2.2.2. HIV-1 exposure

Four-fold serial dilutions were prepared in 15 mL tubes and included: undiluted virus (neat), 1/4, 1/16, 1/64 and 1/256 dilutions. An HIV-unexposed control was also included. Polybrene® (20 µg/mL; Sigma-Aldrich, USA) was added to increase the efficiency of HIV infection (Davis, Morgan, & Yarmush, 2002). Cells were incubated together with the viral dilutions for two hours

at 37°C and 5% CO_2, after which the volume in each well was increased to 1 mL with complete DMEM and incubated overnight. After overnight incubation, HIV-containing supernatants were aspirated from wells and replaced with 1 mL complete DMEM.

6.2.2.2.2.3. GFP expression

GFP expression was analysed 48 hours after HIV-exposure. Ghost cells were dissociated and collected in 2 mL microcentrifuge tubes and centrifuged at 300 x g for 10 min. Supernatants were aspirated and cells were resuspended in 300 µL intracellular (IC) fixation buffer (Thermo Fisher Scientific, USA) and incubated for 10 min at 4°C, as per the manufacturer's instructions. After incubation, the fixation buffer was removed and cells were washed by adding 1 mL PBS to each tube, followed by centrifugation at 300 x g for 10 min. The supernatants were aspirated, and cells were resuspended in 400 µL PBS and transferred to corresponding flow cytometry tubes. The percentage GFP-expressing cells was determined by flow cytometry as described in earlier in this section.

6.2.2.2.2.4. HIV-1 titration calculation

Percentage GFP^+ cells were averaged over triplicates for each dilution. The HIV titers, expressed as infectious units per milliliter (IU/mL), were calculated based on the percentage GFP-expressing cells for the various dilutions (averaged over triplicates) and relating the number to the dilution factor. The following equation was used to calculate the number of infectious units:

Infectious units (IU)/mL = (cell number) × (% GFP^+ cells) × (dilution factor).

6.2.2.3. Nucleic acid detection

6.2.2.3.1. DNA extraction and purification

The QIAamp DNA Micro kit (Qiagen, Germany) was used to perform DNA extractions according to manufacturer's instructions. Cells for DNA extraction were centrifuged at 300 x g for 10 min in 1.5 mL microcentrifuge tubes and cell pellets were resuspended in 50 µL PBS. ATL buffer (100 µL) and 10 µL proteinase K were added to each sample to lyse the cell membranes and digest proteins that could potentially influence DNA extraction. Samples were mixed and

incubated for 10 min in a 56°C pre-heated water bath to allow for proteinase K digestion. Absolute ethanol (50 µL) was added to samples to precipitate proteins after which the samples were incubated for another 3 min. Lysates and precipitated proteins were transferred to QIAamp MinElute columns inserted in 2 mL collection tubes and centrifuged at 6000 x g for 1 min. AW buffer (1500 µL) was applied to the silica membrane of the column and centrifuged for 1 min at 6000 x g to wash bound nucleic acids. The column was washed once more with 500 µL AW2 buffer. The column was transferred to a new 2 mL collection tube and centrifuged at 14 000 x g for 1 min to dry the membrane of the spin column. The column was placed in a new collection tube and 20 – 50 µL of nuclease-free water was added directly to the center of the membrane and incubated for 5 min, followed by centrifugation at 14 000 x g for 1 min to elute the DNA. To maximise DNA recovery, the eluate was placed back onto the membrane for 1 min and centrifuged again at 14 000 x g for 1 min. DNA sample concentrations were measured using the NanoDrop® ND 1000 spectrophotometer (Thermo Fisher Scientific, USA) and stored at -80°C until further use.

A low 260/230 ratio usually indicates contamination. For DNA samples with a low 260/230 ratio, an ethanol precipitation was performed to remove contaminants. Sodium acetate (3 M NaAc; pH 5.2; Sigma-Aldrich, USA) was added to each sample in a 1:10 ratio. Ice-cold absolute ethanol (70 µL; 2.5 x total volume) was added to each tube and mixed gently. The solutions were incubated at -20°C for 1 hour. This incubation allowed salts and protein contaminants to dissolve in the NaAc solution, and the DNA to precipitate in the absolute ethanol. The DNA was pelleted by centrifugation at 14 000 rpm for 30 min, after which the supernatant (containing contaminants) was removed. The DNA pellets were washed once more with 100 µL 70% ethanol and centrifuged for 30 min at 14 000 rpm. As much supernatant as possible was removed. Any remaining ethanol was removed by air-drying in a pre-heated heating block to 50°C for 10 – 15 min. Air-dried pellets were resuspended in 25 µL nuclease-free water before NanoDrop quantification.

6.2.2.3.2. RNA extraction and reverse transcription

Extractions were performed using the Qiagen RNeasy Mini Kit (Qiagen, Germany) according to the manufacturer's instructions. Prior to each use, RLT buffer was prepared by adding

β-mercaptoethanol (Sigma-Aldrich, USA) to a final concentration of 1%, as per manufacturer's instructions. Cells were transferred to 1.5 mL microcentrifuge tubes and pelleted by centrifugation at 300 x g for 10 min. The supernatant was discarded and the pellets were resuspended in 350 µL RLT buffer for each extraction, followed by vortexing for 1 min to homogenise the sample. RLT buffer contains β-mercaptoethanol and guanidine thiocyanate to lyse cells. A total of 350 µL of 70% ethanol was added to the flow-through and mixed by pipetting. The sample was transferred to an RNeasy MinElute spin column inserted in a 2 mL collection tube and centrifuged at 8 000 x g for 15 seconds. After discarding the supernatant, a total of 700 µL RW1 buffer was added to the MinElute column and centrifuged at 8000 x g for 15 seconds. The silica membranes were washed twice by adding 500 µL of buffer RPE to the MinElute column and centrifuged at 8000 x g for 15 seconds. A total volume of 500 µL of 80% ethanol was added to the MinElute column and centrifuged at 8000 x g for 2 min to wash the membrane of the spin column. The flow-through was discarded after each step. The MinElute column was placed in a new 1.5 mL collection tube and 20 – 30 µL of RNase-free water was added directly to the center of the membrane and incubated for 2 min, followed by centrifugation at maximum speed for 1 min to elute the RNA. The eluates were placed back on the column membranes and incubated for 1 min, followed once again by centrifugation at 8000 x g for 1 min to improve the RNA yield. RNA sample concentrations were measured using the NanoDrop ND 1000 spectrophotometer (Thermo Fisher Scientific, USA) and stored at -80^{o}C until further use.

RNA was reverse transcribed to cDNA using the Bioline SensiFast cDNA Synthesis kit (Meridian Life Sciences, Inc, USA) according to the manufacturer's instructions. A no template control (NTC) and no reverse transcriptase control (NRC) were included. Approximately 1 µg RNA was added per reaction, which was prepared on ice in 0.2 mL PCR tubes. The SensiFAST cDNA reagent kit uses a high-fidelity (HiFi) reverse transcriptase enzyme and optimised buffer system. Gene Amp™ thermocycler conditions for reverse transcription were as follows: 25°C for 10 min to anneal primers; 42°C for 15 min for reverse transcription; 85°C for 5 min to inactivate reverse transcriptase; 4°C hold to prepare for sample storage.

6.2.2.3.3. Polymerase chain reaction (PCR)

Polymerase chain reaction mixes were prepared in 0.2 mL PCR tubes to detect HIV-1 proviral DNA using the 2X Kapa Taq ReadyMix PCR kit (Kapa RM; Kapa Biosystems, USA). The final volume of 25 µL included Kapa RM, forward and reverse primers, 50 ng DNA and molecular grade water. Table 6.4 indicates the primers that were used to detect the different HIV-1 isolates. The Gene Amp Thermal Cycler conditions were as follows: 94°C for 2 min to denature the DNA, 94°C for 30 seconds to extend denaturation of the DNA, 58°C for 30 seconds to anneal primers; 72°C for 1 min to elongate primers; 72°C for 10 min to extend the elongation step; 4°C hold. Annealing, elongation and extended elongation were subjected to 30 cycles of amplification.

A positive PCR control containing primers targeting the human large ribosomal subunit (L32) was included with each reaction. Primer sequences for the L32 primer pair were obtained from Dr Patrick Salmon's Lentviral Vectors Lab, University of Geneva, Switzerland. The forward primer sequence is 5'- GTG AAG CCC AAG ATC GTC AA -3' (L32F), and the reverse primer sequence is 5'- TTG GTG ACT CTG ATG GCC AG -3' (L32R). The annealing temperature for L32 primers is 65°C.

Table 6.4. HIV-1C primer pairs for CM9 and SW7.

Isolate and region	Primer name	Sequence (5'-3')	Product size (bp)
CM9 V3-loop	Forward	GGACCATGCAATAATGTCAGC	406
	Reverse	GTGTTGTAATTTCTAGGTCCCC	
CM9 LTR/Gag	Forward	CCCTCAGATGCTGCATATAAGCAGCTGC	983
	Reverse	TCCTTTAACATTTGCATGGCTGCTTG	
SW7 V3-loop	Forward	GGACCATGCCATAATGTCAGC	551
	Reverse	CCTACCCCCTGCCACATG	
SW7 LTR/Gag	Forward	CCCTCAGATGCTGCATATAAGCAGCTGC	978
	Reverse	TCTTTTAACATTTGCATGGCTGCTTG	

For detection of CD4 mRNA expression, forward (CD4F) 5'- GTC CCT TTT AGG CAC TTG CTT CT - 3' and reverse (CD4R) 5'- TCT TTC CCT GAG TGG CTG CT -3' primers for a CD4 region were obtained from the literature (370). Reactions were prepared with approximately 50 ng/µL cDNA per reaction. Thermocycling was performed as described earlier with a temperature of 65°C for primer annealing of CD4 primer pairs. Controls for the PCR included a non-template control (NTC) for the PCR and an L32 positive PCR control.

6.2.2.3.4. Gel electrophoresis

1X BlueJuice™ Gel Loading Buffer (Thermo Fisher Scientific, USA) was added to the PCR products and run using agarose gel eletrophoresis at 90 V for 60 min. The FastRuler™ Low-range DNA ladder (Thermo Fisher Scientific, USA) was included to estimate the amplicon sizes. DNA was visualised under UV light in a Gel Doc™ XR+ Gel Documentation System (Bio-Rad Laboratories Inc., USA).

6.2.3. $CD34^+$ HSPCs for HIV-1 infection

6.2.3.1. Sample collections

Umbilical cord blood and leukapheresis products were used as sources of HSPCs for this study. Informed consent and UCB sample collections were performed as described in Chapter 2. Collected UCB samples were used within 24 hours of collection. Autologous and allogeneic cryopreserved leukapheresis products were donated by the Alberts Cellular Therapy (ACT) unit at a private hospital in Pretoria, after consent and approval was obtained to donate the products (otherwise to be discarded) for research purposes. Leukapheresis products were anonymised by ACT staff. On arrival, sample identities were assigned and products were catalogued and transferred to liquid nitrogen storage.

6.2.3.2. HIV testing

Only HIV-negative samples were included for this study. HIV testing was performed as described in Chapter 3, Section 3.2.3. HIV testing of leukapheresis donors is routinely performed prior to HSCT. Leukapheresis donor HIV status was therefore provided by the ACT staff. Leukapheresis products were not tested using the GeneXpert, since the GeneXpert was validated for testing whole blood samples only.

6.2.3.3. Sample processing

6.2.3.3.1. UCB samples

UCB processing was performed as described in Chapter 3, Section 3.2.4, while counting of UCB-derived HSPCs was performed as described in Chapter 2, Section 2.2.3.1.

6.2.3.3.2. Leukapheresis products

Cryopreserved leukapheresis products were obtained from cryobags that contained 10% DMSO as a cryoprotectant. A thawing protocol supplied by the ACT was used to thaw the leukapheresis products, with some minor adjustments. The leukapheresis collection bags were removed from liquid nitrogen storage and placed on the bench at room temperature for precisely 2 min. The sample bag was then transferred to a 2 L glass beaker in a water bath containing pre-heated (37°C) 1% NaCl (Sigma-Aldrich, USA). The leukapheresis product was thawed for a maximum of 2 min, by gently massaging the bag while submerged in the NaCl. Once an icy consistency was achieved, 10 mL was aliquoted into several 50 mL tubes containing 40 mL cold PBS. Tubes were mixed by inversion and centrifuged at 300 x g for 10 min at 4°C. Cell pellets were resuspended in TP_4 buffer (TP buffer supplemented with 40 µg/mL human albumin) and strained through 70 µM Falcon™ cell strainers (Falcon, Corning, USA). Cells were pelleted once more by centrifugation at 300 x g for 10 min after which several pellets from different tubes were pooled and resuspended in 50 mL RPMI (2% pen/strep) to determine the absolute number of HSPC in the sample. A 100 µL aliquot of the cell suspension was taken for counting. Antibodies were added to samples (approximately 2 – 3 µL per 1 x 10^6 viable cells) and incubated for 15 min prior to counting. A 100 µL of Flow-Count fluorospheres was added to each sample immediately prior to analysis on the flow cytometer. Counting and analysis on the Gallios flow cytometer was performed as described in Chapter 2, Section 2.2.3.1.

6.2.3.4. Enrichment of $CD34^+$ HSPCs

The proportion of $CD34^+$ HSPCs is low in the MNC fraction of UCB and leukapheresis products. To ensure that a pure population of HSPCs is used for downstream experiments, selection for cells expressing CD34 was performed by MACS (Chapter 2, Section 2.2.3.2) and/or FACS (Chapter 2, Section 2.2.3.3) enrichment.

6.2.3.4.1. FACS sorting of leukapheresis products

Fluorescent activated cell sorting was used to sort viable HSPCs directly from either MACS-enriched $CD34^+$ HSPCs or HSPC populations post-infection with HIV. The MACS-enriched $CD34^+$ cell suspension was stained for sorting with an appropriate volume of CD34 PE-Cy7 (Beckman Coulter, USA) and 7AAD based on the absolute number of cells present (approximately 2 – 3 µL per 1 x 10^6 viable cells), determined during counting. Cells were stained for 30 min at 4°C, protected from light. Stained cells were washed with TP buffer and centrifuged at 300 x g for 10 min and the supernatant was aspirated. The pellet was resuspended in the desired volume based on the absolute cell number (final concentration of approximately 7 x 10^6 cells/mL). Cells were sorted based on CD34 expression by FACS using the BD FACSAria™ Fusion cell sorter (BD Biosciences, USA). Viable, MACS-enriched $CD34^+$ HSPCs were sorted into 15 mL centrifuge tubes containing 2 mL pre-warmed (37°C) StemSpan ACF medium (2% pen/strep). HIV-infected $CD34^+$ HSPCs were sorted into 24-well plates containing 1 mL medium StemSpan ACF medium (2% pen/strep) supplemented with five cytokines, FLT3L, SCF, TPO, IL-3 and G-CSF, each at 100 ng/mL.

6.2.3.5. HIV-1 infection and detection

6.2.3.5.1. UCB-derived HSPC infection

A total of 2 x 10^5 $CD34^+$ HSPCs were sorted (as described in Chapter 2) per well into 500 µL RPMI (2% pen/strep, 10% human albumin) of a 24-well plate. Half of the HIV-exposed and unexposed wells were supplemented with 100 ng/mL of the following cytokines: FLT3L, SCF and TPO. Plates were incubated overnight at 37°C and 5% CO_2 to recover (Figure 6.9).

Following overnight incubation, cells from each well were transferred to 1.5 mL microcentrifuge tubes and pelleted by centrifugation at 300 x g for 10 min, after which the supernatant was aspirated. Cell pellets were resuspened in 1 mL appropriate medium as follows. HIV-exposed cells were resuspended in CM9-containing supernatant and HIV-unexposed wells were resuspended in control supernatant and incubated for two hours. Following incubation, cells were incubated at 37°C and 5% CO_2. Half the volume

of each well was collected after 24 hours by gently pipetting to ensure a homogenous cell suspension, and the other half was collected in the same way after 48 hours. Cells collected at each time point were pelleted for DNA extraction by centrifugation at 300 x g for 10 min.

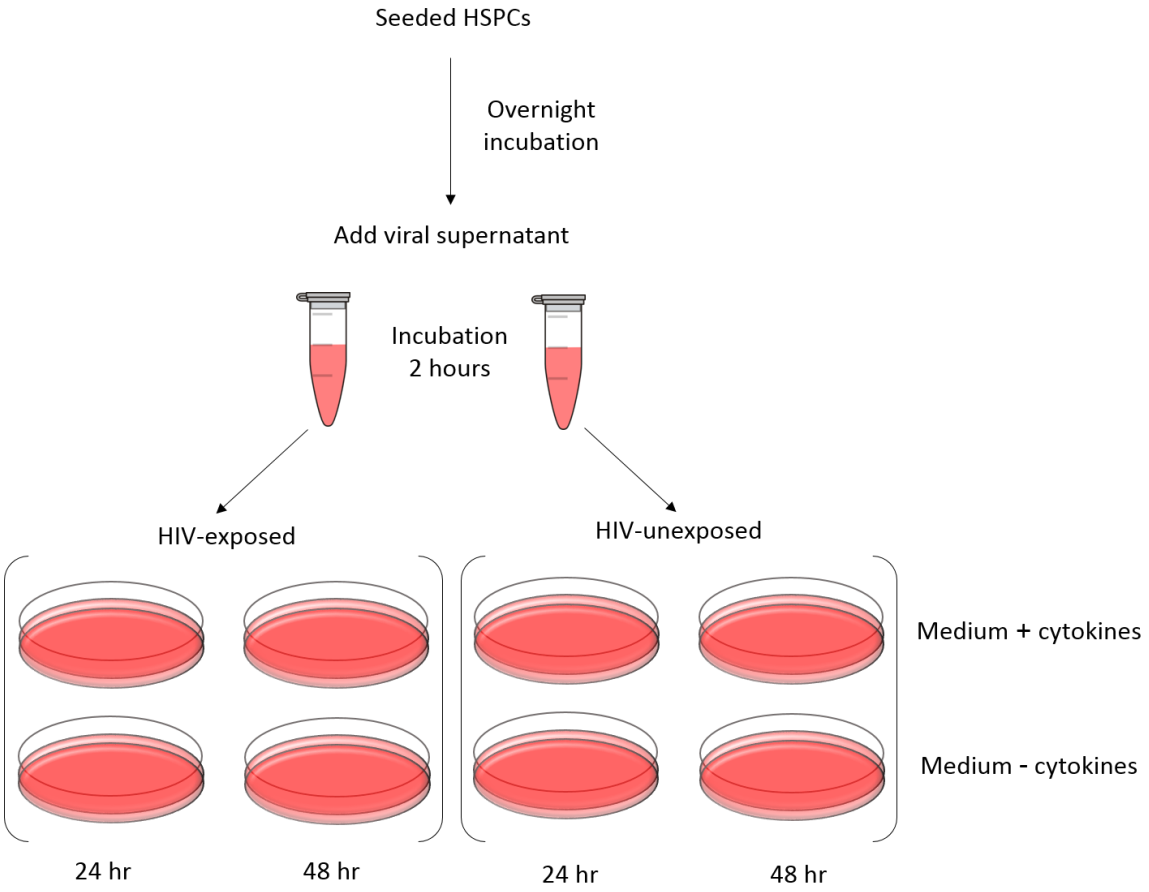


Figure 6.9. HIV-1 infection of UCB-derived HSPCs. Freshly isolated HSPCs were incubated overnight to recover. Cells were aspirated and pelleted by centrifugation in two separate centrifuge tubes. Cell pellets were resuspended in virus-containing supernatant and control supernatant, respectively, and incubated for two hours in a small volume. After the two-hour incubation, cells were transferred to a 24-well plate for 24 and 48 hours. (Figure was created by Juanita Mellet).

6.2.3.5.2. Leukapheresis-derived HSPC infection

Differences were observed in susceptibility of UCB-derived HSPCs when infected with and without the presence of cytokines. Therefore, the HSPCs were exposed to three different infection conditions, including and excluding five cytokines: (1) Control, (2) control supernatant and (3) HIV-1-exposed. For the conditions that included cytokines, 100 ng/mL of each of the following cytokines were added: SCF, TPO, FLT3L, IL-3 and G-CSF. Initial exposure to HIV-1 was performed in a rotating incubator at 37°C for two hours. This was to ensure adequate contact

between virus and cells. After the initial two hour exposure, aliquots were split in half and plated in a 6-well plate. Pre-warmed StemSpan ACF (2% pen/strep) was added to reach a final volume of 1.5 mL per well. The cells were incubated at 37°C and 5% CO_2 for 24 hours.

After 24-hour incubation, cells from each well were collected into separate 2 mL microcentrifuge tubes. Each well was washed with 400 µL PBS and transferred to the corresponding tubes. Cells were centrifuged at 300 x g for 10 min and the supernatants were aspirated. The cell pellets were resuspended in 300 µL TP_4 buffer and 200 µL was used to sort for expansion, while 100 µL was stained to sort for CFU assays. The CFU assays formed part of Ms. Candice Herd's MSc project. The aliquots were stained for sorting with 3 µL CD34 PE-Cy7 and 3 µL 7AAD as described in Section 7.2.4.4.1. Viable $CD34^+$ HSPCs (1 x 10^4) were sorted in triplicate into wells of a 24-well plate containing 1 mL StemSpan ACF (2% pen/strep) per well. The five cytokines were once again added to each well at 100 ng/mL. The plate was incubated for seven days at 37°C with 5% CO_2.

6.2.3.6. HSPC immunophenotypic analysis

The immunophenotypic profiles of sorted $CD34^+$ HSPCs were assessed after sorting, on the day of infection and 24 hours post-infection. Two antibody panels containing different immunophenotypic markers were assessed using the FACSAria Fusion (Table 6.5 and 6.6). Viability dyes, Zombie Violet (panel 1) and 7AAD (panel 2) were included and allowed for the exclusion of non-viable cells. HIV-1 receptor (CD4) and co-receptor (CCR5 and CXCR4) expression was also assessed. The monoclonal antibodies (3 – 5 µL) were added to flow tubes containing 50 µL of the cell suspension. Single-colour antibody staining controls were used to correct for spectral overlap between the fluorochromes (colour compensation).

Table 6.5. Antibody panel 1.

Antibody	Fluorochrome	Clone	Supplier
Lineage cocktail	FITC		
CD3		UCHT1	Biolegend, USA
CD14		HCD14	
CD16		3G8	
CD19		HIB19	
CD20		2H7	
CD56		HCD56	
CCR5	PE	J418F1	Biolegend, USA
Rat IgG2b	PE isotype	RTK4530	Biolegend, USA
CD34	PE-Cy7	581	Beckman Coulter, USA
CD133/2	APC	293C3	Myltenyi Biotec, Germany
Mouse IgG2b	APC isotype	IS6-11E5.11	Myltenyi Biotec, Germany
CD38	APC-Cy7	HIT2	Biolegend, USA
Zombie Violet	BV510	-	Biolegend, USA
CXCR4	BV605	12G5	Biolegend, USA

Table 6.6. Antibody panel 2.

Antibody	Fluorochrome	Clone	Supplier
CD4	BB515	RPA-T4	BD Biosciences, USA
Mouse IgG1	BB515	X40	BD Biosciences, USA
Lineage cocktail	APC		Biolegend, USA
CD3		UCHT1	
CD14		HCD14	
CD16		3G8	
CD19		HIB19	
CD20		2H7	
CD56		HCD56	
CD34	PE-Cy7	581	Beckman Coulter, USA
CCR5	PE	J418F1	Biolegend, USA
CXCR4	BV605	12G5	Biolegend, USA
7AAD	-	-	Beckman Coulter, USA

The gating strategy for antibody panels 1 (Figure 6.10) and 2 (Figure 6.11) was performed as follows. The HSPCs ("Cells" region) was identified using a FS vs. SS plot. A 7AAD vs. FS plot, gated on "Cells" was used to identify viable cells ("Viable" region). $CD34^+$ HSPCs were identified using CD34 PE-Cy7 [Excitation: 488/532/561 nm; Emission: 575/780 nm] and gated on "Viable" cells. Lin^+ and Lin^- cells were identified using the lineage cocktail FITC [Excitation: 488 nm; Emission: 520 nm]. $CD38^+$ and $CD38^-$ cells were identified using CD38 APC-Cy7 [Excitation: 633 nm; Emission: 776 nm]. $CD133^+$ cells were identified using CD133 PE [Excitation: 496 nm; Emission: 578 nm]. $CD4^+$ HSPCs were identified using the CD4 BB515 [Excitation: 490 nm; Emission: 515 nm]. $CCR5^+$ and $CXCR4^+$ HSPCs were identified using CCR5 PE [Excitation: 496 nm; Emission: 578 nm] and CXCR4 BV605 [Excitation: 405 nm; Emission: 603 nm], respectively. All plots were gated on $CD34^+$ cells. The negative and positive regions for each phenotypic marker were created based on positive and negative populations present in a sample (if applicable) or whole blood when the distinction between the negative and positive populations were not clear. The CD4, CCR5, CXCR4 and CD133 regions were created using isotypic controls.

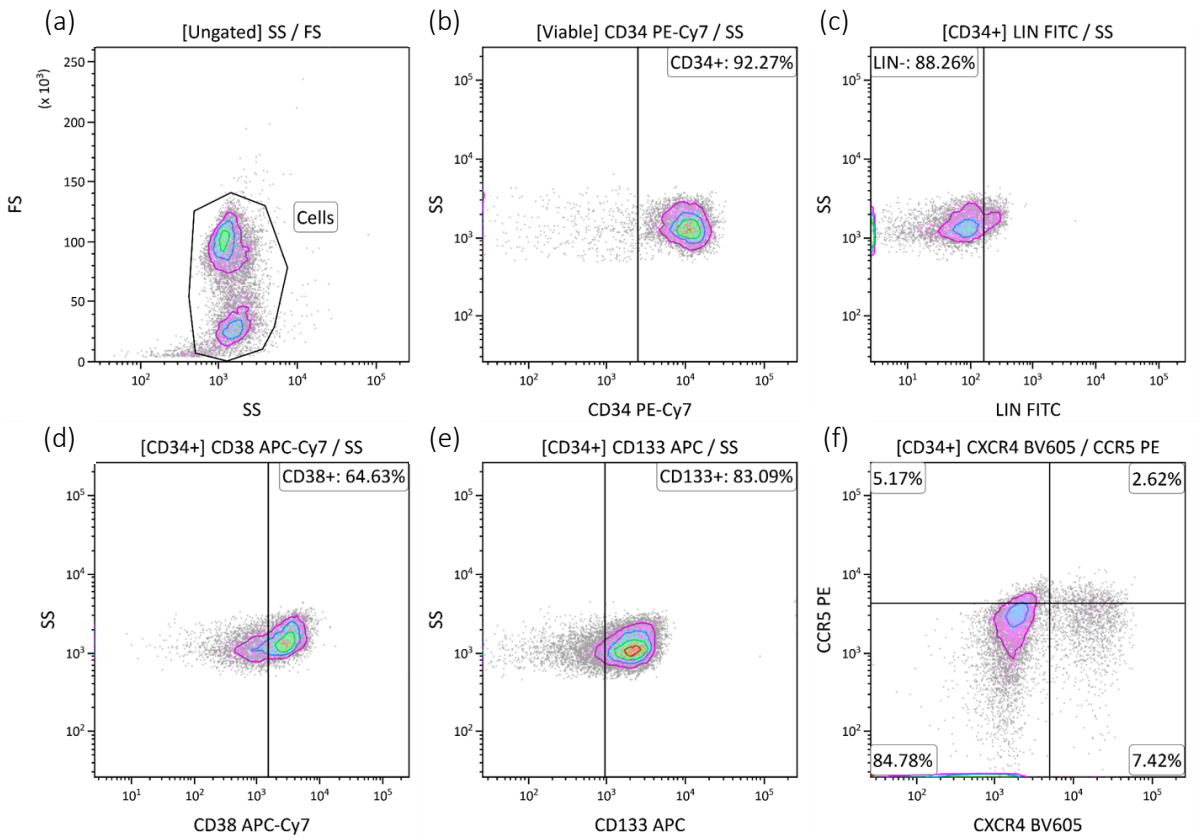


Figure 6.10. Immunophenotypic gating strategy for antibody panel 1. Sequential gating strategy for HSPC-associated immunophenotypic markers. Density plots showing (a) FS vs. SS and cells stained with (b) CD34 PE-Cy7, (c) Lin FITC, (d) CD38 APC-Cy7, (e) CD133 APC and (f) CCR5 PE and CXCR4 BV605. Zombie violet plot not shown in this illustration.

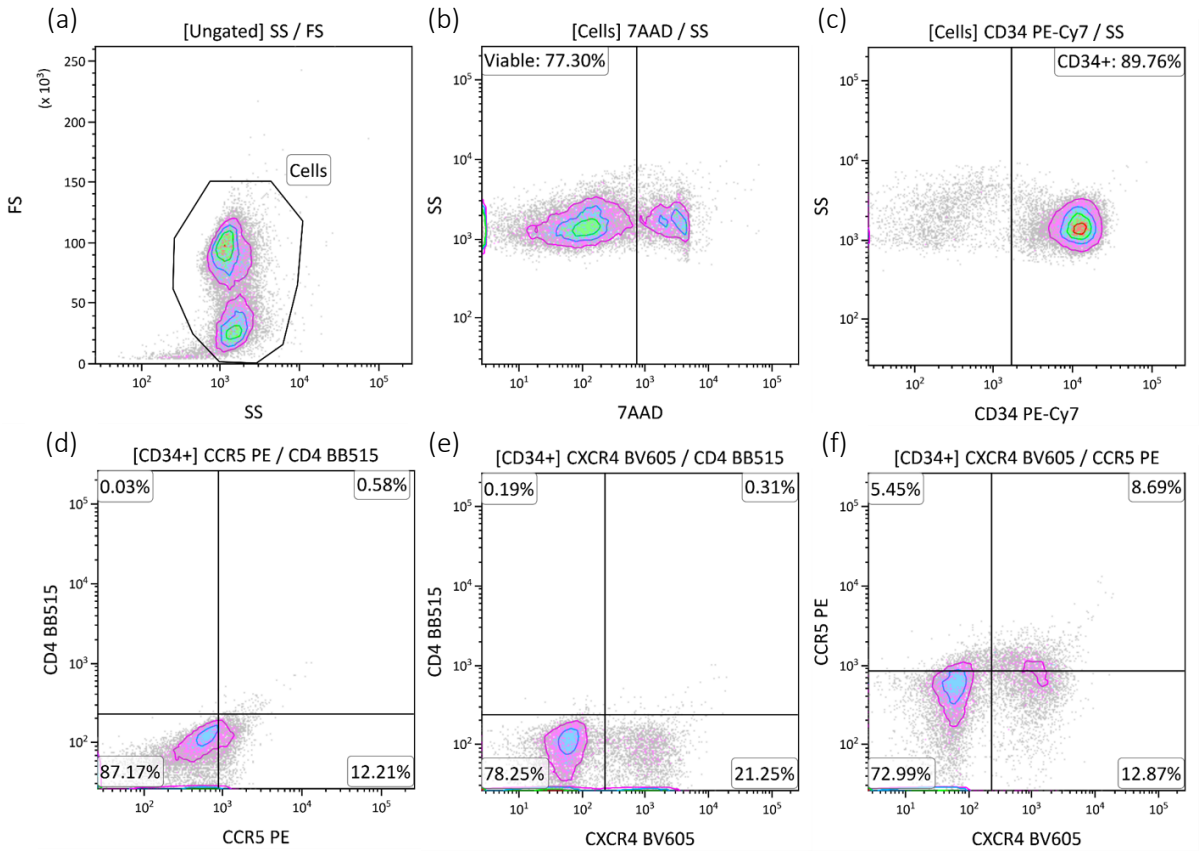


Figure 6.11. Immunophenotypic gating for antibody panel 2. Sequential gating strategy for HSPCs-associated immunophenotypic markers. Density plot showing (a) HSPC cell population (FS vs. SS) stained (b) 7AAD (viable, 7AAD-negative), (c) CD34 PE-Cy7, (d) CD4 BB515 and CCR5 PE, (e) CD4 BB515 and CXCR4 BV605 and (f) CCR5 PE and CXCR4 BV605.

