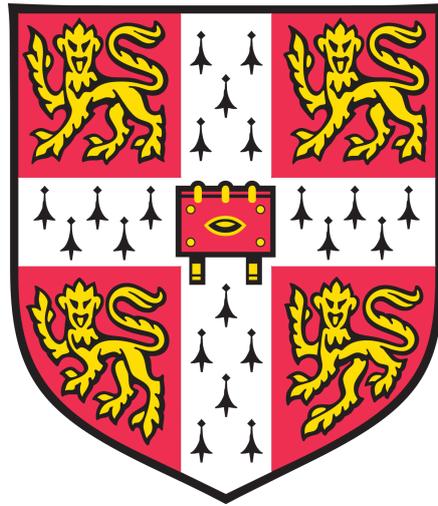


**MOLECULAR CHARACTERISATION OF EXPANDED
MOUSE HAEMATOPOIETIC STEM CELLS USING A
NOVEL *IN VITRO* REPORTER STRATEGY**



LOK CHI JAMES CHE

Hughes Hall
University of Cambridge
September 2020

This thesis is submitted for the degree of Doctor of Philosophy

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text.

It does not exceed the prescribed word limit for the School of Clinical Medicine Degree Committee.

MOLECULAR CHARACTERISATION OF EXPANDED MOUSE HAEMATOPOIETIC STEM CELLS USING A NOVEL *IN VITRO* REPORTER STRATEGY

Lok Chi James Che

Abstract

Haematopoietic stem cells (HSCs) are responsible for the lifelong maintenance of the blood forming system which produces trillions of blood cells daily. They are able to achieve this because of two defining properties: 1) they can give rise to progeny which eventually form all of the blood cell types in an organism and 2) they can create equally potent daughter cells. This latter property of self-renewal has the potential to be harnessed to create unlimited numbers of HSCs outside the body, which would be highly beneficial to cellular and gene therapy. As a result, decades of research have focused on improving *in vitro* HSC expansion efficiency with most studies failing to expand functional HSCs in sufficiently large quantities. Recent efforts in mouse HSC biology achieved a more than 200-fold expansion of functional HSCs; however, single cell cultures in these conditions displayed a large amount of heterogeneity. Using a recently generated HSC reporter mouse, I devised a novel *in vitro* reporter strategy capable of reading out functional HSC activity *in vitro* and also discovered a previously unreported population of lymphoid cells marked by the reporter (Chapter 3). I showed that the *in vitro* reporter strategy could be used to screen for molecules that promote HSC expansion and could prospectively identify single-cell derived cultures that contained large numbers of functional HSCs (Chapter 4). This permitted us to undertake gene expression profiling to determine the molecular identity of expanded HSCs using RNA sequencing. Comparing the transcriptome of these cells and the secretome of these heterogeneous clonal cultures, I identified potentially novel regulators for promoting the expansion of HSCs (Chapter 5).

Preface

Chapter 1

Some of the figures, in particular the depiction of the haematopoietic hierarchy and cell fate choices, were designed and produced by Mairi Shepherd.

Chapter 3

All work carried out was supervised and guided by Dr. David Kent. For animal work, I was trained by David Kent, Caroline Oedekoven, Mairi Shepherd, Miriam Belmonte, Tina Hamilton and Dean Pask. I carried out all of the animal work and experiments, except for the experiment characterising lymphoid tissues (lung, liver, thymus and lymph nodes) and the α -GalCer reactivity experiment, which were performed by Dr. Alyssa Cull and Dr. Jillian Barlow in University of York. All cell sorting, except in York, was carried out at the Flow Cytometry Core Facility at the Cambridge Institute for Medical Research (CIMR) by Dr. Reiner Schulte, Dr. Chiara Cossetti, Gabriela Grondys-Kortaba and Annie Hoxhalli. For RNA-sequencing experiments, I performed the RNA extraction with assistance from Daniel Bode; the library preparation and subsequent sequencing was carried out by the Genomics Core Facility at Cambridge Stem Cell Institute (CSCI) by Maïke Paramor and Vicki Murray. The bioinformatic analysis of the data were mostly performed by Daniel Bode and John Davey, with my involvement.

Chapter 4

All work carried out was supervised and guided by Dr. David Kent. As stated above, I carried out all of the animal work and experiments. Work with EL08 cells were partly guided by Robert Oostendorp. Work with F12-based cultures was done with guidance from Adam Wilkinson. All cell sorting was carried out at the Flow Cytometry Core Facility at CIMR by Dr. Reiner Schulte, Dr. Chiara Cossetti, Gabriela Grondys-Kortaba and Annie Hoxhalli. All proteomics of conditioned media samples were done in the Proteomics Core Facility by Dr. Robin Antrobus at CIMR. I performed the bioinformatic analysis of the dataset which was normalised by Robin Antrobus. The experimental design schematics were designed and produced by Anna Clay and Mairi Shepherd.

Chapter 5

All work carried out was supervised and guided by Dr. David Kent. I carried out all of the animal work and experiments, except for the secondary transplantations in the repeat of transplantation experiment, which were done by Dr. Alyssa Cull and Grace Boyd in the University of York. All

cell sorting was carried out at the Flow Cytometry Core Facility at CIMR by Dr. Reiner Schulte, Dr. Chiara Cossetti, Gabriela Grondys-Kortaba and Annie Hoxhalli. For RNA-sequencing experiments, I performed the RNA extraction with assistance from Daniel Bode; the library preparation and subsequent sequencing was carried out by the Genomics Core Facility at CSCI by Maike Paramor and Vicki Murray. The bioinformatic analysis of the data were mostly performed by Daniel Bode with my involvement. Miriam Belmonte also helped with figure preparation. The proteomic analysis was performed by Daniel Bode in collaboration with Dr. Jyoti Choudhary's group at the ICR in London. The experimental design schematics were designed and produced by Anna Clay and Mairi Shepherd.

Acknowledgements

This thesis would not have been possible without the support from all my family, friends, colleagues and funders.

First and foremost, I would like to thank my supervisor, Dr. David Kent, for providing me this opportunity. Even through two paternity leaves and moving lab location, he always tried his very best to be available for my support and guidance. I will be always be grateful for your teachings and knowledge, and for giving me the freedom to do the many weird and wacky experiments that I have pursued. Lastly my appreciation for his encouragement of me to have a life outside of science.

I would also like to thank Elisa Laurenti for agreeing to supervise me in Dave's absence, and the MRC for funding my research.

Of course, I wouldn't be here without all the amazing members of Kent Lab. Thank you for contributing to such a positive work environment and making it such an enjoyable experience to come to work every day. Special mention goes to Daniel Bode, Miriam Belmonte, Craig McDonald, Mairi Shepherd and Nina Øbro. My gratitude extends to the rest of CAB level 5 members for the enjoyable lunch time conversations.

Professionally, I could not have done this PhD without the help of the CIMR Flow Core Facility. Thank you, Reiner, Chiara, Gabi and Anni for being so good at your jobs.

Amongst my friends, I would especially like to thank my housemates and by extension, the Sunday-dinner crew for providing a wonderful living situation to come home to and for the most wholesome tradition. I would also like to thank Dane Sherburn for being an inspiration and constant source of positivity, and for always being there for coffees in London and interesting discussions.

Last but not least, I would like to thank my family for supporting me; Caitlin Walker, my partner who has supported me wonderfully during these stressful months of finishing my PhD; Charmaine Che, my sister, for being the home away from home and of course, my parents for supporting me throughout my whole life and pushing me to be the best version of myself.

Table of contents

ABSTRACT	III
DECLARATION	II
PREFACE	IV
ACKNOWLEDGEMENTS.....	VI
1 INTRODUCTION	1
1.1 Haematopoiesis.....	1
1.1.1 Myeloid cells.....	3
1.1.2 Lymphoid cells.....	4
1.1.3 Haematopoietic homeostasis.....	5
1.2 Haematopoietic stem cells	5
1.2.1 Early evidence for haematopoietic stem cells.....	7
1.2.2 HSCs in development	7
1.2.3 HSCs in ageing	8
1.2.4 The stem cell niche.....	10
1.2.5 Assays for stem and progenitors	14
1.2.6 HSC reporter mice	18
1.3 HSC heterogeneity.....	21
1.3.1 Cellular heterogeneity.....	21
1.3.2 Molecular heterogeneity.....	23
1.3.3 Haematopoietic hierarchy revised	29
1.4 HSC expansion	29
1.4.1 Clinical significance.....	29
1.4.2 Difficulties and challenges.....	31
1.4.3 HSCs derived from induced pluripotent stem cells or reprogramming.....	31
1.4.4 Harnessing endogenous HSC expansion.....	32
1.4.5 HSC fate choice.....	32
1.4.6 HSC intrinsic self-renewal pathways	34
1.4.7 HSC extrinsic self-renewal regulators.....	35
1.4.8 <i>Ex vivo</i> expansion of HSCs	36
1.4.9 Current state of the <i>ex vivo</i> expansion field.....	43
1.5 Thesis aims	44
2 METHODS.....	45
2.1 Mice.....	45
2.1.1 Genotyping.....	45
2.2 HSC isolation or BM analysis by flow cytometry	46
2.2.1 Bone marrow harvest.....	46
2.2.2 Red cell lysis	46
2.2.3 Lineage depletion	47
2.2.4 Isolating HSCs by FACS	47
2.2.5 Lymphoid and peripheral tissue harvest	48
2.2.6 Flow cytometric analysis or FACS isolation of <i>Fgd5</i> ⁺ EPCR ⁻ cells	49

2.3	<i>In vitro</i> assays	49
2.3.1	Cell culture images	49
2.3.2	CFU assays	49
2.3.3	OP9 assays	49
2.3.4	Intracellular flow cytometry	49
2.3.5	Stemspan (SS) based HSC cultures	50
2.3.6	F12-based 28-day HSC cultures	51
2.3.7	F12-based short-term (<10 days) cultures	51
2.3.8	EL08 conditioned media preparation	51
2.3.9	EL08 CM cultures	51
2.3.10	Flow cytometric analysis of <i>in vitro</i> cultures	51
2.4	Transplantation assays	53
2.4.1	Single cell transplantations	53
2.4.2	Secondary transplantations	53
2.4.3	Peripheral blood analysis	54
2.5	Antibodies	56
2.6	Flow cytometer set up	57
2.7	RNA sequencing	57
2.7.1	RNA-seq analysis	58
2.8	Proteomics	60
2.8.1	EL08 CM proteomic screen	60
2.8.2	F12 PVA culture media proteomic screen	60
2.9	Statistical analysis	61
RESULTS		62
3	IDENTIFICATION OF A NOVEL IMMUNE CELL SUBSET IN THE <i>FGD5</i>^{ZSGREEN+ZSGREEN/+} REPORTER MOUSE STRAIN	62
3.1	More than 50% of <i>Fgd5</i> ⁺ cells are not phenotypic HSCs	62
3.1.1	<i>Fgd5</i> ⁺ CD48dull ESLAM cells are not significantly different from normal ESLAMs	63
3.1.2	FE- cells are not HSCs	65
3.1.3	FE- cells are of haematopoietic lineage	67
3.1.4	FE- cells do not result from different BM preparation methods	68
3.2	Gene expression of FE- cells suggests a lymphoid cell identity	69
3.2.1	Surface marker phenotyping confirms lymphoid marker expression	71
3.2.2	FE- cells do not vary with age and are present in multiple lymphoid tissues.	72
3.2.3	FE- cells can be sub-fractionated with surface markers	73
3.3	Bulk RNA sequencing of CD5 ⁺ and CD244 ⁺ fractions	74
3.3.1	α-Galactosylceramide reactivity confirms that FE-CD5 ⁺ cells are iNKT cells	78
3.3.2	FE-CD5 ⁺ cells secrete interferon-γ	78
4	<i>FGD5</i> AND EPCR MARK HSCS <i>IN VIVO</i> AND <i>IN VITRO</i>	80
4.1.1	<i>Fgd5</i> and EPCR positive cells highly enriches LSK cells <i>in vitro</i>	80
4.1.2	<i>Fgd5</i> and EPCR together mark transplantable HSCs <i>in vitro</i>	82
4.1.3	Single positive <i>Fgd5</i> or EPCR cells contain fewer functional stem cells	84
4.2	<i>Fgd5</i> also marks FL HSCs	85

4.3	Novel reporter strategy validates HSC supportive culture conditions and identifies novel targets for supporting HSC expansion.....	86
4.3.1	EL08 CM improves survival and increases proliferation of HSCs.....	87
4.3.2	Short-term but not extended conditioning of media with EL08 cells support HSCs	87
4.3.3	Proteomic analysis of EL08 CM identifies self-renewal regulators previously discovered by gene expression studies and other novel targets	89
4.3.4	Recombinant proteins have no effect on the survival and clone sizes of HSCs	92
4.3.5	Pleiotrophin supports the expansion of phenotypic HSCs at low concentrations	93
4.4	F12 HSA cultures increase survival but do not support HSCs unless IL-11 is added.....	95
4.4.1	F12-based cultures have higher survival compared to SS-based cultures	96
4.4.2	F12-based cultures create smaller clones due to lower SCF concentrations	96
4.4.3	F12 HSA cultures do not support phenotypic HSCs unless supplemented with IL-11.....	97
4.4.4	Only F12-based cultures can maintain HSCs for 28 days.....	99
4.5	IL-11 is redundant in F12 PVA cultures	100
4.5.1	PVA and HSA have comparable survival rates and clone sizes.....	101
4.5.2	PVA is better than HSA at supporting phenotypic HSC expansion	101
4.5.3	Transplantations confirm that IL-11 has minimal effect on F12-PVA cultures.....	103
4.5.4	Single HSCs cultured in F12 PVA display clonal heterogeneity	104
4.5.5	Transplantation outcome can be retrospectively predicted by reporter strategy.....	105
5	MOLECULAR CHARACTERISATION OF CLONAL HETEROGENEITY IN F12 CULTURES .	108
5.1	Transplantation of 28-day clones from single cell cultures.....	108
5.1.1	Split dose transplantations confirm bona fide HSC expansion and functional heterogeneity...	110
5.1.2	ELSK separates functional HSCs from non-HSCs.....	111
5.1.3	FELSK correlates highly with functional HSC activity.....	112
5.2	Repopulating ELSK cells are molecularly distinct from non-repopulating cells and resemble freshly isolated HSCs at gene expression level.....	113
5.2.1	Non-HSCs are not mature cells but differentiated progenitors	117
5.3	Differential gene expression between ELSK and non-ELSK cells suggest early activation in non-ELSK cells.....	118
5.3.1	Rho GTPase pathways significantly upregulated in phenotypic HSCs.....	120
5.3.2	Differential gene expression between PosELSK and NegELSK	122
5.4	Identification of a molecular programme for “expanded” HSCs	127
5.4.1	Single cell HSPC transcriptome displays heterogeneous expression of signature genes	129
5.4.2	MapK signalling pathway overrepresented in signature genes	131
5.5	Non-ELSKs provide insights to feedback signals.....	132
5.5.1	The transcriptomes of Non-ELSK cells correlate with functional outcome of the clone	132
5.5.2	ROS associated pathways are upregulated in non-repopulating clones.	134
5.6	Secretome analysis identifies targets for positive and negative regulation.....	135
6	DISCUSSION	139
6.1	Context of study	139
6.2	Summary of major findings	140
6.3	Implications and future directions	142
6.3.1	Identification of novel iNKT1 and putative monocytes marked by <i>Fgd5</i> expression	142
6.3.2	<i>Fgd5</i> and EPCR: novel markers for HSCs <i>in vitro</i>	143

6.3.3	Proteomic analysis of conditioned media provides insights to expansion regulators	144
6.3.4	Validated and fully defined F12 and PVA-based expansion conditions	147
6.3.5	Molecular characterisation of <i>ex vivo</i> expanding HSCs.....	148
6.3.6	Unifying molecular signature of HSCs or distinct molecular signature of HSC subtypes	151
6.3.7	Future of <i>ex vivo</i> HSC expansion: translation into human HSCs.....	152
6.4	Concluding remarks.....	152
7	REFERENCES.....	153

List of figures

Figure 1.1 The haematopoietic hierarchy	2
Figure 1.2 Haematopoietic tissue homeostasis and the balance between HSC self-renewal and differentiation	3
Figure 1.3 Drosophila germ stem cells require contact with cap cells for their self-renewal.	11
Figure 1.4 The bone marrow niche	14
Figure 1.5 The emerging continuum of haematopoietic hierarchy.....	27
Figure 1.6 HSC fate choice possibilities.	33
Figure 2.1 Representative image of <i>Fgd5</i> ^{ZsGreen-ZsGreenr/+} genotyping experiment.....	46
Figure 2.2 Representative ESLAM Sca-1 ⁺ gating strategy.....	48
Figure 2.3 Representative images of different liquid culture colony sizes.....	50
Figure 2.4 Representative gating strategy for FELSK cells.....	52
Figure 2.5 Schematic of clonal analysis interpretation.	53
Figure 2.6 Representative gating layout for peripheral blood chimerism analysis.	55
Figure 2.7 MDS plots of batch corrections performed on bulk RNA-seq samples.	58
Figure 3.1 Phenotypic LT-HSCs are <i>Fgd5</i> ⁺	62
Figure 3.2 <i>Fgd5</i> ⁺ cells are not uniformly LT-HSCs.....	63
Figure 3.3 Representative gating layout for <i>Fgd5</i> ⁺ EPCR ⁺ CD150 ⁺ CD48 ^{dull} cells.....	64
Figure 3.4 The CD48 ^{dull} ESLAM phenotype contains a high proportion of functional HSCs. ..	65
Figure 3.5 FE- cells do not generate haematopoietic colonies in HSPC assays.	66
Figure 3.6 FE- cells do not grow in OP9 B cell progenitor assays.	66
Figure 3.7 Representative gating strategy for BM cells from repopulated primary transplantation recipient.	67
Figure 3.8 FE- correlates strongly with T cell output in primary transplantation.	68
Figure 3.9 FE- cells are present in both flushed and crushed bones.....	69
Figure 3.10 FE- cells express various surface markers characteristic of lymphoid and NK cells.	71
Figure 3.11 Frequency of FE- cells in mice of different ages.....	72
Figure 3.12 FE- cells in various lymphoid and peripheral tissues.....	73
Figure 3.13 Co-expression of immune surface markers allows FE- cells to be further subfractionated.....	74
Figure 3.14 SingleR identifies FE-CD5 ⁺ cells as NKT cells and FE-CD244 ⁺ cells as monocytes.	76
Figure 3.15 SingleR identifies FE-CD5 ⁺ cells as iNKT1 cells.	77
Figure 3.16 FE-CD5 ⁺ cells are reactive to α -GalCer.....	78

Figure 3.17 Intracellular flow cytometry reveals cytokine profile expressed by FE-CD5+ cells.	79
Figure 4.1 <i>Fgd5</i> and EPCR enriches for LSK cells in <i>ex vivo</i> cultures.	81
Figure 4.2 Schematic of experimental design to test the ability of <i>Fgd5</i> and EPCR to isolate functional HSCs <i>in vitro</i>	82
Figure 4.3 <i>Fgd5</i> and EPCR marks functional HSCs <i>ex vivo</i>	83
Figure 4.4 Single positive <i>Fgd5</i> or EPCR cells contain fewer functional stem cells.	84
Figure 4.5 <i>Fgd5</i> and EPCR mark functional stem cells even in cultures with FE+CD150- starting cells.	85
Figure 4.6 Foetal Liver cycling HSCs are <i>Fgd5</i> ⁺	86
Figure 4.7 Duration of conditioning alters beneficial effects of EL08 CM.	88
Figure 4.8 Heatmap of proteins identified within EL08 CM suggests a general increase in protein content as conditioning increases.	90
Figure 4.9 Label-free quantification of EL08 CM.	92
Figure 4.10 Pleiotrophin increases phenotypic HSCs in 10-day cultures.	94
Figure 4.11 F12-based cultures increases survival of HSCs <i>in vitro</i>	97
Figure 4.12 F12 cultures supplemented with IL-11 is comparable to SS-based culture.	98
Figure 4.13 F12-based cultures are superior to StemSpan based cultures for long-term HSC expansion.	100
Figure 4.14 HSA vs PVA with or without IL-11.	102
Figure 4.15 F12 PVA based cultures with or without IL-11	104
Figure 4.16 transplantation of single cell clones cultured with IL-11 for 28 days.	105
Figure 4.17 Donor chimerism correlates with phenotypic isolation strategy but not cell numbers.	106
Figure 5.1 Representative gating strategy to separate phenotypic HSCs from culture.	109
Figure 5.2 schematic of experimental design to characterise molecular heterogeneity in PVA cultures.	110
Figure 5.3 Transplantation of ELSK from single clones cultured in F12 PVA media.	111
Figure 5.4 Reconstitution of pooled non-ELSK cells compared to the respective ELSK cells.	112
Figure 5.5 FELSK correlates with chimerism.	113
Figure 5.6 MDS plots show clear separation between repopulating and non-repopulating cells.	115
Figure 5.7 MoIO score is significantly higher in repopulating and ELSK cells.	116
Figure 5.8 SingleR identifies PosELSK samples as most similar LTHSCs.	118

Figure 5.9 Volcano plot showing differentially expressed genes between ELSK and non-ELSK cells.	119
Figure 5.10 Pathways associated with upregulated genes in ELSK cells compared to non-ELSK cells.	121
Figure 5.11 Pathways associated with genes upregulated in non-ELSK cells compared to ELSK cells.	122
Figure 5.12 Volcano plot of genes differentially expressed between PosELSK and NegELSK cells.	123
Figure 5.13 PCA of all samples indicate that PC1 correlates highly with functional outcomes.	127
Figure 5.14 Signature gene score is significantly higher in repopulating and ELSK cells.....	129
Figure 5.15 Gene expression patterns of signature genes across HSPC subpopulations.....	130
Figure 5.16 Pathways associated with signature genes.	131
Figure 5.17 PosNonELSK cells separate from NegNonELSK cells and correlate with chimerism of the clone.	132
Figure 5.18 Volcano plot of differentially expressed genes between NegNonELSK and PosNonELSK cells.....	133
Figure 5.19 Pathways associated with genes upregulated in NegNonELSK cells vs PosNonELSK cells.	135
Figure 5.20 Venn diagram of proteins identified in the secretome of repopulating and non-repopulating clones.....	135
Figure 5.21 Interaction map of unique proteins in media from repopulating clones.	136
Figure 5.22 Interaction map of unique proteins in media from non-repopulating clones. .	137

List of tables

Table 1.1 Classification of NKT and NKT-like cells adapted from Godfrey et al. ¹⁴	5
Table 1.2 Table of HSC reporter mice	19
Table 1.3 Summary of expansion protocols, modified and updated from Walasek et al. ³⁸³	41
Table 2.1 Primer sequences for Fgd5 genotyping.....	45
Table 2.2 List of antibodies used, with respective clone numbers and manufacturer.	56
Table 2.3 Flow cytometer set up for BD LSR Fortessa.....	57
Table 2.4 Flow cytometer set up for BD Influx.....	57
Table 3.1 Expression of cell surface marker genes associated with FE- cells.....	69
Table 3.2 Top 40 Cluster of differentiation genes expressed by FE- cells.....	70
Table 3.3 GO terms associated with top 500 genes expressed in FE- cells.....	70
Table 4.1 Tally of clones above and below FELSK cut-off and Fisher's exact test result.	89
Table 4.2 Tally of clones above and below FELSK cut-off and Fisher's exact test result.	95
Table 4.3 Tally of clones above and below FELSK cut-off and Fisher's exact test result.	99
Table 4.4 Tally of clones above and below FELSK cut-off and Fisher's exact test result.....	103
Table 5.1 12 clones with matched ELSK and non-ELSK samples were chosen for RNA-seq.	114
Table 5.2 Gene ontology terms based on upregulated genes in ELSK cells compared to Non-ELSK cells.	119
Table 5.3 Gene ontology terms based on upregulated genes in Non-ELSK cells compared to ELSK cells.	120
Table 5.4 Manually curated gene ontology terms based on upregulated genes in PosELSK compared to NegELSK.	123
Table 5.5 Manually curated gene ontology terms based on upregulated genes in NegELSK compared to PosELSK.....	124
Table 5.6 Significantly upregulated genes in PosELSK vs NegELSK cells.....	125
Table 5.7 Significantly upregulated genes in NegELSK vs PosELSK cells.....	126
Table 5.8 List of 30 genes identified by PC1 loading plot to be signature genes for expanding HSCs.	128
Table 5.9 Gene ontology terms of upregulated genes in PosNonELSK cells vs NegNonELSK cells.	133
Table 5.10 Gene ontology terms of upregulated genes in NegNonELSK cells vs PosNonELSK cells.	134

List of abbreviations

°C	Celsius
3D	3-dimensional
5-FU	Fluorouracil
7AAD	7-amino-actinomycin D
AF	Alexa Fluor
AGM	Aorta Gonad Mesonephros
AML	Acute myeloid leukaemia
APC	Allophycocyanin
BD	BD Biosciences
BM	Bone Marrow
BME	β -Mercaptoethanol
BrdU	Bromodeoxyuridine
BSA	Bovine serum albumin
BV	Brilliant violet
CB	Cord blood
CD	Cluster of differentiation
CFC	Colony forming cell
CFU-S	Colony forming unit - spleen
CLP	Common lymphoid progenitor
CM	Conditioned media
CM _{10D}	Media conditioned for 10 days
CM _{14D}	Media conditioned for 14 days
CM _{1D}	Media conditioned for 1 day
CM _{3D}	Media conditioned for 3 days
CMP	Common myeloid progenitor
Col1	Collagen 1
CPM	Counts per million
Cre	Cre recombinase
Cy	Cyanine
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
E	Embryonic day
ECM	Extracellular matrix

EDTA	Ethylenediaminetetraacetic acid
EL08	EL08-1D2
ELSK	EPCR ⁺ Lineage ⁻ Sca-1 ⁺ c-Kit ⁺
ESC	Embryonic stem cell
Exp	Experiment
F12	Ham's F-12 Nutrient Mix
FACS	Fluorescenc-activated cell sorting
FC	Fold-change
FCS	Foetal calf serum
FDR	False discovery rate
FE-	<i>Fgd5</i> ⁺ EPCR ⁻
FE+	<i>Fgd5</i> ⁺ EPCR ⁺
FELSK	<i>Fgd5</i> ⁺ EPCR ⁺ Lineage ⁻ Sca-1 ⁺ c-Kit ⁺
FGF	Fibroblast growth factor
F ^{hi} E ^{hi}	<i>Fgd5</i> ^{high} EPCR ^{high}
F ^{hi} E ^{lo}	<i>Fgd5</i> ^{high} EPCR ^{low}
FL	Foetal liver
F ^{lo} E ^{hi}	<i>Fgd5</i> ^{low} EPCR ^{high}
F ^{lo} E ^{lo}	<i>Fgd5</i> ^{low} EPCR ^{low}
Flt3L	Flt3 Ligand
GM	Granulocytes and monocytes
GO	Gene ontology
Gy	Gray
HDGF	Heptaoma-derived growth factor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HLA	Human leukocyte antigen
HPX	Haemopexin
HSA	Human serum albumin
HSC	Haematopoietic stem cell
HSCT	Haematopoietic stem cell transplantation
HSPC	Haematopoietic stem and progenitor cell
IFN-γ	Interferon gamma
IGF	Insulin growth factor
IGFBP	Insulin-like growth factor binding protein
IL	Interleukin

ILC	Innate lymphoid cell
iNKT	Invariant NKT
iPSC	Induced pluripotent stem cells
IT-HSC	Intermediate-term haematopoietic stem cell
ITSX	Insulin-Transferrin-Selenium-Ethanolamine
LDA	Limiting dilution assay
LSK	Lineage ⁻ Sca-1 ^{+c} -Kit ⁺
LT-HSC	Long-term haematopoietic stem cell
MDS	Multiple dimensional scaling
min	Minutes
mL	Millilitre
mM	millimolar
MolO	Molecular overlapping
MPP	Multipotent progenitor
mTOR	Mammalian target of rapamycin
NegELSK	ELSK cells from non-repopulating clones
NegNonELSK	Non-ELSK cells from non-repopulating clones
ng	Nanogram
NGF	Nerve growth factor
NH ₄ Cl	Ammonium chloride
NK cell	Natural killer cell
NKT cell	Natural killer T cell
OncM	Oncostatin M
OPN	Osteopontin
p	P-value
P-restricted	Platelet restricted
PBS	Phosphate buffer saline
PBS2%	PBS supplemented with 2% foetal calf serum
PCA	Principle component analysis
PE	Phycoerythrin
PFA	Formaldehyde
PGE2	Prostaglandin E2
PMA	Phorbol 12-myristate 13-acetate
PosELSK	ELSK cells from repopulating clones
PosNonELSK	Non-ELSK cells from repopulating clones

PTN	Pleiotrophin
PVA	Poly-vinyl alcohol
QC	Quality control
qPCR	Quantitative polymerase chain reaction
Rho	Rhodamine123
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
ROS	Reactive oxygen species
RT	Room temperature
RT-PCR	Reverse transcription- polymerase chain reaction
s	Second
SCF	Stem cell factor
scRNA-seq	Single Cell RNA sequencing
SCT	STEMCELL technologies
SD	Standard deviation
Sigma	Sigma Aldrich
SR1	StemRegenin1
SS	StemSpan Serum-Free Expansion Medium
ST-HSC	Short-term haematopoietic stem cell
TGF- β	Tumour growth factor-beta
TNF- α	Tumour necrosis factor - alpha
TPO	Thrombopoietin
UG26	UG26-1B6
VPA	Valporic acid
Vwf	von Willebrand factor
W41	C57BL/6 ^{W41/W41} -Ly5.1
WBC	White blood cell
WT	Wild type
α -GalCer	α -Galactosylceramide
μ L	Microlitre
μ m	Micrometre
μ M	Micromolar

1 Introduction

1.1 Haematopoiesis

The blood system has two main functions. Firstly, it is vital for transporting nutrients and signalling molecules, such as oxygen, carbon dioxide and hormones to and from cells around the body¹. Secondly, it produces the immune cells that allow the body to ward off infections and mount immune responses¹. Its main component, blood, is a liquid tissue consisting of aqueous plasma and a collection of specialised cells that serve specific functions¹. Blood plasma is made up of mostly water and contains organic molecules such as proteins, nucleic acids, electrolytes, sugars, vitamins and amino acids, which serves as everything the body needs to maintain its function¹. Amongst the cellular portion, there are two main classes of specialised blood cells – myeloid cells (more details in section 1.1.1 below) and lymphoid cells (section 1.1.2 below)². It has been estimated that the body generates between 100 billion to 1 trillion of these diverse blood cells every day³.

The process of forming these blood cellular components is known as haematopoiesis (from ancient Greek: Haima meaning “blood”; Poiesis meaning “to make”)⁴. As depicted in Figure 1.1, haematopoiesis is considered to be a hierarchical process. At the apex of the hierarchy is a rare group of cells residing in the bone marrow (BM) called haematopoietic stem cells (HSCs) that give rise to progenitor cells with limited self-renewal ability and potency, which in turn give rise to the mature specialised cells with very little to no self-renewal ability³. To provide the billions of blood cells required every day, the bulk of the proliferation is done by the progenitor cells⁵. Under homeostatic conditions, HSCs are kept in a quiescent state, also known as G_0 ⁶. The reduced proliferative burden is thought to protect HSCs from accumulating mutations during DNA replication and cell division, increasing their longevity^{6,7}. At the population level, the balance between differentiation and self-renewal must be tightly regulated in order to ensure homeostasis⁵. Too much differentiation can lead to exhaustion of the HSC pool; too much self-renewal accompanied by lack of maturation can lead to haematological malignancies such as myeloproliferative diseases or in more severe cases leukaemia (Figure 1.2)⁵.

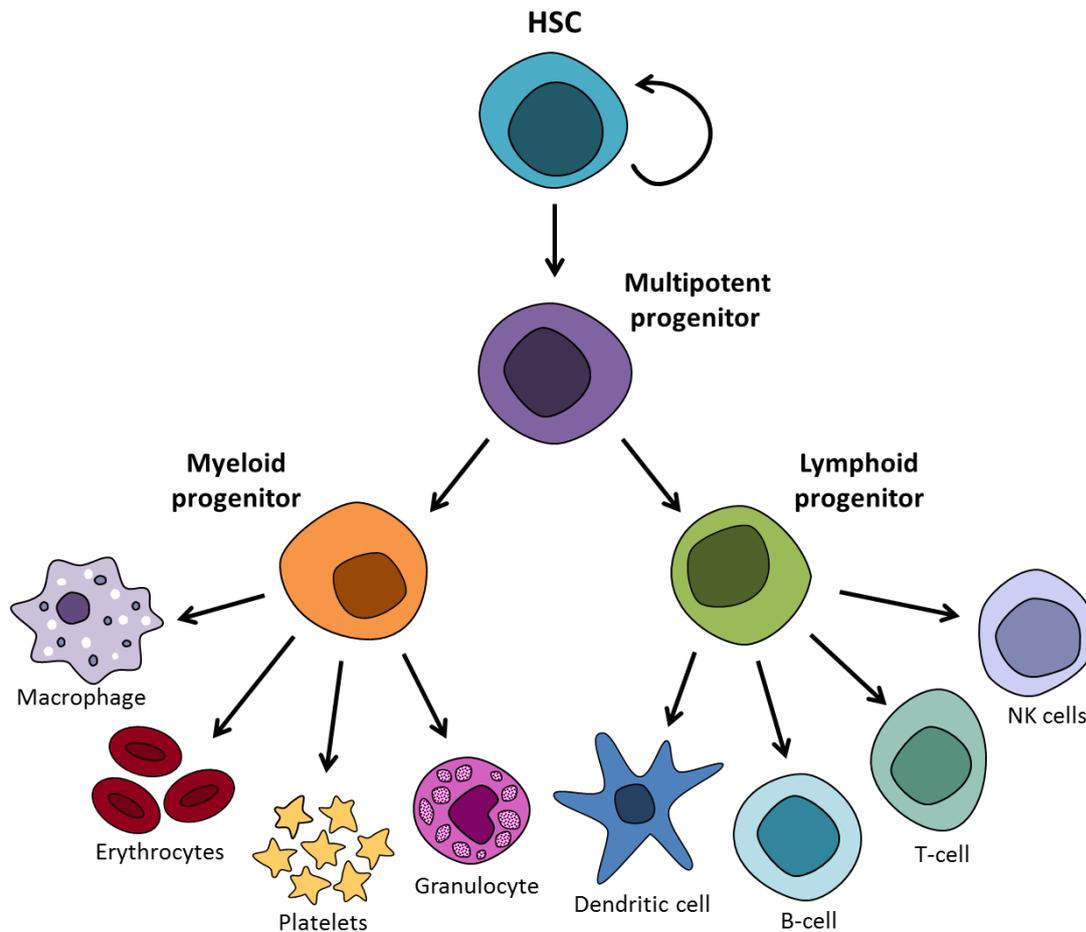


Figure 1.1 The haematopoietic hierarchy

A simplified haematopoietic hierarchy illustrating the concept of stepwise differentiation from HSCs (the apex) to stages of lineage-restricted progenitors and finally to specialised mature cell types. Through tight regulation, all the mature blood cells required throughout life are produced and replenished from the HSC. Illustration by Mairi Shepherd (unpublished).

Historically, to study the various functions of individual cell types, the field has utilised cell surface markers to categorize different subsets of cells within this haematopoietic hierarchy, also known as immunophenotyping⁵. This has led to an increasingly refined differentiation tree with increasing numbers of immune cell subtypes and consequently increasing functional purity of marked populations⁵. Classical categorisations have presented a stepwise process of differentiation started with long-term HSCs (LT-HSCs) to short-term HSCs (ST-HSCs) to multipotent progenitors (MPPs)⁸; the first major branch point in the differentiation tree occurs with lineage restricted progenitors such as the common myeloid progenitor (CMP) and common lymphoid progenitor (CLP)⁸; These oligopotent progenitors give rise to unipotent progenitors and eventually to more mature cell types⁸. Typically, each step along the hierarchy results in more restricted self-renewal and differentiation potential and recent cellular and molecular evidence points towards a continuum of differentiation and self-renewal potential (discussed in detail in section 1.3.3 below)

⁸. That said, the broad concept of the hierarchically organised tissue has remained useful for stem cell biology.

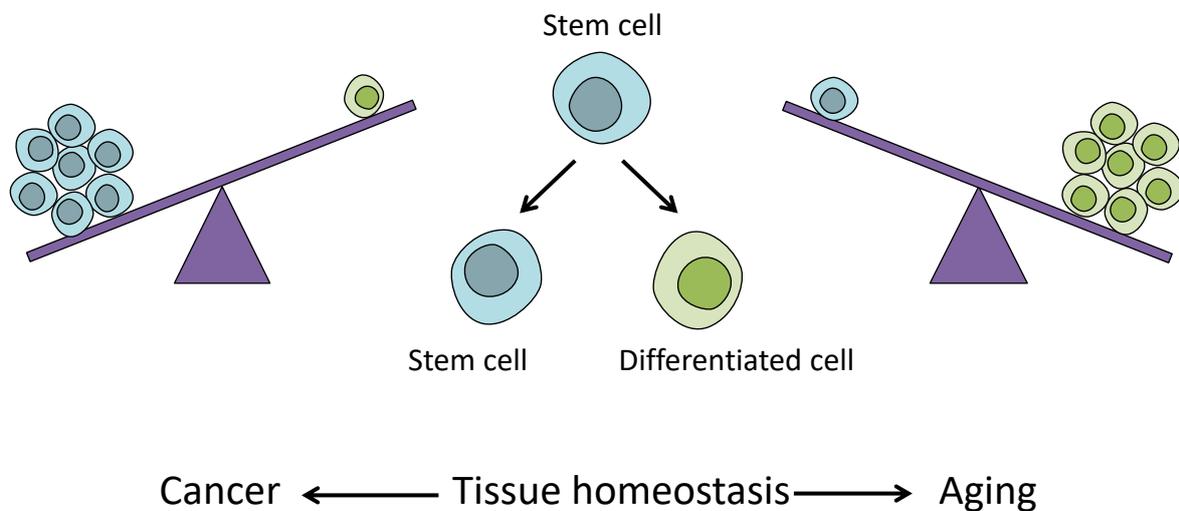


Figure 1.2 Haematopoietic tissue homeostasis and the balance between HSC self-renewal and differentiation

At the population level, HSCs must balance between differentiation and self-renewal in order to ensure proper homeostasis. Too much self-renewal can lead to haematological malignancies like myeloproliferative diseases or in more severe cases leukaemia, provided there is a differentiation block; Too much differentiation can lead to the depletion of the HSC pool. Illustration adapted from Mairi Shepherd (unpublished).

1.1.1 Myeloid cells

Generally speaking, the broad classes of myeloid cells include erythrocytes, monocytes, macrophages, granulocytes, mast cells, megakaryocytes and thrombocytes¹. Each cell lineage has specialised jobs that ensure peripheral tissues are supplied with nutrients and are monitored for infection¹.

Accounting for about 40-45 percent of the total volume of blood¹, erythrocytes, also known as red blood cells, are the most abundant cell type in blood and are specialised to carry oxygen around the body¹. In comparison, the rest of the cellular components (i.e. the white blood cells (WBCs)) described below make up about 1 percent of total blood volume¹. Megakaryocytes are the largest cells in the marrow that break up to form much smaller thrombocytes (or platelets), which are technically not cells but fragments of cells¹. Thrombocytes play an important role in blood clotting, by adhering to each other and onto injured surfaces¹. Monocytes give rise to the macrophages and myeloid dendritic cells. Macrophages are large immune cells capable of phagocytosing pathogens as well as infected cells across many tissues in the body⁹. Similar to macrophages, dendritic cells are specialised antigen presenting cells that phagocytose infectious agents and present the antigen fragments on its surface for cells of the adaptive immune system to recognise¹. Granulocytes,

characterised by the presence of granules in the cytoplasm, are the most numerous types of WBCs¹. There are 3 primary subtypes of granulocytes: neutrophils, basophils and eosinophils; which are separated by the colour of the granules when the cells are stained by a compound dye. Neutrophils and eosinophils are actively phagocytic and respectively function against bacteria and parasites¹, whereas basophils mainly release the contents of their granules, including histamine and leukotrienes¹. Mast cells are similar to basophils in terms of morphology, and also contain granules loaded with histamine and other compounds¹⁰. When activated, mast cells release these immune mediators which induce inflammation¹.

1.1.2 Lymphoid cells

Lymphoid cells, or lymphocytes, constitute about 30-40 percent of WBCs¹. Named for their abundance in the lymphatic system, which include the lymph nodes, spleen, thymus tonsils and lymphoid tissues around the gastrointestinal tract, lymphocytes broadly include 3 classes of cell types: B cells, T cells and Innate lymphoid cells, which includes natural killer cells (NK cells)^{5,11}.

As the name suggests, innate lymphoid cells (ILCs) are innate immune cells that are derived from CLPs¹². They include 5 subclasses: NK cells, ILC1s, ILC2s, ILC3s and LT_i cells¹². Of these, the most well-known are NK cells, which are cytotoxic cells that kill cancerous or virus-infected cells¹². All ILCs excrete an array of cytokines which mediate the immune response¹².

Unlike innate immunity, adaptive immunity is highly targeted and specific to antigens. B and T cells are important components of adaptive or acquired immunity to foreign cells and antigens¹.

B cells are responsible for producing the antibody-driven adaptive immune response. Upon encountering a foreign substance or antigen, B cells differentiate into plasma cells, which secrete antibodies to which bind to the antigen and neutralises it. T cells are important for cell-mediated, cytotoxic adaptive immune response¹. There are two main classes of T cells: T killer and T helper cells. Killer cells are responsible for triggering apoptosis in infected cells that it recognises, while T helper cells secrete cytokines that facilitate the response of T killer cells¹.

More recently described are a rare group of T cells (0.1% of peripheral blood T cells), known as Natural killer T (NKT) cells that share some properties of both T cells and NK cells¹³. The term NKT cell was first described as a subset of T cells that express natural killer cell marker NK1.1¹⁴. It is now generally used to describe CD1d⁺-restricted T cells. They are heterogenous and can further be subdivided into 3 classifications summarised by Godfrey et al. in Table 1.1 (below)¹⁴. Of note, Type 1 NKTs, also known as invariant NKT (iNKT) cells, can be further subdivided into iNKT1, iNKT2 and iNKT17. In Chapter 3, I describe a previously unknown subpopulation of immune cells marked by expression of *Fgd5*, which identifies a novel subset of iNKT1 cells.

Table 1.1 Classification of NKT and NKT-like cells adapted from Godfrey et al.¹⁴

	Type 1 NKT	Type 2 NKT	NKT-like
Other names	classical NKT invariant NKT (iNKT) V α 14i NKT (mouse) V α 24i NKT (human)	non-classical NKT diverse NKT	NK1.1 ⁺ T cells CD3 ⁺ CD56 ⁺ T cells
MHC Restriction	CD1d	CD1d	MHC, other?
α -GalCer reactivity	+	-	-
T-cell-receptor repertoire	V α 14-J α 18: V β 8.2, 7, 2 (mouse) V α 24-J α 18: V β 11 (human)	diverse	diverse
NK1.1	+(resting mature) -(immature or post-activation)	+/-	-
IL4 production	+	+	-
IFN γ production	+	+	+

1.1.3 Haematopoietic homeostasis

Haematopoiesis is a dynamic and responsive process. Over the course of a lifetime, the haematopoietic system has to respond to changes in the environment and adapt to stresses such as infections¹⁵, sudden or chronic losses of blood⁷, malignant haematopoiesis¹⁶ and chemotherapy¹⁷. A clear demonstration of the responsiveness of HSCs to environmental stress is its ability to replenish the haematopoietic system after transplantation into marrow-ablated recipients, perhaps most powerfully illustrated by transplantation of single HSCs in experimental models¹⁸. At each point along the haematopoietic hierarchy, cells have to make one of many possible “fate choices”, such as to proliferate, self-renew, differentiate or to initiate apoptosis (discussed in more detail throughout later sections)⁵. These choices are regulated by extrinsic and intrinsic cues, providing feedback mechanisms that maintain proper functioning haematopoietic homeostasis¹⁹. As the source of the majority of haematopoietic cells, HSCs are of particular interest and will be the focus of the next section.

1.2 Haematopoietic stem cells

Adult stem cells, or tissue-specific stem cells, are cells with the ability to produce specialised cells of a given tissue and retain this ability for the lifetime of an organism. HSCs are one of the best

studied tissue stem cells and, as mentioned, are ultimately responsible for producing all mature blood cell types. They are able to do this because they are multipotent (i.e., they can make the multiple different types of mature blood cells) and they can self-renew (i.e. they can generate daughter cells with the same properties as the parent cells). Self-renewal is particularly important because it gives the HSC the ability to generate one or two equally potent cells upon cell division.

As mentioned, HSCs are usually in a quiescent state. HSCs have also been demonstrated to be in a low metabolic state marked by low mitochondrial activity, which is autophagy dependent^{20,21}. Several metabolic studies have suggested that HSCs rely mainly on glycolysis instead of oxidative phosphorylation²². Unlike senescent cells, however, HSCs can re-enter the cell cycle upon stimulation, although they typically return to quiescence⁶. Contrastingly, haematopoietic progenitors are highly proliferative and metabolically active cells that depend on oxidative metabolism^{7,22,23}.

In label-retaining studies, the most dormant HSCs, which retain the label over months at steady state, have been demonstrated to possess the most robust repopulation potential in transplantation assays^{6,24-27}. Additionally, when stimulated, LT-HSCs exit quiescence and enter cell cycle much later than ST-HSCs^{6,28,29}. In studies looking at Bromodeoxyuridine (BrdU) retention, LT-HSCs have been estimated to divide once every 55-145 days^{6,24}. In stark contrast, ST-HSCs have been shown to divide at least 4 times more often. In a later study, Takizawa et al. showed that in steady state, HSCs with life-long multilineage repopulation potential are present in cycling (>5 times in 7 weeks) and quiescent (no divisions in 14 weeks) cells³⁰. However, these studies assume that the label can stain all HSCs, and thus it is difficult to ascertain the conclusions. In *Tie2Cre-YFP* labelled HSCs, Busch et al. estimated that 1 in 110 LT-HSCs differentiate into ST-HSCs per day³¹. However, they did find that fluorouracil (5-FU) challenge increased the output of HSCs, showing that the division rate is dependent on environmental factors. Controversially, Bernitz et al. had reported that using fluorescent Histone 2B fusion protein (H2B-FP) label retaining assays, HSCs divide four times before entering quiescence permanently²⁵. A very recent paper by Morcos et al. have challenged this, by arguing that there is leaky H2B-FP expression with age and that this background expression can be misinterpreted as label retention. They found that HSCs continue to cycle at slow rates throughout old age³².

In addition to the reduced proliferative burden which helps prevent accumulation of random DNA mutations during cell divisions, HSCs benefit from additional DNA damage protection by

expressing high amounts of DNA damage repair and cytoprotective genes such as ABC/MDR transporter genes³. Studies have further suggested that quiescence is an actively maintained and tightly regulated state, characterised by low levels of protein synthesis and mitochondrial oxidative phosphorylation³³. Low levels of oxidative metabolism are hypothesised as a protective mechanism to generate fewer Reactive Oxygen Species (ROS), which may cause cellular damage³⁴.

1.2.1 Early evidence for haematopoietic stem cells

In seminal experiments during the 1960s, James Till and Ernest McCulloch provided the first formal evidence for the properties of multi-lineage differentiation and self-renewal to be contained within a single cell in the adult haematopoietic system. Building on the discovery that lethally irradiated animals could be rescued by injection of BM cells³⁵, they were able to show experimentally that these cells form macroscopic colonies in the spleen of irradiated mice in the first few weeks after transplantation³⁶. Furthermore, the number of colonies formed on the spleen were proportional to the number of BM cells injected. This technique allowed for an approximation of the frequency of spleen colony forming units (CFU-S) at 1 in 10,000 BM cells³⁶. Within each spleen colony there were cells from multiple lineages and, in a follow up paper, each colony was also demonstrated to be clonal, meaning the cells in the colony had derived from a single starting cell³⁷. Finally, the discovery that these colonies can form more colonies in secondarily transplanted mice proves that such cells are capable of self-renewal³⁸. Thus, the key properties of functional HSCs, self-renewal and multipotency, were established. These early experiments also gave additional insights about HSCs, including the description of heterogeneity in the cell division rates and differentiation potential of different cells capable of making a spleen colony – providing the first glimpse into HSC functional heterogeneity discussed later in section 1.3 below. Furthermore, studies based on tritiated thymidine uptake suggest that CFU-S initiating cells are mostly quiescent³⁹. Research over the last decades has revealed that CFU-S are not the most primitive haematopoietic subset and there exists a more primitive cell type that has much more extensive self-renewal potential, known as the HSC⁴⁰.

1.2.2 HSCs in development

To meet the demands of a growing organism, a natural period of HSC expansion occurs during embryonic development. HSCs isolated from developing embryo have markedly different properties compared to adult HSCs⁴¹, and their consequent cell fate choices are differently balanced. During development in mice, haematopoiesis occurs in two separate waves: the primitive and definitive wave⁴¹. Primitive haematopoiesis first occurs around Embryonic (E) day 7.5 (E17 in humans) in the yolk sac blood islands. In this wave, erythroid progenitors are formed in order

to provide the early embryo with red blood cells and some macrophages that it requires for oxygenation⁴¹. Because these primitive cells lack self-renewal capacity, the primitive wave is transient and transplantable HSCs are not generated¹¹. Definitive haematopoiesis (sometimes referred to as adult haematopoiesis) occurs subsequently at E8.25 (E21 in human) in the aorta gonad mesonephros (AGM) region and placenta and involves the generation of the first transplantable HSCs⁴²⁻⁴⁵. At E9.5, erythromyeloid progenitors appear in the extra embryonic yolk sac, followed by the placenta. At E9 in mice, the foetal liver (FL) begins to be colonised by haematopoietic cells derived from the seeding tissues such as placenta, yolk sac and AGM. Around E11 and E12, the FL has already become the main site of definitive haematopoiesis until the BM is established just before birth with peak HSC production at approximately E14.5. From then on and throughout adulthood, the BM is the primary site of haematopoiesis⁴⁶.

FL HSCs undergo a massive expansion, increasing in numbers by 10-30 fold within 4 days⁴⁷. While there may still be some seeding of HSCs generated from the placenta^{43,48,49}, most of the increase in foetal HSCs is due to their frequent execution of symmetrical self-renewal divisions^{50,51}. This is reflected by their faster cycling rates compared to adult HSCs - all foetal HSCs are cycling while >75% of adult HSCs are quiescent⁵². In the early stages after transplantation, FL HSCs repopulate recipients more quickly than their adult counterparts and generate more HSCs through symmetric self-renewal^{50,53,54}. Around 6 weeks after transplantation, FL HSCs adopt a more adult-like self-renewal capacity⁴¹, resembling the natural transition that occurs between 3 and 4 weeks after birth⁴¹. While certain genes have been identified to be exclusively important for FL HSCs, such as *Sox17*⁵⁵ and *Ezh2*⁵⁶; or vice versa in adult HSCs, such as *Bmi1*⁵⁷, *Gfi1*^{58,59} and *Cebpa*⁶⁰, much remains to be discovered about the molecular regulation of the 3-4 week switch. Thus far, the Lin28b-let-7-Hmga2 axis has been shown to be a key pathway regulating this developmental transition⁶¹.

1.2.3 HSCs in ageing

It is clear that as humans and mammals age, they are generally less able to perform homeostatic functions and one of the hallmarks of ageing that contributes to this is the exhaustion of stem cells⁶²; HSCs are no exception. As HSCs age, they gradually lose self-renewal and regenerative potential⁶³. Studies have shown that when competed against young HSCs in transplantation settings, young HSCs are functionally superior in terms of the number and balance of myeloid and lymphoid cells they produce⁶⁴⁻⁶⁶. In both mouse and humans, ageing is associated with an increase in myeloid cells and studies have shown that aged HSCs are myeloid-biased or rather lymphoid deficient⁶⁷⁻⁶⁹. Interestingly, the absolute numbers of phenotypic HSCs also increases during ageing⁷⁰, which coincides with an increase in heterogeneity of individual HSCs^{64,65,71,72}.

This finding is especially interesting in light of the recently described phenomenon of age-related clonal haematopoiesis, where the majority of blood cells are derived from a small number of clones. This reduced genetic diversity can be observed at the stem cell level and also independently in more mature cells⁶³. This could be a result of an outgrowth of a dominant stem cell, or a decrease in the total number of stem cells, the latter of which would allow mutations to be detected more readily amongst a smaller pool of HSCs. Notably, clonal haematopoiesis occurs in healthy individuals and although it is associated with higher risks of malignancy, it is not in and of itself a sign of disease⁶³.

There are several cell intrinsic mechanisms that have been reported to contribute to HSC ageing, including DNA damage, telomere shortening, epigenetic misregulation and loss of cell polarity⁷³. Defects in DNA repair machinery have been linked to premature stem cell ageing in both mice and patients⁷⁴. The decline in stem cell activity can be accelerated upon replicative stresses, due to the increased likelihoods of random mutations with each replication^{74,75}.

A secondary mechanism, whereby replication can cause DNA damage is by the shortening of telomeres. HSCs express telomerase which should elongate telomeres⁷⁶, however telomeres still shorten during ageing in humans and mice^{77,78}. In both humans and mice, the telomere deficiencies are associated with HSC decline^{77,79,80}. However, the overexpression of telomerase in mice does not fully rescue functional impairment of HSCs, suggesting that shortening telomeres is not the only part of the picture⁸¹. Of note, the telomeres in laboratory mice are much longer than mice found in the wild, and thus, the manipulations of telomerase expression may not have significant observable effects^{82,83}.

In aged HSCs, there are signs of impaired levels of autophagy, which normally functions to recycle organelles such as mitochondria²⁰. Mitochondria are found to accumulate in aged HSCs, which causes an increase in metabolic stress and reactive oxygen species (ROS) production^{34,84}. The accumulation of ROS has been shown to compromise HSC function and the reduction in mitochondrial stress can reverse the loss of stem cell function⁸⁴. One mechanism through which ROS was suggested to affect HSCs was by inducing mitochondrial DNA mutations. It was shown that mitochondrial DNA mutations accumulate with age and partially drive haematopoietic dysfunction, however it was also shown that HSCs are relatively resistant⁸⁵.

Apart from random mutations, clonal haematopoiesis is highly associated with certain somatic mutations, including in epigenetic and transcriptional regulator genes such as *JAK2*, *DNMT3A*, *TET2* and *ASXL1*^{86–90}. This suggests a mechanism of selection at play in clonal haematopoiesis, whereby cells that acquire certain mutations can outcompete other clones either through expansion or increased resilience. However, several papers have also reported coordinated changes to the HSC niche (see section 1.2.4 below) during ageing that are associated with clonal haematopoiesis or malignancy^{91–94}, suggesting the possibility that selection is induced by a changed microenvironment that favours one clone over others. The extent of the role of the ageing niche has yet to be fully understood.

1.2.4 The stem cell niche

The concept of the HSC ‘niche’, the idea that there are specialised sites within the BM that support HSCs, was first proposed by Ray Schofield⁹⁵. He suggested that the interactions with other cells form a unique environment that maintains the self-renewal activity of an HSC. This concept was hugely influential in other stem cell fields and there have been many examples described since (reviewed by Xie et al.⁹⁶). A particularly striking demonstration of this idea came when germ stem cells (GSCs) in *Drosophila melanogaster* were found to require direct contact with the cap cells within the ovary (Figure 1.3)^{97,98}. It was further demonstrated that this physical contact was mediated by drosophila E-cadherin and that cap cells express Dpp, a bone morphogenetic protein (BMP) ligand, which is essential for GSC self-renewal^{98,99}. It was subsequently discovered that BMP signalling is regulated by binding to collagen IV, which is expressed by plasmacytes^{100,101}. Thus, the concept of a physical stem cell niche was supported and understood at the molecular level, which is now referred to as the microenvironment surrounding the stem cell.

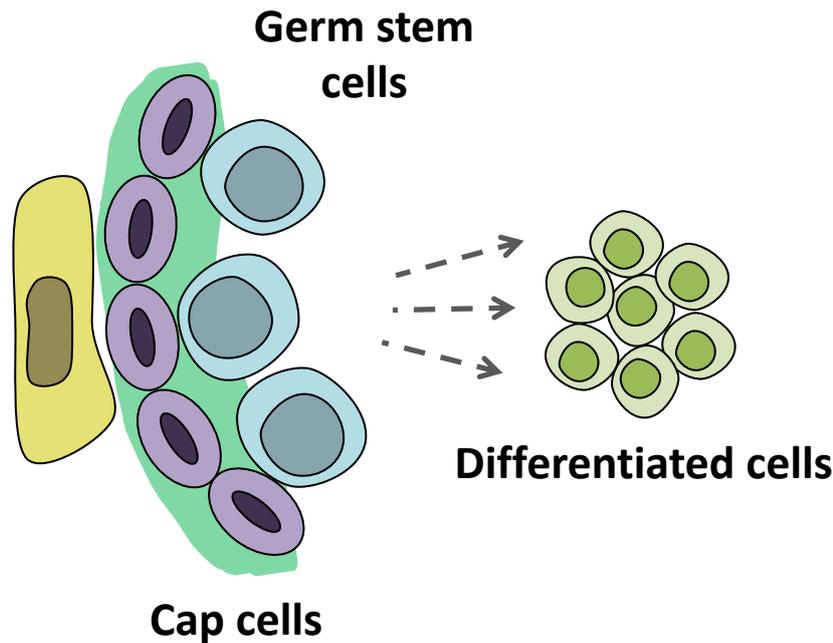


Figure 1.3 *Drosophila* germ stem cells require contact with cap cells for their self-renewal.

GSCs require contact with cap cells that express Dpp in order to maintain their stem cell identity. Once contact is removed, they begin to differentiate. Figure adapted from Mairi Shepherd (unpublished).

Postnatally, HSCs primarily reside in the BM, though they are also sometimes found in circulation in the blood. Nonetheless, they are extremely rare cells that comprise about 0.003% of the total BM cells (1 in every 30,000 cells)¹⁰². While all bones support haematopoiesis in mice, in humans the major site of haematopoiesis is in the axial skeleton (cranium, sternum, ribs, vertebrae and ilium)¹⁰³. However, most of our knowledge of the BM niche comes from studies of long bones in mice¹⁰⁴. Notably in humans, apart for the proximal regions, long bones have limited haematopoietic activity¹⁰⁴.

Many studies have looked at the architecture of the BM to understand what the microenvironment around an HSC would be like. As reviewed by Boulais et al. and Pinho et al.^{104,105}, the BM is a highly vascularised organ, situated within the cortical bone, which mostly functions mechanically as a hard shell. The endosteum lines the inside of the cortical bone, forming the interface between the bone and the marrow. Around the endosteum, osteoblasts produce the bone itself and osteoclasts are responsible for breaking down bone.

Trabecular bone forms the inner bone tissue and is sometimes known as “cancellous” or “spongy bone”. Longitudinal arteries and veins run parallel to the long axis of the cortical bone. The central artery branches into smaller radial arteries and arterioles, often close to the endosteum. Fluid is drained from the marrow through the sinusoidal network, which is distributed throughout the

marrow and merges into a central sinusoid forming the venous circulation. Through this network of arteries and veins, HSCs and other haematopoietic cells can enter and exit the BM into blood circulation around the body. Lastly the BM is interspersed with sympathetic nerves, and parasympathetic fibres may also innervate the distal femoral metaphysis. In adult mouse BM, arterioles are wrapped by sympathetic nerves and perivascular stromal cells, forming structures known as the neuro-reticular complex. These stromal cells are characterised by the expression of pericyte marker neural-glial antigen 2 and the type VI intermediate filament protein Nestin (NES).

With regards to the specific components of the HSC BM niche that regulate the maintenance and differentiation of HSCs, there has been numerous studies that have interrogated the cell types and molecular signals at play. Paradoxically, almost every cellular constituent of the BM has been suggested to play a role in HSC biology^{103,104,106-108}. Some groups have suggested that there are distinct niches for HSC subpopulations¹⁰⁹, although this may be complicated by the fact that HSCs are a heterogeneous population with distinct properties and no HSC reporter yet exists with 100% specificity.

The anatomical location of HSCs and their proximity to various niche cells therefore remains a topic of debate. The earliest investigation of HSC localisation by Lord et al. used a small gauge needle to flush out cells from the centre of the femur and showed that CFU-S are more concentrated near the bone surface¹¹⁰. In early imaging experiments utilising traditional histological and intra-vital microscopy imaging techniques, transplanted Haematopoietic stem and progenitor cells (HSPCs) were found to home towards the endosteum of recipient mice^{111,112}. This finding was supported by functional changes in HSCs upon genetic ablation of osteoblasts found near the endosteum¹¹³. However, these studies are limited by the purity of HSCs in question, the extensive manipulation during HSC isolation and the perturbed nature of the BM during transplantation and irradiation. Many subsequent studies have supported the endosteum being the primary HSC niche^{112,114-119}.

However, more recent studies have challenged this view of the endosteal niche. In these studies, endogenous HSCs marked using surface markers were located close to endothelial cells, pointing towards a perivascular niche^{120,121}. Several studies have pointed to mesenchymal cells, mainly wrapped around sinusoids, as important players in the niche¹²². In particular, it was demonstrated that mesenchymal stem and progenitor cells can form ectopic haematopoietic sites when transplanted subcutaneously or sub-renally^{123,124}. Important mesenchymal cells found in close

proximity to HSCs include Nes-GFP⁺ perivascular cells¹²⁰, Leptin receptor-expressing perivascular cells¹²⁵ and CXCL12-abundant reticular cells^{126,127}, which appear to be highly overlapping populations. HSCs are known to express high levels of chemokine receptor CXCR4, which bind to the chemokine CXCL12 and genetic deletion of CXCR4 severely reduces HSC numbers¹²⁸. Furthermore, many factors known to be important HSC regulatory signals are expressed by perivascular cells, including stem cell factor (SCF)^{102,129,130}, osteopontin (OPN)^{115,116}, CXCL12^{127,128}, granulocyte colony stimulating factor (G-CSF), interleukin 6 (IL-6), pleiotrophin (PTN)^{131,132} and angiopoietin^{114,133}. Genetic ablation of these cells or their cytokine production causes a reduction in HSCs¹⁰⁸. Within the perivascular niche, it remains unclear whether HSCs preferentially reside closer to sinusoids, or arterioles. Some reports have suggested that HSCs that are more quiescent reside closer to arterioles^{109,134,135}. To complicate the matter, even Schwann cells have been implicated in the maintenance of HSCs^{136,137}.

Since imaging techniques are limited by the number of fluorophores used and staining is particularly difficult *in vivo*, the development of HSC reporter mice has been particularly instrumental in the study of HSC niches *in vivo*. In a study of *Hoxb5* reporter mice, marked HSCs were found to be associated with VE-cadherin⁺ endothelial cells¹³⁸. Using the *α-catulin* reporter mice, a recent study suggested that HSCs are not associated with arterial vessels but instead randomly distributed in the BM in a way that is indistinguishable from randomly generated dots¹³⁹.

Recently, quiescent HSCs in aged mice were shown to reside predominantly in perisinusoidal niches, rather than near arterioles⁹¹. Even more recently, a group used a dual genetic strategy to create a reporter called MFG mouse, which uses a *Flt3^{Cre}* allele to cut out the GFP coding sequences in the *Mds1*-GFP reporter marking primitive haematopoietic progenitors.¹⁴⁰ They showed that this subset of HSCs reside almost exclusively next to sinusoids near the endosteal surface, and not arterioles. While previous studies have shown the importance of hypoxic environments to the HSC niche^{141–143}, this study found that HSCs are not in the deepest hypoxic regions, but rather moderately hypoxic environments.

However, as detailed in section 1.2.6, none of these reporters have marked HSCs exclusively. With the improvement of *in vivo* live imaging techniques and more reliable HSC reporter/markers, it should be possible to clarify the controversy. The advancement in characterising niche cell subtypes, using single cell RNA-sequencing (scRNA-seq) technology, will also be useful^{144–146}. Different niches such as the FL niche can also be examined to determine core HSC regulators.

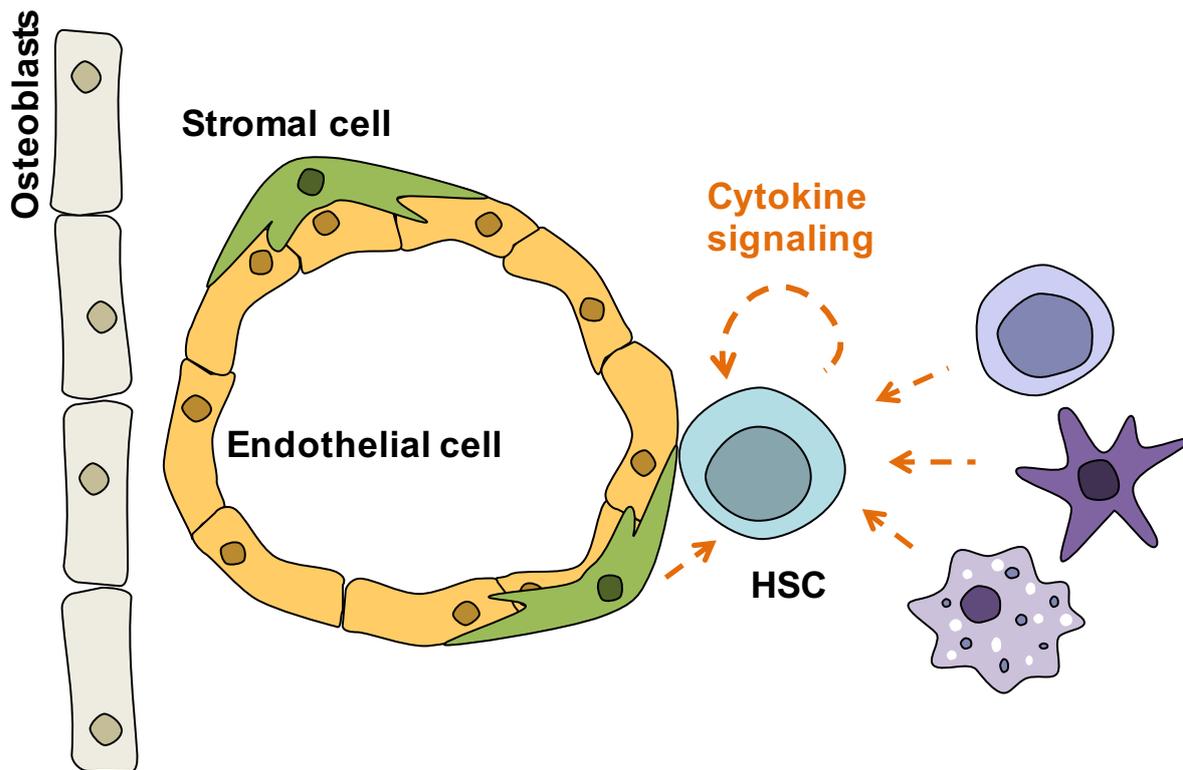


Figure 1.4 The bone marrow niche

The BM is a complex tissue where the HSC resides. It contains arterioles, sinusoids, stromal cells, various haematopoietic cells as well as osteoclasts and osteoblasts. Many of these cells have been implicated in the regulation of HSC function. Illustration by Mairi Shepherd (unpublished).

1.2.5 Assays for stem and progenitors

In vivo HSC transplantation assay

The first *in vivo* assay for stem-cell activity was based on the ability of HSCs to rescue lethally irradiated recipients by BM transplantation³⁵. As mentioned in section 1.2.1, CFU-S assays were used in the early days of the field to track cells that upon transplantation home to the spleen and form macroscopic colonies, providing short-term radioprotection to the mouse¹⁴⁷. To this day, BM transplantation is still used clinically, and it is still the gold standard for testing whether a cell population contains HSCs or when transplanted as a single cell, whether that cell is an HSC. Inherent in this, from the very start, HSCs were given a functional definition: a cell that, when transplanted serially, can sustain long-term multi-lineage engraftment. This has been useful for the field but is bound by the nature of retrospective assays and the common need to perturb the haematopoietic niche.

To track the progeny of HSCs, several groups pioneered the use of random retroviral integration sites as clonal markers to track the lineage output of donor cells^{148–150}. This method was largely superseded by the CD45 congenic system, which typically involves two different CD45 alleles (CD45.1 and CD45.2) on the donor and recipient cells, which can be detected using different monoclonal antibodies¹⁵¹. CD45 encodes a surface protein common to all haematopoietic cells except erythrocytes and platelets, hence it can detect donor contributions to multiple lineages when combined with an antibody panel of other lineage markers, such as B220 (B cell), CD3 (T cell) or Mac-1 and Ly6g (granulocyte and macrophages)¹⁵¹. Recently, the CD45 congenic system has been combined with *Vwf* or *Gata1* transgenic system, in order to study five lineage donor contribution of platelets and erythrocytes additionally¹⁵¹. Because HSC reconstitution kinetics are slow, transplantations are usually accompanied by a dose of helper BM cells that contain haematopoietic progenitors with transient reconstituting potential, thereby securing the short-term survival of the recipients. Commonly, transplantations are combined with a limiting dilution assay (LDA), which allows the number of HSCs to be estimated within the original cell population¹⁵².

The most stringent assay for self-renewal division events is the split doublet transplantation assay, also known as paired daughter cell transplantation assay¹⁹. In this assay, a single HSC is stimulated to divide once *in vitro*, and each daughter cell is then transplanted into different irradiated recipients to assay their respective reconstitution potential¹⁵³.

The definition of an HSC has evolved slightly over the years and typically it is defined as a cell that can sustain multi-lineage engraftment for an extended period of time (>16 weeks) upon serial transplantation into irradiated recipient mice¹⁴⁷. Multilineage engraftment is commonly determined by donor blood lineages exceeding thresholds (usually >1%) and the length of engraftment and ability to repopulate secondary or tertiary recipients are requirements for its self-renewal capacity¹⁵⁴. Secondary transplantation is the most rigorous way to test for LT-HSCs because only self-renewing LT-HSCs can reconstitute primary recipients and still have their progeny be able to do the same¹⁹. Cells that can reconstitute primary recipients for only 16 weeks but fail to reconstitute secondary recipients are termed ST-HSCs. If secondary reconstitution is only transient in nature, the donor cells are termed intermediate-term HSCs (IT-HSCs)¹⁵⁵.

***In vitro* colony assays**

As the transplantation assay is a time consuming and expensive assay, the field has developed *in vitro* assays that assess the various different functional properties of HSCs. The colony forming cell (CFC) assay, also known as the methylcellulose assay, is used to assess a cells ability to form

multilineage colonies in semi-solid media (made of agar or methylcellulose mixtures), upon cytokine stimulation^{156,157}. To detect more primitive cells, the long-term culture initiating cell (LTC-IC) assay cells can be used to assess the number of progenitors made in a longer culture period (as a surrogate for self-renewal) and myeloid and lymphoid differentiation capacity can be assessed^{158,159}. In the assay, cells are typically co-cultured with adherent stromal cells for an extended period of time (minimum 5 weeks)¹⁶⁰⁻¹⁶². After this period, the culture is assayed for the number of CFUs present, which represent the progeny of original LTC-ICs, as input CFUs would have undergone terminal differentiation at this point. Stromal cells can also be genetically modified, for example to express human growth factors to enhance sensitivity of the assay¹⁶³. Similar to the LTC-ICs, cobblestone area-forming cell assays can be used to test for more primitive cells¹⁵⁷. In this assay, stromal feeder layers are also used to support colonies of haematopoietic cells, which adhere and form observable “cobblestone areas”¹⁵⁷. Because CFU assays tend to be insufficient at supporting B and T cell development *in vitro*, stromal cell lines, such as OP9 and OP9-DL1 (OP9 cells overexpressing the Delta1 ligand) have been specifically established to expand B and T cell progenitors¹⁶⁴. However, due to varying stromal cells, interlaboratory variability is often a concern¹⁴⁷. As a result, liquid media culture assays have succeeded them in recent years for both human and murine HSCs^{28,102,165} (detailed in section 1.4).

HSC purification strategies

Multi-parameter fluorescence activated cell sorting (FACS) was a significant breakthrough that allowed the purification of single stem and progenitor cells from the BM⁵. It does this by combining the interrogation of fluorescent molecules attached to antibodies specific for surface proteins with physical isolation of single cells. Isolated cells can then be tested for their ability to contribute to blood cell formation in transplanted recipients for periods of months to years – the current functional definition of an HSC². Steady improvements in the technology and the array of cell surface proteins and antibody combinations have led to the isolation of long term self-renewing HSCs at an ever-increasing purity and these efforts are summarised in this section.

HSCs were first described to not express lineage specific cell surface markers (Lin⁻), such as Mac-1, B220, CD3, CD19, TER119 and Gr-1, and have low amounts of cell surface marker Thy1 (Thy1^{low}Lin⁻)¹⁶⁶. Importantly, this demonstrated that HSCs can be enriched using negative selection of cell surface markers. In 1988, a landmark study appeared in *Science* where Spangrude et al. reported that Stem Cell Antigen 1 (Sca-1) could be used as a positive marker to greatly enrich HSCs¹⁶⁷. Eight years later, the first formal demonstration that a single cell possessed both multilineage reconstitution ability and durable self-renewal was published by Osawa et al. where single

Lin⁻Sca-1⁺c-Kit⁺(LSK)CD34⁻ cells reconstituted lethally irradiated mice on their own¹⁸. Of these CD34⁺LSK cells, about 21% of single cells were reported to have long-term multilineage reconstitution. Subsequently, a number of different strategies using an ever-increasing number of cell surface markers were described to better enrich HSCs including CD34⁻¹⁸, CD38⁺¹⁶⁸, CD105⁺¹⁶⁹, Flt2/Flt3⁻¹⁷⁰, EPCR⁺¹⁷¹, CD49b¹⁵⁵, SLAM family markers (CD48⁻, CD150⁺, CD229⁻, CD244⁻)¹²¹, ESAM¹⁷² and others¹⁷³. Additionally, the ability of HSCs to efflux certain dyes such as Rhodamine-123⁻¹⁷⁴, Hoechst 33342⁻¹⁷⁵ can also be exploited as an isolation strategy. Interestingly, very few of these markers are essential to HSC function⁵.

Similar efforts in humans have been hampered by the lack of equivalent functional transplantation assays. A major breakthrough came with the use of the severe combined immune-deficient (*Scid*) mouse, which lacks B and T cells^{176,177}. The compromised immune system allowed the mouse to be xenotransplanted with human HSCs without complete immune rejection¹⁷⁶. The engraftment efficiency was further improved by increasing the level of immuno-compromisation, as in the non-obese diabetic (NOD)-*Scid* mice, which has added defects in its innate immune system¹⁷⁸. As reviewed by Goyama et al.¹⁷⁹, the best available xenograft model currently uses NOD-*Scid-IL2R β* ^{-/-} (NSG) mice, which has complete T Cell, B cell and NK cell ablation, and express human cytokines to make up for mouse cytokines that are not cross-reactive¹⁷⁶. However, even with the improvement in human xenograft mouse models, the purity of HSCs achieved to date is substantially lower than in the mouse system. The highest enrichment to date uses Lin⁻CD34⁺CD38⁻CD45RA⁻CD90⁺CD49f⁺ cells from umbilical cord blood in a xenograft assay and is reported to contain up to 9.5% long term repopulating multi-lineage cells in primary transplantation¹⁸⁰ (69% could give rise to secondary engraftment). However, this percentage decreases substantially in human BM and peripheral blood^{180,181}. Human samples are an outbred population and therefore also have the problem of being highly variable¹⁷⁶. Moreover, the added variability resulting from different lifestyles of individual humans can affect the behaviour and activity of HSCs¹⁷⁶. Combined with the lack of readily available tissue samples, in the form of BM aspirates or cord blood samples, studying self-renewal mechanisms in single human HSCs is currently not tractable despite numerous studies which have engaged in their molecular profiling^{182,183}. My thesis therefore focuses on the study of mouse HSCs with the vision to apply the findings in future to human HSC biology.

Currently, the most advanced mouse HSC isolation strategies have reported purities (defined with stringent serial repopulating ability) of >40%⁵. My thesis largely uses the strategy published in 2015 involving the ESLAM markers (CD45⁺CD150⁺CD48⁻EPCR^{high})^{102,184}. This strategy obtains mouse

LT-HSCs at >50%, is developmentally stable (e.g., isolates LT-HSCs at high purities from FL through to old age^{184,185}), is easy to obtain from primary mouse BM, and is a relatively simple strategy of just 4 markers, making it easy to combine with various reporter mice. Addition of Sca-1^{bright} to this sorting strategy increases enrichment to 67%¹⁸⁶, although some HSC subtypes are selectively depleted using this strategy and are therefore not appropriate for all assays.

1.2.6 HSC reporter mice

While the functional purities of >50% achieved by surface marker isolation strategies is impressive, especially considering the HSC population represent 0.003% of BM cells¹⁰², one could argue that the inability to purify HSCs completely is due to the limitations of only using cell surface markers. In recent years, several HSC reporter mice have been developed, spurred by efforts to label and study HSCs in the native BM niche. As summarised in Table 1.2 below, several groups have developed mouse models using different genes, such as *Abcg2*¹⁸⁷, *Fgd5*¹⁸⁸, *Vwf*¹⁸⁹, *Cttnal1*¹³⁹ and *Hoxb5*¹³⁸, that are expressed exclusively within the phenotypically defined primitive HSC compartment. In particular, reporter mice generated for *Cttnal1*, *Hoxb5* and *Fgd5* have been reported to enrich for HSCs in primary transplantation to a frequency of 14.9%, 47.6% (in combination with other markers) and 31.2% respectively. Single cell gene expression studies have also validated these genes to be highly enriched in the long-term HSC cluster¹⁹⁰, though the technology is limited by its inability to demonstrate the absence of expression in more mature progenies. In the recently developed MFG mouse¹⁴⁰ that was mentioned in section 1.2.4, only 11% of phenotypic LT-HSCs (SLAM, CD150⁺CD48⁻LSK) expressed the GFP marker; and out of the total GFP cells, 85% are SLAM HSCs. Limiting dilution analysis of MFG cells suggests the purity of serially transplantable HSCs is around 1 in 9 (11%)¹⁴⁰. Although still not perfect at labelling HSCs, these reporters could present useful tools for labelling cultured HSCs.

For my doctoral studies, I utilised the *Fgd5*^{ZsGreen•ZsGreen/+} mouse because the reporter was validated across multiple phenotypically defined HSC populations and was functionally tested in multiple transplantation settings¹⁸⁸. *Fgd5* had not previously been studied in HSC biology and encodes a protein with Guanosine exchange factor (GEF) activity which has been suggested to act through CDC42 in the VEGF activation pathway^{191,192}. Gazit et al. showed that reporter knock-in mice for *Fgd5* had exclusive labelling of phenotypic HSCs (defined as LSKCD48⁻CD150⁺)¹⁸⁸. They also showed that all HSC activity was confined to the small fraction of *Fgd5*⁺ cells in the BM¹⁸⁸. Of note, homozygous knock-ins are embryonic lethal and thus all reporter mice are heterozygous for the gene¹⁸⁸.

Table 1.2 Table of HSC reporter mice adapted from Pinho et al.¹⁰⁵

Mouse strain	Genetic modification	Specificity within the adult haematopoietic compartment	Analysis	Year published	Ref
<i>Hoxb5</i> –Tri-mCherry	Knock-in	Specific to long-term HSCs	• Modified CUBIC clearing and light-sheet microscopy	2016	138
			• Flow cytometry		
			• Transplantation		
<i>Cttnal1</i> –GFP	Knock-in	Restricted to HSPCs; requires KIT staining to enrich for HSCs	• Modified Murray’s clearing, immunostaining and confocal and multiphoton microscopy	2015	139
			• Flow cytometry		
			• Transplantation		
<i>Fgd5</i> –mCherry	Knock-in	Restricted to HSCs; low expression in haematopoietic progenitors	• Flow cytometry	2014	188
			• Transplantation		
<i>Vwf</i> –GFP	Transgenic	Labels platelet-biased and myeloid-biased HSCs, megakaryocyte progenitors, megakaryocytes and platelets	• Flow cytometry	2013	189
			• Transplantation		
			• Immunostaining and confocal microscopy		
<i>Msi2</i> –GFP	Knock-in	Labels haematopoietic progenitors	• Confocal microscopy	2016	193
			• Flow cytometry		
<i>Pdzk1ip1</i> –GFP	Transgenic	Enriches for highly purified HSCs but also labels a small subpopulation of haematopoietic progenitors and mature granulocytes	• Doxycycline chase	2016	194
			• Transplantation		
			• Flow cytometry		
<i>Evi1</i> –GFP	Knock-in	Labels haematopoietic progenitors	• Flow cytometry	2011	195
			• Transplantation		
<i>Scf</i> –tTA–H2B–GFP	Transgenic	H2B–GFP ^{high} label-retaining cells are enriched in quiescent long-term HSCs; also labels haematopoietic progenitors	• Doxycycline chase	2008	6
			• Transplantation		
			• Flow cytometry		
			• Immunostaining and confocal microscopy		
<i>Krt7</i> –GFP	Knock-in	Specific to HSCs	• Flow cytometry	2017	196
Continued...					

Mouse strain	Genetic modification	Specificity within the adult haematopoietic compartment	Analysis	Year published	Ref
<i>Tie2</i> -GFP	Transgenic	Enriches for highly purified HSCs but also labels haematopoietic progenitors	<ul style="list-style-type: none"> • Flow cytometry • Transplantation • Immunostaining and whole-mount confocal and multiphoton microscopy 	2016	197
<i>Gprc5c</i> -GFP	Transgenic	Enriches for dormant HSCs but also labels haematopoietic progenitors	<ul style="list-style-type: none"> • Flow cytometry • Transplantation 	2017	198
<i>Hdc</i> -GFP	Transgenic	Enriches for myeloid-biased HSCs but also labels haematopoietic progenitors and mature myeloid cells	<ul style="list-style-type: none"> • Flow cytometry • Transplantation • Immunostaining and confocal microscopy 	2017	199
<i>Gata2</i> -GFP	Knock-in	Enriches for HSCs and haematopoietic progenitors; SCA-1 or lineage staining is required for HSC selectivity	<ul style="list-style-type: none"> • Flow cytometry • Transplantation • Immunostaining and microscopy 	2006	200
<i>Hoxb4</i> -YFP	Knock-in	Labels haematopoietic progenitors	<ul style="list-style-type: none"> • Flow cytometry • Transplantation 	2011	201
<i>Mds1</i> -GFP/ <i>Flt3Cre</i> (MFG)	Knock-in	Specific to LT-HSCs	<ul style="list-style-type: none"> • Flow cytometry • Transplantation • Immunostaining and microscopy 	2020	140
<i>Abcg2</i> -YFP	Transgenic	Labels stem cell and progenitors for blood, small intestine and testicular germ cells.	<ul style="list-style-type: none"> • Flow cytometry • Lineage tracing 	2011	187
<i>Flt3Cre</i> /mT/mG (Flkswitch)	Transgenic	Labels LT and some ST-HSCs.	<ul style="list-style-type: none"> Flow cytometry Transplantation 	2011	202

1.3 HSC heterogeneity

1.3.1 Cellular heterogeneity

As early as 1964, it was shown that individual CFU-S had highly variable secondary colony forming ability¹⁴⁰. In later *in vitro* culture experiments, heterogeneity in colony production and differentiation was also observed even though external stimuli were kept the same²⁰³. Heterogeneity *in vivo* was further suggested through retroviral marking of HSCs which permitted tracking over extended periods of time, thereby mapping different kinetics of functional activity in transplantation^{204,205}. However, it was never entirely clear whether the observed heterogeneity was due to an inability to purify HSCs to 100%, which might mean the observed heterogeneity was due to contaminating non-HSCs. As HSC isolation strategies improved though, it became feasible to interrogate cellular heterogeneity through single cell transplantation. Single cell transplantation studies have revealed an immense amount of information about HSC heterogeneity in terms of differences in self-renewal potential, differentiation potential, repopulation kinetics and quiescence¹⁴⁸.

The first description of HSC heterogeneity at the single cell level was from Muller-Sieberg et al. where it was shown that lineage contributions and repopulation kinetics of two daughter HSCs deriving from the same clone were very similar²⁹. Strikingly, the study showed that only a small fraction of HSCs had a “balanced” multilineage output with equal contributions to all mature blood cell types and the majority of HSCs had a bias towards the cell types they generate, even though they produced all cell types measured. This conflicted with the classical model of hierarchical haematopoiesis at the time, where each HSC was implicitly assumed to behave similarly in generating mature haematopoietic progeny⁸. Interestingly, this lineage bias was conserved even in secondary transplantation, suggesting a cell intrinsic role in lineage biases^{29,206–208}, though it could be argued that the initial seeding of an HSC in a primary mouse results in distinct niches exerting influence on the production of daughter HSCs^{29,185,206,209}. Indeed it was later discovered by Benz et al. that some myeloid biased HSCs can produce balanced HSCs and vice versa^{109,135}, although the former was far more prevalent. A more intensive characterisation by Dykstra et al. revealed that “myeloid-biased” HSCs tended to have similar myeloid output to balanced HSCs, and are lymphoid deficient rather than myeloid-biased¹⁸⁵. Interestingly, lymphoid biased HSCs were found to contain only finite self-renewal (unable to repopulate secondary mice)²⁰⁶. While this could suggest that the heterogeneity in lineage outputs and self-renewal ability is linked, it could also be explained by the fact that lymphoid cells have a longer lifespan and thus

lymphoid cells that are detected at 16 weeks could have been generated weeks before and not produced from an actively contributing HSC. Importantly these findings have been supported by barcoding experiments, suggesting that lineage biased HSCs exist not just in transplantation settings but in native haematopoiesis as well²⁰⁸. Notably, myeloid biased HSCs have been shown to be enriched based on high CD150 levels, low CD49b levels or high Hoechst dye efflux ability, although not to homogeneity^{210,211}. Based on the categorisation presented by Dykstra et al.²⁰⁶, the myeloid biased HSCs are called α -HSCs and are defined with a myeloid to lymphoid ratio, measured 16 weeks post-transplant, of >2 . Lymphoid biased HSCs, with finite self-renewal, are called γ -HSCs or δ -HSCs, with myeloid to lymphoid ratios of <0.25 , the difference being δ -HSCs have no detectable myeloid outputs. The balanced HSCs lie in-between with ratios between 0.25 and 2^{65,155}. This categorisation will also be adopted in the interpretation of single cell transplantation experiments appearing later in this thesis.

Because CD45 is not expressed in erythrocytes and megakaryocytes, these two lineages have not been traditionally analysed in the context of HSC lineage biases. Benveniste et al. had previously used glucose phosphate isomerase 1 isoforms to track erythrocyte output²¹². However, the system distinguishes the two isoforms by electrophoresis, which is much less convenient than traditional flow cytometric methods. Recently the Kusabira Orange transgenic mouse²¹³ and *Vwf*-tdTomato/*Gata1*-eGFP double reporter mouse²¹⁴ were developed to overcome this limitation, allowing for flow cytometric analysis of 5 blood lineage outputs.

Collectively, these mouse models bolstered evidence of a myeloid-restricted stem cell (MySC) population, consistent with highly biased α -HSCs. However, a platelet restricted population of HSCs (termed P-restricted HSCs) was also discovered that had exclusive and sustained platelet reconstitution above 0.1% even up to 44 weeks¹⁸⁹. Upon secondary transplantation, these platelet-restricted HSCs had sustained high levels of platelet production, though some had low levels of myeloid and lymphoid lineage reconstitution⁶⁹. It is important to note, however, detecting even a low amount of repopulation is made easier because of how numerous platelets are in the blood. P-restricted HSCs isolated from primary and secondary recipients were capable of generating myeloid and lymphoid cells *in vitro*, suggesting they were bona fide multipotent and self-renewing HSCs²¹⁴. Interestingly it has been recently suggested that these distinct HSC subtypes are located in different niches within the BM, suggesting a level of extrinsic regulation not previously appreciated¹⁰⁹.

In line with previous reports of lineage coupling, one study identified several classes of progenitors: myeloid-restricted repopulating progenitor (MyRPs) cells that only generated platelets, known as megakaryocytes repopulating progenitors (MkRPs); platelet-erythrocytes, known as megakaryocyte-erythrocytes repopulating progenitors (MERPs); and platelet-erythrocyte-granulocyte-macrophages, known as common myeloid repopulating progenitors (CMRPs)²¹⁵. All of these progenitors were able to reconstitute primary recipients up to 6 months but were unable to repopulate secondary recipients.

1.3.2 Molecular heterogeneity

Historically, HSC heterogeneity has been characterised by functional assays, but relatively little information about the molecular mechanisms was coupled to the observed heterogeneity, largely due to an inability to prospectively isolate HSC subtypes. An understanding of the molecular state of HSC subtypes and the mechanisms driving self-renewal and differentiation fate choices would lead to improved clinical outcomes through understanding leukemogenesis and expanding or producing blood cells outside the body.

Bulk gene expression studies

Early knowledge of the molecular programme of HSCs came from early knockout mice studies, which examined the roles of nearly 200 genes (reviewed by Rossi et al.²¹⁶). Several critical self-renewal genes will be described in detail in section 1.4.6 below, while this section focuses on global gene expression studies that have been useful in understanding the HSC-specific molecular state. The first global gene expression analysis of HSCs was published by Phillips et al. where subtractive hybridization was used to screen for FL HSC-specific genes²¹⁷. Soon after, the same method was applied to adult BM HSCs²¹⁸. They identified expression of many previously reported genes, but also novel genes specific to HSCs, including transcription factors, membrane proteins and secreted molecules. Interestingly, there was considerable overlap but also differences between FL-HSCs and BM HSCs genes. In an attempt to find a common stem cell signature, several studies compared the molecular profiles of HSCs with other stem cell populations, such as ESCs or neural stem cells²¹⁹⁻²²¹. However, this approach was not particularly successful, as among the 3 studies, only 1 gene (integrin alpha-6; *Itga6*) was expressed commonly between stem cell populations. That said, these early studies did expand the number of genes known to be associated with HSC molecular signatures.

Using microarray gene expression analysis, Venezia et al. looked at the gene expression profile of adult HSCs before and after 5-FU stimulation, and also compared them to naturally cycling FL-

HSCs²²². What they found was a distinct molecular profile characterising quiescent and activated HSCs. Similar studies have profiled HSCs as well as their various differentiated progeny^{121,184,223}.

Single-cell PCR assays

Due to the limitations of gene expression assays at the time, much of the early work described was completed using bulk HSPC populations. However, because cell fate decisions are made at the single cell level, researchers began pursuing functional and molecular assays at single cell resolution, in order to understand the mechanisms. The earliest single-cell quantification of gene expression in haematopoietic cells was achieved in 1990 by Brady et al.^{224,225}. Many groups have subsequently profiled single cells in the haematopoietic system and found early evidence of heterogeneity in gene expression in seemingly homogeneous cell populations^{224,226–228}. These studies also established the idea of lineage priming, where multipotent cells express lineage-specific genes at low levels^{8,229,230}. However, single cell RT-PCR was still very limited by the small number of genes it could assay.

The combination of RT-PCR with microfluidic-based methods expanded the number of genes assayable and led to many subsequent molecular studies of hundreds of single HSPCs. By looking at the expression of 43 known haematopoietic genes, Glotzbach et al. reported that phenotypically defined LT-HSCs, which were putatively homogeneous, actually contained subpopulations with different transcriptional fingerprints²³¹. Consistent with this, Moignard et al. looked at the expression of 18 key haematopoietic transcription factors in 596 single primary HSPC and also found considerable heterogeneity²³². Interestingly some genes were observed to exhibit a bimodal expression, allowing for states of high, medium or no expression. Several transcriptional regulatory networks were discovered, where certain haematopoietic genes are positively or negatively correlated with each other – indicating that different transcriptional programmes are adopted by different cells. Examples include inhibitory relationships between transcription factors PU.1 with GATA1²³³, and GFI1 with GFI1B^{234,235}. In a later study, Schütte et al. validated these transcriptional programmes and showed that the gene expression patterns were stable over time²³⁶.

Apart from regulatory networks, single-cell gene expression profiling has also led to the identification of novel progenitor populations and new insights with cellular hierarchies^{237,238}. Guo et al. identified surface markers that separate transcriptionally and functionally distinct myeloerythroid progenitor populations and provided the early evidences for megakaryocytic priming in highly enriched phenotypic HSC populations²³⁸.

Additionally, single cell profiling could be used to identify molecular signatures in cell populations. A study by Wilson et al. combining single cell functional assays with gene expression analysis of 48 genes, using flow cytometric index sorting, found a molecular signature associated with the transcriptional overlap between 4 differently defined phenotypic LT-HSC populations (MolO gene signature)¹⁸⁶. Importantly, this molecular signature can be used to predict HSCs within scRNA-seq datasets. Interestingly, a study by Petriv et al. profiling of microRNA across the haematopoietic hierarchy found that microRNA profiles could directly infer cell lineages, suggesting that microRNA expression is tightly regulated across differentiation²³⁹.

Single-cell RNA-sequencing

The advent of transcriptome-wide scRNA-seq was a substantial breakthrough. Compared to RT-PCR methods, which could only quantify the expression of up to 200 genes, scRNA-seq was demonstrated to measure up to 10,000 genes in each single cell, offering an unprecedented resolution to transcriptional states²⁴⁰. As reviewed by Watcham et al.²⁴⁰, there are two predominant methods of scRNA-seq that have their own advantages and disadvantages, thereby influencing the possible conclusions that can be drawn. 1) Droplet-based methods, such as Drop-Seq, InDrops or 10X genomics, provide a high throughput method to sequence more than 100,000 cells but typically only detect between 1000-3000 expressed genes per cell. 2) Plate-based methods, such as Smart-Seq2, CelSeq2 or mcSCRB-seq are lower throughput, usually profiling up to 10,000 cells but can detect over 5000 genes per cell²⁴¹⁻²⁴³.

Alongside these technological advances, the past decade has seen an explosion of scRNA-seq studies of haematopoietic cells. In the blood system, the earliest studies profiled well-defined haematopoietic populations using plate-based methods. The first scRNA-seq profiling of HSCs was reported in the above mentioned paper by Wilson et al., in which the MolO gene signature was validated¹⁸⁶. This was followed by a comprehensive study from Paul et al. analysing myeloid progenitors²⁴⁴. This study identified several novel transcription factors involved in lineage priming of myeloid differentiation, providing a glimpse of the power of the technique. Soon after, Nestorowa et al. published the transcriptome of 1600 phenotypically defined HSPCs, including LT-HSCs alongside 10 other progenitor populations¹⁹⁰. The study saw dramatic expression changes associated with early differentiation, in particular with genes associated with cell cycle and metabolism. In later studies, because droplet-based methods require less defined input populations, they allowed for less biased profiling of cells, ultimately providing a more holistic transcriptional landscape of haematopoiesis that lends itself to more exploration of the transcriptional landscape. One of the first studies was by Zheng et al., who profiled more than

180,000 peripheral blood mononuclear cells and BM mononuclear cells and demonstrated the ability of the technique to identify subpopulations in human acute myeloid leukaemia (AML) samples²⁴⁵. Subsequently, Dahlin et al. profiled 44,802 single HSPCs and defined transcriptomic maps containing entry points to 8 different blood lineages²⁴⁶.

However, because of aforementioned limitations of read depth with droplet-based RNA-sequencing (RNA-seq) technologies, there is a need for such global profiling of HSPC with plate-based methods. A recent paper by Dong et al. presented the single cell transcriptome (using Smart-seq2) of 28 haematopoietic cell types - the most comprehensive transcriptomic reference to-date²⁴⁷. Interestingly, the transcriptomes of HSCs, 7 days after transplantation, was also profiled. The authors found little evidence of substantial HSC expansion, rather most of the transplanted progeny became cells that resembled multipotent progenitors at the transcriptomic level²⁴⁷. However, this could be driven by the expression of cell cycling genes.

The emerging view of scRNA-seq studies is that haematopoiesis is a continuous process, which differs from the classical view containing distinct homogeneous populations. In this model, HSC functional heterogeneity has been linked to transcriptional lineage priming, where immature cells exhibit distinct gene expression patterns that are similar to mature cell types which they bias towards^{6,7,52,248,249}. Accordingly, the transcriptional landscape portrays a continuum of such low-primed HSPCs slowly acquiring uni-lineage specificity, without a major multipotent or bipotent stage. In support of this model, several recent studies have reported that phenotypically defined MPPs actually contain mostly unipotent progenitors of corresponding lineages^{244,250,251}.

There are, however, two related observations that need to be reconciled. One is that distinct functional subpopulations of cells can be isolated easily by the use of cell surface markers; and two is that these cells reside in a presupposed transitory region within single-cell transcriptomic landscapes. As reviewed by Laurenti et al.⁸, a possible explanation is that RNA and protein levels do not correspond well with each other because of additional post transcriptional regulation. However transcriptomic and proteomic studies of bulk HSPCs have suggested that, in general, the correlation is reasonable²⁵². Another explanation could be that scRNA-seq does not capture certain information, such as epigenetic or spatial changes, which can be important in determining cellular states. Additionally, there are limitations with ordering scRNA-seq data along pseudotime, as reviewed by Weinreb et al²⁵³. Detecting jumps in cell states are very difficult with pseudotime

modelling, because it assumes that differentiation is a continuous process, which may self-reinforce such a conclusion²⁴⁰.

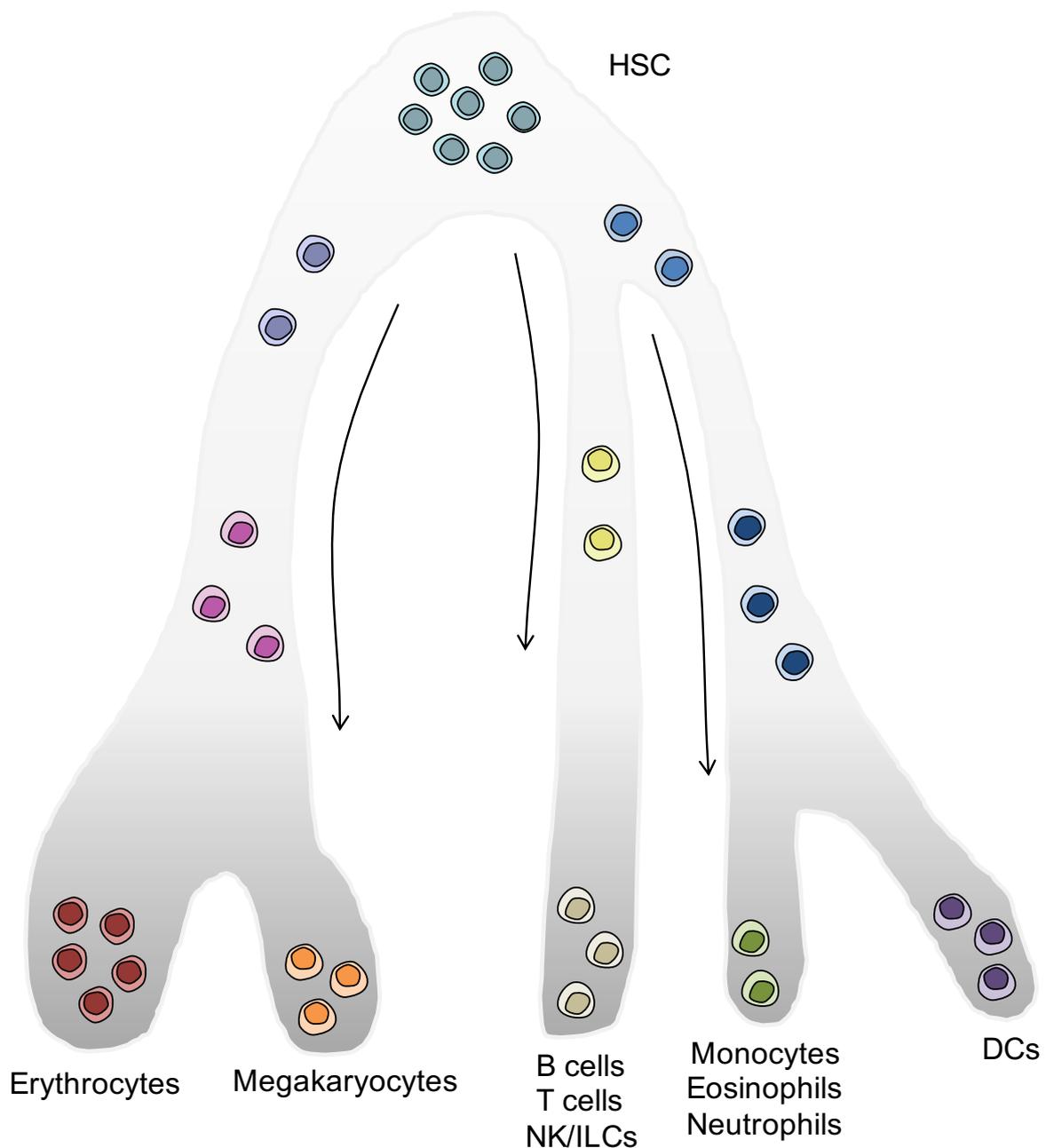


Figure 1.5 The emerging continuum of haematopoietic hierarchy.

The result of recent scRNA-seq studies point towards a continuum of haematopoietic differentiation. HSCs remain at the apex of the hierarchy, but the steps towards their terminal fates are blurred by incremental transitional changes in their transcriptome.

Epigenetics

At least in the embryonic stem cell (ESC) system, it's been shown using single cell ChIPseq that chromatin states can define different subpopulations. A recent paper by Yu et al. found that HSC clonal heterogeneity is associated with different epigenetic patterns²¹¹. Using bisulfite sequencing

and transposase-accessible chromatin sequencing (ATAC-seq) on individual clones, they found that DNA methylation and chromatin accessibility patterns drive differences in HSC function²¹¹. More recently a study by Florian et al. used single-cell 3D confocal imaging, single-cell transplantation, scRNA-seq and single-cell ATAC-seq to show that aged HSCs lose polarity, which is controlled by Cdc42²⁵⁴. Apolar HSCs were shown to preferentially divide symmetrically into daughter cells with less self-renewal capacity, while young polar HSCs tended to divide asymmetrically, maintaining at least one potent daughter HSC²⁵⁴. In particular, the self-renewal potential of daughter HSCs was linked to epigenetic mark H4K16ac and the amount of open chromatin, suggesting an epigenetic mechanism driving stem cell fates. In a 2014 study, Cabezas-Wallscheid et al. integrated proteomic, transcriptomic and epigenetic (DNA methylation) analysis of HSPCs²⁵². They found that during differentiation, certain differentially methylated regions showed continuous gain or loss of methylation. Confirming previous studies, they also found that DNA methylation is inversely correlated with gene expression, suggesting that at the very least this continuous change in methylation pattern is part of the mechanism regulating differentiation²⁵⁵. Combined with the fact that several key HSC genes are epigenetic regulators^{256,257}, it would seem future single cell epigenetic studies will be crucial in revealing whether epigenetics is the true driver behind HSC heterogeneity and differentiation.

Proteome

Because proteomic approaches have traditionally required very large cell numbers (>100,000 cells), and HSCs represent such a rare population of cells, most research has focused on transcriptomes of HSCs. In the same study by Cabezas-Wallscheid et al. above, the authors find a high concordance between transcriptomes and proteomes²⁵². Accordingly, many proteins found to be differentially highly expressed in HSCs have been previously identified at the transcriptional level. However, because proteomic methods don't have a 3' capture bias, there will always be proteins that are excluded by transcriptomic analysis²⁵⁸. Future HSC expansion cultures and increased assay sensitivity for proteins, may improve current proteomic studies. Already the cell numbers required for proteomic studies is decreasing with recent papers reporting using single cells, using tandem mass tag (TMT) labelling^{259,260}. However, the number of proteins identified remained low (~1600 proteins), and the cell types used had high total protein content.

To reliably achieve single cell protein quantifications, low throughput methods such as CyTOF²⁶¹ or single-cell western blot²⁶² are still available. Though both methods require antibodies and known targets, reducing the potential new discoveries. Nonetheless, Knapp et al. had used CyTOF to quantify the changes in up to 43 markers, including surface markers, transcription factors and signalling intermediates, upon stimulation with haematopoietic growth factors²⁶³.

1.3.3 Haematopoietic hierarchy revised

As mentioned above, recent molecular studies have challenged the classical hierarchical differentiation tree of haematopoietic progenitors. As reviewed by Laurenti et al.⁸, cellular barcoding and lineage tracing studies have challenged the contribution of HSCs to everyday haematopoiesis, by showing that that longevity of progenitor cell contributions are vastly underestimated due to their inability to engraft. These studies have suggested that under non-transplantation settings, HSCs contribute minimally to native haematopoiesis and instead the majority of blood homeostasis is said to be maintained by MPPs^{182,229,230,252,264,265}. The exception to this was the megakaryocytes, half of which appear to be derived directly from the HSC compartment still^{210,266}. This also contradicts the classical view of the haematopoietic hierarchy because it suggests that HSCs can generate lineage-restricted progenitors directly, bypassing multipotent progenitors.

The more that is discovered about the molecular mechanisms underpinning haematopoiesis and HSC self-renewal, the closer we will be to achieving regenerative therapies. The next section turns to the clinical implications of HSC biology and one of its longstanding goals of *ex vivo* HSC expansion.

1.4 HSC expansion

1.4.1 Clinical significance

HSCs are ideal candidates for the treatment of haematological disorders because of their durable self-renewal and multipotency, allowing them to replenish fully a patient's blood system²⁶⁷. In fact, the first successful HSC transplantation (HSCT) was performed more than 60 years ago, and it remains the only curative therapy for a number of haematological malignancies and disorders^{15,210}. HSCT can be either allogeneic or autologous. Allogeneic HSCT is more common and utilizes healthy donor HSCs, that are human leukocyte antigen (HLA) matched. Autologous HSCT utilises the patient's own HSCs to prevent BM aplasia after high dose chemotherapy or radiotherapy or as a histocompatible source of cells for gene therapy¹⁹. The outcome of the procedure has improved considerably over the years, mainly due to better HLA matching, however, the availability of HLA matched donors is limited, with no more than 60% of patients finding a suitable donor and approximately 37,000 people worldwide still on the waiting list²⁶⁸. Combined with the fact that the probability of success for the procedure is largely correlated with the dose of stem

cells injected ($3\text{-}4 \times 10^6$ CD34⁺ cells/kg of human body weight required)²⁶⁹, there is an immense clinical need to expand HSC numbers *ex vivo*.

Additionally, HSCT can be combined with gene therapy strategies to cure congenital haematological illnesses. In gene therapy, targeting HSCs is crucial because it ensures durable lifetime supply of gene-corrected blood cell progenies after transplantation²⁷⁰. Gene therapy has been successfully demonstrated to treat diseases such as Wiskott–Aldrich syndrome²⁷¹, leukodystrophy²⁷² and recently sickle cell anaemia²⁷³. Clinical trials for T cell-corrected immunotherapy to treat leukaemia are also hopeful²⁷⁰. However, current gene therapy protocols require HSPCs to be cultured *ex vivo* for 2-4 days, while being exposed to retroviral or lentiviral gene correction vectors, therefore the loss of self-renewal potential during culture causes a significant decrease in the longevity of the graft²⁷⁴. Current approaches address this problem by using a large amount of input cells, to increase the chances of having durable HSCs included in the final product²⁷⁰. Thus, *ex vivo* expansion or efficient maintenance of HSCs can significantly improve the gene therapy product delivered to patients.

Apart from the direct transplantation as a curative therapy, it is also possible to use HSCs to generate functional mature blood cells *in vitro* which, when scaled up, could potentially meet the transfusion demands of patients. The cell types that have been derived successfully include red blood cells^{275,276}, megakaryocytes and platelets^{181,271,277}, neutrophils^{278,279} and T cells^{280,281}. However most of these studies have not yet produced large enough yields to replace current sources²⁸². Furthermore, because of limited HSC sources, these studies have largely utilise ESCs or induced pluripotent stem cells (iPSCs), which have their own set of limitations (discussed in section 1.4.3 below). Therefore, having a readily available and expanding source of HSCs *in vitro* could be very useful to scale up production of blood cells.

Because of this immense clinical potential, supply of HSCs often fails to meet the demand of patients and researchers. However, despite tremendous efforts, there have been very few breakthroughs¹⁹. Only very recently, has robust *ex vivo* expansion of murine HSCs in long-term cultures been achieved (detailed in section 1.4.9 below)²⁸³. The next section summarises the vast literature on HSC self-renewal *in vivo* and expansion cultures *ex vivo*, and also why expansion of HSCs has been particularly difficult.

1.4.2 Difficulties and challenges

There have been two main barriers that have hampered progress regarding HSC expansion – the lack of robust markers for functional stem cells *in vitro* and the lack of understanding of the molecular pathways that support expanding adult HSCs.

As mentioned in section 1.2.5, the gold standard for the functional verification of HSCs still requires serial transplantations. This is extremely time consuming and expensive, especially considering the number of potential self-renewal factors, factor combinations and different concentrations that need to be tested. As a result, many studies utilise *in vitro* assays, such as with phenotypic markers or CFC assays, to assess HSC expansion. However, while surface marker isolation strategies can isolate HSCs to about 50% purity *in vivo*, many surface markers for isolating fresh HSCs have been reported to be unreliable markers of HSCs in culture²⁸⁴. For example, Tie-2 loses its expression in culture, and endoglin and Mpl are only expressed in a portion of cultured HSCs²⁸⁴. Similarly, CD49f and CD38 expression changes in cultured human HSCs¹⁹. As a result, most groups still use LSK as a readout of phenotypic HSCs *in vitro*. Recently, two papers by Fares et al. and Tomellini et al. respectively demonstrated that EPCR and integrin- α 3 could be potentially reliable markers for human HSCs in culture^{285,286}. However, the purity is still below 5% in both cases and it is unclear whether the two markers work for cultured murine HSCs.

Furthermore, although the HSC state is now increasingly well characterised *in vivo*, there is very little known about the molecular state of expanding adult HSCs after they have been cultured *ex vivo*. This would be an important undertaking in order to add to our current understanding of the molecular states of HSCs and potentially lead to new surface markers for HSCs expanding *ex vivo*, or the identification of novel molecular pathways that could be targeted to increase expansion efficiency.

1.4.3 HSCs derived from induced pluripotent stem cells or reprogramming

To circumvent these challenges, groups have attempted to differentiate ES and iPSCs into HSCs and mature blood cells *in vitro* in an efficient manner. There are still two main limitations with using iPSCs to derive blood cells for clinical purposes. Traditionally, iPSCs are generated using genetic over-expression of the reprogramming factors (e.g., *Oct4*, *Klf4*, *Sox2* and *Myc*), which poses a translational barrier due to the risk associated with genetic insertions via viral vectors²⁶⁷. Recent advances have allowed iPSCs to be generated without viral transduction, using soluble factors; however, the reprogramming efficiency using these methods remains very low. The first major efforts have therefore focused on generating platelets due to their lack of DNA and safer clinical

profile in terms of cancer risks. Secondly, whilst strategies to differentiate iPSCs have been successful in generating aforementioned platelets and other haematopoietic cells such as erythrocytes¹⁷⁶, T cells²⁷¹, NK cells²⁸⁷ and macrophages²⁸⁸, it has been challenging to generate bona fide HSCs that are capable of serial and multi-lineage reconstitution upon transplantation. Recently adult mouse endothelial cells were reprogrammed into HSCs via ectopic expression of four transcription factors (*Fosb*, *Gfi1*, *Runx1* and *Spi1*)²⁸⁹, however methods without genetic manipulation are still unavailable. As a result, a majority of groups have focused efforts on expanding primary HSCs via stimulating self-renewal programmes *in vitro*.

1.4.4 Harnessing endogenous HSC expansion

Unlike FL HSCs, which expand massively in numbers during development⁴⁷, the number of adult HSCs in mouse BM stays relatively constant, around 1 in every 30,000 (0.003%)²⁸⁸. However, when they are transplanted as single cells, they can increase in numbers by approximately 1000-fold²⁹⁰. Interestingly, when transplanting a bulk population of HSCs, HSC expansion was inversely proportional to the initial number of HSCs transplanted and plateaus at a certain point²⁹¹. This finding was supported by transplantation data of cultured HSCs from a later study by Sekulovic et al.²⁹², suggesting that there may be negative feedback mechanisms to limit self-renewal – perhaps as a defence mechanism against cancer. HSCs can then be taken from the initially reconstituted mice and transplanted again into new irradiated recipients to achieve a further expansion. At least four successive transplantations have been done and these indicate that HSCs can undergo at least 8400-fold increase cumulatively over the initial number of input HSCs^{70,293}. A caveat is that self-renewal activity is generally reduced after each successive transplantation, and self-renewal ability is eventually lost after 6 serial transplantations⁶⁹. This demonstrates that although adult HSCs seem to have a limited self-renewal capacity in this transplantation setting, dictated both by intrinsic and extrinsic mechanisms, they are still fundamentally capable of expanding substantially beyond what they would be expected to during the course of an organism's natural lifespan.

1.4.5 HSC fate choice

There are many different fate choices that an HSC can make. This is not to imply that HSCs have agency, but rather there are many different possible outcomes for an HSC, and it must “decide”, based on an array of cell-extrinsic or cell-intrinsic factors, which fate it “chooses”. When considering how to expand HSCs, it is imperative to understand these different decisions summarised in Figure 1.6 below. The balance between proliferation and quiescence, survival and cell death, and self-renewal and differentiation all contribute to the number of HSCs and specialised cells in the system – a crucial part of maintaining life-long haematopoiesis⁵. For

example, too much differentiation leads to HSC pool exhaustion, but too much self-renewal can lead to oncogenic transformation via a differentiation block. These decisions can be co-ordinated by intrinsic molecular regulators (e.g., transcription factors, epigenetic modulators or microRNAs) or extrinsic factors that provide appropriate signals to activate intrinsic pathways^{19,267}. Therefore, in order to expand HSCs, it becomes important to understand - and subsequently maximise - survival, proliferation and self-renewal. In certain cases, cell fate can also be coupled through a gradual restriction of cell potential. The next section specifically focuses on the molecular regulation of self-renewal.

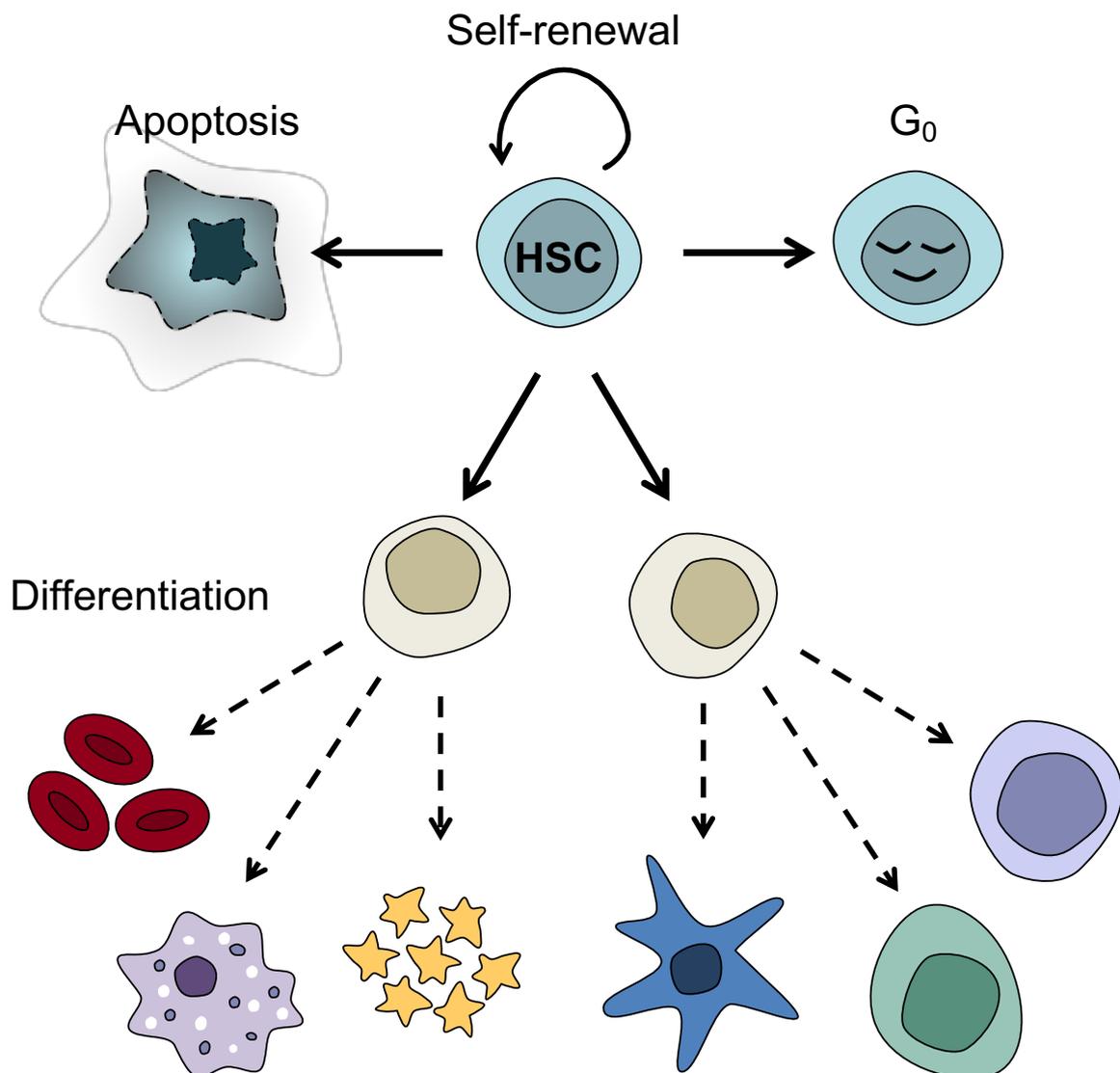


Figure 1.6 HSC fate choice possibilities.

Conceptually, there are only a finite number of outcomes that can occur to an individual HSC. Most commonly, it stays in quiescence or G₀. However, when activated, it can either divide asymmetrically to produce one daughter HSC and one differentiated progeny; it can also divide symmetrically to produce two daughter HSCs or two differentiated progenies. Finally, it can undergo apoptosis. In order to expand HSCs *ex vivo*, symmetric self-renewal divisions must be achieved in large numbers.

1.4.6 HSC intrinsic self-renewal pathways

Molecular regulators of self-renewal have been extensively explored over the last three decades in an effort to inform *in vitro* expansion efforts. Through genetic manipulation studies, molecules critical for HSC self-renewal have been identified, mostly in FL and adult BM HSCs^{19,267,294}. This list includes a wide array of transcription factors, epigenetic modifiers, cell cycle regulators and microRNAs¹⁹. The three major regulators are summarised below.

Transcription factors

Transcription factors, resulting from their ability to mediate expression of target genes and activate self-renewal programmes, are important regulators of HSC self-renew. A myriad of transcription factors have been shown to be essential for HSC self-renewal, including β -catenin⁵³, MYC²⁹⁵, CEBP α ²⁹⁶, HOXB4¹³¹, GATA2^{297–299}, PU.1³⁰⁰, RUNX1³⁰¹, JUNB^{302,303}, MEIS1³⁰⁴, PBX1³⁰⁵, GFI1⁵⁹, EVI1¹⁹⁵ and more³⁰⁶. Interestingly, as detailed in section 1.2.2, foetal (actively cycling) and adult HSCs (quiescent) appear to have different transcription factor programmes including SOX17 and Hmga2^{41,61}, which have been shown to be important for FL HSCs, but not for adult HSCs^{55,307}.

Epigenetic modifiers

The process of differentiation requires changes in the epigenetic programme of cells. Therefore, it is not surprising that many epigenetic modifiers are implicated in HSC self-renewal²⁹⁹. Interestingly, DNA methyltransferases (e.g. DNMT3a^{257,308} and DNMT3b³⁰⁹) and also TET-enzymes (e.g. TET2)^{310,311}, which are involved in DNA methylation, are both implicated in self-renewal. As mentioned, DNMT3a and TET2 are also commonly mutated in clonal haematopoiesis in humans, suggestive of its role in regulating self-renewal in human HSCs⁶³. Chromatin remodelers such as the Polycomb group family of proteins also harbour HSC self-renewal regulators such as EZH2, BMI1³¹², Mel-18³¹³ and CBX7³¹⁴. Loss of function mutation in *Bmi1* leads to defective self-renewal and accelerated differentiation while overexpression increases self-renewal in HSCs³¹⁵.

MicroRNAs

MicroRNAs have also been implicated in self-renewal regulation in HSCs³¹⁶. miR125 was shown to increase HSC cell numbers by 8 fold³¹⁷ and overexpression of miR29a causes myeloid progenitors to reacquire self-renewal properties^{318,319}.

Interestingly, transcription factors, epigenetic regulators and microRNAs can interact with each other to form self-renewal pathways. A notable example in mouse foetal HSCs involves the aforementioned RNA-binding protein LIN28B, the microRNA *let7*, and high motility group AT-hook protein 2 (HMGIC, more commonly known as the gene product of *Hmga2*)⁶¹. In FL HSCs, LIN28B binds to *let7* and prevents it from inhibiting *Hmga2* expression⁶¹.

1.4.7 HSC extrinsic self-renewal regulators

The intrinsic self-renewal regulators mentioned above are regulated by a number of extrinsic factors. Several reviews have highlighted several of these upstream signalling pathways including inflammation^{320–322}, hypoxia^{141–143,323}, Wnt signalling^{267,295,324} and Notch signalling^{325–327}. There are less studied extrinsic influences on HSC self-renewal such as diet^{328–331} and mechanosensing^{332–334} of the ECM. In the next section the role of immune cytokines, probably the most well studied extrinsic regulators of self-renewal, are highlighted as they are particularly relevant in this thesis.

Immune cytokines

Haematopoietic cytokines have been widely studied because of their important role in blood lineage development. Some cytokines directly push differentiation towards specific lineages, such as G-CSF and Macrophage-colony stimulating factor (M-CSF)³³⁵; Others have been implicated in self-renewal³³⁶. Stem cell factor (SCF) was first discovered as the ligand and activator of KIT, a receptor tyrosine kinase expressed by all HSCs³³⁷. The importance in HSC function for this tandem of proteins has been shown in mutational studies, where mutations in KIT cause defects in HSC numbers^{336,338} and mutations in SCF causes defects in the microenvironment supporting HSCs³³⁹. Further evidence for the importance of SCF and KIT signalling comes from a recent study of sprout-related, EVH1 domain-containing protein 1 (SPRED1)³⁴⁰. SPRED1 is a negative regulator of KIT signalling and SPRED1-deficient HSCs show increased self-renewal potential³⁴⁰. Interestingly, foetal and adult HSCs seem to have different sensitivity to SCF¹³⁰.

Thrombopoietin (TPO) is the main cytokine regulating megakaryocyte and platelet development. Interestingly, it has also been shown to affect HSC self-renewal³⁴¹. Knock out mice studies have shown that TPO-null mice have decreased numbers of repopulating HSCs³⁴². Similarly, genetic deletion of *Mpl* (the receptor for TPO) also reduces HSC self-renewal potential³⁴³. Additionally, genetic perturbation of LNK, which is a negative regulator of TPO signalling, has been shown to increase HSC-self-renewal³⁴⁴. LNK mainly acts through negatively regulating JAK2, which is a receptor tyrosine kinase downstream of many different cytokines, including TPO, IL-6 and IL-11³⁴⁵. A single LNK-deficient HSC can expand approximately 3000-fold after transplantation²⁹⁰.

GP130 receptor signalling has been known to be important for HSC self-renewal, because GP130 null mice have reduced numbers of HSCs³⁴⁶. A variety of cytokines are known to stimulate GP130, including IL-11, IL-6, Leukaemia inhibitory factor and Oncostatin M³⁴¹. Although GP130 is ubiquitously expressed, its activation depends on the binding to alpha receptors that are specific to each of these cytokines. In this way, GP130 is able to exert cytokine specific effects³⁴¹. Of these,

IL-6-deficient mice have HSCs with reduced self-renewal capability³⁴⁷, and IL-11 overexpressing HSCs have increased self-renewal capacity *in vivo*³⁴⁸.

TGF- β has been shown to induce HSC quiescence³⁴⁹. Freshly isolated HSCs have active TGF- β signalling and TGF- β receptor deficient HSCs have impaired reconstitution ability and increased proliferative activity³⁵⁰. The same study also demonstrated that non-myelinating Schwann cells were responsible for the activating of TGF- β in the BM, thus inducing HSC quiescence¹³⁷. Interestingly the effects of TGF- β seems to be HSC subtype dependent²⁶⁵.

1.4.8 *Ex vivo* expansion of HSCs

Cytokine combinations

As mentioned, cytokines are crucial to HSC function. Experimentally, they also have many advantages including their convenience of use, and the fact that they exhibit their effects reversibly, without the need for permanent DNA manipulations. As a result, early work in developing HSC *ex vivo* cultures have focused heavily on optimising cytokine conditions. In the 1990s and early 2000s, an array of studies focused on the *in vitro* effects of SCF, activators of GP130, TPO and Flt-3 ligand (Flt3L), the latter of which was demonstrated to be dispensable for murine HSC self-renewal^{1337,341,349,351–353}. Miller et al. demonstrated that HSCs can be maintained serum-free *ex vivo* for 10 days when cultured in combination with SCF, Flt3L and IL-11³⁵³. Interestingly SCF in combination with Flt3L alone was shown to be sufficient to stimulate HSCs to survive and proliferate in culture. However, retention of stem cell activity in culture was shown to further require additional activation of gp130 pathway¹³⁰. In mice, this is often achieved by addition of IL-11³⁴¹, whereas in humans IL-6 is typically used³⁵⁴.

In order to systematically optimise and assess the effects of these 4 distinct cytokines, Audet et al. performed a two-level factorial analysis, by testing every possible factor combination at two different concentrations respectively³³⁷. They found that SCF and IL-11 were the most potent stimulators of HSC expansion. Additionally, they found that IL-11 has an optimal concentration of 20ng/mL with any higher concentrations being detrimental to HSC expansion, whereas SCF has no effective maximum concentration. Based on this and the study by Miller et al.³⁵³, the group concluded that TPO offers no beneficial effect to stem cell expansion. However, other groups have published differing results that suggest that TPO is crucial for HSC self-renewal *in vitro* and many groups continue to use TPO alongside SCF^{342,343,349}. An explanation of the differences could be that the other media components, including base media and media supplements, have differing

effects on cytokine cocktails. Very recently, Wilkinson et al. reported that low SCF (10ng/mL) combined with high TPO (100ng/mL) achieves the best expansion results²⁸³, although any combinations of SCF and TPO above 50ng/mL also performed well (details in section 1.4.9). The precise optimal concentration of cytokines remains to be determined; however it is clear from these decades of studies that the expansion of functional HSCs occurs largely independently of effects on proliferation, and cell number becomes a poor readout for HSC expansion. In fact, cytokine conditions that stimulate rapid proliferation and yield high cell numbers often do not have the most functional HSCs³⁵⁵.