6.3. RESULTS

6.3.1. Propagated HIV-1 isolates and stocks

HIV-1C primary isolates were propagated using activated PBMCs. HIV RNA is reverse transcribed to DNA upon viral entry into target cells. The error-prone reverse transcriptase introduces mutations that lead to unpredictable loss-of-function, ultimately resulting in low levels of infectious virus produced. HIV was required for several HIV-related projects in our research group. Therefore, multiple HIV propagations were performed to produce sufficient HIV for all experiments. We aimed to use HIV produced from single productions where possible, to keep the HIV genomic variations to a minimum. Producing sufficient HIV-1C was particularly challenging. Although initial R5-tropic (CM1) productions yielded infectious virus, multiple production rounds resulted in non-infectious virus being produced (Appendix D). Despite attempts to pre-concentrate HIV, infectious units for CM1 remained low. We

attempted to produce R5-tropic virus from several other HIV-1C isolates, which was not successful. We were able to produce sufficient CM9 and SW7 over multiple production rounds. Table 6.7 summarises the HIV produced during this study, together with the respective functional titres as determined by the GHOST cell assay.

Table 6.7. HIV-1C stocks.

Isolate	Stock description	Date harvested	Functional titer (IU/mL)	Standard deviation (SD)
CM1	CM1-250217	25-02-2017	1.60×10^5	5.18×10^4
	CM1-110618	11-06-2018	1.20×10^5	5.60×10^4
CM9	CM9-290118	29-01-2018	9.47×10^5	2.47×10^5
	CM9-280218_P	28-02-2018	2.39×10^5	3.20×10^4
	CM9-020518	02-05-2018	1.05×10^6	6.28×10^5
	CM9-270618	27-06-2018	3.60×10^5	2.25×10^5
SW7	SW7-240618	24-06-2018	7.74×10^4	2.72×10^4
	SW7-110718	11-07-2018	2.87×10^5	1.26×10^5
	SW7-010818	01-08-2018	4.71×10^5	4.70×10^4
	SW7-220918	22-09-2018	7.90×10^3	6.34×10^2
Du422F	Du422F-240718	24-07-2018	2.60×10^4	5.95×10^3

6.3.2. Immunophenotyping of GHOST cells

Immunophenotypic analysis of GHOST cells, which should express high CD4, CCR5 and CXCR4, revealed high expression of CCR5 and CXCR4. However, low expression of CD4 was detected using CD4 FITC (Figure 6.12).

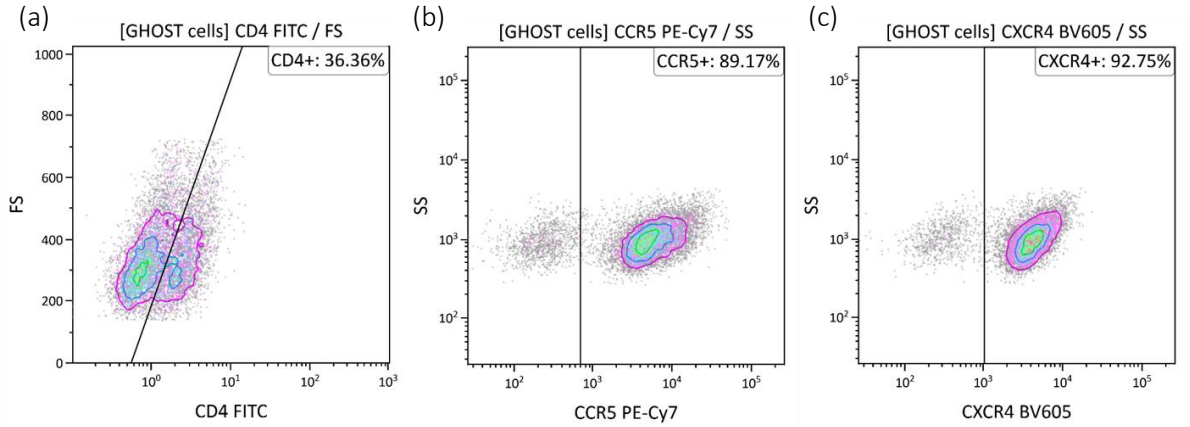


Figure 6.12. HIV receptor and co-receptor expression on GHOST cells. HIV receptor (CD4) and co-receptor (CCR5 and CXCR4) expression on GHOST cells. Density plots showing GHOST cells stained with (a) CD4 FITC, (b) CCR5 PE-Cy7 and (c) CXCR4 BV605. The negative/positive region boundaries were created for each phenotypic marker using an isotypic controls.

The low expression observed was postulated to be due to destabilisation of the CD4 construct (MX-CD4). If the low CD4 expression observed was due to construct destabilisation, then theoretically there should be less mRNA in $CD4^{low}$- compared to $CD4^{high}$-GHOST cells. To test this hypothesis, $CD4^{low}$- and $CD4^{high}$-expressing GHOST cells were sorted, mRNA isolated and mRNA was reverse transcribed into cDNA and PCR was performed.

Three negative controls were included (NTC_{PCR}, NTC_{RT}, and NRC) and as expected, no amplicons were produced in any of the negative control samples. The NTC_{PCR} controlled for DNA contamination, the NTC_{RT} controlled for RNA contamination and the NRC controlled for DNA contamination during RNA preparation. TZM-bl cells and PBMCs were included as positive controls since these cells express high CD4, while HEK293T cells were included as a negative control. An L32 control was included to ensure that the reagents and thermocycler were functioning. PCR of CD4 was expected to yield a single amplicon of 67 bp in size. Bands observed in TZM-bl cells, PBMCs, $CD4^{high}$ and $CD4^{low}$ GHOST cells were of the expected size (approximately 200 bp) (Figure 6.13). The intensities of the CD4 amplicons in $CD4^{low}$ and $CD4^{high}$ GHOST cells were similar, which suggested that decreased CD4 expression as observed by flow cytometry, was not due to destabilisation of the CD4 construct.

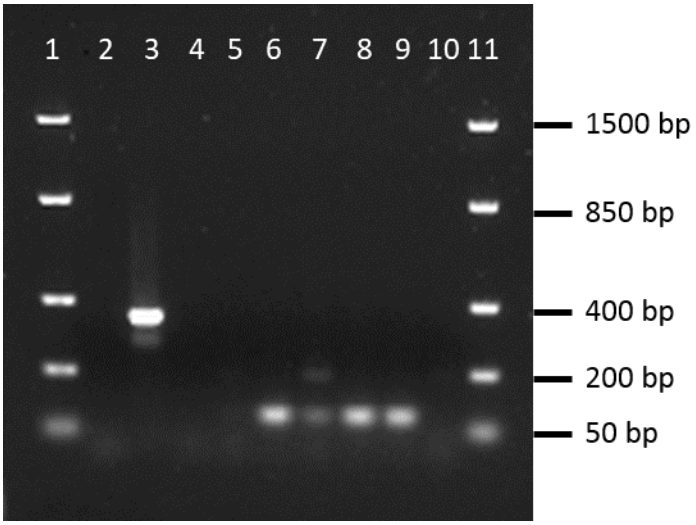


Figure 6.13. Gel image of CD4 mRNA in GHOST cells. Lanes 1 and 11 represent the FastRuler LR DNA ladder for size estimation. Lanes 2 and 3 represent the NTC_{PCR} and L32 PCR positive control, respectively. Lanes 4 and 5 represent the NTC_{RT} and NRC reverse transcription controls, respectively. Lanes 6 – 10 represent CD4 amplicons for TZM-bl (lane 6), PBMC (lane7), GHOST $CD4^{high}$ (lane 8), GHOST $CD4^{low}$ (lane 9), and HEK293T (lane 10) cells.

Since no difference was observed at the mRNA level between $CD4^{low}$ and $CD4^{high}$ cells, flow cytometric detection and dissociation methods were investigated. We initially used the CD4 FITC monoclonal antibody to detect CD4 expression. Antibody-conjugated fluorochromes differ in relative brightness, resulting in differences in detection even when the same antibody and cells are used. A general rule applied in flow cytometry is to visualise/analyse weakly expressed epitopes (low level expressed on the cell surface) using brightly excited fluorochromes. FITC is a weak fluorochrome and could restrict discrimination between CD4 positive and negative populations if weakly expressed on GHOST cells. Applying the last mentioned rule, we opted for CD4 conjugated to a brighter fluorochrome (APC) to detect CD4-expressing GHOST cells.

Preliminary results indicated reduced CD4 expression on enzymatically dissociated compared to non-enzymatically dissociated GHOST cells. We designed an experiment that simultaneously compared FITC- and APC-conjugated CD4 detection on GHOST cells, and determined the rate of CD4 recovery following dissociation (non-enzymatic vs. enzymatic) (Figure 6.14). Enzymatic (trypsin-EDTA) dissociation was investigated and CD4 expression was measured every 24 hours for four days using CD4 FITC and CD4 APC. Five biological replicates at different passages were

analysed for this purpose (Table 6.8). Three biological replicates from our laboratory were compared to two biological replicates kindly donated by Dr Janine Scholefield, a senior researcher at the Council for Scientific and Industrial Research (CSIR), South Africa. Assessment of the average CD4 recovery revealed improved detection of CD4 using a brighter fluorochrome (APC). CD4 FITC analysis indicated that only 56.3% (± 9.3%) of GHOST cells express high CD4 after dissociation, which decreased upon culturing to 13.5% (± 5.8%). CD4 APC analysis showed 94.3% (± 2.7%) expression of CD4 on GHOST cells after dissociation, which further increased upon culturing to 99.4% (± 0.3%). These results indicate that prior knowledge of expected levels of expression and fluorochrome choice are important considerations when enumerating CD4-expressing GHOST cells.

Table 6.8. GHOST cell cultures used for CD4 phenotyping.

Replicate	Passage	Date frozen	Laboratory
BR1	P4	09-03-2017	ICMM
BR2	P7	23-07-2017	ICMM
BR3	P3	39-06-2016	ICMM
BR4	P14	01-01-2017	CSIR
BR5	P11	08-09-2014	CSIR

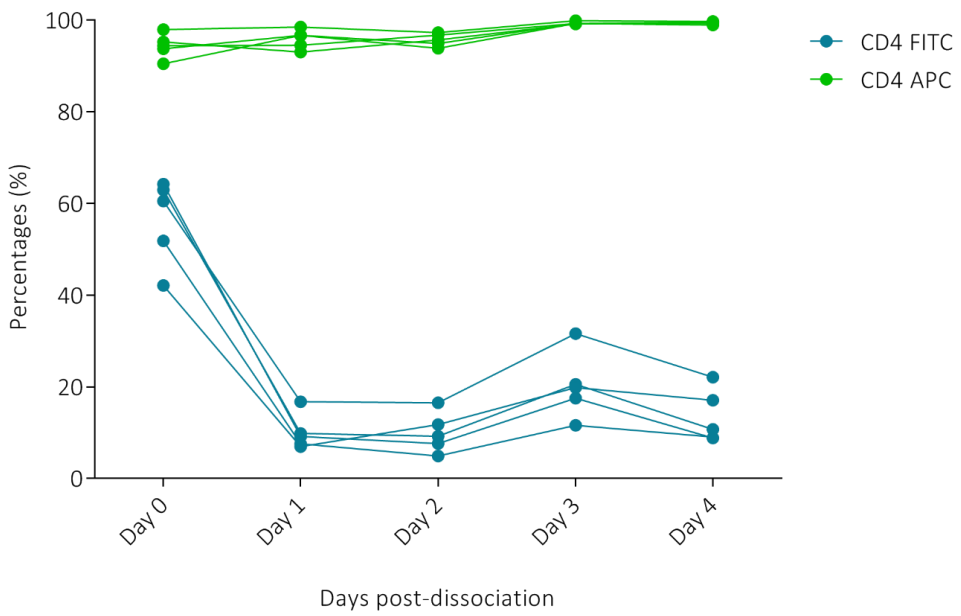


Figure 6.14. GHOST cell CD4 recovery following dissociation. Flow cytometric CD4 expression on GHOST cells after enzymatic dissociation using CD4 FITC (blue) and CD4 APC (green), respectively.

6.3.3. PCR limit to detect HIV-1

A limit of detection (LOD) refers to the lowest quantity of a substance to be reliably detected by an assay (371). The extent of susceptibility of HSPCs to HIV-1 is unknown and failure to identify the presence of HIV-1 proviral DNA through PCR could be due to the inability of the assay to detect low levels of HIV-1 proviral DNA. For that reason, a LOD experiment was performed using different proportions of HIV-infected (GFP^+) to HIV unexposed (GFP^-) GHOST cells (Table 6.9). GHOST cells were infected with two different HIV-1C tropic viruses, namely SW7 (X4-tropic isolate) and CM9 (R5X4-tropic isolate). The DNA concentrations extracted from the different proportions of HIV unexposed/HIV-infected GHOST cells are summarised in Table 6.10.

Table 6.9. The number of HIV-infected to HIV unexposed GHOST cells.

	0.1% HIV^+	1% HIV^+	5% HIV^+	10% HIV^+	20% HIV^+
HIV-infected (GFP^+) cells	100	1000	5000	10 000	20 000
HIV unexposed (GFP^-) cells	99 900	99 000	95 000	90 000	80 000

Table 6.10. DNA samples for PCR limit of detection. DNA samples with their respective DNA concentrations, 260/280 and 260/230 ratios.

Isolate	Sample description	DNA concentration (ng/µL)	260/280 ratio	260/230 ratio
CM9	0.1% HIV^+ cells	53.47	2.41	1.72
	1% HIV^+ cells	55.26	2.64	2.23
	5% HIV^+ cells	44.08	2.74	1.93
	10% HIV^+ cells	49.91	2.42	1.38
	20% HIV^+ cells	51.5	2.47	1.87
SW7	0.1% HIV^+ cells	53.92	2.50	1.98
	1% HIV^+ cells	54.14	2.29	1.62
	5% HIV^+ cells	57.04	2.35	1.53
	10% HIV^+ cells	51.71	2.66	1.08
	20% HIV^+ cells	58.54	2.24	1.56

Primers specific for the LTR-*gag* region of HIV were used to detect proviral DNA. Detection of 5000 infected cells in 1 x 10^5 total cells (5%) was established to be the limit of detection of our PCR method to detect proviral DNA. This is assuming that each cell is infected with a single HIV proviral genome copy. When visualised on a gel, the amplicons of CM9-exposed cells were better defined compared to the bands observed for the amplicons from SW7-exposed cells (Figure 6.15).

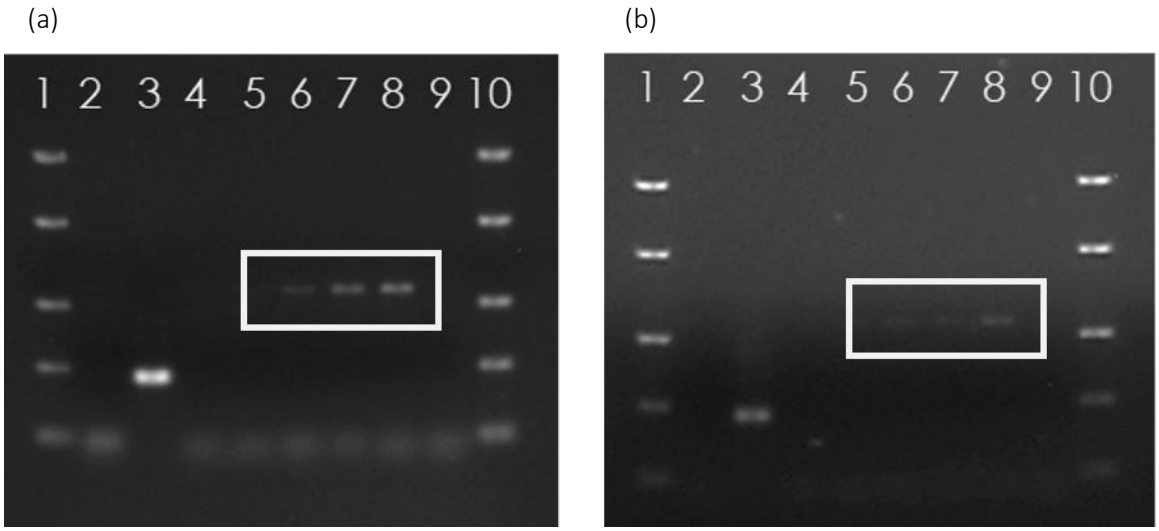


Figure 6.15. Gel images for PCR limit of detection. Gel electrophoresis images representing LTR-*gag* amplicons (white boxes) from different concentrations of HIV-infected GHOST cells to determine the PCR limit of detection for HIV-1C isolates, (a) CM9 and (b) SW7. Lanes 1 and 10 represent the FastRuler™ MR, lane 2 was loaded with the no template control (NTC) for each primer pair, and lane 3 was loaded with the L32 positive control performed on HIV-unexposed GHOST cell DNA. Lanes 4 – 9 were loaded with the respective isolate-specific LTR-*gag* PCR products containing 0.1%, 1%, 5%, 10%, 20%, and no HIV-infected cell DNA, respectively.

6.3.4. HIV-1 infection of UCB-derived HSPCs

In previous chapters, enrichment of $CD34^+$ HSPCs from UCB was achieved using FACS. However, for this study, an increased number of cells was required (> 5 x 10^6 $CD34^+$ HSPCs), which exceeded the number of cells that could be isolated from a single UCB unit. As shown in Chapter 2, Section 2.3.1, FACS enrichment results in increased cell loss, while MACS results in decreased purity of $CD34^+$ HSPCs. We therefore decided to pre-enrich for $CD34^+$ HSPCs using MACS prior to FACS. However, the already low number of $CD34^+$ HSPCs present in a single UCB unit further decreased upon sequential MACS and FACS enrichment, which ultimately resulted

in insufficient cell numbers to perform the experiments for this study using a single UCB unit. Multiple UCB units (three to five) were therefore pooled in an attempt to ensure isolation of a sufficient number of $CD34^+$ HSPCs for these experiments.

A single experiment was performed using pooled UCB-derived HSPCs from five independent donors to determine the optimal multiplicity of infection (MOI), exposure time and infection medium. Two MOIs (MOI3 and MOI10), two exposure times (24 and 48 hours), and the effect of cytokine stimulation on HSPC susceptibility to infection was simultaneously tested. Infection of HSPCs with CM9 was performed in medium with and without cytokines (FLT3L, SCF, TPO, IL-3 and G-CSF), which were included to determine whether the addition of cytokines affect the susceptibility of HSPCs to HIV infection *in vitro*. HSPC infection was determined by PCR of proviral DNA. DNA concentrations were high considering the number of cells from which DNA was extracted (Table 6.11). Protein contamination (260/280 ratio) was within range, while salt and small molecule contamination (260/230 ratio) was higher than typically accepted for PCR applications.

Table 6.11. DNA samples for HIV-1 proviral detection. The respective DNA concentrations, 260/280 and 260/230 ratios of the DNA samples analysed.

Sample description			DNA concentration (ng/µL)	260/280 ratio	260/230 ratio
24 hours	Cytokines	HIV^- (MOI3)	26.38	1.80	0.14
		HIV^+ (MOI3)	30.92	1.70	0.26
		HIV^- (MOI10)	26.37	1.96	0.12
		HIV^+ (MOI10)	24.45	1.76	0.63
	No cytokines	HIV^- (MOI3)	12.37	1.43	0.23
		HIV^+ (MOI3)	34.73	1.72	0.16
		HIV^- (MOI10)	11.23	1.58	0.17
		HIV^+ (MOI10)	28.24	1.78	0.46
48 hours	Cytokines	HIV^- (MOI3)	17.89	1.88	0.05
		HIV^+ (MOI3)	16.23	2.10	0.05
		HIV^- (MOI10)	21.15	1.98	0.07
		HIV^+ (MOI10)	23.31	1.80	0.05
	No cytokines	HIV^- (MOI3)	15.44	1.85	0.09
		HIV^+ (MOI3)	19.05	1.94	0.08
		HIV^- (MOI10)	33.24	2.05	0.10
		HIV^+ (MOI10)	28.83	1.98	0.09

HIV^+ = HIV-positive; HIV^- = HIV-negative; MOI = multiplicity of infection.

A gel image illustrating the data for 24 hours post-infection samples and the respective controls is shown in Figure 6.16. Proviral DNA was detected after 24 hours in cells infected at both MOI3 and MOI10 when cytokines were included. Proviral DNA was detected at MOI10 irrespective of cytokine addition. These results suggest that 24 hours post-infection is sufficient to detect HIV-1 proviral DNA. Interestingly, no proviral DNA could be detected 48 hours post-infection in samples exposed to HIV (data not shown).

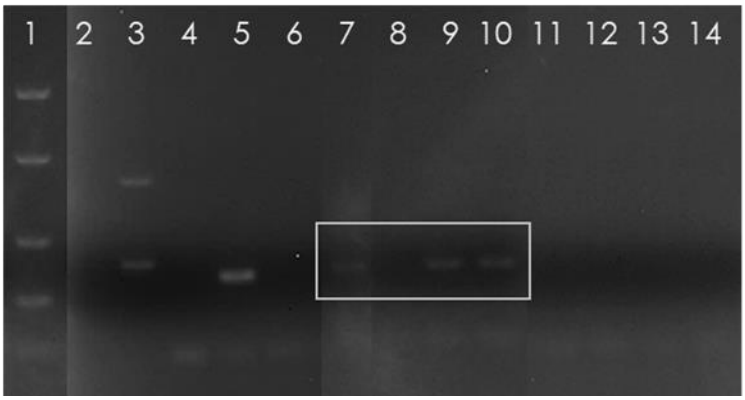


Figure 6.16. Gel image for detection of proviral DNA in HSPCs. Gel electrophoresis image representing detection of HIV-1 proviral DNA (LTR-*gag* amplicons, white box) in CD34$^+$ HSPCs at MOI3 and MOI10, 24 hr post-infection, with and without cytokines. Lane 1 represents the FastRuler™ LR DNA ladder. Lanes 2 and 4 represent a no template control (NTC) for the L32, positive control (lane 2) and LTR (lane 4) primer pairs. Lane 3 represents PCR products amplified using the L32 primer as the positive reaction control. Positive (CM9-infected PBMC DNA, lane 5) and negative (HIV-unexposed PBMC DNA, lane 6) controls were included to show primer specificity. Lanes 7 – 10 represent PCR products for MOI3 +CYT (lane 7), HIV+ MOI3 –CYT (lane 8), HIV+ MOI10 +CYT (lane 9) and HIV+ MOI10 –CYT (lane 10). Lanes 11 – 14 represent PCR products for HIV- MOI3 +CYT (lane 11), HIV– MOI3 –CYT (lane 12), HIV– MOI10 +CYT (lane 13) and HIV– MOI10 –CYT (lane 14).

Initial infection of HSPCs was performed using multiple UCB units. However, the logistical challenges associated with obtaining multiple UCB units on the same day resulted in the abandonment of UCB as a source of HSPCs for subsequent HIV experiments. The number of CD34$^+$ HSPCs in leukapheresis products is higher compared CD34$^+$ HSPCs in UCB (372), and leukapheresis products were therefore selected as an alternative source of CD34$^+$ HSPCs.

6.3.5. CD34$^+$ HSPCs from leukapheresis products

Initial attempts to thaw cryopreserved leukapheresis products yielded low viability, resulting in low numbers of viable CD34$^+$ HSPCs. In addition, the high number of non-viable cells present in leukapheresis products resulted in clot formation, which blocked the magnetic columns used for MACS and/or the sample line of the FACSAria cell sorter. It was therefore important to improve sample viability prior to MACS and/or FACS enrichment before we could proceed with experiments. In order to improve post-thaw viability the following parameters were optimised: processing temperature, resuspension solution and dead cell removal.

6.3.5.1. Processing temperatures

Initial thawing of leukapheresis products was performed in a 37°C incubator which took approximately 5 – 10 min. Subsequent processing steps were carried out at room temperature and solutions were pre-heated to 37°C. An optimised protocol from the ACT suggested using a waterbath, which decreased thawing time to less than 4 min. Sample processing at colder temperatures generally halts the metabolic activity of cells (373) and decreases the cytotoxicity of DMSO exposure. In addition, processing leukapheresis products on ice versus using cold (4°C) solutions and keeping cells at 4°C during processing was explored to improve post-thaw viability.

Leukapheresis products from two different donors were independently thawed in a 37°C water bath. The first product was diluted on ice, while the second product was diluted in pre-cooled PBS (4°C) and cell pellets were resuspended in cold (4°C) TP_4 buffer. Viability was assessed by flow cytometry as described in Chapter 2, Section 2.2.3.1. Post-thaw viability was significantly improved at 4°C processing (using pre-cooled solutions) compared to processing on ice (Figure 6.17). It is, however, important to note that two independent samples from two different donors were used to test processing temperatures. The improved viability could thus be sample specific. Sample processing at 4°C was therefore adopted for leukapheresis products and the trend of improved viability was sustained after switching from processing on ice to processing at 4°C.

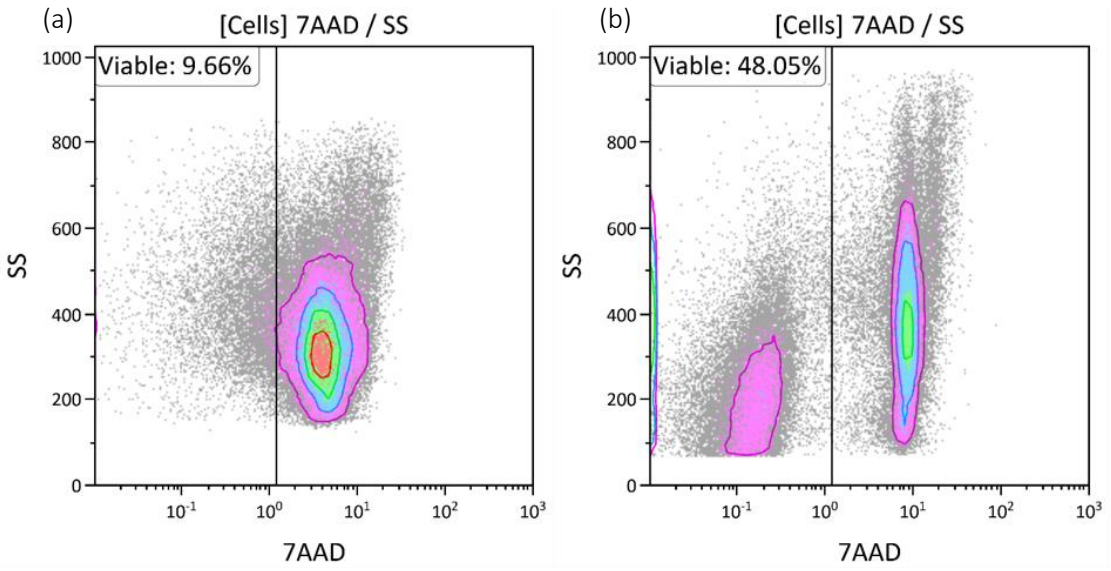


Figure 6.17. The effect of processing temperatures on leukapheresis product viability. Viability of cells processed and resuspended on (a) ice and (b) pre-cooled (4°C) solutions.

6.3.5.2. Resuspension solution

Two leukapheresis bags were thawed on two separate days. After the first centrifugation using pre-cooled (4°C) PBS, the supernatant was aspirated. One half of the cell pellets was resuspended in 50 mL cold RPMI and the other half in 50 mL cold TP_4 buffer. Sample viability was again determined by flow cytometry as before. Representative results are illustrated in Table 6.12 and Figure 6.18, and indicate that no significant difference in viability (7AAD-negative population) between RPMI and TP_4 buffer. However, different cell profiles were observed for cells resuspended in RPMI compared to TP_4 buffer, indicating a clear population of intact (7AAD-negative) cells present in RPMI-resuspended cells. RPMI was therefore used as the resuspension medium for future experiments.

Table 6.12. The effect of resuspension solutions on viability.

	Resuspension solution	
	RPMI	**TP4 buffer**
AP181122-01	88.87%	89.10%
AP181211	75.42%	89.94%

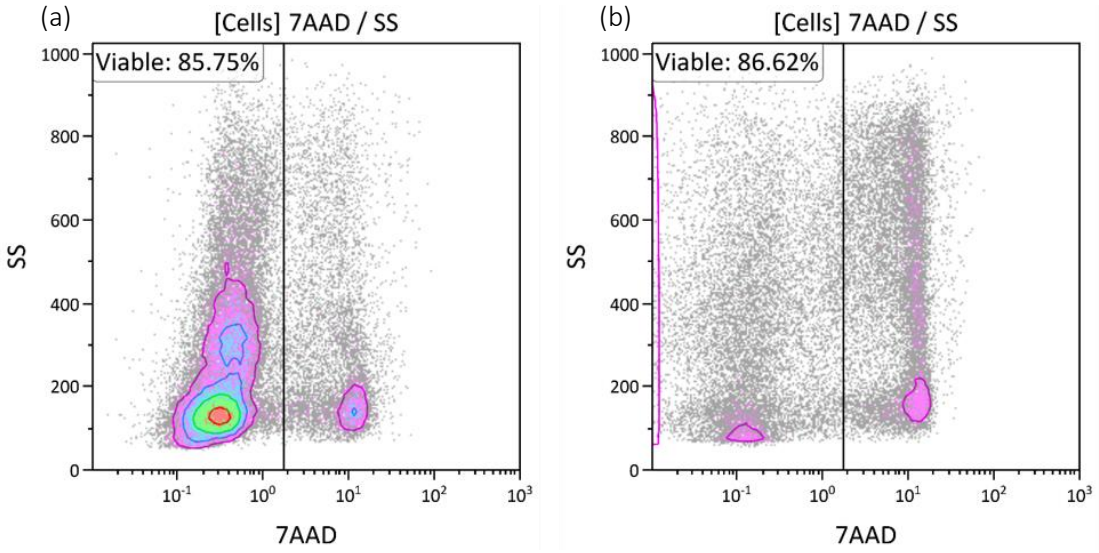


Figure 6.18. Effect of post-thaw resuspension solution on viability. Viability of cells resuspended in (a) RPMI and TP_4 buffer after thawing.

6.3.5.3. Removal of non-viable cells

Two leukapheresis products were thawed on two separate days to test post-thaw removal of dead cells by separately performing density-gradient- and slow-centrifugation steps after viability assessment (as described in the previous section). In previous experiments, poor separation was achieved after density-gradient centrifugation when cells were resuspended in TP_4 buffer, possibly due to the density of TP_4 buffer. Cells were therefore resuspended in RPMI and layered onto Histopaque-1077 (1:2 ratio). Mononuclear cells were collected after centrifugation. For the slow centrifugation, cells were resuspended in TP_4 buffer and centrifuged at 300 x g for 10 min without brake in an attempt to separate debris and intact cells.

Sample viability and percentage $CD34^+$ HSPCs were assessed to determine the number of $CD34^+$ HSPCs lost in the additional steps. Results comparing proportions of cells post-histopaque and post-slow centrifugation are shown in Table 6.13. Viability was not improved using either of the two techniques and no significant effect on the respective leukocyte or $CD34^+$ HSPC proportions was observed. For this reason, neither of the two techniques were included in the leukapheresis processing protocol for this study.

Table 6.13. Viability, leukocyte and CD34⁺ HSPC percentages after dead cell removal methods.

	Histopaque-1077			Slow centrifugation		
	Viability	Leukocytes	CD34⁺	Viability	Leukocytes	CD34⁺
AP181122-01	92.67%	89.96%	0.08%	88.38%	78.04%	0.02%
AP181211	76.14%	87.22%	0.65%	72.57%	92.42%	0.24%

6.3.6. HIV-1 infection of leukapheresis-derived CD34⁺ HSPCs

Due to time constraints and the numerous challenges encountered, only a single experiment was performed to determine the susceptibility of leukapheresis-derived HSPCs to HIV-1 infection. One leukapheresis product bag was thawed and processed as described in Section 6.2.4.3.2. Post-thaw counting indicated 76.57% viable cells, and viable $CD34^{+}CD45^{dim}$ HSPCs were present at a frequency of 1.81% (Figure 6.19a – d). The sample contained a total of 4.24×10^{9} viable leukocytes, and 3.02×10^{7} viable $CD34^{+}CD45^{dim}$ HSPCs. HSPCs were enriched by MACS and post-MACS enumeration indicated 87.20% viable $CD34^{+}CD45^{dim}$ HSPCs with a purity of 91.12% (Figure 6.19e – h). A total of 3.7×10^{6} $CD34^{+}$ HSPCs were FACS sorted from the MACS-enriched product as described in Section 6.2.4.4.

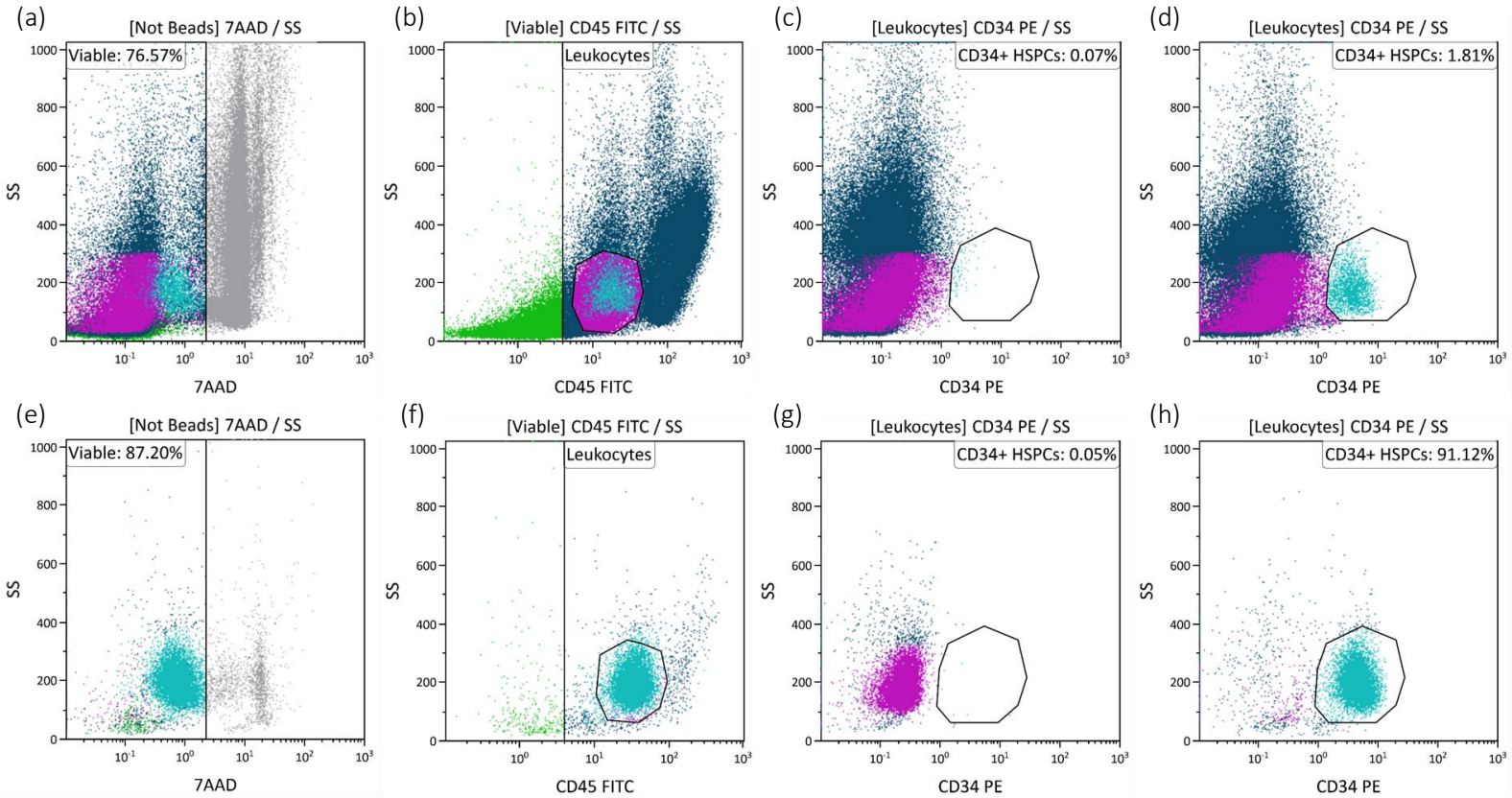


Figure 6.19. Viability and HSPC percentages after thawing and MACS-enrichment. Flow cytometry protocol and analysis of $CD34^+$ HSPCs after thawing and MACS-enrichment, respectively. (a) Colour dot plots showing viable (7AAD-negative) cells after thawing and cells stained with (b) CD45 FITC, (c) CD34 isotype and (d) CD34 PE after thawing. (e) Colour dot plot showing viable (7AAD-negative) cells after MACS enrichment. Colour dot plots showing cells stained with (f) CD45 FITC, (g) CD34 isotype and (h) CD34 PE after MACS-enrichment. Colour dot plots allow for tracking of cells in all plots of the flow cytometry protocol used. Turquoise cell population indicate HSPCs; grey cell population indicate dead (7AAD-positive) cells; purple cell population represents lymphocytes; dark blue cell population is representative of all leukocytes; green cell population indicated all viable cells. Colour tracking is based on the following hierarchy precedence (referring to the preferred visualisation of the population): HSPCs (turquoise) > Lymphocytes (purple) > Leukocytes (blue) > Viable cells (green), all cells (grey).

Immunophenotypic analysis was performed on MACS and FACS enriched cells as described in Section 6.2.4.7. As expected, 85.55% of the cells were viable and 93.42% expressed CD34 after sorting (Figure 6.20a). Most of the CD34$^+$ cells did not express any lineage markers (99.69%, Figure 6.20b), while 83.09% expressed CD133 (Figure 6.20c) and 64.92% expressed CD38 (Figure 6.20d). CD34$^+$ HSPCs expressed low CXCR4 (21.56%) and CCR5 (12.79%) (Figure 6.20e). Even though CD4 expression was not detected in these CD34$^+$ HSPCs, a small fraction co-expressed CXCR4 and CCR5 (8.19%). The negative/positive boundaries for CD4, CCR5 and CXC4 were set using the appropriate isotypic controls and were kept the same for analysis of the sample aliquot stained with the respective mouse anit-human antibodies. However, we acknowledge that isotypic controls are not always "perfect" controls and the 8.19% co-expression might be an overestimate (based on the lack in evidence of a clear separation between negative and positive populations).

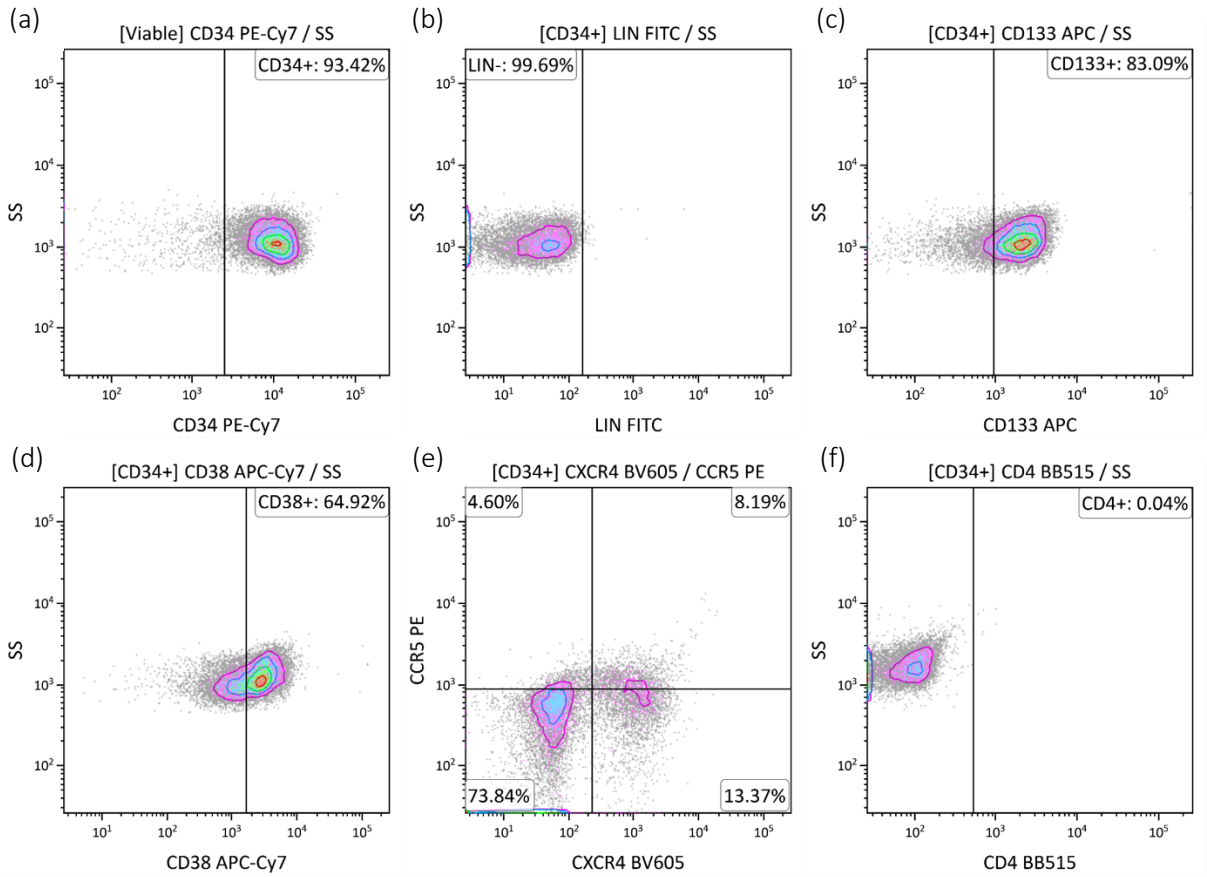


Figure 6.20. Immunophenotype of CD34$^+$ HSPCs on the day of isolation. Immunophenotype of CD34$^+$ HSPCs after MACS- and FACS-enrichment. Density plots showing cells stained with (a) CD34 PE-Cy7, (b) LIN FITC, (c) CD4133 APC, (g) CD38 APC-Cy7, (h) CCR5 PE, CXCR4 BV605 and (f) CD4 BB515.

The sorted CD34⁺ HSPCs were placed in medium and incubated overnight to recover from the sequential MACS and FACS. After overnight recovery, CD34⁺ HSPCs were infected with control medium (unexposed control medium) and CM9-containing supernatant. Comparison using different HIV-1C tropic viruses would have required more HSPCs than we were able to isolate for this experiment. Previous experiments indicated an MOI of 3 to be sufficient for HSPC infections. However, the low volumes of virus available and propagations resulting in non-infectious virus, resulted in us having to use a lower MOI (MOI2). Immunophenotypic analysis on the day of infection showed similar HSPC marker expression compared to the day of isolation. However, CXCR4 expression increased to > 70% (Figure 6.21b) overnight and only 9.31% of the cells were viable based on 7AAD staining (Figure 6.21a).

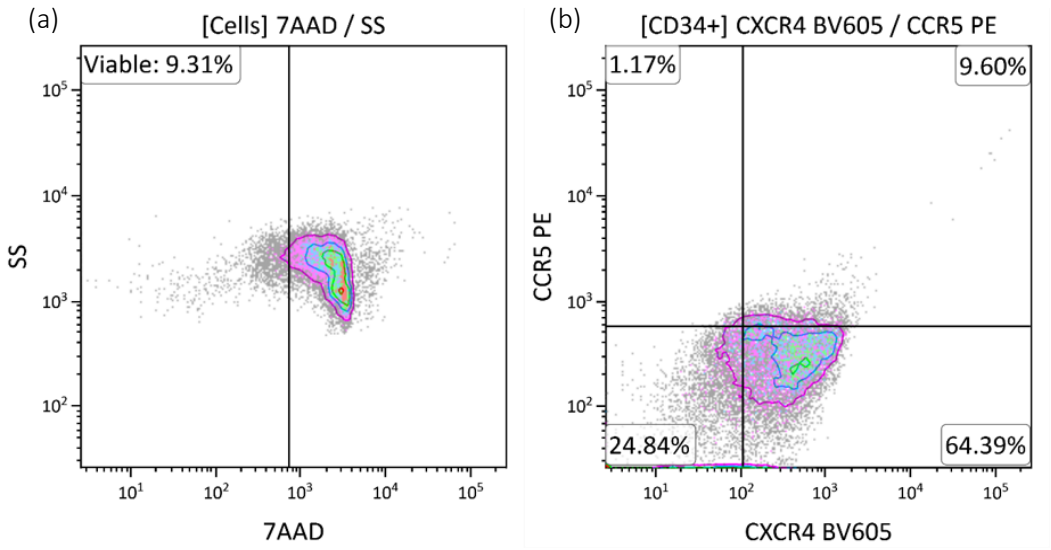


Figure 6.21. Viability and immunophenotype of CD34+ HSPCs after overnight incubation. Viability and immunophenotypic analysis of CD34⁺ HSPCs after overnight incubation. Density plots showing cells stained with (a) 7AAD and (b) CCR5 PE and CXCR4 BV605.

Early apoptotic cells can recover in appropriate medium with the addition of cytokines (248). Even though 7AAD detects necrotic cells and not early apoptotic cells, the cells were still sorted for expansion in the hope that they would recover. Unfortunately, the cells that were seeded for expansion did not expand and the experiment was terminated. Since the initial viability of the sample was good, failure to expand could be due to a decrease in viability observed after sequential MACS and FACS enrichment. The cells that remained after expansion and CFU sorting were pooled for HIV-1 nucleic acid detection as described in Section 6.2.2.3. The DNA

concentrations ranged from between 19 to 27 ng/µL, with relatively high and low 260/280 and 260/230 ratios, respectively (Table 6.14). DNA purification resulted in improved 260/280 and 260/230 ratios. However, the DNA concentrations were significantly decreased and therefore these DNA samples could not be used to detect HIV-1 proviral DNA. The DNA concentrations, 260/280 and 260/230 ratios of the purified DNA are shown in Table 6.14.

Table 6.14. DNA concentration of HSPC sample before and after purification. DNA samples with their respective DNA concentrations, 260/280 and 260/230 ratios.

		Pre-purification			Post-purification		
		DNA (ng/µL)	260/280	260/230	DNA (ng/µL)	260/280	260/230
Cytokines	Control	21.50	3.32	0.79	13.92	2.07	2.05
	Control medium	23.85	3.01	0.87	3.46	2.24	1.48
	HIV supernatant	22.55	3.09	0.51	1.90	1.60	1.12
No cytokines	Control	19.92	3.72	0.72	2.71	2.07	0.93
	Control medium	22.38	3.36	0.76	1.35	22.09	1.36
	HIV supernatant	27.86	2.85	0.89	13.08	2.31	1.89

Due to the reduced viability after sequential MACS and FACS enrichment, we decided to repeat the experiment in an attempt to determine the cause for the decreased viability observed. The thawed leukapheresis product was split, one half was used to enrich for $CD34^+$ HSPCs using MACS, while the other half was used to enrich for $CD34^+$ HSPCs using FACS. MACS-enrichment yielded 3 x 10^6 $CD34^+$ cells with a low purity (50%), while a total of 2.5 x 10^5 $CD34^+$ HSPC were FACS sorted with a higher purity (> 90%). The purified $CD34^+$ HSPCs were cultured for three days in the presence of the five cytokines as before. Viability was determined after 72 hours and most of the cells were non-viable. Due to time constraints, the cause of the reduced viability of $CD34^+$ HSPCs from leukapheresis products after enrichment with either MACS and/or FACS could not be resolved.

6.4. DISCUSSION AND CONCLUSION

The purpose of this study was to determine whether HSPCs are susceptible to HIV-1 infection and whether a subset of HSPCs exist that are resistant to infection.

6.4.1. HIV-1C isolates

Most HIV-related studies make use of HIV-1B viruses or use patient samples from cohorts infected with HIV-1B. However, individuals infected with HIV-1B only represent a fraction of the infections worldwide. Most HIV-1 infections in India, Brazil and sub-Saharan Africa are due to HIV-1C (142–144). Primary HIV-1C viruses were used for this study with varying levels of infectivity (Table 6.12). It has been estimated that only 1% of HIV particles are infectious (374). Various factors can affect the mutation frequency and assembly of HIV, resulting in different proportions of infectious virus (375). The use of primary HIV-1 isolates represent viral populations replicating *in vivo.* However, to retain high viral infectivity and reduce genomic variation, HIV-1C molecular clones will need to be considered for future experiments, and these experiments planned for the near future. HIV-1 molecular clones are replication-competent HIV-moieties that are cloned into plasmids to produce clonally identical infectious HIV particles when transfected into a packaging cell line (376). Reverse transcription does not occur during the production of molecular clones by transfection, since the viral genome is encoded by a plasmid which eliminates loss-of-function mutations (377). The use of HIV-1C molecular clones would reduce the proportion of non-infectious or replication-incompetent virus produced. Two studies have demonstrated successful production of HIV-1C molecular clones from primary isolates (378,379).

Large volumes of viral supernatant were added to $CD34^+$ HSPCs to reach the desired MOI despite attempts to concentrate the virus during propagation. As the volume surrounding the cells increases, the efficiency of infection decreases (380). There are several methods available to concentrate virus, which include: polyethylene glycol (PEG) precipitation (381), ultrafiltration (380) and ultracentrifugation (382). Use of these methods in the future might enable us to concentrate the propagated virus and reduce the volumes added to cells in future experiments.

6.4.2. HIV-1 detection methods

Several HIV-related methods were established and optimised in our laboratory for the detection and titration of HIV. The p24 ELISA was not used to titrate the virus, since p24 is released from all cells and could originate from defective HIV (383). It was therefore used rather to determine productive infection. p24 ELISA is a more suitable method for titrating HIV-1 molecular clones, assuming that each HIV-1 clone produced is infectious. Since primary HIV-1C was used for this study and production of non-infectious virus is well-described (375), p24 ELISA was considered to be unsuitable for titrations. Functional titrations to determine the number of infectious units per millilitre were performed using the GHOST infectivity assay.

A PCR-based HIV-1 proviral detection method was optimised with HIV-1C-specific primers. The lowest limit of detection of this method was established to be 5000 infected cells per 10^5 total cells. A study by Izopet *et al.* (384), was able to detect four proviral genomes per 10^6 total cells. The infection frequency of highly susceptible $CD4^+$ T cells *in vivo* is low (< 1%) (385–387). Since the frequency of HSPC infection is not known, but expected to be less than $CD4^+$ T cells, lack of detection of the method used in this study (PCR) is likely not to be sensitive enough to detect few proviral genomes. Sensitivity of our current method therefore needs to be improved or a similar more sensitive method needs to be established to detect proviral DNA in HSPCs. Some studies suggest that HSPCs are susceptible to infection, but that a pre-integration block prevents the HIV genome from integrating the host cell genome (188,388). An integration detection method such as *Alu-gag* PCR, which allows for amplification of HIV DNA only when integrated, still needs to be established in our laboratory.

Resting $CD4^+$ T cells are a well-described HIV reservoir (176,177). HSPCs are another long-lived cell type able to self-renew and if/when infected will produce cell progeny that harbour integrated HIV. Long-lived cells that harbour replication-competent provirus make eradication of the virus difficult, since cART does not act on stable integrated provirus (389). Most proviral genomes are inactive and detectable by PCR-based methods. However, PCR provides no information on whether the provirus is replication-competent or not. The viral outgrowth assay

is currently the method of choice to determine replication compentency of virus produced. Establishing the viral outgrowth assay in our laboratory would be important for future HIV research in our laboratory.

6.4.3. HIV-1 infection of $CD34^+$ HSPCs

Controversy exists in the literature on whether HSPCs are susceptible to infection with HIV. Several studies have shown that HSPCs express receptor and co-receptors and are therefore susceptible to infection with HIV and serve as a latent reservoir (183,184,189,192), while other studies suggest that these cells are resistant to infection (185,186,188). $CD34^+$ HSPCs encompass a heterogeneous population of cells including HSCs, early and late progenitors. HIV integration and replication *in vitro* is dependent on the activation state of the target cells (367). However, very few T cells, the primary target of HIV, are active *in vivo* (367). This is also true for freshly isolated HSPCs from UCB, where > 90% of the cells are dormant (309).

Cytokines regulate hematopoiesis and are often added to cultures for the purpose of expansion (Chapter 4). The addition of cytokines to *in vitro* cultures activates HSPCs to divide. Stevenson *et al.* (367) have shown that HIV is able to enter target cells and remain extrachromosomal until the cell is activated. HIV is only able to integrate into the host cell genome upon cell activation. Once cells divide they are considered to be active and more susceptible to infection with HIV. The initial infection experiment of UCB-derived HSPCs was unable to detect HIV proviral DNA in HSPCs infected without cytokines. It is therefore possible that these cells are not susceptible to infection initially, but rather become susceptible upon culturing. Whether susceptibility is due to the addition of cytokines is currently unknown and needs further investigation. Infection of HSPCs has been performed with (190,192,196,390) and without cytokines (188,364,388), possibly contributing to the lack of consensus regarding the susceptibility of HSPCs to HIV infection.

This study was unable to detect CD4 on $CD34^+$ HSPCs; however, HSPC subsets expressing CXCR4 and CCR5 were identified. The subset of CCR5-expressing HSPCs was smaller than the subset of CXCR4-expressing HSPCs. Contrary to the findings of this study, Sebastian *et al.* (184) identified $CD34^+$ progenitors expressing CD4 that harbour X4- and R5-tropic HIV. Nixon *et al.*

(189) also showed that common myeloid progenitors (CMP) and granulocyte-megakaryocyte progenitors (GMP) express CCR5 and can become infected with R5-tropic virus. Carter *et al.* (192) indicated that X4-tropic virus is more likely to infect HSPCs. The use of CD4 in combination with either co-receptors is well-described. However, several CD4-independent mechanisms have been described and therefore these cells cannot be completely disregarded as target cells (170,172,391,392). Another possibility is that HSPCs express CD4, but at very low levels, undetectable with the detection approaches used in this study.

This study revealed expression of CXCR4 on $CD34^+$ HSPCs of approximately 20% on the day of isolation which increased to > 70% after overnight incubation. Carter *et al.* (192) demonstrated *in vivo* and *in vitro* susceptibility of HSPCs to HIV infection. However, increased expression of CXCR4 upon culturing would result in increased susceptibility of HSPCs to infection, which is not inherent. A study by Skinner *et al.* (393) likewise showed upregulation of CXCR4 cell surface expression on murine HSPCs after overnight culture. It has been suggested that $CD34^+CXCR4^-$ HSPCs represent a more primitive population of HSPCs compared to the $CD34^+CXCR4^+$ cells, which have already initiated commitment towards the lymphoid lineage (394). A study by Rutella *et al.* (395) showed that G-CSF upregulates CXCR4, since CXCR4 was found to be expressed on > 90% of mobilised peripheral blood $CD34^+$ HSPCs, which is inconsistent with the findings from this study. The results from this study are based on a single experiment using a single biological replicate. Therefore, these experiments would need to be repeated. If this observation can be repeated then *in vitro* infection of HSPCs would need to be evaluated. HSPCs from HIV-infected individuals for this type of experiment would be an advantage and a true reflection of HSPC susceptibility and response to HIV.

In conclusion, failure of the infection experiment was likely due to loss of viability, which occurred between MACS and FACS enrichment and the day of infection. It was not caused by the combination of MACS and FACS enrichment, since the last experiment performed MACS and FACS separately, which still resulted in reduced viability. Before infection experiments are repeated, the decreased viability of the leukapheresis products need to be resolved. Different overnight media could also be tested to determine whether this contributed to the reduced viability. For future *in vitro* experiments, the use of primary isolates versus molecular clones

and the source of HSPCs would need to be considered. Primary HIV-1 isolates acquire mutations upon viral replication, but might be considered more clinically relevant, whereas the use of molecular clones will produce virus that is less mutation prone. UCB is an easy accessibly source of HSPCs. However, the number of $CD34^+$ HSPCs that can be isolated from these products is low and often not sufficient for experimental purposes. Leukapheresis products generally contain more $CD34^+$ HSPCs. However, the loss of viability after enrichment would need to be resolved before further experiments are conducted.

Hematopoietic stem cell gene therapy is a promising approach to treat HIV. A pilot clinical trial has reported on the safety of stem cell-based gene therapy for HIV treatment (396). Genetically modified and non-modified autologous HSPCs were transplanted and $CD34^+$ HSPCs in the non-modified graft exceeded the number of $CD34^+$ HSPCs in the modified graft, resulting in low level gene marking *in vivo* (396). Although promising, this type of approach would require the presence of significant numbers of long-term repopulating HSPCs to enable successful long-term engraftment of gene-modified cells. One aspect that could result in this approach not succeeding is the presence of proviral DNA in HSPCs. It would therefore be important to identify a population of HSPCs resistant to HIV infection or inactivate latent proviral genomes from infected HSPCs for gene therapy.

CHAPTER 7. CONCLUDING REMARKS AND FUTURE PERSPECTIVES

This study consisted of multiple aims relating to the use of HSPCs and their clinical applications. In order to achieve the first aim, an isolation method producing pure and viable $CD34^+$ HSPCs was needed. This included testing two isolation techniques, MACS and FACS. Second, serum- and animal-component free culture conditions and optimal cytokine combination needed to be determined for the *ex vivo* culturing of HSPC. Different serum-free media and two commonly used cytokine combinations, FLT3L, SCF, TPO, IL-3 and FLT3L, SCF, TPO, IL-6, were tested with and without the addition of G-CSF to determine their influence on HSPC expansion.

Comparing MACS and FACS for the isolation of HSPCs indicated that it is not essential to perform MACS before FACS for enrichment of HSPCs, as is usually done in the field. MACS yields greater HSPC cell numbers with acceptable sample viability, while FACS results in pure and viable HSPCs, but with increased cell loss. Therefore, for experiments that require high HSPC numbers, MACS isolation should be the preferred method of choice, whereas, if increased purity is desired, FACS isolation would be better suited. Cell loss is inevitable and a great disadvantage regardless the isolation technique.

Expansion of HSPCs aims to increase both short- and long-term repopulating cells to ultimately enable successful short- and long-term engraftment following HSCT. The three different serum-free culture conditions demonstrated improved expansion of HSPCs compared to DMEM supplemented with FBS when cultured for seven days. Expansion was measured by viability, proliferation, HSPC immunophenotype and SP analysis. We found that StemSpan ACF is better suited for *ex vivo* culturing of HSPCs and allows compliance with GMP standards. We do, however, acknowledge that this study did not include all serum-free media commercially available. This study further revealed that heterogeneity exists in response to different media conditions, which might be due to different serum-free medium formulations favouring and increasing certain populations of cells. Greater expansion was observed in cultures supplemented with FLT3L, SCF, TPO, IL-3 and G-CSF and was therefore used for subsequent HSPC expansion and culturing. This study did not include functional *in vitro* and *in vivo* assays, which would need to be performed to confirm the presence of LT-HSCs following *in vitro*

expansion of HSPCs. Expansion of HSPCs could be applied to other non-stem cell-based therapies, such as cases of prolonged neutropenia, where patients have increased risk of infection. This particular type of treatment would require a large number of cells and could therefore benefit from HSPC expansion. The cytokines and combinations of cytokines used would depend on individual applications. However, it will be essential to move towards global standardisation of *ex vivo* HSPC culture conditions in the future.

To achieve the second aim, the optimal concentration of SR1 first needed to be determined which was once again measured by viability, proliferation, HSPC immunophenotype and SP analysis. A seven-day *ex vivo* expansion of UCB-derived $CD34^+$ HSPCs, showed an SR1 concentration-dependent increase in $CD34^+$ sub-populations, while a decrease was observed in the number of $CD34^-$ HSPCs. No functional assays were included during this study; several studies have however shown the engraftment potential of SR1-treated HSPCs in *in vivo* mouse models and in a recent Phase I/II clinical trial where SR1-expanded HSPCs were co-transplanted with non-expanded cells. Based on these results, 1 µM SR1 was used for the expansion of $CD34^+$ HSPCs for gene expression experiments.

The effect of SR1 on gene expression of HSPCs has previously been evaluated for shorter culture periods. To the authors' knowledge, this is the first time that the effect of SR1 has been reported on the transcriptome of seven-day expanded HSPCs. This is relevant since SR1 is currently being used to expand HSPCs *ex vivo* for extended periods. This study did not report differential expression among the AhR pathway-related genes, such as *AHRR and HSP90AA1,* as other studies have shown. This might suggest that SR1 initially acts on the above-mentioned genes, but that this effect is restored after seven days. Cytochrome P450 1B1 (*CYP1B1)* and erythrocyte membrane protein band 4.1-like 3 (*EPB41L3*) were most significantly reduced after seven-day expansion with SR1. Downregulation of *CYP1B1* has previously been observed at earlier time points, but was even further reduced in this study possibly due to continuous downregulation over a longer period compared to shorter time periods assessed by other studies. This suggests that the effect of SR1 on *CYP1B1* seems to be long lasting, since it remains

significantly downregulated after seven days of expansion with SR1. The effect of SR1 is reversible, suggesting that *CYP1B1* expression would return to normal once SR1 is removed from the culture medium.

This is the first time that gene expression analysis has been compared between SR1-expanded and non-expanded $CD34^+$ HSPCs. The gene expression profiles between seven-day expanded $CD34^+$ HSPCs and non-expanded $CD34^+$ HSPCs were noticeably different. GO classification revealed upregulated genes in expanded vs. non-expanded $CD34^+$ HSPCs enriched for processes including regulation of cell cycle, macrophage activation, DNA replication, neutrophil activation and positive regulation of population proliferation. This indicates that *ex vivo* expansion of cells activates genes involved in proliferation and differentiation, as expected, since several cytokines were included in addition to SR1 to promote expansion of HSPCs. GO classification revealed downregulated genes in expanded vs. non-expanded $CD34^+$ cells to be enriched for processes such as chromatin silencing at recombinant DNA (rDNA), nucleosome assembly and regulation of myeloid cell differentiation. The limitations of the study include that these experiments were performed using bulk RNA, which represents the average expression of the cells in the population of interest, while the effect on smaller populations is often masked. Another limitation is that an increased number of cells is generally required to obtain sufficient good quality RNA. To obtain enough cells and RNA from smaller sub-populations, such as the $CD34^-$ population of HSPCs, is challenging. It would therefore be beneficial to study the effect of SR1 on $CD34^+$ and $CD34^-$ HSPC populations at a single-cell transcriptome level. This would allow us to determine whether the effect of SR1 is more pronounced in specific sub-populations. The use of $CD34^+$ HSPCs exclusively expanded with SR1 would be useful in cases where the HSPC cell dose in the initial harvested cell therapy products is suboptimal and therefore not a feasible option for HSCT when used alone. The use of SR1 for *ex vivo* expansion of cell therapy products is promising, however, the long-term effects are still unknown and need to be elucidated. Novel expansion agents with similar expansion and engraftment properties to SR1 are likely to emergence in future and may provide greater benefit than those currently available.

To achieve the third aim, a total of 122 single cells were analysed and four definitive clusters were identified. The clusters identified and the genes expressed within the various clusters are in strong agreement with previously defined clusters. We identified two populations of earlier undifferentiated HSPCs (Cluster 1 and 4). Two additional populations had gene signatures associated with GMP and neutrophil development (Cluster 2) and erythrocyte progenitors (Cluster 3). Our data therefore suggest that the transcriptomic profiles of progenitor cells are sufficiently distinct from early, primitive HSPC populations to allow resolution of these populations with a limited number of single cells analysed. However, a clear pitfall of this study is the limited number of cells analysed compared to other similar studies, which did not allow discovery of the total extent of heterogeneity within the $CD34^+$ HSPC population. The use of more efficient single-cell platforms with the ability to capture more cells will be considered for future experiments in our laboratory in order to be competitive in this rapidly-growing field.

Several other studies have explored this heterogeneity and have revealed diversity with regard to HSPC differentiation potential and functional ability, which is not exclusive to the $CD34^+$ population but extends to the sub-populations. Even though the hierarchical model of hematopoiesis has been studied for many years, single-cell technologies are challenging the classical HSPC hierarchy model, which is more complex than previously assumed. This fast-evolving technology provides a powerful tool to dissect cellular heterogeneity and uncover the uniqueness of each cell, and also to identify new markers and pathways that previously could not be identified. New emerging single-cell techniques enable the simultaneous study of multiple omics (genome, transcriptome and proteome) in the same cell. This will further contribute to and improve our understanding of the heterogeneity of cell populations in the future.