Until recently, even the best expansion attempts with haematopoietic cytokines have only achieved maintenance for a week or two at most¹³⁶, suggesting that cytokines are limited in their capacity to maintain long-term self-renewal. As a result, numerous other strategies were developed in attempts to expand HSCs *ex vivo*, and these are discussed below.

Transgene overexpression

The Hox family of transcription factors were one of the first group of genes implicated in HSC self-renewal^{294,302}. Thus, they were also the first genes used in overexpression studies to expand HSCs *ex vivo*. Overexpression of *Hoxb4* resulted in a 40-fold increase in HSCs with full multilineage potential²⁹⁷. HOXB4 can also be transiently delivered into the cell as a TAT-HOXB4 fusion protein which was shown to expand HSCs by at least 4-fold²⁹⁸. Other examples include *Fbxw7*³⁵⁶ and *Dppa5*³⁵⁷, which provide 2-fold and 6-fold expansion respectively when overexpressed.

Structural chromosomal rearrangements of the nucleoporin 98 gene (NUP98) is commonly associated with haematopoietic malignancies such as AML³⁵⁸. It was discovered that the NUP98-HOXA9 fusion gene product can induce AML in mice³⁵⁸. Subsequently, the fusion protein was also found to increase *in vitro* self-renewal and proliferation of human HSCs³⁵⁹ as well as *in vitro* self-renewal divisions in mouse HSCs²⁹².

In a comprehensive study, Deneault et al. performed a gain-of-function screen on 104 nuclear factors, of which 18 had an effect on HSC activity³⁶⁰. Overexpression of 10 of them, *Smarcc1*, *Vps72*, *Fos*, *Trim27*, *Sox4*, *Klf10*, *Ski*, *Prdm16*, *Erdr1*, and *Sfp1*, were found to be equivalent or better than *Hoxb4* at expanding HSCs. These studies were arduous and represented a large body of work, and although they were important in elucidating many various HSC self-renewal regulators, there is still a missing link with our understanding of upstream signal transduction that upregulates these

self-renewal genes. As a result, HSC expansion without genetic modification was still elusive for a very long time.

Supportive co-cultures

Inspired by the concept of the HSC niche, many studies have cultured HSCs with various feeder cells derived from different sources that naturally support HSC expansion such as the AGM¹⁰³, urogenital ridge (UG)³⁶¹, FL⁴⁵ and BM³⁶². In particular cell lines, such as AGM-S3, AFT024, UG26-1B6 (UG26) and EL08-1D2 (EL08), have been shown to support the survival and maintenance of HSCs for at least 6 weeks in culture³⁶³. Notably, Oostendorp et al. demonstrated that the supportive effect of UG26 and EL08 does not require direct cell to cell contact with the HSCs, suggesting that secreted factors were sufficient^{361,364,365}. Several studies have also suggested that mesenchymal stem and progenitor cells can support HSC activity in co-cultures^{120,366}.

To provide a clinically useful source of expanded HSCs, introducing genetic changes that expand HSC numbers could have unknown additional consequences (including development of cancer), and co-cultures would introduce another biological product that would require significant screening and biosafety regulations. As a result, feeder-free and serum-free methods have now superseded these earlier attempts.

Soluble factors

Many groups have reported various soluble factors that can expand HSCs when supplemented with cytokines, summarised by several excellent reviews^{267,367,368}. The notable conditions are summarised in Table 1.

By examining differentially expressed proteins between supportive cell lines, such as UG26 and EL08, and other non-supportive cell lines, Buckley et al. identified WNT5A as a key protein capable of maintaining HSCs in culture⁴⁵. Similar strategies have identified several soluble factors that can expand HSCs, such as insulin growth factor binding protein (IGFBP)-2^{325,368-370}, angiopoietin-like proteins (Angptl)³⁷¹, PTN¹³² and nerve growth factor (NGF)¹³³, with reports of expansion of up to 48-fold. Insulin-like growth factor (IGF)-2 was isolated from supportive CD3⁺Ter119⁻ FL cells¹⁵³ and subsequently, Angptl3 was also found to have potent HSC expansion ability³⁶². Culturing with Angptl3 and IGF-2 supplemented with SCF, TPO and FGF-1 resulted in an 30-fold increase in HSCs¹³³. Adding IGFBP-2, found to be expressed by a tumorigenic cell line, to such a cocktail was reported to expand HSCs by 48-fold after 21 days of culture^{133,284}. In support of this, deletion of activating transcription factor 4(ATF4) was shown to be detrimental for foetal

HSCs and the effect seems to be dependent on its role in inducing *Angptl3* expression in endothelial and stromal cells within the FL³⁷². From primary human brain endothelial cells (HUBECs) that support expansion of human HSCs, Himburg et al. identified PTN as a potent regulator of HSC expansion¹³². Mouse HSCs cultured with PTN for 7 days display a 4-fold increase in CRUs¹³². Wohrer et al. identified NGF and Collagen 1 (Col 1), amongst other factors secreted by UG26, that were able to expand HSCs in culture by 4-fold when supplemented with SCF and IL-11¹⁵³. However, upon secondary transplants, it was shown that NGF and Col 1 were only able to maintain the number of input HSCs. This highlights the importance of secondary transplants to assess true durability in self-renewal.

Many other factors and strategies have also been investigated in the context of HSC expansion. As mentioned, Wnt signalling has been implicated in the extrinsic regulation of self-renewal. GSK-3b is a well-known b-catenin inhibitor, and thus an inhibitor of canonical Wnt signalling³⁷³. GSK-3b inhibitors such as CHIR99021 have been shown to expand murine HSCs when combined with rapamycin, an inhibitor of mammalian target of rapamycin (mTOR)³⁷³. Interestingly, Wnt signalling through b-catenin seems to be dispensable, and over-activation even detrimental *in vivo*^{374–377}. Nuclear receptors have also been investigated for roles in HSC expansion. Stimulation of the retinoic acid receptor seems to be beneficial for mouse HSCs^{198,378}, but not for human HSCs^{379,380}. Another strategy to expand HSCs involves targeting epigenetic machinery such as by inhibiting histone acetylation or deacetylation^{381–384}. Valproic acid (VPA), which is a histone deacetylase inhibitor, was shown to increase HSC expansion in both mouse and human^{382–384}. Recently, low calcium has also been reported to improve HSC maintenance *ex vivo*³⁸⁵. However, this is contradicted by another report describing high intracellular calcium to be associated with quiescence and increased repopulation ability²⁴⁸.

In recent years, there have been several molecules that seem very promising with human HSCs expansions. In unbiased screens, StemRegenin1 (SR1) and UM171 have been identified to increase human LT-HSCs by 2 and 13-fold respectively, measured by 20-week primary engraftment after 7 days of culture^{386,387}. Notably, the expansion increases to 30-fold when both molecules are used in combination. While SR1 was shown to antagonize the aryl hydrocarbon receptor (AHR), UM171-mediated expansion was shown to be AHR independent. Interestingly both molecules are ineffective at expanding mouse HSCs³⁸⁶, and their detailed mechanisms have yet to be elucidated. Prostaglandin E2 (PGE2) is another promising molecule and has been shown to increase multi-lineage engraftment of cord blood cells^{388,389}. Interestingly PGE2 has been shown to interact with

the Wnt signalling pathway by influencing B-catenin breakdown³⁹⁰. In support of Wnt signalling playing an important role in human HSC expansion, short-term cultures of human cord blood (CB) CD34⁺ cells with 6-bromoindirubin 3'-oxime (BIO), a GSK-3b inhibitor, were shown to enhance engraftment potential³⁹¹. Notably, Notch stimulation has also been used to expand human cord blood CD34⁺ cells^{327,392}. Targeting the upregulation of glycolytic pathways, Guo et al. used peroxisome proliferator-activated receptor- γ antagonist, GW9662, to achieve 4-fold expansion of human CB HSC³⁹³. Similar to aforementioned VPA, Garcinol is a plant-derived histone acetyltransferase (HAT) inhibitor that was shown to expand human HSCs by 2.5-fold³⁸¹.

Bioengineering the HSC niche

It is conceivable that, despite decades of research, the lack of success in HSC *ex vivo* expansion is in part due to the failure of traditional liquid cultures or stromal co-cultures, to satisfy a 3-dimensional aspect of the stem cell niche. In recent years, there have been attempts to bioengineer a mimic of HSC niches, in order to improve expansion efficiencies. Tiwari et al. used MS-5 stromal cells to produce extracellular matrices *in vitro*, which are then used as a scaffold for culturing CD34 enriched human cord blood cells³⁹⁴. They demonstrated that the acellular scaffolds increased phenotypic HSCs and CFUs by 80-fold. Following this, several other studies have reported other ways of producing artificial niches, sometimes functionalised with molecules such as CXCL12, to improve *ex vivo* culture^{395–398}. Accompanying this line of research, several studies have also investigated the role of ECM proteins such as collagen, fibronectin, dystroglycan, heparin sulfate, proteoglycans, osteopontin and laminins^{115,116,399–401}. Some studies have even suggested that substrate elasticity can influence HSC self-renewal^{332,333}.

Table 1.3 Summary of expansion protocols, modified and updated from Walasek et al.³⁸⁴.

Molecule	Species	Input cells	Supplements	Culture time	Assay	Fold change		Ref
						Over control	Over input	
FGFs	M	BM		3w	HSC frequency		-	402
IGF-2	M	BM SP CD45 ⁺ Sca-1 ⁺	SCF FGF-1 TPO	10d	CRU (HSC frequency)		7.8	362
IGFBP-2	M	BM SP Sca-1 ⁺	FGF SCF TPO Angplt-3	21d	CRU (HSC frequency)		48	371,403
Angplts (2,3,5,7)	M	SP CD45 ⁺ Sca-1 ⁺	SCF TPO FGF-1 IGF-2	10d	CRU (HSC frequency)		30	133
Pleiotrophin	M	LSK CD34 ⁻	SCF Flt3L TPO	7d	CRU frequency (12 weeks p-tpx)	6	4	132
					% engraftment secondary tpx	10	10	
	H	Lin-34 ⁺ 38 ⁻ UCB	SCF Flt3L TPO	% engraftment (4 weeks)	3	3		
ATRA	M	LSK	Serum SCF Flt3L IL-6 IL-11	7d	Donor reconstitution (per 10 ⁵ cells)	5		378
TEPA	H	CD34 ⁺ CD133 ⁺ UCB	TPO Flt3L IL-6		% engraftment			404
VPA	M	LSK	SCF Flt3L TPO IL3	14d	% engraftment	2.2		382
	H	UCB HSC	SCF Flt3L TPO IL3	14d	SRC frequency	6.0		
Chlamydocin	H	CD34 ⁺ MPB	SCF Flt3L TPO	24h	% engraftment (SRC)	4.0	2.0	405
5aza and TSA	H	CD34 ⁺ UCB	Serum SCF IL-3 Flt3L MGDF	9d	SRC frequency	40	9	406
BIO	H	CD34 ⁺ UBC	Serum SCF Flt3L TPO	5d	Engraftment of expanded cells	2	0	391
					% engraftment	0		407
			Serum SCF Flt3L TPO	5d	SRC frequency (6 weeks)	2.5		
SR1	H	CD34 ⁺ MPB UBC	SCF Flt3L TPO IL-6	3–5w	SRC frequency	14	17	408
					Secondary SRC frequency	4	12	
PGE2	M	WBM		2h	CFU-S (d12)	3		409
	H	LSK			HSC frequency (CRU)	2-4		

Continued...						Fold change		
Molecule	Species	Input cells	Supplements	Culture time	Assay	Over control	Over input	Ref
NR-101	H	CD34 ⁺ 38 ⁻ UCB		7d	SRC frequency	2.3	2.9	410
Immobilised Delta1	M	LSK	Serum SCF IL-6 IL-11 Flt3L IL7	28-35d	HSC frequency (CRU)			392
	H	CD34 ⁺ CD38 ⁻ UCB	SCF Flt3L IL-6 TPO IL-3 LLP	21d	% engraftment (SRC)	5.3		411
Wnt3a	M	LSK Thy1 ^{Lo}	Serum	6d	% engraftment CRU	5		412
Wnt5a	M	Lin- BM cells		3d	% engraftment CRU	-	-	45
Ssh protein	H	CD34 ⁺ CD38- Lin-	SCF, G-CSF Flt3-L IL-3 IL-6	7d	% engraftment (SRC)	3		413
BMP4	H	CD34 ⁺ CD39-Lin-	SCF Flt3L G-CSF IL-3 IL-6	6d	% engraftment (SRC)	-		414
UM171	H	CD34 ⁺	SCF Flt3L TPO	12d	SRC frequency		13	386
NGF	M	ESLAM	Col1 SCF IL-11	7d	HSC frequency		4	153
					Secondary HSC frequency		0	
GW9662	H	UCB CD34 ⁺	SCF Flt3L TPO	4d	SRC frequency	5	4	393

FGFs, fibroblast growth factors; IGF-2, insulin growth factors 2; Angplts, angiopoietin-like proteins; ATRA, all-*trans* retinoic acid; TEPA, tetra-ethylenepentamine; VPA, valproic acid; TSA, trichostatin A; 5aza, 5-aza-2'-deoxycytidine; BIO, 6-bromoindirubin-3'-oxime; SR1, StemRegenin1; PGE₂, prostaglandin E₂; Ssh, sonic hedgehog; BMP4, Bone morphogenic protein; NGF, Nerve growth factor; Col1, Collagen 1; M, mouse; H, human; BM, bone marrow; SP, side population; UCB, umbilical cord blood; MPB, mobilized peripheral blood; SCF, stem cell factor; TPO, thrombopoietin; Flt3L, Flt3 ligand; IL, interleukin; LLP, low density lipoprotein; MGDF, megakaryocyte growth and development factor; w, week; d, day; CFU(-S), colony forming cell (spleen); CRU, competitive repopulating unit; SRC, NOD/SCID-repopulating cell;

Negative feedback control

HSCs and their progeny secrete many proteins that signal to surrounding cells⁴¹⁵. As mentioned, some of these molecules positively regulate HSC function and some of them are detrimental. A potential way to improve HSC expansion efficiency is to limit the negative feedback that occurs within HSC cultures. To address this, Csaszar et al. developed a way to control the amount of inhibitory feedback signalling, which they termed “fed-batch” cultures⁴¹⁶. In this system, medium is constantly fed into the culture, diluting any negative regulators that are secreted. Leading to an 11-fold increase in functional HSCs after 12 days of culture. The paper also measured and detailed the accumulation of several known HSC regulators during the course of the culture, such as TGF- β , IL-6, CCL2, CCL3 and CCL4⁴¹⁶. These factors were also later confirmed to be present in mouse HSC cultures²⁸³. Both studies used low-throughput antibody-based detection platforms, thus only monitoring a limited number of pre-selected secreted factors. Future global secretome analysis of HSC cultures could be highly beneficial to determining the positive and negative regulators of HSC expansion secreted by HSCs or their progeny.

1.4.9 Current state of the *ex vivo* expansion field

Despite the substantial expansion capacities of the various strategies described above, relatively few of these approaches have been widely adopted (or verified) by researchers in subsequent studies. This is particularly noticeable when compared to cytokines such as TPO and SCF which have stood the test of time and continue to be included in most expansion combinations. Several possible explanations for this exist, including the fact that many culture conditions are not fully defined, containing serum or bovine serum albumin (BSA) that may have profound batch effects⁴¹⁷. Although many of them included functional transplants, they needed to have secondary transplants and long enough time points to ensure that the donor cells did not come from a durable MPPs. Furthermore, purity of input cells was not high for most of the studies, which could cause indirect secondary effects through contaminating cells. Although not all studies suffer from these potential issues, as observed in chapter 4, where highly purified LT-HSCs were used in serum/feeder free conditions, very few of these putative expansion factors have a significant effect on expanding HSCs. Even the study by Wohrer et al¹⁵³, that used highly enriched HSCs and performed the appropriate secondary transplantation assays, was not reproducible in our lab⁴¹⁸.

A major breakthrough was recently achieved when Wilkinson et al. published a paper describing a completely defined, serum-free and albumin-free culture condition that allows 200-900 fold increases in functional HSCs over a period of a month²⁸³. As detailed in chapter 4 and 5, these

findings were reproducible. Interestingly, the system does not require the addition of any extra soluble factors apart from SCF and TPO, and it achieves expansion by replacing commonly used BSA with poly-vinyl alcohol (PVA). In a previous report, the same authors demonstrated that BSA could be replaced by recombinant human serum albumin (HSA), to increase HSC maintenance *in vitro*⁴¹⁷. In this study, they showed that contaminants from the production of recombinant HSA cause an increased inflammatory response that induces differentiation, and by replacing it with fully synthetic PVA, they were able to reduce this response. Although no other soluble factors were added to the culture, the authors did show that the use of fibronectin coated plates significantly increased expansion. However, it is still unclear whether fibronectin helps with better cell adhesion and retention or by active fibronectin signalling pathways. Notably, the system also relies on frequent and complete media changes, in order to minimise the effects of negative regulators secreted endogenously. Interestingly, clonal cultures under this PVA culture condition displayed substantial heterogeneity in terms of differences in repopulation potential as well as lineage biases after transplantation. Future studies that elucidate the molecular mechanisms of heterogeneity would be important to further improve culture yields and clinical outcomes.

1.5 Thesis aims

The field has lacked robust markers for functional HSCs *in vitro* and this has hampered progress in a wide array of investigations. Recent developments in HSC reporter mice provide an exciting opportunity to study new, mostly intracellular markers that were previously inaccessible.

HSC *ex vivo* expansion has been a longstanding goal in the field, yet only very recently has there been successful long-term cultures that are capable of robustly expanding HSCs. However, still very little is known about the heterogeneity within and between clonal cultures.

The primary aim of this PhD thesis was to develop a novel reporter system to mark functional HSCs *in vitro*. Using this novel reporter system, the second objective was to screen various known and novel factors for their ability to expand HSCs in *ex vivo* cultures. Finally, the reporter system was applied to recently published PVA expansion cultures in order to identify HSC-containing clones and to molecularly profile the HSCs and non-HSCs from positive and negative single cell-derived cultures in an effort to determine the molecular regulators of expanding HSCs.

2 Methods

2.1 Mice

Fgd5^{ZsGreen-ZsGreen/+} knock in/knock out mice were purchased from Jackson Laboratories and wild-type (WT) mice were either *Fgd5*^{+/+} litter mates or CD45.2 C57BL/6 under 1 year of age. All transplantation recipients were C57BL/6^{W41/W41}-Ly5.1 (W41). All mice were kept in microisolator cages in Central Biomedical Service animal facility of Cambridge University and York University, and provided continuously with sterile food, water, and bedding. All mice were kept in specified pathogen-free conditions, and all procedures performed according to the United Kingdom Home Office regulations, in accordance with the Animal Scientific Procedure Act.

2.1.1 Genotyping

Ear biopsies were taken from individual mice and placed in individual Eppendorf tubes. To extract the DNA, 50µL alkaline lysis buffer (24.4mM NaOH (Sigma Aldrich, St. Louis, MO, USA (Sigma)) and 195µM disodium EDTA (Life Technologies, Carlsbad, CA, USA) in water) was added and the tubes were put on a heated shaker at 95°C, 1000rpm for 20 minutes (min). 50µL of neutralizing buffer (0.04M Tris-HCL (Sigma) in water) was added and samples were kept in 4°C until PCR amplification.

Briefly, per sample, 12.5µL of Kapa2G mastermix (Kapa biosystems), 10.9µL of nuclease free water and 0.6µL of 10µM of primers (WT or Mutant, Table 2.1) were added. After amplification, PCR products were run on 1% agarose gel (Biorad) with 1:10000 GelRed Nucleic Acid Gel Stain (Biotum). A representative image of a result with a WT, *Fgd5*^{ZsGreen-ZsGreen/+}, Water control and *Fgd5*^{ZsGreen-ZsGreen/+} control is shown in Figure 2.1.

Table 2.1 Primer sequences for Fgd5 genotyping.

Primer	Sequence
<i>Fgd5</i> MUT reverse	GCG GTT GCC GTA CAT GAA G
<i>Fgd5</i> WT reverse	ATG ACC TCA TTG GGG AAG G
<i>Fgd5</i> forward	GGA AGC TCC AGA TGA AGA GG

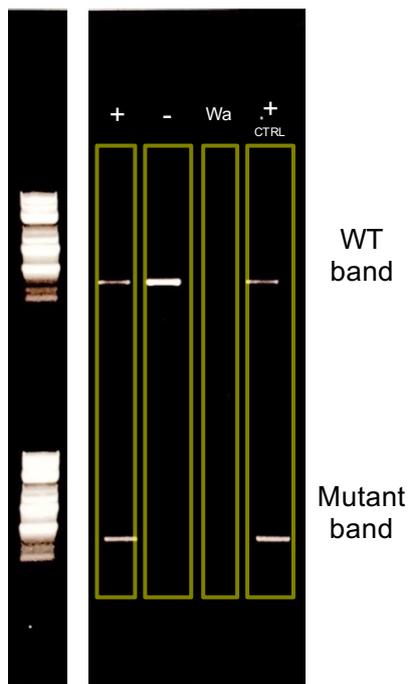


Figure 2.1 Representative image of *Fgd5*^{ZsGreen-ZsGreen}/₊ genotyping experiment.

Double mutants are embryonic lethal, hence positive mice are heterogenous with both WT and mutant bands. 100kb ladder on the left. Wa, water control; + CTRL, Positive control.

2.2 HSC isolation or BM analysis by flow cytometry

2.2.1 Bone marrow harvest

Mice were sacrificed by dislocation of the neck. BM cells were isolated from the tibia, femur and sternum of both hind legs, by crushing the bone in PBS2% - 2% Foetal Calf Serum (FCS (Sigma) or STEMCELL Technologies (SCT)) in PBS (Phosphate-buffered saline, Sigma). Samples were filtered through 20 μ m sterile filters before further processing.

2.2.2 Red cell lysis

Red cell lysis was performed using ammonium chloride (NH₄Cl, SCT). The cells were first concentrated by spinning down and pelleting at 300g for 5 min. The supernatant was removed carefully, and the cells were resuspended in 3mL PBS2%. 5mL of NH₄Cl was added and cells were incubated for 5 min on ice. After a short vortex to ensure pellet is fully resuspended, cells were again incubated for another 5 min on ice. The cells were washed with 12mL of PBS2% (spun down at 300g for 5 min, supernatant removed) and then resuspended in 500 μ l and transferred to a FACS tube for lineage depletion.

2.2.3 Lineage depletion

HSPC were enriched using EasySep Mouse Hematopoietic Progenitor Cell Enrichment Kit (SCT). Briefly, the 500 μ L of cell suspension was incubated with 10 μ L of EasySep Mouse Hematopoietic Progenitor Cell Isolation cocktail for 15 min on ice. 20 μ L of EasySep Streptavidin Rapid Spheres 50001 was added for a further 15 min incubation on ice. 2mL of PBS2% was added and the tube was placed in the EasySep magnet for an incubation of 3 min at room temperature. Carefully whilst still in the magnet, the supernatant was poured out into a new tube and the magnetic step was repeated another time.

2.2.4 Isolating HSCs by FACS

ESLAM Sca-1⁺ cells were isolated as described previously¹⁸⁴ using CD45 BV421, EPCR PE, CD150 PE/Cy7, CD48 APC, Sca-1 BV605 (Table 2.2) and 7-Aminoactinomycin D (7AAD) (Life Technologies). The cells were sorted on an Influx (BD) using the following filter sets 530/40 (for *Fgd5*), 585/29 (for PE), 670/30 (for APC), 460/50 (for BV421), 670/30 (for 7AAD) and 610/20 (for BV605) (Table 2.4). When single HSCs were required, the single-cell deposition unit of the sorter was used to place 1 cell into the wells of 96-well plates, each well having been preloaded with 50 μ L or 100 μ L medium (described below).

In all experiments the ESLAM Sca-1⁺ gating strategy (Figure 2.2) was used to sort HSCs, otherwise where indicated in the text, *Fgd5*⁺EPCR⁺CD45⁺ cells were sorted instead of ESLAM Sca-1⁺ cells.

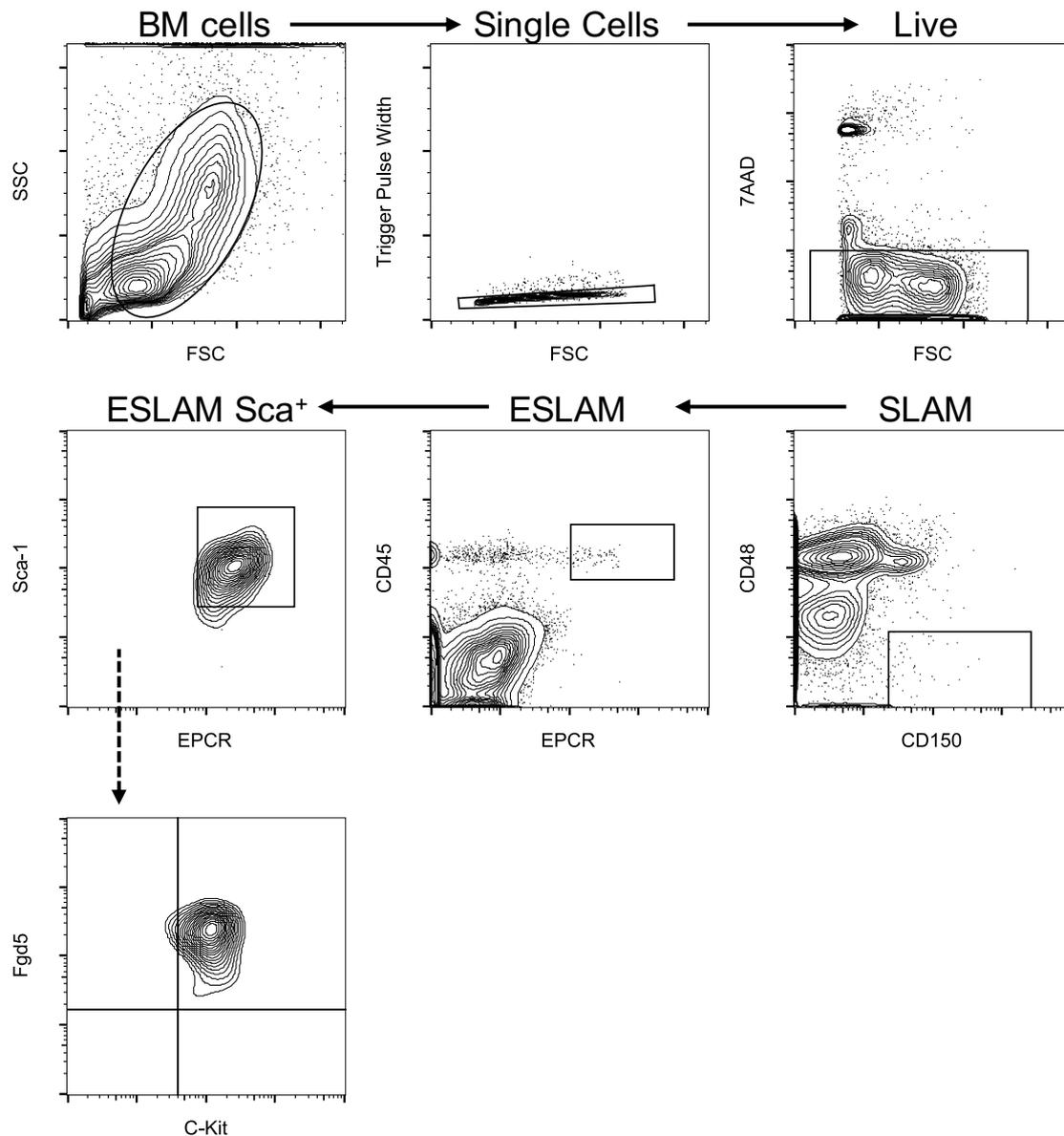


Figure 2.2 Representative ESLAM Sca-1⁺ gating strategy.

EPCR⁺ CD45⁺ CD150⁺ CD48⁻ Sca-1⁺ HSCs are isolated from lineage depleted BM by fluorescence assisted cell sorting. Although not a part of the standard gating strategy, these cells are mostly *Fgd5* and c-Kit positive, as indicated by the dotted line.

2.2.5 Lymphoid and peripheral tissue harvest

Cell suspensions of spleen, mesenteric lymph nodes and thymus were obtained by passing the tissues through a 70 μ m strainer. Lung and liver samples were digested with 750U/mL Collagenase I (Life Technologies) and 0.3mg/mL DNase I (Sigma). Following digestion, liver cells were then passed through a 70 μ m strainer and mononuclear cells were isolated using a 33% gradient of Percoll (Merck) at 690g for 12 min with minimal break. Spleen, liver, thymus and lung cell suspensions were then treated with NH₄Cl (SCT) to remove red blood cells.

2.2.6 Flow cytometric analysis or FACS isolation of *Fgd5*⁺EPCR⁻ cells

Cell suspensions of BM cells were prepared as above. For sorting of FE⁻ cells, cell suspensions were stained with EPCR PE, CD45 BV421, CD5 PE/Cy7, 7AAD and/or CD244 AF647 and/or NK1.1 APC/Cy7 and/or Sca-1 BV605 (Table 2.2). For surface marker phenotyping, the additional antibodies are listed in Table 2.2.

2.3 *In vitro* assays

2.3.1 Cell culture images

Cell culture images were taken with an Olympus CKX41 microscope and an Olympus SC50 camera.

2.3.2 CFU assays

100 or 200 cells were resuspended in 600 μ L of semi-solid MethoCult GF M3434 (SCT) and plated into two wells in a 6-well SmartDish (SCT). Cells were then cultured for 14 days at 37°C with 5% CO₂. Colonies were imaged and counted using STEMvision (SCT).

2.3.3 OP9 assays

OP9 cells were plated into tissue culture-treated 96-well flat bottom plates (Corning) at 2000 cells per well in 100 μ L of Opti-MEM supplemented with 10% FCS, 1% Penicillin/Streptomycin (Sigma) and 0.2% BME. Single HSCs or FE⁻ cells were sorted a day later into each well. 100 μ L of media (of above supplements) was immediately added to give final concentrations of 25ng/mL Flt3L, 25ng/mL SCF and 25ng/mL IL-7. Cells were cultured for 14 days with half media change after 7 days. At the end of the experiment, clones were harvested and analysed by flow cytometry for B220 APC, CD19 PE/Cy7, NK1.1 APC/Cy7, CD45 BV421, Mac-1 BV605 (Table 2.2) and 7AAD (Life Technologies). Cells were considered clonogenic if CD45⁺7AAD⁻ cells are present.

2.3.4 Intracellular flow cytometry

BM suspensions were prepared as above (section 2.2). Cells were stained with EPCR PE, CD45 BV421, CD5 BV711 (Table 2.2) and 7AAD (Life Technologies). FE-CD5⁺ and FE-CD244⁺ cells were sorted separately for stimulation. To stimulate the cells, they were cultured in 96-well advanced RPMI supplemented with 10% FCS, 1% Penicillin/Streptomycin (Sigma), 1% L-Glutamine (L-Glut, Sigma) with 0.2% β -Mercaptoethanol (BME, Life technologies), 500ng/mL Phorbol 12-myristate 13-acetate (PMA, Sigma) and 500ng/mL Ionomycin (Sigma) for 4 hours at 37°C with 5% CO₂. 0.6 μ L of GolgiStop (BD) was also added to each well. As unstimulated control,

cells were cultured as above without PMA and Ionomycin. Cells were stained with Zombie Aqua (Biolegend) before being fixed with 2% formaldehyde (PFA, Thermo Fisher Scientific (Thermo)) and permeabilised with FoxP3 perm buffer (eBiosciences). Cells were stained with IFN γ APC/Cy7, IL-5 APC and TNF α PerCP/cy5.5 (Table 2.2). Flow cytometry was performed on an LSRFortessa (BD) and all data were analysed using FlowJo v10 (Treestar, Ashland, OR, USA).

2.3.5 Stemspan (SS) based HSC cultures

Single or bulk HSCs were cultured in 96 well U-bottom plates (Corning) containing 100 μ L of StemSpan Serum-Free Expansion Medium (SS, SCT) supplemented with 1% Penicillin/Streptomycin (Sigma), 1% L-Glut (Sigma), 0.2% BME (Life technologies), 300ng/mL of mouse SCF (SCT or Bio-Techne) and 20ng/mL human IL-11 (SCT or Bio-Techne) at 37°C with 5% CO₂. All SS-based cultures are performed serum-free, apart from the 6-day liquid cultures in Section 3.1.2, where the above media makeup is supplemented with 10% FCS (Sigma). Where detailed in the text, indicated concentrations of OPN (Sigma), PTN (Bio-Techne), IGFBP2 (Bio-Techne) and IGFBP4 (Bio-Techne) were added to this media makeup.

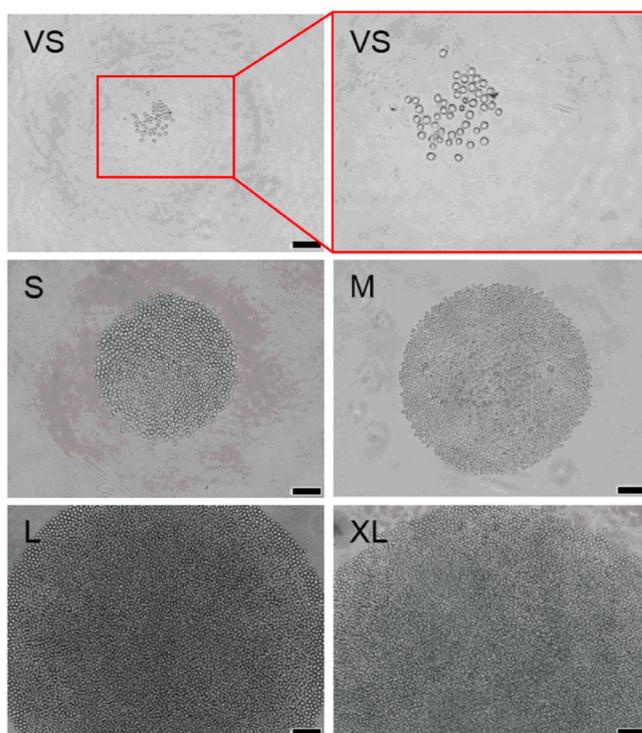


Figure 2.3 Representative images of different liquid culture colony sizes.

Single HSCs are cultured *in vitro* for 10 days. On day 10, clones are assessed for survival and clone size and then harvested for flow cytometric analysis. Scale bar represents 100 μ m. Figure adapted from Caroline Oedekoven⁴¹⁸.

2.3.6 F12-based 28-day HSC cultures

F12-based cultures performed as described previously⁴¹⁹. Briefly single or bulk HSCs were cultured on BioCoat fibronectin 96 well plates (Corning) in 200 μ L of Ham's F12 nutrient mix (Thermo) supplemented with 1% Insulin-Transferrin-Selenium-Ethanolamine (ITSX, Gibco), 10mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, Gibco), 1% Penicillin/Streptomycin/L-Glutamate (P/S/G, Gibco), 100ng/mL mouse TPO (Preprotech), 10ng/mL mouse SCF (Peprotech) and 0.1% PVA (Sigma) or HSA (Albumin Bioscience) at 37°C with 5% CO₂. Where indicated, 20ng/mL of human IL-11 (SCT or Bio-Techne) was also used. Complete medium changes were made every 2-3 days after the first 5-6 days.

Where indicated, 10% of the cultures were taken out for flow cytometric analysis detailed below.

2.3.7 F12-based short-term (<10 days) cultures

For short-term cultures up to 10-days, cells were cultured as above (section 2.3.6), except 96 well U-bottom plates (Corning) were used and no media changes were performed. Where indicated, varying concentrations of SCF and TPO were used.

2.3.8 EL08 conditioned media preparation

EL08 CM was prepared as described previously¹⁵³. Briefly, EL08 cells were cultured in sterile tissue culture flasks (Greiner Bio-One Ltd (Greiner)) in α -MEM supplemented with 10% FCS (sigma), 10% horse serum (Sigma), 1% Penicillin/Streptomycin (Sigma) and 1% L-Glut (Sigma) until 90% confluent. Flasks were then irradiated at 30 Gy and then washed twice carefully with PBS (Sigma) and SS would be added. After indicated duration of incubation at 37°C with 5% CO₂, CM is harvested and filtered through 0.2 μ m sterile filters and kept at -20°C until further use.

2.3.9 EL08 CM cultures

As above, single HSCs were cultured in 96 well U-bottom plates (Corning) containing 100 μ L of EL08 CM supplemented with 1% Penicillin/Streptomycin (Sigma), 1% L-Glut (Sigma), 0.2% BME, 300ng/mL of mouse SCF (SCT or Bio-Techne) and 20ng/mL human IL-11 (SCT or Bio-Techne) at 37°C with 5% CO₂.

2.3.10 Flow cytometric analysis of *in vitro* cultures

Cells (cultured from bulk or single clones) were stained with EPCR PE, Sca-1 BV605, Mac-1 APC, Gr-1 PE/Cy7, c-Kit APC/Cy7, CD45 BV421 (Table 2.2) and 7AAD (Life Technologies). To enumerate cells, a defined number of fluorescent beads (Trucount Control Beads, BD) were added to each well and each sample was back calculated to the proportion of the total that were run

through the cytometer. Flow cytometry was performed on an LSRFortessa (BD) with a High Throughput Sampler (BD) (for single clone analysis) and all data were analysed using FlowJo v10. Representative gates for FELSK ($Fgd5^+EPCR^+Lineage\ Sca-1^+c-Kit^+$) cells are shown in Figure 2.4.

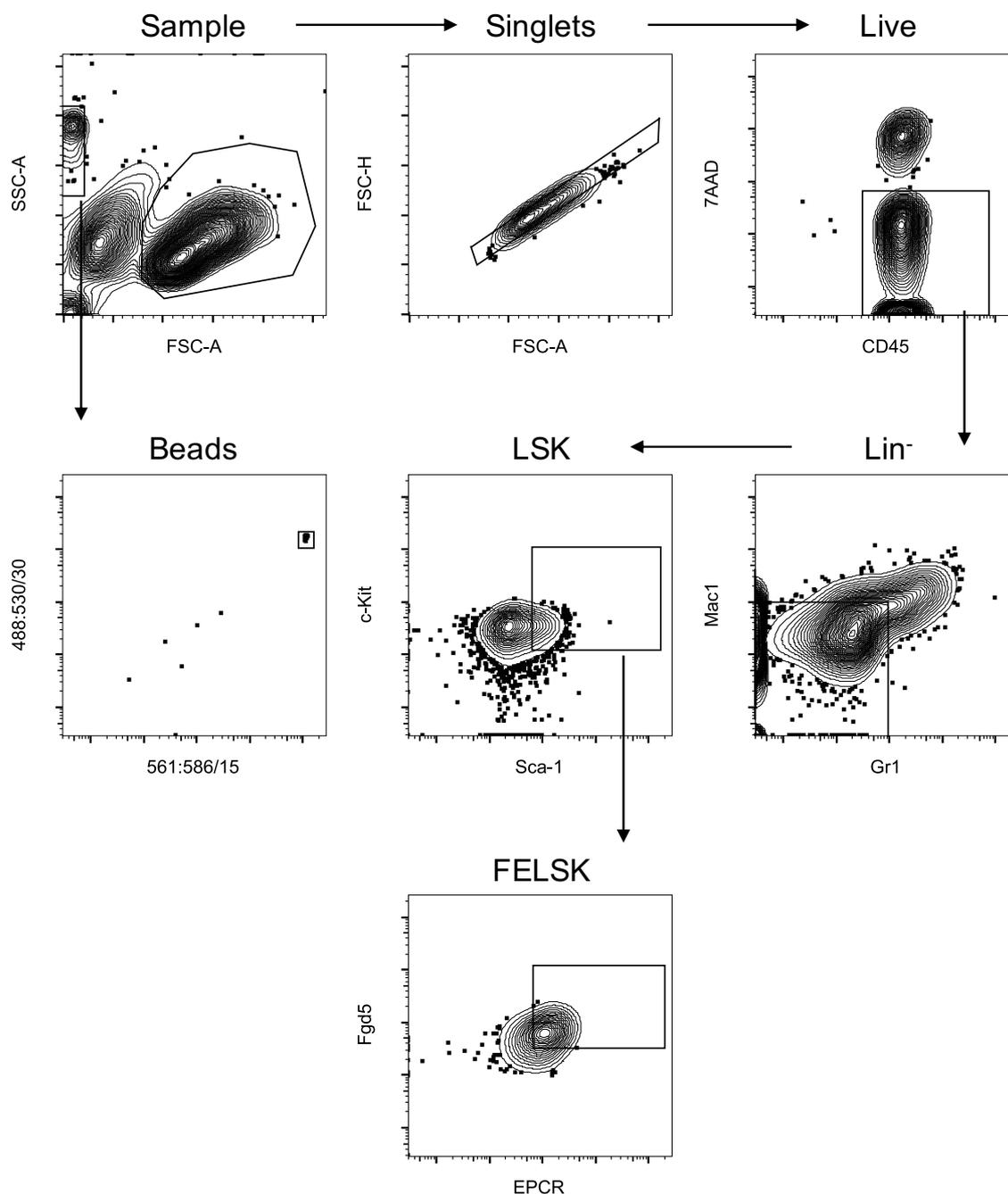


Figure 2.4 Representative gating strategy for FELSK cells.

To analyse single cell clones, the percentage of phenotypic HSCs (FELSK) is plotted against the number of live cells in each clone (Figure 2.5).

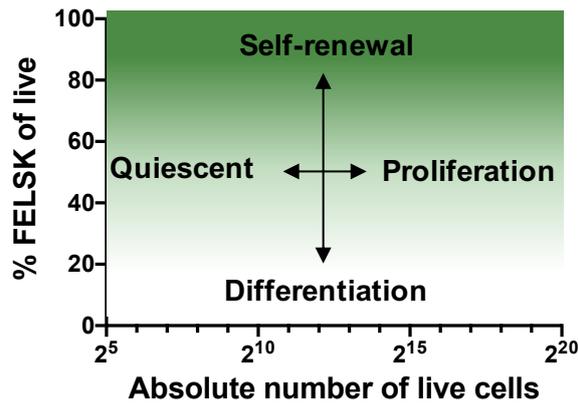


Figure 2.5 Schematic of clonal analysis interpretation.

Single clones are analysed by flow cytometry. A defined number of fluorescent count beads are used to back-calculate the total number of cells in each clone. By plotting against the percentage of phenotypic HSCs defined as *Fgd5*⁺EPCR⁺LSK (FELSK) cells. Culture conditions can be compared for their ability to induce proliferation and maintain HSC self-renewal. Because technical variations may occur, a control condition (SS or F12) is always used in each experiment.

2.4 Transplantation assays

All donor cells were obtained from *Fgd5*^{ZsGreen-ZsGreen/+} mice that were between 8 and 16 weeks of age. All recipient mice were W41 mice that were at least 8 weeks of age. Recipient mice given a sublethal dose of radiation (4 Gy) using a caesium source. After sorting, cells were diluted to desired cell doses with PBS. If needed, cells were spun down to ensure that they were suspended in the final transplantation volume of 200-300 μ L. Transplantations were performed by intravenous tail vein injection using a 29.5G insulin syringe (Terumo).

2.4.1 Single cell transplantations

Single-cell transplantations were performed by tail vein injection of sublethally irradiated W41 mice as previously described²⁰⁶. Briefly, single cells were sorted into 100 μ L of medium in a 96-well U-bottom plate and spun down at 300g for 1 minute. Wells were visually inspected for the presence of a cell. To the wells containing a cell, 100 μ L of PBS was mixed with the well and all liquid was subsequently aspirated into the insulin syringe. After removal of air bubbles, all liquid was injected through the tail vein.

2.4.2 Secondary transplantations

For secondary transplantations, BM from the primary recipient was harvested from the femur tibia and hips by flushing with PBS2%, as above. Red cell lysis was performed as above, and cells were immediately frozen in FCS with 10% DMSO (Thermo) or transplanted into W41 mice as

recipients. Each recipient mice received an equivalent cell dose of one femur. Secondary mice were monitored for at least 16 weeks.

2.4.3 Peripheral blood analysis

Peripheral blood samples were collected from the tail vein of all mice every 4 weeks up until 16 weeks post transplantation. Blood was collected in EDTA coated microvette tubes (Sarstedt AG & Co, Nuembrecht, Germany). Red blood cell lysis was performed as before. Samples were stained with surface marker antibodies for CD45.1 AF700, CD45.2 APC/Cy7, Mac-1 BV605, B220 APC, Ly6G PE/Cy7 and CD3 PE (Table 2.2). As before, 7AAD (Life Technologies) was used as a viability dye and samples were analysed on a BD LSR Fortessa flow cytometer. Representative gating strategies are shown in Figure 2.6. Briefly, after gating for viable and singlets, donor and recipient cells were distinguished by their expression of CD45.1 or CD45.2. B cells are defined as B220⁺ cells, T cells are defined as CD3⁺ cells and GM cells are defined as Mac-1⁺ cells. Animals with at least 1% donor WBCs at 16 after transplantation were considered to be repopulated with long-term reconstituting cells. Animals with at least 1% contribution to all three lineages are considered multi-lineage.

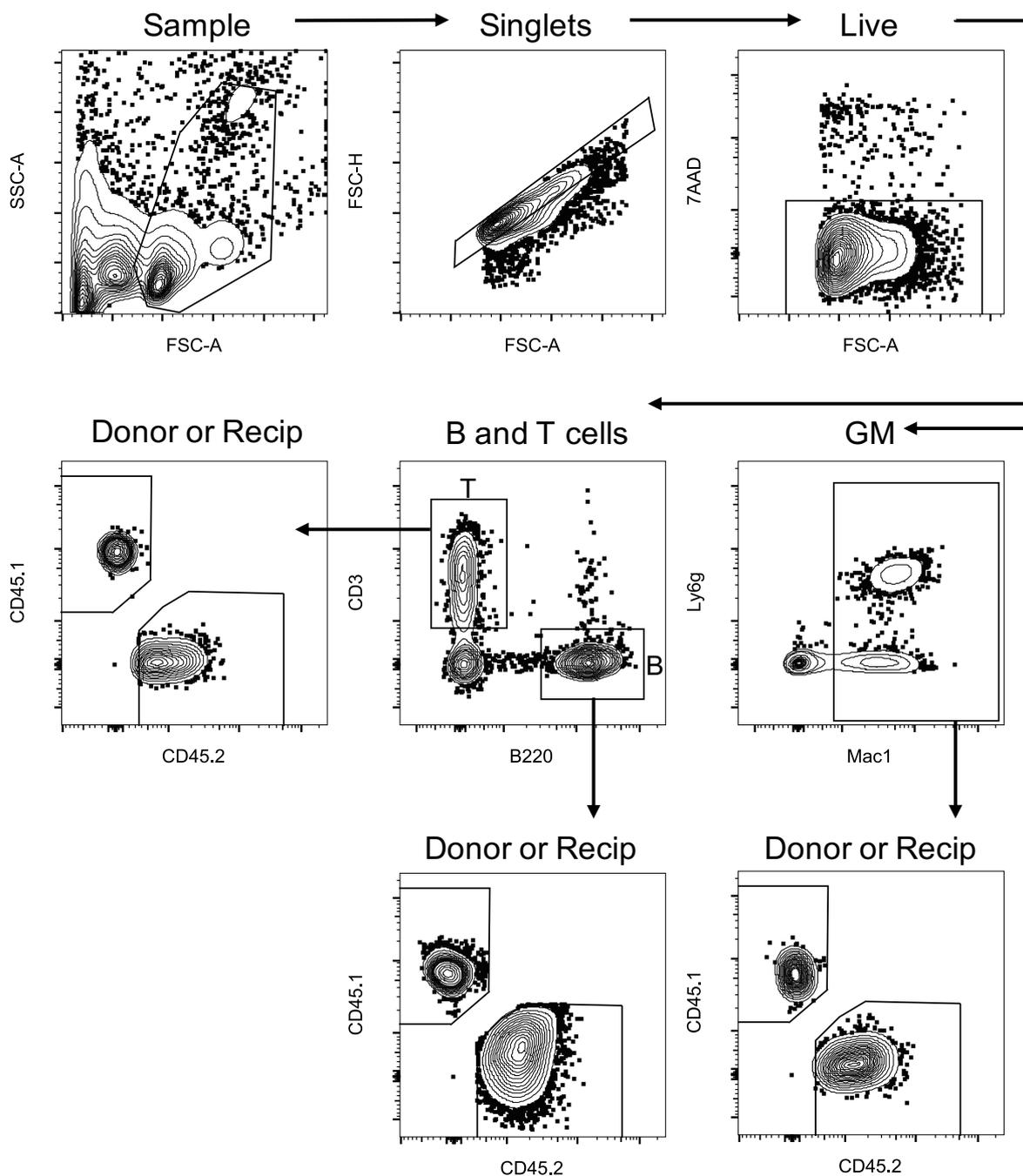


Figure 2.6 Representative gating layout for peripheral blood chimerism analysis.

2.5 Antibodies

Table 2.2 List of antibodies used, with respective clone numbers and manufacturer.

Antigen	Fluorochrome	Clone	Manufacturer
B220	APC	RA3-6B2	Biolegend
CD150	PE/Cy7	TC15-12F12.2	Biolegend
CD19	PE/Cy7	6D5	Biolegend
CD1d-PBS57 or unloaded	APC	-	NIH Tetramer Core
CD244.2	AF647	m2B4 (B6)458.1	Biolegend
CD25	PerCP/Cy5.5	PC61	Biolegend
CD3	PE or APC/Cy7	17A2	Biolegend
CD335	PE/Cy7	29A1.4	Biolegend
CD4	BV605	RM4-5	BD
CD45	PB or BV421	30-F11	Biolegend
CD45.1	AF700	A20	eBioscience
CD45.2	APC/Cy7	104	Biolegend
CD48	APC	HM48-1	Biolegend
CD49b	BV510	HMa2	BD
CD5	PE/Cy7 or BV711	53-7.3	Biolegend
CD5	PE/Cy7	53-7.3	Biolegend
CD8a	BV510	53-6.7	BD
c-Kit	APC/Cy7	2B6	Biolegend
EPCR	PE	RMEPCR1560	SCT
IFN- γ	APC/Cy7	XMG1.2	Biolegend
IL-5	APC	TRFK5	Biolegend
IL-7 α	BV785	A7R34	Biolegend
Ly6C (Gr-1)	PE/Cy7	RB6-8C5	Biolegend
Ly6G	PB/BV421	1A8	Biolegend
Mac-1	BV605	M1/70	Biolegend
Mouse HSPC cocktail	Biotin	145-2C11; M1/70; 6D5; RA3-6B2; RB6-8C5; TER-119	SCT
Ncam	APC	MEM188	Abcam
NK1.1	APC/Cy7	PK136	Biolegend
PD1	BV510	29F.1A12	Biolegend
Sca-1	BV605 or BV421	D7	Biolegend
Streptavidin	BV510 or BV605	-	Biolegend
TCR- β	BV785	H57-597	Biolegend
Ter119	PE/Cy7	Ter119	Biolegend
TNF- α	PerCP/Cy5.5	MP6-XT22	Biolegend
$\gamma\delta$ -TCR	PE/Cy7	GL3	Biolegend

2.6 Flow cytometer set up

Table 2.3 Flow cytometer set up for BD LSR Fortessa.

Fluorochrome	DF/Band-pass filter
BV421 or PB	405 450/50
BV510	405 525/50
BV605	405 610/20
ZsGreen	488 515/20
PE	532 586/15
7AAD	532 710/50
PE/Cy7	532 780/60
APC	640 670/14
AF700	640 730/45
APC/Cy7	640 780/60

Table 2.4 Flow cytometer set up for BD Influx.

Fluorochrome	DF/Band-pass filter
BV421/PB	405 460/50
BV605	405 650LP
ZsGreen	488 530/40
PE	561 585/29
7AAD	561 670/30
PE/Cy7	561 750LP
APC	640 670/30

2.7 RNA sequencing

RNA was extracted using the Picopure RNA Isolation Kit (Thermo) according to manufacturer protocol. Bulk RNA-seq was carried out at the Genomics core facility of the Cambridge Stem Cell Institute upon submission of sample RNA. Briefly, libraries were prepared using the SMARTer Stranded Total RNA-seq Kit v2 – Pico Input mammalian (Takara Bio, CA, USA) according to manufacturer protocol. Quality control (QC) steps were performed using Qubit RNA HS Assay Kit and bioanalyzer. Sequencing was run at the Cancer Research UK Cambridge Institute Genomics core, on a Novaseq (Illumina).

RNA-seq data was analysed by Daniel Bode (Kent lab) and John Davey (University of York).

2.7.1 RNA-seq analysis

Pre-processing and batch correction

Raw counts were processed using edgeR (version 3.28.1)^{420,421}. Firstly, lowly expressed genes were excluded from downstream analysis. Here, genes with fewer than two libraries expressing a minimum of 1 CPM (counts per million) were considered lowly expressed. Subsequently, read counts were normalised using the trimmed mean of M values (TMM) method⁴²². Where there is multiple sequencing runs across an experiment, technical replicates were used to inform batch correction, performed with Limma (version 3.42.2)⁴²³. With little variation between Batch1 and Batch2, batch correction was performed on Batch1 and Batch3, where a significant variation of technical replicates was identified (Figure 2.7) Log-transformed and batch corrected values were subsequently used for downstream analysis.

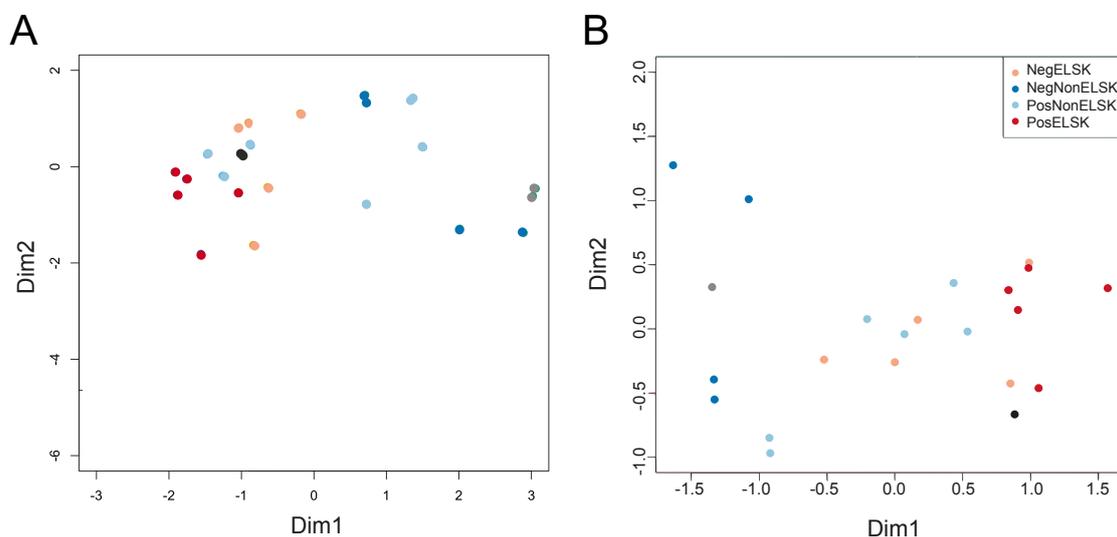


Figure 2.7 MDS plots of batch corrections performed on bulk RNA-seq samples.

3 runs of RNA-seq was performed with two separate technical replicates (black and grey). Coloured in by groups defined in section 5.2.

A) Before batch correction, with minor technical noise displayed by the grey replicates.

B) After batch correction, the technical repeats are virtually in different.

Differential gene expression analysis

Differential expression was performed using a likelihood ratio test approach. For this purpose, a negative binomial generalised linear model (GLM) was fitted. Multidimensional scaling (MDS) plots were computed using Limma (version 3.42.2). Genes were considered differentially expressed when a $\text{LogFC} \geq 2$ and $\text{FDR} < 0.05$.

GO term analysis

To compute gene ontology (GO) enrichment, gene symbols were converted to Entrez gene identifiers, using the mouse genome annotation database (org.Mm.eg.db, version 3.10.0). GO terms were extracted from the GO annotation database (GO.db, version 3.10.0) and GO term enrichment was computed using the Limma package (version 3.42.2). Biological process GO terms with a *p-value* < 0.05 were considered enriched.

Principal component analysis

Principal component analysis (PCA) was performed using the PCAtools R package (version 1.2.0). To ensure a Gaussian distribution of gene expression values for PCA computations, lowly expressed genes were removed based on a cumulative cut-off >40CPM across all samples per gene. During PCA computation, 10% of the most non-variable genes were excluded from analysis. To identify key genes driving separation of principal components, loadings plots were computed using the top 15% variable genes. Subsequently, a 0.05 cut-off irrespective of directionality was applied to select genes. Pearson correlation coefficients and the respective r^2 values were computed to determine the correlation of transplantation metadata with principal components.

MolO score and signature gene score analysis

A molecular overlap (MolO) gene signature, associated with freshly isolated LT-HSCs was previously described¹⁸⁶. MolO signature genes which passed the threshold for expressed genes (minimum 1 CPM in at least 2 libraries) were extracted from the dataset. The geometric mean was computed on log-transformed expression values for all MolO genes to derive the MolO score for each sample. A geometric mean was also computed for a novel repopulation gene signature, derived from the loading plots of the PCA.

SingleR correlation analysis

To identify dominant cell types of each sample library, the sc RNA-seq-based cell type recognition tool SingleR (version 1.0.6) was repurposed and applied to the bulk RNA-seq dataset at hand⁴²⁴. Default parameters were used to compute the correlation of each sample against the curated ImmGen reference dataset^{425–427}. In particular, subtypes within the broad haematopoietic stem cell compartment were used as reference. For NKT subtypes, published datasets were used⁴²⁸.

Pathway analysis

Pathway analysis was performed based on the curated Reactome pathway database, using the ReactomePA tool (version 1.30.0)⁴²⁹. Entrez gene identifiers for genes of interest were used as input. A *p-value* cut-off <0.05 was applied.

Visualisations on scRNA-seq dataset

Previous scRNA-seq characterisation of the haematopoietic hierarchy revealed distinct molecular signature of different cell types¹⁹⁰. The expression of transplantation signature genes was visualised in the Nestorowa et al. dataset¹⁹⁰.

2.8 Proteomics

2.8.1 EL08 CM proteomic screen

Samples were run by Robin Antrobus from the proteomics core in CIMR. Briefly 10 µl of each sample was digested in solution using 500ng trypsin. The resulting peptides were dried down, re-suspended in 15 µl MS solvent (3 % MeCN, 0.1 % TFA) for analysis on an Orbitrap Q Exactive mass spectrometer (Thermo). Peptides were fractionated using a Dionex RSLC nano3000 (Thermo) with solvent A comprising 0.1 % formic acid and solvent B comprising 80 % MeCN / 0.1 % formic acid. Peptides were loaded onto a 50 cm EASYSpray PepMap C18 column (Thermo) and eluted into the mass spectrometer using a gradient rising from 10 % to 40% solvent B by 55 min. MS data were acquired in the Orbitrap at 70,000 fwhm between m/z 400 and 1500 with a maximum AGC of 1×10^5 . Peptides were isolated and fragmented using HCD at 30 % collision energy. MSMS spectra were acquired in the ion trap with a maximum AGC of 1×10^4 . Raw files were processed in Maxquant 1.5.2.8. Heatmap was generated on R.

2.8.2 F12 PVA culture media proteomic screen

All media from media changes during 28-day cultures were collected in 96-well PCR plates and stored at -20°C. To allow for a temporal comparison and sufficient protein content, the following timepoints (in days) were pooled: 9 and 12 (early); 16 and 19 (middle); 23 and 26 (late). These were collected from three replicates of repopulating clones and three from non-repopulating clones, thus totalling 18 samples. Samples were quantified using a colourimetric peptide assay (Pierce), following the manufacturer's protocol. Sample volumes were extracted to accord to 5µg total protein and diluted in 100mM triethylammonium bicarbonate buffer (TEAB). Sample clean-up and desalting were performed using SDB-XC and C18 silica filters, manually packaged in tips and pre-conditioned. Pre-conditioning of C18 filters was performed using 100% acetonitrile (ACN), followed by washes with 80%ACN/0.1% trifluoroacetic acid (TFA) and equilibration with 0.1%TFA/dH₂O. All washes and equilibration were performed in triplicate. Prior to application to the filter, the sample was acidified with a final concentration of 0.25% TFA. After application to the filter, samples were washed with 0.1%TFA/dH₂O (triplicate) and eluted with 80%ACN/0.1%TFA. A total 5µg total protein was used per sample. A pooled control, comprising

samples from repopulating and non-repopulating clones, was prepared. Two separate multiplexing sets were setup with a pooled control in both sets. TMT10plex isobaric labelling reagents (Thermo) were freshly resuspended in dry ACN and subsequently added to each 10-plex, as indicated above. After incubation for 1h, the reaction was quenched using 5% hydroxylamine for 15min and subsequently pooled within their respective 10-plex. Samples were dried in a SpeedVac at 45°C prior to desalting. SDB-XC filters were washed with 100% ACN, followed by 77%ACN/0.1%TFA and equilibrated with 0.1%TFA/5%ACN/dH₂O. All washes and equilibration were performed in triplicate. After application to the filter, samples were desalted using 0.1%TFA/5%ACN/dH₂O (triplicate) and eluted with 77%ACN/0.1%TFA. Following desalting, samples were dried using a SpeedVac at 45°C, resuspended in 0.5% formic acid and sonicated for 10 seconds. Liquid chromatography – tandem mass spectrometry (LC-MS/MS) was subsequently performed using an Orbitrap Fusion mass spectrometer, coupled with a Dionex Ultimate 3000 UHPLC system. Raw data were subject to a database search in Proteome Discoverer 2.2 using default settings. Sample preparation was performed by Daniel Bode and LC-MS/MS was run by Dr Lu Yu, Jyoti Choudhary Lab, Institute of Cancer Research (ICR), London, UK.

2.9 Statistical analysis

All ANOVAs, Tukey's multiple comparisons test, t-Tests, Fisher's exact tests, Wilcoxon matched-pairs signed rank tests and Pearson's correlation statistics were calculated, and graphs were generated on Graphpad Prism6 or R.

Results

3 Identification of a novel immune cell subset in the *Fgd5*^{ZsGreen•ZsGreen/+} reporter mouse strain

As mentioned in section 1.4.1, HSC expansion has been a longstanding-goal in the field and progress has been partly hampered by the lack of robust *in vitro* markers. Recent advances in the development of HSC reporter mouse strains, spurred by better molecular characterisation of HSCs, provide a potential opportunity for novel *in vitro* markers. As mentioned in more detail in section 1.2.6, the *Fgd5*^{ZsGreen•ZsGreen/+} reporter mouse was chosen because the original data looked most promising. In the original report, Gazit et al. suggested that all phenotypic HSCs defined as Lineage⁻, Flk2⁻, CD34^{lo/-}, CD150⁺ were *Fgd5*⁺ and vice versa, that the vast majority of *Fgd5*⁺ cells are Lineage⁻, c-Kit⁺, Sca-1⁺, CD48⁻ and CD150⁺. As stated, this allowed for a single colour identification and purification of HSCs¹⁸⁸. Therefore, the *Fgd5*^{ZsGreen•ZsGreen/+} reporter mouse was used to test whether they can be useful markers of HSCs *in vitro*.