To achieve the fourth aim, several experimental components were performed before HSPCs could be infected with HIV. Primary HIV-1C virus stocks needed to be produced with sufficient titres to infect target cells. Detection of target cell infection was required to determine whether the effect observed was die to direct infection of HSPCs with HIV.

HIV productions were performed based on the Montefiori protocol and adapted to suit the unique replication dynamics of our primary isolates. This enabled successful production of two primary isolates, CM9 (R5X4-tropic) and SW7 (X4-tropic) with relatively low infectivity. Despite several attempts to produce infectious R5-tropic virus from various isolates, severe loss of infectivity possibly due to the mutation frequency of HIV resulted in us only using the above-mentioned isolates. The use of primary HIV-1 isolates represents viral populations replicating *in vivo.* However, to retain high viral infectivity and reduce genomic variation, HIV-1C molecular clones need to be considered for future experiments. Reverse transcription does not occur during the production of molecular clones by transfection, since the viral genome is encoded by a plasmid which eliminates loss-of-function mutations. The use of HIV-1C molecular clones would reduce the proportion of non-infectious or replication-incompetent virus produced.

Several HIV-related methods were established and optimised in our laboratory for the detection and titration of HIV. Functional titrations to determine the number of infectious units per millilitre were performed using the GHOST infectivity assay. These HIV-susceptible target cells were also useful to optimise the PCR detection method and establish the limit of detection of this assay. The lowest limit of detection was established to be 5000 infected cells per 10^5 total cells (5%). The infection frequency of HSPCs is not known, but is expected to be less than $CD4^+$ T cells, which is roughly 1%. Lack of detection using our current PCR method would therefore be due to the inability of this assay to detect few proviral genomes. Sensitivity of the current method would need to be improved or a similar more sensitive method needs to be established to detect proviral DNA in HSPCs. Some studies suggest that HSPCs are susceptible to infection, but that a pre-integration block prevents the HIV genome from integrating into the host cell genome. An integration detection method such as *Alu-gag* PCR, which allows for amplification of HIV DNA only when integrated, still needs to be established in our laboratory. However, PCR provides no information on whether the provirus is replication-competent, which is achieved through viral outgrowth assays. This method would allow us to determine whether HSPCs are latently infected with replication-competent provirus.

Controversy exists in literature as to whether HSPCs are susceptible to infection with HIV. Several studies have shown that HSPCs express the receptor and co-receptors and are therefore susceptible to infection with HIV and serve as a latent reservoir, while other studies suggest that these cells are resistant to infection. The initial infection experiment of UCB-derived HSPCs was unable to detect HIV proviral DNA in HSPCs infected without cytokines. It is therefore possible that these cells are not susceptible to infection initially, but rather become susceptible upon culturing. Whether susceptibility is due to the addition of cytokines is still not completely known. Although UCB is an easy accessible source of HSPCs and was used initially for infection experiments, the number of $CD34^+$ HSPCs is low and the logistical challenges associated with obtaining multiple UCB units resulted in the abandonment of UCB as a source of HSPCs for these experiments. The number of $CD34^+$ HSPCs in leukapheresis products is generally higher compared to $CD34^+$ HSPCs in UCB which led us to use leukapheresis products as an alternative source of HSPCs.

A single experiment was performed to determine the susceptibility of leukapheresis-derived HSPCs to HIV-1 infection. HSPCs were enriched by sequential MACS and FACS and left to recover overnight. Viability analysis the following day indicated that only a small fraction (9.31%) of the cells were viable. A similar experiment was performed to determine the cause for the decreased viability observed. A thawed leukapheresis product was split, one half was used to enrich for $CD34^+$ HSPCs using MACS and the other half using FACS. Again, most of the cells were non-viable after 72 hours. Due to time constraints, the reduced viability of $CD34^+$ HSPCs from leukapheresis products after enrichment with either MACS and/or FACS could not be resolved. Before infection experiments are repeated, the decreased viability of the leukapheresis products need to be resolved.

Expression of receptor and co-receptors on HSPCs was assessed by flow cytometry at various timepoints. We were unable to detect CD4 on $CD34^+$ HSPCs; however, HSPC subsets expressing CXCR4 and CCR5 were present. The subset of CCR5-expressing HSPCs was smaller compared to CXCR4-expressing HSPCs. Interestingly, expression of CXCR4 on $CD34^+$ HSPCs was approximately 20% on the day of isolation and increased to > 70% after overnight incubation. Increased expression of CXCR4 upon culturing would result in increased susceptibility of HSPCs

to infection, which is not inherent. However, these findings are based on a single experiment using a single biological replicate and would therefore need to be verified. If this observation is confirmed then *in vitro* infection of HSPCs need to be evaluated. HSPCs from HIV-infected individuals for this type of experiment would be an advantage and a true reflection of HSPC susceptibility and response to HIV.

Hematopoietic stem cell gene therapy is a promising approach to treat HIV. A pilot clinical trial has reported on the safety of stem cell-based gene therapy for HIV treatment (396). Genetically modified and non-modified autologous HSPCs were transplanted, and $CD34^+$ HSPCs in the non-modified graft exceeded the number of $CD34^+$ HSPCs in the modified graft, resulting in low level gene marking *in vivo* (396). Although promising, this type of approach would require the presence of significant numbers of long-term repopulating HSPCs to enable successful long-term engraftment of gene-modified cells. One aspect that could result in this approach not succeeding is the presence of proviral DNA in HSPCs. It would therefore be important for gene therapy to identify a population of HSPCs that is resistant to HIV infection or to inactivate latent proviral genomes from infected HSPCs.

The ability to expand HSPCs while maintaining their stem cell properties has several advantages. Expansion of HSPCs to clinically relevant cell numbers in cases where insufficient numbers are available, will further extend the use and benefits of HSCT-based therapies. Exploring the heterogeneity of both expanded and non-expanded stem cell products will be key to the future success of these therapies in treating a variety of diseases including HIV. In addition, a better understanding of the balance of interactions in the HSPC microenvironment, and how proliferation and differentiation of HSPCs are affected by HIV, could result in enhanced treatment strategies in future.

APPENDIX A. RUNNING THE TUBE CONTROLS

A.1. RNA EXTRACTION AND PURIFICATION

The RNeasy Micro Plus kit (Qiagen, Germany) was used to perform RNA extractions from $CD34^+$ HSPCs. Prior to starting, 10 µL of β-Mercaptoethanol (Sigma-Aldrich, USA) was added to 1 mL RLT buffer. A total of 350 µL RLT buffer was added to 20 µL (100 – 200 cells/µL) of cells. The mixture was vortexed for 1 min. The lysate was transferred to a gDNA Eliminator column in a 2 mL collection tube and centrifuged at 8000 x g for 30 seconds. This removed the genomic DNA from the sample. The gDNA column was discarded after centrifugation. 70% Ethanol (350 µL) was added to the flow-through and mixed by pipetting. The sample was transferred to an RNeasy MinElute spin column in a 2 mL collection tube and centrifuged at 8 000 x g for 15 seconds and the flow-through was discarded. RW1 buffer (700 µL) was added to the MinElute column and centrifuged at 8000 x g for 15 seconds and the flow-through was again discarded. Buffer RPE (500 µL) was added to the MinElute column and centrifuged at 8000 x g for 15 seconds and the flow-through was once again discarded. 80% Ethanol (500 µL) was added to the MinElute column and centrifuged at 8000 x g for 2 min to wash the spin column membrane. After centrifugation the collection tube with the flow-through was discarded. The RNeasy MinElute column was transferred to a new collection tube and centrifuged at maximum speed for 5 min to dry the membrane of the spin column. The collection tube with the flow-through was discarded. The MinElute column was placed in a new 1.5 mL collection tube and 14 µL of RNase-free water was added directly to the center of the membrane and centrifuged at maximum speed for 1 min to elute the RNA.

A.2. LYSIS, REVERSE TRANSCRIPTION AND AMPLIFICATION

The lysis mixture was prepared as described in Chapter 5. A positive control (tube 1) and a no template control (NTC, tube 2) were prepared as indicated in Table A.1 below. The lysis thermal cycle (Table A.2) was then run in a Thermal Cycler (Gene Amp®, PCR System 9700, Applied Biosystems, Life Technologies, Thermo Fisher Scientific, USA).

Table A.1. Tube control lysis mixture.

Components	Tube 1: Positive control	Tube 2: NTC
RNA	1 µL	-
Cell wash buffer	-	1 µL
Lysis mixture	2 µL	2 µL
Total	**3 µL**	**3 µL**

Table A.2. Lysis thermal cycle.

Temperature	Time
72°C	3 min
4°C	10 min
25°C	1 min
4°C	hold

The reverse transcription mixture was prepared as described in Chapter 5, combined with the lysis product (3 µL) from the previous step (Table A.3) and mixed briefly. The reverse transcription reaction (Table A.4) was run in a Thermal Cycler as before.

Table A.3. Tube control reverse transcription mixture.

Components	Tube 1: Positive control	Tube 2: NTC
Cell lysis product	3 µL	3 µL
Reverse transcription mixture	4 µL	4 µL
Total	**7 µL**	**7 µL**

Table A.4. Reverse transcription cycle.

Temperature	Time
42°C	90 min
70°C	10 min
4°C	hold

The PCR mixture was prepared as described in Chapter 5 and combined with 1 µL of the reverse transcription reaction product from the previous step (Table A.5) and mixed briefly. The PCR reaction (Table A.6) was run in a Thermal Cycler as before.

Table A.5. Tube control PCR mixture.

Components	Tube 1: Positive control	Tube 2: NTC
PCR mixture	9 µL	9 µL
Reverse transcription product	1 µL	1 µL
Total	**10 µL**	**10 µL**

Table A.6. PCR cycle.

Temperature	Time	Cycles
95°C	1 min	1
95°C	20 sec	5
58°C	4 min	
68°C	6 min	
95°C	20 sec	9
64°C	30 sec	
68°C	6 min	
95°C	30 sec	7
64°C	30 sec	
68°C	7 min	
72°C	10 min	1
4°C	hold	

A.3. DILUTE PRODUCTS

The prepared products were briefly vortexed and spun down to collect the content at the bottom of the tube. The final PCR-amplified product was diluted by combining it with the C1 DNA Dilution Reagent as shown in Table A.7.

Table A.7: Final dilution.

Components	Tube 1: Positive control	Tube 2: NTC
PCR product	1 µL	1 µL
C1 DNA Dilution Reagent	45 µL	45 µL
Total	**46 µL**	**46 µL**

The final diluted tube controls were quantified using the Quantus and the cDNA fragment size distribution was determined using the TapeStation 2200, as described in Chapter 5.

APPENDIX B. SINGLE-CELL R SCRIPTS

Title: "Single-cell HSPC analysis"
Author: "J Mellet"
Date: "September 2019"

```
# Activate/install necessary packages
> library(Seurat)
> library(dplyr)
> library(Matrix)
> library(Cairo)

# Import datasets
> sc.HSC <- read.delim("E:/Juanita/Single-cell Experiments/R/SC_HSC_090419/renamed_annotated_combined_genes.counts", header = TRUE, sep = "\t" )
> sc.HSC
> rownames(sc.HSC) <- sc.HSC$id
> sc.HSC <- sc.HSC[ -c(1) ]

# Remove pseudogenes (Columns 58678 to 58735) from the dataframe

Several pseudogenes skewed the data and were therefore removed

> sc.HSC <- sc.HSC[ -c(19308, 41144, 4172, 45380, 31391, 26419, 58677:58735), ]

# Convert dataframe to matrix

> sc.HSC <- as(as.matrix(sc.HSC), "dgCMatrix")

# Create and correct dataframes for every dataset

Additional/incorrect columns were present in the dataframe, which were removed.

Cells that did not pass the QC from the MultiQC report was also removed.

> V1 <- sc.HSC[, -c(108:177)]

> V1


> V4 <- sc.HSC[, -c(1:107, 177)]

> V4
```

Create a Seurat Object for every dataset

Genes need to be expressed in a minimum of 3 cells and cells need to express a minimum of 200 genes to be considered for downstream analysis

```
> sc.V1 <- CreateSeuratObject(V1, min.cells = 3, min.genes = 200, project="scHSC")
> sc.V1
> sc.V1@meta.data[, "dataset"] <- 1
> head(x=sc.V1@meta.data)

> sc.V4 <- CreateSeuratObject(V4, min.cells = 3, min.genes = 200, project="scHSC")
> sc.V4
> sc.V4@meta.data[, "dataset"] <- 4
> head(x=sc.V4@meta.data)
```

Pre-processing of the data

Determine the number of mitochondrial genes present in each Seurat object

Dying/low quality cells often express high percentages mitochondrial genes.

Mitochondrial genes start with MT-, making it easy to identify them in the dataset.

There are 37 human mitochondrial genes.

```
> mito.genesV1 <- grep(pattern = "^MT-", x = rownames(x = V1), value = TRUE)
> mito.genesV1
> percent.mitoV1 <- Matrix::colSums(V1[mito.genesV1, ])/Matrix::colSums(V1)
> percent.mitoV1
> sc.V1 <- AddMetaData(object = sc.V1, metadata = percent.mitoV1, col.name = 'percent.mito')
head(sc.V1@meta.data)

> mito.genesV4 <- grep(pattern = "^MT-", x = rownames(x = V4), value = TRUE)
> mito.genesV4
> percent.mitoV4 <- Matrix::colSums(V4[mito.genesV4, ])/Matrix::colSums(V4)
> percent.mitoV4
> sc.V4 <- AddMetaData(object = sc.V4, metadata = percent.mitoV4, col.name = 'percent.mito')
> head(sc.V4@meta.data)
```

Determine cell cycle scores for each Seurat object

A list of cell cycle genes are imported and separated into S phase genes and G2M phase genes

The gene expression levels of the cell cycle genes were scored in every cell.

Each cell was then categorised as being either in the S phase or G2M phase.

```
> cc.genes <- readLines(con = "E:/Juanita/Single-cell Experiments/R/SC_HSC_090419/regev_lab_cell_cycle_genes.txt")
> cc.genes
> s.genes <- cc.genes[1:43]
> g2m.genes <- cc.genes[44:97]

> sc.V1 <- CellCycleScoring(object = sc.V1, s.genes = s.genes, g2m.genes = g2m.genes, set.ident = FALSE)
> head(x=sc.V1@meta.data)

> sc.V4 <- CellCycleScoring(object = sc.V4, s.genes = s.genes, g2m.genes = g2m.genes, set.ident = FALSE)
> head(x=sc.V4@meta.data)
```

Filter, normalize and scale the data for each Seurat object

The minimum and maximum number of genes included was 200 and 10 000 per cell, respectively.

Cells with high number of mitcohondrial genes (>10%) were removed.

Seurat employs a global-scale normalisation (LogNormalize) that normalises the gene expression measurements for each cell by the total expression, multiplies this by a scale factor (10 000) and log-transform the results.

```
> sc.V1 <- FilterCells(object = sc.V1, subset.names = c("nGene", "percent.mito"),
                       low.thresholds = c(200, -Inf), high.thresholds = c(10000, 0.10))
> sc.V1 <- NormalizeData(object = sc.V1, normalization.method = "LogNormalize", scale.factor = 10000)
> sc.V1 <- ScaleData(object = sc.V1, vars.to.regress = c("percent.mito", "S.Score", "G2M.Score"))
```

```
> sc.V4 <- FilterCells(object = sc.V4, subset.names = c("nGene", "percent.mito"),
                       low.thresholds = c(200, -Inf), high.thresholds = c(10000, 0.10))
> sc.V4 <- NormalizeData(object = sc.V4, normalization.method = "LogNormalize", scale.factor = 10000)
> sc.V4 <- ScaleData(object = sc.V4, vars.to.regress = c("percent.mito", "S.Score", "G2M.Score"))
```

Identify variable genes

This function calculates the average expression and dispersion for each gene, places these genes into bins, and then calculates a z-score for dispersion within each bin.

```
> sc.V1 <- FindVariableGenes(object = sc.V1, mean.function = ExpMean, dispersion.function = LogVMR, x.low.cutoff = 0.0125, x.high.cutoff = 2.5, y.cutoff = -1.5, do.plot = FALSE)
> length(x = sc.V1@var.genes)

> sc.V4 <- FindVariableGenes(object = sc.V4, mean.function = ExpMean, dispersion.function = LogVMR, x.low.cutoff = 0.0125, x.high.cutoff = 2.5, y.cutoff = -1.5, y.high.cutoff = 3, do.plot = FALSE)
> length(x = sc.V4@var.genes)
```

Genes to use for CCA

Only the top 5000 highly variable genes will be used for downstream analysis.

```
> g.V1 <- head(rownames(sc.V1@hvg.info), 5000)
> g.V4 <- head(rownames(sc.V4@hvg.info), 5000)
> genes.use <- unique(c(g.V1, g.V4))
> genes.use <- intersect(genes.use, rownames(sc.V1@scale.data))
> genes.use <- intersect(genes.use, rownames(sc.V4@scale.data))
```

Canonical Correlation Analysis (CCA)

This is the suggested approach by Seurat to circumvent batch effect.

CCA identifies common sources of variation between datasets.

> sc.combined <- RunCCA(sc.V1, sc.V4, genes.use = genes.use, num.cc = 30)

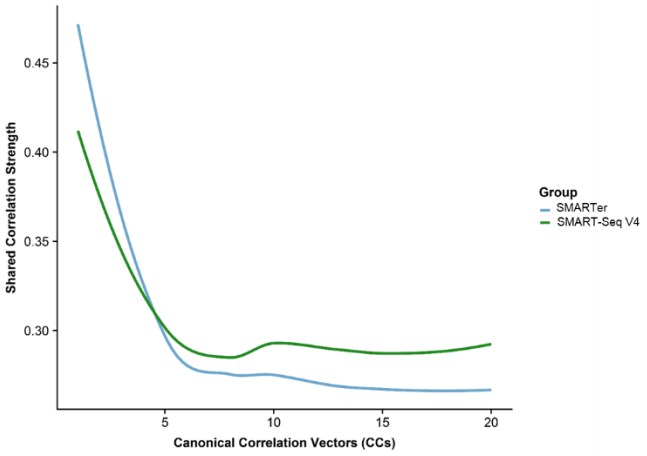
Choose the CCs for CCA

This function will choose the CCs for downstream analysis.

Either the MetageneBicorPlot or the CC Heatmaps can be used to identify the significant CCs.

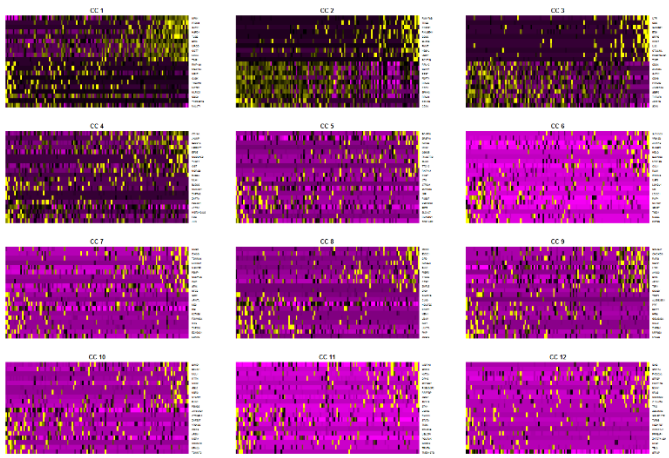
The MetageneBicorPlot examines a measure of correlation strength for each CC, which generally saturates after a reasonable number of CCs.

> X <- MetageneBicorPlot(object=sc.combined, grouping.var = "dataset", dims.eval = 1:20, display.progress = FALSE)



> CC_heatmaps <- DimHeatmap(object = sc.combined, reduction.type = "cca", cells.use = NULL, remove.key = FALSE, num.genes = 20, dim.use = 1:12, do.balanced = TRUE, do.return = TRUE, cexRow = 0.5)

> CC_heatmaps



Alignment

The 'AlignSubspace' function aligns subspaces across a given grouping variable.

This also provides a new dimentional reduction called cca.aligned that is used for downstream analysis such as clustering.

> sc.combined <- AlignSubspace(sc.combined, reduction.type = "cca", grouping.var = "dataset", dims.align = 1:5)

tSNE clustering

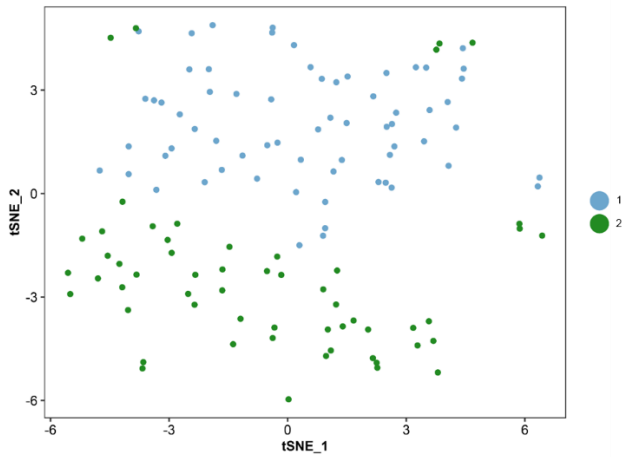
Different resolutions were run to determine the number of distinct and stable clusters.

> sc.combined <- FindClusters(sc.combined, reduction.type = "cca.aligned",
dims.use = 1:5, force.recalc = TRUE, save.SNN = TRUE, resolution = 0.8)

> sc.combined <- RunTSNE(sc.combined, reduction.use = "cca.aligned",
dims.use = 1:5, check_duplicates = FALSE)

> TSNEPlot(sc.combined, do.return = T, pt.size = 1.5, group.by = "dataset")

> TSNEPlot



Identify differentially expressed genes for each cluster

Identifies the differentially expressed genes and displays the top 20 differentially expressed genes in the form of a heatmap.

Determine the presence of well-described HSPC markers in distinct clusters.

> sc.markers <- FindAllMarkers(object = sc.combined, only.pos = TRUE, logfc.threshold = 0.5,

test.use = "wilcox", return.thresh = 0.05, min.pct = 0.1)

> top20 <- sc.markers %>% group_by(cluster) %>% top_n(20, avg_logFC)

top20

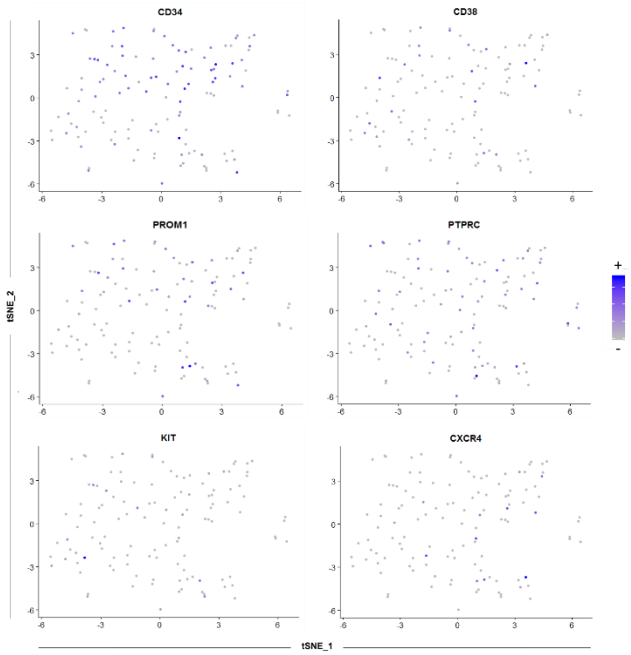
> Heatmap_2C <- **DoHeatmap**(object = sc.combined, genes.use = top20$gene, draw.line = TRUE,

slim.col.label = TRUE, cex.col = 6, group.cex = 10,

cex.row = 6, remove.key = TRUE,

do.plot = TRUE)

> HSPC.markers <- **FeaturePlot**(object = sc.combined, features.plot = c("CD34", "CD38", "PROM1", "PTPRC"), cols.use = c("grey", "blue"), reduction.use = "tsne")



APPENDIX C. DIFFERENTIALLY EXPRESSED GENES

Table C.1. A full list of all the differentially expressed genes in Cluster 1 (resolution 0.7 and 0.8).

GSTP1	LDHA	COX5B	EDF1	ARHGAP15	BEX2
RPL31	CD74	APRT	ATOX1	SH3BGRL3	UXT
RPL15	FDFT1	TRBC2	RPS12	IFITM3	TRAPPC1
CD34	EAF2	LRRC75A-AS1	PSME1	CYTL1	SNRPB2
RPS14	BAALC	RFLNB	ANGPT1	LAPTM4B	ACTB
SPINK2	CD37	CD99	SAP18	EMC4	MGST2
GAPDH	HLA-DPA1	ATP6V1F	MORF4L1	ACTG1	ASAH1
RPS7	MRPL57	CD81	NDUFB11	ATP5MD	RPS5
PRDX1	RPL7	CLIC1	RACK1	EIF3D	RPS2
FAM30A	HSPB1	RPL13	TRAPPC6A	C1QTNF4	ESD
UBB	SKP1	APEX1	LAMP2	CNBP	HSPA8
HOPX	SMIM24	MYL12B	RPUSD3	PFDN1	UBE2L3
FXYD5	ARPC2	PKM	TSG101	MRPL16	IGLL1
NDUFA2	SOD1	NPR3	PPCS	TMEM230	PCNP
SNRPD2	LAT2	TMEM106C	CRHBP	PRMT1	TXN2
RPS6	TALDO1	EGR1	CYC1	NDUFB10	RGS19
GNG11	SELL	CNPY3	MRPS18C	RAB4A	CAPG
BTF3	PSMA4	ATP6V1G1	ATRAID	TSPAN3	CD164
MDK	SSBP1	IL1B	RNPS1	VAMP8	EIF3I
OST4	PSMB7	UBE2L6	TFPI	PUF60	BCAT1
SNX3	EIF5	PHGDH	DYNLT1	GSTM5	TMEM147
MAT2B	TIMM50	NSA2	GHITM	AKR1A1	FAR2
BID	GNAI1	JUP	RALY	ATP6V0E1	NIPAL3
LAMTOR1	CCT3	COA3	STMN1	CAPZA2	TINF2
CMTM7	SERPINB1	MDH2	GAS5	PRDX6	DNAJC8
MYD88	RPS4X	PRR14L	ATP8B4	COMMD10	ADH5
NREP	ERI3	ATP5PB	LAPTM4A	MTCH2	STXBP3
SPINT2	C19orf48	ATP5F1A	VCP	VDAC3	GAPT
HEMGN	VPS72	PCBD1	VPS35	MPI	RUVBL2
MRPS21	HM13	PARVG	ERG28	FAM219B	IL12RB2
CORO1A	SYPL1	MTCH1	MAD2L2	CHD1L	AKR1B1
CARD19	ACAT2	PSMD8	SPTBN1	SLC25A20	IWS1
EXOSC5	OSGEP	RAB7A	MCFD2	DEPP1	OXA1L
NBN	PRDX4	RTCB	KLF10	MGLL	IDI1
IFRD2	WDR54	BSG	LINC01278	MARCKSL1	ZNF711
MDH1	CSRP1	PAN3	CLNS1A	SMAD4	STOML2
ADA	EIF3H	RBM22	NAT9	IGFBP7	FAHD2A
ST13	DFFA	ZNF862	RNF217	TRO	EIF3E
ETHE1	TRAF3IP2	NAE1	FIBP	UFC1	FAM50B
RPS4Y1	PCID2	PHPT1	TIMM17B	PSMC3	H2AFZ
TMEM263	TRIM22	PDCD6IP	POLD2	SLC30A4	ELP6
UNC50	DNTTIP2	USP47	PPT1	SLC30A4	ELP6
UBE2J1	PDE7A	HIST1H2AC	MKLN1	PHF10	SPOP
C2orf68	PHF10	SPOP			

Table C.2. A full list of all the differentially expressed genes in Cluster 2 (resolution 0.7 and 0.8).

AIF1	EXOC6B	FAM178B	NDUFB9	NDUFA8	USP14
TRIB2	FOSB	VNN2	TRIM23	CCNE1	GOLGA4
MXD1	BTN3A2	JHY	LINC00899	ALMS1	SETD5
CD36	LCOR	CSF1R	PHACTR1	SHLD3	FNDC3A
ZFP36L1	S100A9	HERC1	KLF6	LINC00635	PHTF1
ITPR2	TPRKB	STX10	PHF7	MEFV	PACRG
PPAT	NOL9	TET2	ITPR1	ELOVL5	S100A6
RNASEH1	C19orf38	TMPO	STAT5B	SYT14	DNAJB4
MALAT1	NPAT	FPR1	TENT5C	PNMT	TOP2B
S100A8	C6orf89	FRYL	HBB	CLEC4E	HIPK3
FBXO42	WWTR1	SIRPG	CBWD1	ATP2B1	EFCAB8
CNRIP1	ESR1	PRUNE1	ANAPC16	EFCAB8	

Table C.3. A full list of all the differentially expressed genes in Cluster 1 (resolution 0.9 and 1.0).

GSTP1	APRT	EVL	HHEX	SCCPDH	ARF5
RPS14	RPS2P5	LAT2	RPUSD3	PSMC6	RAB32
C1QTNF4	PRDX1	CFAP97	TIMM10	TECR	NAT9
FDFT1	CYBA	LIMD2	PPP4C	ATRAID	POLD2
UBB	APEX1	CD37	MSRB3	MRPL43	TRIM22
FXYD5	SH3BGRL3	SLC9A3R1	CAPG	FAM219B	PUF60
CD74	VAMP8	SELENOH	LAPTM4B	ZNF862	SERP1
BAALC	MZB1	RPS5	SIRT5	DDAH2	ADH5
CD34	UBC	GNPDA1	ZNF672	PLEKHO1	ZSCAN26
RPL13	NAALADL1	RBM22	NUDT14	BCAP31	MYD88
HLA-DPA1	TRAPPC6A	IPO13	NKG7	TMEM106C	CPNE2
MDK	TAGLN2	CCT3	UBE2L6	MPI	IRF5
ACTG1	TALDO1	TCTEX1D1	TMEM147	COA3	TFEC
SPINK2	SPINT2	SRSF7	ADA	IGFBP7	ATP6V0E1
CD99	LDHA	NDUFAF3	RAB7A	TM2D2	KLF10
GAPDH	SELL	HOPX	TMEM230	DNAJB12	LOXL3
IFITM3	PHGDH	CYYR1	FIS1	METTL26	GLA
FAM30A	MAT2B	SLC25A5	VPS35	GRAMD1A	FDPS
ATP6V1F	APRT	R3HCC1	HHEX	SCCPDH	SNAP47
GSTP1	RPS2P5	PYCARD	RPUSD3	PSMC6	RRAGD
RPS14	PRDX1	CFAP97	TIMM10	TECR	DDOST
C1QTNF4	CYBA	LIMD2	PPP4C	ATRAID	HM13
FDFT1	APEX1	CD37	MSRB3	MRPL43	HTRA2
UBB	SH3BGRL3	SLC9A3R1	CAPG	FAM219B	ZNF16
FXYD5	VAMP8	SELENOH	LAPTM4B	ZNF862	CORO1A
CD74	MZB1	RPS5	SIRT5	DDAH2	HIST1H2AC
BAALC	UBC	GNPDA1	ZNF672	PLEKHO1	PSMD8
CD34	NAALADL1	RBM22	NUDT14	BCAP31	PFDN1
RPL13	TRAPPC6A	IPO13	NKG7	TMEM106C	KEAP1
HLA-DPA1	TAGLN2	CCT3	UBE2L6	MPI	ISYNA1
MDK	TALDO1	TCTEX1D1	TMEM147	COA3	PRKCH
ACTG1	SPINT2	SRSF7	ADA	IGFBP7	EIF3I
SPINK2	LDHA	NDUFAF3	RAB7A	TM2D2	WDR54
CD99	SELL	HOPX	TMEM230	DNAJB12	FAM50B
GAPDH	PHGDH	CYYR1	FIS1	METTL26	MRPS26
IFITM3	MAT2B	SLC25A5	VPS35	GRAMD1A	DGKE
FAM30A	EIF3D	R3HCC1	CYB5R3	CTSW	INTS13
ATP6V1F	VPS28	PYCARD	RALY	CNPY3	COPRS
BST2	PSMB7	TSG101	CMTM7	SULT1A1	BSG
CCND3	PSMB6	TAZ	RABIF	ARL16	HGSNAT
IGHM	EDF1	MTCH2	AKR1A1	BUD23	FOXO1
RGS19	S100A4	LAMP2	SERPINB1	DEPP1	DRAM1
RPS2	CLIC1	MDH2	TG	SDHA	CANX
EAF2	IGLL1	ETHE1	NDUFB10	CARD9	USE1
C1GALT1	MPL				

Table C.4. A full list of all the differentially expressed genes in Cluster 2 (resolution 0.9 and 1.0).

S100A8	FRMD8	CD36	ALG6	TMCC1	RAC2
S100A9	MEFV	SLC43A2	SAR1A	CHN2	COPS6
CBWD1	PDE11A	ZNF439	SCN3A	ATP6V0B	MIPEP
CLEC4E	RBM4	DDX59	HSP90B1	RPL36A	SLC15A4
AIF1	TOP2B	SRP72	IL6ST	PPME1	LRRK2
WRNIP1	NUTF2	KCNN4	CSF1R	TPRKB	EXOC6B
MALAT1	ZSCAN31	ALMS1	LRRC1	ATM	ANAPC16
FOSB	NEDD8	SNRPF	SMPD1	SIRPG	RANBP9
SERPINA1	FPR1	VNN2	ZFP36L1	BTN3A2	CCNE1
MT-RNR2	NRIP1	ZNF33A	LINC01618	RNASEH1	BPTF
URB2	IRF8	NPAT	MXD1	RN7SL174P	LINC00635
ZNFX1	UBR1	KPNA6	CCR7	UBR4	CACUL1
BCL2	BCL6	NPC2	NCF2	PCYT1A	KCNQ1OT1
PARP4	CD55	AKNA	POGZ	TOX4	FAM171B
NEAT1	RHEBP2	MAP3K1	MXD1		

Table C.5. A full list of all the differentially expressed genes in Cluster 3 (resolution 0.9 and 1.0).

ACSM3	UPF3B	ABCF1	USP3	NUDT19	TFRC
GATA1	HBD	FECH	AKAP10	ROCK1	TFIP11
CNRIP1	HNRNPA2B1	CLPX	CENPJ	MPC2	PPM1A
POLE2	CHD7	CSE1L	BBS4	GINS2	SEC24A
ING3	PHAX	CASP3	PJA2	RALA	REXO2
MRFAP1	HBB	IARS	USP33	CCNB2	SCFD1
DNAJA4	ANAPC4	RCL1	HELLS	CBLL1	NKTR
NOC3L	BLZF1	SLC25A17	ZNF33B	ZNF605	TBK1
BRI3BP	KDM1A	AP1G2	RBBP4	NAA25	CLTC
PSMD10	NELFCD	GAB3	NCAPG2	NOP58	SMC3
CHP1	ZNF28	SFXN4	COX15	ZNF701	TAF1D
ZBTB16	PLXNC1	MTHFD2	PSMD6-AS2	TRIB2	HNRNPD
SAR1A	TTC1	MBD5	SMARCAD1	PIK3C3	IVNS1ABP
NIPSNAP2	CCT2	NRDC	EZR	MAP4K4	RABGGTB
MGME1	GDAP2	GNB4	CCL20	EIF2S1	PHTF1
THOC7	MBNL2	VPS9D1	DDX41	UFD1	SETD2
MYL4	CCDC88B	PATL1	GUSB	PA2G4	TMEM87A
PRKAR2B	ZNF714	ZNF738	FEN1	PRPF4B	DNTTIP2
ADK	CTNNBL1	KDM5A	CREBZF	DNAJA2	GLUD1
TIMM17A	MRPS9	PLK3	TAF7	NORAD	TMEM14C
IKBIP	CEP162	TMEM181	MLH3	HNRNPK	GGNBP2
AHCTF1	ANKRD12	EPGN	ANXA2	TTC19	COASY
PDS5A	ZNF24	DNAJB4	YIPF4	CAND1	BAZ2B
TTC13	RAF1	DHX9	ETF1	BCL7B	HNRNPA0
RHBDF2	PTBP3	MAT2A	CSNK1A1	MS4A2	SOS1
BCAP29	MUT	GPN1	NCL	GTF3C4	SHLD3
ST7	VPS35L	LARP4B	MTX2	RIF1	SLC30A4
ZNF12					

Table C.6. A full list of all the differentially expressed genes in Cluster 4 (resolution 0.9 and 1.0).

IKZF1	RAD1	ABRACL	CD164	MARCH7	NUP133
FAM111A	OSBPL2	FLI1	CASP4	SLITRK4	GPBP1L1
EIF3E	ZC3H13	TIAL1	CNOT4	TYW5	HOXA9
DCUN1D1	TBL1XR1	ADGRG6	FNBP1	MMRN1	NAP1L1
AP1S2	CLEC9A	PLCB2	PPIP5K2	PSMA3	TIMM23
NEURL1	PER2	IQGAP2	SCN9A	HIST1H2BC	COPS8
ADSS	SELENOF	METTL18	VPS13C	TMEM41B	RAPGEF2
ADPRM	GNG11	GAPT	LRMDA	PRR14L	PCNP
MALSU1	PAN3	RNF217	TPT1-AS1	RBBP7	RBM5
RPS4X	MAPK1IP1L	ZDHHC17	NSA2	TRIM21	RAB22A
MTRR	CDC42SE2	FDXACB1	LINC01138	MAPK6	UTP3
TOR1AIP2	MAVS	ARHGAP15	IRF2BP2	CAPZA2	NFATC2
ANKRD28	STK32C	KIN	DTWD1	CRHBP	CHD1L
TMEM173	PDE1A	TAP1	SLC15A2	MRTFB	ETNK1
CA8	EIF2S3	SENP7	SLC35E3	MPP6	ZNF649
MLLT3	SCOC	ATP6V1H	PSMA3-AS1	TAX1BP1	CAMKK2
TXLNA	ARID1B	PTEN	WDR12	VPS29	TPP2
ANXA4	TFPI	TPST2	PWWP2A	VDAC3	THUMPD1
DDX50	ZBTB14	HEMGN	ING3	IFI16	ATP8B4
DPPA4	GOLPH3L	ANGPT1	HADHB	CLNS1A	CCP110
ZNF84	DAP3	NOG	ABCD4	MEIS1	NBEAL2
SLA	TXNDC9	MTPAP	OGT	LEPROT	P4HTM
HIVEP1	KIAA0087	HMGN4	LAIR1	EBPL	SSBP1
UBE2W	PSMC3				

APPENDIX D. OPTIMISING THE VIRUS PRODUCTION PROTOCOL

In vitro HIV experiments require the use and production of infectious HIV. Viral stocks needed to be propagated to produce increased quantities for infection experiments. The production of R5X4-tropic (CM9) and X4-tropic (SW7) strains was successfully achieved through modifying the Montefiori protocol to accommodate the unique replication dynamics of each isolate. This protocol makes use of activated PBMCs which contain a large number of susceptible host cells for the virus to infect. Upon infection, susceptible PBMCs have the ability to produce and release viral particles into the culture supernatant which can then be harvested for subsequent experiments. Although this is the most physiological model for the production of viral particles, viral mutations occur naturally during each round of replication resulting in an array of viral quasispecies present after production (397). For this study, the p24 ELISA and GHOST cell assay were used to determine the presence of replication competent virus and infectious units per mL, respectively.

D.1. PROPAGATION OF R5- AND R5X4-TROPIC VIRUS

Primary HIV-1C R5-tropic (CM1) and R5X4-tropic (CM9) isolates were used to establish and optimise the HIV production protocol. Activating PBMCs, infecting activated PBMCs with HIV-1C isolates, feeding co-cultures with newly activated PBMCs and harvesting virus-containing supernatants and HIV-infected cells were carried out as described in Section 6.2.1.3. The CM9 isolate was used for initial production performed as described in the Montefiori protocol. During this production experiment, HIV-1C virus was harvested on Day 7 and Day 10. Aliquots of harvested supernatant were tested using p24 ELISA to determine whether viral replication had occurred. Viral p24 protein was clearly present in aliquots from both days, but was noticeably higher on Day 10 (Table D.1). We therefore decided to increase the length of production time during the next production. Viral isolate CM1 was used for the next production and virus was harvested on Day 10 and Day 15. The p24 ELISA results indicated higher p24 protein present in Day 10 supernatant compared to Day 15 (Table D.1). Increased production time did not result in increased virus produced. Since harvesting on Day 10 yielded good overall virus for both CM9 and CM1, virus was harvested on Day 10 for future productions.

Table D.1. p24 ELISA results for initial CM1 and CM9 productions.

Isolate	Sample description	Raw OD	Normalised OD	p24 concentration (pg/mL)
CM9	Blank	0.364	1	-
	Day 7	1.116	3.066	67.7
	Day 10	3.577	9.826	1057.1
CM1	Blank	0.322	1	-
	Day 10	3.824	10.875	1210.7
	Day 15	2.514	6.807	615.3

In an attempt to further increase viral output, the co-culture was split into two separate flasks at the first feed. Each flask was given a fresh/new pool of 10 x 10^6 activated PBMCs every three days, which is double the number of target cells usually available for the virus to infect. Increasing the number of PBMCs on feeding days minimises the possibility of depleting the target cells to be infected. This strategy improved the concentration of virus produced (Table D.2). The produced virus from the two separate flasks of a single production was quite different. Since flasks were kept separate, viral divergence could be different in each flask, resulting in different quasispecies. Therefore, in the adapted protocol, the contents from the two separate flasks were pooled before each feed to homogenise the viral quasispecies and ensure that replication-competent virus is present in each flask after each feed.

Table D.2. p24 ELISA results for modified CM1 and CM9 productions.

Sample description	Raw OD	Normalised OD	p24 concentration (pg/mL)
Blank	0.212	1	-
CM1 Day 10 (Flask 1)	3.596	16.962	2101.6
CM1 Day 10 (Flask 2)	OUT*	-	-
CM9 Day 10 (Flask 1)		18.792	2369.5
CM9 Day 10 (Flask 2)	0.484	2.281	47.4

OUT* = OD value is higher than the higher limit of detection of the spectrophotometer.

Despite our efforts to increase viral production, the number of infectious units remained low and required the addition of large volumes of virus to achieve the desired MOI. Efficiency of infection decreases as the volume surrounding target cells increases (380). Since the initial productions were performed in 30 mL medium, we decided to reduce the volume to 15 mL and later 10 mL. Decreasing the volume pre-concentrated the virus and also increased the viral yield before harvesting, as detected through the GHOST cell assay (Table D.3). Another attempt to concentrate the virus included a 3000 x g spin overnight and collecting the virus at the bottom of the tube, while the virus-depleted supernatant at the top was discarded. As detected by the GHOST cell assay, both these methods improved viral yield and increased infectious units per mL. The pre-concentration step resulted in increased IU/mL and was therefore included in the modified protocol.

Table D.3. GHOST cell assay results for centrifugation and pre-concentration methods.

	Centrifugation		Pre-concentration	
	% GFP	IU/mL	% GFP	IU/mL
Neat	24.34	1.1×10^6	33.81	1.5×10^6
1/4 dilution	2.10	3.8×10^5	4.94	8.9×10^5
1/16 dilution	0.88	6.4×10^5	1.47	1.1×10^6
Average		7.1×10^6		1.2×10^6

Although initial CM1 productions were successful, this viral isolate lost infectivity following several rounds of production. Despite numerous attempts to produce infectious CM1, this was unsuccessful. Various R5-tropic stocks were obtained from the NICD and since inclusion of an R5-tropic isolate is important, small productions were initiated using four different R5-tropic isolates (COT1, Du123F, Du156 and Du422) over three days to determine their production potential. Virus-containing supernatants were harvested from each isolate on Day 3 for p24 ELISA (Table D.4). Isolates that showed increased p24 on Day 3 (COT1 and Du156) were taken through the full ten-day production. Despite the promising p24 results from Day 3 supernatants for COT1 and Du156, little to no infectivity was achieved, as determined by the GHOST cell assay (Figure D.1).

Table D.4. p24 ELISA results for small productions of R5-tropic isolates.

Sample description	Raw OD	Normalised OD	p24 concentration (pg/mL)
Blank	0.199	1	-
COT1	1.867	9.382	288.2
Du123F	0.676	3.397	78.7
Du156	1.767	8.879	270.6
Du422	0.402	2.020	30.8

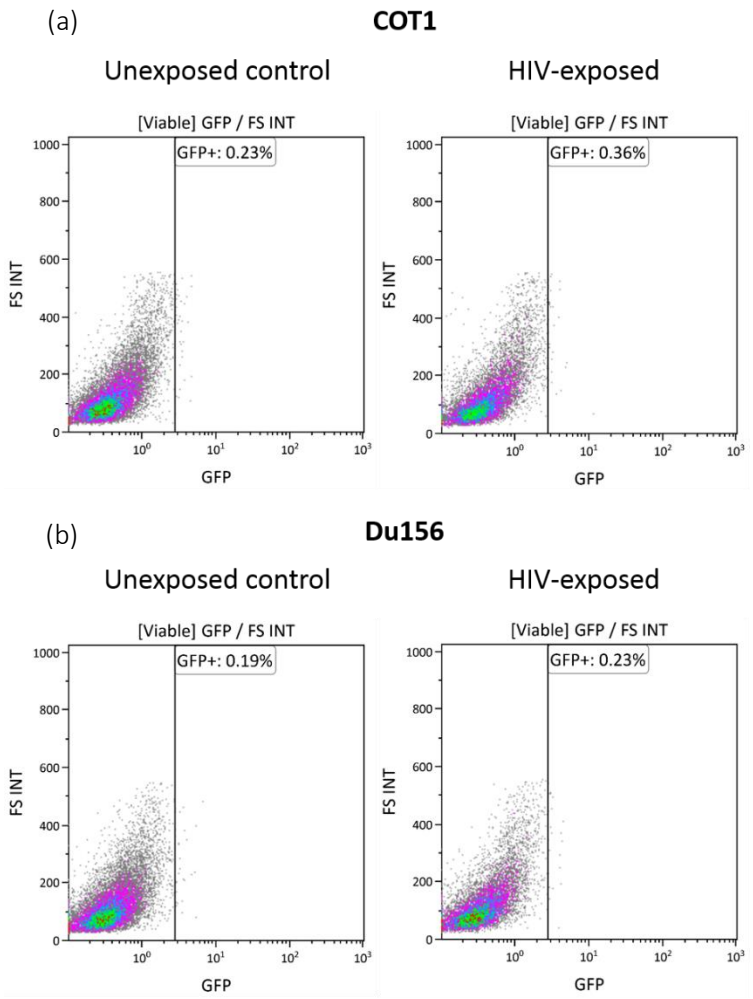


Figure D.1. GHOST assay results for COT1 and DU156 isolates. Density plots (FS vs. GFP) illustrating the GFP expression of HIV-1 isolates, (a) COT1 and (b) Du156, during a Ghost cell assay after full (ten-day) productions. Results for unexposed controls (left) and HIV-exposed samples (neat, undiluted viral stock) (right) were selected for this illustration. Both unexposed controls and HIV-exposed samples were GFP-negative. It was therefore concluded that these HIV-1 isolates were non-infectious.

D.2. PROPAGATION OF X4-TROPIC VIRUS

Propagation of HIV-1C X4-tropic isolate (SW7) resulted in massive cell death. During the initial production, virus-containing supernatant was harvested on Day 4 and Day 10. Although Day 4 supernatant indicated viral production, this was significantly decreased in Day 10 supernatant. This suggested that SW7 is more cytotoxic and therefore results in more rapid target cell depletion when compared to the other isolates. This observation directed us to include more frequent feeds with an increased number of PBMCs in a seven-day production for the production of SW7 isolates. The shortened protocol maintained viral productions at various timepoints, with the highest viral yield on Day 7. These results were also achieved for subsequent productions (Table D.5).

Table D.5. p24 ELISA results for initial and modified X4-tropic productions.

Sample description		Raw OD	Normalised OD	p24 concentration (pg/mL)
Initial production	Blank	0.166	1	-
	Day 4 (1/10 dilution)	0.404	2.434	45.26
	Day 10 (1/10 dilution)	0.24	1.446	10.71
Modified production_1	Blank	0.169	1	-
	Day 3 (1/10 dilution)	0.227	1.343	7.13
	Day 5 (1/10 dilution)	0.297	1.757	21.61
	Day 7 (1/10 dilution)	1.099	6.503	187.54
Modified production_2	Blank	0.199	1	-
	Day 7 (1/5 dilution)	3.855	19.372	637.5

APPENDIX E. INFORMED CONSENT DOCUMENT

UMBILICAL CORD BLOOD COLLECTION

PATIENT INFORMATION LEAFLET AND INFORMED CONSENT

(Each patient must receive, read and understand this document before the start of the study)

STUDY TITLE

The susceptibility of hematopoietic stem and progenitor cells to acute and latent HIV-1 infection

Secondary study title (to improve some of the laboratory techniques):

Novel approaches to conventional quality assurance parameters usually performed before freezing and after thawing cord blood units

INTRODUCTION

This information leaflet is to help you to decide whether you would like to participate in a research project. Before agreeing to take part in the study, please make sure you understand and is comfortable with all procedures involved and please do not hesitate to ask the investigator about anything you might be uncertain of. You should not agree to take part unless you are completely satisfied about all procedures involved.

WHAT IS THE PURPOSE OF THE STUDY?

Researchers at the University of Pretoria would like to investigate the properties of hematopoietic stem cells. Hematopoietic stem cells are very early (young) cells that have the potential to become any of the cells (more mature) in your bone marrow and blood. These hematopoietic stem cells can be found in the afterbirth (placenta), peripheral (circulating) blood, and bone marrow. There are more hematopoietic stem cells in cord blood than in peripheral (circulating) blood. It is important to realize that hematopoietic stem cells can only be collected from peripheral blood in sufficient numbers, after the donor was treated with a drug that will stimulate the production and release of these hematopoietic stem cells in the circulating blood. This is usually done when a donor donate hematopoietic stem cells to be used in treatment of blood-related cancers and other blood diseases. However, the higher numbers of hematopoietic cells in umbilical cord occurs naturally and is not due to any treatment received by the expecting mother. Please note that this consent form serves for the purpose of umbilical cord blood collection only. Hematopoietic stem cells could potentially be used to treat patients with various kinds of diseases such as cancer, blood diseases and HIV. In order to use these cells to treat humans in the future, researchers must first study these cells in a laboratory. The research also involves improving the current techniques in the laboratory in order to detect these very rare cells more efficiently and optimally. As mentioned, these cells are very rare and researchers also look in to ways to store these precious cells optimally as well as how to expand (increase) the cell numbers in the

laboratory. Please be assured that the collection of these cells and the research will not use any unethical procedures.

Hematopoietic stem cells make up only approximately 1% of all the cells present in the umbilical cord blood. After isolation of the hemapoietic stem cells the rest of these cells are discarded, but these cells can be very useful for other research projects. For example, monocytes (one of the cell types present in the blood) can be very useful to researchers that are doing tuberculosis research and lymphocytes, especially T-lymphocytes, are very useful to researchers that are doing HIV research. Blood cells are usually mature. Therefore, the genetic make-up of these cells are often used by researchers as a reference point of how a mature cell looks like and to what extend the genetic make-up of immature (young) cells, such as hematopoietic cells, differ from the mature cells.

By signing this document you give the researchers permission to use the other cells (otherwise regarded as biological waste) in research projects. Please note that researchers are not allowed to do any research with these cells, before they received permission from the relevant ethics committees for their individual research projects.

WHEN AND HOW IS THE BLOOD COLLECTED?

Under normal circumstances, the umbilical cord and placenta is discarded after a baby is born, since they serve no further purpose for either the mother or the baby. Researchers could however use the normally discarded blood for research purposes. After the birth of the baby, the umbilical cord is cut and cord blood will then be collected from the cord. Collecting the blood does not harm the mother or the baby in any way. The collection can only take place at the time of delivery and is performed by your doctor.

WHAT WILL BE EXPECTED OF ME?

Researchers will have access to your medical information, however, all your personal information will be regarded as confidential and anonymous during the study.

By signing this document you give the researchers permission to request your HIV status from your doctor, since it is important for the study that only HIV-negative blood is used. If your HIV status is unknown, please indicate below whether you give the researchers permission to test for HIV. If you do not wish us to know the result of your HIV test, we will not be able to include you in the study. In the case of an HIV positive result, you will be counselled by qualified medical personnel.

In some isolated cases it might however be important for the doctors or researchers involved in the study to convey medical information to medical personnel or appropriate Research Ethics committees. In such a case, you by signing this document, give permission to the investigator to release your medical records to regulatory health authorities or an appropriate Research Ethics committee. If necessary, these medical professionals will discuss the results with your doctor and everyone will act in your best interest.

HAS THE STUDY RECEIVED ETHICAL APPROVAL?

A protocol for the study was submitted to the Research Ethics Committee of the Faculty of Health Sciences at the University of Pretoria and received written approval from the committee. The study is structured in accordance with the Declaration of Helsinki, which deals with the recommendations guiding doctors in biomedical research involving humans.

WHAT ARE MY RIGHTS AS A PARTICIPANT IN THE STUDY?

Your participation in the study is entirely voluntary and you can refuse to participate or withdraw consent at any time without stating any reason. Your withdrawal will not affect your access to medical care or the quality of medical care that you or your baby will receive.

IS THERE ANY RISK OR DISCOMFORT DURING COLLECTION?

No. The cord blood is collected after your baby has been born. The collection is painless, easy, and does not involve you or your baby, and takes about 5 minutes.

IS THERE FINANCIAL GAIN / LOSS FOR MY ACCOUNT IN THE STUDY?

There will be no gain or loss for you or your baby should you participate/withdraw from the study. The research could potentially lead to future profitable treatments. Neither you nor your baby will have access to these profits. There will be no additional financial costs for you to participate in the study.

SOURCE OF ADDITIONAL INFORMATION

If at any time you have any questions about the study, please do not hesitate to contact Ms. Juanita Mellet (079-523-6401), Mr. Carlo Jackson (082-259-4553), Dr. Chrisna Durandt (084-484-5561) or Prof. Michael Pepper (012-319-2190). You are also welcome to contact the Faculty of Health Sciences Ethics Committee at the University of Pretoria if you have any concerns or questions. Their contact details are:

The Research Ethics Office:
Tel: 012 - 354 1330 or 012 - 354 1677
Fax: 012 - 354 1367

For information regarding the collection procedure, you may contact your doctor.

CONFIDENTIALITY

No personal information will be collected from you, the participant. Each participant will be assigned a code and this code will be the only information that the researchers will have. Research reports and articles in scientific journals will not include any information which identifies you or your baby in the study. Only anonymous details which include date of birth, ethnicity and gender will be used in the study.

INFORMED CONSENT

I confirm that the person asking for my consent to take part in the study has told me about the nature, conduct, benefits and risks of the study. I have also received, read and understood the above written information (Patient Information Leaflet and Informed Consent) regarding the study.

I am aware that the results of the study, including anonymous personal details (ethnicity, gender, age, and date of birth) will be processed into a study report.

I hereby give the researchers permission to request my HIV status from my doctor.

I may, at any stage, withdraw my consent and participation in the study. I have had time to ask questions and have no objection to participating in the study.

I hereby give the researchers permission to request my HIV status from my doctor.

YES	NO

Laboratory tests sometime require the isolation of genetic material, also known as DNA and RNA. Genetic material contains information about the cell that only can be revealed if researchers perform specialized tests on the genetic material. These tests are often needed in order to completely understand the characteristics of cells. Genetic information also allows researchers to look into what effect infectious agents, such as *Mycobacterium tuberculosis* (cause tuberculosis in humans) and HIV, might have on these cells. In addition, molecular biology tests (tests that make use of DNA or RNA) are often the most sensitive tests available to detect if cells are infected with bacteria (such as *Mycobacterium tuberculosis)* and/or viruses (such as HIV). Please indicate below if you give the researchers permission to perform these tests. Please note that as soon as the genetic material is extracted from the cells, it would receive an anonymous code to protect your identity.

I hereby give consent for genetic tests and studies to be performed on the source donated.

YES	NO

Patient's name _____ (Please print)
Patient's signature _____ Date _____

I herewith confirm that the above patient has been informed fully about the nature, conduct and risks of the above study.

Investigator's name _____ (Please print)
Investigator's signature _____ Date _____

Doctor’s name _____ (Please print)
Doctor’s signature _____ Date _____

APPENDIX F. AMENDED INFORMED CONSENT DOCUMENT

UMBILICAL CORD AND CORD BLOOD COLLECTION

PATIENT INFORMATION LEAFLET AND INFORMED CONSENT

(Each patient must receive, read and understand this document before the start of the study)

Dear Patient/Participant: _____

This information leaflet provides information on why you have been approached to donate umbilical cord blood/umbilical cord tissue for research purposes. Please note that your decision to donate cord blood/umbilical cord tissue for research purposes is completely voluntary, and we hope the information below will be sufficient to allow you to make an informed decision regarding donation for research purposes.

STUDY TITLE

The samples you donate will be used in one or more of the research studies listed below that are underway at the Institute for Cellular and Molecular Medicine (ICMM), University of Pretoria:

Study titles:

1. The susceptibility of hematopoietic stem and progenitor cells to acute and latent HIV-1 infection;
2. Evaluating the effects of HIV-1 infection on haematopoietic stem cell colony formation
3. Novel approaches to conventional quality assurance parameters usually performed before freezing and after thawing cord blood units
4. The role of Pref-1 in *in vitro* adipogenic differentiation of mesenchymal stromal/stem cells
5. Correlation of hematopoietic stem cell transplantation quality assurance parameters with engraftment success in multiple myeloma patients

INTRODUCTION

This information leaflet is to help you decide whether or not you would like to donate your samples for use in a research project. Before agreeing to take part, please make sure you understand and are comfortable with all procedures involved. Please do not hesitate to ask the investigator about anything about which you may be uncertain. You should not agree to take part unless you are completely satisfied that you fully understand all procedures involved.

WHAT IS THE PURPOSE OF THE STUDY?

Researchers at the University of Pretoria would like to investigate the properties of stem cells. These researchers are interested in two types of stem cells. Hematopoietic stem cells which are very early

(young) cells that have the potential to become any of the more mature cells in your blood. These hematopoietic stem cells can be found in the afterbirth (placenta), peripheral (circulating) blood, and bone marrow. Hematopoietic stem cells could potentially be used to treat patients with various kinds of diseases such as cancer, blood diseases and HIV. In order to use these cells to treat humans in the future, researchers must first study these cells in a laboratory. The research also involves improving the current techniques in the laboratory in order to detect the haematopoietic stem cells more efficiently and optimally. Researchers are looking in to ways to store these precious cells optimally as well as how to expand (increase) cell numbers in the laboratory. Please be assured that the collection of these cells and the research will not use any unethical procedures.

Hematopoietic stem cells make up only approximately 1% of all the cells present in the umbilical cord blood. After isolation of the hematopoietic stem cells the rest of these cells are discarded, but these cells can be very useful for other research projects (see below). For example, monocytes (one of the cell types present in the blood) can be very useful to researchers that are doing tuberculosis research and lymphocytes, especially T-lymphocytes, are very useful to researchers that are doing HIV research. Blood cells are usually mature. Therefore, the genetic make-up of these cells is often used by researchers as a baseline for maturity, and to determine the extent to which the genetic make-up of immature (young) cells, such as hematopoietic stem and progenitor cells, differs from mature cells.

Researchers at the University of Pretoria are also investigating the biological properties of mesenchymal stromal/stem cells. These cells are isolated from the umbilical cord tissue itself. Researchers at the University of Pretoria would like use these mesenchymal stem cells to investigate the process of fat cell maturation (adipogenesis), as obesity is becoming an increasing problem worldwide. The mesenchymal stem cells isolated from the umbilical cord can be used to investigate this maturation process. However, in order to better understand fat cell maturation, researchers must first study the behaviour and growth of the stem cells isolated from the umbilical cord in the laboratory and/or in animal models. The collection of adult stem cells does not make use of any unethical procedures.

Please note that this consent relates to umbilical cord and umbilical cord blood collection only.

By signing this document you give the researchers permission to use cells from the cord and cord blood (otherwise regarded as biological waste) in research projects. Please note that researchers are only allowed to do any research with these cells for which they have obtained permission from the relevant ethics committees for their individual research projects.

WHEN AND HOW IS THE BLOOD AND UMBILICAL CORD COLLECTED?

There is no risk or discomfort involved in the collection of cord or cord blood, for you or for your baby. The birth of the child would still render the normal risks and discomforts associated with your normal birth procedure.

Under normal circumstances, the umbilical cord and placenta are discarded after a baby is born, since they serve no further purpose for either the mother or the baby. Researchers could however use the normally discarded blood and tissue for research purposes. After the birth of the baby, the cord blood will be collected from the cord and a 15 cm long piece of cord will be cut and collected. The remainder

of the cord will be discarded as normal. Collecting the blood does not harm the mother or the baby in any way. The collection can only take place at the time of delivery and is performed by your doctor.

WHAT WILL BE EXPECTED OF ME?

Researchers will have access to your medical information. However, all your personal information will be regarded as confidential and anonymous during the study.

Laboratory tests will be performed on the cord blood and umbilical cord samples if you wish to donate them to the study. Laboratory tests sometimes require the isolation of genetic material. Genetic material is the means by which cells store information and transfer "instructions" from one generation to the next. Genetic material is stored within the cell as molecules called DNA and RNA. The information contained within the genetic material of the cell can only be revealed if researchers perform specialized tests on the genetic material. These tests are often needed in order to completely understand the characteristics of cells. Genetic information also allows researchers to look into what effect infectious agents, such as *Mycobacterium tuberculosis* (cause tuberculosis in humans) and HIV, might have on these cells. In addition, molecular biology tests (tests that make use of DNA or RNA) are often the most sensitive tests available to detect if cells are infected with bacteria (such as *Mycobacterium tuberculosis)* and/or viruses (such as HIV). By signing this document, you give researchers permission to perform these tests. Please note that as soon as the genetic material is extracted from the cells, it will receive an anonymous code to protect your identity.

It is important that we only use HIV-negative samples for our research. We thus perform a test for HIV in our laboratory as one of our quality assurance steps. By signing this document you give the researchers permission to test the sample for HIV. Should there be a positive HIV test result from our laboratory, we will convey this information to your doctor. Please take note that our laboratory is a research laboratory and not an accredited diagnostic facility. Any test that is positive on our testing platform will have to be confirmed by an accredited diagnostic laboratory.

HAS THE STUDY RECEIVED ETHICAL APPROVAL?

Protocol for the studies mentioned herein were submitted to the Research Ethics Committee of the Faculty of Health Sciences at the University of Pretoria and have received written approval from the Committee to proceed. The studies are structured in accordance with the Declaration of Helsinki, which deals with the recommendations guiding doctors in biomedical research involving humans.

WHAT ARE MY RIGHTS AS A PARTICIPANT IN THE STUDY?

Your participation in the study is entirely voluntary and you can refuse to participate or withdraw consent at any time without stating any reason. Your withdrawal will not affect your access to medical care nor the quality of medical care that you or your baby will receive.

IS THERE ANY RISK OR DISCOMFORT DURING COLLECTION?

No. The cord blood is collected after your baby has been born. The collection is painless, easy, and does not involve you or your baby, and takes about 5 minutes.

IS THERE FINANCIAL GAIN / LOSS FOR MY ACCOUNT IN THE STUDY?

There will be no gain or loss for you or your baby should you participate/withdraw from the study. The research could potentially lead to future profitable treatments. Neither you nor your baby will have access to these profits. There will be no financial costs for you to participate in the study.

SOURCE OF ADDITIONAL INFORMATION

If at any time you have any questions about the study, please do not hesitate to contact Ms. Juanita Mellet (079-523-6401), Dr. Chrisna Durandt (084-484-5561) or Prof. Michael Pepper (012-319-2190). You are also welcome to contact the Faculty of Health Sciences Research Ethics Committee at the University of Pretoria if you have any concerns or questions. Their contact details are:

The Research Ethics Office:
Tel: 012 - 356 3085

For information regarding the collection procedure, you may contact your doctor.