3.1 More than 50% of *Fgd5*⁺ cells are not phenotypic HSCs

Upon obtaining the strain, we first undertook experiments to confirm that the *Fgd5* reporter marked LT-HSCs. Fresh BM was harvested from *Fgd5*^{ZsGreen•ZsGreen/+} mice and analysed by flow cytometry for phenotypic markers of LT-HSCs (EPCR⁺CD150⁺CD48⁻CD45⁺, ESLAM, see section 1.2.5 for more details). In agreement with the original report, 98.71% ± SD 2.286 of ESLAM cells were *Fgd5*⁺ in 7 mice tested (Figure 3.1 below).

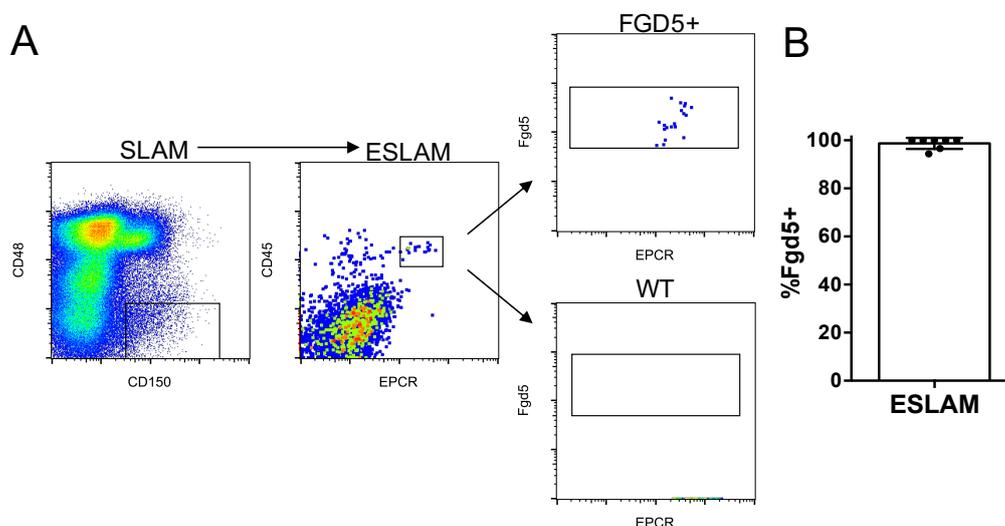


Figure 3.1 Phenotypic LT-HSCs are *Fgd5*⁺.

A) Representative gating strategy for ESLAM cells and *Fgd5* expression in *Fgd5^{ZsGreen+}* mice and WT littermates.

B) Percentage of ESLAM cells that are *Fgd5*⁺ n=7. Error bars represent data \pm SD.

Contrary to the original report, however, we observed that a substantial fraction of cells falling in the initial *Fgd5*⁺ gate was not present in the ESLAM gate (Figure 3.2A). Despite *Fgd5* significantly enriching for ESLAM cells to 38.37% \pm SD 9.058% on its own, there are still a large proportion of non-ESLAM cells. Intriguingly, these non-ESLAM cells, comprising 61.63% \pm SD 9.058% of *Fgd5*⁺ cells, clearly form a separate population that are EPCR⁻ and CD45^{high}, henceforth referred to as FE- cells (*Fgd5*⁺EPCR⁻). The FE- cells also express high levels of CD48⁺ and are CD150⁻ (Figure 3.2), further suggesting they do not have a phenotypic LT-HSC identity. The *Fgd5*⁺EPCR⁺ (FE+) cells conversely are predominantly CD150⁺ and CD48⁻, demonstrating that FE+ on its own represents a considerable enrichment of LT-HSCs with a two-marker strategy.

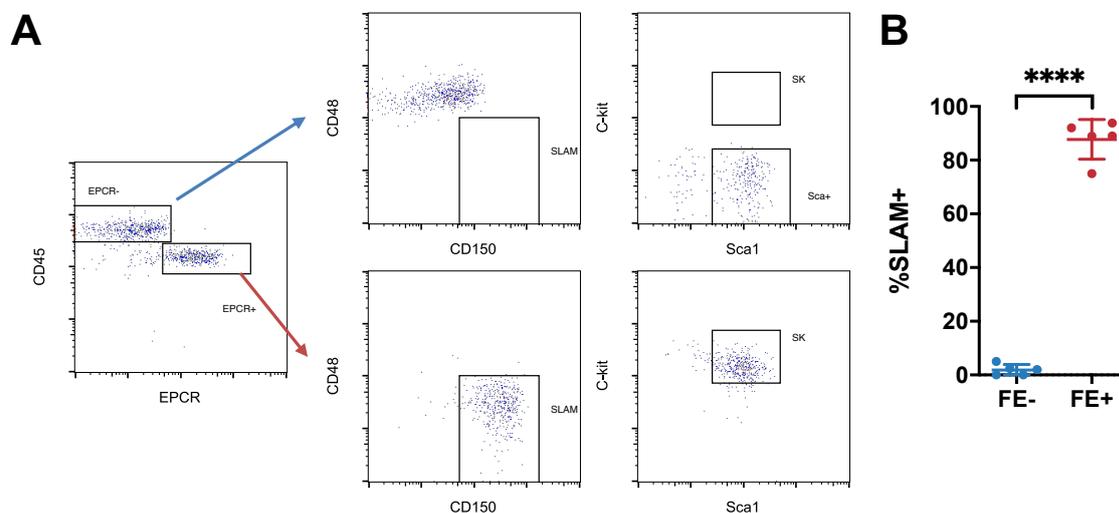


Figure 3.2 *Fgd5*⁺ cells are not uniformly LT-HSCs.

A) Representative flow plots of cells, pre-gated for *Fgd5*⁺ live singlets. *Fgd5*⁺ cells can be fractionated into an EPCR⁺ (FE+) and an EPCR⁻CD45^{hi} (FE-) population.

B) FE+ cells are mostly phenotypically CD150⁺CD48⁻ by percentage whereas FE- cells are not. n= 5 paired T-test two tailed. **** = p < 0.0001. Error bars represent data \pm SD.

3.1.1 *Fgd5*⁺ CD48^{dull} ESLAM cells are not significantly different from normal ESLAMs

Interestingly, the two-marker FE+ strategy identified a population of cells that were CD48^{dull} as opposed to the CD48⁻ gate reported in the original SLAM marker paper¹²¹. These cells would normally be removed in a standard ESLAM gating strategy¹⁸⁴, but the co-expression of EPCR and *Fgd5* suggested that this gate might unfairly exclude some functional HSCs. To investigate whether *Fgd5* and EPCR alone could be a two-colour alternative to marking LT-HSCs, we assessed the CD48^{dull} fraction specifically using single cell transplantations (n=21) (Gating strategy shown in Figure 3.3).

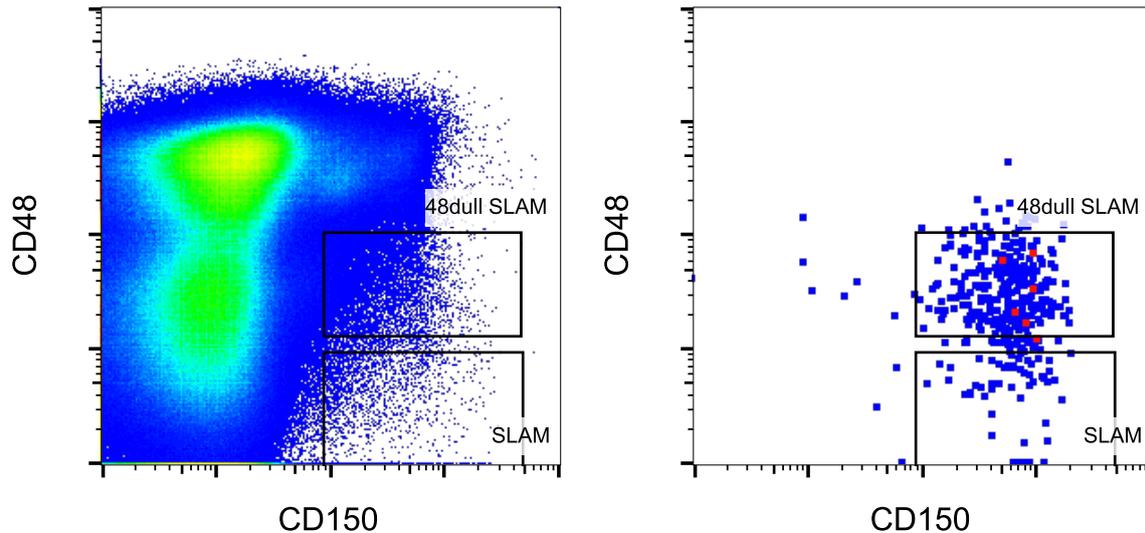


Figure 3.3 Representative gating layout for *Fgd5*⁺EPCR⁺CD150⁺CD48^{dull} cells.

The left panel shows the two gates for SLAM cells and CD48^{dull} SLAM cells, overlaid on live cells and the right panel shows the gates overlaid on FE+ cells. *Fgd5*⁺EPCR⁺CD150⁺CD45⁺CD48^{dull} cells represent 0.018% of live BM cells, whilst standard *Fgd5*⁺ ESLAM cells represent 0.0023% of live BM cells.

7 of 21 single cells had long term (16 weeks) reconstitution in primary mice. 5 of these 7 single cells (24% of the total) displayed multi-lineage reconstitution at week 16 after primary transplantation, 4 of which had the robust GM-contribution characteristic of an HSC with durable self-renewal potential (α and β -HSC subtypes, see section 1.3.1)(Figure 3.4)²⁰⁶. Another 7 out of 21 clones had transient reconstitution, where chimerism above 1% was detected at one or more earlier time points, but not at week 16. This suggests that a substantial number of additional LT-HSCs could be captured by extending the CD48 gate. However, 24% is lower than the typical ESLAM phenotype (56%)¹⁸⁴, meaning that extending the gate might not be appropriate for a number of assays (e.g., single cell molecular assays). Also, as only 21 mice were assayed resulting in a p-value of 0.081 (Fisher's exact test), it is still formally possible that there is a significant difference between the two populations, and this would require additional transplantations to make robust conclusions. Together these data demonstrate that *Fgd5* and EPCR on their own mark a population of cells highly enriched for functional HSC activity and might allow more LT-HSCs to be obtained from individual animals.

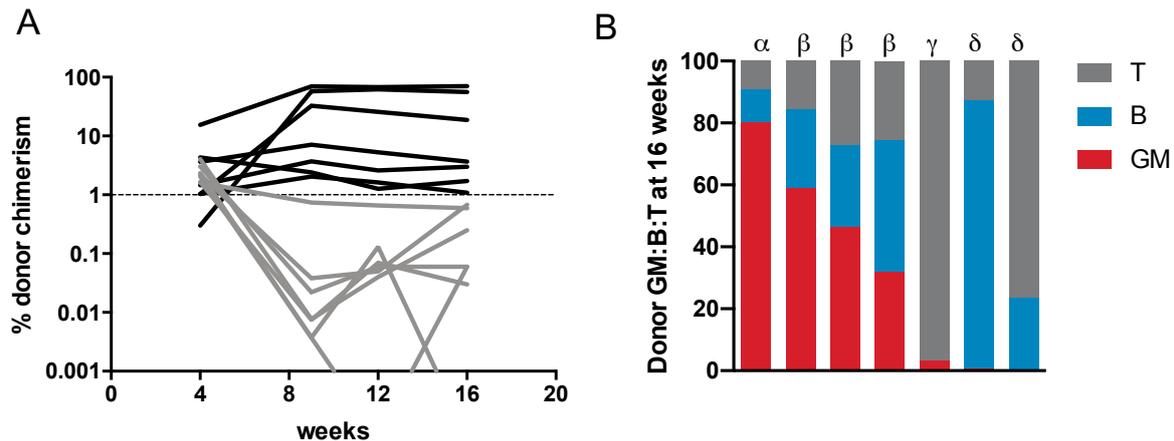


Figure 3.4 The CD48^{dull} ESLAM phenotype contains a high proportion of functional HSCs.

- A) Donor chimerism of singly transplanted CD48^{dull} ESLAM cells (n=21); repopulated >1% at week 16 (black), transient >1% at week 4,8 or 12 (grey). The dotted horizontal line represents 1% donor chimerism. Recipients with no detectable reconstitution is not shown.
- B) Ratio of lineage contribution of 7 positively-reconstituted donor cells at week 16 and its assigned HSC subtype above. Only the α and β -HSCs would be expected to give robust secondary transplantations indicative of HSCs with durable self-renewal potential.

3.1.2 FE- cells are not HSCs

If *Fgd5* were to be used as an HSC marker *in vivo*, it becomes vital to understand the other cells that do not express the traditional HSC markers. This also becomes important for *in vitro* cultures in order to exclude the possibility that *Fgd5*⁺ non-HSCs might arise in the culture. The remainder of this results chapter focuses on the detailed cellular and molecular characterisation of the FE-cell population.

In order to characterise FE- cells, we first undertook *in vitro* stem/progenitor cell assays. Colony forming cell (CFC) assays and 10-day single HSC liquid cultures were undertaken as described previously¹⁶⁵. As shown in Figure 3.5, FE- cells were unable to grow in LT-HSC liquid culture conditions and did not form any myeloid haematopoietic colonies *in vitro*. Since ~70% of FE- cells expressed Sca-1, a known stem/progenitor cell marker, we first tested whether they might represent non-terminal lymphoid progenitors by undertaking OP9 co-culture assays. Again, FE- cells failed to form colonies compared to ESLAM cells, suggesting that they are not lymphoid cell progenitors either (Figure 3.6).

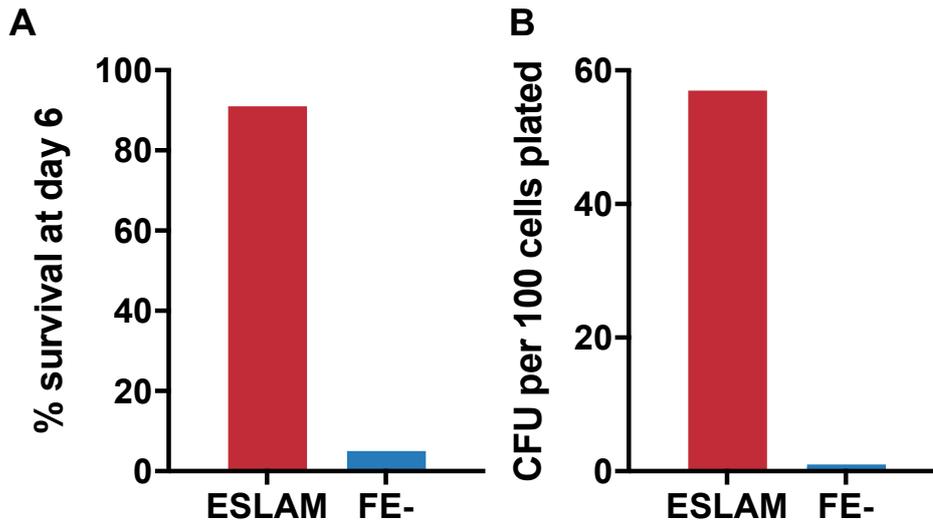


Figure 3.5 FE- cells do not generate haematopoietic colonies in HSPC assays.

- A) Clonal survival rates of ESLAM vs FE- cells in liquid cultures of SS supplemented with 10% FCS, 300ng/mL SCF and 20ng/mL IL-11. (n= 122 FE- cells n=192 ESLAM cells)
- B) CFC assay. ESLAM cells, n = 100 cells plated; FE-, n = 300 cells plated

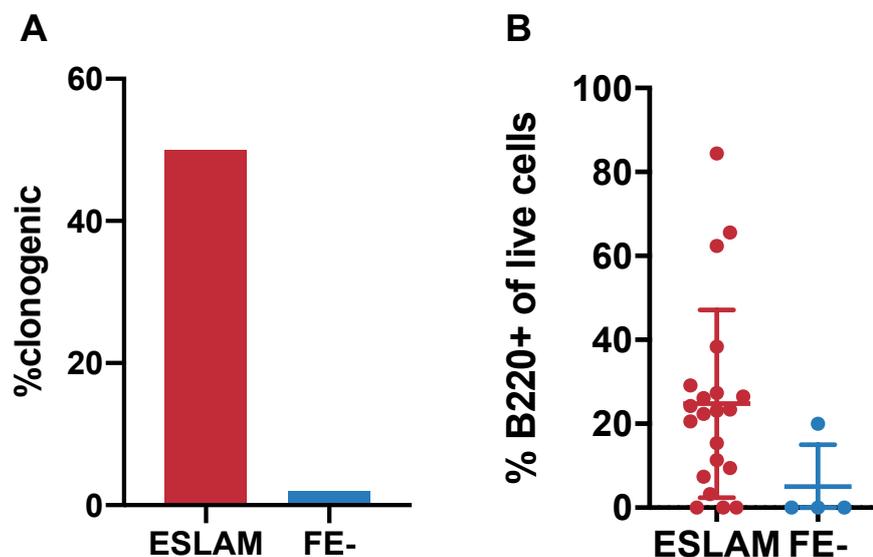


Figure 3.6 FE- cells do not grow in OP9 B cell progenitor assays.

Single ESLAM or FE- cells were cultured with OP9 stromal cells and analysed for their ability to form colonies and produce B220⁺ cells.

- A) Clonal survival of FE- cells compared to ESLAM cells after 14 days on OP9 cells.
- B) Percentage of B220⁺ cells in each surviving clone derived from ESLAM or FE- cells.
ESLAM, n=48; FE-, n=48.

3.1.3 FE- cells are of haematopoietic lineage

To confirm if FE- cells are of haematopoietic origin, recipient BM from primary single cell transplantations including the CD48^{dull} experiment detailed in section 3.1.1, were analysed for presence of FE- cells. As shown in Figure 3.7, *Fgd5*⁺ cells were present in repopulated marrow more than 16 weeks post transplantation. Reassuringly, both FE- and FE+ cells were found in the BM, proving that single HSCs can give rise to new FE+ and FE- cells. The proportion of FE- and FE+ cells was observably different between recipients; some BM samples only had FE- cells and vice versa, some only contained FE+ cells. Therefore, the proportions of FE- and FE+ cells were correlated with the lineage output of the donor cells.

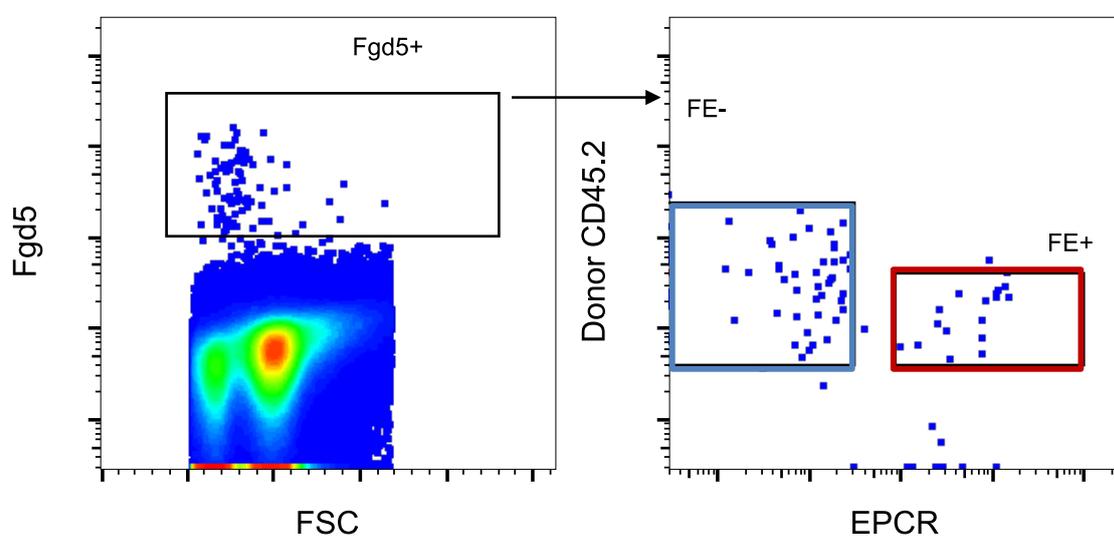


Figure 3.7 Representative gating strategy for BM cells from repopulated primary transplantation recipient.

BM of recipients of single clone-derived 1-3 cell transplants, 16 weeks post transplantation. Left panel is pre-gated for live and singlets; *Fgd5*⁺ cells exist in BM of recipient. Right panel shows the gating strategy for FE- cells and FE+ cells in recipient BM. This particular mouse had cells in both gates, while some recipients do not have any *Fgd5*⁺ cells or cells in only one gate.

As expected, the percentage of FE+ cells, presumably representing phenotypic LT-HSCs, correlates strongly with donor GM lineage contribution ($r = 0.7570$, $p=0.0007$), which is typically associated with secondary reconstitution capability²⁰⁶. Interestingly, the frequency of FE- cells is strongly correlated ($r = 0.7857$, $p=0.0003$) with the donor contribution to T cell lineages, but not with B cell ($p=0.0956$) or GM lineages ($p=0.1144$) (Figure 3.8). This suggests that FE- cells are coupled to T lineage fates.

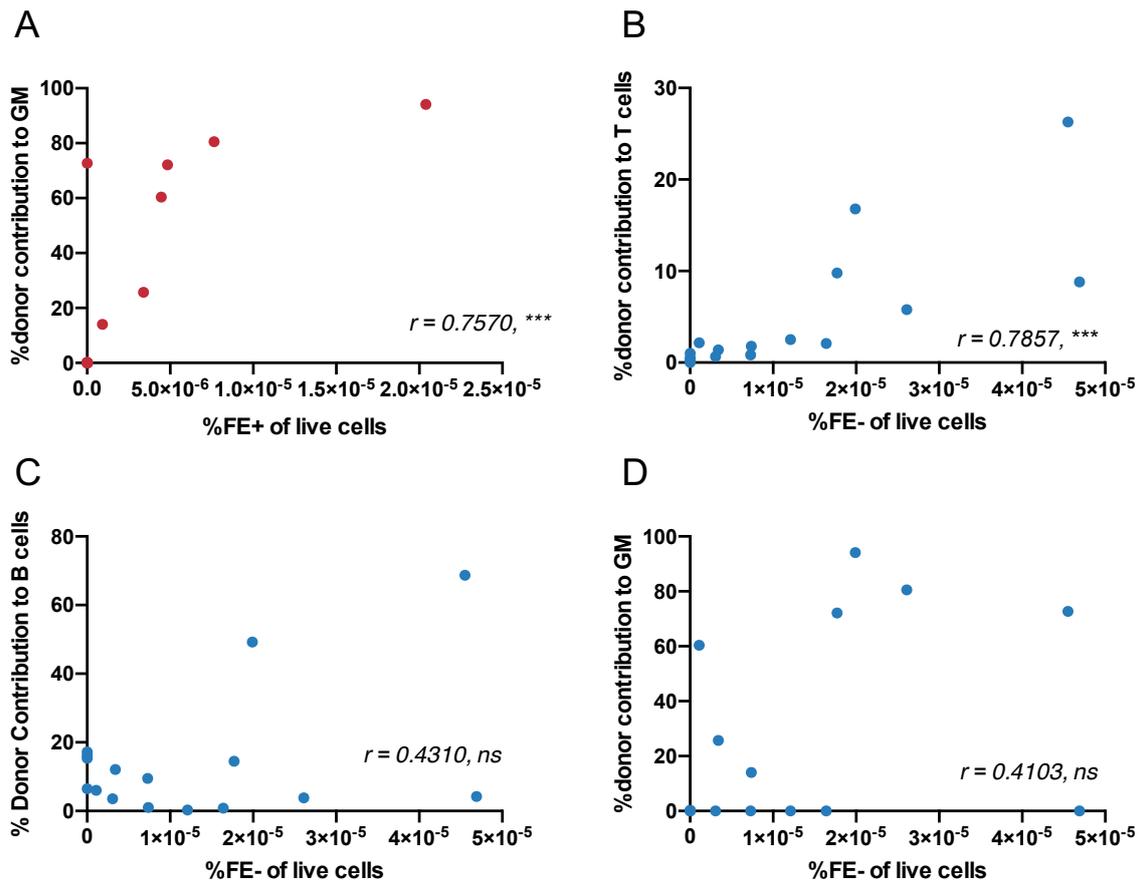


Figure 3.8 FE- correlates strongly with T cell output in primary transplantation.

Analysis of BM of primary recipients of 1-3 fresh or cultured HSCs, 16 weeks post transplantation.

A) Percentage of donor GM cells correlated to percentage of FE+ cells.

B-D) Percentage of donor T, B and GM cells correlated with percentage of FE- cells.

Pearson correlation, ***= $p < 0.0005$, exp=2, n = 16.

3.1.4 FE- cells do not result from different BM preparation methods

Next we tried to understand why these cells were not present in the original paper by Gazit et al.¹⁸⁸

Up until this point, all the BM samples were prepared by crushing the bones with a pestle and mortar. Since flushing bones with a syringe is another common way to prepare BM cell suspensions, we hypothesised that FE- cells might result from the technique of BM crushing. To test this, two legs of the mice were prepared independently by crushing and flushing the bones, and the cells were stained for *Fgd5* and EPCR. As shown in Figure 3.9, FE- cells are present in both preparation techniques.

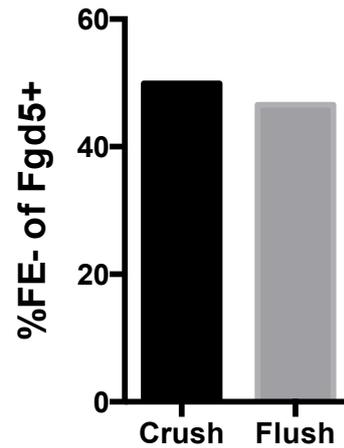


Figure 3.9 FE- cells are present in both flushed and crushed bones.

Percentage of FE- cells in crushed or flushed bones from the same mouse. Showing that FE- cells are present and in similar proportions from both preparation methods. Exp = 1 n=1.

3.2 Gene expression of FE- cells suggests a lymphoid cell identity

As we have already shown, FE- cells express high levels of CD45 and CD48, both of which are found in a wide array of haematopoietic cell types. Because of the ambiguous nature of this cell population, we performed bulk RNA-seq on 4415 FE- cells to gain more insights into their identity and detected 29876 genes.

As expected, surface markers such as CD45 (*Ptprc*) and CD48 were highly expressed, validating observations by flow cytometry (Table 3.1). Although not at an extremely high level relative to other detected genes (5107th out of 29876 genes) *Fgd5* is expressed (12.6 CPM) and so was *ly6a* (Sca-1) (2846th highest, 21.7 CPM). HSC marker genes that were not expected to be expressed in this population lacking functional HSC activity were indeed lowly expressed, such as c-Kit (*Kit*) (3.6CPM), EPCR(*Procr*) (2.33 CPM) and CD150 (*Slamf9*) (4.5 CPM). This adds confidence to the veracity of the sequencing data.

Table 3.1 Expression of cell surface marker genes associated with FE- cells

Gene name	CPM	Rank
<i>Ptprc</i>	444.60897	43
<i>Cd48</i>	76.5069955	475
<i>Ly6a</i>	21.7349717	2846
<i>Fgd5</i>	12.5754137	5106
<i>Slamf1</i>	4.45091407	11411
<i>Kit</i>	3.62101417	13255
<i>Procr</i>	2.32734868	18197

As a first broad pass, the top (Cluster of differentiation) CD marker genes were filtered from the dataset. Consistent with findings from section 3.1.3 of coupling to T cell lineage, FE- cells highly express T lymphocyte markers (Table 3.2), including CD3, CD2, CD6, CD27, CD5 and CD4¹². Several known Natural Killer cell markers (NK cell), such as CD244 and CD226, are also highly expressed, though CD244 can also be expressed by monocytes and dendritic cells. Interestingly, there are also genes that are associated with monocytes, such as CD68 and CD36.

Table 3.2 Top 40 Cluster of differentiation genes expressed by FE- cells

CD genes	CPM	Rank	CD genes	CPM	Rank
<i>Cd52</i>	314.0330746	68	<i>Cd27</i>	40.10012543	1269
<i>Cd82</i>	169.1285354	148	<i>Cd3d</i>	37.39708161	1381
<i>Cd68</i>	154.357035	172	<i>Cd300a</i>	37.12284799	1403
<i>Cd3g</i>	137.738457	206	<i>Cd274</i>	33.23289836	1633
<i>Cd74</i>	129.3725532	218	<i>Cd40lg</i>	33.16195634	1638
<i>Cd2</i>	115.8900714	255	<i>Cd300lb</i>	31.20067375	1775
<i>Cd53</i>	103.0490123	310	<i>Cd244a</i>	30.4359084	1826
<i>Cd300ld</i>	95.6324585	352	<i>Cd24a</i>	29.16623209	1954
<i>Cd44</i>	86.936797	401	<i>Cd5l</i>	26.87848233	2185
<i>Cd48</i>	76.50699551	475	<i>Cd81</i>	24.63688996	2455
<i>Cd47</i>	74.19312876	503	<i>Cd7</i>	22.05367041	2798
<i>Cd3e</i>	71.54343145	526	<i>Cd163</i>	19.43293165	3254
<i>Cd300c2</i>	67.32175138	582	<i>Cd6</i>	18.55928239	3423
<i>Cd37</i>	65.98612553	607	<i>Cd5</i>	17.28239149	3700
<i>Cd180</i>	62.46169044	655	<i>Cd96</i>	16.62978885	3862
<i>Cd300e</i>	61.92829097	671	<i>Cd83</i>	14.93512864	4329
<i>Cd84</i>	53.91835165	825	<i>Cd4</i>	14.79806235	4374
<i>Cd9</i>	51.75818646	879	<i>Cd160</i>	14.20456069	4557
<i>Cd36</i>	50.71924938	913	<i>Cd302</i>	13.89522458	4638
<i>Cd226</i>	40.19040659	1265	<i>Cd177</i>	12.86234296	4977

To gain a broader understanding of the cells, GO term analysis was performed on the top 500 ranked genes (Table 3.3). Amongst the most significant GO terms, many were associated with Immune system processes and T cell responses. Although GO terms are not conclusive, this further suggests that FE- cells are T lymphocytes.

Table 3.3 GO terms associated with top 500 genes expressed in FE- cells

GO Term	P-value	FDR
immune system process	1.01E-16	2.00E-13
innate immune response	5.74E-10	1.01E-06
intracellular signal transduction	5.74E-10	1.01E-06
leukocyte cell-cell adhesion	2.13E-09	3.77E-06

adaptive immune response	2.07E-07	3.65E-04
T cell receptor signalling pathway	2.78E-07	4.92E-04
cellular response to tumour necrosis factor	1.20E-05	0.02126408
negative thymic T cell selection	1.25E-05	0.02200169
positive regulation of T cell proliferation	1.93E-05	0.03416226
positive regulation of T cell activation	3.63E-05	0.06409375

3.2.1 Surface marker phenotyping confirms lymphoid marker expression

Next, FE- cells were immunophenotyped for common T cell- lineage markers. Because NK cells can also exhibit T cell markers, NK cell markers were also tested; alongside other commonly used haematopoietic lineage markers. As shown in Figure 3.10, many of the highly expressed genes for surface markers are indeed present on a proportion of FE- cells, including CD3 (46% \pm SD 15.3%), CD5 (40.75 \pm SD 9.43%), and CD244 (55.25% \pm SD 12.53). Interestingly, NK1.1, which is a gene associated with NK cells, was also present in a proportion of cells (30.25% \pm 8.958). Notably, CD4 was present but in a smaller proportion of FE- cells (21 \pm SD 13.54).

As expected, common markers for other lineages, such as B220 (B cell), Ly6g (Granulocytes), and Ter119 (Erythrocytes) were negligibly expressed (Figure 3.10). Overall, this supports the characterisation that these cells are lymphoid in nature but raises the possibility that they are NKT cells.

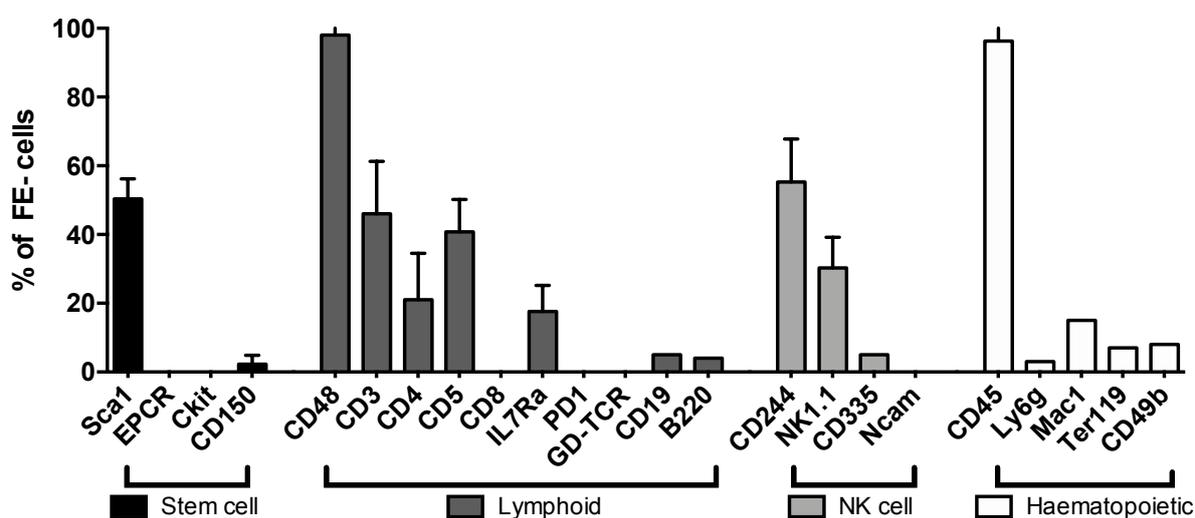


Figure 3.10 FE- cells express various surface markers characteristic of lymphoid and NK cells.

Percentage of FE- cells expressing various surface markers grouped by the cell types that they are most commonly associated with. Error bars represent data \pm SD.

Sca-1, EPCR, exp=5; c-Kit, CD150, CD48, CD3, CD4, CD5, CD244, NK1.1, CD45, exp=4; IL7Ra, exp=3; CD8, CD19 exp=2; PD1, GD-TCR, B220, CD335, Ncam, Ly6g, Mac-1, Ter119, CD49b, exp=1.

3.2.2 FE- cells do not vary with age and are present in multiple lymphoid tissues.

Since multiple experiments were done on *Fgd5^{ZsGreen⁺ZsGreen/+}* mice of varying ages, we reanalysed the flow cytometry data to look for changes in FE- frequency with age. Because some of the experiments had used lineage depleted BM samples, a head-to-head comparison of BM samples with or without lineage depletion was performed. Surprisingly, there was no significant difference ($p=0.0547$) before and after lineage depletion (Figure 3.11). This may be explained by the relatively low proportion of *Fgd5⁺* cells compared to the overwhelming proportion of lineage positive cells that the enrichment kit removes, the vast majority of which would be *Fgd5⁺*. As shown in Figure 3.11, there is a decline in proportion of FE- cells with age, although this is not statistically significant.

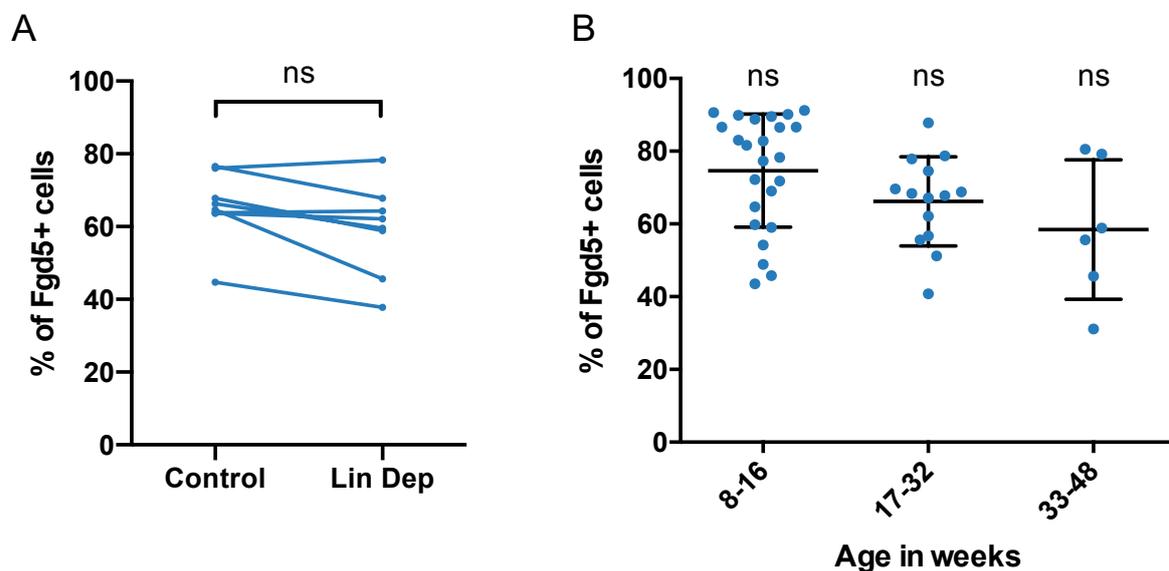


Figure 3.11 Frequency of FE- cells in mice of different ages.

- A) Percentage of FE- cells out of *Fgd5⁺* cells with or without lineage depletion. Exp=5; n=8. Wilcoxon matched-pairs signed rank test (two tailed).
- B) Percentage of FE- cells out of *Fgd5⁺* against different ages. Weeks 8-16, n=24; weeks 17-32, n= 14; weeks 33-48, n=6. Tukey's multiple comparisons test (two tailed). Error bars represent data \pm SD.

As FE- cells were presumed to be lymphoid cells, various lymphoid and peripheral tissues were analysed for the presence of FE- cells. Indeed, FE- cells were found in non-BM lymphoid tissues, and in particular there was a significantly increased frequency of FE- cells in the liver ($p=0.01$), lungs ($p=0.0003$) and thymus ($p=0.0276$) compared to the BM (Figure 3.12). As expected, the vast majority of *Fgd5⁺* cells in these lymphoid tissues are FE- cells (Figure 3.12), reflecting the absence of stem cell population in these lymphoid tissues.

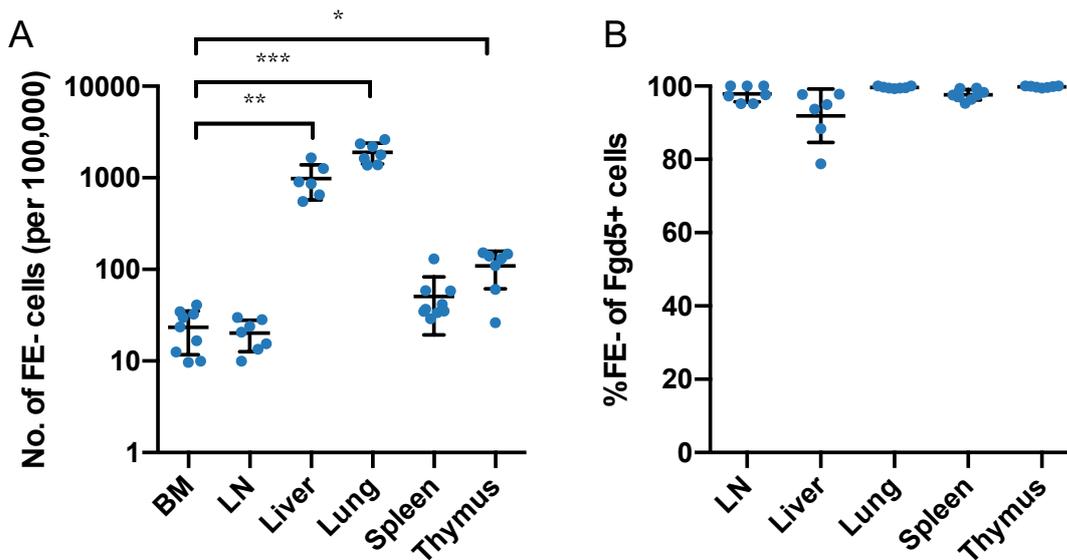


Figure 3.12 FE- cells in various lymphoid and peripheral tissues.

A) Frequency of FE- cells in various lymphoid and peripheral tissues compared to BM. Error bars represent data \pm SD.

B) Percentage of FE- cells out of *Fgd5*⁺ cells in the same lymphoid and peripheral tissues.

BM, n=8; LN, Lymph node, n=7; Liver, n = 6; Lung, n=7; Spleen, n=8; Thymus, n=7. Exp=2. Error bars represent data \pm SD.

3.2.3 FE- cells can be sub-fractionated with surface markers

Following the establishment that a proportion of FE- cells express CD3, CD5, CD244 and NK1.1 individually, these markers were evaluated for co-expression, in order to better understand the cell subpopulations within the FE- fraction. As shown in Figure 3.13, FE- cells can be cleanly separated into two subpopulations: CD5⁺ and CD244⁺. Interestingly, CD5⁺ cells co-express CD3, NK1.1 and Sca-1, suggesting that they might be similar to NKT cells. Further characterisation shows that the small proportion of CD4⁺ and IL7R α ⁺ cells belong to a subfraction of these CD5/CD3/NK1.1/Sca-1⁺ subpopulation (data not shown), henceforth referred to as FE-CD5⁺ cells. The other population will be referred to as FE-CD244⁺ cells.

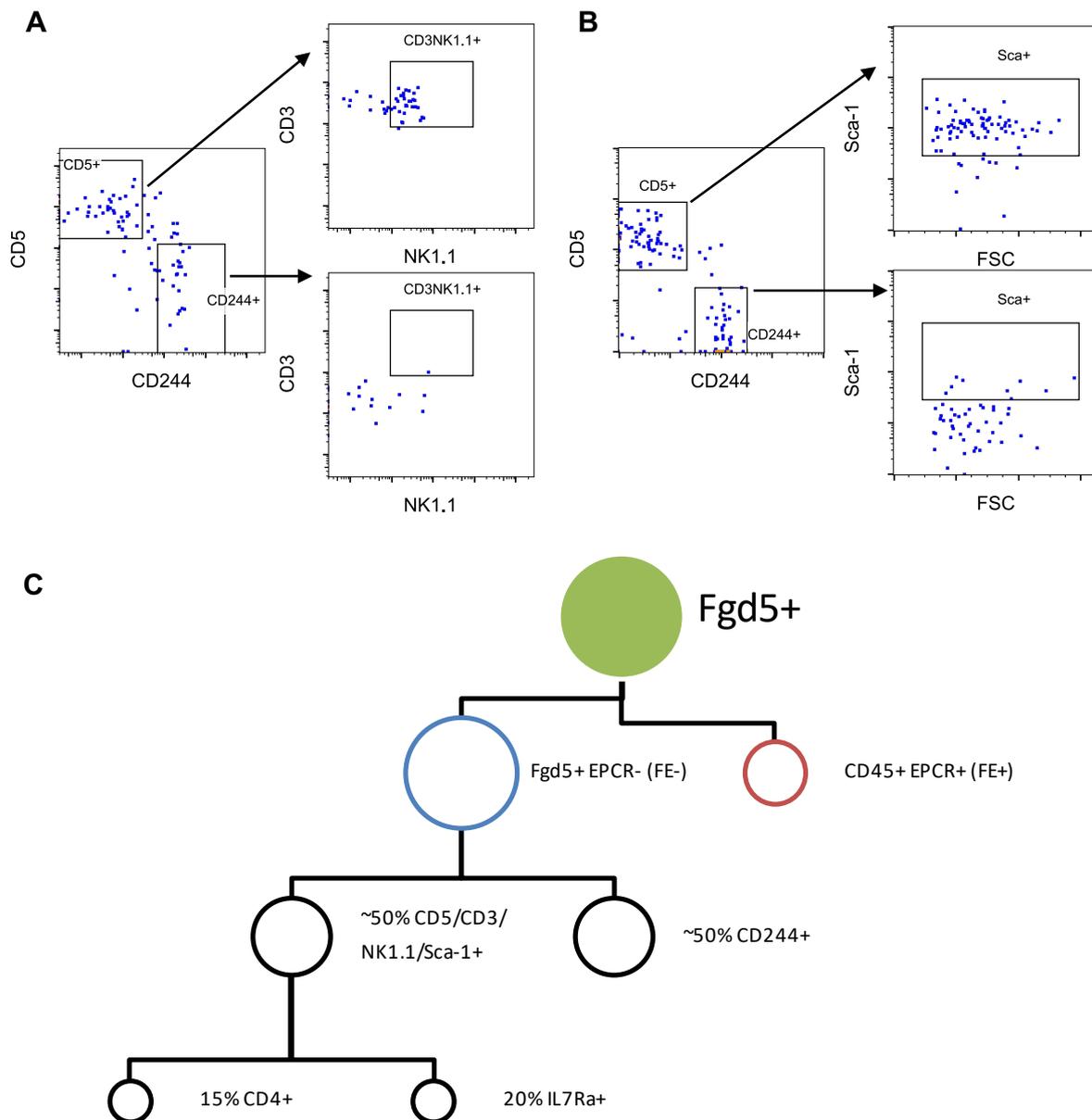


Figure 3.13 Co-expression of immune surface markers allows FE- cells to be further subfractionated.

A,B) Representative flow cytometric layout of FE- cells and their co-expression of CD5, CD3, NK1.1 and Sca-1.

C) Deduced hierarchy of marker co-expression in *Fgd5*⁺ cells. The area of each circle represents the approximate proportion of that population to *Fgd5*⁺ cells.

3.3 Bulk RNA sequencing of CD5⁺ and CD244⁺ fractions

In order to further resolve the identity of the two subpopulations of FE- cells, FE-CD5⁺ cells (11,495 and 12,969 cells) and FE-CD244⁺ cells (6,814 and 5,666 cells) were sorted and bulk RNA-seq was performed again.

Differential gene expression analysis of the 14,614 detectable genes was performed. 2724 genes were significantly upregulated in FE-CD5⁺ cells, including the genes encoding the surface markers already shown to be exclusively expressed in FE-CD5⁺ cells, such as *Cd5*, *Cd3*, *Ly6a* (Sca-1), *Klrb1c* (NK1.1), *Cd4*, *I17r*. Conversely, 3135 genes were significantly upregulated in FE-CD244⁺ cells, including *Cd244a* itself. Overall, this lends confidence to the robustness of the sequencing data.

As a broad first pass, we utilised the Immunological Genome Project (ImmGen) to compare our data against. ImmGen is a public resource created by immunologists and computational biologists containing standardised gene expression data of the entire mouse immune system. Including different maturation, tissue localisation and activation states, there are over 250 cell types in the resource. Using data mined from ImmGen, our novel cell populations were compared to various haematopoietic populations for transcriptome similarities using SingleR⁴²⁴. Single R is a computation method developed by Aran et al. that annotates transcriptomes by correlating them to reference bulk transcriptomes such as the ImmGen database.

As shown in Figure 3.14, SingleR identified FE-CD5⁺ cells as NKT cells and FE-CD244⁺ cells as monocytes. To further understand which subtypes of NKT cells they are, FE-CD5⁺ cells were further compared to the bulk and single cell transcriptomes of various iNKT subtypes generated by Engel et al.⁴²⁸. As shown in Figure 3.15, FE-CD5⁺ cells are clearly most similar to iNKT1 cells, in both bulk and single cell reference datasets.

To further investigate the identity of FE-CD244⁺ cells, they were compared to bulk RNA-seq datasets of monocytes derived from different mouse tissues⁴³⁰. However, the correlation score was low and inconclusive with all the reference datasets compared (data not shown). As monocytes can be diverse, and FE-CD244⁺ cells may not be homogeneous, their identity remains elusive and require further investigation. Therefore, the focus shifted towards FE-CD5⁺ cells.

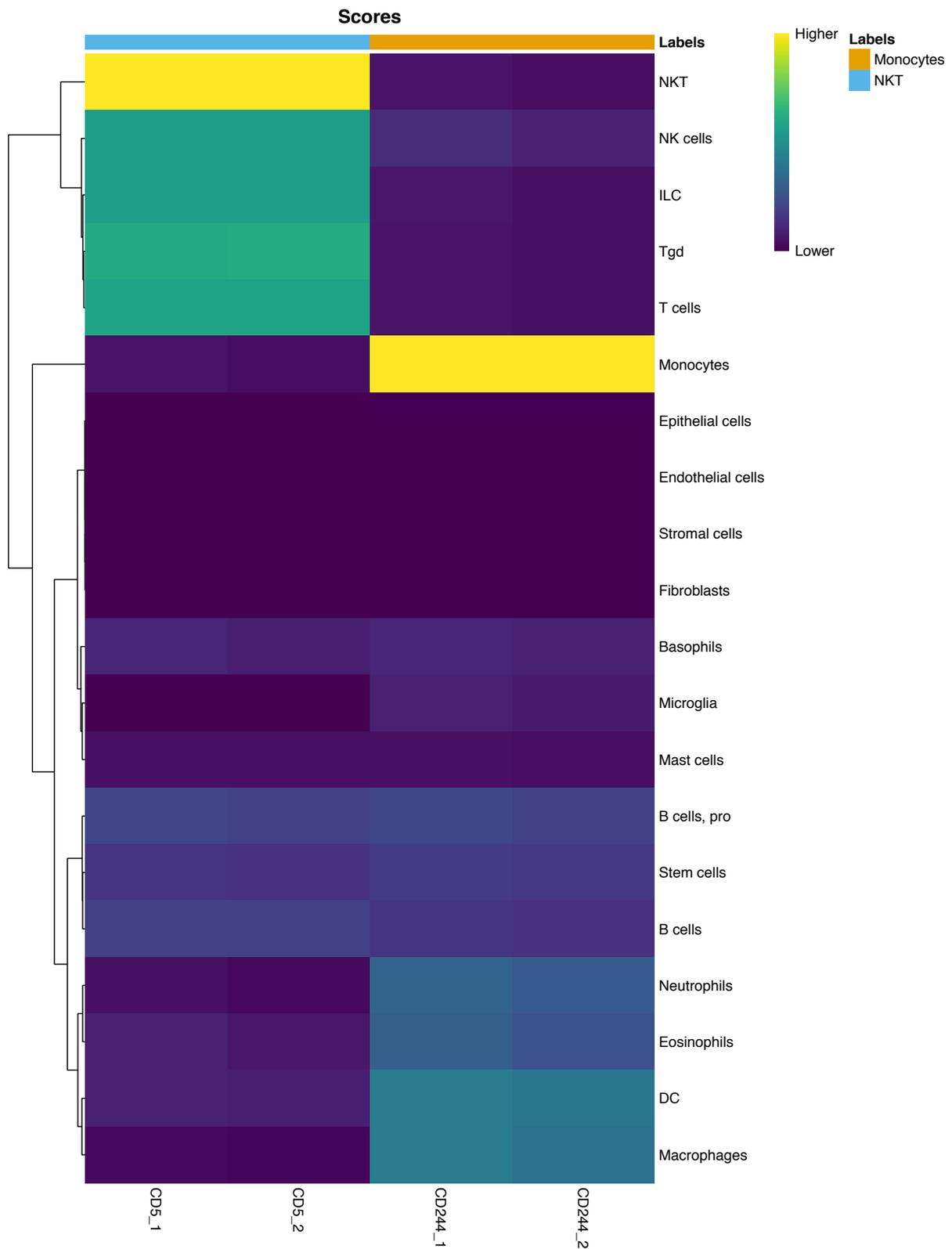


Figure 3.14 SingleR identifies FE-CD5+ cells as NKT cells and FE-CD244+ cells as monocytes.
 SingleR analysis of two technical repeats of bulk FE-CD5+ and FE-CD244+ cells using ImmGen database as reference. Correlation scores were scaled to highest (yellow) to lowest (purple). The population with the highest correlation is labelled on the top.

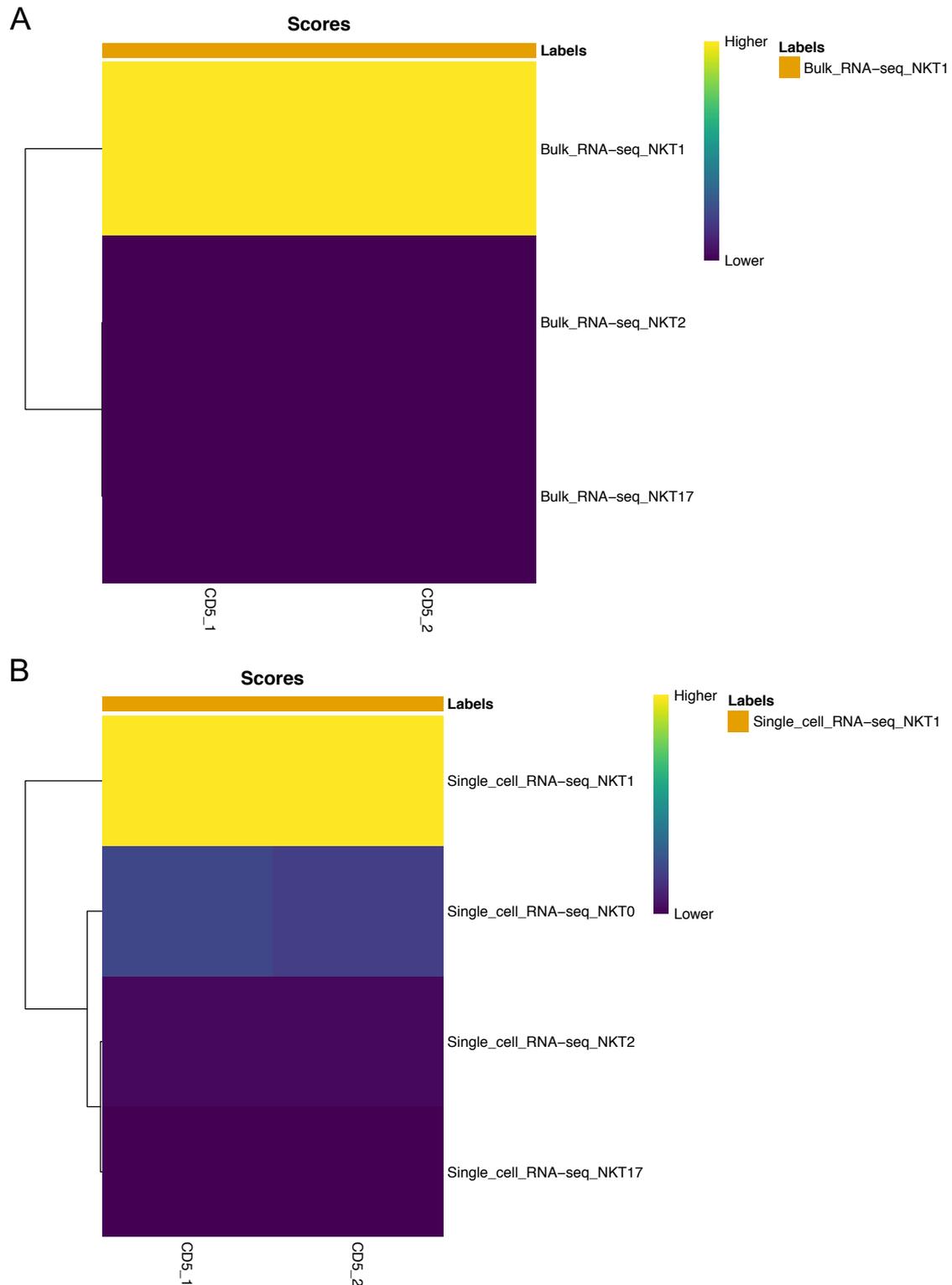


Figure 3.15 SingleR identifies FE-CD5+ cells as iNKT1 cells.

A) SingleR analysis of FE-CD5+ cells using bulk RNA-seq dataset of iNKT subtypes as reference⁴²⁸. The reference dataset contains 3 replicates each for NKT1, NKT2 and NKT17 cell types. The population with the highest correlation is labelled on the top.

B) SingleR analysis of FE-CD5+ cells using scRNA-seq dataset of iNKT subtypes as reference⁴²⁸. The reference dataset contains 203 cells; 45 NKT0, 46 NKT1, 68 NKT2 and 44 NKT17 cells. The population with the highest correlation is labelled on the top.

3.3.1 α -Galactosylceramide reactivity confirms that FE-CD5+ cells are iNKT cells

As mentioned in section 1.1.2, iNKT cells are distinguished from other NKT subtypes by their ability to bind and react with α -Galactosylceramide (α -GalCer). To test whether these cells were indeed iNKT cells, BM and thymus cells were analysed for CD1d-tetramer (CD1d)-PBS57 (an analogue to α -GalCer) binding. As shown in Figure 3.16, ~75% of FE-CD5+ cells in both BM and thymus were able to bind to Cd1d-PBS57 and express intermediate levels of TCR β , strongly suggesting that they are indeed iNKT cells. As the FE-CD5+ cells co-express NK1.1, which distinguishes iNKT1 cells from other iNKT subtypes, it highly suggests that FE-CD5+ cells are iNKT1 cells, as SingleR predicted. Interestingly only about 20% of CD1d⁺, TCR β ⁺, NK1.1⁺ cells are *Fgd5*⁺, suggesting that *Fgd5* expression marks a subset of NKT1 cells (Figure 3.16). It remains unclear whether the *Fgd5*⁺ subset of NKT1 cells have a distinct function to *Fgd5*⁻ iNKT cells.

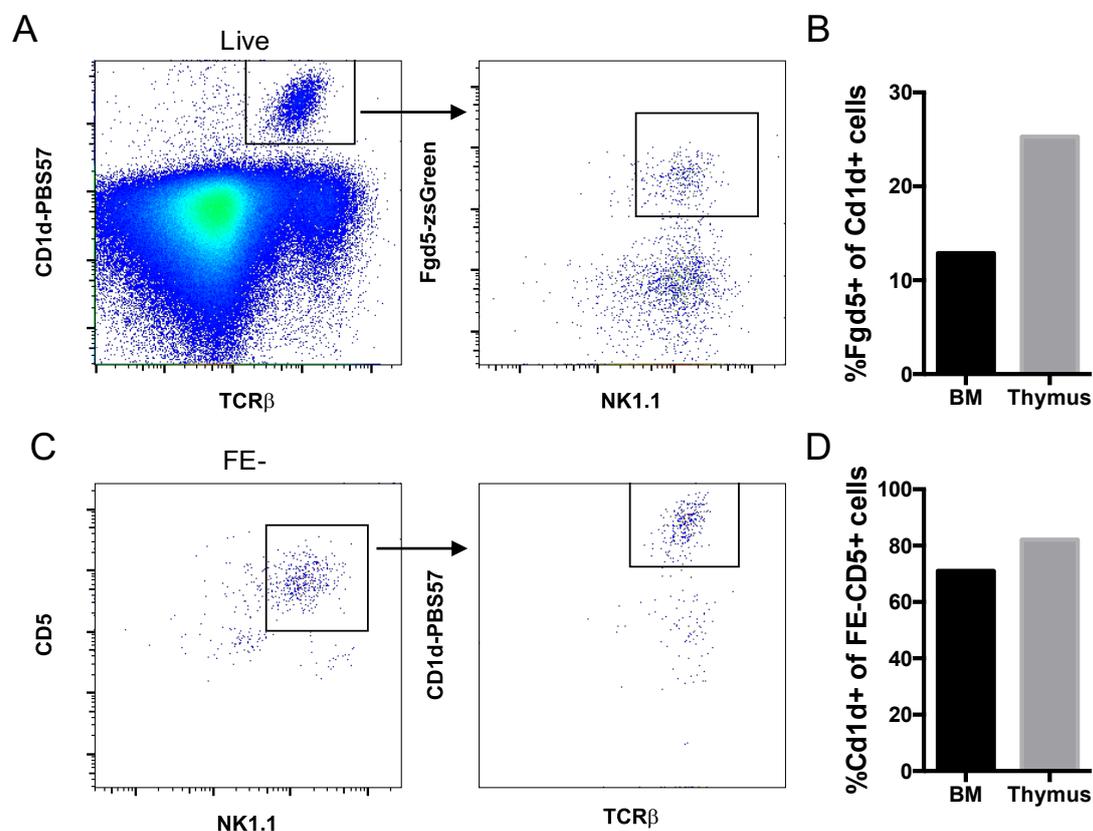


Figure 3.16 FE-CD5+ cells are reactive to α -GalCer.

- A) Representative gating layout of *Fgd5*⁺ BM cells. Left panel is gated out of live singlets.
 B) Percentage of *Fgd5*⁺ cells in CD1d-PBS57/TCR β ⁺ cells from BM and thymus. Exp=1. n=1.
 C) Representative gating layout of *Fgd5*⁺ BM cells. Left panel is gated out of live, singlet, FE- cells
 D) Percentage of CD1d⁺/TCR β ⁺ cells in FE-CD5⁺ cells from BM and thymus. Exp=1. n=1.

3.3.2 FE-CD5+ cells secrete interferon- γ

As mentioned in section 1.1.2, iNKT cells are known to secrete specific cytokines upon stimulation. In particular NKT1 cells are known to predominantly secrete interferon- γ (IFN- γ). In

order to access the cytokine profiles of FE-CD5⁺ cells, intracellular flow cytometry was used to assess the expression of IFN- γ and other common immune cytokines, such as tumour necrosis factor (TNF α) and interleukin 5 (IL-5). FE-CD5⁺ cells were sorted and stimulated with PMA and ionomycin and then fixed for intracellular antibody staining.

As shown in Figure 3.17 below, FE-CD5⁺ cells produce IFN γ , TNF α and IL5 after stimulation. Overall, this confirms that the FE-CD5⁺ population belongs to a subset of NKT1 cells. While less is known about NKT1 cells and secretion of TNF α and IL-5, further investigation is needed to determine if *Fgd5*⁻ NKT1 cells express them.

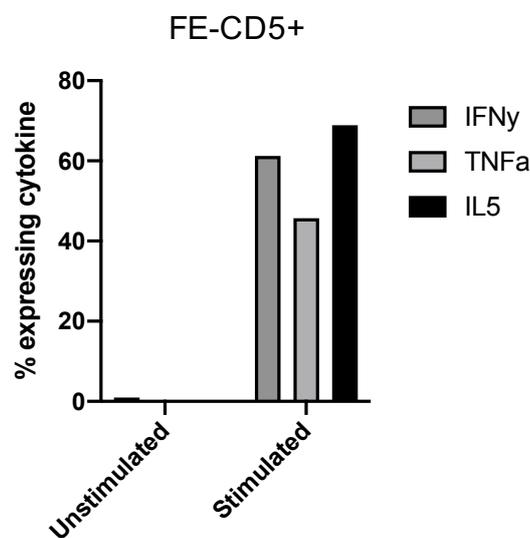


Figure 3.17 Intracellular flow cytometry reveals cytokine profile expressed by FE-CD5⁺ cells.

Percentage of cells expressing IFN- γ , TNF α and IL-5 after 4hr stimulation with 500ng/mL PMA and Ionomycin for FE-CD5⁺ cells Exp=1.

4 *Fgd5* and EPCR mark HSCs *in vivo* and *in vitro*

As shown in Chapter 3, although *Fgd5* expression may not be able to mark HSCs exclusively *in vivo*, it can highly enrich for functional HSCs when combined with EPCR. However, this doesn't necessarily mean that they would be reliable markers of HSCs *in vitro*. As detailed in section 1.4.2, many HSC markers change their expression during culture²⁸⁴. Even though EPCR has been shown to mark human HSCs *in vitro*, this has not been validated in mouse HSCs. Therefore, in this chapter, *Fgd5* and EPCR will be studied for their ability to mark HSCs *in vitro*. The reporter system will also be validated against published conditions that support HSCs.

4.1.1 *Fgd5* and EPCR positive cells highly enriches LSK cells *in vitro*

To test whether *Fgd5* and EPCR might identify HSCs *in vitro*, single ESLAM HSCs were cultured in serum-free medium supplemented with 300ng/mL of SCF and 20ng/mL of IL-11 (SS), conditions that maintain HSC numbers out to 10 days^{102,130}. At day 10, cells were visually inspected and grouped by size (Figure 2.3). Small and medium colonies were then pooled respectively and analysed for expression of HSC markers, such as *Fgd5*, EPCR, Sca-1 and c-Kit and also lineage markers, such as Mac-1 and Gr-1 by flow cytometry (Figure 4.1). Large colonies were individually analysed but none of them had any Lineage⁻, Sca-1⁺ and c-Kit⁺ (LSK) nor *Fgd5* and EPCR positive cells remaining (Figure 4.1). As mentioned in section 1.2, previous work has demonstrated an inverse relationship between proliferation rate and self-renewal potential. Therefore, as expected, smaller colonies retained a higher proportion of LSK cells compared to medium sized clones (48.43 ± SD 17.91 vs 19.07% ± SD 8.879, p=0.488) (Figure 4.1). However, whether in small or medium colonies, the percentage of LSK cells in the *Fgd5*^{high}/EPCR^{high} (F^{hi}E^{hi}) fraction is approximately 50% higher than that from the total clone, demonstrating that F^{hi}E^{hi} cells are enriched for LSK cells (Figure 4.1). In contrast, *Fgd5*^{low}/*EPCR*^{low} (F^{lo}E^{lo}) cells were almost entirely non-LSK (small colonies: 11.55% ± SD 9.542; medium colonies: 5.653% ± SD 4.452), indicating that cells lost *Fgd5* and EPCR expression when they differentiated and expressed markers of mature blood cells.

To further dissect single cell heterogeneity in 10-day cultures, single colonies were analysed individually instead of pooled together. The results confirm that the percentage of FE positivity is highly correlated with the percentage of LSK cells ($r = 0.8779$, $p < 0.0001$) (Figure 4.1). Interestingly, there were colonies that had high LSK percentages, but low FE+ percentages, suggesting that FE+ may be even more selective than LSK *in vitro*.

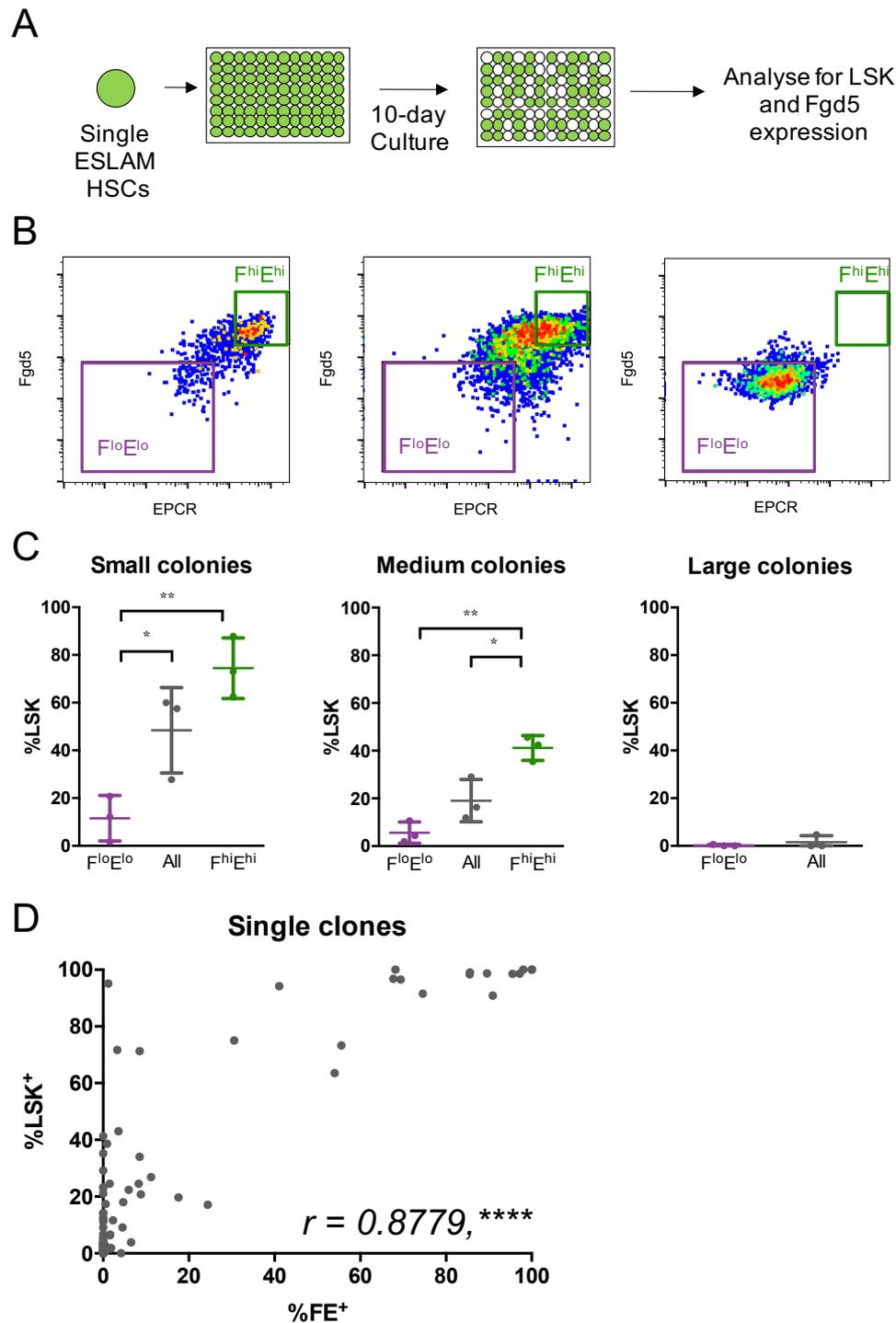


Figure 4.1 *Fgd5* and EPCR enriches for LSK cells in *ex vivo* cultures.

- A) Schematic of experimental design. Single ESLAM cells were cultured for 10 days in SS with 300ng/mL SCF and 20ng/mL IL-11, colonies were then pooled by size and analysed by flow cytometry.
- B) Representative gating layouts for small (left), medium (middle) and large (right) colonies.
- C) Percentage of LSK cells within respective gates of pooled small, medium and large colonies. Tukey's multiple comparisons test. **= $p < 0.005$. Exp = 3. Error bars represent mean \pm SD.
- D) Percentage of FE⁺ cells correlated with percentage of LSK cells in individual colonies. Pearson correlation. ****= $p < 0.0001$. Exp=2; n=90.

4.1.2 *Fgd5* and EPCR together mark transplantable HSCs *in vitro*

While *in vitro* assays are useful, the robustness of *Fgd5* and EPCR as *in vitro* markers of LT-HSC activity requires validation by *in vivo* functional assays. Therefore, 3018 FE+ cells were isolated and cultured in SS for 3 days. By day 3, most cells remained positive for *Fgd5* and EPCR although a range of expression levels was detectable and on average the levels of *Fgd5* correlated well ($r=0.707$) with the expression of EPCR (Figure 4.3). In order to determine which cell fraction retained HSC activity, three populations were isolated and transplanted into irradiated recipient mice: 1) bulk live cells (152 cells) 2) $F^{hi}E^{hi}$ (582 cells) and 3) $F^{lo}E^{lo}$ (1512 cells) (Figure 4.2). Overall, the $F^{hi}E^{hi}$ fraction represented $\sim 7\%$ and the $F^{lo}E^{lo}$ fraction represented $\sim 24\%$ of all cultured cells.

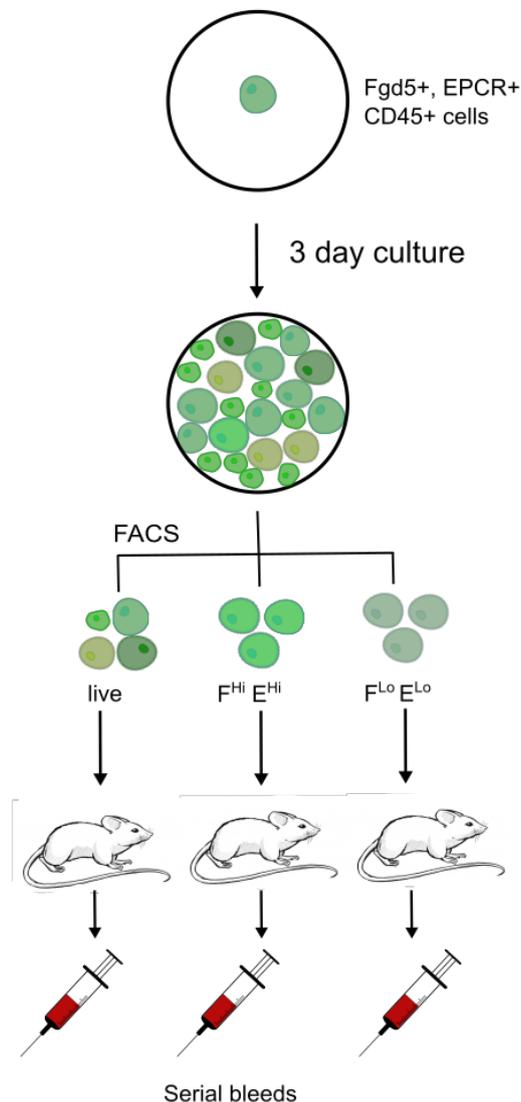


Figure 4.2 Schematic of experimental design to test the ability of *Fgd5* and EPCR to isolate functional HSCs *in vitro*.

FE+ cells were cultured for 3 days in serum free conditions, supplemented with 300ng/mL SCF and 20ng/mL IL-11. At day 3, cells were resorted for *Fgd5*^{hi} EPCR^{hi} ($F^{hi}E^{hi}$) and *Fgd5*^{lo} EPCR^{lo} ($F^{lo}E^{lo}$) cells and transplanted into recipients. 3 mice per condition. Live cells were also sorted and transplanted into 2 recipients.

All transplanted mice showed reconstitution at 16 weeks post-transplantation; however, the mice transplanted with $F^{hi}E^{hi}$ cells showed significantly higher chimerism (Figure 4.3) compared to bulk cells ($p = 0.0059$) and $F^{lo}E^{lo}$ cells ($p = 0.00016$). Additionally, all of the mice injected with $F^{hi}E^{hi}$ donor cells displayed multilineage reconstitution, defined as having $>1\%$ contribution of each GM, B cell and T cell lineages. Strikingly, none of the $F^{lo}E^{lo}$ cells transplanted had multilineage reconstitution, all lacking contribution to the GM lineages, which is strongly associated with secondary reconstitution ability. This, in combination with the fact that more $F^{lo}E^{lo}$ cells (~ 2.5 fold) were transplanted, indicates that *Fgd5* and EPCR expression can robustly identify HSCs in short-term *in vitro* cultures.

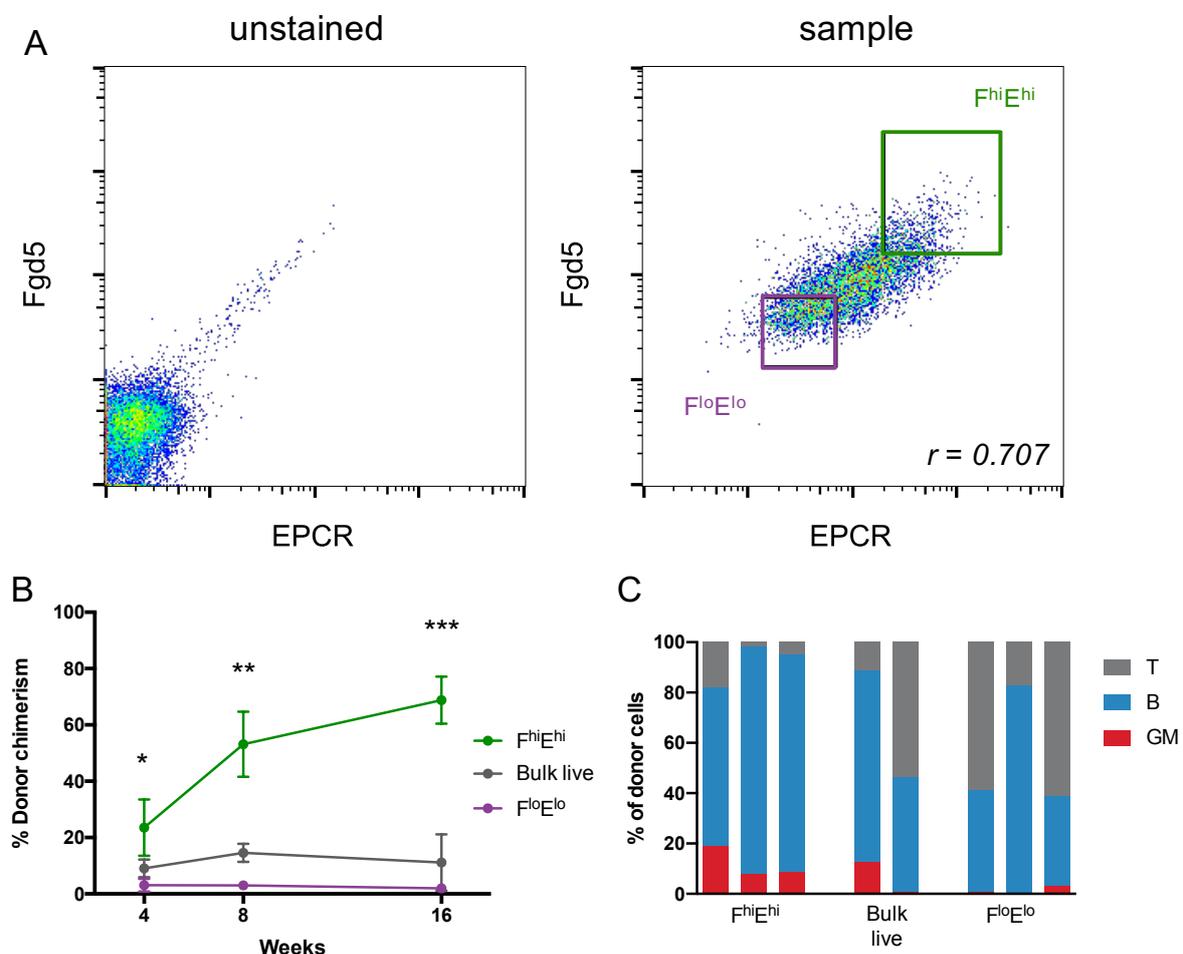


Figure 4.3 *Fgd5* and EPCR marks functional HSCs *ex vivo*.

- A) Gating layout for re-sort. Left panel is the unstained sample and right panel is sorted sample (3d cultured HSCs). Pearson correlation.
- B) Donor chimerism tracked for 16 weeks. Unpaired t-Test between $F^{hi}E^{hi}$ and $F^{lo}E^{lo}$ cells. ***= $p < 0.0005$, Exp=1; n=3.
- C) Lineage output of donor cells, presented as a percentage of donor cells.

Because the bulk live cells should contain $F^{hi}E^{hi}$ cells, it may be seen as slightly unexpected that they had significantly lower chimerism than $F^{hi}E^{hi}$ cells. However, this was likely attributable to the lower equivalent cell dose of $F^{hi}E^{hi}$ cells that was transplanted (i.e., 582 cells would be an ~ 8000 bulk cell starting equivalent as opposed to the 152 transplanted).

4.1.3 Single positive *Fgd5* or EPCR cells contain fewer functional stem cells

Next, *Fgd5* and EPCR were tested as single markers for their ability to enrich for HSCs in culture. Single positive *Fgd5* or single positive EPCR cells were re-sorted after culture for transplantation (Figure 4.4). Notably, a more enriched starting population of ESLAM cells was used to initiate the culture, compared to the FE+ cells in the transplantation experiment described previously (Figure 4.3). In addition, $Fgd5^{+}EPCR^{+}CD150^{-}$ cells (FE+CD150-) were also cultured to validate the two markers on a less pure population. Following 3 days of culture in SS, the FE+CD150- fraction had reduced *Fgd5* and EPCR expression compared to ESLAM cells (Figure 4.5 and Figure 4.4). As expected, the composite of these two plots looks similar to the flow plot in Figure 4.3. From the cultured ESLAM cells, two fractions were resorted, 1) $F^{hi}E^{lo}$ (representing $\sim 9\%$ of the population) and 2) $F^{lo}E^{hi}$ (representing $\sim 12.6\%$ of the population). These single marker^{high} fractions were transplanted into 3 mice each. Only one out of the 3 mice injected with the $F^{lo}E^{hi}$ cells had detectable donor chimerism at 16 weeks (Figure 4.4). Unfortunately, one of the mice injected with the $F^{hi}E^{lo}$ fraction had to be culled after 8 weeks due to significant weight loss. The donor cells in that particular mouse had strong GM contributions at week 8, suggestive that it would also be positive at week 16 (Figure 4.4). Even taking this into account, *Fgd5* and EPCR as single markers would appear to be less robust for isolating HSCs *in vitro*.

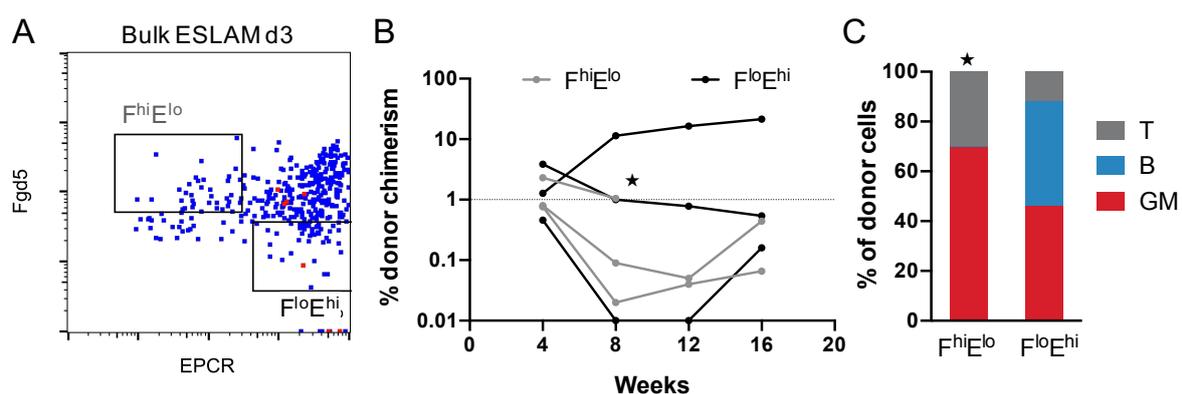


Figure 4.4 Single positive *Fgd5* or EPCR cells contain fewer functional stem cells.

- A) Gating layout for single positive *Fgd5* and EPCR cells.
 B) Donor chimerism tracked for 16 weeks. The star (★) indicates recipient was culled for health reasons. Dotted line indicates where chimerism is 1%. Exp=1 n=3 each.
 C) Lineage output of donor cells for the recipients that had above 1% chimerism for their last timepoint.

From the cultured CD150⁻ fraction, F^{hi}E^{hi} cells (representing 3% of the cells) were sorted and injected into 2 mice to test if *Fgd5* and EPCR marked HSCs in a much less pure population of cells (Figure 4.5). Interestingly, both mice were positive for donor cells at week 16, and both were multilineage (Figure 4.5). This is strong evidence that, regardless of starting population, *Fgd5* and EPCR together are strong distinguishers of long-term repopulating cells *in vitro*. Overall, as EPCR and *Fgd5* expression is highly correlated, it seems possible to use EPCR on WT cells as a surrogate readout for *Fgd5* expression. However, it still seems beneficial to include *Fgd5* to make gating a bit easier, as well as remove single positive EPCR cells.

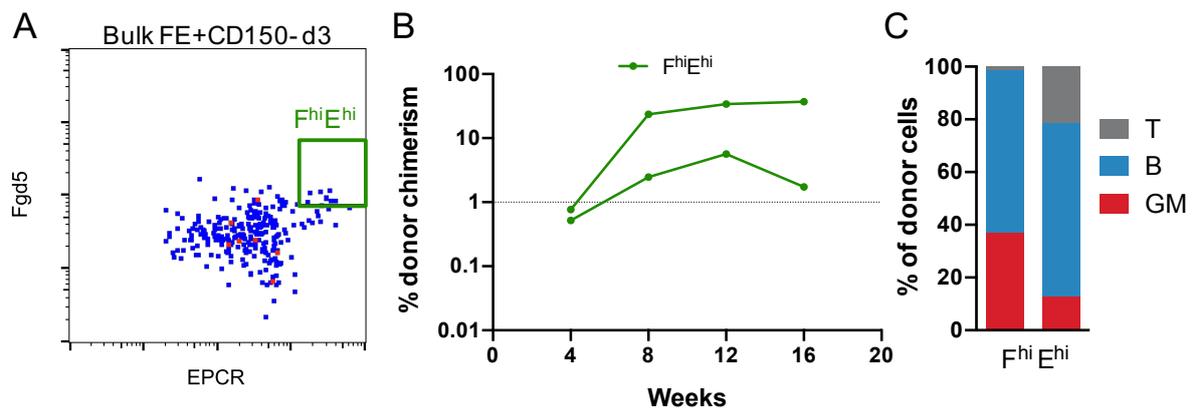


Figure 4.5 *Fgd5* and EPCR mark functional stem cells even in cultures with FE+CD150-starting cells.

- A) Gating layout for F^{hi}E^{hi} cells in 3-day cultures of FE+CD150⁻ cells.
 B) Donor chimerism tracked for 16 weeks. Dotted line indicates where chimerism is 1%.
 C) Lineage output of donor cells at week 16, as a percentage of donor cells.
 Exp=1 n=2.

4.2 *Fgd5* also marks FL HSCs

The ESLAM HSC isolation strategy is known to be able to isolate HSCs from all developmental stages, including E14.5 FL¹⁸⁵. Whilst *Fgd5* expression has been shown to mark quiescent adult HSCs in the BM, in order for it to be a robust marker *ex vivo*, it must also mark cycling and expanding HSCs. In order to test this, FL HSCs, which are actively cycling (section 1.2.2), were analysed for the expression of *Fgd5*. As shown in Figure 4.6, the FL contains *Fgd5*⁺ cells and indeed all FL ESLAM HSCs are *Fgd5*⁺. Interestingly, there is a significant population of *Fgd5*⁺ cells that are not EPCR⁺, resembling the adult BM. Overall this demonstrates that *Fgd5* can mark cycling HSCs, which is a prerequisite for it being an *in vitro* marker for expanding HSCs.

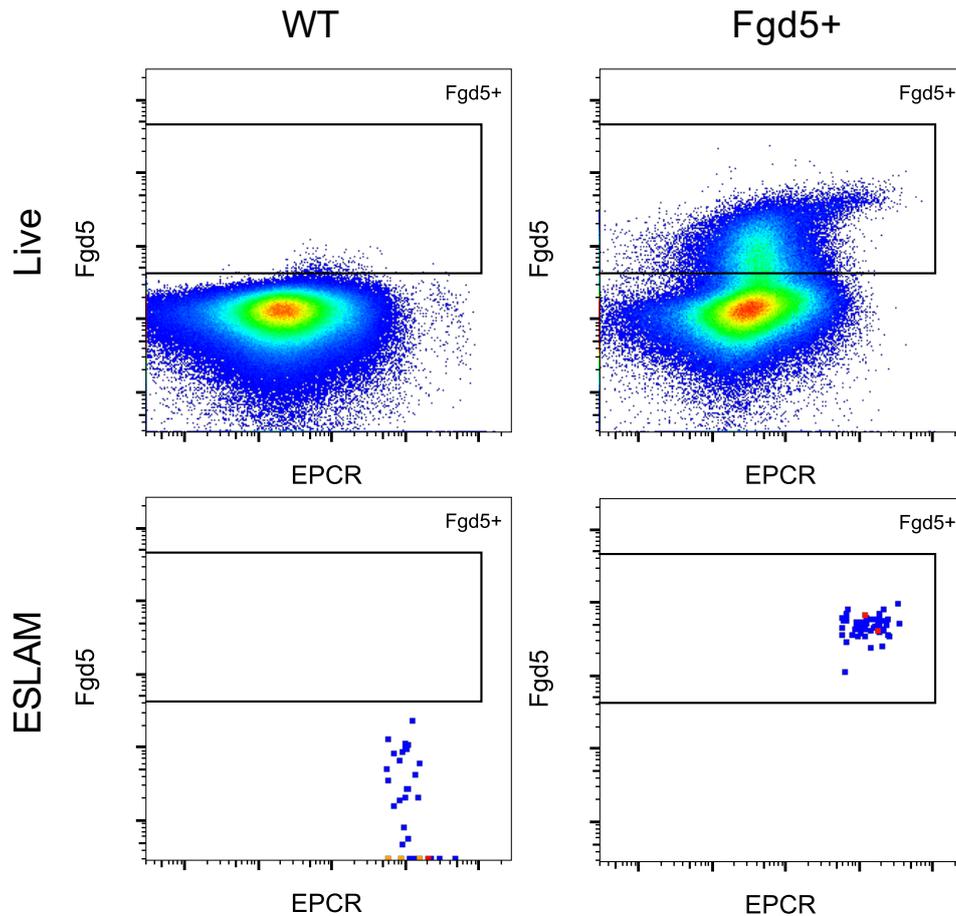


Figure 4.6 Foetal Liver cycling HSCs are $Fgd5^+$.

$Fgd5$ expression levels in WT (left) and $Fgd5^+$ (right) FL cells, clearly showing that all phenotypic HSCs in the FL are $Fgd5^+$. Top panel is gated for live cells and bottom panel is gated for ESLAM cells, showing clearly where $Fgd5$ expression begins. Showing concatenated data from 4 FL for each group.

4.3 Novel reporter strategy validates HSC supportive culture conditions and identifies novel targets for supporting HSC expansion

Having confirmed that $Fgd5$ and EPCR can both mark cycling HSCs, we next applied this *in vitro* reporter strategy to test existing cell culture conditions for HSC expansion and screen for potentially new supportive factors. As mentioned in section 1.4.8, co-culturing HSCs with supportive stromal cells has been a promising strategy for maintaining and supporting HSCs. In particular, the EL08 cell line has been shown to support HSCs *ex vivo*, through non-physical contact mechanisms³³⁵.

We therefore tested conditioned medium (CM) from EL08 for its ability to produce $Fgd5^+EPCR^+$ HSCs in culture, and for molecules within it that support HSCs. Single LT-HSCs were cultured in medium conditioned by EL08 cells for a varying length of time (1, 3, 10 or 14 days). As a control,

cells were also cultured in SS (as in section 4.1.2) without any CM added¹⁰². We hypothesised that the different lengths of conditioning time would support HSCs differently and by comparing the secretomes of these varying conditioning times, the individual factors involved in supporting HSC expansion might be identified. At day 10 after start of culture, the clones were harvested and individually analysed by flow cytometry. Whilst the previous experiments have established that *Fgd5* and EPCR in combination can provide a robust two-colour isolation strategy for HSCs *in vivo* and *in vitro*, we also added antibodies for Sca-1, c-Kit and lineage markers (Gr-1 and Mac-1) to further enhance our ability to distinguish HSC-containing clones. Defining the HSC phenotype as *Fgd5*⁺, EPCR⁺, lineage⁻, Sca-1⁺ and c-Kit⁺ (FELSK), we compared phenotypic HSC content across conditions.

4.3.1 EL08 CM improves survival and increases proliferation of HSCs

As shown in

Figure 4.7, HSCs cultured in CM had a significant survival advantage ($p=0.0012$ for the least significant comparison with media conditioned for 14 days) compared to non-conditioned SS control irrespective of the length of conditioning time. The average survival rate in SS control was $26.33\% \pm \text{SD } 8.083$ whereas all of the CM averaged above 80%, with decreasing survival as length of conditioning increases. EL08 CM_{1D} (media conditioned for 1 day), had the highest survival rates of $96\% \pm \text{SD } 2.828\%$. With respect to proliferation, HSCs cultured in EL08 CM_{1D} gave rise to the largest clones and were significantly larger ($p=0.026$) compared to CM_{3D} (Figure 4.7). However, the difference in clone sizes were not significant in all other comparisons.

4.3.2 Short-term but not extended conditioning of media with EL08 cells support HSCs

Consistent with previous findings, when investigating the proportion of FELSK cells compared to the respective clone sizes, larger clones had a smaller proportion of phenotypic HSCs ($r=-0.3109$, $p=0.0124$ for SS control) (Figure 4.7). Although survival was high, medium conditioned for 10 and 14 days were not significantly different from SS control in terms of average percentage of FELSK cells. Of note, the average percentage of phenotypic HSCs was the highest in clones cultured in CM_{3D} ($40.48\% \pm \text{SD } 36.18$). This is significantly higher than all other CM and the SS control ($p<0.0001$). Overall this suggests that the optimal duration for generating EL08 conditioned medium is 3 days and that prolonged conditioning is actually detrimental to HSC maintenance, perhaps due to the accumulation of secreted factors that stimulate proliferation and differentiation.

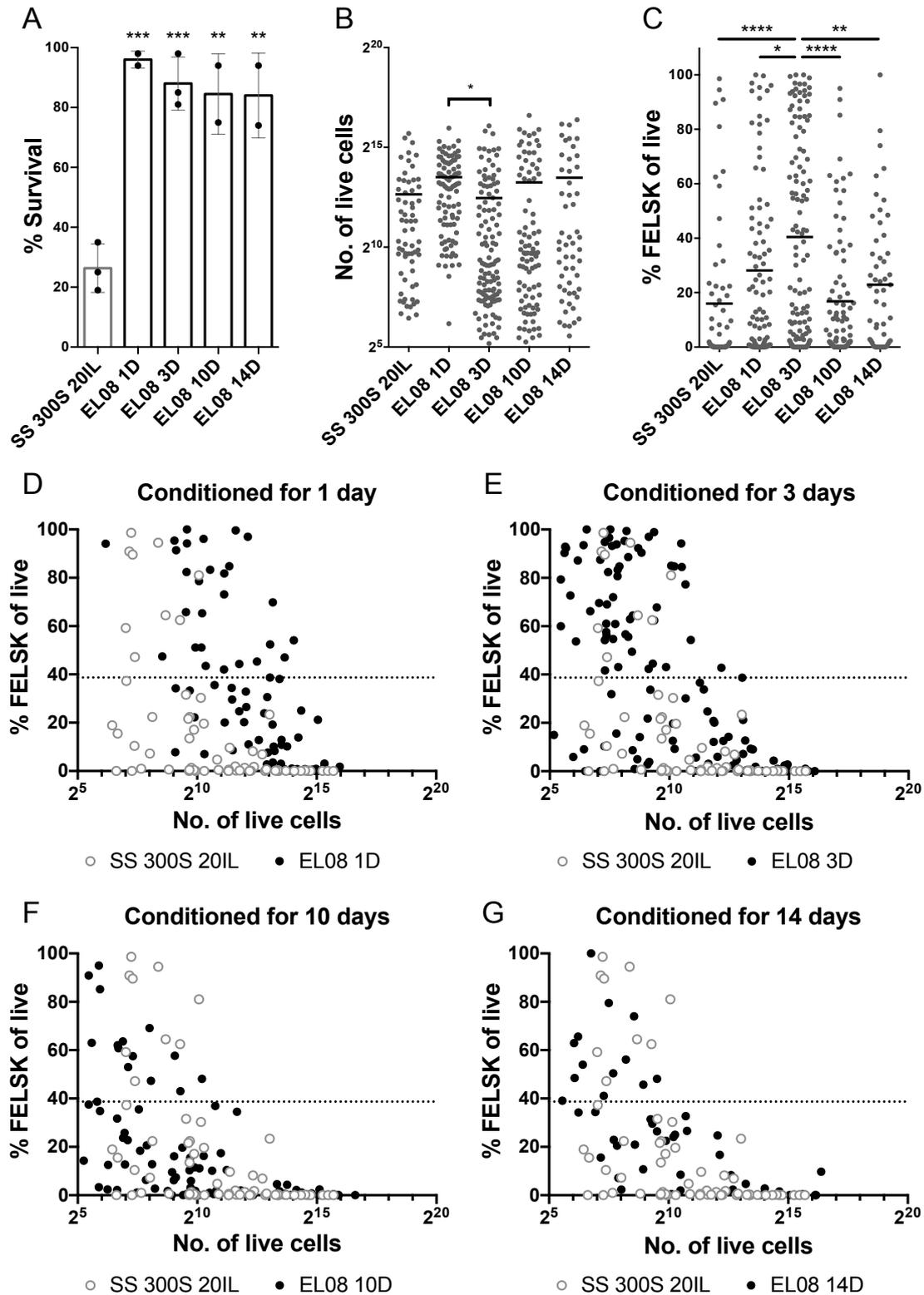


Figure 4.7 Duration of conditioning alters beneficial effects of EL08 CM.

A) Clonal survival at day 10, out of total cells plated. One-way ANOVA. **= $p < 0.01$. ***= $p < 0.001$.

B) Mean clone sizes of cultures. One-way ANOVA. **= $p < 0.01$. *= $p < 0.05$.

C) Mean percentage of FELSK in individual clones. One-way ANOVA. ****= $p < 0.0001$. **= $p < 0.01$. *= $p < 0.05$.

D-G) Clonal outcomes of cells cultured in SS 300S 20IL (n=271); EL08 1D (n=96); EL08 3D (n=192); EL08 10D (n=144); EL08 14D (n=144).

It is possible that the averaging of the phenotypic HSC content of clones might hide differences between individual clones (i.e., it is possible that only clones with the very highest percentages of FELSK cells contain HSCs). If so, it would be important to look at the proportion of clones generated that are nearly entirely comprised of phenotypic HSCs. As a first pass, we used the average %FELSK of 10-day clones with fewer than 500 cells (38.7% in base SS conditions (Figure 4.7)) as an arbitrary cut-off, since small clones have been previously demonstrated to have the most HSC activity. This percentage changes slightly during each experiment because of technical variations such as cytometer and gate settings, hence the cut-off point was recalculated for each experiment with the control SS conditions. The proportion of clones above this cut-off point was then compared by Fisher's exact test to determine whether the proportion of clones with a high proportion of phenotypic HSCs was different compared to the SS control. As expected, proportions of phenotypic HSCs created using CM_{3D} was significantly higher than SS control ($p < 0.0001$) whilst CM_{10D} and CM_{14D} were not (Table 4.1). Interestingly, although 1-day clones are larger in size, there remains a significantly larger proportion of clones above the cut-off compared to SS control ($p = 0.0332$) (Table 4.1). Overall, this reaffirms that EL08 CM conditioned for short periods of time (1-3 days) supports HSCs and that extended period of conditioning is detrimental.

Table 4.1 Tally of clones above and below FELSK cut-off and Fisher's exact test result.

	SS 300S 20IL	EL08 1D	EL08 3D	EL08 10D	EL08 14D
Above 38.7%	9	27	57	14	12
Below 38.7%	55	65	62	77	42
p-value vs SS		0.0332	<0.0001	1	0.3346

4.3.3 Proteomic analysis of EL08 CM identifies self-renewal regulators previously discovered by gene expression studies and other novel targets

Whilst previous studies have identified certain genes that are highly expressed in EL08 cells, they have used gene expression assays (microarray), which do not actually measure the active functional protein that is secreted by the cells. In order to identify the proteins that support HSC maintenance, CM from the time course experiment was analysed by mass spectrometry (Label-free quantification)⁴³¹. With label-free methods, the presence and absence of proteins may not be as reliable because peptide signals may be hidden by the high dynamic range created by media proteins such as serum albumin⁴³². Therefore, we focused our analysis on 214 abundant proteins identified with high confidence (at least 2 unique peptides) across all samples. As shown by the heatmap in Figure 4.8, there is an overall increase in total protein content at the global level as the

length of conditioning increases with very few proteins present in higher quantities in earlier timepoints compared with later timepoints. Since CM_{D3} outperformed both an earlier and both later timepoints in terms of functional output of phenotypic HSCs, we hypothesised that there would be a combination of stimulatory and inhibitory factors of HSC self-renewal. However, since the majority of proteins increased in concentration over the course of conditioning, it was difficult to pinpoint targets for inhibitory factors.

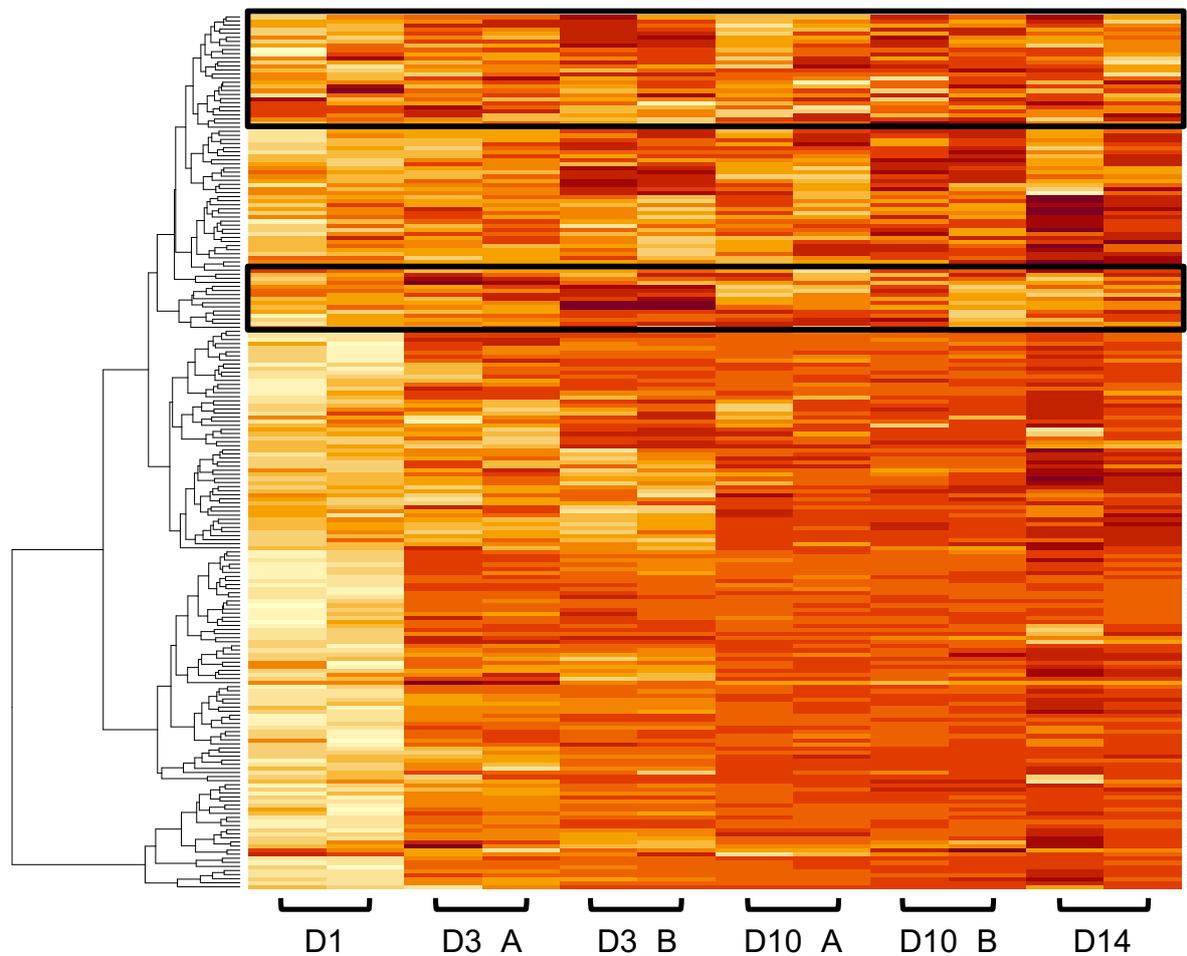


Figure 4.8 Heatmap of proteins identified within EL08 CM suggests a general increase in protein content as conditioning increases.

Heatmap of 214 proteins identified with more than 1 unique peptide and filtered out potential contaminants. Each column represents an EL08 CM sample, technical repeats are displayed alongside each other, marked by the brackets. Different batches of EL08 $CM_{3D/10D}$ are denoted by A and B. Key: darker red means higher LFQ relative to other, i.e. the colours are scaled by row. The two black boxes are added to highlight the relatively few proteins that are more abundant in $CM_{1D/3D}$ (earlier timepoints).

As highlighted by the proteins within the black boxes in Figure 4.8, there were a small cluster of proteins that were present in higher quantities in $CM_{D1/D3}$ compared to $CM_{D10/D14}$. Most of these proteins were of little interest because of their limited ability to be signalling proteins for other cells, such as 40S ribosomal protein S20 or beta-actin. Interestingly, haemopexin (Hpx) was one

of the proteins that display this quantification pattern (Figure 4.8 and Figure 4.9) and has been recently identified to promote HSC maintenance⁴¹⁷. Of interest as well is zyxin and caldesmon, which both show a similar pattern (Figure 4.8 and Figure 4.9). Zyxin is a zinc binding phosphoprotein that concentrates at focal adhesions, with a proline-rich domain that may interact with SH3 domains and spur signal transduction⁴³³. Caldesmon is a calmodulin binding protein⁴³⁴, which may be interesting considering the recent studies implication calcium regulation and HSC maintenance³⁸⁵. However, as shown in Figure 4.9, in general these proteins are less abundant compared to other proteins, which we turned our focus to.

As detailed in Figure 4.9, many previously reported upregulated genes, identified by microarray analysis, are present, including OPN, PTN, IGFBP4, IGFBP6, thrombospondin-1, thrombospondin-2 and a disintegrin and metalloproteinase with thrombospondin motifs 1 (Adamts1)^{131,335}. Also of note are a number of extracellular matrix (ECM) proteins, such as Collagen and Nidogen (Figure 4.9). In fact, Collagen α 1(I) chain had the highest LFQ value consistently across all samples; not to mention the many other collagen chains identified within the dataset (data not shown), suggesting a role for ECM proteins that are increasingly becoming appreciated in HSC biology⁴³⁵.

OPN has been previously studied as a component of the endosteal niche, and has been implicated in HSC localisation as well as negative regulation of HSC proliferation¹¹⁵. PTN has also been implicated in multiple studies as a promoter of self-renewal *in vitro* for both mouse and human HSCs^{131,132,436,437}. Whilst IGFBP4 or IGFBP6 has not been studied in the context of HSC self-renewal, IGFBP2 has been reported in multiple studies to be a supporter of HSC maintenance and expansion^{371,403,438}. Therefore, we carried forward with OPN, PTN and IGFBP2 and IGFBP4 as initial factors to validate for their individual functional effect on promoting HSC growth.

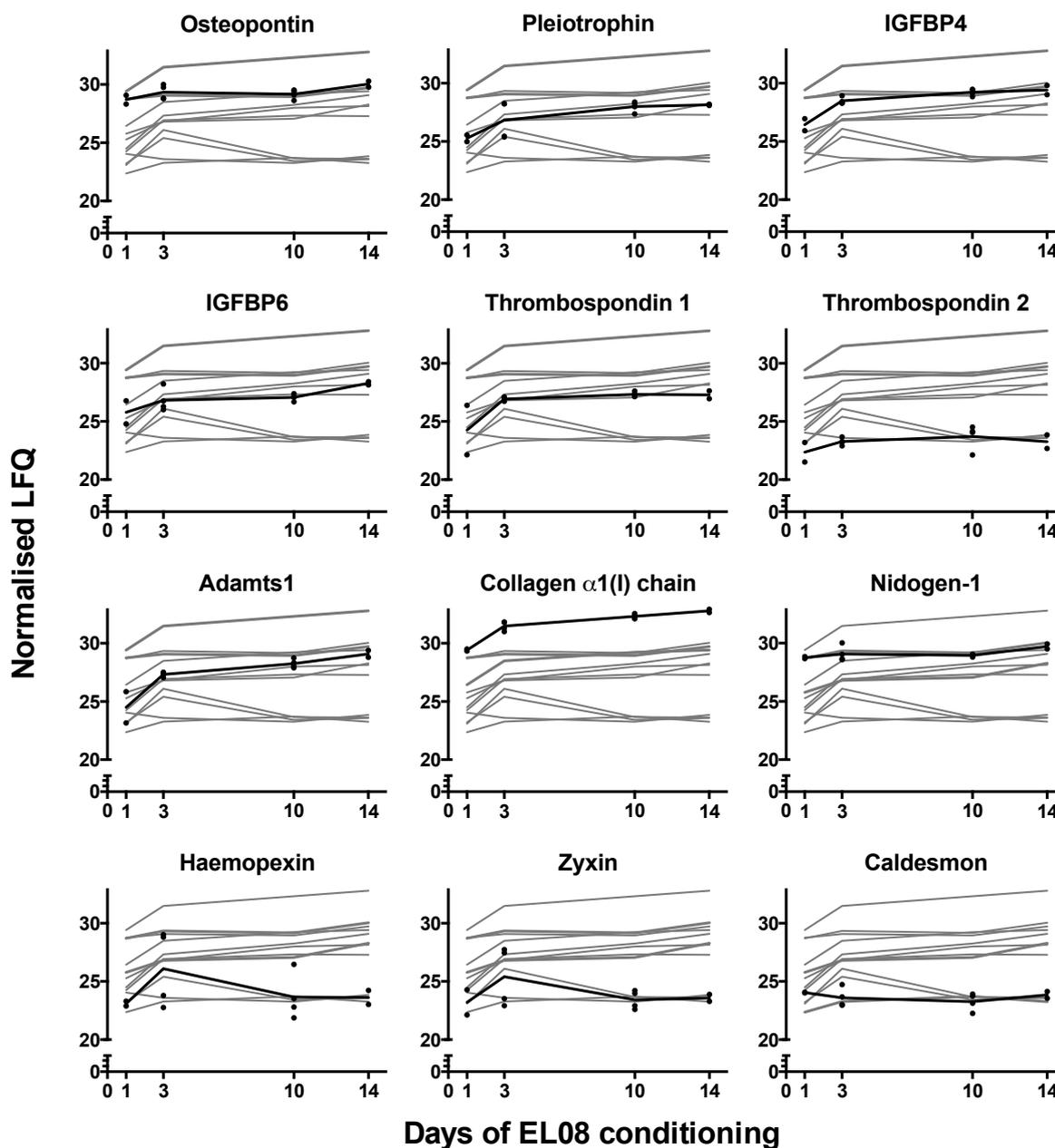


Figure 4.9 Label-free quantification of EL08 CM.

Normalised LFQ of a manually curated list of 12 proteins of interest, over the course of conditioning. The black line indicates the mean LFQ of the protein of interest titled in each graph. The grey lines represent the other 11 proteins in this matrix. Each dot represents the normalised LFQ of each run of CM sample, including technical repeats and different batches of CM as in Figure 4.8.

4.3.4 Recombinant proteins have no effect on the survival and clone sizes of HSCs

To test their impact on promoting HSC production, single HSCs were cultured in the above SS-based cultures for 10 days with or without the individual factors. Clones were grown and then harvested for flow cytometry analysis to assess the proportion of phenotypic HSCs. The concentrations tested were determined through literature searches on previously performed *in vitro* experiments^{116,371,436} and accompanied by a lower concentration to gain some insight as to whether

there was a dose response. For one pair of factors, OPN and PTN, we combined them to determine whether there would be any additive or synergistic effects. None of the factors tested, at high or low concentrations, offered a significant survival benefit to HSCs in 10-day cultures (Figure 4.10A). Interestingly, none of the factors significantly influenced the average clone size either (Figure 4.10C), suggesting that other factors in EL08 CM are responsible for the observed increases in survival and proliferation.

4.3.5 Pleiotrophin supports the expansion of phenotypic HSCs at low concentrations

In terms of the average FELSK content of the individual clones, OPN, IFGBP2 and IGFBP4 exhibited no effect compared to SS control. However, cultures containing 100ng/mL of PTN were significantly better than SS control (vs. 100ng/mL PTN, $p=0.0327$; vs. 100ng/mL OPN and 100ng/mL PTN, $p=0.0077$). Interestingly, the effect was not seen in cultures with 400ng/mL of PTN, suggesting that higher concentrations are not beneficial to HSCs.

As in section 4.3.2, the proportions of clones above and below the average FELSK content of SS control small clones (recalculated in these experiments to be 53.1%) were compared by Fisher's exact test. As expected, the proportions of clones above this cut-off were significantly higher in 100ng/mL PTN ($p=0.0164$) and 100ng/mL OPN/PTN ($p=0.0174$) (Table 4.2). The proportions were significantly lower compared to control when 800ng/mL of OPN and 400ng/mL of PTN were added, suggesting that high concentrations of OPN and PTN combined lead to a detrimental effect on HSC maintenance. As the overall protein quantities increased during extended conditioning (see section 0), this could be one reason why EL08 medium cultured for a longer duration is detrimental to HSC maintenance.

Based on these preliminary results, more investigation would be required to fully elucidate the role of PTN at low concentrations for HSC maintenance. It is further important to note that some of the experiments described here have not been repeated sufficiently to determine if their effects (or lack thereof) are reproducible.

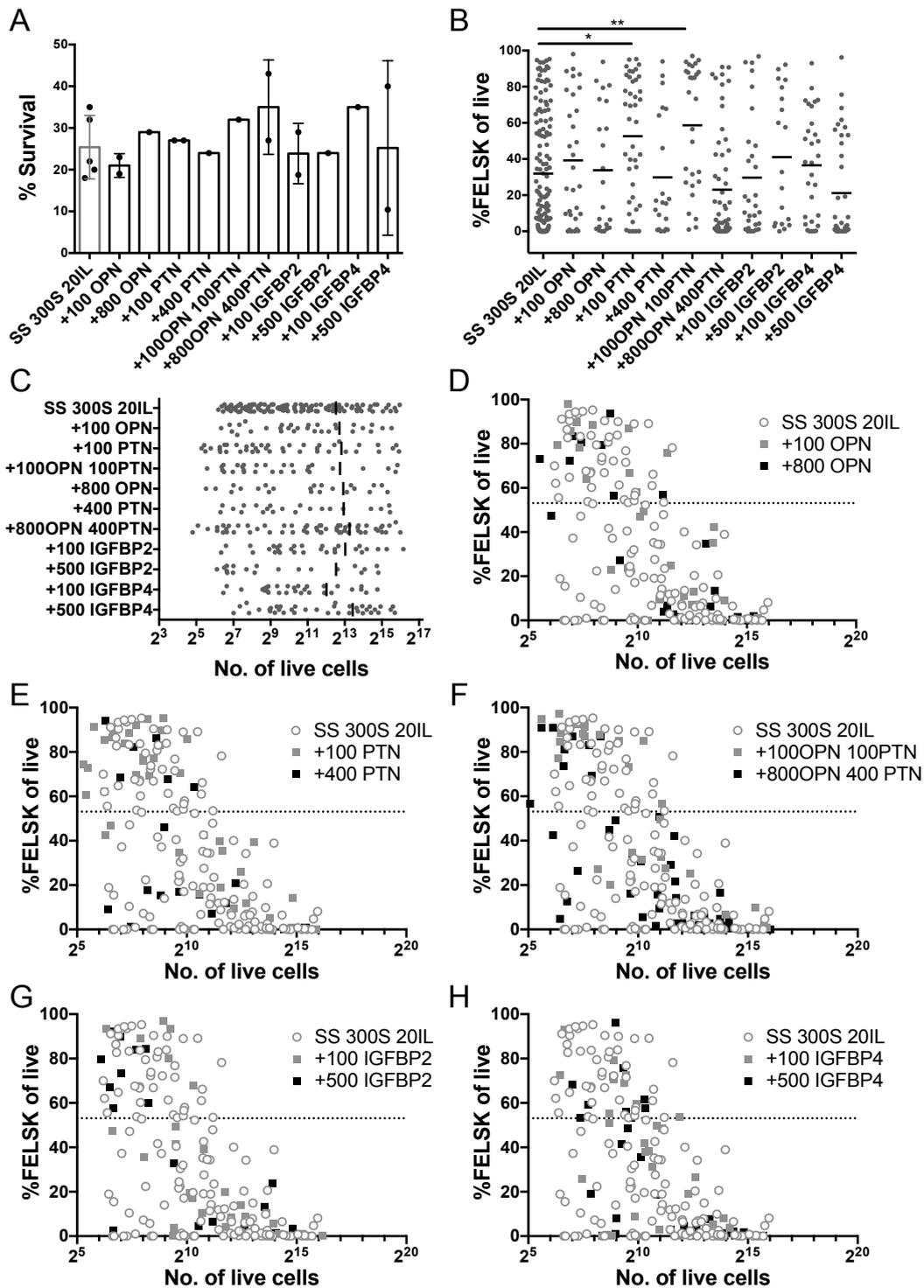


Figure 4.10 Pleiotrophin increases phenotypic HSCs in 10-day cultures.

A) Day 10 survival rates of single HSCs cultured in SS with 300ng/mL SCF and 20ng/mL IL-11 supplemented with indicated factors and concentrations (ng/mL). One-way ANOVA between conditions with at least two repeats, compared to control.

B) Mean percentage of FELSK in individual clones. One-way ANOVA. **= $p < 0.01$. *= $p < 0.05$.

C) Mean clone sizes of cultures. One-way ANOVA. Not significant.

D-H) Clonal outcomes of cells cultured in SS 300S 20IL (n=843); 100 OPN (n=192); 800OPN (n=96); 100PTN (n=96); 100 PTN (n=192); 400 PTN (n=96);100 OPN 100PTN (n=96); 800 OPN 400 PTN (n= 192); 100 IGFBP2 (n=144); 500 IGFBP2 (n=96); 100 IGFBP4 (n=96); 500 IGFBP4 (n=144).

Table 4.2 Tally of clones above and below FELSK cut-off and Fisher's exact test result.

	SS 300S 20IL	+100 OPN	+100 PTN	+100OPN 100PTN	+800 OPN	+400 PTN	+800OPN 400PTN	+100 IGFBP2	+500 IGFBP2	+100 IGFBP4	+500 IGFBP4
Above 53.1%	39	11	22	16	8	6	10	6	9	10	9
Below 53.1%	82	19	19	12	14	15	47	23	10	20	28
P-value vs SS		0.6684	0.0164	0.0174	0.8058	0.8058	0.0482	0.2653	0.2048	1	0.4185

4.4 F12 HSA cultures increase survival but do not support HSCs unless IL-11 is added

As we were performing the above studies, substantial advances were being made in the field of murine HSC expansion by the Yamazaki group in Tokyo. The first of these studies reported an all recombinant culture system that could replace BSA with recombinant HSA⁴¹⁷. As detailed in section 1.4.9, the authors argued that contaminants within BSA drive variability in HSC cultures, and that this was problematic for the systematic study of HSC self-renewal factors. Instead of pressing on with our culture system described above, we initiated a collaboration with the authors to utilise our *Fgd5* EPCR reporter strategy to study their culture conditions. During this time, they had also already begun work that was ultimately published in a landmark study in 2019²⁸³ and were able to share these conditions with us prior to publication. As mentioned in section 1.4.9, this study reported an ~200-900-fold expansion of HSCs by replacing HSA with PVA and titrating the amounts of SCF and TPO in the culture condition. During the course of the experiments described below, we began with HSA based conditions and eventually evolved into PVA based conditions to rapidly adopt the most optimal system for expanding HSCs.

As published in their latter study, in HSA based cultures, low SCF (10ng/mL) was found to synergise with high TPO (100ng/mL) to support HSCs in 7-day cultures. Apart from the differences in cytokine concentrations, this culture system also utilised a different base media – F12 Ham's nutrient mix (F12), supplemented common media additives such as ITSX, Penicillin/Streptomycin and L-glutamine. If F12-based cultures are superior to SS-based cultures at maintaining HSCs, it would be important to switch to using it as a base medium for testing self-renewal factors as above. Therefore, we compared the F12-based culture system to our initial SS culture system described above (section 4.3).

4.4.1 F12-based cultures have higher survival compared to SS-based cultures

In order to compare the two conditions, single HSCs were cultured for 10 days in either SS (300ng/mL SCF and 20 ng/mL IL-11) or F12 (HSA) with varying SCF and TPO concentrations as published. Although Wilkinson et al, has shown that cytokine responsiveness varies depending on the batch of BSA used, they did not investigate the effects of IL-11 on cultures with recombinant HSA. As mentioned in section 1.4.7, IL-11 has been shown to be crucial for HSC self-renewal in SS-based media. Therefore, IL-11 was added to investigate its effect on F12 HSA based cultures. First, we compared the survival rates of the cultures. After 10 days, irrespective of the cytokine concentration, F12-based cultures had dramatically higher survival rates ($p < 0.0001$) compared to SS-based cultures ($85.57\% \pm \text{SD } 15.55\%$ vs $25.67\% \pm \text{SD } 5.123\%$) (Figure 4.11), suggesting that HSCs generally favour these conditions.

4.4.2 F12-based cultures create smaller clones due to lower SCF concentrations

Next we compared the differences in proliferation between SS and F12-based cultures. In terms of the colony sizes at day 10, F12-based cultures in general had smaller colony sizes, though the difference was not statistically significant (Figure 4.11B). Adding IL-11 to F12-based cultures had no significant effect on clone sizes (Figure 4.11B), which is expected because it is not a known mitogen. Rather one would expect that mitogens such as SCF and TPO would have a significant effect on the clone sizes. Interestingly the effects of SCF on colony size is seemingly greater than the effects of TPO. Within F12-based cultures, cells cultured in 10ng/mL SCF generated significantly more small colonies than cells cultured in 100ng/mL SCF, irrespective of the TPO concentration (Figure 4.11B). Likewise, the cells cultured in 100ng/mL of SCF generated significantly more large colonies at day 10 compared to cells cultured in 10ng/mL of SCF, regardless of TPO concentration (Figure 4.11B). In general, because smaller colonies are associated with higher self-renewal capability, this would be consistent with the idea that higher SCF concentrations are detrimental to HSC self-renewal.

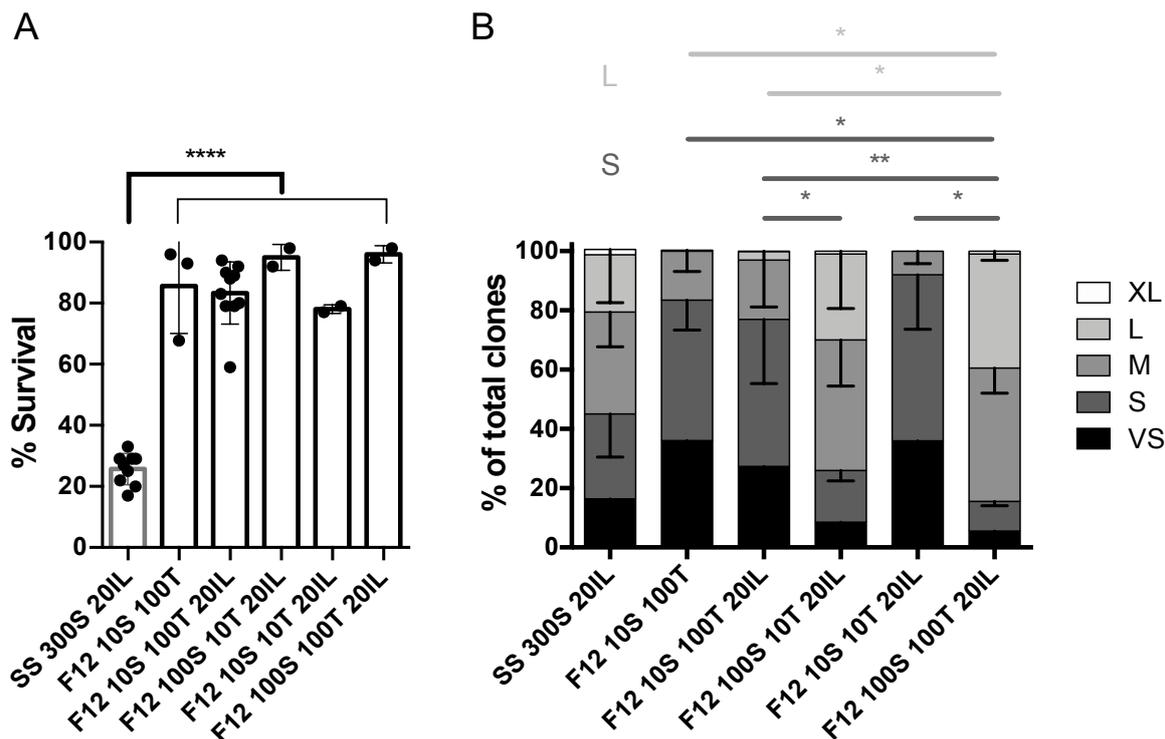


Figure 4.11 F12-based cultures increases survival of HSCs *in vitro*.

- A) Day 10 survival rates of single HSCs cultured in SS-based media (300ng/mL SCF and 20ng/mL IL-11) and F12-based media with various cytokine concentrations. One-way ANOVA. ****= $p < 0.0001$.
- B) Colony sizes of single HSCs clones cultured in above conditions at day 10. Colony sizes are categorised as XL (Extra large), L (Large), M (Medium), S (Small) and VS (Very small) (Figure 2.3). Two-way ANOVA. Tukey's multiple comparison's test. *= $p < 0.05$; **= $p < 0.01$.

SS 300S 20IL, n=864; F12 10S 100T, n=288; F12 10S 100T 20IL, n=960; Rest, n=192.

4.4.3 F12 HSA cultures do not support phenotypic HSCs unless supplemented with IL-11

Next, SS-based cultures were compared to F12-based cultures in terms of the FELSK percentages of its clones. As before, the clones of 10-day single cell cultures were analysed individually by flow cytometry for FELSK content. Interestingly, although not significant in the above categorical clone size analysis; when analysed individually, the average clone size in SS-based cultures ($5730 \pm \text{SD } 5183$), was significantly higher than that of F12 cultures with IL-11 ($613.5 \pm \text{SD } 840.3$; $p < 0.0001$) or without ($720 \pm \text{SD } 1651$; $p < 0.0001$) (Figure 4.12). This is not surprising as SS-based cultures contain much higher concentrations of SCF.

The average FELSK content of F12 cultures with IL-11 were not different ($p = 0.9864$) compared to cultures from SS-based culture (Figure 4.12). However, F12 with IL-11 was significantly higher FELSK content ($p = 0.0094$) than cultures without IL-11 (Figure 4.12). As in section 4.3.2, the average %FELSK of SS control small clones was recalculated based on these experiments to be 60%. Using the cut-off to categorise the clones, as expected, the proportion of clones above this

cut-off was significantly higher in F12 cultures with IL-11 compared to without IL-11 ($p=0.0497$) (Table 4.3). Interestingly the proportional differences were also significant between SS control and F12 cultures without IL-11 ($p=0.0156$) (Table 4.3). Notably, the SS control was not proportionally different compared to F12 with IL-11 ($p=0.555$) (Table 4.3). This suggests that IL-11, like in SS-based cultures, is beneficial to HSC self-renewal in F12-based cultures. However, this experiment was unable to determine whether F12 cultures with IL-11 would be better than SS-based cultures. Considering that F12-based cultures has significantly higher survival, it may represent an improvement compared to SS-based cultures. Therefore, to definitively compare SS-based cultures with F12-based cultures, transplantation experiments were performed and detailed in the next section.