CONFIDENTIALITY

Each participant's sample and the material derived therefrom will be assigned a specific code and this code will be used in all research studies. Certain information, including race, ethnicity, gender and medical history, may be important for scientific reasons. This information will be linked to the sample code and not to your identity. Research reports and articles in scientific journals will not include any information that may lead to your personal identification.

In some isolated cases it might however be important for the doctors or researchers involved in the study to convey medical information to medical personnel or appropriate Research Ethics Committees. In such a case, by signing this document, you give permission to the investigators to release your medical records to regulatory health authorities or an appropriate Research Ethics Committee. If necessary, these medical professionals will discuss the results with your doctor and everyone will act in your best interest.

INFORMED CONSENT

I confirm that the person asking my consent to take part in this study has told me about the nature, process, risks, discomforts and benefits of the study.

I have received, read and understood the above written information (Information Leaflet) regarding the study.

I am aware that the results of the study, including personal details, will be anonymously processed into research reports.

I am participating willingly. I have had time to ask questions and have no objection to participate in the study.

I understand that there is no penalty should I wish to discontinue with the study and my withdrawal will not affect my access to medical care or the quality of the medical care I will receive.

I hereby give consent for genetic tests specifically related to the current research projects to be performed on the source donated.

YES	NO

I hereby give the researchers permission to test the sample for HIV for experimental purposes.

YES	NO

In the case of a positive HIV test result I would like my doctor to disclose the result of the test to me.

YES	NO

I hereby consent to donate the source indicated below to the present study (please tick appropriate box):

Umbilical cord ……..	
Umbilical cord blood ……	

I have received a copy of the accompanying information leaflet.

Participant full names (print): _____

Participant signature: _____ Date: _____

Investigator full names (print): _____

Investigator signature: _____ Date: _____

Doctor full names (print): _____

Doctor signature: _____ Date: _____

APPENDIX G. UP ETHICS APPROVAL

The Research Ethics Committee, Faculty Health Sciences, University of Pretoria complies with ICH-GCP guidelines and has US Federal wide Assurance.
- FWA 00002567, Approved dd 22 May 2002 and Expires 20 Oct 2016.
- IRB 0000 2235 IORG0001762 Approved dd 22/04/2014 and Expires 22/04/2017.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Health Sciences Research Ethics Committee

30/10/2014

Approval Certificate
New Application

Ethics Reference No: 410/2014

Title: The susceptibility of hematopoietic stem and progenitor cells to acute and latent HIV-1 infection.

Dear Miss Juanita Mellet

The **New Application** as supported by documents specified in your cover letter for your research received on the 26/09/2014, was approved by the Faculty of Health Sciences Research Ethics Committee on the 29/10/2014.

Please note the following about your ethics approval:
- Ethics Approval is valid for 4 years.
- Please remember to use your protocol number (**410/2014**) on any documents or correspondence with the Research Ethics Committee regarding your research.
- Please note that the Research Ethics Committee may ask further questions, seek additional information, require further modification, or monitor the conduct of your research.

Ethics approval is subject to the following:
- The ethics approval is conditional on the receipt of 6 monthly written Progress Reports, and
- The ethics approval is conditional on the research being conducted as stipulated by the details of all documents submitted to the Committee. In the event that a further need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.

We wish you the best with your research.

Yours sincerely

Dr R Sommers; MBChB; MMed (Int); MPharMed.
Deputy Chairperson of the Faculty of Health Sciences Research Ethics Committee, University of Pretoria

The Faculty of Health Sciences Research Ethics Committee complies with the SA National Act 61 of 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 and 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki, the South African Medical Research Council Guidelines as well as the Guidelines for Ethical Research: Principles Structures and Processes 2004 (Department of Health).

☎ 012 354 1677 🖷 0866516047 deepeka.behari@up.ac.za http://www.healthethics-up.co.za
✉ Private Bag X323, Arcadia, 0007 - 31 Bophelo Road, HW Snyman South Building, Level 2, Room 2.33, Gezina, Pretoria

The Research Ethics Committee, Faculty Health Sciences, University of Pretoria complies with ICH-GCP guidelines and has US Federal wide Assurance.

* FWA 00002567, Approved dd 22 May 2002 and Expires 20 Oct 2016.

* IRB 0000 2235 IORG0001762 Approved dd 22/04/2014 and Expires 22/04/2017

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Health Sciences Research Ethics Committee

30-Jul-2018

Approval Certificate

New Application

Ethics Reference No.: 410/2014

Title: OLD: The susceptibility of hematopoietic stem and progenitor cells to acute and latent HIV-1 infection.
New: Hematopoietic stem and progenitor cell heterogeneity and susceptibility to HIV-1

Dear Miss Juanita Mellet

The **New Application** as supported by documents specified in your cover letter for your research received on the , was approved by the Faculty of Health Sciences Research Ethics Committee on the 28-Mar-2018.

Please note the following about your ethics approval:

- Ethics Approval is valid from to .
- Please remember to use your protocol number (410/2014) on any documents or correspondence with the Research Ethics Committee regarding your research.
- Please note that the Research Ethics Committee may ask further questions, seek additional information, require further modification, or monitor the conduct of your research.

Ethics approval is subject to the following:

- The ethics approval is conditional on the receipt of 6 monthly written Progress Reports, and
- The ethics approval is conditional on the research being conducted as stipulated by the details of all documents submitted to the Committee. In the event that a further need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.

We wish you the best with your research.

Yours sincerely

Dr R Sommers
MBChB MMed(Int) MPharMed
Deputy Chairperson: Faculty of Health Sciences Research Ethics Committee

The Faculty of Health Sciences Research Ethics Committee complies with the SA National Act 61 of 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 and 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki, the South African Medical Research Council Guidelines as well as the Guidelines for Ethical Research: Principles Structures and Processes 2004 (Department of Health).

(012 3563084 - deepika.behari@up.ac.za - http://www.up.ac.za/healthethics
Private Bag X323, Arcadia, 0007 - Tswelopele Building, Level 4, Room 4-60 (opposite the BMW Building), Dr Savage Road, Gezina, Pretoria

APPENDIX H. NETCARE RESEARCH OPERATIONS COMMITTEE APPROVAL

Netcare Management (Pty) Limited

Tel: + 27 (0)11 301 0000
Fax: Corporate +27 (0)11 301 0499
76 Maude Street, Corner West Street, Sandton, South Africa
Private Bag X34, Benmore, 2010, South Africa

RESEARCH OPERATIONS COMMITTEE FINAL APPROVAL OF RESEARCH

Approval number: UNIV-2014-0058

Ms Juanita Mellet

E mail: juanitamellet@yahoo.co.uk

Dear Ms Mellet

RE: THE SUSCEPTIBILITY OF HEMATOPOIETIC STEM AND PROGENITOR CELLS TO ACUTE AND LATENT HIV-1 INFECTIONS

The above-mentioned research was reviewed by the Research Operations Committee's delegated members and it is with pleasure that we inform you that your application to conduct this research at Netcare Femina Hospital, has been approved, subject to the following:

i) Research may now commence with this FINAL APPROVAL from the Netcare Research Operations Committee.

ii) All information regarding Netcare will be treated as legally privileged and confidential.

iii) Netcare's name will not be mentioned without written consent from the Netcare Research Operations Committee.

iv) All legal requirements regarding patient / participant's rights and confidentiality will be complied with.

v) The research will be conducted in compliance with the GUIDELINES FOR GOOD PRACTICE IN THE CONDUCT OF CLINICAL TRIALS IN HUMAN PARTICIPANTS IN SOUTH AFRICA (2006)

vi) Netcare must be furnished with a STATUS REPORT on the progress of the study at least annually on 30th September irrespective of the date of approval from the Netcare Research Operations Committee as well as a FINAL REPORT with reference to intention to publish and probable journals for publication, on completion of the study.

vii) A copy of the research report will be provided to the Netcare Research Operations Committee once it is finally approved by the relevant primary party or tertiary institution, or once complete or if discontinued for any reason whatsoever prior to the expected completion date.

REFERENCES

1. Becker AJ, McCulloch EA, Till JE. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. Nature. 1963;197:452–4.
2. Siminovitch L, McCulloch EA, Till JE. The distribution of colony-forming cells among spleen colonies. J Cell Comp Physiol. 1963;62:327–36.
3. Seita J, Weissman IL. Hematopoietic stem cell: self-renewal versus differentiation. Wiley Interdiscip Rev Syst Biol Med. 2010;2(6):640–53.
4. Daley GQ. Stem cells and the evolving notion of cellular identity. Philos Trans R Soc B Biol Sci. 2015;370(1680):20140376.
5. Zakrzewski W, Dobrzyński M, Szymonowicz M, Rybak Z. Stem cells: past, present, and future. Stem Cell Res Ther. 2019;10(1):68.
6. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663–76.
7. Godini R, Lafta HY, Fallahi H. Epigenetic modifications in the embryonic and induced pluripotent stem cells. Gene Expr Patterns. 2018;29:1–9.
8. Yamanaka S. Induced pluripotent stem cells: past, present, and future. Cell Stem Cell. 2012;10(6):678–84.
9. Spangrude G, Heimfeld S, Weissman I. Purification and characterization of mouse hematopoietic stem cells. Science. 1988;241(4861):58–62.
10. Osawa M, Hanada K, Hamada H, Nakauchi H. Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. Science. 1996;273(5272):242–5.
11. Schofield R. The relationship between the spleen colony-forming cell and the haemopoietic stem cell. Blood Cells. 1978;4(1–2):7–25.
12. Kiel MJ, Morrison SJ. Uncertainty in the niches that maintain haematopoietic stem cells. Nat Rev Immunol. 2008;8(4):290–301.
13. Gómez-López S, Lerner RG, Petritsch C. Asymmetric cell division of stem and progenitor cells during homeostasis and cancer. Cell Mol Life Sci. 2014;71(4):575–97.
14. Szade K, Gulati GS, Chan CKF, Kao KS, Miyanishi M, Marjon KD, *et al.* Where hematopoietic stem cells live: the bone marrow niche. Antioxid Redox Signal. 2018;29(2):191–204.
15. Beerman I, Luis TC, Singbrant S, Lo Celso C, Méndez-Ferrer S. The evolving view of the hematopoietic stem cell niche. Exp Hematol. 2017;50:22–6.
16. Nilsson SK, Johnston HM, Coverdale JA. Spatial localization of transplanted hemopoietic stem cells: inferences for the localization of stem cell niches. Blood. 2001;97(8):2293–9.

17. Zhang J, Niu C, Ye L, Huang H, He X, Tong W-G, *et al*. Identification of the haematopoietic stem cell niche and control of the niche size. Nature. 2003;425(6960):836–41.
18. Calvi LM, Adams GB, Weibrecht KW, Weber JM, Olson DP, Knight MC*, et al.* Osteoblastic cells regulate the haematopoietic stem cell niche. Nature. 2003;425(6960):841–6.
19. Ma YD, Park C, Zhao H, Oduro KA, Tu X, Long F, *et al.* Defects in osteoblast function but no changes in long-term repopulating potential of hematopoietic stem cells in a mouse chronic inflammatory arthritis model. Blood. 2009;114(20):4402–10.
20. Lymperi S, Horwood N, Marley S, Gordon MY, Cope AP, Dazzi F. Strontium can increase some osteoblasts without increasing hematopoietic stem cells. Blood. 2007;111(3):1173–81.
21. Mendelson A, Frenette PS. Hematopoietic stem cell niche maintenance during homeostasis and regeneration. Nat Med. 2014;20(8):833–46.
22. Nombela-Arrieta C, Pivarnik G, Winkel B, Canty KJ, Harley B, Mahoney JE, *et al*. Quantitative imaging of haematopoietic stem and progenitor cell localization and hypoxic status in the bone marrow microenvironment. Nat Cell Biol. 2013;15(5):533–43.
23. Sacchetti B, Funari A, Michienzi S, Di Cesare S, Piersanti S, Saggio I, *et al.* Self-renewing osteoprogenitors in bone marrow sinusoids can organize a hematopoietic microenvironment. Cell. 2007;131(2):324–36.
24. Calvi LM, Bromberg O, Rhee Y, Weber JM, Smith JNP, Basil MJ, *et al.* Osteoblastic expansion induced by parathyroid hormone receptor signaling in murine osteocytes is not sufficient to increase hematopoietic stem cells. Blood. 2012;119(11):2489–99.
25. Nakamura Y, Arai F, Iwasaki H, Hosokawa K, Kobayashi I, Gomei Y, *et al.* Isolation and characterization of endosteal niche cell populations that regulate hematopoietic stem cells. Blood. 2010;116(9):1422–32.
26. Greenbaum A, Hsu Y-MS, Day RB, Schuettpelz LG, Christopher MJ, Borgerding JN, *et al.* CXCL12 in early mesenchymal progenitors is required for haematopoietic stem-cell maintenance. Nature. 2013;495(7440):227–30.
27. Ding L, Morrison SJ. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. Nature. 2013;495(7440):231–5.
28. Gomariz A, Helbling PM, Isringhausen S, Suessbier U, Becker A, Boss A, *et al.* Quantitative spatial analysis of haematopoiesis-regulating stromal cells in the bone marrow microenvironment by 3D microscopy. Nat Commun. 2018;9(1):2532.
29. Westerterp M, Gourion-Arsiquaud S, Murphy AJ, Shih A, Cremers S, Levine RL, *et al.* Regulation of hematopoietic stem and progenitor cell mobilization by cholesterol efflux pathways. Cell Stem Cell. 2012;11(2):195–206.

30. Baryawno N, Przybylski D, Kowalczyk MS, Kfoury Y, Severe N, Gustafsson K, *et al.* A cellular taxonomy of the bone marrow stroma in homeostasis and leukemia. Cell. 2019;177(7):1915-1932.e16.
31. Kondo M. Lymphoid and myeloid lineage commitment in multipotent hematopoietic progenitors. Immunol Rev. 2010;238(1):37–46.
32. Yamamoto R, Morita Y, Ooehara J, Hamanaka S, Onodera M, Rudolph KL, *et al.* Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. Cell. 2013;154(5):1112–26.
33. Sanjuan-Pla A, Macaulay IC, Jensen CT, Woll PS, Luis TC, Mead A, *et al.* Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature. 2013;502(7470):232–6.
34. Scala S, Aiuti A. In vivo dynamics of human hematopoietic stem cells: novel concepts and future directions. Blood Adv. 2019;3(12):1916–24.
35. Velten L, Haas SF, Raffel S, Blaszkiewicz S, Islam S, Hennig BP, *et al.* Human haematopoietic stem cell lineage commitment is a continuous process. Nat Cell Biol. 2017;19(4):271–81.
36. Psaila B, Barkas N, Iskander D, Roy A, Anderson S, Ashley N, *et al.* Single-cell profiling of human megakaryocyte-erythroid progenitors identifies distinct megakaryocyte and erythroid differentiation pathways. Genome Biol. 2016;17(1):83.
37. Görgens A, Radtke S, Möllmann M, Cross M, Dürig J, Horn PA, *et al.* Revision of the human hematopoietic tree: granulocyte subtypes derive from distinct hematopoietic lineages. Cell Rep. 2013;3(5):1539–52.
38. Civin CI, Strauss LC, Brovall C, Fackler MJ, Schwartz JF, Shaper JH. Antigenic analysis of hematopoiesis. III. A hematopoietic progenitor cell surface antigen defined by a monoclonal antibody raised against KG-1a cells. J Immunol. 1984;133(1):157–65.
39. AbuSamra DB, Aleisa FA, Al-Amoodi AS, Jalal Ahmed HM, Chin CJ, Abuelela AF, *et al.* Not just a marker: CD34 on human hematopoietic stem/progenitor cells dominates vascular selectin binding along with CD44. Blood Adv. 2017;1(27):2799–816.
40. Abe T, Matsuoka Y, Nagao Y, Sonoda Y, Hanazono Y. CD34-negative hematopoietic stem cells show distinct expression profiles of homing molecules that limit engraftment in mice and sheep. Int J Hematol. 2017;106(5):631–7.
41. Anjos-Afonso F, Currie E, Palmer HG, Foster KE, Taussig DC, Bonnet D. CD34− cells at the apex of the human hematopoietic stem cell hierarchy have distinctive cellular and molecular signatures. Cell Stem Cell. 2013;13(2):161–74.

42. Manfredini R, Zini R, Salati S, Siena M, Tenedini E, Tagliafico E, *et al.* The kinetic status of hematopoietic stem cell subpopulations underlies a differential expression of genes involved in self-renewal, commitment, and engraftment. Stem Cells. 2005;23(4):496–506.
43. Hemmoranta H, Hautaniemi S, Niemi J, Nicorici D, Laine J, Yli-Harja O, *et al*. Transcriptional profiling reflects shared and unique characters for CD34+ and CD133+ cells. Stem Cells Dev. 2006;15(6):839–51.
44. Takahashi M, Matsuoka Y, Sumide K, Nakatsuka R, Fujioka T, Kohno H, *et al.* CD133 is a positive marker for a distinct class of primitive human cord blood-derived CD34-negative hematopoietic stem cells. Leukemia. 2014;28(6):1308–15.
45. Notta F, Doulatov S, Laurenti E, Poeppl A, Jurisica I, Dick JE. Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. Science. 2011;333(6039):218–21.
46. Doulatov S, Notta F, Eppert K, Nguyen LT, Ohashi PS, Dick JE. Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. Nat Immunol. 2010;11(7):585–93.
47. Chávez-González A, Dorantes-Acosta E, Moreno-Lorenzana D, Alvarado-Moreno A, Arriaga-Pizano L, Mayani H. Expression of CD90, CD96, CD117, and CD123 on different hematopoietic cell populations from pediatric patients with acute myeloid leukemia. Arch Med Res. 2014;45(4):343–50.
48. Majeti R, Park CY, Weissman IL. Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. Cell Stem Cell. 2007;1(6):635–45.
49. Fortunel NO, Otu HH, Ng HH, Chen J, Mu X, Chevassut T, *et al.* Comment on " 'Stemness': transcriptional profiling of embryonic and adult stem cells" and "a stem cell molecular signature". Science. 2003;302(5644):393; author reply 393.
50. Czechowicz A, Palchaudhuri R, Scheck A, Hu Y, Hoggatt J, Saez B, *et al.* Selective hematopoietic stem cell ablation using CD117-antibody-drug-conjugates enables safe and effective transplantation with immunity preservation. Nat Commun. 2019;10(1):617.
51. Ratajczak MZ, Adamiak M. Membrane lipid rafts, master regulators of hematopoietic stem cell retention in bone marrow and their trafficking. Leukemia. 2015;29(7):1452–7.
52. Kollet O, Spiegel A, Peled A, Petit I, Byk T, Hershkoviz R, *et al.* Rapid and efficient homing of human CD34(+)CD38(-/low)CXCR4(+) stem and progenitor cells to the bone marrow and spleen of NOD/SCID and NOD/SCID/B2m(null) mice. Blood. 2001;97(10):3283–91.

53. Plett PA, Frankovitz SM, Wolber FM, Abonour R, Orschell-Traycoff CM. Treatment of circulating CD34(+) cells with SDF-1alpha or anti-CXCR4 antibody enhances migration and NOD/SCID repopulating potential. Exp Hematol. 2002;30(9):1061–9.
54. Naumann U, Cameroni E, Pruenster M, Mahabaleshwar H, Raz E, Zerwes HG, *et al.* CXCR7 functions as a scavenger for CXCL12 and CXCL11. Pockley G, editor. PLoS One. 2010;5(2):e9175.
55. Dar A, Kollet O, Lapidot T. Mutual, reciprocal SDF-1/CXCR4 interactions between hematopoietic and bone marrow stromal cells regulate human stem cell migration and development in NOD/SCID chimeric mice. Exp Hematol. 2006;34(8):967–75.
56. Karpova D, Ritchey JK, Holt MS, Abou-Ezzi G, Monlish D, Batoon L, *et al.* Continuous blockade of CXCR4 results in dramatic mobilization and expansion of hematopoietic stem and progenitor cells. Blood. 2017;129(21):2939–49.
57. Ye H, Wang X, Li Z, Zhou F, Li X, Ni Y, *et al.* Clonal analysis reveals remarkable functional heterogeneity during hematopoietic stem cell emergence. Cell Res. 2017;27(8):1065–8.
58. Gatti RA, Meuwissen HJ, Allen HD, Hong R, Good RA. Immunological reconstitution of sex-linked lymphopenic immunological deficiency. Lancet. 1968;292(7583):1366–9.
59. Dausset J. Iso-Leuko-Antibodies. Vox Sang. 1958;3(1):40–1.
60. Ahmed S, Kanakry JA, Ahn KW, Litovich C, Abdel-Azim H, Aljurf M, *et al.* Lower graft-versus-host disease and relapse risk in post-transplant cyclophosphamide–based haploidentical versus matched sibling donor reduced-intensity conditioning transplant for Hodgkin lymphoma. Biol Blood Marrow Transplant. 2019;
61. Lemos NE, Farias MG, Kubaski F, Scotti L, Onsten TGH, Brondani L de A, *et al.* Quantification of peripheral blood CD34+ cells prior to stem cell harvesting by leukapheresis: a single center experience. Hematol Transfus Cell Ther. 2018;40(3):213–8.
62. Tucci F, Frittoli M, Barzaghi F, Calbi V, Migliavacca M, Ferrua F, *et al.* Bone marrow harvesting from paediatric patients undergoing haematopoietic stem cell gene therapy. Bone Marrow Transplant. 2019;
63. Burman J, Tolf A, Hägglund H, Askmark H. Autologous haematopoietic stem cell transplantation for neurological diseases. J Neurol Neurosurg Psychiatry. 2018;89(2):147–55.
64. Swart JF, Delemarre EM, van Wijk F, Boelens J-J, Kuball J, van Laar JM, *et al.* Haematopoietic stem cell transplantation for autoimmune diseases. Nat Rev Rheumatol. 2017;13(4):244–56.
65. Hacein-Bey-Abina S, Garrigue A, Wang GP, Soulier J, Lim A, Morillon E, *et al.* Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J Clin Invest. 2008;118(9):3132–42.

66. Booth C, Romano R, Roncarolo MG, Thrasher AJ. Gene therapy for primary immunodeficiency. Hum Mol Genet. 2019;
67. DiGiusto DL, Cannon PM, Holmes MC, Li L, Rao A, Wang J, *et al.* Preclinical development and qualification of ZFN-mediated CCR5 disruption in human hematopoietic stem/progenitor cells. Mol Ther - Methods Clin Dev. 2016;3:16067.
68. Caillat-Zucman S, Le Deist F, Haddad E, Gannagé M, Dal Cortivo L, Jabado N, *et al.* Impact of HLA matching on outcome of hematopoietic stem cell transplantation in children with inherited diseases: a single-center comparative analysis of genoidentical, haploidentical or unrelated donors. Bone Marrow Transplant. 2004;33(11):1089–95.
69. Rocha V, Labopin M, Sanz G, Arcese W, Schwerdtfeger R, Bosi A, *et al.* Transplants of umbilical-cord blood or bone marrow from unrelated donors in adults with acute leukemia. N Engl J Med. 2004;351(22):2276–85.
70. Chalmers IM, Janossy G, Contreras M, Navarrete C. Intracellular cytokine profile of cord and adult blood lymphocytes. Blood. 1998;92(1):11–8.
71. Gluckman E, Broxmeyer HA, Auerbach AD, Friedman HS, Douglas GW, Devergie A, *et al.* Hematopoietic reconstitution in a patient with Fanconi's anemia by means of umbilical-cord blood from an HLA-identical sibling. N Engl J Med. 1989;321(17):1174–8.
72. Barker JN, Weisdorf DJ, DeFor TE, Blazar BR, McGlave PB, Miller JS, *et al.* Transplantation of 2 partially HLA-matched umbilical cord blood units to enhance engraftment in adults with hematologic malignancy. Blood. 2005;105(3):1343–7.
73. Scaradavou A, Brunstein CG, Eapen M, Le-Rademacher J, Barker JN, Chao N, *et al.* Double unit grafts successfully extend the application of umbilical cord blood transplantation in adults with acute leukemia. Blood. 2013;121(5):752–8.
74. Staba Kelly S, Parmar S, De Lima M, Robinson S, Shpall E. Overcoming the barriers to umbilical cord blood transplantation. Cytotherapy. 2010;12(2):121–30.
75. Brunstein CG, Barker JN, Weisdorf DJ, DeFor TE, Miller JS, Blazar BR, *et al.* Umbilical cord blood transplantation after nonmyeloablative conditioning: impact on transplantation outcomes in 110 adults with hematologic disease. Blood. 2007;110(8):3064–70.
76. Rocha V, Gluckman E, Eurocord-Netcord registry and European Blood and Marrow Transplant group. Improving outcomes of cord blood transplantation: HLA matching, cell dose and other graft- and transplantation-related factors. Br J Haematol. 2009;147(2):262–74.
77. Horwitz ME, Chao NJ, Rizzieri DA, Long GD, Sullivan KM, Gasparetto C, *et al*. Umbilical cord blood expansion with nicotinamide provides long-term multilineage engraftment. J Clin Invest. 2014;124(7):3121–8.

78. Wagner JE, Brunstein CG, Boitano AE, DeFor TE, McKenna D, Sumstad D, *et al.* Phase I/II trial of StemRegenin-1 expanded umbilical cord blood hematopoietic stem cells supports testing as a stand-alone graft. Cell Stem Cell. 2016;18(1):144–55.
79. Anand S, Thomas S, Hyslop T, Adcock J, Corbet K, Gasparetto C, *et al.* Transplantation of ex vivo expanded umbilical cord blood (NiCord) decreases early infection and hospitalization. Biol Blood Marrow Transplant. 2017;23(7):1151–7.
80. Kittler EL, McGrath H, Temeles D, Crittenden RB, Kister VK, Quesenberry PJ. Biologic significance of constitutive and subliminal growth factor production by bone marrow stroma. Blood. 1992;79(12):3168–78.
81. Lisovsky M, Braun SE, Ge Y, Takahira H, Lu L, Savchenko VG, *et al.* Flt3-ligand production by human bone marrow stromal cells. Leukemia. 1996;10(6):1012–8.
82. Guerriero A, Worford L, Holland HK, Guo GR, Sheehan K, Waller EK. Thrombopoietin is synthesized by bone marrow stromal cells. Blood. 1997;90(9):3444–55.
83. Ueda T, Tsuji K, Yoshino H, Ebihara Y, Yagasaki H, Hisakawa H, *et al.* Expansion of human NOD / SCID-repopulating cells by stem cell factor , Flk2 / Flt3 ligand , thrombopoietin , IL-6 , and soluble IL-6 receptor. J Clin Invest. 2000;105(7):1013–21.
84. Delaney C, Heimfeld S, Brashem-Stein C, Voorhies H, Manger RL, Bernstein ID. Notch-mediated expansion of human cord blood progenitor cells capable of rapid myeloid reconstitution. Nat Med. 2010;16(2):232–6.
85. Li L, Piloto O, Kim K-T, Ye Z, Nguyen HB, Yu X, *et al.* FLT3/ITD expression increases expansion, survival and entry into cell cycle of human haematopoietic stem/progenitor cells. Br J Haematol. 2007;137(1):64–75.
86. Keller U, Götze K, Duyster J, Schmidt B, Rose-John S, Peschel C. Murine stromal cells producing hyper-interleukin-6 and Flt3 ligand support expansion of early human hematopoietic progenitor cells without need of exogenous growth factors. Leukemia. 2002;16(10):2122–8.
87. Bowie MB, Kent DG, Copley MR, Eaves CJ. Steel factor responsiveness regulates the high self-renewal phenotype of fetal hematopoietic stem cells. Blood. 2007;109(11):5043–8.
88. Wang L, Guan X, Wang H, Shen B, Zhang Y, Ren Z, *et al.* A small-molecule/cytokine combination enhances hematopoietic stem cell proliferation via inhibition of cell differentiation. Stem Cell Res Ther. 2017;8(1):169.
89. Solar GP, Kerr WG, Zeigler FC, Hess D, Donahue C, de Sauvage FJ, *et al.* Role of c-mpl in early hematopoiesis. Blood. 1998;92(1):4–10.

90. Sitnicka E, Lin N, Priestley G V, Fox N, Broudy VC, Wolf NS, *et al.* The effect of thrombopoietin on the proliferation and differentiation of murine hematopoietic stem cells. Blood. 1996;87(12):4998–5005.
91. Yagi M, Ritchie KA, Sitnicka E, Storey C, Roth GJ, Bartelmez S. Sustained ex vivo expansion of hematopoietic stem cells mediated by thrombopoietin. Proc Natl Acad Sci U S A. 1999;96(14):8126–31.
92. Bordeaux-Rego P, Luzo A, Costa FF, Olalla Saad ST, Crosara-Alberto DP. Both interleukin-3 and interleukin-6 are necessary for better ex vivo expansion of CD133+ cells from umbilical cord blood. Stem Cells Dev. 2010;19(3):413–22.
93. Norkin M, Lazarus HM, Wingard JR. Umbilical cord blood graft enhancement strategies: has the time come to move these into the clinic? Bone Marrow Transplant. 2013;48(7):884–9.
94. Shpall EJ, Quinones R, Giller R, Zeng C, Baron AE, Jones RB, *et al.* Transplantation of ex vivo expanded cord blood. Biol Blood Marrow Transplant. 2002;8(7):368–76.
95. Zhang CC, Kaba M, Ge G, Xie K, Tong W, Hug C, *et al.* Angiopoietin-like proteins stimulate ex vivo expansion of hematopoietic stem cells. Nat Med. 2006;12(2):240–5.
96. Zhang CC, Kaba M, Iizuka S, Huynh H, Lodish HF. Angiopoietin-like 5 and IGFBP2 stimulate ex vivo expansion of human cord blood hematopoietic stem cells as assayed by NOD/SCID transplantation. Blood. 2008;111(7):3415–23.
97. Zheng J, Huynh H, Umikawa M, Silvany R, Zhang CC. Angiopoietin-like protein 3 supports the activity of hematopoietic stem cells in the bone marrow niche. Blood. 2011;117(2):470–9.
98. Chou S, Lodish HF. Fetal liver hepatic progenitors are supportive stromal cells for hematopoietic stem cells. Proc Natl Acad Sci. 2010;107(17):7799–804.
99. González-Castillo C, Ortuño-Sahagún D, Guzmán-Brambila C, Pallàs M, Rojas-Mayorquín AE. Pleiotrophin as a central nervous system neuromodulator, evidences from the hippocampus. Front Cell Neurosci. 2014;8:443.
100. Himburg HA, Harris JR, Ito T, Daher P, Russell JL, Quarmyne M, *et al.* Pleiotrophin regulates the retention and self-renewal of hematopoietic stem cells in the bone marrow vascular niche. Cell Rep. 2012;2(4):964–75.
101. Dexter TM, Allen TD, Lajtha LG. Conditions controlling the proliferation of haemopoietic stem cells in vitro. J Cell Physiol. 1977;91(3):335–44.
102. Huang X, Zhu B, Wang X, Xiao R, Wang C. Three-dimensional co-culture of mesenchymal stromal cells and differentiated osteoblasts on human bio-derived bone scaffolds supports active multi-lineage hematopoiesis in vitro: functional implication of the biomimetic HSC niche. Int J Mol Med. 2016;38(4):1141–51.