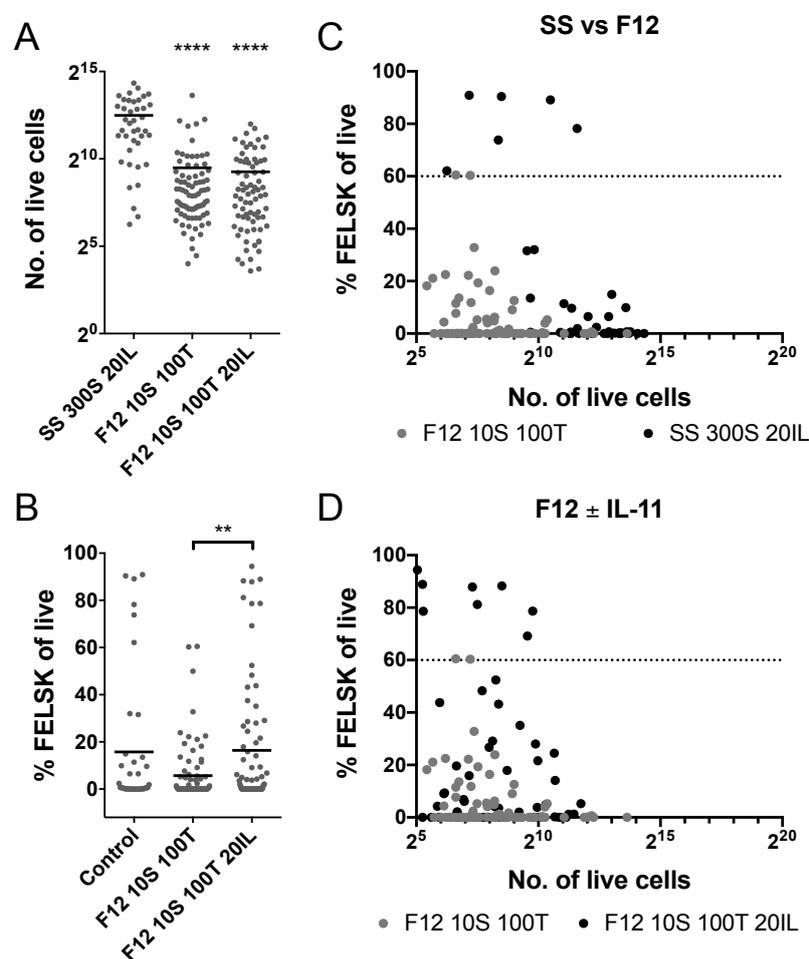


Figure 4.12 F12 cultures supplemented with IL-11 is comparable to SS-based culture.

Single HSCs cultured in either SS-based media or F12-based media for 10 days. SS-based media is supplemented with 300ng/mL SCF and 20ng/mL IL-11. F12-based media is supplemented with 10ng/mL SCF, 100ng/mL TPO and with or without 20ng/mL of IL-11 as indicated. For all conditions: Exp=2 n=192.

A) 10-day clone sizes. One-way ANOVA. ****= $p<0.0001$.

B) Percentage FELSK out of live cells in each clone from each condition. One-way ANOVA. **= $p<0.01$.

C-D) Relationship between clone size and percentage FELSK in C) SS vs. F12 w/o IL-11 and D) F12 ± IL-11. Dashed line indicates the average FELSK percentage in small colonies derived from SS-based cultures (60%).

Table 4.3 Tally of clones above and below FELSK cut-off and Fisher's exact test result.

	SS Control	F12 HSA	F12 HSA IL-11
Above 60%	6	2	8
Below 60%	34	79	67
P-value vs SS		0.0156	0.555
P-value vs F12 HSA			0.0497

4.4.4 Only F12-based cultures can maintain HSCs for 28 days

To fully understand if F12-based cultures retained/produced more HSCs than SS-based cultures, the two cultures were tested in a transplantation setting. Previous studies have shown that SS-based cultures could maintain HSCs up to 10 days and Wilkinson et al. has shown that F12-based cultures could maintain HSCs for up to 28 days²⁸³. Since the longer culture period has a more dramatic expansion of HSC number, we compared sets of 50 HSCs (n=3) cultured for 28 days in either F12 (10ng/mL SCF, 100ng/mL TPO and 20ng/mL IL-11) or SS (300ng/mL SCF and 20ng/mL IL-11). On day 28, the progenies of these cultures were harvested and 90% of the content of was transplanted into a recipient with the remaining 10% analysed by flow cytometry for FELSK content.

As shown in Figure 4.13A-B, the resulting cultures do not differ dramatically by eye at 28 days with respect to cell death and differentiation. Interestingly, the F12-based culture produced qualitatively more “large” cells (Figure 4.13B) which we assumed to be megakaryocytes and likely the result of TPO addition. When phenotypic HSC content was assessed however, none of the SS cultures had any remaining FELSK cells at 28 days compared to consistent proportions of FELSK cells in F12 cultures. Upon transplantation, this phenotypic data was verified with no detectable stem cell activity left in SS cultures, compared to 2 of 3 F12-based cultures with >1% chimerism at week 16. Of note, the culture with the higher FELSK percentage (13.8%) gave rise to the highest donor chimerism when transplanted (20.5% at week 16), providing further validation of our novel *in vitro* HSC reporter strategy. While this did not match the level of expansion reported in the Wilkinson et al. paper, it did suggest that F12-based cultures were better at maintaining HSCs *ex vivo* and prompted us to replace SS as the base medium for all future experiments. Indeed, with technical improvements, such as better media exchanges and the switch from HSA to PVA, the expansion level became closer to reported figures.

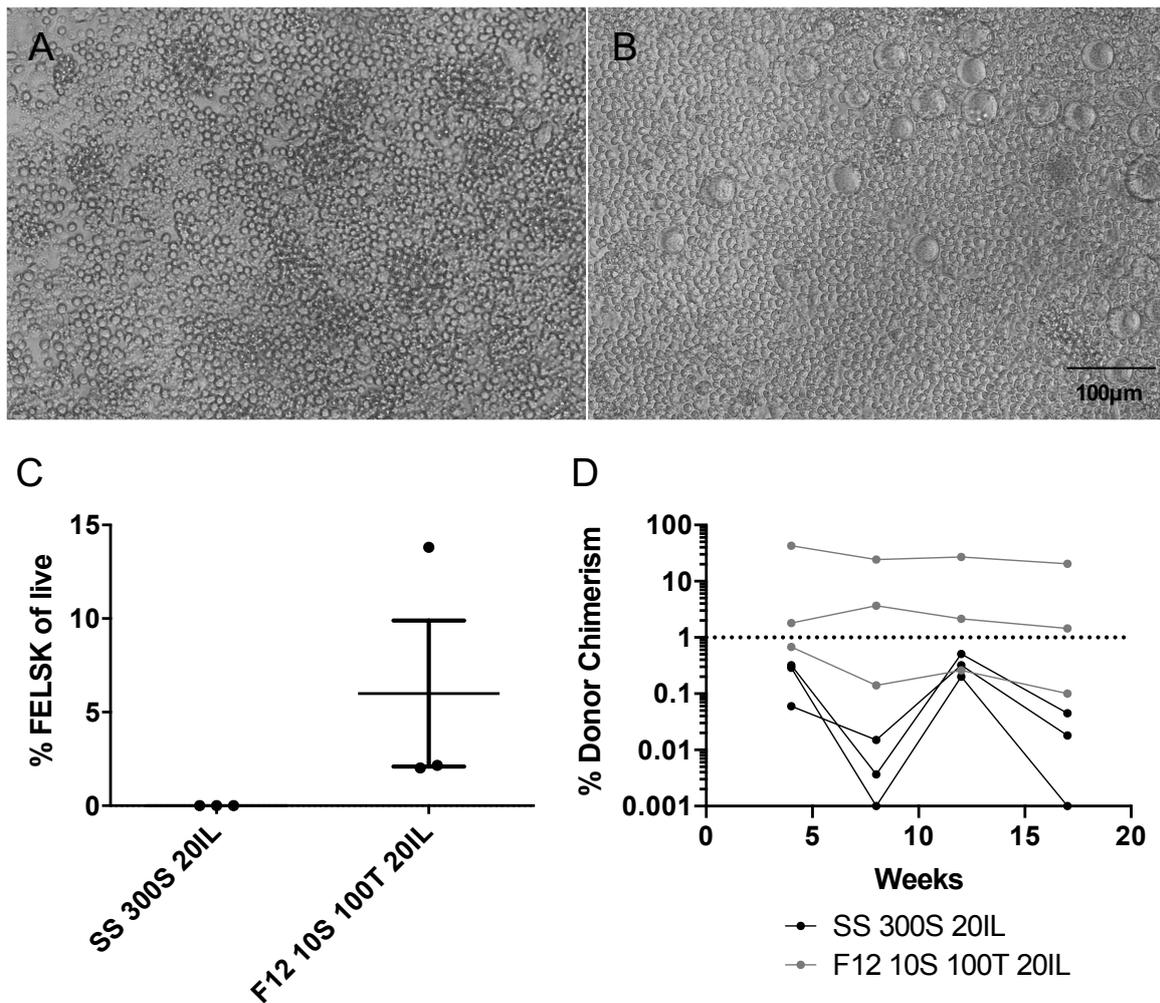


Figure 4.13 F12-based cultures are superior to StemSpan based cultures for long-term HSC expansion.

ESLAM-Sca-1⁺ HSCs were cultured for 28 days in SS (300ng/mL SCF and 20ng/mL IL-11) or F12 (10ng/mL SCF, 100ng/mL TPO and 20ng/mL IL-11)-based media with media changes every 2-3 days.

A-B) Representative light microscopy images of A) SS-based cultures and B) F12-based cultures at day 28. 10x magnification.

C) Day 28 flow cytometric analysis of cultures. Error bars represent mean \pm SEM. Unpaired T test. No significance.

D) Donor chimerism up to 16 weeks post transplantation. Dotted line represents 1% chimerism. Multiple T tests. n=3 each.

4.5 IL-11 is redundant in F12 PVA cultures

As mentioned above, the 2019 Wilkinson paper demonstrated that F12 cultures could be further improved by replacing HSA with PVA and that this affords a 200-900-fold increase in functional HSCs after 28-days of culture²⁸³. To validate whether PVA could replace HSA, single HSCs were again cultured for 10 days in F12 media with either HSA or PVA. As IL-11 was shown to be beneficial for cultures with HSA, single HSCs were also cultured with or without 20ng/mL of IL-

11, to test whether the benefit extends to PVA based cultures. As before, clone size and survival were measured, and each clone was analysed individually after 10 days for FELSK content.

4.5.1 PVA and HSA have comparable survival rates and clone sizes

Similar to HSA cultures, F12 PVA cultures also have significantly higher survival rates ($p=0.01$) compared to SS-based cultures ($79\% \pm \text{SD } 2.83$ vs $23.5\% \pm \text{SD } 2.12$) (Figure 4.14A). The survival rates when IL-11 is added was also not different from just F12 PVA cultures without IL-11 ($79\% \pm \text{SD } 5.657\%$ vs $79\% \pm \text{SD } 2.828\%$) (Figure 4.14A).

In terms of total clone size, there is considerable clonal heterogeneity in F12-HSA or -PVA cultures, which is consistent with published data by Wilkinson et al.²⁸³ (Figure 4.14). Consistent with previous findings, cells cultured in PVA, with or without IL-11, are significantly smaller ($p<0.0001$) than cells cultured in SS (Figure 4.14).

4.5.2 PVA is better than HSA at supporting phenotypic HSC expansion

Consistent with findings from Wilkinson et al., replacing HSA with PVA does indeed increase the percentage of FELSK cells, though the effect is only significant when IL-11 is added to PVA ($p=0.003$). Interestingly, although IL-11 significantly improved the mean FELSK percentage in F12 HSA cultures, adding IL-11 to F12 PVA cultures does not significantly improve the percentage of FELSK cells ($p=0.3065$) (Figure 4.14), suggesting that the inclusion of PVA replaces the need for IL-11 stimulation. Furthermore, F12 PVA cultures were not significantly better compared to F12 HSA cultures with IL-11 ($p>0.9999$) (Figure 4.12). Of note, the % of FELSK cells is not significantly different in F12 PVA cultures compared to SS control, suggesting that the impact of the F12 medium on HSC phenotype was evidenced in the latter portion of the cell culture (days 10-28).

As before, the average %FELSK of small clones derived from SS control was used as a cut-off (for these experiments, 60%). When comparing the proportions above and below this cut-off, there is a significantly higher proportion of high FELSK clones in F12 PVA cultures compared to F12 HSA cultures ($p=0.0059$) (Table 4.4). As before, there was no significant difference between F12 PVA cultures and SS-based cultures ($p=1.0$) (Table 4.4). Also, there was no significant difference between the proportions of phenotypic HSCs in F12 PVA cultures with or without IL-11 ($p=0.5368$), reinforcing the lack of benefit that IL-11 affords to F12 PVA cultures.

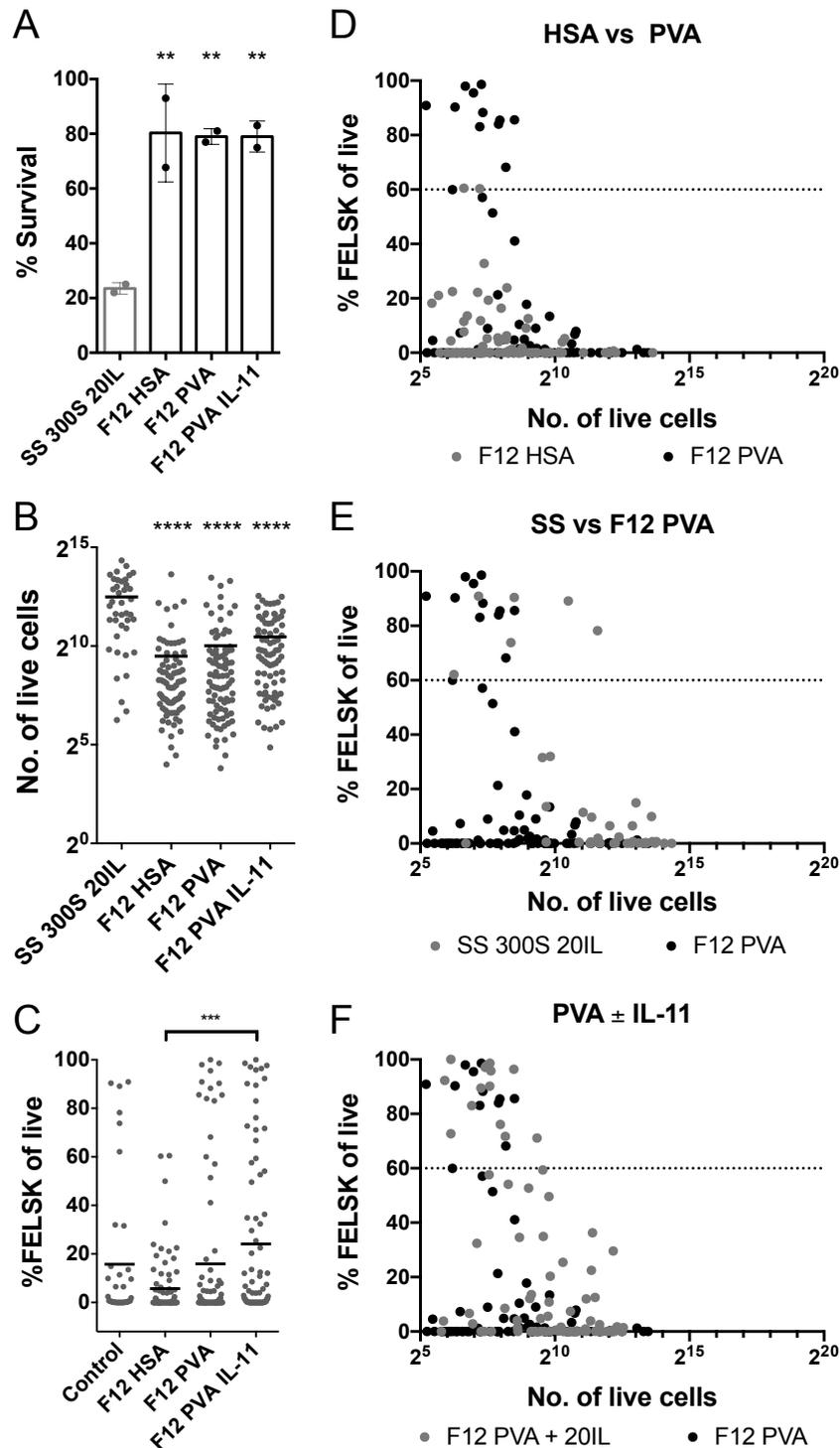


Figure 4.14 HSA vs PVA with or without IL-11

10-day cultures of single HSCs in SS (300ng/mL SCF and 20ng/mL IL-11) and F12 with HSA or PVA (10ng/mL SCF, 100ng/mL TPO and ± 20ng/mL IL-11). Exp=2 n=192 each condition.

A) Survival rate of single cell clones at day 10 of culture in respective conditions. One-way ANOVA. **= $p < 0.01$.

B) 10-day clone sizes. One-way ANOVA. ****= $p < 0.0001$.

C) Percentage of FELSK in individual clones. One-way ANOVA. ***= $p < 0.001$.

D-F) Relationship between clone size and percentage FELSK in (D) F12 HSA vs F12 PVA (10S&100T), (E) SS vs F12 PVA and (F) F12 PVA ± IL-11. Dashed line indicates the average FELSK percentage in small colonies derived from SS-based cultures (60%).

Table 4.4 Tally of clones above and below FELSK cut-off and Fisher's exact test result

	SS Control	F12 HSA	F12 PVA	F12 PVA IL-11
above 60%	6	2	13	15
below 60%	34	79	76	65
P-value vs SS		0.0156	1	0.7996
P-value vs F12 HSA			0.0059	0.0007
P-value vs F12 PVA				0.5368

4.5.3 Transplantations confirm that IL-11 has minimal effect on F12-PVA cultures

To confirm that addition of IL-11 to PVA based cultures has a redundant effect on retaining functional HSCs, 50 HSCs were cultured for 28 days in F12 PVA based cultures supplemented with or without IL-11, and then transplanted into recipient animals. As shown in Figure 4.15, with or without IL-11, all cultures had above 1% chimerism at week 16. 2 of 3 mice transplanted with cells in the absence of IL-11 had above 1% donor contribution to GM cells, compared to 1 of 3 in the presence of IL-11. Since the proportion of GM cells in an engrafted mouse has been previously linked to durable HSC self-renewal²⁰⁶ these data suggest that, if anything, IL-11 addition slightly impairs HSC expansion and support the conclusion that IL-11 does not dramatically improve F12 PVA cultures.

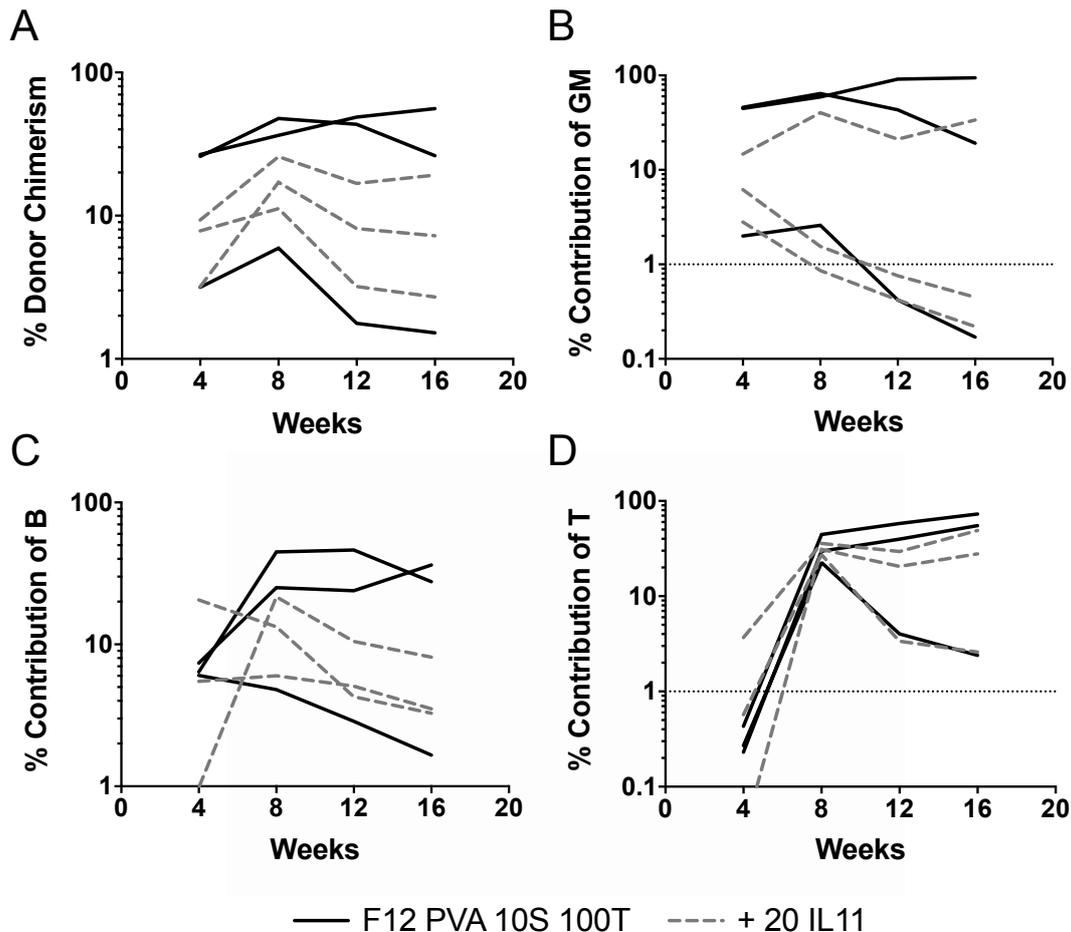


Figure 4.15 F12 PVA based cultures with or without IL-11

Transplantation of the progeny of 50 ESLAM HSCs cultured in F12 PVA supplemented with 10ng/mL SCF and 100ng/mL TPO, with (Grey dashed) or without IL-11 (Black) for 28 days. Exp=1 n=3.

A) Percentage donor chimerism.

B-D) Percentage donor contribution to GM, B and T cells. Straight dotted line indicates where contribution is at 1%.

4.5.4 Single HSCs cultured in F12 PVA display clonal heterogeneity

Single HSC initiated cultures were also transplanted and, consistent with the heterogeneity in phenotypic outcomes observed by Wilkinson et al.²⁸³ and previous 10-day cultures, clones cultured with IL-11 displayed substantial heterogeneity. Compared to Wilkinson et al. where 4 of 14 clones (28.5%) displayed long-term multilineage engraftment in the absence of IL-11, clones containing IL-11 were capable of long-term reconstitution in 3 of 8 cases (37.5%) (Figure 4.16). This is a proportion that is not statistically different as calculated by Fisher's exact test ($p=1.0$), which again suggests that there is marginal impact in adding IL-11 to these cultures. Based on these data, from this point onwards, the base medium was selected to be F12 PVA without IL-11, in order to be consistent with published conditions by Wilkinson et al.²⁸³

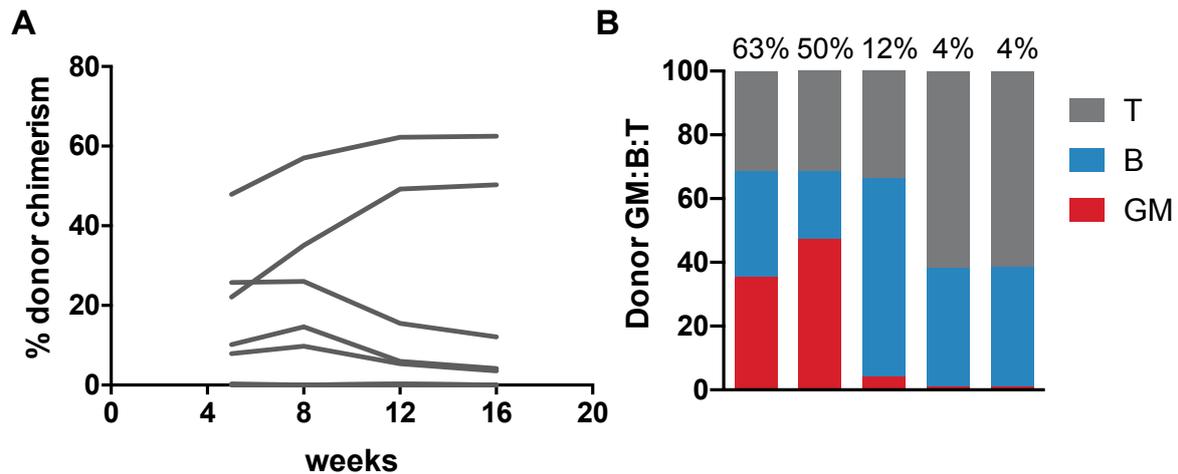


Figure 4.16 transplantation of single cell clones cultured with IL-11 for 28 days.

- A) Donor chimerism up until 16 weeks in primary recipients of single clones cultured in F12 PVA supplemented with 10ng/mL SCF, 100ng/mL TPO and 20ng/mL IL-11. n=8.
- B) Ratio of donor contribution to GM, B and T cell lineages. Ordered by overall donor chimerism on top of each bar.

4.5.5 Transplantation outcome can be retrospectively predicted by reporter strategy

In order to further validate the FELSK reporter strategy, 10% of the above single cell cultures were removed on day 27 and analysed by flow cytometry, allowing transplantation outcomes to be correlated with surface marker phenotype in each culture. As shown in Figure 4.17A,B, there is large accompanying heterogeneity in the phenotypic outcomes of 28-day clonal cultures. Interestingly there is a strong correlation between the phenotypic outcomes and functional outcomes. As shown in Figure 4.17F, the percentage of LSK cells in each clone correlates highly with its donor chimerism at week 16 ($r = 0.86$; $p=0.0056$). As expected, the correlation is slightly stronger with percentage of FELSK cells ($r=0.92$; $p=0.0012$). Interestingly, there is no significant correlation ($p=0.9116$) between the size of the clone and its repopulating ability, again demonstrating a decoupling between proliferation and self-renewal and supporting the need for a reporter strategy to isolate functional HSCs in a robust manner.

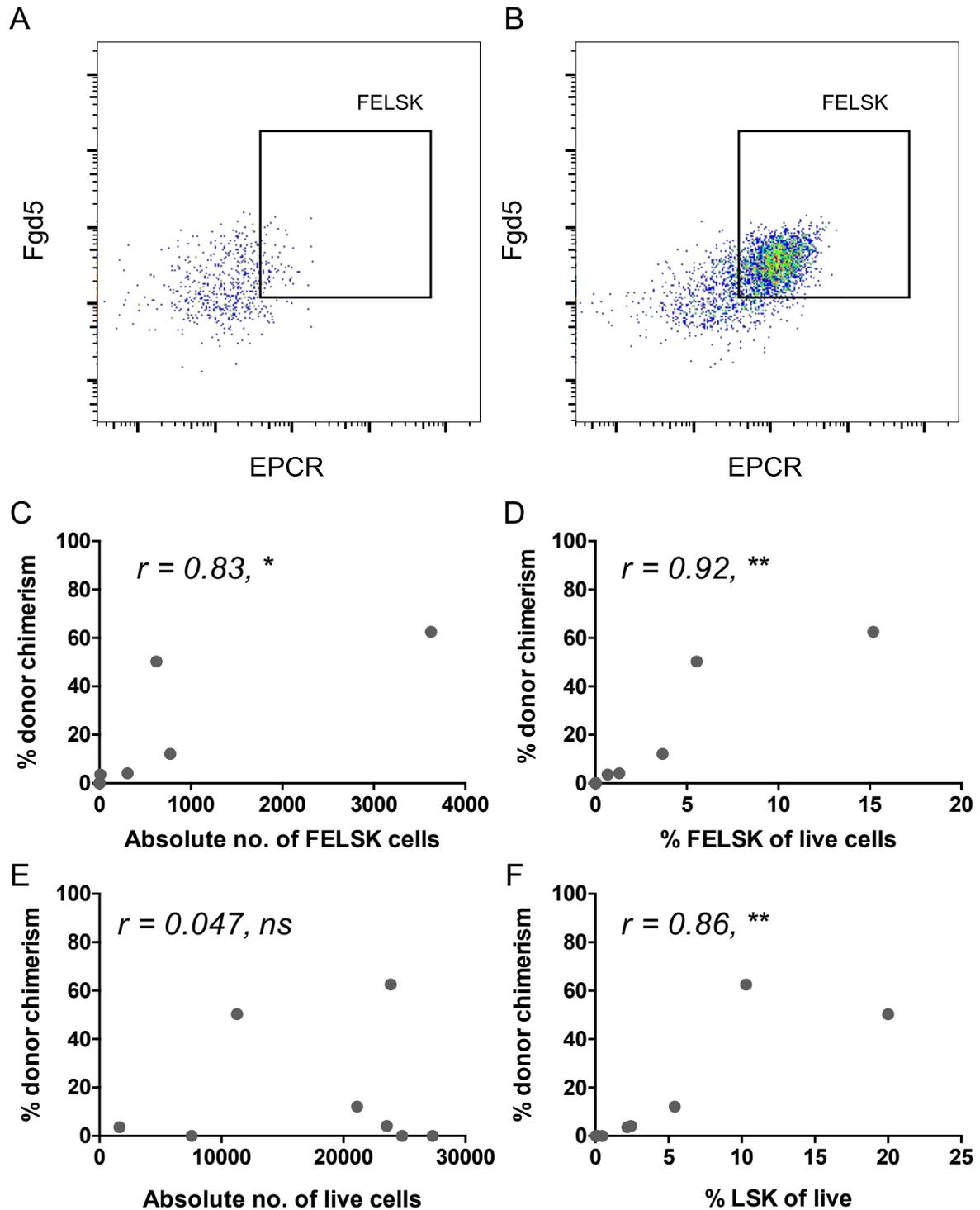


Figure 4.17 Donor chimerism correlates with phenotypic isolation strategy but not cell numbers.

10% of clones from Figure 4.16 were analysed by flow cytometry a day before transplantation.

A-B) Representative gating layout of FLSK cells (Gated out of LSK gate), in clones with low FLSK% (A) and high FLSK% (B).

C-F) Correlation between donor chimerism and C) number of FLSK cells, D) percentage of FLSK cells out of live, E) number of live cells and F) percentage of LSK cells. Pearson correlation. $*$ = $p < 0.05$. $**$ = $p < 0.01$. ns= not significant. $n=8$.

Overall, these studies have shown that *Fgd5* and EPCR are reliable markers for HSCs *in vitro* - in short-term cultures, the vast majority of repopulation potential is contained within cells expressing high amounts of *Fgd5* and EPCR; likewise, *Fgd5* and EPCR expression is strongly correlated with traditional *in vitro* HSC markers; and can be used altogether in combination as a novel reporter strategy to screen HSC supportive conditions. This strategy validated both EL08 CM and F12 PVA cultures. In the process, PTN was identified as one of the secreted factors in EL08 CM that is supporting HSC self-renewal. These experiments also confirmed and demonstrated the clonal heterogeneity in long-term HSC expansion cultures, which becomes the focus of next chapter's investigation.

5 Molecular characterisation of clonal heterogeneity in F12 cultures

The F12 PVA culture system described in 2019 is capable of achieving 200- and 900-fold expansion of HSCs over a 28-day period²⁸³. Interestingly, when individual clones were selected to examine the expansion of single HSCs over the time course, there was substantial heterogeneity with only 4 of 14 clones (28.5%) displaying long-term multilineage engraftment. These data suggest that the vast majority of HSC expansion present in bulk cultures was achieved by a minority of clones. In Chapter 4, I described the utility of our FELSK reporter system for identifying cell cultures rich in functional mouse HSCs and Chapter 5 focuses on using this reporter strategy to shed light on the molecular basis for the observed clonal heterogeneity.

5.1 Transplantation of 28-day clones from single cell cultures

Having also observed significant heterogeneity in the functional outcomes of single HSC derived cultures in section 4.5.4, we were interested in whether we could resolve the molecular differences by using the *in vitro* FELSK strategy to identify HSCs. To achieve this, single HSCs were cultured for 28-days and then re-sorted for two fractions: phenotypic HSCs (EPCR⁺ Lineage⁻ Sca-1⁺ and c-Kit⁺ (ELSK)) and all the remaining cells (non-ELSK) (Figure 5.1). For these experiments, *Fgd5* was not used for the gating strategy for two reasons: 1) EPCR and *Fgd5* levels correlate strongly and 2) Reliance on the reporter mouse decreased the applicability of the method including the practical ability to achieve sufficient numbers of starting cells for rapid molecular profiling. HSCs were still isolated from *Fgd5*^{ZsGreen•ZsGreen/+} mice so the immunophenotypic profiling of individual cultures could still be performed. In total, we assessed 20 clones from two independent experiments for matched functional and molecular profiling and the overall experimental design is illustrated in Figure 5.2. On average, 22382 ± SD 34278 ELSK cells were sorted for each clone compared to 90700 ± SD 85383 non-ELSK cells (p=0.0027, paired t-test) (data not shown).

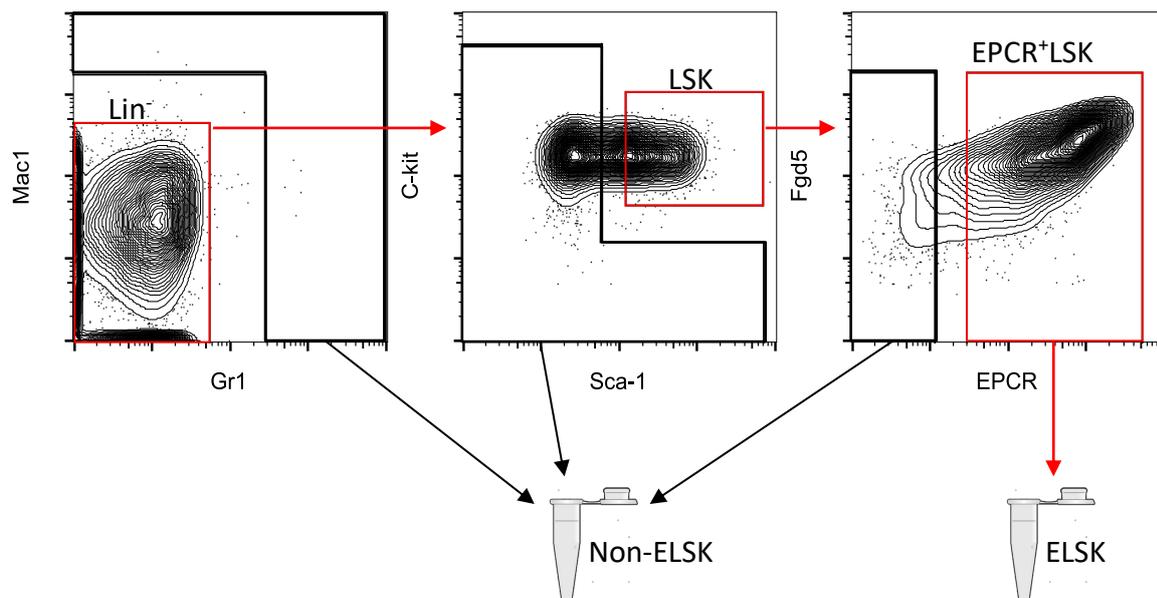


Figure 5.1 Representative gating strategy to separate phenotypic HSCs from culture.

Single HSCs were cultured for 28 days in F12 PVA, supplemented with 10ng/mL SCF and 100ng/mL TPO, with media changes. At day 28, single cell clones were each re-sorted for two fractions. Gating strategy for EPCR⁺ Lineage⁻ Sca-1⁺ and c-kit⁺ cells (ELSK) indicated by red gates. Non-ELSK cells are captured by the black gates, which are intentionally spaced away from red gates to reduce the chances of ELSK cells being captured.

In order to match the functional outcomes of the clones with their molecular profiles, the re-sorted ELSK and non-ELSK samples were divided in two halves; One half was stored for later bulk RNA-seq, while the other half was transplanted into irradiated mice (Figure 5.2).

Specifically, all ELSK samples, presumed to be the fraction containing HSCs, were transplanted into at least one recipient per clone (50% dose), with some (7 clones) being transplanted into two recipients at a 45% and 5% dose. Due to limited numbers of recipients available at the time of experiment, the non-ELSK cells were pooled into three groups that were each transplanted into a single recipient mouse.

In order to analyse secreted factors within the clonal cultures, the medium that was taken out during media changes was also stored for proteomic analysis (Figure 5.2). Therefore overall, we generated matched datasets for single expanded HSC that contained functional assessment of HSC activity, gene expression of phenotypic HSCs and non-HSCs by RNA-seq, and proteomic analysis of secreted factors throughout the culture.

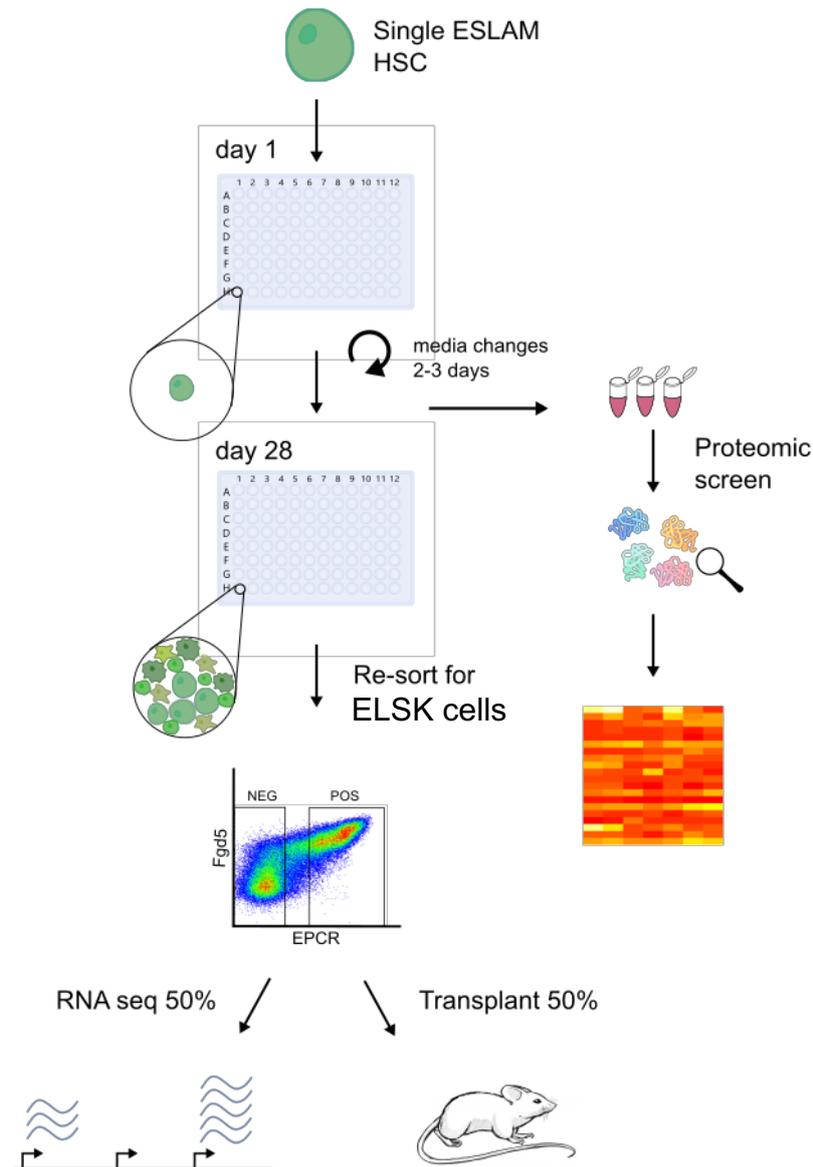


Figure 5.2 schematic of experimental design to characterise molecular heterogeneity in PVA cultures.

Single HSCs were cultured for 28 days in F12 PVA cultures supplemented with 10ng/mL SCF and 100ng/mL TPO. On day 28, clones were resorted for ELSK and non ELSK cells as shown in Figure 5.1. Resorted samples were separated in half; one fraction (or pooled with other clones) was transplanted into recipient mice; the other fraction was stored for RNA-seq.

Media changes were performed every 2-3 days, the media were collected and analysed by mass spectrometry for proteins.

5.1.1 Split dose transplantations confirm bona fide HSC expansion and functional heterogeneity

Consistent with our previous transplantation data and published findings, there was considerable clonal heterogeneity in primary transplantation outcome. Compared to the originally reported percentage of clones with high level multilineage engraftment (4 of 14, 28.5%) and our initial IL-

11 stimulated cultures (3 of 8, 37.5%), we again observed a similar frequency (8 out of 20, 40%), with no statistical difference ($p=0.717$) (Figure 5.3). Of the 5% doses, all 7 recipients had $>1\%$ donor chimerism; However, only 2 of 7 clones had robust ($>1\%$) contribution to GM cells, confirming at least one self-renewal division in single HSCs.

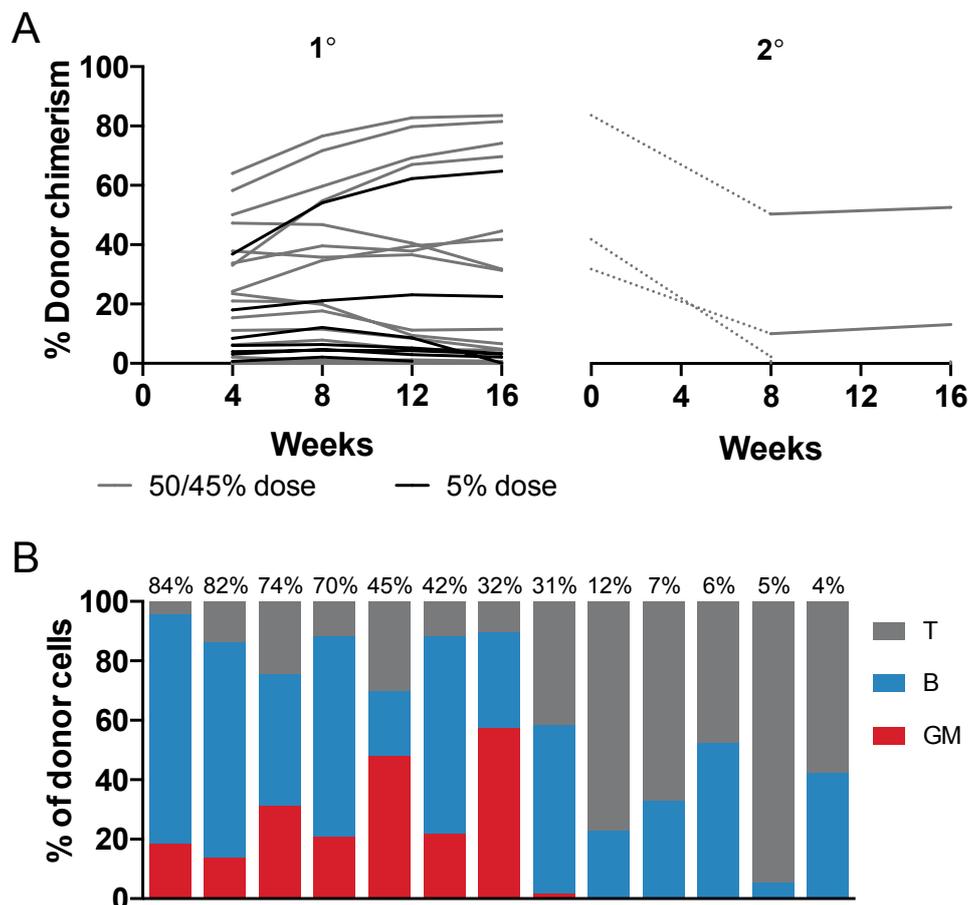


Figure 5.3 Transplantation of ELSK from single clones cultured in F12 PVA media.

A) Donor chimerism from primary and secondary transplantation of ELSK cells from single cell clones cultured in F12 PVA with 10ng/mL SCF and 100ng/mL TPO for 28 days. Grey line represents 50%/45% cell doses $n=20$. Black line represents 5% cell dose $n=7$. Exp=2 (2 mice were culled for non-experimental reasons before experimental end point). Awaiting 16-week timepoint from secondaries from 2nd experiment, here showing secondaries from 3 positive primary mice.

B) Lineage output of positively repopulated clones ($>1\%$ chimerism at week 16) from 45 or 50% doses. Ordered by percentage donor chimerism.

5.1.2 ELSK separates functional HSCs from non-HSCs

Despite transplanting on average 26-fold more cells per mouse and representing multiple clones, non-ELSK cells mostly did not have robust multilineage reconstitution (2 of 3 groups had below 1% contribution to GM) (Figure 5.4). Upon deeper analysis, the group that had successfully engrafted was pooled from clones that included 4 out of 8 of the positively repopulating clones, including the highly expanded clones with robust chimerism even at the 5% dose, making it not

unlikely that a few HSCs may have been mistakenly sorted into the non-ELSK population (Figure 5.4). Overall, this gives confidence that the ELSK cells can robustly separate functional HSCs from non-HSCs.

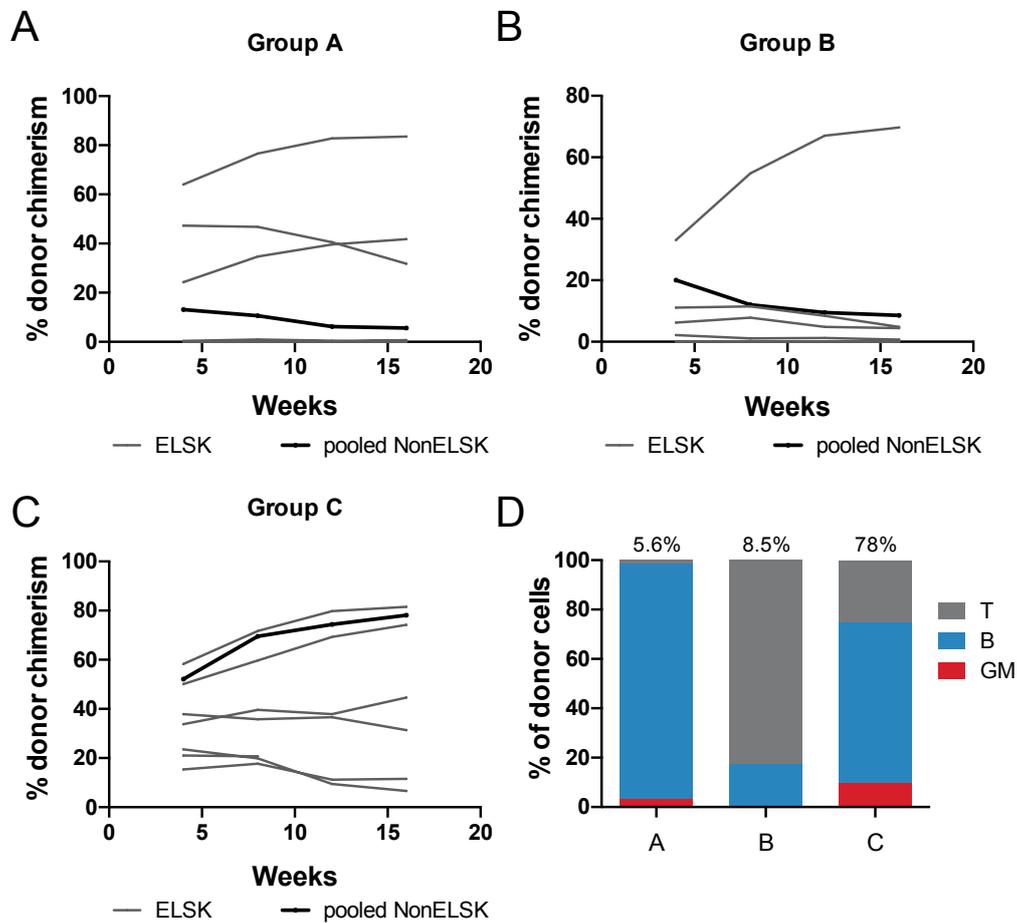


Figure 5.4 Reconstitution of pooled non-ELSK cells compared to the respective ELSK cells.

The non-ELSK cells from the same clones transplanted in Figure 5.3 were pooled into 3 separate groups and each transplanted into an individual recipient. Two independent transplantations: Group A was from experiment 1, n=7; Group B and C were from experiment 2, n=6 and 7 respectively.

A-C) Donor chimerism of pooled non-ELSK (black) and the ELSK cells (grey) from their respective groups in primary transplantations. n=3.

D) Lineage output of pooled non-ELSK cells. 16-week chimerism is labelled above respective group.

5.1.3 FELSK correlates highly with functional HSC activity

From the information generated from the re-sort on the day of transplantation, the percentage of donor chimerism can be correlated with the FELSK percentage present in individual clones. As shown in Figure 5.5 there was again a high correlation between %FELSK and donor chimerism ($r=0.9296$, $p<0.0001$). Of note, all of the five clones that were lower than 1% FELSK, were also below 1% donor chimerism at week 16. Interestingly, there was a lower but still significant correlation between absolute numbers of FELSK cells and donor chimerism.

Also, as identified in Chapter 4, there was no significant correlation between the total clone size and chimerism ($r=-0.098$, $p=0.6898$), again affirming that there is a negative relationship between proliferation and HSC self-renewal. A very high correlation ($r=0.9383$, $p<0.0001$) was observed between %FELSK of the clone and its contribution to GM cells, which as mentioned is a strong indicator of robust self-renewal (Figure 5.5). Interestingly, a pattern can be observed that suggests that there is a threshold of %FELSK cells, above which contribution to GM would be high.

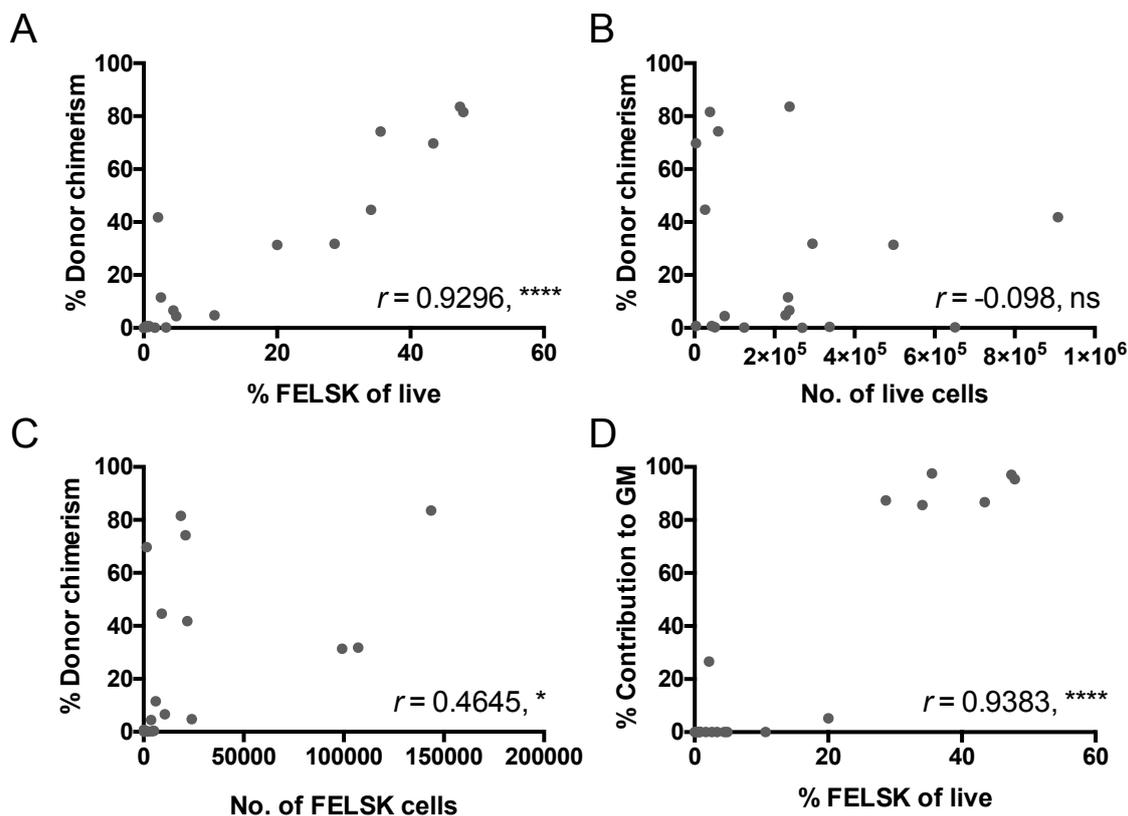


Figure 5.5 FELSK correlates with chimerism.

A-D) Clones from transplantations detailed in Figure 5.3 were analysed for their FELSK content and clone size, which is correlated with respective chimerism and contributions to GM cells. Pearson Correlation. ****= $p<0.0001$. *= $p<0.05$. ns = not significant. $n=19$.

5.2 Repopulating ELSK cells are molecularly distinct from non-repopulating cells and resemble freshly isolated HSCs at gene expression level

To study the molecular state of *in vitro* expanded HSCs and the potential drivers of “good” versus “bad” clones, RNA-seq was performed on 12 clones, representing the first known RNA-seq dataset of expanded HSCs. In total, 24 samples were sequenced over three runs including matched ELSK and non-ELSK fractions (Table 5.1). Clones were selected for a range of functional output and FELSK percentages. To simplify the analysis, the samples were categorised into 4 groups: 1) ELSK cells from clones that repopulated mice (PosELSK), 2) ELSK cells from clones that did

not repopulate mice (NegELSK), 3) Non-ELSK cells from clones that repopulated mice (PosNonELSK), and 4) Non-ELSK cells from clones that did not repopulate mice (NegNonELSK). Repopulation was defined as having >1% donor chimerism and >1% contribution to GM at 16 weeks.

Table 5.1 12 clones with matched ELSK and non-ELSK samples were chosen for RNA-seq.

Clone	Repopulation	Fraction	Group	16w Donor Chimerism	16w contribution to GM	QC
A	Y	*ELSK	*PosELSK	83.50%	97.00%	
		Non-ELSK	PosNonELSK			
B	Y	ELSK	PosELSK	81.50%	95.40%	
		Non-ELSK	PosNonELSK			
C	Y	*ELSK	*PosELSK	74.20%	97.50%	
		Non-ELSK	PosNonELSK			
D	Y	ELSK	PosELSK	44.60%	85.60%	
		Non-ELSK	PosNonELSK			
E	Y	ELSK	PosELSK	41.80%	26.60%	
		Non-ELSK	PosNonELSK			
F	Y	ELSK	PosELSK	31.80%	87.40%	
		Non-ELSK	PosNonELSK			
G	N	ELSK	NegELSK	0.67%	0.02%	
		Non-ELSK	NegNonELSK			
H	N	ELSK	NegELSK	0.40%	0.02%	
		Non-ELSK	NegNonELSK			
I	N	ELSK	NegELSK	0.22%	0%	
		*Non-ELSK	*NegNonELSK			
J	N	ELSK	NegELSK	0.17%	0.02%	
K	N	ELSK	NegELSK	0.11%	0%	
		Non-ELSK	NegNonELSK			
L	N	ELSK	NegELSK	0%	0%	Failed
		Non-ELSK	NegNonELSK			

* indicates sample was run twice as a technical repeat.

Of the 24 cell fractions, one Neg-ELSK sample failed quality control (Table 5.1). As a control, we ran two samples as technical repeats (Table 5.1) in all three sequencing runs to allow us to identify any batch variation. As shown in method section 2.7.1, the technical replicate was extremely similar in the two runs, which allowed for a simple batch correction. After filtering out low expressing genes, there were 16648 genes detected across the 22 unique samples. As visualised in Figure 5.6, multiple dimensional scaling (MDS) shows a clear separation between samples originating from clones which repopulated mice and clones that did not. Samples could be further resolved by whether they were ELSK cells or Non-ELSK cells. Notably, there was an overlap between

PosNonELSK cells and NegELSK cells, which suggests that even though NegELSK cells are phenotypically more similar to PosELSK cells, they are molecularly more similar to non-ELSK cells. As shown in Figure 5.6, the molecular separation within ELSK cells can be further resolved by percentage chimerism and GM contribution of the clone.

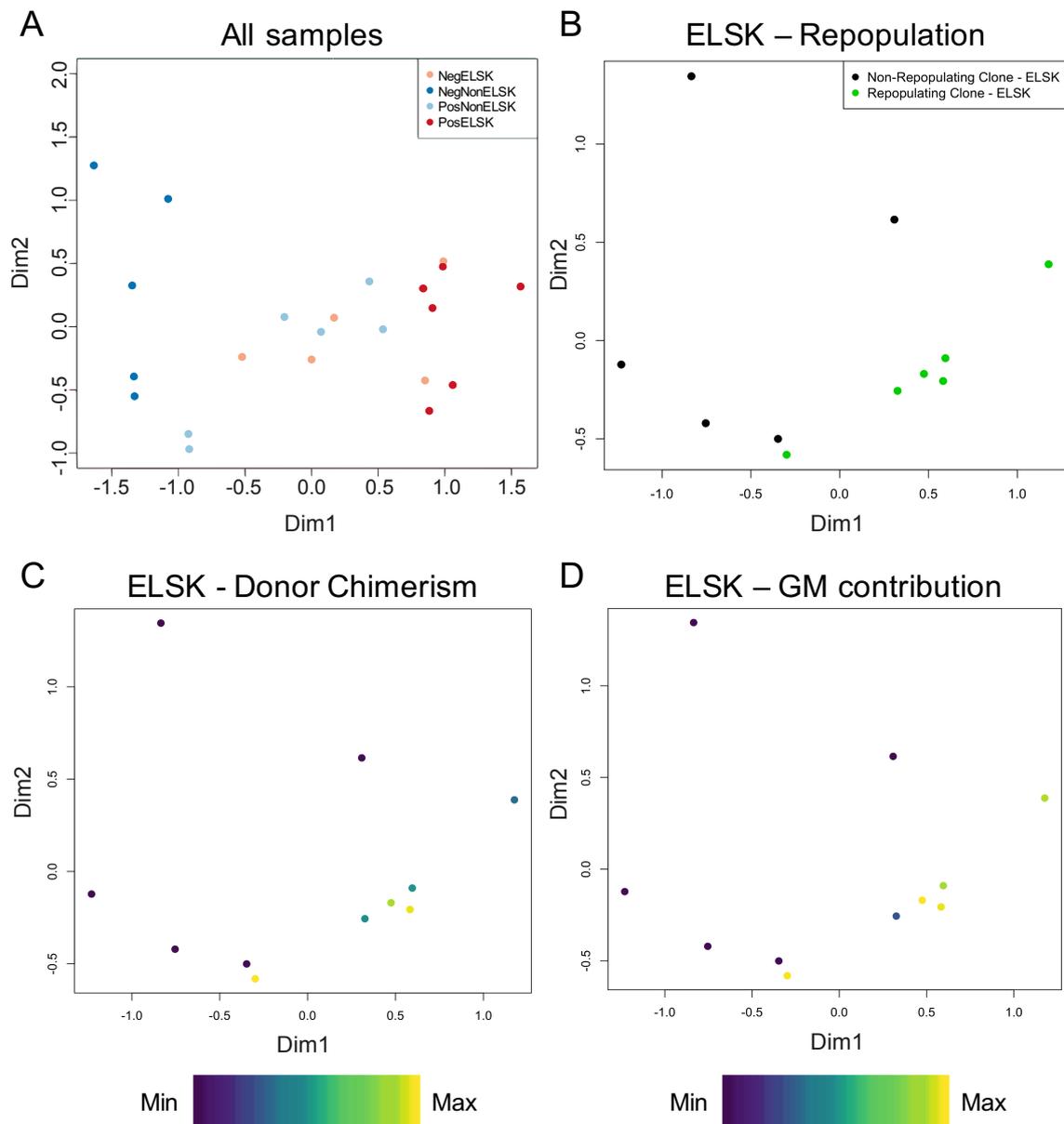


Figure 5.6 MDS plots show clear separation between repopulating and non-repopulating cells.

A) MDS plots of all 22 unique samples, coloured by the groupings into PosELSK, NegELSK, PosNonELSK and NegNonELSK.

B-D) MDS plots of all 11 ELSK samples coloured by B) whether they can repopulate recipients or not, C) the respective 16-week donor chimerism in primary transplantation and D) the respective 16-week contribution to GM lineage. Colours are scaled to the samples from highest (yellow) to lowest (purple).

As detailed in section 1.3.2, a molecular signature (MoIO) for freshly isolated quiescent HSCs from mouse BM was developed by Wilson et al.¹⁸⁶ Using this molecular signature, our samples were scored for their combinatorial expression of the MoIO genes. As shown in Figure 5.7, ELSK cells have a significantly higher MoIO score than non-ELSK cells ($p < 0.05$). Moreover, repopulating clones had higher MoIO scores compared to non-repopulating clones ($p < 0.0001$) (Figure 5.7). Interestingly, there was no significant difference between NegELSK cells and NegNonELSK cells (Figure 5.7), suggesting that the MoIO signature correlates with actual HSC function rather than the ELSK phenotype itself. Of note, amongst the MoIO signature genes, several genes were below the minimum expression limits in the QC for all the samples (*Cldn10*, *Ramp2*, *Smtnl1*, *Sox18* and *Sqrdl*), suggesting that although these genes are expressed in freshly isolated HSCs (and may play a biological role there), they are unnecessary for *ex vivo* cultured HSCs. Overall, this confirms that not only can the transcriptomic information separate functionally distinct samples, it also demonstrates that expanded HSCs retain a gene expression signature reminiscent of freshly isolated HSCs.

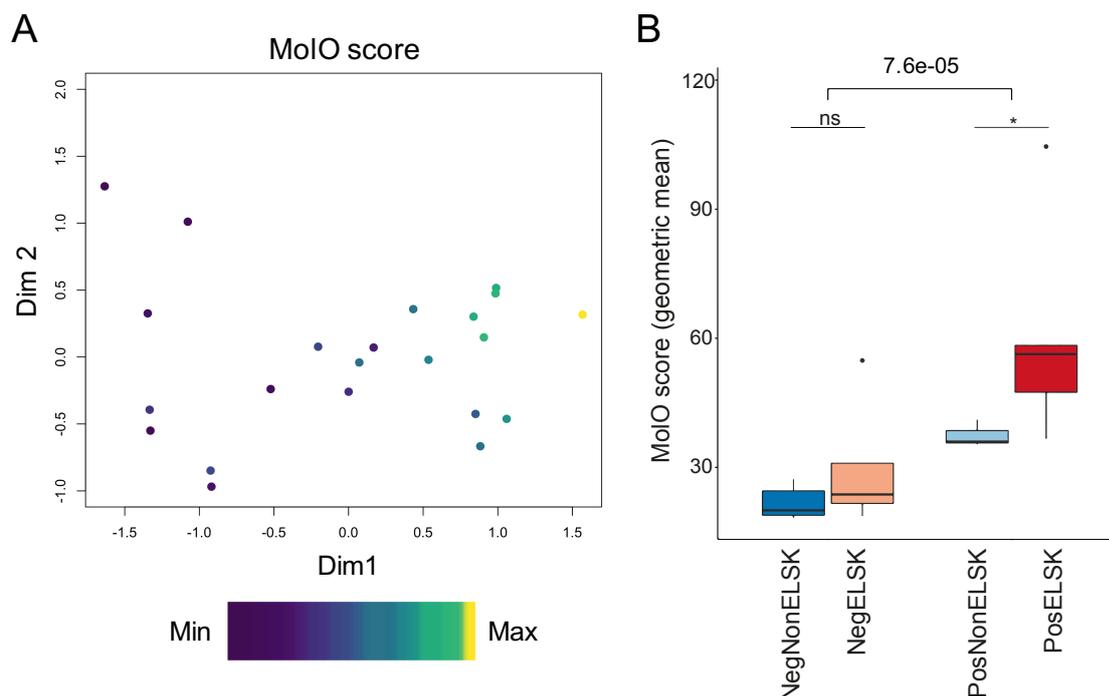


Figure 5.7 MoIO score is significantly higher in repopulating and ELSK cells.

A) MDS plot of all samples from Table 5.1 with the MoIO score of each sample coloured in. Scaled to the samples, max (yellow) and min (purple)

B) MoIO score (geometric mean) of the four sample groups. t-Test. $* = p < 0.05$. ns= not significant.

5.2.1 Non-HSCs are not mature cells but differentiated progenitors

Next, in order to achieve a more global comparison of the samples with known haematopoietic datasets, the samples were compared to the various cell types on ImmGen, by SingleR analysis. As described in section 3.3, ImmGen is a gene expression database containing 253 haematopoietic cell types and SingleR is a programme that is designed to compare sample datasets with reference cell types. As shown in Figure 5.8, all of the samples, including non-ELSK cells are more similar to HSCs and multipotent progenitors than they are to any specific mature cell fraction, suggesting that even the clones that no longer contain functional HSC activity are still comprised of relatively immature haematopoietic cells. This is also consistent with our previous data and the original data from Wilkinson et al. that there are very few lineage positive cells and that mature cells are not well-supported in F12, PVA-based cultures.

Of the 22 samples, only two were classified as LTHSCs and both of these were PosELSK clones (Figure 5.8). Interestingly, amongst the 7 samples labelled “MLP” (Multipotent lymphoid progenitors, Lineage-CD19⁻IgM⁻CD43⁺CD24⁻AA4.1⁺CD45R⁻CD117⁺IL7R⁻), all of them were NegNonELSK cells, suggesting that even the most differentiated samples are not mature cells (Figure 5.8). The remaining samples are most similar to ST-HSCs (CD34⁺Flk2⁻LSK cells), however, as shown in Figure 5.8, the PosELSK samples are clustered together and contain some similarities with LT-HSCs. Overall this confirms previous findings that even though the cells in the ELSK fraction have been cultured for 28 days, their transcriptomes still largely resemble their freshly isolated counterparts and likely contain the essential transcriptional networks for durable HSC activity.

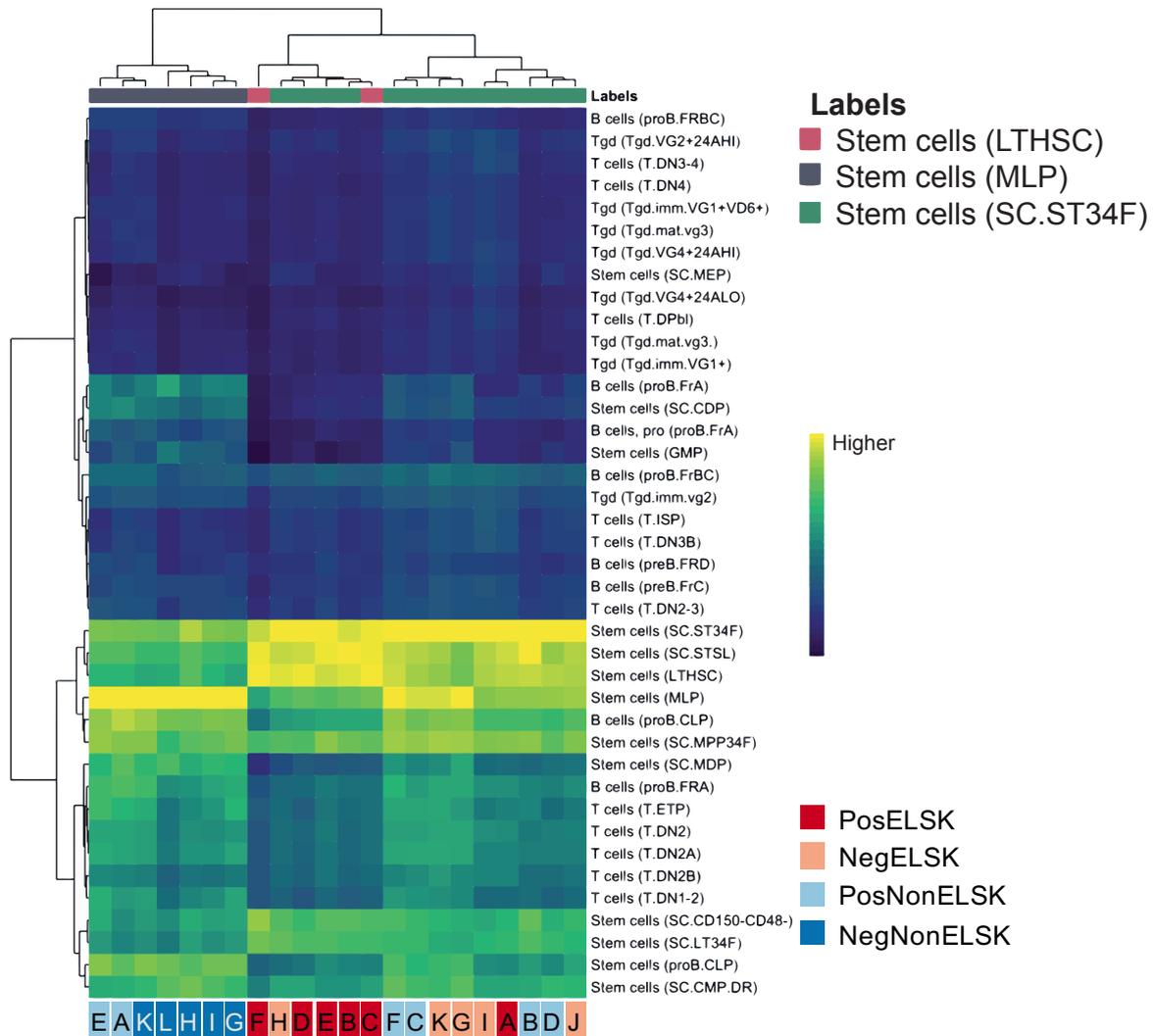


Figure 5.8 SingleR identifies PosELSK samples as most similar LTHSCs.

The 22 samples were compared to ImmGen gene expression database of 253 cell types. The 40 samples with the highest correlation score are shown, coloured from highest (yellow) to lowest (blue). The cell type with the highest correlation is labelled above the heatmap and the corresponding clone and sample type (detailed in Table 5.1) is labelled below.

5.3 Differential gene expression between ELSK and non-ELSK cells suggest early activation in non-ELSK cells

We next undertook differential gene expression analysis in order to better understand the genes that drive the separation between ELSK and non-ELSK cells. Amongst the 16648 genes analysed, 2398 genes were differentially expressed (Log fold change (LogFC) ≥ 2 and FDR < 0.05) between ELSK cells and non-ELSK cells (1158 genes upregulated in ELSK cells and 1250 upregulated in non-ELSK cells) (Figure 5.9).

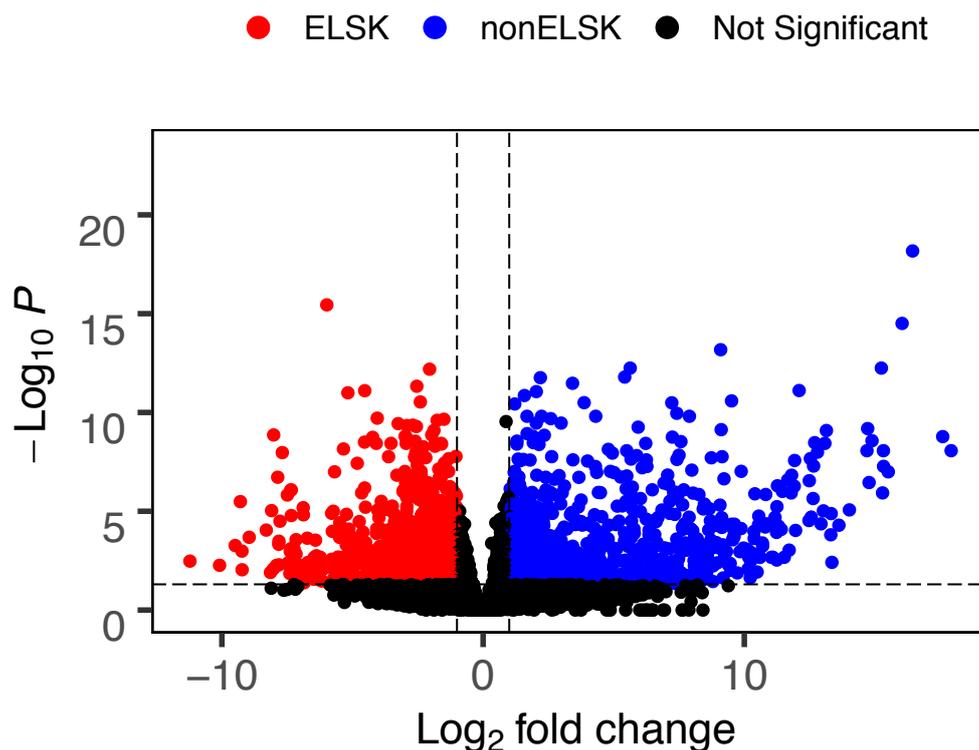


Figure 5.9 Volcano plot showing differentially expressed genes between ELSK and non-ELSK cells.

LogFC and p-value of all 16648 genes. Red dots represent genes upregulated in ELSK cells (1158 genes) and blue dots represent genes upregulated in PosNonELSK cells (1250 genes). Black dots represent non-significant genes. Dashed lines represent the cut-offs.

Gene ontology (GO) terms of genes upregulated in ELSK cells suggest an enrichment in genes involved in developmental processes and cytoskeletal reorganisation (Table 5.2). Consistent with their identity as more differentiated cells, GO terms from non-ELSK upregulated genes suggest involvement in immune responses, inflammation and cell activation (Table 5.3).

Table 5.2 Gene ontology terms based on upregulated genes in ELSK cells compared to Non-ELSK cells.

GO term	Genes DE	P-value
developmental process	418	2.61E-27
anatomical structure development	395	5.41E-26
localization	390	2.65E-23
intracellular signal transduction	215	6.23E-23
cytoskeleton organization	132	2.97E-21
regulation of signal transduction	215	1.55E-19
organelle organization	255	8.62E-19
regulation of response to stimulus	265	6.12E-17
cell cycle	141	1.96E-16
cell projection organization	135	1.08E-15

Table 5.3 Gene ontology terms based on upregulated genes in Non-ELSK cells compared to ELSK cells.

GO term	Genes DE	P-value
immune system process	304	5.51E-54
localization	463	3.62E-37
positive regulation of biological process	474	7.81E-37
regulation of response to stimulus	344	1.17E-36
intracellular signal transduction	259	1.87E-34
defense response	189	3.74E-34
cell activation	149	2.24E-33
regulation of immune system process	177	7.07E-33
immune response	188	7.37E-32
inflammatory response	109	4.85E-30

5.3.1 Rho GTPase pathways significantly upregulated in phenotypic HSCs

To gain a deeper understanding of molecular mechanisms, pathway analysis was also performed on the differentially expressed genes between ELSK cell and non-ELSK cells. As shown in Figure 5.10, signalling by Rho GTPases seem to be particularly involved in ELSK cells. Interestingly, several cell cycle pathways have also been identified and is suggested to interact with Rho GTPase pathways. Rho GTPases have been previously implicated in HSC function^{439,440} and in particular, *Cdc42* has been recently implicated in regulating HSC polarity and self-renewal divisions²⁵⁴. If validated, Rho GTPases can potentially represent an attractive therapeutic target pathway to increase expansion of HSCs.

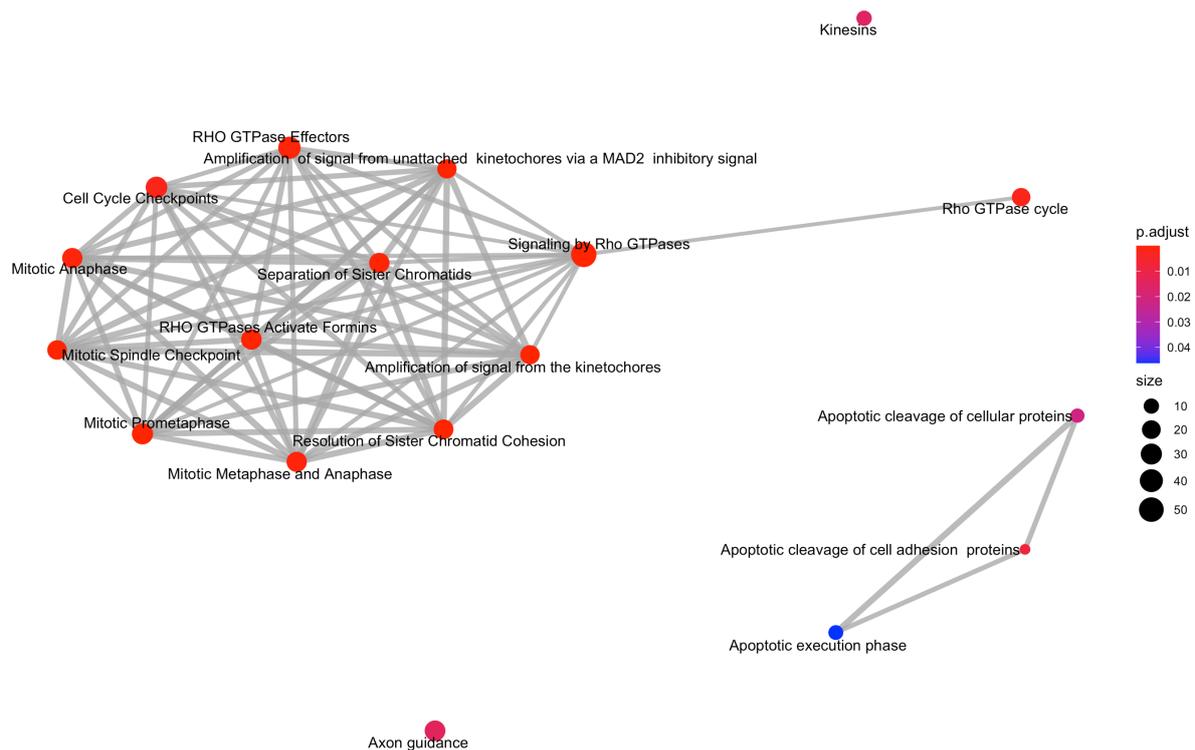


Figure 5.10 Pathways associated with upregulated genes in ELSK cells compared to non-ELSK cells.

Reactome pathway analysis of genes significantly upregulated in ELSK cells compared to non-ELSK cells. Lines connect pathways that have shared genes, and pathways are coloured by adjusted p-values from more significant (red) to less (blue). The main cluster relates to Rho GTPase activity.

In non-ELSK cells, pathways involved in receptor signalling are overrepresented (Signalling by SCF-KIT, Interleukin-3, Interleukin-5 and GM-CSF signalling) (Figure 5.11). As shown in Figure 5.11, there is also a cluster of pathways involved in ribosomal activity. Overall suggesting an increased activation in these cells, which supports conclusions from GO term analysis.

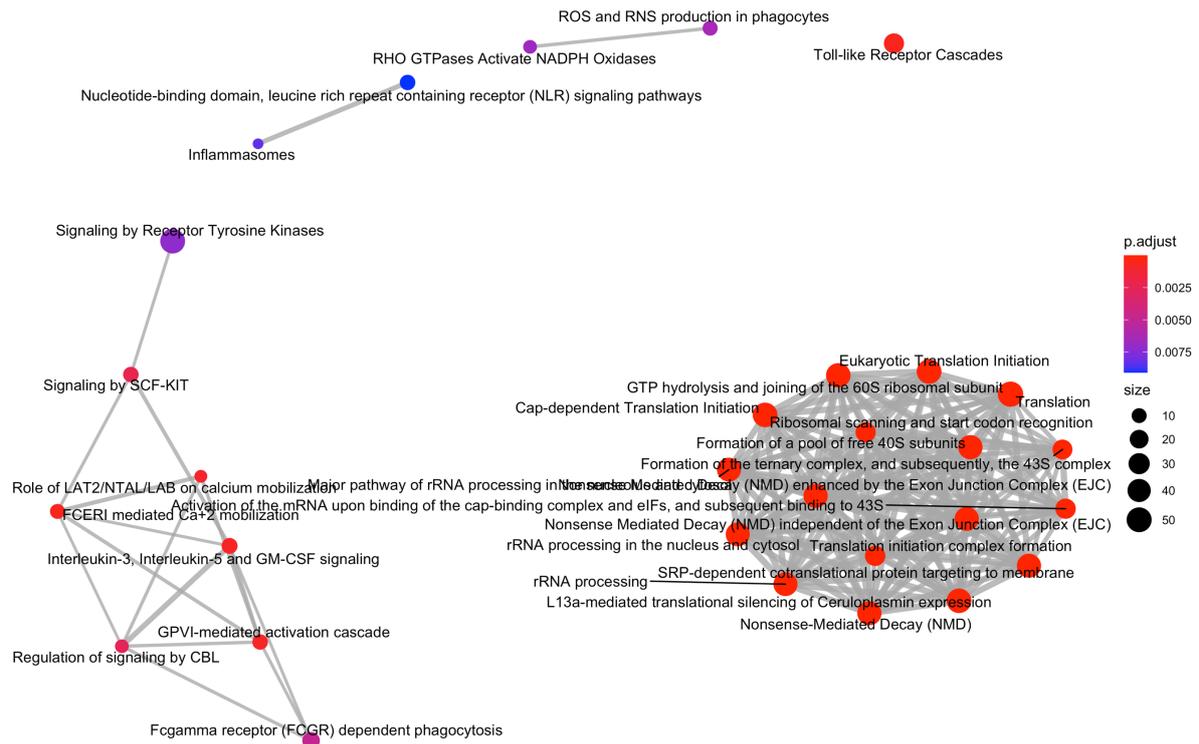


Figure 5.11 Pathways associated with genes upregulated in non-ELSK cells compared to ELSK cells.

Reactome pathway analysis of genes significantly upregulated in non-ELSK cells compared to ELSK cells. Lines connect pathways that have shared genes, and pathways are coloured by adjusted p-values from more significant (red) to less (blue). The two main clusters involve cytokine signalling and ribosomal activity.

5.3.2 Differential gene expression between PosELSK and NegELSK

The comparison between ELSK cells and non-ELSK cells was useful to establish the broad differences between primitive and more differentiated cells in the culture. However, to understand the molecular state of functional HSCs *in vitro*, a more interesting comparison is between PosELSK and NegELSK cells. Of the 16648 genes, only 26 genes were upregulated, and 35 genes were downregulated in PosELSK cells versus NegELSK cells, suggesting a high degree of molecular similarity despite dramatic differences in functional output (Figure 5.12).

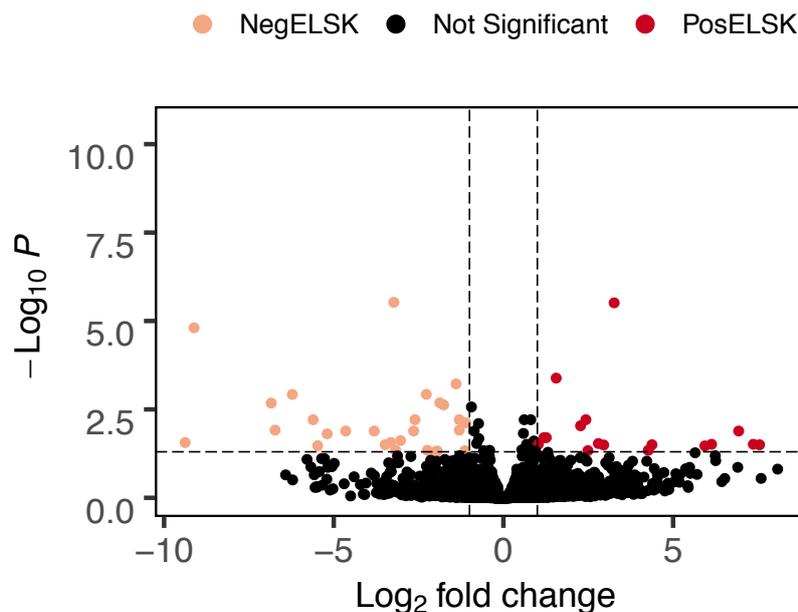


Figure 5.12 Volcano plot of genes differentially expressed between PosELSK and NegELSK cells.

LogFC and p-values of all 16648 genes. Pink dots represent genes upregulated in NegELSK (35 genes) and light blue dots represent genes upregulated in PosELSK cells (26 genes). Black dots represent non-significant genes. Dashed lines represent the cut-offs.

The GO terms associated with repopulating ELSKs were also involved in developmental processes (Table 5.4). However, interestingly, repopulating ELSKs also seemed to upregulate genes involved in cell adhesion (Table 5.4). Again, immune responses are associated with genes upregulated in NegELSK cells, and genes involved in myeloid leukocyte differentiation are also upregulated (Table 5.5).

Table 5.4 Manually curated gene ontology terms based on upregulated genes in PosELSK compared to NegELSK.

GO term	Genes DE	P-value
integrin binding	2	0.00698628
biological adhesion	5	0.00799701
anchored component of membrane	2	0.01516967
developmental process	11	0.02558864
cell adhesion molecule binding	2	0.02323818
multicellular organism development	10	0.02393103
regulation of cellular protein metabolic process	6	0.03034427
cell surface receptor signalling pathway	6	0.03050336
cell adhesion	4	0.03616079
metabolic process	15	0.04267028

Table 5.5 Manually curated gene ontology terms based on upregulated genes in NegELSK compared to PosELSK.

GO term	Genes DE	P-value
myeloid leukocyte differentiation	5	7.31E-06
regulation of immune system process	10	8.57E-06
response to external stimulus	12	3.02E-05
immune system process	12	5.37E-05
myeloid cell differentiation	5	0.00014
cell surface receptor signalling pathway	11	0.000153
myeloid leukocyte activation	4	0.000153
response to stress	13	0.000258
cell cycle phase	2	0.000389
inflammatory response	5	0.001655

Because the GO terms are not meant to be used as a guide rather than to form definitive conclusions, the differentially expressed genes are listed in Table 5.6 and Table 5.7.

Of the 26 upregulated genes in PosELSK cells, three were predicted genes. Interestingly, *Vwf* has been an extensively studied marker of HSCs¹⁸⁴. In fact, a reporter for *Vwf* has been generated to mark putative platelet-primed HSCs¹⁰⁹ (section 1.2.6). *Pld3* has also been previously identified to be upregulated in highly enriched LT-HSCs compared to haematopoietic cells with less self-renewal capacity¹⁸⁴. Lipoprotein lipase (*Lpl*) has recently been shown to be required for HSPC maintenance, by regulating free fatty acids supply⁴⁴¹. Lamin A (*Lmna*) is also interesting as it was reported to regulate epigenetic and chromatin architecture upon HSC ageing⁴⁴², and is important for other adult stem cell systems⁴⁴³. *Tcf7l1* has been implicated in BCR-ABL acute lymphoblastic leukaemia (ALL)⁴⁴⁴.

The upregulation of integrin beta5 (*Itgb5*) is interesting as this particular integrin has not previously been associated with HSC function, whereas integrin beta1 for example has been shown to be important for HSCs. This suggests that HSCs in culture may require different integrins and associated signalling. Similarly, *Serpinc6b* is a protease that has not been associated with HSCs; The main serpin associated in haematopoiesis is serpin A.

Table 5.6 Significantly upregulated genes in PosELSK vs NegELSK cells.

Genes	logFC	logCPM	P-value	FDR
<i>Gm53</i>	7.54891634	-0.1565622	8.41E-05	0.0315636
<i>Ifit3b</i>	7.36482925	-0.1737339	7.73E-05	0.03063083
<i>Rgmb</i>	6.93313664	-1.2617855	2.04E-05	0.01292519
<i>Lpl</i>	6.13401797	1.16789263	7.72E-05	0.03063083
<i>Htra4</i>	5.9367492	-0.8681886	9.82E-05	0.03404528
<i>Prok1</i>	4.37913488	0.05053911	8.36E-05	0.0315636
<i>Gm18066</i>	4.26948991	-0.5231257	0.00014398	0.04543924
<i>Lmna</i>	3.26377273	2.94085242	3.68E-10	3.06E-06
<i>Tcf7l1</i>	2.95415442	1.55752451	8.86E-05	0.03207164
<i>Rab3a</i>	2.79958824	0.0901563	6.65E-05	0.02911376
<i>Gm45033</i>	2.48857355	0.16267108	0.00015428	0.04645133
<i>Vwf</i>	2.42917114	4.59764933	5.19E-06	0.00618873
<i>Klhl3</i>	2.27400105	3.09622884	1.10E-05	0.00917956
<i>Serpinb6b</i>	1.55443811	4.1173494	9.90E-08	0.00041207
<i>Itgb5</i>	1.2620512	4.81018371	3.70E-05	0.01986446
<i>Plxnb2</i>	1.19192701	6.57904222	3.63E-05	0.01986446
<i>Gbp3</i>	1.02659679	3.74272348	7.44E-05	0.03063083
<i>Pld3</i>	0.95016956	3.88342083	0.00017308	0.04802355
<i>Fam43a</i>	0.89811432	5.6574484	5.04E-05	0.0246611
<i>Gpr146</i>	0.82723086	4.16874348	0.00016481	0.04730663
<i>Glb1l</i>	0.81507663	4.50963369	0.00014739	0.04543924
<i>Arhgap31</i>	0.80671095	5.62295713	6.18E-06	0.00618873
<i>Retreg1</i>	0.68237705	5.23711245	7.60E-05	0.03063083
<i>Aplp2</i>	0.61341178	6.15472426	6.32E-06	0.00618873
<i>Nfic</i>	0.5890611	6.87835405	2.54E-05	0.0151006
<i>Ncoa7</i>	0.54166476	6.06302169	0.00012966	0.0440531

Of the genes upregulated in NegELSK cells, one of them is a long noncoding RNA (lncRNA), *6530402F18Rik*, and three of them are predicted genes. Of interest, *Cebpa* is upregulated, which is an important haematopoietic transcription factor that regulates myeloid lineage commitment⁴⁴⁵. The upregulation of *Myc* is also interesting as it's a well known oncogene involved in cell proliferation²⁹⁶. Of note, *Cdk6* is also interesting because of it's role in regulating HSC exit from quiescence and has previously been shown to be upregulated in ST-HSCs²⁸. The upregulation of *Cd69* is also interesting, as it is considered an early activation marker in HSCs and other immune cells, and involved in T cell differentiation⁴⁴⁶. As mentioned before, *Cd244* is one of the SLAM family surface receptors that is upregulated upon HSC transition to MPPs¹²¹, however it is also expressed by multiple mature blood cell types. Of note, other genes associated with differentiated

cell types include *Mpo*, which encodes for myeloperoxidase expressed highly in neutrophils and granulocytes and *Cpa3*, which encodes for Carboxypeptidase A3 is gene expressed highly in mast cells. Interestingly, while *Itgb5* is upregulated in PosELSK cells, *Itgb3* is upregulated in NegELSK cells, suggesting a difference in integrin signalling.

Table 5.7 Significantly upregulated genes in NegELSK vs PosELSK cells.

Genes	logFC	logCPM	P-value	FDR
<i>Mdn1</i>	-0.4064294	8.7985989	1.45E-04	0.04543924
<i>Myc</i>	-0.5815024	8.02111435	1.58E-04	0.04645133
<i>6530402F18Rik</i>	-0.7244995	4.96814152	4.06E-05	0.0211374
<i>Epop</i>	-0.7328131	5.14398147	9.06E-06	0.00793693
<i>Eif2b3</i>	-0.7894685	4.44702733	6.58E-05	0.02911376
<i>Cdk6</i>	-0.861117	9.59079299	2.10E-05	0.01292519
<i>Arhgap6</i>	-0.9415149	5.4337907	1.7775E-06	0.00269023
<i>Ntng2</i>	-1.1088195	4.31091602	8.20E-06	7.58E-03
<i>Zfp462</i>	-1.134887	4.60539109	1.59E-04	0.04645133
<i>Cd244a</i>	-1.298679	5.17576124	1.60E-05	0.01222766
<i>Cd69</i>	-1.2999675	6.61899075	5.3148E-06	0.00618873
<i>Trps1</i>	-1.3953718	5.18779564	1.81E-07	0.00060269
<i>Slc24a3</i>	-1.7611088	4.20162248	1.42E-06	0.00236132
<i>Itgb3</i>	-1.8730119	5.43261622	9.91E-07	0.00206229
<i>Slc39a4</i>	-1.9497932	2.66102158	1.68E-04	0.04745925
<i>Cpa3</i>	-2.0051984	4.69596414	1.82E-04	0.04954771
<i>Dtx4</i>	-2.2401976	4.26511532	1.38E-04	0.04543924
<i>Cst7</i>	-2.2708136	4.90415677	4.7617E-07	0.00119204
<i>Siglecf</i>	-2.6064022	4.26864586	6.11E-06	0.00618873
<i>Rab44</i>	-2.6456331	7.28881497	1.8507E-05	0.01292519
<i>Tmem156</i>	-3.031765	2.81119377	4.7656E-05	0.02404159
<i>Gm34589</i>	-3.1779078	1.19905711	1.44E-04	0.04543924
<i>Cebpa</i>	-3.2290337	6.11901204	1.78E-10	2.9695E-06
<i>Hsd11b1</i>	-3.3213592	1.31883217	5.77E-05	0.02735447
<i>Ctsg</i>	-3.4840521	6.01169637	8.53E-05	0.0315636
<i>Rgs1</i>	-3.8061441	3.77574096	2.0526E-05	0.01292519
<i>Mpo</i>	-4.6488139	11.0051616	2.0531E-05	0.01292519
<i>Ccr1</i>	-5.1952601	1.66837887	2.70E-05	0.01551143
<i>Rnf43</i>	-5.473591	1.61545879	9.48E-05	0.03356179
<i>Fcgr3</i>	-5.6082288	3.5186246	6.24E-06	0.00618873
<i>Gm7967</i>	-6.2207428	3.62027945	5.01E-07	0.00119204
<i>Gm15581</i>	-6.7281744	0.29748443	1.62E-05	0.01222766
<i>Hp</i>	-6.8432125	3.79517351	1.13E-06	0.00209799
<i>Hdc</i>	-9.1165786	3.82239234	2.8155E-09	1.5624E-05
<i>Syce1</i>	-9.3790474	1.89716573	5.92E-05	2.74E-02

Overall, this confirms the functional outcomes and associated GO terms, that PosELSK cells are enriched for HSC genes and that NegELSK cells have a gene expression profile suggestive of activated ST-HSCs or more restricted progenitors. In the meantime, many of these differentially regulated genes have not been previously associated with HSC function or haematopoiesis in general. These novel genes identified could be taken forward in functional validations to see if they have essential functions in HSC expansion or used in *in vitro* qPCR screens for expanding HSCs.

5.4 Identification of a molecular programme for “expanded” HSCs

In order to identify the genes that associate most strongly with ex vivo expanded HSCs, the samples were first visualised by principal component analysis (PCA). As shown in Figure 5.13, the resulting distributions in the PCA are very similar to above MDS plots. Importantly, PC1 is responsible for 35.6% of the variances in gene expression and it also correlates very significantly with donor chimerism ($r^2 = 0.27$, $p < 0.05$), contribution to GM ($r^2 = 0.41$, $p < 0.01$), contribution to T cells ($r^2 = 0.2$, $p < 0.05$), contribution to B cells ($r^2 = 0.19$, $p < 0.05$), repopulation ($r^2 = 0.47$, $p < 0.001$) and MoIo score ($r^2 = 0.55$, $p < 0.0001$).

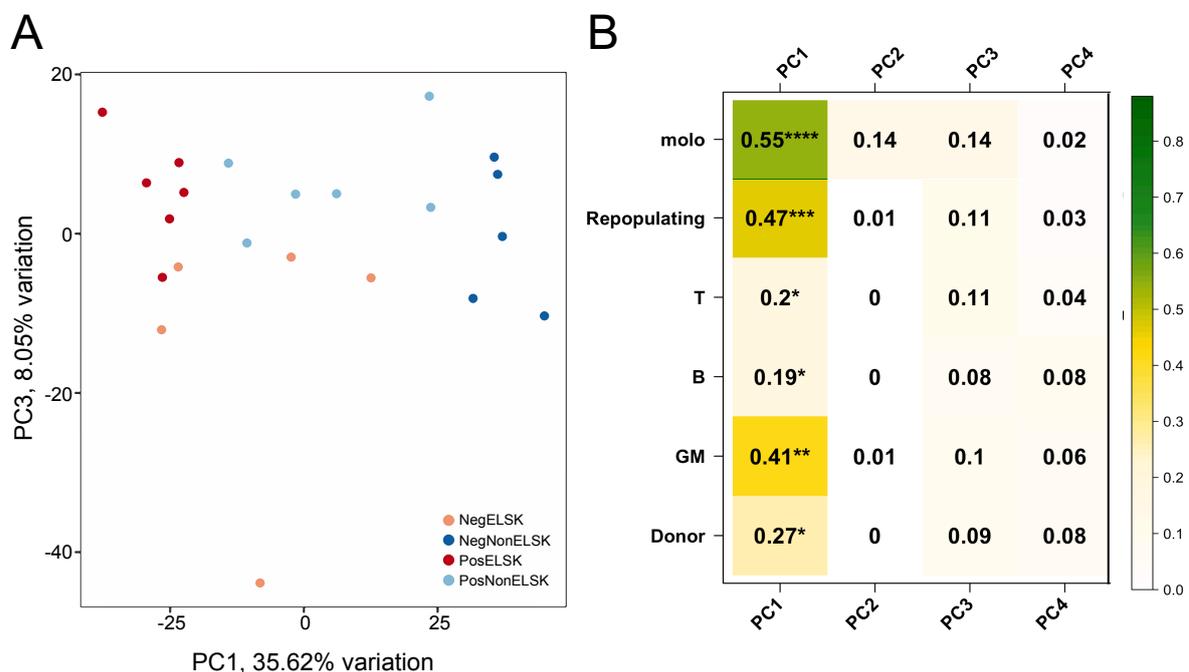


Figure 5.13 PCA of all samples indicate that PC1 correlates highly with functional outcomes.

A) PCA of all samples coloured by their respective groups.

B) Pearson correlation (r^2) between each principle component and functional or statistical outcome. Molo, MoIo signature score; Repopulating, binary yes or no correlation for samples; Donor, Donor chimerism.

By using the loading plots for PC1, the genes that drive differences in PC1 can be extracted. As listed in Table 5.8, 30 genes were significantly driving (component loading <-0.05) the negative PC1 vector. Using these PC1 loading genes, a repopulation signature was created in order to be able to predict and/or monitor the functional HSC content of an expansion culture (Table 5.8).

Table 5.8 List of 30 genes identified by PC1 loading plot to be signature genes for expanding HSCs.

<i>Epb41l3</i>	<i>Ncam2</i>	<i>Ptk2</i>	<i>Esam</i>	<i>Hlf</i>
<i>Camk2b</i>	<i>CCm2l</i>	<i>Procr</i>	<i>Gm38197</i>	<i>Arx</i>
<i>Robo1</i>	<i>Slamf1</i>	<i>Klhl4</i>	<i>Nrk</i>	<i>Bcam</i>
<i>Neur1b</i>	<i>Tgfbr3</i>	<i>Myof</i>	<i>Zfp532</i>	<i>Palld</i>
<i>Vill</i>	<i>Ryk</i>	<i>Sel1l3</i>	<i>Dlg2</i>	<i>Mecom</i>
<i>Gimap4</i>	<i>Mpdz</i>	<i>Tcf15</i>	<i>Prdm16</i>	<i>Prex2</i>

As shown in Figure 5.14, the signature score was significantly higher in repopulating cells compared to non-repopulating cells as well as ELSK cells compared to non-ELSK cells. As shown by the overlay of the signature score on the MDS plot in Figure 5.14, the signature score can better predict repopulation categories compared to MolO score (Figure 5.7).

Unsurprisingly, EPCR (*Procr*) was on this list, as well as other known HSC cell surface markers such as CD150 (*Slamf1*)¹²¹ and ESAM (*Esam*) (Table 5.8)¹⁷². Interestingly, there were 3 other surface membrane protein genes, *Robo1*, *Ncam2* and *Myof*, that had not been implicated with HSC function before. These genes could represent potential new markers for monitoring or isolating *ex vivo* expanded HSCs and might serve to replace *Fgd5*, thereby potentially averting the need for a reporter strain. Also, amongst the list are several genes well-accepted to be transcriptional regulators of HSC self-renewal, including *Mecom*⁴⁴⁷, *Hlf* and *Prdm16*, which have also recently been shown to be upregulated in expanded human HSCs²⁸⁵. In a separate study, *Hlf* has also been recently reported in a preprint journal to be expressed in expanded human HSCs⁴⁴⁸. Overexpression of *Prdm16*, also known as Mel1, has been shown to lead to expansion of HSCs comparable to that achieved by *Hoxb4* overexpression³⁶⁰.

Given that TGF- β has been implicated in HSC biology, it is perhaps not surprising to find on the list TGF- β receptor 3 (*Tgfbr3*)^{137,449,450}. *Camk2b*, a calcium/calmodulin-dependent protein kinase, is interesting because calcium signalling has recently been implicated in HSC maintenance³⁸⁵. *Klhl4*, is the second Kelch family protein to appear in this thesis, with *Klhl3* found to be upregulated in PosELSK cells vs NegELSK cells in the previous section, suggesting an undiscovered role for this

protein family. Interestingly, a considerable number of genes have not been previously implicated in HSCs or haematopoietic function.

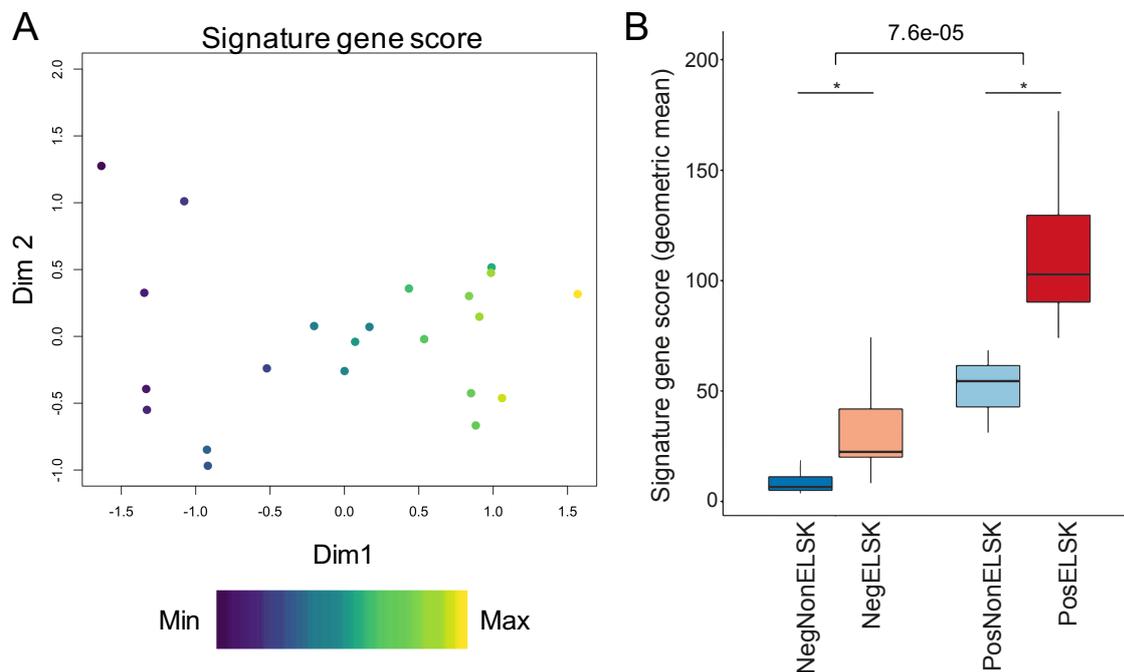


Figure 5.14 Signature gene score is significantly higher in repopulating and ELSK cells.

A) MDS plot of all samples coloured by their signature gene score.

B) Geometric mean of signature gene score of the categories. T test. *= p<0.05.

5.4.1 Single cell HSPC transcriptome displays heterogeneous expression of signature genes

To further explore the role that the 30 signature genes play in HSC function, a single cell HSPC transcriptomic dataset generated by Nestorowa et al¹⁹⁰ was mined for their expression. Over nine HSPC populations (LTHSC, STHSC, MPP1, MPP2, MPP3, CMP, GMP, LMPP and MEP), the gene expression of signature genes in individual cells are shown in an array of violin plots (Figure 5.15). 15 of the genes were below expression limits, perhaps suggestive of the differences between fresh and cultured cells.

Consistent with originally published findings, several known HSC markers such as *Procr*, *Slamf1* and *Esam* are more highly expressed in phenotypically more primitive HSPC populations. As expected, the genes that were not previously associated with HSC function such as *Robo1*, *Ptk2*, *Gimap4* and *Mpdz*, are lowly expressed across all cell types, reaffirming their limited role in native HSPC biology. Whereas genes such as *Vill* is relatively broadly expressed across the haematopoietic compartment.

Whether by the surface markers identified here or by using qPCR, these signature genes, when validated and proven, might be beneficial for screening *ex vivo* cultures of HSCs to determine if cultures contain large numbers of HSCs or not.

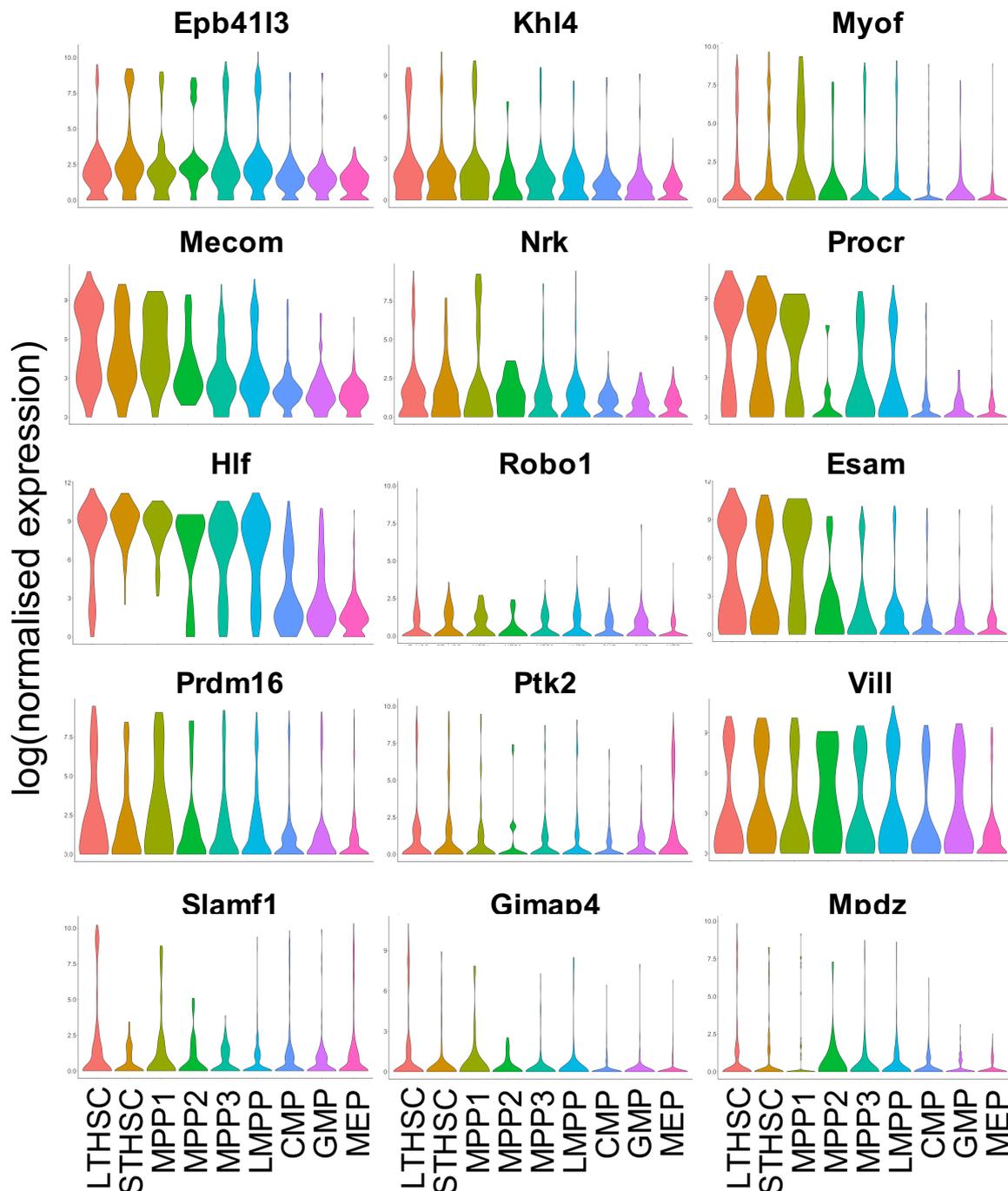


Figure 5.15 Gene expression patterns of signature genes across HSPC subpopulations.

The expression of 30 signature genes are visualised in the scRNA-seq dataset generated by Nestorowa et al.¹⁹⁰. Only 15 genes are displayed because 15 genes were not detected in the dataset. The HSPC populations are ordered by level of differentiation from lowest (LTHSC) to highest (MEP).

5.4.2 MapK signalling pathway overrepresented in signature genes

Knowing the genes that define expanding HSCs are important and useful for screening cultures, however, it does not necessarily provide the knowhow to increase expansion efficiencies. To this end, pathway analysis was again used to complement the analysis and provide insights to HSC expansion signalling pathways.

As shown in Figure 5.16, only 12 pathways were significantly associated with the 30 signature genes. Interestingly, Map Kinase (MapK) signalling pathway seems to be highly associated with signature genes. Functional validations by inhibiting or stimulating key players of MapK signalling cascade may be interesting to test if the pathway is indeed important for HSC expansion. Of note, some pathways typically associated with the neuronal system were also identified, including NMDA receptor regulation, protein-protein interactions at the synapses and neurexins and neuroligins. It is possible that such pathways that have been historically identified in the neuronal system, have a secondary effect in the haematopoietic system. More investigation would be needed to tease apart this connection.

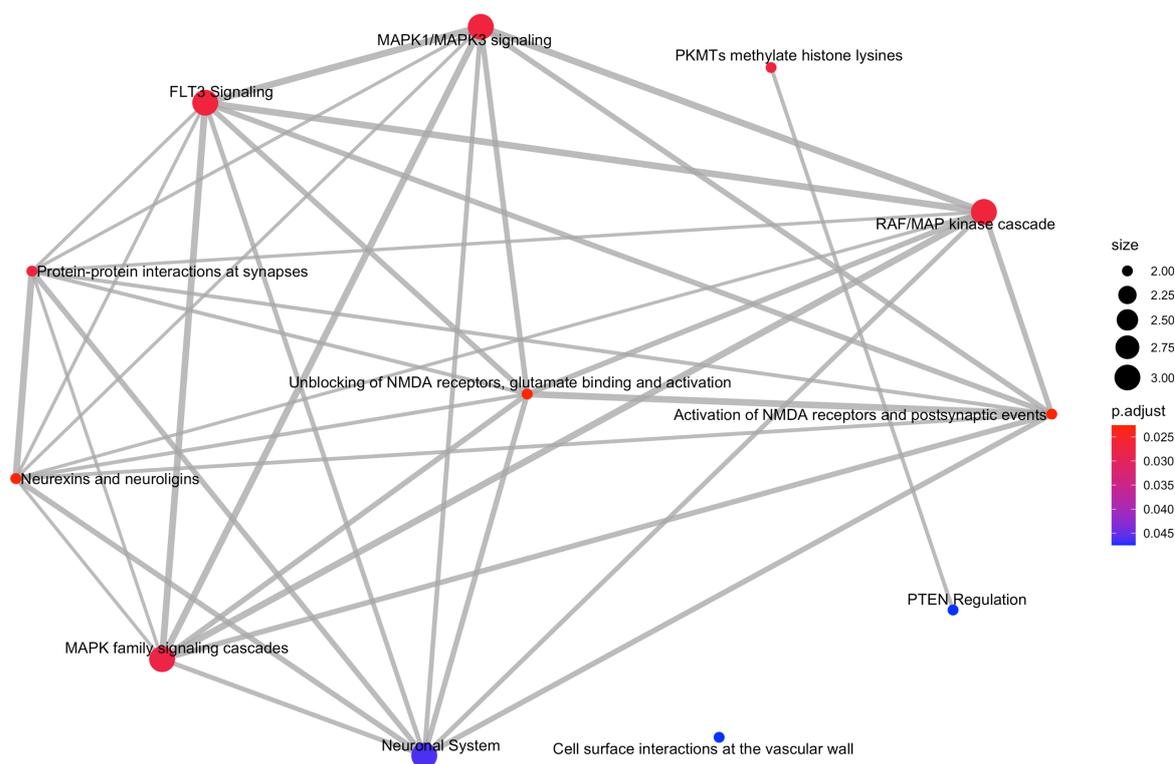


Figure 5.16 Pathways associated with signature genes.

Reactome pathways significantly associated with the list of signature genes in Table 5.8. Lines connect pathways that have shared genes, and pathways are coloured by adjusted p-values from more significant (red) to less (blue).

5.5 Non-ELSKs provide insights to feedback signals

As mentioned above, in all cultures, the non-ELSK cells vastly outnumber ELSK cells. Even the few exceptional clones with very high percentages of ELSK cells, still have a majority (over 50%) non-ELSK cells by phenotype. Thus, if negative feedback in cultures is an important aspect of HSC expansion, as has been previously suggested⁴¹⁶, then it would be crucial to look at non-ELSK cells, as they might be expected to produce the majority of feedback signals within the culture, whether by secretory proteins or cell-to-cell contact.

5.5.1 The transcriptomes of Non-ELSK cells correlate with functional outcome of the clone

An MDS of non-ELSK cells was plotted with the samples coloured in by the respective donor chimerism of their ELSK counterparts (Figure 5.17). Interestingly, even within non-ELSK cells, there is a separation between repopulating clones and non-repopulating clones.

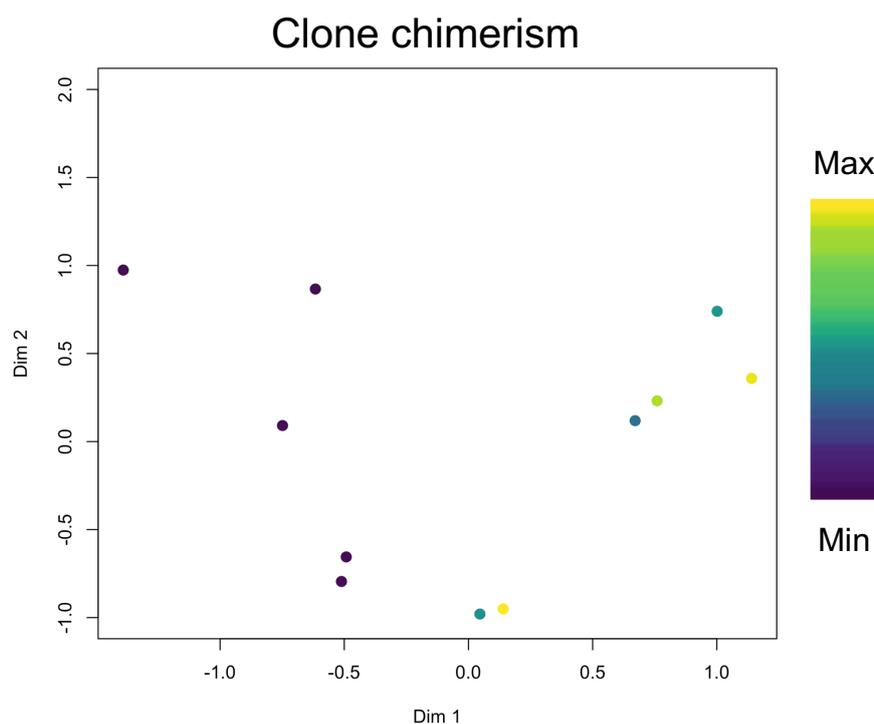


Figure 5.17 PosNonELSK cells separate from NegNonELSK cells and correlate with chimerism of the clone.

MDS plot of just the non-ELSK cells, coloured by their clone's respective donor chimerism from matched ELSK cells. Colours are scaled by donor chimerism, from max (yellow) to lowest (purple).

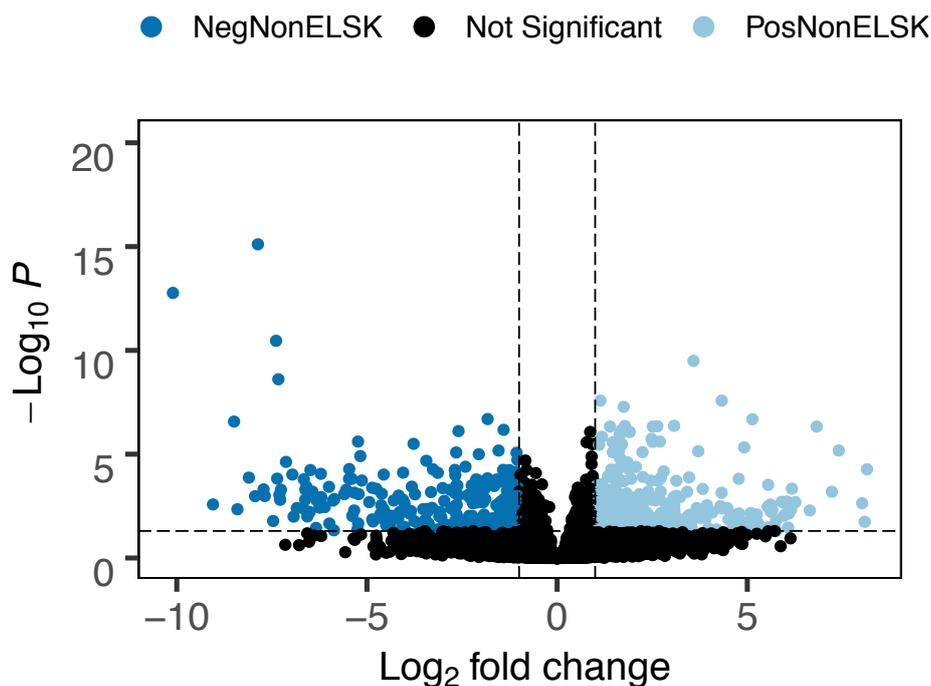


Figure 5.18 Volcano plot of differentially expressed genes between NegNonELSK and PosNonELSK cells.

LogFC and p-value of all 16648 genes. Dark blue dots represent genes upregulated in NegNonELSK cells (426 genes) and light blue dots represent genes upregulated in PosNonELSK cells (530 genes). Black dots represent non-significant genes. Dashed lines represent the cut-offs.

Of the 16648 genes, 956 were differentially expressed between PosNonELSK cells (530 genes upregulated) and NegNonELSK cells (426 genes upregulated) (Figure 5.18). GO term analysis suggests that NegNonELSK cells are more differentiated, with GO terms clusters for immune activation and cell activation. In contrast, PosNonELSK cells are upregulated for GO terms associated with developmental processes, signal transduction, cell localisation and adhesion.

Table 5.9 Gene ontology terms of upregulated genes in PosNonELSK cells vs NegNonELSK cells.

GO term	Genes DE	P-value
developmental process	205	1.31E-16
anatomical structure development	196	1.42E-16
intracellular signal transduction	113	2.64E-16
multicellular organism development	183	3.40E-16
regulation of signal transduction	114	9.44E-15
regulation of cell communication	128	2.00E-14
cellular developmental process	154	8.66E-14
localization	189	1.25E-13
cell adhesion	63	3.20E-11
biological adhesion	63	4.68E-11

Table 5.10 Gene ontology terms of upregulated genes in NegNonELSK cells vs PosNonELSK cells.

GO term	Genes DE	P-value
immune system process	111	4.56E-23
defense response	83	8.33E-23
response to stress	127	6.46E-20
regulation of immune system process	72	2.56E-18
cell migration	70	1.43E-17
response to external stimulus	96	3.73E-17
cell motility	73	5.07E-17
localization of cell	73	5.07E-17
cell activation	59	6.63E-17
inflammatory response	46	9.39E-17

5.5.2 ROS associated pathways are upregulated in non-repopulating clones.

Interestingly, pathways associated with genes upregulated in NegNonELSK suggests that the cells are involved in ROS and RNS production in phagocytes and antimicrobial peptides (Figure 5.19). This supports the idea that they are more differentiated immune cells that may lead to secretion of immune molecules. ROS has also been previously linked to HSC differentiation³⁴, and this may be a mechanism by which Non-ELSK cells influence the fate of HSCs within F12 PVA cultures.

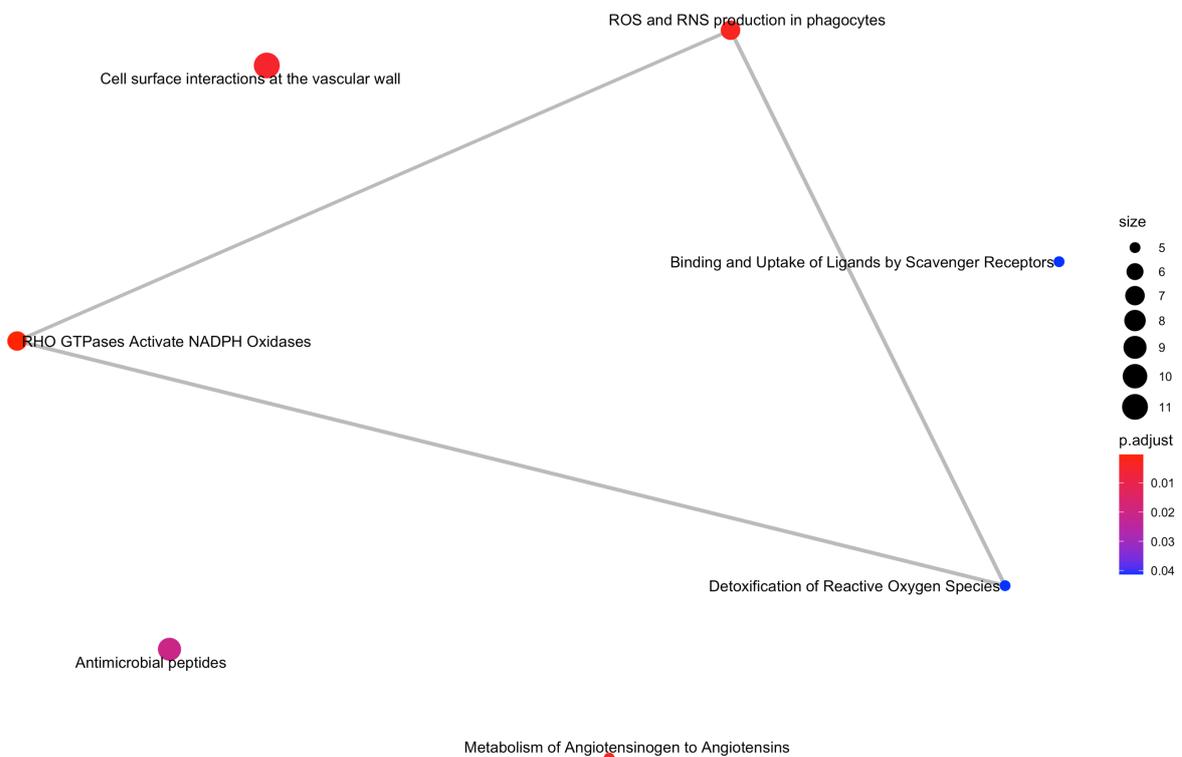


Figure 5.19 Pathways associated with genes upregulated in NegNonELSK cells vs PosNonELSK cells.

Reacomme pathway analysis of genes upregulated in NegNonELSK cells. Lines connect pathways that have shared genes, and pathways are coloured by adjusted p-values from more significant (red) to less (blue). The regulation of ROS is amongst the interesting pathways identified.

5.6 Secretome analysis identifies targets for positive and negative regulation

As mentioned earlier during the explanation of the experimental design, the conditioned media was collected during every medium change over the 28 days. The media was analysed by mass spectrometry in order to identify the proteins in the media. In the first experiment with 2 runs, CM from 6 clones (3 repopulated and 3 non-repopulated) were used, comprising of early, middle and late timepoints for a total of 18 samples. Additionally, a pooled control was used in both runs to allow for comparisons.

Due to high variability in the dynamic range between samples, the quantification of protein abundances was deemed unreliable. Therefore, the presence and absence of proteins were investigated between repopulating clones and non-repopulating clones. A very stringent cut-off was used to filter for reliable protein identifications (>3 unique peptide sequence), after which only 477 proteins were identified. Amongst them, 37 were uniquely identified in repopulating clones compared to non-repopulating clones and 40 were uniquely identified in non-repopulating clones (Figure 5.20).

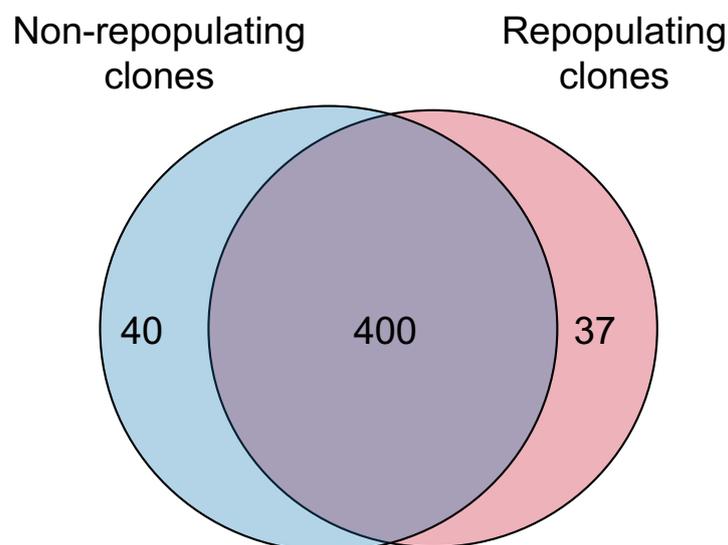


Figure 5.20 Venn diagram of proteins identified in the secretome of repopulating and non-repopulating clones.

From two TMT labelled runs, 477 proteins were identified in 18 conditioned media samples; 6 clones with 3 different timepoints (early, middle and late). 40 were uniquely identified in non-repopulating clones and 37 were uniquely identified in repopulating clones. Strict cut-offs were used to define reliably identified proteins, with a minimum of 3 unique peptide for each protein.