103. Aqmasheh S, Shamsasanjan K, Akbarzadehlaleh P, Pashoutan Sarvar D, Timari H. Effects of mesenchymal stem cell derivatives on hematopoiesis and hematopoietic stem cells. Adv Pharm Bull. 2017;7(2):165–77.
104. Breems DA, Blokland EA, Siebel KE, Mayen AE, Engels LJ, Ploemacher RE. Stroma-contact prevents loss of hematopoietic stem cell quality during ex vivo expansion of CD34+ mobilized peripheral blood stem cells. Blood. 1998;91(1):111–7.
105. Dumont N, Boyer L, Émond H, Çelebi-Saltik B, Pasha R, Bazin R, *et al*. Medium conditioned with mesenchymal stromal cell–derived osteoblasts improves the expansion and engraftment properties of cord blood progenitors. Exp Hematol. 2014;42(9):741-752.e1.
106. Tepliashin AS, Korzhikova S V, Sharifullina SZ, Chupikova NI, Rostovskaia MS, Savchenkova IP. Characteristics of human mesenchymal stem cells isolated from bone marrow and adipose tissue. Tsitologiia. 2005;47(2):130–5.
107. Mohamed-Ahmed S, Fristad I, Lie SA, Suliman S, Mustafa K, Vindenes H, *et al.* Adipose-derived and bone marrow mesenchymal stem cells: a donor-matched comparison. Stem Cell Res Ther. 2018;9(1):168.
108. De Toni F, Poglio S, Youcef A Ben, Cousin B, Pflumio F, Bourin P, *et al.* Human adipose-derived stromal cells efficiently support hematopoiesis in vitro and in vivo: a key step for therapeutic studies. Stem Cells Dev. 2011;20(12):2127–38.
109. Gimble JM, Katz AJ, Bunnell BA. Adipose-derived stem cells for regenerative medicine. Circ Res. 2007;100(9):1249–60.
110. Arulmozhivarman G, Kräter M, Wobus M, Friedrichs J, Bejestani EP, Müller K, *et al.* Zebrafish in-vivo screening for compounds amplifying hematopoietic stem and progenitor cells: preclinical validation in human CD34+ stem and progenitor cells. Sci Rep. 2017;7(1):12084.
111. Peled T, Shoham H, Aschengrau D, Yackoubov D, Frei G, Rosenheimer G N, *et al*. Nicotinamide, a SIRT1 inhibitor, inhibits differentiation and facilitates expansion of hematopoietic progenitor cells with enhanced bone marrow homing and engraftment. Exp Hematol. 2012;40(4):342-355.e1.
112. Boitano AE, Wang J, Romeo R, Bouchez LC, Parker AE, Sutton SE, *et al.* Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells. Science. 2010;329(5997):1345–8.
113. Thordardottir S, Hangalapura BN, Hutten T, Cossu M, Spanholtz J, Schaap N, *et al.* The aryl hydrocarbon receptor antagonist StemRegenin 1 promotes human plasmacytoid and myeloid dendritic cell development from CD34+ hematopoietic progenitor cells. Stem Cells Dev. 2014;23(9):955–67.

114. Tao L, Togarrati PP, Choi KD, Suknuntha K. StemRegenin 1 selectively promotes expansion of multipotent hematopoietic progenitors derived from human embryonic stem cells. J Stem Cells Regen Med. 2017;13(2):75–9.
115. Fares I, Chagraoui J, Gareau Y, Gingras S, Ruel R, Mayotte N, *et al.* Cord blood expansion. Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. Science. 2014;345(6203):1509–12.
116. Bari S, Zhong Q, Fan X, Poon Z, Lim AST, Lim TH, *et al*. Ex vivo expansion of CD34+ CD90+ CD49f+ hematopoietic stem and progenitor cells from non-enriched umbilical cord blood with azole compounds. Stem Cells Transl Med. 2018;7(5):376–93.
117. Milner LA, Kopan R, Martin DI, Bernstein ID. A human homologue of the Drosophila developmental gene, Notch, is expressed in CD34+ hematopoietic precursors. Blood. 1994;83(8):2057–62.
118. Carlesso N, Aster JC, Sklar J, Scadden DT. Notch1-induced delay of human hematopoietic progenitor cell differentiation is associated with altered cell cycle kinetics. Blood. 1999;93(3):838–48.
119. Kunisato A, Chiba S, Nakagami-Yamaguchi E, Kumano K, Saito T, Masuda S, *et al*. HES-1 preserves purified hematopoietic stem cells ex vivo and accumulates side population cells in vivo. Blood. 2003;101(5):1777–83.
120. Varnum-Finney B, Brashem-Stein C, Bernstein ID. Combined effects of Notch signaling and cytokines induce a multiple log increase in precursors with lymphoid and myeloid reconstituting ability. Blood. 2003;101(5):1784–9.
121. Benveniste P, Serra P, Dervovic D, Herer E, Knowles G, Mohtashami M, *et al*. Notch signals are required for in vitro but not in vivo maintenance of human hematopoietic stem cells and delay the appearance of multipotent progenitors. Blood. 2014;123(8):1167–77.
122. Buske C. Deregulated expression of HOXB4 enhances the primitive growth activity of human hematopoietic cells. Blood. 2002;100(3):862–8.
123. Antonchuk J, Sauvageau G, Humphries RK. HOXB4-induced expansion of adult hematopoietic stem cells ex vivo. Cell. 2002;109(1):39–45.
124. Krosl J, Austin P, Beslu N, Kroon E, Humphries RK, Sauvageau G. In vitro expansion of hematopoietic stem cells by recombinant TAT-HOXB4 protein. Nat Med. 2003;9(11):1428–32.
125. Alonso-Ferrero ME, van Til NP, Bartolovic K, Mata MF, Wagemaker G, Moulding D, *et al*. Enhancement of mouse hematopoietic stem/progenitor cell function via transient gene delivery using integration-deficient lentiviral vectors. Exp Hematol. 2018;57:21–9.

126. Ohta H, Sekulovic S, Bakovic S, Eaves CJ, Pineault N, Gasparetto M, *et al.* Near-maximal expansions of hematopoietic stem cells in culture using NUP98-HOX fusions. Exp Hematol. 2007;35(5):817–30.
127. Pineault N, Abramovich C, Ohta H, Humphries RK. Differential and common leukemogenic potentials of multiple NUP98-Hox fusion proteins alone or with Meis1. Mol Cell Biol. 2004;24(5):1907–17.
128. Agard M, Asakrah S, Morici LA. PGE2 suppression of innate immunity during mucosal bacterial infection. Front Cell Infect Microbiol. 2013;3.
129. Goessling W, Allen RS, Guan X, Jin P, Uchida N, Dovey M, *et al.* Prostaglandin E2 enhances human cord blood stem cell xenotransplants and shows long-term safety in preclinical nonhuman primate transplant models. Cell Stem Cell. 2011;8(4):445–58.
130. Cutler C, Multani P, Robbins D, Kim HT, Le T, Hoggatt J, *et al.* Prostaglandin-modulated umbilical cord blood hematopoietic stem cell transplantation. Blood. 2013;122(17):3074–81.
131. Hütter G, Nowak D, Mossner M, Ganepola S, Müssig A, Allers K, *et al*. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. N Engl J Med. 2009;360(7):692–8.
132. Gupta RK, Abdul-Jawad S, McCoy LE, Mok HP, Peppa D, Salgado M, *et al.* HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation. Nature. 2019;568(7751):244–8.
133. Durack DT. Opportunistic infections and Kaposi's sarcoma in homosexual men. N Engl J Med. 1981;305(24):1465–7.
134. Masur H, Michelis MA, Greene JB, Onorato I, Stouwe RA, Holzman RS, *et al.* An outbreak of community-acquired Pneumocystis carinii pneumonia: initial manifestation of cellular immune dysfunction. N Engl J Med. 1981;305(24):1431–8.
135. UN Joint Programme on HIV/AIDS (UNAIDS). Global Report: UNAIDS report on the global AIDS epidemic: 2019. 2019.
136. Deeks SG, Lewin SR, Havlir D V. The end of AIDS: HIV infection as a chronic disease. Lancet. 2013;382(9903):1525–33.
137. Daw MA, El-Bouzedi A, Ahmed MO, Dau AA. Molecular and epidemiological characterization of HIV-1 subtypes among Libyan patients. BMC Res Notes. 2017;10(1):170.
138. Nyamweya S, Hegedus A, Jaye A, Rowland-Jones S, Flanagan KL, Macallan DC. Comparing HIV-1 and HIV-2 infection: lessons for viral immunopathogenesis. Rev Med Virol. 2013;23(4):221–40.
139. Santoro MM, Perno CF. HIV-1 genetic variability and clinical implications. ISRN Microbiol. 2013;2013:1–20.

140. Alessandri-Gradt E, De Oliveira F, Leoz M, Lemee V, Robertson DL, Feyertag F, *et al*. HIV-1 group P infection. AIDS. 2018;32(10):1317–22.
141. Abecasis AB, Wensing AM, Paraskevis D, Vercauteren J, Theys K, Van de Vijver DA, *et al*. HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. Retrovirology. 2013;10(1):7.
142. Geretti AM. HIV-1 subtypes: epidemiology and significance for HIV management. Curr Opin Infect Dis. 2006;19(1):1–7.
143. Siddappa NB, Dash PK, Mahadevan A, Desai A, Jayasuryan N, Ravi V, *et al.* Identification of unique B/C recombinant strains of HIV-1 in the southern state of Karnataka, India. AIDS. 2005;19(13):1426–9.
144. Soares MA, de Oliveira T, Brindeiro RM, Diaz RS, Sabino EC, Brigido L, *et al.* A specific subtype C of human immunodeficiency virus type 1 circulates in Brazil. AIDS. 2003;17(1):11–21.
145. Bjorndal A, Sonnerborg A, Tscherning C, Albert J, Fenyo EM. Phenotypic characteristics of human immunodeficiency virus type 1 subtype C isolates of Ethiopian AIDS patients. AIDS Res Hum Retroviruses. 1999;15(7):647–53.
146. De Wolf F, Hogervorst E, Goudsmit J, Fenyö EM, Rübsamen-Waigmann H, Holmes H, *et al*. Syncytium-inducing and non-syncytium-inducing capacity of human immunodeficiency virus type 1 subtypes other than B: phenotypic and genotypic characteristics. WHO network for HIV isolation and characterization. AIDS Res Hum Retroviruses. 1994;10(11):1387–400.
147. Choge I, Cilliers T, Walker P, Taylor N, Phoswa M, Meyers T, *et al.* Genotypic and phenotypic characterization of viral isolates from HIV-1 subtype C-infected children with slow and rapid disease progression. AIDS Res Hum Retroviruses. 2006;22(5):458–65.
148. Coetzer M, Cilliers T, Ping LH, Swanstrom R, Morris L. Genetic characteristics of the V3 region associated with CXCR4 usage in HIV-1 subtype C isolates. Virology. 2006;356(1–2):95–105.
149. Tscherning C, Alaeus A, Fredriksson R, Björndal Å, Deng H, Littman DR, *et al.* Differences in chemokine coreceptor usage between genetic subtypes of HIV-1. Virology. 1998;241(2):181–8.
150. Ball SC, Abraha A, Collins KR, Marozsan AJ, Baird H, Quinones-Mateu ME, *et al.* Comparing the ex vivo fitness of ccr5-tropic human immunodeficiency virus type 1 isolates of subtypes B and C. J Virol. 2003;77(2):1021–38.
151. Marozsan AJ, Moore DM, Lobritz MA, Fraundorf E, Abraha A, Reeves JD, *et al.* Differences in the fitness of two diverse wild-type human immunodeficiency virus type 1 isolates are related to the efficiency of cell binding and entry. J Virol. 2005;79(11):7121–34.

152. Baeten JM, Chohan B, Lavreys L, Chohan V, McClelland RS, Certain L, *et al.* HIV-1 subtype D infection is associated with faster disease progression than subtype a in spite of similar plasma HIV-1 loads. J Infect Dis. 2007;195(8):1177–80.

153. Kiwanuka N, Laeyendecker O, Robb M, Kigozi G, Arroyo M, McCutchan F, *et al.* Effect of human immunodeficiency virus type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. J Infect Dis. 2008;197(5):707–13.

154. Vasan A, Renjifo B, Hertzmark E, Chaplin B, Msamanga G, Essex M, *et al.* Different rates of disease progression of HIV type 1 infection in Tanzania based on infecting subtype. Clin Infect Dis. 2006;42(6):843–52.

155. Frost SDW, Liu Y, Pond SLK, Chappey C, Wrin T, Petropoulos CJ, *et al.* Characterization of human immunodeficiency virus type 1 (HIV-1) envelope variation and neutralizing antibody responses during transmission of HIV-1 subtype B. J Virol. 2005;79(10):6523–7.

156. Derdeyn CA, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, Denham SA, *et al*. Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. Science. 2004;303(5666):2019–22.

157. Travers SAA, O'Connell MJ, McCormack GP, McInerney JO. Evidence for heterogeneous selective pressures in the evolution of the env gene in different human immunodeficiency virus type 1 subtypes. J Virol. 2005;79(3):1836–41.

158. Choisy M, Woelk CH, Guegan J-F, Robertson DL. Comparative study of adaptive molecular evolution in different human immunodeficiency virus groups and subtypes. J Virol. 2004;78(4):1962–70.

159. Patel MB, Hoffman NG, Swanstrom R. Subtype-specific conformational differences within the V3 Region of subtype B and subtype C human immunodeficiency virus type 1 Env proteins. J Virol. 2008;82(2):903–16.

160. de Oliveira T, Engelbrecht S, Janse van Rensburg E, Gordon M, Bishop K, zur Megede J, *et al.* Variability at human immunodeficiency virus type 1 subtype C protease cleavage sites: an indication of viral fitness? J Virol. 2003;77(17):9422–30.

161. Korber BT, Allen EE, Farmer AD, Myers GL. Heterogeneity of HIV-1 and HIV-2. AIDS. 1995;9 Suppl A:S5-18.

162. Cilliers T, Nhlapo J, Coetzer M, Orlovic D, Ketas T, Olson WC, *et al*. The CCR5 and CXCR4 coreceptors are both used by human immunodeficiency virus type 1 primary isolates from subtype C. J Virol. 2003;77(7):4449–56.

163. Renjifo B, Gilbert P, Chaplin B, Msamanga G, Mwakagile D, Fawzi W, *et al.* Preferential in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D. AIDS. 2004;18(12): 1629–36.

164. Iordanskiy S, Waltke M, Feng Y, Wood C. Subtype-associated differences in HIV-1 reverse transcription affect the viral replication. Retrovirology. 2010;7:85.

165. Rodenburg CM, Li Y, Trask SA, Chen Y, Decker J, Robertson DL, *et al.* Near full-length clones and reference sequences for subtype C isolates of HIV type 1 from three different continents. AIDS Res Hum Retroviruses. 2001;17(2):161–8.

166. McCormick-Davis C, Dalton SB, Singh DK, Stephens EB. Sequence note: comparison of Vpu sequences from diverse geographical isolates of HIV type 1 identifies the presence of highly variable domains, additional invariant amino acids, and a signature sequence motif common to subtype C isolates. AIDS Res Hum Retroviruses. 2000;16(11):1089–95.

167. Abraha A, Nankya IL, Gibson R, Demers K, Tebit DM, Johnston E, *et al.* CCR5- and CXCR4-tropic subtype C human immunodeficiency virus type 1 isolates have a lower level of pathogenic fitness than other dominant group M subtypes: implications for the epidemic. J Virol. 2009;83(11):5592–605.

168. Arien KK, Abraha A, Quinones-Mateu ME, Kestens L, Vanham G, Arts EJ. The replicative fitness of primary human immunodeficiency virus type 1 (HIV-1) group M, HIV-1 group O, and HIV-2 isolates. J Virol. 2005;79(14):8979–90.

169. De Silva FS, Venturini DS, Wagner E, Shank PR, Sharma S. CD4-independent infection of human B cells with HIV type 1: detection of unintegrated viral DNA. AIDS Res Hum Retroviruses. 2001;17(17):1585–98.

170. Liu Y, Liu H, Kim BO, Gattone VH, Li J, Nath A, *et al.* CD4-independent infection of astrocytes by human immunodeficiency virus type 1: requirement for the human mannose receptor. J Virol. 2004;78(8):4120–33.

171. Zerhouni B, Nelson JAE, Saha K. Isolation of CD4-independent primary human immunodeficiency virus type 1 isolates that are syncytium inducing and acutely cytopathic for CD8+ lymphocytes. J Virol. 2004;78(3):1243–55.

172. Yoshii H, Kamiyama H, Goto K, Oishi K, Katunuma N, Tanaka Y, *et al.* CD4-independent human immunodeficiency virus infection involves participation of endocytosis and cathepsin B. Zheng JC, editor. PLoS One. 2011;6(4):e19352.

173. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, *et al*. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc Natl Acad Sci. 2008;105(21):7552–7.

174. Schuitemaker H, Koot M, Kootstra NA, Dercksen MW, de Goede RE, van Steenwijk RP, *et al*. Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus population. J Virol. 1992;66(3):1354–60.

175. Arif MS, Hunter J, Léda AR, Zukurov JPL, Samer S, Camargo M, *et al.* Pace of coreceptor tropism switch in HIV-1-infected individuals after recent infection. Kirchhoff F, editor. J Virol. 2017;91(19).

176. Chun TW, Stuyver L, Mizell SB, Ehler LA, Mican JAM, Baseler M, *et al*. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. Proc Natl Acad Sci. 1997;94(24):13193–7.

177. Churchill MJ, Deeks SG, Margolis DM, Siliciano RF, Swanstrom R. HIV reservoirs: what, where and how to target them. Nat Rev Microbiol. 2016;14(1):55–60.

178. Han Y, Wind-Rotolo M, Yang HC, Siliciano JD, Siliciano RF. Experimental approaches to the study of HIV-1 latency. Nat Rev Microbiol. 2007;5(2):95–106.

179. Haldar M, Murphy KM. Origin, development, and homeostasis of tissue-resident macrophages. Immunol Rev. 2014;262(1):25–35.

180. Persaud D, Gay H, Ziemniak C, Chen YH, Piatak M, Chun TW, *et al.* Absence of detectable HIV-1 viremia after treatment cessation in an infant. N Engl J Med. 2013;369(19):1828–35.

181. Henrich TJ, Hu Z, Li JZ, Sciaranghella G, Busch MP, Keating SM, *et al.* Long-term reduction in peripheral blood HIV type 1 reservoirs following reduced-intensity conditioning allogeneic stem cell transplantation. J Infect Dis. 2013;207(11):1694–702.

182. Henrich TJ, Hatano H, Bacon O, Hogan LE, Rutishauser R, Hill A, *et al.* HIV-1 persistence following extremely early initiation of antiretroviral therapy (ART) during acute HIV-1 infection: an observational study. Bekker LG, editor. PLOS Med. 2017;14(11):e1002417.

183. Carter CC, Onafuwa-Nuga A, McNamara LA, Riddell J, Bixby D, Savona MR, *et al.* HIV-1 infects multipotent progenitor cells causing cell death and establishing latent cellular reservoirs. Nat Med. 2010;16(4):446–51.

184. Sebastian NT, Zaikos TD, Terry V, Taschuk F, McNamara LA, Onafuwa-Nuga A, *et al.* CD4 is expressed on a heterogeneous subset of hematopoietic progenitors, which persistently harbor CXCR4 and CCR5-tropic HIV proviral genomes in vivo. Krausslich HG, editor. PLOS Pathog. 2017;13(7):e1006509.

185. Durand CM, Ghiaur G, Siliciano JD, Rabi SA, Eisele EE, Salgado M, *et al.* HIV-1 DNA is detected in bone marrow populations containing CD4+ T cells but is not found in purified CD34+ hematopoietic progenitor cells in most patients on antiretroviral therapy. J Infect Dis. 2012;205(6):1014–8.

186. Josefsson L, Eriksson S, Sinclair E, Ho T, Killian M, Epling L, *et al*. Hematopoietic precursor cells isolated from patients on long-term suppressive HIV therapy did not contain HIV-1 DNA. J Infect Dis. 2012;206(1):28–34.

187. Scadden DT, Shen H, Cheng T. Hematopoietic stem cells in HIV disease. J Natl Cancer Inst Monogr. 2001;(28):24–9.

188. Zhang J, Scadden DT, Crumpacker CS. Primitive hematopoietic cells resist HIV-1 infection via p21Waf1/Cip1/Sdi1. J Clin Invest. 2007;117(2):473–81.

189. Nixon CC, Vatakis DN, Reichelderfer SN, Dixit D, Kim SG, Uittenbogaart CH, *et al.* HIV-1 infection of hematopoietic progenitor cells in vivo in humanized mice. Blood. 2013;122(13):2195–204.

190. McNamara LA, Ganesh JA, Collins KL. Latent HIV-1 infection occurs in multiple subsets of hematopoietic progenitor cells and is reversed by NF-κB activation. J Virol. 2012;86(17):9337–50.

191. Pace M, O'Doherty U. Hematopoietic stem cells and HIV infection. J Infect Dis. 2013;207(12):1790–2.

192. Carter CC, McNamara LA, Onafuwa-Nuga A, Shackleton M, Riddell J, Bixby D, *et al.* HIV-1 utilizes the CXCR4 chemokine receptor to infect multipotent hematopoietic stem and progenitor cells. Cell Host Microbe. 2011;9(3):223–34.

193. Opie J. Haematological complications of HIV infection. S Afr Med J. 2012;102(6):465–8.

194. Calis JCJ, Phiri KS, Vet RJWM, de Haan RJ, Munthali F, Kraaijenhagen RJ, *et al*. Erythropoiesis in HIV-infected and uninfected Malawian children with severe anemia. AIDS. 2010;24(18):2883–7.

195. Li G, Zhao J, Cheng L, Jiang Q, Kan S, Qin E, *et al.* HIV-1 infection depletes human CD34+CD38- hematopoietic progenitor cells via pDC-dependent mechanisms. PLoS Pathog. 2017;13(7):e1006505.

196. Redd AD, Avalos A, Essex M. Infection of hematopoietic progenitor cells by HIV-1 subtype C, and its association with anemia in southern Africa. Blood. 2007;110(9):3143–9.

197. Moses A V, Williams S, Heneveld ML, Strussenberg J, Rarick M, Loveless M, *et al.* Human immunodeficiency virus infection of bone marrow endothelium reduces induction of stromal hematopoietic growth factors. Blood. 1996;87(3):919–25.

198. Gill V, Shattock RJ, Freedman AR, Robinson G, Griffin GE, Gordon-Smith EC, *et al*. Macrophages are the major target cell for HIV infection in long-term bone marrow culture and demonstrate dual susceptibility to lymphocytotropic and monocytotropic strains of HIV-1. Br J Haematol. 1996;93(1):30–7.

199. Isgrò A, Aiuti A, Mezzaroma I, Addesso M, Riva E, Giovannetti A, *et al.* Improvement of interleukin 2 production, clonogenic capability and restoration of stromal cell function in human immunodeficiency virus-type-1 patients after highly active antiretroviral therapy. Br J Haematol. 2002;118(3):864–74.

200. Bahner I, Kearns K, Coutinho S, Leonard EH, Kohn DB. Infection of human marrow stroma by human immunodeficiency virus-1 (HIV-1) is both required and sufficient for HIV-1-induced hematopoietic suppression in vitro: demonstration by gene modification of primary human stroma. Blood. 1997;90(5):1787–98.

201. Schwartz GN, Kessler SW, Szabo JM, Burrell LM, Francisš ML. Negative regulators may mediate some of the inhibitory effects of HIV-1 infected stromal cell layers on erythropoiesis and myelopoiesis in human bone marrow long term cultures. J Leukoc Biol. 1995;57(6):948–55.

202. Zauli G, Vitale M, Gibellini D, Capitani S. Inhibition of purified CD34+ hematopoietic progenitor cells by human immunodeficiency virus 1 or gp120 mediated by endogenous transforming growth factor beta 1. J Exp Med. 1996;183(1):99–108.

203. Rameshwar P, Denny TN, Gascón P. Enhanced HIV-1 activity in bone marrow can lead to myelopoietic suppression partially contributed by gag p24. J Immunol. 1996;157(9):4244–50.

204. Kulkosky J, Laptev A, Shetty S, Srinivasan A, BouHamdan M, Prockop DJ, et al. Human immunodeficiency virus type 1 Vpr alters bone marrow cell function. Blood. 1999;93(6):1906–15.

205. Gibellini D, Vitone F, Buzzi M, Schiavone P, De Crignis E, Cicola R, *et al.* HIV-1 negatively affects the survival/maturation of cord blood CD34+ hematopoietic progenitor cells differentiated towards megakaryocytic lineage by HIV-1 gp120/CD4 membrane interaction. J Cell Physiol. 2007;210(2):315–24.

206. Ramirez CM, Sinclair E, Epling L, Lee SA, Jain V, Hsue PY, *et al.* Immunologic profiles distinguish aviremic HIV-infected adults. AIDS. 2016;30(10):1553–62.

207. Vishnu P, Aboulafia DM. Haematological manifestations of human immune deficiency virus infection. Br J Haematol. 2015;171(5):695–709.

208. Enawgaw B, Alem M, Addis Z, Melku M. Determination of hematological and immunological parameters among HIV positive patients taking highly active antiretroviral treatment and treatment naïve in the antiretroviral therapy clinic of Gondar University Hospital, Gondar, Northwest Ethiopia: a com. BMC Hematol. 2014;14(1):8.
209. Koh PS, Keeble WW, Faulkner GR, Lewis P, Moses AV, Bagby GC. Influence of HIV-1 on induced expression of granulopoietic factors. Exp Hematol. 2000;28(7):52.
210. Bryder D, Rossi DJ, Weissman IL. Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. Am J Pathol. 2006;169(2):338–46.
211. Panteleev AV, Vorob'ev IA. Expression of early hematopoietic markers in cord blood and mobilized blood. Tsitologiia. 2012;54(10):774–82.
212. van der Valk J, Mellor D, Brands R, Fischer R, Gruber F, Gstraunthaler G, *et al.* The humane collection of fetal bovine serum and possibilities for serum-free cell and tissue culture. Toxicol In Vitro. 2004;18(1):1–12.
213. Mannello F, Tonti GA. Concise review: no breakthroughs for human mesenchymal and embryonic stem cell culture: conditioned medium, feeder layer, or feeder-free; medium with fetal calf serum, human serum, or enriched plasma; serum-free, serum replacement nonconditioned medium, o. Stem Cells. 2007;25(7):1603–9.
214. Gstraunthaler G, Lindl T, van der Valk J. A plea to reduce or replace fetal bovine serum in cell culture media. Cytotechnology. 2013;65(5):791–3.
215. Usta SN, Scharer CD, Xu J, Frey TK, Nash RJ. Chemically defined serum-free and xeno-free media for multiple cell lineages. Ann Transl Med. 2014;2(10):97.
216. Ieyasu A, Ishida R, Kimura T, Morita M, Wilkinson AC, Sudo K, *et al.* An all-recombinant protein-based culture system specifically identifies hematopoietic stem cell maintenance factors. Stem Cell Reports. 2017;8(3):500–8.
217. Mellor DJ, Diesch TJ, Gunn AJ, Bennet L. The importance of 'awareness' for understanding fetal pain. Brain Res Rev. 2005;49(3):455–71.
218. van der Valk J, Bieback K, Buta C, Cochrane B, Dirks WG, Fu J, *et al.* Fetal bovine serum (FBS): past - present - future. ALTEX. 2018;35(1):99–118.
219. Brunner D, Frank J, Appl H, Schöffl H, Pfaller W, Gstraunthaler G. Serum-free cell culture: the serum-free media interactive online database. ALTEX. 2010;27(1):53–62.
220. Coecke S, Balls M, Bowe G, Davis J, Gstraunthaler G, Hartung T, *et al.* Guidance on good cell culture practice. a report of the second ECVAM task force on good cell culture practice. Altern Lab Anim. 2005;33(3):261–87.

221. Pamies D, Bal-Price A, Simeonov A, Tagle D, Allen D, Gerhold D, *et al.* Good cell culture practice for stem cells and stem-cell-derived models. ALTEX. 2017;34(1):95–132.
222. Giancola R, Bonfini T, Iacone A. Cell therapy: cGMP facilities and manufacturing. Muscles Ligaments Tendons J. 2012;2(3):243–7.
223. Wasnik S, Kantipudi S, Kirkland MA, Pande G. Enhanced ex vivo expansion of human hematopoietic progenitors on native and spin coated acellular matrices prepared from bone marrow stromal cells. Stem Cells Int. 2016;2016:1–13.
224. Spanholtz J, Tordoir M, Eissens D, Preijers F, van der Meer A, Joosten I, *et al.* High log-scale expansion of functional human natural killer cells from umbilical cord blood CD34-positive cells for adoptive cancer immunotherapy. PLoS One. 2010;5(2):e9221.
225. Sangeetha VM, Kale VP, Limaye LS. Expansion of cord blood CD34 cells in presence of zVADfmk and zLLYfmk improved their in vitro functionality and in vivo engraftment in NOD/SCID mouse. PLoS One. 2010;5(8):e12221.
226. Tarunina M, Hernandez D, Kronsteiner-Dobramysl B, Pratt P, Watson T, Hua P, *et al*. A Novel high-throughput screening platform reveals an optimized cytokine formulation for human hematopoietic progenitor cell expansion. Stem Cells Dev. 2016;25(22):1709–20.
227. Jaime-Pérez JC, Monreal-Robles R, Rodríguez-Romo LN, Mancías-Guerra C, Herrera-Garza JL, Gómez-Almaguer D. Evaluation of volume and total nucleated cell count as cord blood selection parameters: a receiver operating characteristic curve modeling approach. Am J Clin Pathol. 2011;136(5):721–6.
228. Philip J, Kushwaha N, Chatterjee T, Mallhi R. Optimizing cord blood collections: assessing the role of maternal and neonatal factors. Asian J Transfus Sci. 2015;9(2):163.
229. Goodell MA, Rosenzweig M, Kim H, Marks DF, DeMaria M, Paradis G, *et al.* Dye efflux studies suggest that hematopoietic stem cells expressing low or undetectable levels of CD34 antigen exist in multiple species. Nat Med. 1997;3(12):1337–45.
230. Jackson CS, Durandt C, Janse van Rensburg I, Praloran V, Brunet de la Grange P, Pepper MS. Targeting the aryl hydrocarbon receptor nuclear translocator complex with DMOG and Stemregenin 1 improves primitive hematopoietic stem cell expansion. Stem Cell Res. 2017;21:124–31.
231. Leugers SL, Allison DW, Liang Y, Swiderski C, Zant G Van, Donahue LM. *Ex vivo* expansion of CD34+ cells in Stemline™ II hematopoietic stem cell expansion medium generates a large population of functional early and late progenitor cells. In: Cell Technology for Cell Products. Dordrecht: Springer Netherlands; p. 339–41.

232. Zonari E, Desantis G, Petrillo C, Boccalatte FE, Lidonnici MR, Kajaste-Rudnitski A, *et al.* Efficient ex vivo engineering and expansion of highly purified human hematopoietic stem and progenitor cell populations for gene therapy. Stem Cell Reports. 2017;8(4):977–90.
233. Wisniewski D, Affer M, Willshire J, Clarkson B. Further phenotypic characterization of the primitive lineage− CD34+CD38−CD90+CD45RA− hematopoietic stem cell/progenitor cell sub-population isolated from cord blood, mobilized peripheral blood and patients with chronic myelogenous leukemia. Blood Cancer J. 2011;1(9):e36–e36.
234. Nakamura Y, Ando K, Chargui J, Kawada H, Sato T, Tsuji T, *et al.* Ex vivo generation of CD34(+) cells from CD34(-) hematopoietic cells. Blood. 1999;94(12):4053–9.
235. Goncalves KA, Hoban MD, Proctor JL, Adams HL, Hyzy SL, Boitano AE, *et al.* Phenotype does not always equal function: HDAC inhibitors and UM171, but not SR1, lead to rapid upregulation of CD90 on non-engrafting CD34+CD90-negative human cells. Biol Blood Marrow Transplant. 2018;24(3):S479.
236. Eaves CJ. Hematopoietic stem cells: concepts, definitions, and the new reality. Blood. 2015;125(17):2605–13.
237. Barker JN. Umbilical cord blood (UCB) transplantation: an alternative to the use of unrelated volunteer donors? Hematology. 2007;2007(1):55–61.
238. Ballen KK, Gluckman E, Broxmeyer HE. Umbilical cord blood transplantation: the first 25 years and beyond. Blood. 2013;122(4):491–8.
239. Oran B, Shpall E. Umbilical cord blood transplantation: a maturing technology. Hematol Am Soc Hematol Educ Progr. 2012;2012:215–22.
240. Sotnezova E V, Andreeva ER, Grigoriev AI, Buravkova LB. *Ex vivo* expansion of hematopoietic stem and progenitor cells from umbilical cord blood. Acta Naturae. 2016;8(3):6–16.
241. Gadina M, Gazaniga N, Vian L, Furumoto Y. Small molecules to the rescue: inhibition of cytokine signaling in immune-mediated diseases. J Autoimmun. 2017;85:20–31.
242. de Lima M, McMannis J, Gee A, Komanduri K, Couriel D, Andersson BS, *et al*. Transplantation of ex vivo expanded cord blood cells using the copper chelator tetraethylenepentamine: a phase I/II clinical trial. Bone Marrow Transplant. 2008;41(9):771–8.
243. Jaroscak J, Goltry K, Smith A, Waters-Pick B, Martin PL, Driscoll TA, *et al*. Augmentation of umbilical cord blood (UCB) transplantation with ex vivo-expanded UCB cells: results of a phase 1 trial using the AastromReplicell System. Blood. 2003;101(12):5061–7.
244. de Lima M, McNiece I, Robinson SN, Munsell M, Eapen M, Horowitz M, *et al*. Cord-blood engraftment with ex vivo mesenchymal-cell coculture. N Engl J Med. 2012;367(24):2305–15.

245. Bernitz JM, Daniel MG, Fstkchyan YS, Moore K. Granulocyte colony-stimulating factor mobilizes dormant hematopoietic stem cells without proliferation in mice. Blood. 2017;129(14):1901–12.
246. Dreger P, Haferlach T, Eckstein V, Jacobs S, Suttorp M, Löffler H, *et al.* G-CSF-mobilized peripheral blood progenitor cells for allogeneic transplantation: safety, kinetics of mobilization, and composition of the graft. Br J Haematol. 1994;87(3):609–13.
247. Ghalaut PS, Sen R, Dixit G. Role of granulocyte colony stimulating factor (G-CSF) in chemotherapy induced neutropenia. J Assoc Physicians India. 2008;56:942–4.
248. Tang HM, Tang HL. Anastasis: recovery from the brink of cell death. R Soc open Sci. 2018;5(9):180442.
249. Radtke S, Görgens A, Kordelas L, Schmidt M, Kimmig KR, Köninger A, *et al.* CD133 allows elaborated discrimination and quantification of haematopoietic progenitor subsets in human haematopoietic stem cell transplants. Br J Haematol. 2015;169(6):868–78.
250. Telford WG, Bradford J, Godfrey W, Robey RW, Bates SE. Side population analysis using a violet-excited cell-permeable DNA binding dye. Stem Cells. 2007;25(4):1029–36.
251. Telford WG. Stem cell side population analysis and sorting using DyeCycle Violet. Curr Protoc Cytom. 2010;Chapter 9:Unit9.30.
252. Pearce DJ, Bonnet D. The combined use of Hoechst efflux ability and aldehyde dehydrogenase activity to identify murine and human hematopoietic stem cells. Exp Hematol. 2007;35(9): 1437–46.
253. Storms RW, Goodell MA, Fisher A, Mulligan RC, Smith C. Hoechst dye efflux reveals a novel CD7(+)CD34(-) lymphoid progenitor in human umbilical cord blood. Blood. 2000;96(6):2125–33.
254. Guo Y, Follo M, Geiger K, Lübbert M, Engelhardt M. Side-population cells from different precursor compartments. J Hematother Stem Cell Res. 2003;12(1):71–82.
255. Ohmizono Y, Sakabe H, Kimura T, Tanimukai S, Matsumura T, Miyazaki H, *et al*. Thrombopoietin augments ex vivo expansion of human cord blood-derived hematopoietic progenitors in combination with stem cell factor and flt3 ligand. Leukemia. 1997;11(4):524–30.
256. Petzer AL, Zandstra PW, Piret JM, Eaves CJ. Differential cytokine effects on primitive (CD34+CD38-) human hematopoietic cells: novel responses to Flt3-ligand and thrombopoietin. J Exp Med. 1996;183(6):2551–8.
257. Bari S, Seah KKH, Poon Z, Cheung AMS, Fan X, Ong S-Y, *et al.* Expansion and homing of umbilical cord blood hematopoietic stem and progenitor cells for clinical transplantation. Biol Blood Marrow Transplant. 2015;21(6):1008–19.

258. Almici C, Carlo-Stella C, Wagner JE, Mangoni L, Garau D, Rizzoli V. Biologic and phenotypic analysis of early hematopoietic progenitor cells in umbilical cord blood. Leukemia. 1997;11(12):2143–9.

259. Ninos JM, Jefferies LC, Cogle CR, Kerr WG. The thrombopoietin receptor, c-Mpl, is a selective surface marker for human hematopoietic stem cells. J Transl Med. 2006;4:9.

260. Hannum C, Culpepper J, Campbell D, McClanahan T, Zurawski S, Kastelein R, *et al*. Ligand for FLT3/FLK2 receptor tyrosine kinase regulates growth of haematopoietic stem cells and is encoded by variant RNAs. Nature. 1994;368(6472):643–8.

261. Broxmeyer HE, Lu L, Cooper S, Ruggieri L, Li ZH, Lyman SD. Flt3 ligand stimulates/costimulates the growth of myeloid stem/progenitor cells. Exp Hematol. 1995;23(10):1121–9.

262. Shah AJ, Smogorzewska EM, Hannum C, Crooks GM. Flt3 ligand induces proliferation of quiescent human bone marrow CD34+CD38- cells and maintains progenitor cells in vitro. Blood. 1996;87(9):3563–70.

263. Mascarenhas MI, Bacon WA, Kapeni C, Fitch SR, Kimber G, Cheng SWP, *et al*. Analysis of Jak2 signaling reveals resistance of mouse embryonic hematopoietic stem cells to myeloproliferative disease mutation. Blood. 2016;127(19):2298–309.

264. Levac K, Karanu F, Bhatia M. Identification of growth factor conditions that reduce ex vivo cord blood progenitor expansion but do not alter human repopulating cell function in vivo. Haematologica. 2005;90(2):166–72.

265. Qiu L, Meagher R, Welhausen S, Heye M, Brown R, Herzig RH. Ex vivo expansion of CD34+ umbilical cord blood cells in a defined serum-free medium (QBSF-60) with early effect cytokines. J Hematother Stem Cell Res. 1999;8(6):609–18.

266. Paek JY, Kim DY. The Effect of IL-3 in Ex vivo expansion of UCB CD34+ cells. Blood. 2005;106(11):4201.

267. Lam AC, Li K, Zhang XB, Li CK, Fok TF, Chang AM, *et al.* Preclinical ex vivo expansion of cord blood hematopoietic stem and progenitor cells: duration of culture; the media, serum supplements, and growth factors used; and engraftment in NOD/SCID mice. Transfusion. 2001;41(12):1567–76.

268. Schuettpelz LG, Borgerding JN, Christopher MJ, Gopalan PK, Romine MP, Herman AC, *et al.* G-CSF regulates hematopoietic stem cell activity, in part, through activation of Toll-like receptor signaling. Leukemia. 2014;28(9):1851–60.

269. Strauss LC, Rowley SD, La Russa VF, Sharkis SJ, Stuart RK, Civin CI. Antigenic analysis of hematopoiesis. V. Characterization of My-10 antigen expression by normal lymphohematopoietic progenitor cells. Exp Hematol. 1986;14(9):878–86.

270. Tajima F, Sato T, Laver JH, Ogawa M. CD34 expression by murine hematopoietic stem cells mobilized by granulocyte colony-stimulating factor. Blood. 2000;96(5):1989–93.
271. Josefsen D, Forfang L, Dyrhaug M, Blystad AK, Stokke T, Smeland EB, *et al.* Side population cells in highly enriched CD34-positive cells from peripheral blood progenitor cells identify an immature subtype of hematopoietic progenitor cells but do not predict time to engraftment in patients treated with high-dose therapy. Eur J Haematol. 2011;87(6):494–502.
272. Go MJ, Takenaka C, Ohgushi H. Forced expression of Sox2 or Nanog in human bone marrow derived mesenchymal stem cells maintains their expansion and differentiation capabilities. Exp Cell Res. 2008;314(5):1147–54.
273. Yonemura Y, Ku H, Hirayama F, Souza LM, Ogawa M. Interleukin 3 or interleukin 1 abrogates the reconstituting ability of hematopoietic stem cells. Proc Natl Acad Sci. 1996;93(9):4040–4.
274. Nitsche A, Junghahn I, Thulke S, Aumann J, Radonić A, Fichtner I, *et al*. Interleukin-3 promotes proliferation and differentiation of human hematopoietic stem cells but reduces their repopulation potential in NOD/SCID mice. Stem Cells. 2003;21(2):236–44.
275. Hirayama F, Clark SC, Ogawa M. Negative regulation of early B lymphopoiesis by interleukin 3 and interleukin 1 alpha. Proc Natl Acad Sci. 1994;91(2):469–73.
276. Rossmanith T, Schröder B, Bug G, Müller P, Klenner T, Knaus R, *et al*. Interleukin 3 improves the ex vivo expansion of primitive human cord blood progenitor cells and maintains the engraftment potential of scid repopulating cells. Stem Cells. 2001;19(4):313–20.
277. Pébusque M, Faÿ C, Lafage M, Sempéré C, Saeland S, Caux C, *et al*. Recombinant human IL-3 and G-CSF act synergistically in stimulating the growth of acute myeloid leukemia cells. Leukemia. 1989;3(3):200–5.
278. Bodine DM, Seidel NE, Orlic D. Bone marrow collected 14 days after in vivo administration of granulocyte colony-stimulating factor and stem cell factor to mice has 10-fold more repopulating ability than untreated bone marrow. Blood. 1996;88(1):89–97.
279. Schuettpelz LG, Link DC. Regulation of hematopoietic stem cell activity by inflammation. Front Immunol. 2013;4:204.
280. Winkler IG, Wiercinska E, Barbier V, Nowlan B, Bonig H, Levesque JP. Mobilization of hematopoietic stem cells with highest self-renewal by G-CSF precedes clonogenic cell mobilization peak. Exp Hematol. 2016;44(4):303-314.e1.
281. Li S, Zou D, Li C, Meng H, Sui W, Feng S, *et al.* Targeting stem cell niche can protect hematopoietic stem cells from chemotherapy and G-CSF treatment. Stem Cell Res Ther. 2015;6:175.

282. Jie Z, Zhang Y, Wang C, Shen B, Guan X, Ren Z, *et al*. Large-scale ex vivo generation of human neutrophils from cord blood CD34+ cells. Subbiah S kumar, editor. PLoS One. 2017;12(7):e0180832.

283. Torres-Acosta MA, Harrison RP, Csaszar E, Rito-Palomares M, Brunck MEG. Ex vivo manufactured neutrophils for treatment of neutropenia-A process economic evaluation. Front Med. 2019;6:21.

284. Wuchter P, Ran D, Bruckner T, Schmitt T, Witzens-Harig M, Neben K*, et al*. Poor mobilization of hematopoietic stem cells-definitions, incidence, risk factors, and impact on outcome of autologous transplantation. Biol Blood Marrow Transplant. 2010;16(4):490–9.

285. Nguyen NT, Nakahama T, Nguyen CH, Tran TT, Le VS, Chu HH, *et al.* Aryl hydrocarbon receptor antagonism and its role in rheumatoid arthritis. J Exp Pharmacol. 2015;7:29–35.

286. Laiosa MD, Tate ER, Ahrenhoerster LS, Chen Y, Wang D. Effects of developmental activation of the aryl hydrocarbon receptor by 2,3,7,8-tetrachlorodibenzo-p-dioxin on long-term self-renewal of murine hematopoietic stem cells. Environ Health Perspect. 2016;124(7):957–65.

287. Smith BW, Rozelle SS, Leung A, Ubellacker J, Parks A, Nah SK, *et al*. The aryl hydrocarbon receptor directs hematopoietic progenitor cell expansion and differentiation. Blood. 2013;122(3): 376–85.

288. Bennett JA, Singh KP, Welle SL, Boule LA, Lawrence BP, Gasiewicz TA. Conditional deletion of Ahr alters gene expression profiles in hematopoietic stem cells. PLoS One. 2018;13(11):e0206407.

289. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25–9.

290. Sandoval-Montes C, Santos-Argumedo L. CD38 is expressed selectively during the activation of a subset of mature T cells with reduced proliferation but improved potential to produce cytokines. J Leukoc Biol. 2005;77(4):513–21.

291. Jackson CS, Durandt C, Janse van Rensburg I, Praloran V, Brunet de la Grange P, Pepper MS. Targeting the aryl hydrocarbon receptor nuclear translocator complex with DMOG and Stemregenin 1 improves primitive hematopoietic stem cell expansion. Stem Cell Res. 2017;21:124–31.

292. Koide R, Kulkeaw K, Tanaka Y, Swain A, Nakanishi Y, Sugiyama D. Aryl hydrocarbon receptor antagonist StemRegenin 1 promotes the expansion of human promyelocytic leukemia cell line, NB4. Anticancer Res. 2016;36(7):3635–43.

293. Dooley DC. Analysis of primitive CD34- and CD34+ hematopoietic cells from adults: gain and loss of CD34 antigen by undifferentiated cells are closely linked to proliferative status in culture. Stem Cells. 2004;22(4):556–69.

294. Dao MA. Reversibility of CD34 expression on human hematopoietic stem cells that retain the capacity for secondary reconstitution. Blood. 2003;101(1):112–8.
295. Teng CL, Yu JT, Cheng SB, Yang Y, Chang KH, Hwang WL. Circulating hematopoietic progenitors and CD34+ cells predicted successful hematopoietic stem cell harvest in myeloma and lymphoma patients: experiences from a single institution. J Blood Med. 2016;5.
296. Gori JL, Butler JM, Kunar B, Poulos MG, Ginsberg M, Nolan DJ, *et al.* Endothelial cells promote expansion of long-term engrafting marrow hematopoietic stem and progenitor cells in primates. Stem Cells Transl Med. 2017;6(3):864–76.
297. Saliba AE, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res. 2014;42(14):8845–60.
298. Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA. The technology and biology of single-cell RNA sequencing. Mol Cell. 2015;58(4):610–20.
299. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, *et al.* Tumour evolution inferred by single-cell sequencing. Nature. 2011;472(7341):90–4.
300. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, *et al*. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods. 2009;6(5):377–82.
301. Saliba AE, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res. 2014;42(14):8845–60.
302. Wang Y, Navin NE. Advances and applications of single-cell sequencing technologies. Mol Cell. 2015;58(4):598–609.
303. Qiu S, Luo S, Evgrafov O, Li R, Schroth GP, Levitt P, *et al.* Single-neuron RNA-Seq: technical feasibility and reproducibility. Front Genet. 2012;3.
304. Lovatt D, Ruble BK, Lee J, Dueck H, Kim TK, Fisher S, *et al.* Transcriptome in vivo analysis (TIVA) of spatially defined single cells in live tissue. Nat Methods. 2014;11(2):190–6.
305. Tang F, Barbacioru C, Bao S, Lee C, Nordman E, Wang X, *et al*. Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. Cell Stem Cell. 2010;6(5):468–78.
306. Xue Z, Huang K, Cai C, Cai L, Jiang C, Feng Y, *et al*. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature. 2013;500(7464):593–7.
307. Li Y, Xu X, Song L, Hou Y, Li Z, Tsang S, *et al.* Single-cell sequencing analysis characterizes common and cell-lineage-specific mutations in a muscle-invasive bladder cancer. Gigascience. 2012;1(1):12.
308. Yu C, Yu J, Yao X, Wu WK, Lu Y, Tang S, *et al.* Discovery of biclonal origin and a novel oncogene SLC12A5 in colon cancer by single-cell sequencing. Cell Res. 2014;24(6):701–12.

309. Sumide K, Matsuoka Y, Kawamura H, Nakatsuka R, Fujioka T, Asano H, *et al.* A revised road map for the commitment of human cord blood CD34-negative hematopoietic stem cells. Nat Commun. 2018;9(1):2202.

310. Pellin D, Loperfido M, Baricordi C, Wolock SL, Montepeloso A, Weinberg OK, *et al.* A comprehensive single cell transcriptional landscape of human hematopoietic progenitors. Nat Commun. 2019;10(1):2395.

311. Zheng S, Papalexi E, Butler A, Stephenson W, Satija R. Molecular transitions in early progenitors during human cord blood hematopoiesis. Mol Syst Biol. 2018;14(3):e8041.

312. Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. Nat Rev Genet. 2015;16(3):133–45.

313. Mcgettigan PA. Transcriptomics in the RNA-seq era. Curr Opin Chem Biol. 2013;17(1):4–11.

314. Slomovic S. Polyadenylation of ribosomal RNA in human cells. Nucleic Acids Res. 2006;34(10):2966–75.

315. Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD. Reverse transcriptase template switching: a SMART™ approach for full-length cDNA library construction. Biotechniques. 2001;30(4): 892–7.

316. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.

317. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357–60.

318. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30.

319. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000 Research. 2015;4:1521.

320. Engström PG, Steijger T, Sipos B, Grant GR, Kahles A, Rätsch G, *et al*. Systematic evaluation of spliced alignment programs for RNA-seq data. Nat Methods. 2013;10(12):1185–91.

321. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14(4):417–9.

322. Haque A, Engel J, Teichmann SA, Lönnberg T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. Genome Med. 2017;9(1):75.

323. AlJanahi AA, Danielsen M, Dunbar CE. An Introduction to the analysis of single-cell RNA-sequencing data. Mol Ther Methods Clin Dev. 2018;10:189–96.

324. Tutar Y. Pseudogenes. Comp Funct Genomics. 2012;2012:1–4.

325. Buettner F, Natarajan KN, Casale FP, Proserpio V, Scialdone A, Theis FJ, *et al.* Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. Nat Biotechnol. 2015;33(2):155–60.

326. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, *et al*. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016;352(6282):189–96.

327. Nyamundanda G, Poudel P, Patil Y, Sadanandam A. A novel statistical method to diagnose, quantify and correct batch effects in genomic studies. Sci Rep. 2017;7(1):10849.

328. Waltman L, van Eck NJ. A smart local moving algorithm for large-scale modularity-based community detection. Eur Phys J B. 2013;86:471.

329. Jordan CT, Yamasaki G, Minamoto D. High-resolution cell cycle analysis of defined phenotypic subsets within primitive human hematopoietic cell populations. Exp Hematol. 1996;24(11):1347–55.

330. Yahata T, Muguruma Y, Yumino S, Sheng Y, Uno T, Matsuzawa H, *et al.* Quiescent human hematopoietic stem cells in the bone marrow niches organize the hierarchical structure of hematopoiesis. Stem Cells. 2008;26(12):3228–36.

331. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36(5):411–20.

332. He X, Gonzalez V, Tsang A, Thompson J, Tsang TC, Harris DT. Differential gene expression profiling of CD34+ CD133+ umbilical cord blood hematopoietic stem progenitor cells. Stem Cells Dev. 2005;14(2):188–98.

333. Lai S, Huang W, Xu Y, Jiang M, Chen H, Cheng C, *et al*. Comparative transcriptomic analysis of hematopoietic system between human and mouse by microwell-seq. Cell Discov. 2018;4(1):34.

334. Amici SA, Young NA, Narvaez-Miranda J, Jablonski KA, Arcos J, Rosas L, et al. CD38 is robustly induced in human macrophages and monocytes in inflammatory conditions. Front Immunol. 2018;9.

335. Hartman WR, Pelleymounter LL, Moon I, Kalari K, Liu M, Wu TY, *et al.* CD38 expression, function, and gene resequencing in a human lymphoblastoid cell line-based model system. Leuk Lymphoma. 2010;51(7):1315–25.

336. Back J, Allman D, Chan S, Kastner P. Visualizing PU.1 activity during hematopoiesis. Exp Hematol. 2005;33(4):395–402.

337. Rekhtman N, Radparvar F, Evans T, Skoultchi AI. Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. Genes Dev. 1999;13(11):1398–411.

338. Duff C, Smith-Miles K, Lopes L, Tian T. Mathematical modelling of stem cell differentiation: the PU.1–GATA-1 interaction. J Math Biol. 2012;64(3):449–68.
339. Min IM, Pietramaggiori G, Kim FS, Passegué E, Stevenson KE, Wagers AJ. The transcription factor egr1 controls both the proliferation and localization of hematopoietic stem cells. Cell Stem Cell. 2008;2(4):380–91.
340. Hittinger M, Czyz ZT, Huesemann Y, Maneck M, Botteron C, Kaeufl S, *et al.* Molecular profiling of single Sca-1+/CD34+,−cells—the putative murine lung stem cells. Singla DK, editor. PLoS One. 2013;8(12):e83917.
341. Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. Cell. 2008;135(2):216–26.
342. Zhou X, Crow AL, Hartiala J, Spindler TJ, Ghazalpour A, Barsky LW, *et al.* The genetic landscape of hematopoietic stem cell frequency in mice. Stem Cell Reports. 2015;5(1):125–38.
343. Lin CC, Hsu YC, Li YH, Kuo YY, Hou HA, Lan KH, *et al.* Higher HOPX expression is associated with distinct clinical and biological features and predicts poor prognosis in de novo acute myeloid leukemia. Haematologica. 2017;102(6):1044–53.
344. Paul F, Arkin Y, Giladi A, Jaitin DA, Kenigsberg E, Keren-Shaul H, *et al.* Transcriptional heterogeneity and lineage commitment in myeloid progenitors. Cell. 2015;163(7):1663–77.
345. Kurotaki D, Sasaki H, Tamura T. Transcriptional control of monocyte and macrophage development. Int Immunol. 2017;29(3):97–107.
346. Tamura A, Hirai H, Yokota A, Kamio N, Sato A, Shoji T, *et al*. C/EBPβ is required for survival of Ly6C − monocytes. Blood. 2017;130(16):1809–18.
347. Zhu YP, Padgett L, Dinh HQ, Marcovecchio P, Blatchley A, Wu R, *et al.* Identification of an early unipotent neutrophil progenitor with pro-tumoral activity in mouse and human bone marrow. Cell Rep. 2018;24(9):2329-2341.e8.
348. Leonard M, Brice M, Engel JD, Papayannopoulou T. Dynamics of GATA transcription factor expression during erythroid differentiation. Blood. 1993;82(4):1071–9.
349. Grass JA, Boyer ME, Pal S, Wu J, Weiss MJ, Bresnick EH. GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. Proc Natl Acad Sci. 2003;100(15):8811–6.
350. Dore LC, Crispino JD. Transcription factor networks in erythroid cell and megakaryocyte development. Blood. 2011;118(2):231–9.
351. Hu YL, Fong S, Ferrell C, Largman C, Shen WF. HOXA9 modulates its oncogenic partner Meis1 to influence normal hematopoiesis. Mol Cell Biol. 2009;29(18):5181–92.

352. Calvanese V, Nguyen AT, Bolan T, Galic Z, Mikkola H. MLLT3 sustains human HSC self-renewal and engraftment. Exp Hematol. 2017;53:S42–3.

353. Tan YT, Ye L, Xie F, Beyer AI, Muench MO, Wang J, *et al*. Respecifying human iPSC-derived blood cells into highly engraftable hematopoietic stem and progenitor cells with a single factor. Proc Natl Acad Sci. 2018;115(9):2180–5.

354. Lawrence HJ, Christensen J, Fong S, Hu YL, Weissman I, Sauvageau G, *et al*. Loss of expression of the Hoxa-9 homeobox gene impairs the proliferation and repopulating ability of hematopoietic stem cells. Blood. 2005;106(12):3988–94.

355. Tang L, Bergevoet SM, Gilissen C, de Witte T, Jansen JH, van der Reijden BA, *et al*. Hematopoietic stem cells exhibit a specific ABC transporter gene expression profile clearly distinct from other stem cells. BMC Pharmacol. 2010;10:12.

356. McGuckin CP, Forraz N, Baradez MO, Lojo-Rial C, Wertheim D, Whiting K, *et al*. Colocalization analysis of sialomucins CD34 and CD164. Stem Cells. 2003;21(2):162–70.

357. Watt SM, Butler LH, Tavian M, Bühring HJ, Rappold I, Simmons PJ, *et al*. Functionally defined CD164 epitopes are expressed on CD34(+) cells throughout ontogeny but display distinct distribution patterns in adult hematopoietic and nonhematopoietic tissues. Blood. 2000;95(10):3113–24.

358. Picelli S. Single-cell RNA-sequencing: The future of genome biology is now. RNA Biol. 2017;14(5):637–50.

359. Picelli S, Björklund ÅK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods. 2013;10(11):1096–8.

360. Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc. 2014;9(1):171–81.

361. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, *et al*. Comparative analysis of single-cell RNA sequencing methods. Mol Cell. 2017;65(4):631-643.e4.

362. Allers K, Hütter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, *et al*. Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation. Blood. 2011;117(10):2791–9.

363. Davis BR, Schwartz DH, Marx JC, Johnson CE, Berry JM, Lyding J, *et al*. Absent or rare human immunodeficiency virus infection of bone marrow stem/progenitor cells *in vivo*. J Virol. 1991;65(4):1985–90.

364. Shen H, Cheng T, Preffer FI, Dombkowski D, Tomasson MH, Golan DE, *et al*. Intrinsic human immunodeficiency virus type 1 resistance of hematopoietic stem cells despite coreceptor expression. J Virol. 1999;73(1):728–37.

365. Williamson C, Morris L, Maughan MF, Ping LH, Dryga SA, Thomas R, *et al*. Characterization and selection of HIV-1 subtype C isolates for use in vaccine development. AIDS Res Hum Retroviruses. 2003;19(2):133–44.

366. Beck Z, Jagodzinski LL, Eller MA, Thelian D, Matyas GR, Kunz AN, *et al.* Platelets and erythrocyte-bound platelets bind infectious HIV-1 in plasma of chronically infected patients. Stoddart CA, editor. PLoS One. 2013;8(11):e81002.

367. Stevenson M, Stanwick TL, Dempsey MP, Lamonica CA. HIV-1 replication is controlled at the level of T cell activation and proviral integration. EMBO J. 1990;9(5):1551–60.

368. Maréchal V, Clavel F, Heard JM, Schwartz O. Cytosolic Gag p24 as an index of productive entry of human immunodeficiency virus type 1. J Virol. 1998;72(3):2208–12.

369. Davis HE, Morgan JR, Yarmush ML. Polybrene increases retrovirus gene transfer efficiency by enhancing receptor-independent virus adsorption on target cell membranes. Biophys Chem. 2002;97(2–3):159–72.

370. Kolte L, Gaardbo JC, Skogstrand K, Ryder LP, Ersbøll AK, Nielsen SD. Increased levels of regulatory T cells (T regs) in human immunodeficiency virus-infected patients after 5 years of highly active anti-retroviral therapy may be due to increased thymic production of naive T regs. Clin Exp Immunol. 2009;155(1):44–52.

371. Armbruster DA, Pry T. Limit of blank, limit of detection and limit of quantitation. Clin Biochem Rev. 2008;29 Suppl 1:S49-52.

372. Panch SR, Szymanski J, Savani BN, Stroncek DF. Sources of hematopoietic stem and progenitor cells and methods to optimize yields for clinical cell therapy. Biol Blood Marrow Transplant. 2017;23(8):1241–9.

373. Clarke A, Fraser KPP. Why does metabolism scale with temperature? Funct Ecol. 2004;18(2):243–51.

374. Dimitrov DS, Willey RL, Sato H, Chang LJ, Blumenthal R, Martin MA. Quantitation of human immunodeficiency virus type 1 infection kinetics. J Virol. 1993;67(4):2182–90.

375. Marozsan AJ, Fraundorf E, Abraha A, Baird H, Moore D, Troyer R, *et al.* Relationships between infectious titer, capsid protein levels, and reverse transcriptase activities of diverse human immunodeficiency virus type 1 isolates. J Virol. 2004;78(20):11130–41.

376. Tebit DM, Zekeng L, Kaptué L, Kräusslich HG, Herchenröder O. Construction and characterisation of a full-length infectious molecular clone from a fast replicating, X4-tropic HIV-1 CRF02.AG primary isolate. Virology. 2003;313(2):645–52.

377. Bebenek K, Abbotts J, Wilson SH, Kunkel TA. Error-prone polymerization by HIV-1 reverse transcriptase. Contribution of template-primer misalignment, miscoding, and termination probability to mutational hot spots. J Biol Chem. 1993;268(14):10324–34.

378. Ndung'u T, Renjifo B, Novitsky VA, McLane MF, Gaolekwe S, Essex M. Molecular cloning and biological characterization of full-length HIV-1 subtype C from Botswana. Virology. 2000;278(2):390–9.

379. Wang Z, Hong K, Zhang J, Zhang L, Li D, Ren L, *et al*. Construction and characterization of highly infectious full-length molecular clones of a HIV-1 CRF07_BC isolate from Xinjiang, China. Jiang S, editor. PLoS One. 2013;8(11):e79177.

380. Miller D. A rapid and efficient method for concentration of small volumes of retroviral supernatant. Nucleic Acids Res. 1996;24(8):1576–7.

381. Adams A. Concentration of Epstein-Barr virus from cell culture fluids with polyethylene glycol. J Gen Virol. 1973;20(3):391–4.

382. Ichim CV, Wells RA. Generation of high-titer viral preparations by concentration using successive rounds of ultracentrifugation. J Transl Med. 2011;9(1):137.

383. Geraerts M, Willems S, Baekelandt V, Debyser Z, Gijsbers R. Comparison of lentiviral vector titration methods. BMC Biotechnol. 2006;6:34.

384. Izopet J, Tamalet C, Pasquier C, Sandres K, Marchou B, Massip P, *et al.* Quantification of HIV-1 proviral DNA by a standardized colorimetric PCR-based assay. J Med Virol. 1998;54(1):54–9.

385. Josefsson L, King MS, Makitalo B, Brannstrom J, Shao W, Maldarelli F, *et al*. Majority of CD4+ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule. Proc Natl Acad Sci. 2011;108(27):11199–204.

386. Brinchmann JE, Albert J, Vartdal F. Few infected CD4+ T cells but a high proportion of replication-competent provirus copies in asymptomatic human immunodeficiency virus type 1 infection. J Virol. 1991;65(4):2019–23.

387. Psallidopoulos MC, Schnittman SM, Thompson LM, Baseler M, Fauci AS, Lane HC, *et al*. Integrated proviral human immunodeficiency virus type 1 is present in CD4+ peripheral blood lymphocytes in healthy seropositive individuals. J Virol. 1989;63(11):4626–31.

388. Griffin DO, Goff SP. Restriction of HIV-1-based lentiviral vectors in adult primary marrow-derived and peripheral mobilized human CD34+ hematopoietic stem and progenitor cells occurs prior to viral DNA integration. Retrovirology. 2016;13(1):14.

389. Van Lint C, Bouchat S, Marcello A. HIV-1 transcription and latency: an update. Retrovirology. 2013;10(1):67.

390. Chelucci C, Hassan HJ, Locardi C, Bulgarini D, Pelosi E, Mariani G, *et al.* In vitro human immunodeficiency virus-1 infection of purified hematopoietic progenitors in single-cell culture. Blood. 1995;85(5):1181–7.

391. Xiao P, Usami O, Suzuki Y, Ling H, Shimizu N, Hoshino H, *et al.* Characterization of a CD4-independent clinical HIV-1 that can efficiently infect human hepatocytes through chemokine (C-X-C motif) receptor 4. AIDS. 2008;22(14):1749–57.

392. Saha K, Zhang J, Gupta A, Dave R, Yimen M, Zerhouni B. Isolation of primary HIV-1 that target CD8+ T Lymphocytes using CD8 as a receptor. Nat Med. 2001;7(1):65–72.

393. Skinner AM, O'Neill SL, Grompe M, Kurre P. CXCR4 induction in hematopoietic progenitor cells from Fanca(-/-), -c(-/-), and -d2(-/-) mice. Exp Hematol. 2008;36(3):273–82.

394. Ishii T, Nishihara M, Ma F, Ebihara Y, Tsuji K, Asano S, *et al.* Expression of stromal cell-derived factor-1/pre-B cell growth-stimulating factor receptor, CXC chemokine receptor 4, on CD34+ human bone marrow cells is a phenotypic alteration for committed lymphoid progenitors. J Immunol. 1999;163(7):3612–20.

395. Rutella S, Pierelli L, Bonanno G, Scambia G, Leone G, Rumi C. Homogeneous expression of CXC chemokine receptor 4 (CXCR4) on G-CSF-mobilized peripheral blood CD34+ cells. Blood. 2000;95(12):4015–6.

396. DiGiusto DL, Krishnan A, Li L, Li H, Li S, Rao A, *et al.* RNA-based gene therapy for HIV with lentiviral vector-modified CD34+ cells in patients undergoing transplantation for AIDS-related lymphoma. Sci Transl Med. 2010;2(36):36ra43-36ra43.

397. Meyerhans A, Cheynier R, Albert J, Seth M, Kwok S, Sninsky J, *et al.* Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. Cell. 1989;58(5):901–10.