Notably, HPX, which was previously identified within EL08 CM, was also found in this dataset. However, further runs will be needed to quantify and compare abundances in repopulating and non-repopulating clones. Other notable proteins identified includes TGF- β 1, which has been previously targeted as a negative regulator of human HSC expansion⁴¹⁶; IL-16, an immunomodulatory cytokine⁴⁵³ and myeloid derived growth factor (MYDGF), which has no known effect on HSCs⁴⁵⁴.

Although the majority of proteins identified are predicted to be intracellular proteins, it is conceptually possible that they can exert an effect on HSCs. However, more samples would have to be run in order to find targets with higher confidence. Overall this serves as a proof of principle that proteomic approaches can be used to discover novel feedback signals in media of HSCs.

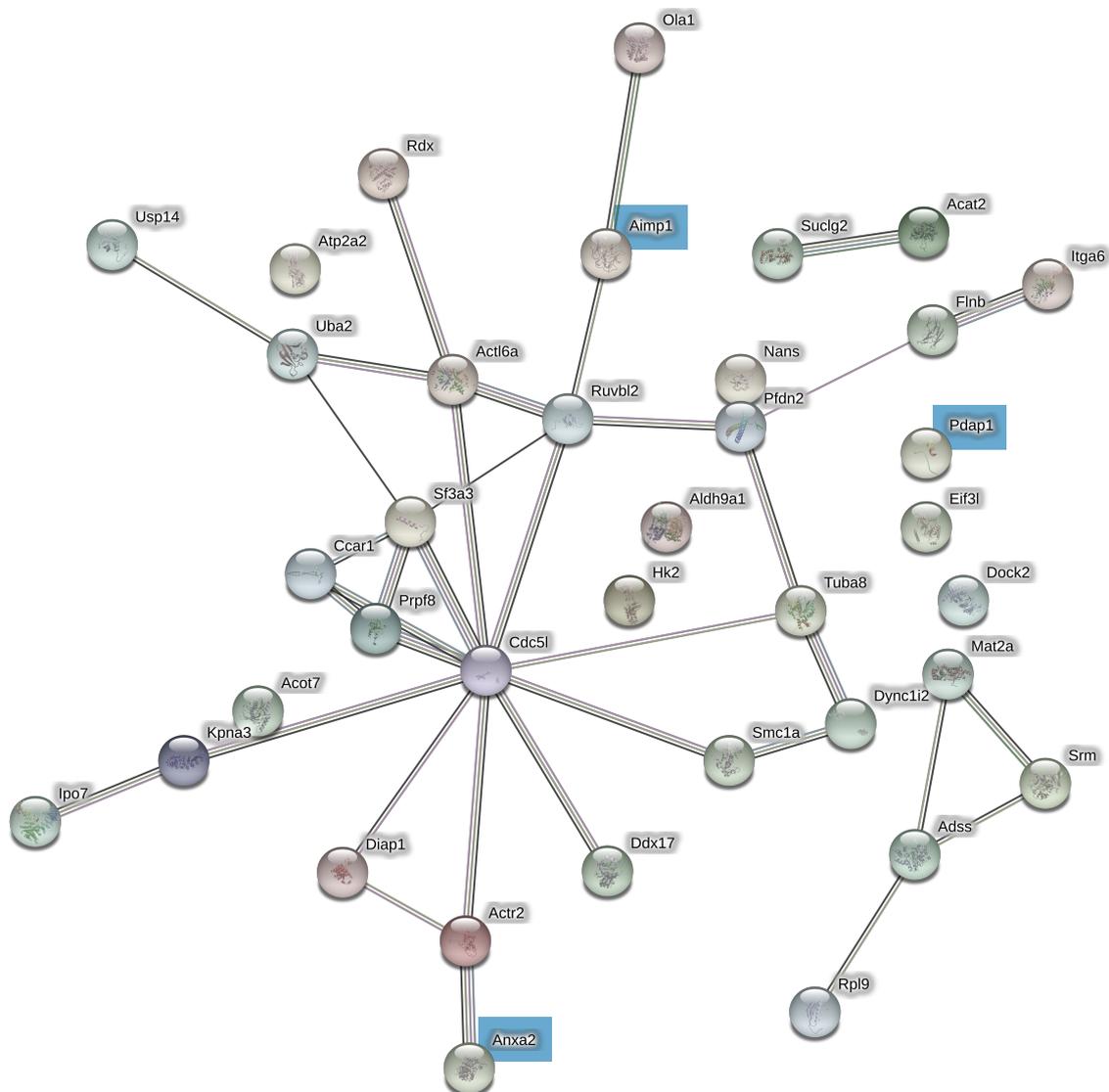


Figure 5.22 Interaction map of unique proteins in media from non-repopulating clones.

40 proteins found to be unique in media from non-repopulating clones, with at least 3 unique peptides identified. Predicted extracellular proteins are highlighted in blue.

Overall, the experiments in this chapter confirm that *Fgd5* and EPCR robustly marks HSCs *in vitro*. By using this reporter strategy to separate out repopulating and non-repopulating cells in long-term cultures, the molecular state of expanded HSCs was characterised. Here, I show that the molecular profile of expanded HSCs resembles that of freshly isolated HSCs, but with some differences that can provide novel markers for expanding HSCs and insights into self-renewal regulation. By looking at the secretome of repopulating clones vs non-repopulating clones, potential HSC regulators were also discovered

6 DISCUSSION

6.1 Context of study

In vitro HSC expansion has been a long-standing goal in the field, with substantial clinical implications for improving HSCT, *ex vivo* production of mature blood cells, and gene therapy. Moreover, the increased knowledge of HSC expansion *ex vivo* may yield clinical insights to how HSCs expand pathologically *in vivo* as in the case of haematological malignancies. Decades of research, mostly through gene overexpression studies, have elucidated many intrinsic regulators of HSC cell fate *in vivo*, including *Hoxb4*²⁹⁷, *Fbxw7*³⁵⁶, *Dppa5*³⁵⁷, *Prdm16*³⁶⁰ amongst others^{267,360}. However, strategies requiring genetic integration have the caveat of risking leukaemic initiation, due to undesirable genetic integration or permanent activation of self-renewal programmes. In some cases, such as for NUP98-HOX fusion genes, near limitless HSC expansion can be achieved *in vitro*, suggesting that it is possible to culture and expand HSCs for long periods in the absence of a haematopoietic niche. Furthermore, numerous extrinsic regulators, especially haematopoietic cytokines and growth factors have been investigated for their effects on HSC self-renewal¹⁹ with various combinations suggesting that HSC maintenance can be achieved in short term cultures with some reporting modest expansion. Despite this progress, however, the exact signalling mechanisms that link extrinsic signals and intrinsic self-renewal regulators - and ultimately drive the HSC self-renewal programmes - remain elusive.

One of the main barriers that has hindered the study of HSC expansion has been the lack of robust markers to isolate stem cells. Although HSC isolation strategies have continually improved, it is still currently not possible to isolate HSCs to 100% purity, and most gating strategies that achieve a high purity (>10%), will also exclude some HSCs^{121,184,293}. Compounding this challenge of marking HSCs *in vivo*, is the added difficulty of marking HSCs *in vitro*, since many of the surface markers change their expression during *ex vivo* culture²⁸⁴. There is therefore a substantial unmet need to develop a larger repertoire of reliable *in vitro* markers to report for functional HSCs in culture.

Another barrier to identifying HSC expansion conditions, is the lack of defined conditions. When something as minute as differences in the batch of bovine serum albumin used, can cause drastic differences in *in vitro* outcomes⁴¹⁷, it becomes difficult to systematically test additional self-renewal

factors. Very recently, a study by Wilkinson et al. demonstrated a culture system with completely chemically defined components that is able to expand HSCs by ~200 to 900- fold in a duration of 28-days²⁸³, something that has never been achieved before. Even more surprisingly, the culture system only relies on two cytokines, SCF and TPO, without additional self-renewal regulators, which sets the stage for it becoming a standardised base culture medium for future investigations into additional self-renewal regulators. Even though this represents a significant breakthrough in the HSC expansion field, there are still some outstanding questions. For one, single HSCs cultured in these conditions display substantial clonal heterogeneity in terms of proliferation and transplantation outcomes. The driver(s) of this clonal heterogeneity are still poorly understood. For this, it is extremely important to understand the molecular basis behind differences in repopulating HSCs and non-repopulating HSCs in expansion cultures, which allows for targeting of specific pathways to increase expansion efficiency. Excitingly, the ability to expand HSCs has also afforded new opportunities to molecularly characterise HSCs with techniques, such as global proteomics, gene editing, metabolomics and immunoprecipitation, that were previously not possible in the field. In addition, the molecular characterisation of this new culture system would have wider implications such as understanding the activation and quiescence re-entry programmes that the cells must execute in clinical transplantations.

In this thesis, I focus on developing a novel *in vitro* HSC reporter system using the $Fgd5^{ZsGreen^*ZsGreen/+}$ reporter mouse. Using this tool, my thesis focuses on the identification of key molecules which define the expanded HSC state. I showed that, in combination, $Fgd5$ and EPCR expression were reliable markers for functional HSCs both *in vivo* and *in vitro*. I further discovered through this that $Fgd5$ expression alone is not sufficient to isolate HSCs *in vivo*, leading to identification and characterisation of a previously unreported population of iNKT1 cells marked by $Fgd5$ expression. In chapter 4, I demonstrated the utility of $Fgd5$ and EPCR expression for screening HSC maintenance culture conditions. Finally, I exploited the reporter strategy to separate functional HSCs from non-stem cells in current state-of-the-art HSC expansion conditions and provided the first molecular characterisation of *ex vivo* expanding HSCs.

6.2 Summary of major findings

In chapter 3, the $Fgd5^{ZsGreen^*ZsGreen/+}$ reporter mouse was investigated for its ability to mark HSCs. Consistent with originally published findings¹⁸⁸, I showed that nearly 100% of phenotypically defined LT-HSCs (ESLAM²⁹³), are positive for $Fgd5$ expression. However, contrary to published evidence¹⁸⁸, $Fgd5^+$ cells were not all phenotypic HSCs. In fact, ~60% of the cells were discovered

to be non-stem cells and were EPCR⁻, CD48⁺, CD150⁻ and c-Kit⁻. These cells were unable to grow in standard HSC culturing conditions and were unable to produce cells in a wide range of myeloid and lymphoid progenitor cell assays. Further phenotypic characterisation determined that this population was not homogeneous and contained at least two fractions of cells: CD5⁺CD3⁺NK1.1⁺Sca-1⁺ cells and CD244⁺ cells (negative for the other markers). Through multiple characterisation experiments, including surface marker phenotyping, gene expression profiling and *in vitro* functional assays, I showed that the first of these fractions represented a subset of iNKT1 cells. These cells are reactive to α -GalCer CD1d tetramers and express high levels of IFN γ both of which are hallmarks of iNKT1 cells. Interestingly of total phenotypic iNKT1 cells, only 20% are *Fgd5*⁺, suggesting that *Fgd5*⁺ iNKT1 cells might represent a novel subfraction of iNKT1 cells. To our knowledge, this is also the first time that *Fgd5* has been implicated in NKT cell biology.

Chapter 4 looks at the HSC fraction of *Fgd5*⁺ cells, where in combination with EPCR, it can be used as an effective two-colour strategy to enrich relatively pure HSCs. *In vitro*, *Fgd5* and EPCR expression correlated highly with LSK percentages – a traditionally used *in vitro* measurement of HSCs. After culturing, only F^{hi}E^{hi} cells contained multilineage reconstitution ability compared to F^{lo}E^{lo} cells. Even from less pure culture-initiating cells (*Fgd5*⁺EPCR⁺CD150⁻ cells), *Fgd5* and EPCR were able to mark the stem cells when the vast majority of cells had lost *Fgd5* and EPCR expression.

Utilising this novel reporter strategy, I screened EL08 CM, which has been shown to support HSCs^{335,455}, and showed that prolonged conditioning is detrimental to its supportive effects. Even so, all CM, regardless of length of conditioning afforded a survival benefit to HSCs compared to non-conditioned control medium. Interestingly, CM_{1D} increased HSC clone sizes, whilst other CM had no significant effects on clone sizes. Having established that shorter conditioning times are better at maintaining HSCs, I analysed the various CM using mass spectrometry and identified previously reported HSC self-renewal regulators as well as potentially novel ones. This is also the first proteomic screen of EL08 CM to be reported. In initial validation experiments, OPN, PTN, IGFBP2 and IGFBP4 were tested for their ability to expand HSCs. While none of the factors tested, gave a survival benefit to HSCs, compared to control, HSCs cultured in 100ng/mL of PTN had higher FLSK content. During this period, a new more powerful protocol was developed for mouse HSC expansion²⁸³ and the novel reporter system was immediately useful in screening and validating the system. I showed that IL-11 is beneficial in HSA cultures but redundant in PVA

cultures. Most importantly, functional HSC activity correlated strongly with phenotypic HSCs marked by the reporter system.

Chapter 5 dissects the observed clonal heterogeneity within these HSC expansion cultures. In serial transplantation experiments, I showed that the percentage of phenotypic HSCs in 28-day cultures, defined by this reporter system, strongly predicts *in vivo* outcomes in transplantation experiments. By using the reporter system to separate repopulating from non-repopulating cells, I present the first molecular characterisation of *ex vivo* expanded HSCs. The molecular states of these cells suggest that *ex vivo* expanding HSCs resemble freshly isolated HSCs, but also contain minor differences. I present a molecular gene signature for expanding HSCs, containing previously described HSC self-renewal regulator genes as well as novel genes not previously implicated in HSC biology. Amongst the genes upregulated in cultured HSCs, Rho GTPase and MapK pathways are overrepresented. Finally, I present an initial secretome analysis of repopulating clones compared to non-repopulating clones and identified HDGF as a potentially novel self-renewal regulator, and PDAP1 as a potential negative self-renewal regulator.

6.3 Implications and future directions

6.3.1 Identification of novel iNKT1 and putative monocytes marked by *Fgd5* expression

Amongst the new and exciting HSC reporter mouse strains that are being developed, we chose to investigate the *Fgd5*^{ZsGreen*ZsGreen/+} reporter mouse because, amongst other reasons, the transplantation data validating its robustness was especially promising¹⁸⁸. *Fgd5* expression was also validated in scRNA-seq datasets to be highly enriched in the LT-HSC compartment¹⁹⁰. Therefore, it is surprising to find that *Fgd5* expression also marks a population non-stem cells (FE- cells) that was non-existent in the data originally published by Gazit et al¹⁸⁸. In our data, these FE- cells contain at least two fractions, one of which is an iNKT1 population and the other is most probably a CD244⁺ monocyte population as suggested by its gene expression profile. A potential explanation was that these cells arise from different BM preparation methods, however, in my data, the cells were present in both crushed and flushed bones. As iNKT1 cells are immune cells, differences in mouse facilities could have led to the differences in immune composition of the mice and could explain the appearance of these cells. To test this, it would be interesting to look for the presence of these cells in multiple mouse facilities and also compare the frequencies of the population, especially in lymphoid tissues such as the thymus, liver and lungs where they are slightly more abundant. Indeed, in the original paper by Gazit et al., only the BM was surveyed. Furthermore, it would also be worthwhile to look for these iNKT1 cells in other reporter mice

generated for the *Fgd5* gene, such as the *Fgd5^{mCherry/+}* reporter mouse, to confirm that this is not a phenomenon exclusively associated with the *Fgd5^{ZsGreen*ZsGreen/+}* reporter mouse.

As only about 20% of phenotypic iNKT1 cells are *Fgd5⁺*, the biggest outstanding question is whether *Fgd5⁺* iNKT1 cells are functionally distinct from *Fgd5⁻* iNKT1 cells and represent a unique novel immune population. Furthermore, does *Fgd5* have an important function for these iNKT1 cells? The next steps would be to perform more detailed intracellular flow experiments on the two fractions of iNKT1 cells to determine the cytokine profiles of each. It would also be interesting to challenge the *Fgd5^{ZsGreen*ZsGreen/+}* mice with certain pathogens to see how the subsets of iNKT1 cells respond. Similarly, as these iNKT1 cells were found in peripheral lymphoid tissues such as the spleen, lungs, thymus and liver, it would be interesting to see if the cells from various tissues are functionally different compared to each other. Moving forward, knock-out studies could offer potential insights into the function of *Fgd5* in iNKT1 cells. Even if *Fgd5⁺* iNKT1 cells are not functionally different, having a reporter for iNKT1 cells is potentially exciting for future *in vivo* imaging studies in the NKT cell field.

The FE-CD244⁺ population was transcriptionally most similar to monocytes. However, when compared to distinct monocyte transcriptomes, there doesn't seem to be a strong match with distinct subsets. Additionally, not much else is known about them in terms of the surface markers they express, thus it is entirely plausible that this population itself is not a homogeneous population, or alternatively they don't belong to any of the monocyte subsets that they were tested against. Further characterisation would be needed to narrow down the precise nature of these cells, and the first place to start would be to validate the list of surface marker genes that are highly expressed.

The discovery of these *Fgd5⁺* immune cells have broader implications in HSC biology. Namely that groups should be careful when using these reporter mice in niche imaging studies^{140,456}, and be aware that immune cells could potentially contaminate putative HSC populations, especially if the mice are housed in different mouse facilities. As HSC expansion cultures would very unlikely support NKT cells, it is not a strong concern that some of the *Fgd5⁺* cells within these cultures are contaminated. However, this would be a good reason for all experiments to be co-stained for EPCR.

6.3.2 *Fgd5* and EPCR: novel markers for HSCs *in vitro*

Our results show that *Fgd5* and EPCR together are useful for marking HSCs *in vivo* and *in vitro*. Interestingly, EPCR has been recently shown to mark expanded human HSCs *in vitro*²⁸⁵ and our

results accord with this finding in mouse HSCs. The data collected so far do not clarify whether *Fgd5* or EPCR is a better single marker. Since their expression is highly correlated, most of the *Fgd5*^{high} cells would also be EPCR^{high}. This means that for the lack of *Fgd5* reporter mice, EPCR can do a good job alone. However, having *Fgd5* would make the gating a bit easier. To formally assess the markers separately, additional transplantations with more stringent gating and higher cell numbers could be performed to quantify the frequency and proportion of HSCs in each fraction. In the future, LDAs would also be required to quantify the exact purity of re-sorted cells.

Overall, the addition of two reliable *in vitro* markers would facilitate the investigation of several ambitions such as: screening HSC self-renewal regulators and finding the molecular signature of expanding HSCs, both of which were attempted in this thesis. Whilst I applied the reporter system to screening recombinant factors, there is no reason why genetic screens such as CRISPR screens could not be used with this system. Additionally, this reporter system can potentially act as a quality control tool to assess the health of long-term mouse cultures.

In the future, it would definitely be worthwhile to test other reporter mice in the same way that we have for the *Fgd5*^{ZsGreen*ZsGreen/+} reporter mouse. In particular, the data for the recently published *Mds1*^{GFP/+}*Flt3*^{Cre} mice mouse (see section 1.2.6) looks especially promising as well. The authors who made this reporter mouse never validated whether the reporter works for *ex vivo* cultured HSCs. However, from a conceptual level, having a *Cre*-mediated reporter-quenching mechanism could be applied to other gene reporters such as *Fgd5* to facilitate more selective marking.

6.3.3 Proteomic analysis of conditioned media provides insights to expansion regulators

Co-culture with stromal cells has been a common strategy to support and maintain HSCs *ex vivo*. In particular, EL08 cells, isolated from the embryonic liver has been of particular interest due to its demonstrated ability to expand HSCs without cell-to-cell contact¹⁵³. Consistent with conditioning lengths used by previous studies on UG26 cell lines¹⁵³, clones cultured in CM_{3D} had significantly higher percentages of phenotypic HSCs (marked by the novel reporter strategy) than all other CM or non-conditioned control. In my data, I found that extended conditioning in general was detrimental to HSC support. Short conditioning time was also associated with larger clones. Whether the increased clone size is due to faster proliferation or better survival of resulting progenitors is unclear. However, considering that EL08 cells were derived from the embryonic liver, where HSCs are cycling constantly, this finding was not unexpected. As the conditioned media was filtered through 0.2µm filters, the effects observed can be fairly certainly attributed to soluble factors within the CM. However, to rule out other mechanisms, a possible experiment to

perform would be to heat inactivate the CM in order denature any potential self-renewal factors within.

Previous studies on supportive stromal cells have used qPCR and microarrays to identify potential self-renewal regulators³³⁵. Proteomic methods have also been used to profile EL08 cells⁴⁵⁷. However, these methods do not directly measure the proteins present in the supportive media. Thus, we decided to complement these studies with proteomic analysis of conditioned media. Indeed, many proteins in our proteomic dataset were identified from previous studies. Though importantly, novel proteins were also identified. The protocol could definitely be optimised. Because media contains such high concentrations of serum albumin, the high dynamic range in mass spectrometry data can hide peptides in lower concentrations⁴³². As a result, it is almost certain that many proteins of interest were not captured. The overall protein content also generally increased as conditioning period was extended, which increases the difficulty of selecting target factors that support HSCs. Nonetheless, the screen did identify several putative HSC self-renewal regulators, such as PTN, OPN, IGFBP4 and HPX. While PTN and IGFBP4 have been previously associated with EL08 cells before³³⁵, this is the first time OPN and HPX has been found to be expressed by EL08 cells, affirming the utility of having both proteomic and gene expression approaches.

The abundance of ECM proteins identified in the CM is particularly interesting as a growing number of studies are starting to look at the role of the ECM at regulating HSC self-renewal^{400,435,458,459}. In particular, collagen has been identified as an HSC supporting factor in a recent paper looking at UG26 supportive stromal cells¹⁵³. A few novel targets such as Nidogen-1 that have not been implicated in HSC biology previously were also discovered, prompting a need for future investigation.

In our initial validation experiments, we tested OPN, PTN, IGFBP2 and IGFBP4 and found that 100ng/mL of PTN was beneficial for HSC maintenance but not at higher concentrations. This is consistent with findings from Himburg et al. who initially discovered PTN as an HSC self-renewal regulator¹³², even though their cultures contained different concentrations of cytokines (20ng/mL TPO, 125ng/mL SCF and 50ng/mL Flt-3L) and the addition of serum. This study also used CD34-LSK as initiating cells, which, as the authors noted, have a purity of about 1 HSC in 39 cells. Further screening will be necessary to narrow down PTN concentrations that are optimal for HSC expansion. Furthermore, the result would ideally be validated by transplantation experiments.

While PTN alone could not account for all the benefits afforded by EL08 CM, additional factors in the proteomic screen could be tested for synergistic effects with PTN. In the future, these candidate factors can also be used on new expansion culture system utilising F12 and PVA. A limitation of this initial screen and of recombinant protein screens in general is whether they have the same functional activity as their physiological counterparts. Some research suggests that for many proteins, their function relies on specific post translational modifications such as glycosylation that may be lacking in certain recombinant protein expression systems⁴⁶⁰. Therefore, the lack of observed effect from tested factors may not necessarily mean the factors don't have an effect physiologically. Furthermore, it still begs the question of whether these effects observed are batch dependent. As mentioned in 1.4.9, in an interesting study by Wilkinson et al., cytokines exhibited varying effects when different batches of BSA was used⁴¹⁷. This concern is partially addressed by the recent development of all recombinant, BSA free culture conditions by Wilkinson et al²⁸³, which was why we chose to focus our investigation on it.

In the secretome analysis of media from 28-day F12 PVA cultures, we improved previous protocols by introducing TMT labelling of peptides, which allowed the increase in detection sensitivity by allowing samples to be run in tandem. Combined with the above EL08 CM screen, the initial run presented in this thesis is a proof-of-principle experiment that it would be possible to identify potential HSC regulators in the media that the cells are grown in. Unsurprisingly, very few of the targets identified in EL08 CM were found in this dataset, probably indicative of the vast differences in cell types used to condition the media. Of note, HPX was also identified, however it was present in media from both repopulating and non-repopulating clones. As this was the first run of media samples from only 6 clones, future runs would be needed to strengthen this dataset. As it currently stands, the two proteomic datasets generated here are not very comparable.

The majority of studies so far have suggested that the bulk of the signalling generated in *ex vivo* cultures are negative feedback signals^{283,416,417}. Conceptually this also makes sense, as unchecked expansion of HSCs *in vivo* would lead to unwanted leukaemia. However, it cannot be ruled out that some positive feedback signalling exists. Here we identify HDGF as a unique protein expressed only by clones that are able to repopulate recipient mice. Future work would be needed to validate HDGF as a positive regulator of HSC self-renewal, however, reports have demonstrated that it is highly expressed in FL⁴⁵¹, a site of HSC expansion. In non-repopulating clones, several unique proteins were identified including PDAP1. It would be interesting to apply the fed-batch culture

system developed by Csaszar et al.⁴¹⁶, to automate negative feedback control by specifically targeting potential negative regulators, such as PDAP1.

6.3.4 Validated and fully defined F12 and PVA-based expansion conditions

As mentioned, the HSC expansion field has overseen incremental improvements in culture strategies, from serum-based cultures to stromal co-cultures, which is slightly more defined³⁶¹; to serum-free cultures that can maintain HSCs up to 10 days¹⁰²; to all recombinant-based cultures containing HSA instead of the less defined BSA that maintains HSCs up to 28 days⁴¹⁷; ultimately now to the breakthrough, fully-defined PVA-based cultures that can fully expand HSCs for more than 28 days²⁸³.

From my data, F12-based conditions had much higher survival than standard SS cultures. To put this into context, SS-based conditions would only achieve similarly high survival if serum is added or the media is conditioned by supportive cells such as EL08 cells. Considering these F12-based cultures are serum free, just the survival rate alone is quite impressive. It would be interesting to understand the specific media formulation differences in F12 that allows HSCs to survive better. Higher SCF concentration, rather than TPO concentration, was associated with faster proliferation of clones. This also explains why F12-based cultures had lower proliferation rates compared to SS-based cultures, as the SCF concentration was 30-fold lower. Future scalable strategies of HSC expansion would need to balance the proliferation rates and HSC self-renewal, in order to achieve the optimal expansion rates. Just from looking at percentage FLSK, there doesn't seem to be a major difference between F12-based conditions and SS-based conditions in 10 days of culture. However, 28-day cultures indicate that F12-based cultures are better at maintaining stem cells. Alternatively, this would also be consistent with major expansion of HSCs in F12 cultures after 10 days of culture.

It would be interesting to assess HSC expansion at different timepoints throughout extended F12-based cultures, as the data so far suggests that expansion is minimal at earlier timepoints. This also raises the question of whether F12-based conditions have a direct or indirect effect on HSCs, by supporting HSC-supportive, but non-HSC cell types. However, data of both Wilkinson et al. and mine would suggest that F12-based culture conditions are not supportive of mature cells. Of note, ESLAM Sca-1⁺ cells were used in this thesis to initiate long-term cultures, which is slightly more enriched than the CD34⁺CD150⁺LSK cells used by Wilkinson et al²⁸³. This would support the idea that expansion effects are cell autonomous and not due to non-stem cells in less pure gating strategies. However, an alternatively consistent possibility is that SS-based cultures support pro-

differentiating cells that ultimately cause HSCs to fail beyond 10 days, when such cells are generated.

One possible experiment to directly compare the propensity of self-renewal divisions between these culture conditions is a paired daughter transplantation experiment. As mentioned in section 1.2.5, this works by allowing single HSCs to divide *in vitro* and then transplanting the two daughters into two separate recipients.

Our data suggests that IL-11 has a beneficial effect on HSCs in HSA based cultures, but not in PVA-based cultures. It would be interesting to see if PVA has a functional role, and signals through pathways that are redundant with IL-11. Currently, PVA is thought to serve a carrier protein function, however, as this has been extremely recent, there has been very little knowledge of its mechanism of action. As PVA is completely chemically defined, it would be interesting to reevaluate the effects of cytokines in a systematic way. However, as mentioned, recombinant proteins and cytokines could still have batch-dependent differences which would cloud any sort of systematic analysis. Future development of chemically synthesised cytokine analogues, such as the TPO analogue, butyzamide, may be useful to reduce batch effects of recombinant cytokines⁴⁶¹. Furthermore, it would be interesting to test the PTN and other factors identified in the EL08 secretome screen, using F12 with PVA as the base media. In order for this to be translated into clinical expansion of HSCs, the recombinant factors would need to be produced by methods compatible with animal free and good manufacturing practice.

6.3.5 Molecular characterisation of *ex vivo* expanding HSCs

An interesting observation from Wilkinson et al. was the clonal heterogeneity observed in F12 cultures²⁸³. Single cells cultured in F12 PVA-based expansion cultures seemingly have widely different outcomes, from different clone sizes to different lineage outputs upon transplantation. However, they didn't attempt to characterise the expanding HSCs molecularly. Understanding the molecular basis for this clonal heterogeneity can potentially help efforts to increase expansion efficiency. Can HSCs that fail to expand be nudged towards self-renewal and can expanding HSCs expand even more? In order to answer these questions, we need to first understand the molecular differences that separate expanding and non-repopulating HSCs. Therefore, in Chapter 5, I presented the molecular profile of *ex vivo* expanded single HSCs. Our data is consistent with data from Wilkinson et al., demonstrating that single HSCs can expand in F12 PVA based conditions, with extensive clonal heterogeneity in 28-day cultures.

In our 28-day single cell cultures, the percentage of phenotypic HSCs, marked by our novel reporter strategy, correlated extremely highly with repopulation when assessed retrospectively. To consolidate this finding, it would be important to prospectively predict the repopulation ability of clones based on the ELSK percentage. Interestingly, total cell numbers did not correlate significantly with repopulation, further affirming the negative relationship between HSC self-renewal and proliferation. As mentioned previously, LDA experiments would be needed in order to fully quantify the purity of functional HSCs within phenotypic HSCs.

In the comparisons of the transcriptomes of clones with ranging functional outcomes, it was clear that repopulating clones have more similar transcriptomes to each other than non-repopulating clones. Consistent with findings by Wilkinson et al. that differentiated cells are less supported in F12 PVA cultures²⁸³, we found that even in non-phenotypic HSCs, their transcriptomes resembled ST-HSCs or MPPs rather than mature cells. Interestingly, based on visual assessment, F12-based cultures do seem to induce some level Megakaryocytic differentiation, which was also noted by Wilkinson et al⁴¹⁹. However, there is not a strong megakaryocytic signature in the transcriptomes of non-ELSK cells, possibly due to the low percentage of total cells they represent.

By looking at the genes that drive transcriptomic differences, we generated a signature gene list for repopulating ex vivo expanded HSCs. Interestingly, although the MoIO genes¹⁸⁶, which are signature genes for quiescent HSCs *in vivo*, are upregulated in ex vivo expanded HSCs, they were not as strongly upregulated as these new signature genes, which support the idea that although *ex vivo* expanding HSCs largely resemble freshly isolated BM HSCs, they exhibit distinct molecular states. Reassuringly, some of the signature genes have already been implicated in HSC self-renewal and unsurprisingly includes *Procr*, which was used to separate repopulating and non-repopulating cells. In particular, one of the genes, *Prdm16* was shown to expand HSCs in overexpression studies³⁶⁰.

Moving forward, extensive functional validation will be required to test signature genes by qPCR on new cultures to see if they can predict clonal repopulation. A robust gene signature for expanding HSCs can be extremely useful - One can imagine performing a simple qPCR quality assessment of the overall HSC content within cultures. Our gene signature also identified several cell surface marker genes, such as *Robo1*, *Esam* and *Ncam2*, which should also be tested by flow cytometry on new cell cultures. Of note, ESAM has already been validated as an HSC marker *in*

*vivo*¹⁷², but not *in vitro*. Importantly, if these markers can further subfractionate repopulating cells within the phenotypic HSCs, higher purities of expanding HSCs can be potentially enriched.

Interestingly, another set of genes was identified through differential gene expression analysis of repopulating phenotypic HSCs vs. non-repopulating phenotypic HSCs. Most notable within this list are *Vnf* and *Pld3*, both of which have been previously found to be upregulated in highly enriched HSCs¹⁸⁴.

Of the pathways upregulated in expanding HSCs, Rho GTPase and MapK associated pathways seem to play an important role. Future experiments would be important to narrow down the specific key players in the signalling pathways, however there is plenty of evidence suggesting the importance of Rho GTPases in the haematopoietic system^{254,439,440}. As FGD5 is a RhoGEF previously linked to the activation of CDC42 via the VEGF pathway^{192,462,463}, it is particularly interesting to speculate whether it plays a functional role in HSC self-renewal. If indeed it does, then it would strengthen the case for *Fgd5* as a marker for HSCs. However, it would also demand more rigorous validations of whether the heterozygous reporter knock in has adverse effects on HSC function. Alongside FGD5, in the same study that identified *Vnf* and *Pld3* as HSC associated genes, *RhoB*, which is a Rho GTPase, was also identified¹⁸⁴. As Rho GTPases are known to moderate signalling from the ECM, it is interesting to speculate whether the upregulated pathway is associated with the use of fibronectin coated plates. Indeed, different integrins, which have been shown to interact with fibronectin are upregulated in repopulating and non-repopulating cells.

There are of course limitations to our strategy. For a start, an obvious limitation is that although our reporter strategy greatly enriches for functional HSCs, they are by no means 100% pure. This is demonstrated by the fact that certain clones with small numbers of phenotypic HSCs were not able to repopulate mice. However, even with this increased noise in the transcriptomic data, the pattern for repopulation is still very clear.

Our strategy also underappreciates any potential cellular heterogeneity in the non-HSC populations, by binning many cell phenotypes into a single group. Whilst we chose bulk sequencing, because of the increased depth of sequencing as well as the ability to pair the samples to their functional outcomes through transplantation, it would be interesting to supplement this dataset with scRNA-seq using 10X genomics where we could gain a more complete picture of the individual cell identities comprising the non-HSC fraction.

6.3.6 Unifying molecular signature of HSCs or distinct molecular signature of HSC subtypes

Despite our knowledge of HSC self-renewal regulators and the advancement in scRNA-seq technologies, the field has still been unable to find a common molecular programme for HSCs. A 2015 study by Wilson et al. was the best attempt at this question so far¹⁸⁶. By profiling HSCs isolated from multiple isolating strategies, the authors were able to find the MoIO signature genes mentioned above. However, even within the highly enriched populations of phenotypic HSCs, there is heterogeneity in HSC subtypes, with differing lineage output and self-renewal ability^{154,206}, further obfuscating a common “molecular programme” for HSCs. At present, it is impossible to prospectively isolate HSC subtypes by surface markers and importantly, it remains uncertain whether the different HSC gating strategies enrich for certain subtypes of HSCs⁴⁶⁴.

The ability to expand HSCs has enabled several new opportunities to tackle this goal. First of all, the above molecular study only looked at freshly isolated BM HSCs, which does not sufficiently encompass all the diverse molecular states of HSCs. Previous studies have compared murine FL-HSCs with adult BM HSCs⁴⁶⁵, however, none so far have used whole transcriptome RNA-seq. Very recently, Popescu et al. published a 10X dataset for human FL cells. More of such approaches will be necessary in the future in order to decode common molecular programmes. Most interestingly, with expanding HSCs now possible and the new discovery of *in vitro* cultures capable of maintaining HSCs in quiescence⁴¹⁸, we can now interrogate the common molecular states between 4 different HSC populations.

In the future, this could be combined with mathematical modelling methods to understand regulatory mechanisms for HSC self-renewal fate choice. Although Wilkinson et al. did not attempt this in his seminal paper, the transcriptomic profile of expanded HSCs presented in this thesis may be the first step towards this ambition. Previous attempts of modelling HSC self-renewal have primarily focused on *in vivo* divisional kinetics and population dynamics of HSCs and progenitors^{31,466,467}; however, it would be interesting to explore similar avenues with *in vitro* expanded HSCs. To this end, it would be useful to generate more HSC profiles under different culturing conditions for comparisons, for example with different cytokine concentrations such as high SCF or low TPO.

Apart from self-renewal, another outstanding question is whether the intrinsic lineage biases of single HSCs are maintained within long-term *ex vivo* culture, as they are commonly in serial transplantation²⁰⁶. If so, we can potentially begin to unravel the molecular differences between

myeloid-biased HSCs and lymphoid-biased HSCs by linking functional outcomes of clones to other molecular assays such as ATAC-seq, Bisulphite sequencing, CHIP-seq and proteomics, in order to analyse their epigenome and proteome for associated molecular differences. As mentioned in section 1.3.2, this area may hold crucial information that is missed by the differentiation landscapes generated by scRNA-seq.

6.3.7 Future of *ex vivo* HSC expansion: translation into human HSCs

Ultimately, this is an exciting new era for HSC expansion. For the first time, murine HSCs can be expanded *ex vivo* robustly and durably. Ultimately this new culture system will only serve as the base in which improvements are made. By adding combinations of self-renewal regulators, or integrating fed-batch negative feedback regulation, or even engineering artificial 3D niches with ECM proteins and functionalised hydrogels⁴⁶⁸. In the future, the ultimate goal would be to translate these findings into clinically expanding human cells. Importantly, in the signature genes identified in this thesis, at least one of them, *Hlf*, has been reported in a pre-print journal to mark expanded human HSCs in culture⁴⁴⁸. As mentioned, EPCR has also already been shown to mark human HSCs *in vitro*²⁸⁵, suggesting that many of the identified genes have parallel functions in human HSCs. Wilkinson et al. had already demonstrated that human HSCs were able to grow in F12 PVA-based cultures. Already fed-batch systems are being applied with novel small molecules such as UM171 and SR1³⁸⁶ to achieve modest levels of expansion. Combining such promising avenues will undoubtedly lead to success in future clinical scale human HSC expansion.

6.4 Concluding remarks

In this PhD thesis, I showed that the *Fgd5^{ZsGreen*ZsGreen/+}* reporter mouse is not only a marker for murine HSCs, but also for a subset of iNKT1 cells. When combined with EPCR, the reporter can mark HSCs both *in vivo* and *in vitro*. This helped identify putative HSC supporting factors in stromal conditioned media and validated current state-of-the-art expansion protocols. Finally, the *in vitro* reporter system was used to identify cultures rich in HSCs and provided the molecular characterisation of expanding HSCs by isolating HSC-enriched fractions. By comparing HSCs and non-HSCs, I present a molecular gene signature that can be used to detect HSC content in *ex vivo* cultures. These findings may have important implications for improving *ex vivo* cultures of HSCs, first by providing a novel set of *in vitro* markers for HSCs and secondly by providing an increased understanding of the molecular pathways and programmes that regulate *ex vivo* expanding HSCs. When translated into human HSCs, the limitless availability of HSCs will benefit clinical applications such as stem cell therapies and protocols to produce mature blood cells *ex vivo*.

7 References

1. Schwartz, R. S. & Conley, C. L. Blood. *Britannica* (2019).
2. Kent, D. G. & Eaves, C. J. Adult Hematopoiesis. *Encyclopedia of Immunobiology* vol. 1 15–25 (2016).
3. Bryder, D., Rossi, D. J. & Weissman, I. L. Hematopoietic Stem Cells The Paradigmatic Tissue-Specific Stem Cell. *Curr. Protoc. Stem Cell Biol.* **169**, 211–215 (2006).
4. Aiuti, A., Scala, S. & Chabannon, C. Biological properties of HSC: Scientific basis for HSCT. in *The EBMT Handbook: Hematopoietic Stem Cell Transplantation and Cellular Therapies* 49–56 (Springer International Publishing, 2018).
5. Eaves, C. J. Hematopoietic stem cells: concepts, definitions, and the new reality. *Blood* (2015).
6. Wilson, A. *et al.* Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair. *Cell* **135**, 1118–1129 (2008).
7. Walter, D. *et al.* Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. *Nature* **520**, 549–552 (2015).
8. Laurenti, E. & Göttgens, B. From haematopoietic stem cells to complex differentiation landscapes. *Nature* **553**, 418–426 (2018).
9. Luo, Y. *et al.* M1 and M2 macrophages differentially regulate hematopoietic stem cell self-renewal and ex vivo expansion. *Blood Adv.* **2**, 859–870 (2018).
10. da Silva, E. Z. M., Jamur, M. C. & Oliver, C. Mast Cell Function: A New Vision of an Old Cell. *J. Histochem. Cytochem.* **62**, 698–738 (2014).
11. Orkin, S. H. & Zon, L. I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631–644 (2008).
12. Spits, H. & Cupedo, T. Innate Lymphoid Cells: Emerging Insights in Development, Lineage Relationships, and Function. *Annu. Rev. Immunol.* **30**, 647–675 (2012).
13. Jerud, E. S., Bricard, G. & Porcelli, S. A. CD1d-Restricted Natural Killer T Cells: Roles in Tumor Immunosurveillance and Tolerance. *Transfus. Med. Hemotherapy* **33**, 18–36 (2006).
14. Godfrey, D. I., MacDonald, H. R., Kronenberg, M., Smyth, M. J. & Van Kaer, L. NKT cells: What's in a name? *Nat. Rev. Immunol.* **4**, 231–237 (2004).
15. Haas, S. *et al.* Inflammation-Induced Emergency Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors. *Cell Stem Cell* **17**, 422–434 (2015).
16. Sykes, S. M., Kokkalis, K. D., Milsom, M. D., Levine, R. L. & Majeti, R. Clonal evolution of preleukemic hematopoietic stem cells in acute myeloid leukemia. *Exp. Hematol.* **43**, 989–992 (2015).
17. Batsivari, A. *et al.* Dynamic responses of the haematopoietic stem cell niche to diverse stresses. *Nat. Cell Biol.* **22**, 7–17 (2020).
18. Osawa, M., Hanada, K., Hamada, H. & Nakauchi, H. Long-Term Lymphohematopoietic Reconstitution by a Single CD34-Low/Negative Hematopoietic Stem Cell. *Science (80-.)*. **273**, 242 LP – 245 (1996).
19. Wilkinson, A. C., Igarashi, K. J. & Nakauchi, H. Haematopoietic stem cell self-renewal in vivo and ex vivo. *Nat. Rev. Genet.* **21**, 541–554 (2020).
20. Ho, T. T. *et al.* Autophagy maintains the metabolism and function of young and old stem cells. *Nature* **543**, 205–210 (2017).
21. Xie, S. Z. *et al.* Sphingolipid Modulation Activates Proteostasis Programs to Govern Human Hematopoietic Stem Cell Self-Renewal. *Stem Cell* **25**, 639-653.e7 (2019).
22. Ito, K. & Suda, T. Metabolic requirements for the maintenance of self-renewing stem cells. *Nat. Rev. Mol. Cell Biol.* **15**, 243–56 (2014).
23. Vannini, N. *et al.* Specification of haematopoietic stem cell fate via modulation of mitochondrial activity. *Nat. Commun.* **7**, (2016).

24. Foudi, A. *et al.* Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells. *Nat. Biotechnol.* **27**, 84–90 (2009).
25. Bernitz, J. M., Kim, H. S., MacArthur, B., Sieburg, H. & Moore, K. Hematopoietic Stem Cells Count and Remember Self-Renewal Divisions. *Cell* **167**, 1296–1309.e10 (2016).
26. Takizawa, H., Regoes, R. R., Boddupalli, C. S., Bonhoeffer, S. & Manz, M. G. Dynamic variation in cycling of hematopoietic stem cells in steady state and inflammation. *J. Exp. Med.* **208**, 273–284 (2011).
27. Qiu, J., Papatsenko, D., Niu, X., Schaniel, C. & Moore, K. Divisional history and hematopoietic stem cell function during homeostasis. *Stem Cell Reports* **2**, 473–490 (2014).
28. Laurenti, E. *et al.* CDK6 levels regulate quiescence exit in human hematopoietic stem cells. *Cell Stem Cell* **16**, (2015).
29. Müller-Sieburg, C. E., Cho, R. H., Thoman, M., Adkins, B. & Sieburg, H. B. Deterministic regulation of hematopoietic stem cell self-renewal and differentiation. *Blood* **100**, 1302–1309 (2002).
30. Takizawa, H. & Manz, M. G. In vivo divisional tracking of hematopoietic stem cells. *Ann. N. Y. Acad. Sci.* **1266**, 40–46 (2012).
31. Busch, K. *et al.* Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature* **518**, 542–546 (2015).
32. Morcos, M. N. F. *et al.* Continuous mitotic activity of primitive hematopoietic stem cells in adult mice. *J. Exp. Med.* **217**, (2020).
33. Khajuria, R. K. *et al.* Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopoiesis. *Cell* **173**, 90–103.e19 (2018).
34. Ito, K. *et al.* Reactive oxygen species act through p38 MAPK to limit the lifespan of hematopoietic stem cells. *Nat. Med.* **12**, 446–451 (2006).
35. Jacobson, L. O., Simmons, E. L., Marks, E. K. & Eldredge, J. H. Recovery from radiation injury. *Science* (80-). **113**, 510–511 (1951).
36. Till, J. E. & McCulloch, E. A. A Direct Measurement of the Radiation Sensitivity of Normal Mouse Bone Marrow Cells A Direct Measurement of the Radiation Sensitivity of Normal Mouse Bone Marrow Cells'. *Source Radiat. Res. Radiat. Res.* **14**, 213–222 (1961).
37. Becker, A. J., McCulloch E. A. & Till J. E. Cytological Demonstration of the Clonal Nature of Spleen Colonies Derived from Transplanted Mouse Marrow Cells. *Nature* (1963).
38. Siminovitch, L., McCulloch, E. A. & Till, J. E. The Distribution of Colony-forming Cells Among Spleen Colonies. *J. Cell. Comp. Physiol.* **62**, (1963).
39. BECKER, A. J., MCCULLOCH, E. A., SIMINOVITCH, L. & TILL, J. E. THE EFFECT OF DIFFERING DEMANDS FOR BLOOD CELL PRODUCTION ON DNA SYNTHESIS BY HEMOPOIETIC COLONY-FORMING CELLS OF MICE. *Blood* **26**, 296–308 (1965).
40. Ploemacher, R. E. & Brons, R. H. Separation of CFU-S from primitive cells responsible for reconstitution of the bone marrow hemopoietic stem cell compartment following irradiation: evidence for a pre-CFU-S cell. *Exp. Hematol.* **17**, 263–6 (1989).
41. Bowie, M. B. *et al.* Identification of a new intrinsically timed developmental checkpoint that reprograms key hematopoietic stem cell properties. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 5878–5882 (2007).
42. Ana Cumano, Françoise Dieterian-Lievre & Isabelle Godin. Lymphoid Potential, Probed before Circulation in Mouse, Is Restricted to Caudal Intraembryonic Splanchnopleura. *Cell* **86**, 907–916 (1996).
43. Gekas, C., Dieterlen-Lièvre, F., Orkin, S. H. & Mikkola, H. K. A. The Placenta Is a Niche for Hematopoietic Stem Cells. *Dev. Cell* **8**, 365–375 (2005).
44. Dzierzak, E. Hematopoietic stem cells and their precursors: developmental diversity and lineage relationships. *Immunol. Rev.* **187**, 126–138 (2002).
45. Buckley, S. M. *et al.* Maintenance of HSC by Wnt5a secreting AGM-derived stromal cell line. *Exp Hematol* **39**, 114–123 (2011).
46. Rowe, R. G., Mandelbaum, J., Zon, L. I. & Daley, G. Q. Engineering Hematopoietic Stem Cells: Lessons

- from Development. *Cell Stem Cell* (2016).
47. Ema, H. & Nakauchi, H. Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. *Blood* **95**, 2284–2288 (2000).
 48. Rhodes, K. E. *et al.* The Emergence of Hematopoietic Stem Cells Is Initiated in the Placental Vasculature in the Absence of Circulation. *Cell Stem Cell* **2**, 252–263 (2008).
 49. Ottersbach, K. & Dzierzak, E. The murine placenta contains hematopoietic stem cells within the vascular labyrinth region. *Dev. Cell* **8**, 377–387 (2005).
 50. Bowie, M. B. *et al.* Hematopoietic stem cells proliferate until after birth and show a reversible phase-specific engraftment defect. *J. Clin. Invest.* **116**, 2808–2816 (2006).
 51. Morrison, S. J., Hemmati, H. D., Wandycz, A. M. & Weissman, I. L. The purification and characterization of fetal liver hematopoietic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 10302–10306 (1995).
 52. Cheshier, S. H. *et al.* In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. *Proc. Natl. Acad. Sci.* **96**, 3120–3125 (1999).
 53. Pawliuk, R., Eaves, C. & Humphries, R. K. Evidence of Both Ontogeny and Transplant Dose-Regulated Expansion of Hematopoietic Stem Cells In Vivo. *Blood* (1996).
 54. Micklem, H. S., Ford, C. E., Evans, E. P., Ogden, D. A. & Papworth, D. S. Competitive in vivo proliferation of foetal and adult haematopoietic cells in lethally irradiated mice. *J. Cell. Physiol.* **79**, 293–298 (1972).
 55. Kim, I., Saunders, T. L. & Morrison, S. J. Sox17 Dependence Distinguishes the Transcriptional Regulation of Fetal from Adult Hematopoietic Stem Cells. *Cell* **130**, 470–483 (2007).
 56. Mochizuki-Kashio, M. *et al.* Dependency on the polycomb gene Ezh2 distinguishes fetal from adult hematopoietic stem cells. *Blood* **118**, 6553–6561 (2011).
 57. Park, I. K. *et al.* Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* **423**, 302–305 (2003).
 58. Hock, H. *et al.* Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. *Nature* **431**, 1002–1007 (2004).
 59. Wilson, N. K. *et al.* Gfi1 expression is controlled by five distinct regulatory regions spread over 100 kilobases, with Scf/Tal1, Gata2, PU.1, Erg, Meis1, and Runx1 acting as upstream regulators in early hematopoietic cells. *Mol. Cell. Biol.* **30**, 3853–3863 (2010).
 60. Ye, M. *et al.* C/EBP α controls acquisition and maintenance of adult haematopoietic stem cell quiescence. *Nat. Cell Biol.* **15**, 385–394 (2013).
 61. Copley, M. R. *et al.* The Lin28b--let-7--Hmga2 axis determines the higher self-renewal potential of fetal haematopoietic stem cells. *Nat. Cell Biol.* **15**, (2013).
 62. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194 (2013).
 63. Konieczny, J. & Arranz, L. Updates on old and weary haematopoiesis. *Int. J. Mol. Sci.* **19**, (2018).
 64. Dykstra, B., Olthof, S., Schreuder, J., Ritsema, M. & Haan, G. De. Clonal analysis reveals multiple functional defects of aged murine hematopoietic stem cells. *J. Exp. Med.* **208**, 2691–2703 (2011).
 65. Beerman, I. *et al.* Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 5465–5470 (2010).
 66. Verovskaya, E. *et al.* Heterogeneity of young and aged murine hematopoietic stem cells revealed by quantitative clonal analysis using cellular barcoding. *Blood* **122**, 523–532 (2013).
 67. Kuranda, K. *et al.* Age-related changes in human hematopoietic stem/progenitor cells. *Aging Cell* **10**, 542–546 (2011).
 68. Rundberg Nilsson, A., Soneji, S., Adolfsson, S., Bryder, D. & Pronk, C. J. Human and Murine Hematopoietic Stem Cell Aging Is Associated with Functional Impairments and Intrinsic Megakaryocytic/Erythroid Bias. *PLoS One* **11**, e0158369 (2016).
 69. Yamamoto, R. *et al.* Large-Scale Clonal Analysis Resolves Aging of the Mouse Hematopoietic Stem Cell

- Compartment. *Cell Stem Cell* **22**, 600–607.e4 (2018).
70. Pang, W. W. *et al.* Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *PNAS* (2011).
 71. Sudo, K., Ema, H., Morita, Y. & Nakauchi, H. Age-associated characteristics of murine hematopoietic stem cells. *J. Exp. Med.* **192**, 1273–1280 (2000).
 72. Cho, R. H., Sieburg, H. B. & Muller-Sieburg, C. E. A new mechanism for the aging of hematopoietic stem cells: Aging changes the clonal composition of the stem cell compartment but not individual stem cells. *Blood* **111**, 5553–5561 (2008).
 73. Geiger, H., de Haan, G. & Carolina Florian, M. The ageing haematopoietic stem cell compartment. *Nat. Publ. Gr.* **13**, (2013).
 74. Rossi, D. J. *et al.* Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature* **447**, 725–729 (2007).
 75. Flach, J. *et al.* Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature* **512**, 198–202 (2014).
 76. Morrison, S. J., Prowse, K. R., Ho, P. & Weissman, I. L. Telomerase activity in hematopoietic cells is associated with self-renewal potential. *Immunity* **5**, 207–216 (1996).
 77. Vaziri, H. *et al.* Evidence for a mitotic clock in human hematopoietic stem cells: Loss of telomeric DNA with age. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 9857–9860 (1994).
 78. Allsopp, R. C., Cheshier, S. & Weissman, I. L. Telomere shortening accompanies increased cell cycle activity during serial transplantation of hematopoietic stem cells. *J. Exp. Med.* **193**, 917–924 (2001).
 79. Rudolph, K. L. *et al.* Longevity, stress response, and cancer in aging telomerase-deficient mice. *Cell* **96**, 701–712 (1999).
 80. Allsopp, R. C., Morin, G. B., DePinho, R., Harley, C. B. & Weissman, I. L. Telomerase is required to slow telomere shortening and extend replicative lifespan of HSCs during serial transplantation. *Blood* **102**, 517–20 (2003).
 81. Allsopp, R. C. *et al.* Effect of TERT over-expression on the long-term transplantation capacity of hematopoietic stem cells [2]. *Nature Medicine* vol. 9 369–371 (2003).
 82. Weinstein, B. S. & Ciszek, D. The reserve-capacity hypothesis: Evolutionary origins and modern implications of the trade-off between tumor-suppression and tissue-repair. *Exp. Gerontol.* **37**, 615–627 (2002).
 83. Hemann, M. T. & Greider, C. W. Wild-derived inbred mouse strains have short telomeres. *Nucleic Acids Res.* **28**, 4474–4478 (2000).
 84. Mohrin, M. *et al.* A mitochondrial UPR-mediated metabolic checkpoint regulates hematopoietic stem cell aging. *Science (80-.)*. **347**, 1374–1377 (2015).
 85. Norddahl, G. L. *et al.* Accumulating mitochondrial DNA mutations drive premature hematopoietic aging phenotypes distinct from physiological stem cell aging. *Cell Stem Cell* **8**, 499–510 (2011).
 86. Zink, F. *et al.* Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. *Blood* **130**, 742–752 (2017).
 87. Xie, M. *et al.* Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* **20**, 1472–1478 (2014).
 88. Jaiswal, S. *et al.* Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *N. Engl. J. Med.* **371**, 2488–2498 (2014).
 89. Genovese, G. *et al.* Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).
 90. Van Den Akker, E. B. *et al.* Uncompromised 10-year survival of oldest old carrying somatic mutations in DNMT3A and TET2. *Blood* **127**, 1512–1515 (2016).
 91. Saçma, M. *et al.* Haematopoietic stem cells in perisinusoidal niches are protected from ageing. *Nat. Cell Biol.* **21**, 1309–1320 (2019).

92. Ho, Y. H. *et al.* Remodeling of Bone Marrow Hematopoietic Stem Cell Niches Promotes Myeloid Cell Expansion during Premature or Physiological Aging. *Cell Stem Cell* **25**, 407–418.e6 (2019).
93. Zambetti, N. A. *et al.* Mesenchymal Inflammation Drives Genotoxic Stress in Hematopoietic Stem Cells and Predicts Disease Evolution in Human Pre-leukemia. *Cell Stem Cell* **19**, 613–627 (2016).
94. Kovtonyuk, L. V., Fritsch, K., Feng, X., Manz, M. G. & Takizawa, H. Inflamm-aging of hematopoiesis, hematopoietic stem cells, and the bone marrow microenvironment. *Front. Immunol.* **7**, 502 (2016).
95. Schofield, R. The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood Cells* **4**, 7–25 (1978).
96. Xie, T. & Li, L. Stem cells and their niche: an inseparable relationship. *Development* **134**, 2001–2006 (2001).
97. Xie, T. & Spradling, A. C. A niche maintaining germ line stem cells in the *Drosophila* ovary. *Science (80-.)*. **290**, 328–330 (2000).
98. Lin, H. The stem-cell niche theory: lessons from flies. *Nat. Rev. Genet.* **3**, 931–940 (2002).
99. Song, X., Zhu, C. H., Doan, C. & Xie, T. Germline stem cells anchored by adherens junctions in the *Drosophila* ovary niches. *Science (80-.)*. **296**, 1855–1857 (2002).
100. Van De Bor, V. *et al.* Companion Blood Cells Control Ovarian Stem Cell Niche Microenvironment and Homeostasis. *Cell Rep.* **13**, 546–560 (2015).
101. Ashe, H. L. Type IV collagens and Dpp positive and negative regulators of signaling. *Fly (Austin)*. **2**, 313–315 (2008).
102. Kent, D. G., Dykstra, B. J., Cheyne, J., Ma, E. & Eaves, C. J. E. Steel factor coordinately regulates the molecular signature and biologic function of hematopoietic stem cells. *Blood* **112**, (2008).
103. Morrison, S. J. & Scadden, D. T. The bone marrow niche for haematopoietic stem cells. *Nature* **505**, 327–34 (2014).
104. Boulais, P. E. & Frenette, P. S. Making sense of hematopoietic stem cell niches. *Blood* **125**, 2621–2629 (2015).
105. Pinho, S. & Frenette, P. S. Haematopoietic stem cell activity and interactions with the niche. *Nat. Rev. Mol. Cell Biol.* **20**, 303–320 (2019).
106. Asada, N., Takeishi, S. & Frenette, P. S. Complexity of bone marrow hematopoietic stem cell niche. *Int. J. Hematol.* **106**, 45–54 (2017).
107. Baryawno, N., Severe, N. & Scadden, D. T. Hematopoiesis: Reconciling Historic Controversies about the Niche. *Cell Stem Cell* **20**, 590–592 (2017).
108. Pinho, S. & Frenette, P. S. Haematopoietic stem cell activity and interactions with the niche. *Nature Reviews Molecular Cell Biology* vol. 20 303–320 (2019).
109. Pinho, S. *et al.* Lineage-Biased Hematopoietic Stem Cells Are Regulated by Distinct Niches. *Dev. Cell* **44**, 634–641.e4 (2018).
110. Lord, B. I., Testa, N. G. & Hendry, J. H. The relative spatial distributions of CFUs and CFUc in the normal mouse femur. *Blood* **46**, 65–72 (1975).
111. Lo Celso, C. *et al.* Live-animal tracking of individual haematopoietic stem/progenitor cells in their niche. *Nature* **457**, 92–96 (2009).
112. Nilsson, S. K., Johnston, H. M. & Coverdale, J. A. Spatial localization of transplanted hemopoietic stem cells: Inferences for the localization of stem cell niches. *Blood* **97**, 2293–2299 (2001).
113. Lévesque, J. P., Helwani, F. M. & Winkler, I. G. The endosteal osteoblastic niche and its role in hematopoietic stem cell homing and mobilization. *Leukemia* **24**, 1979–1992 (2010).
114. Arai, F. *et al.* Tie2/angiopoietin-1 signaling regulates hematopoietic stem cell quiescence in the bone marrow niche. *Cell* **118**, 149–161 (2004).
115. Nilsson, S. K. *et al.* Osteopontin, a key component of the hematopoietic stem cell niche and regulator of primitive hematopoietic progenitor cells. *Blood* **106**, 1232–9 (2005).
116. Stier, S. *et al.* Osteopontin is a hematopoietic stem cell niche component that negatively regulates stem cell pool size. *J. Exp. Med.* **201**, 1781–91 (2005).

117. Xie, Y. *et al.* Detection of functional haematopoietic stem cell niche using real-time imaging. *Nature* **457**, 97–101 (2009).
118. Adams, G. B. *et al.* Stem cell engraftment at the endosteal niche is specified by the calcium-sensing receptor. *Nature* **439**, 599–603 (2006).
119. Nakamura, Y. *et al.* Isolation and characterization of endosteal niche cell populations that regulate hematopoietic stem cells. *Blood* **116**, 1422–1432 (2010).
120. Méndez-Ferrer, S. *et al.* Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature* **466**, 829–834 (2010).
121. Kiel, M. J. *et al.* SLAM Family Receptors Distinguish Hematopoietic Stem and Progenitor Cells and Reveal Endothelial Niches for Stem Cells. *Cell* **121**, 1109–1121 (2005).
122. Mendes, S. C., Robin, C. & Dzierzak, E. Mesenchymal progenitor cells localize within hematopoietic sites throughout ontogeny. *Development* **132**, 1127–1136 (2005).
123. Sacchetti, B. *et al.* Self-Renewing Osteoprogenitors in Bone Marrow Sinusoids Can Organize a Hematopoietic Microenvironment. *Cell* **131**, 324–336 (2007).
124. Chan, C. K. F. *et al.* Endochondral ossification is required for haematopoietic stem-cell niche formation. *Nature* **457**, 490–494 (2009).
125. Ding, L., Saunders, T. L., Enikolopov, G. & Morrison, S. J. Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature* **481**, 457–462 (2012).
126. Peled, A. & Nagler, A. Role of the CXCR4/CXCL12 Axis in Hematopoietic Stem Cell Trafficking. in *Novel Developments in Stem Cell Mobilization* 71–85 (Springer US, 2012).
127. Porecha, N. K., English, K., Hangoc, G., Broxmeyer, H. E. & Christopherson, K. W. Enhanced Functional Response to CXCL12/SDF-1 Through Retroviral Overexpression of CXCR4 on M07e Cells: Implications for Hematopoietic Stem Cell Transplantation. *Stem Cells Dev.* **15**, 325–333 (2006).
128. Sugiyama, T., Kohara, H., Noda, M. & Nagasawa, T. Maintenance of the Hematopoietic Stem Cell Pool by CXCL12-CXCR4 Chemokine Signaling in Bone Marrow Stromal Cell Niches. *Immunity* **25**, 977–988 (2006).
129. Barker, J. E. Sl/Sld hematopoietic progenitors are deficient in situ. *Exp. Hematol.* **22**, 174–177 (1994).
130. Bowie, M. B., Kent, D. G., Copley, M. R. & Eaves, C. J. Steel factor responsiveness regulates the high self-renewal phenotype of fetal hematopoietic stem cells. *Blood* **109**, 5043–5048 (2007).
131. Istvanffy, R. *et al.* Stromal pleiotrophin regulates repopulation behavior of hematopoietic stem cells. *Blood* **118**, 2712–2722 (2011).
132. Himburg, H. A. *et al.* Pleiotrophin regulates the expansion and regeneration of hematopoietic stem cells. *Nat. Med.* **16**, 475–482 (2010).
133. Zhang, C. C. *et al.* Angiopoietin-like proteins stimulate ex vivo expansion of hematopoietic stem cells. *Nat. Med.* **12**, 240–5 (2006).
134. Kunisaki, Y. *et al.* Arteriolar niches maintain haematopoietic stem cell quiescence. *Nature* **502**, 637–643 (2013).
135. Itkin, T. *et al.* Distinct bone marrow blood vessels differentially regulate haematopoiesis. *Nature* **532**, 323–328 (2016).
136. Yamazaki, S. & Nakauchi, H. Bone marrow Schwann cells induce hematopoietic stem cell hibernation. *Int. J. Hematol.* **99**, 695–698 (2014).
137. Yamazaki, S. *et al.* Nonmyelinating Schwann Cells Maintain Hematopoietic Stem Cell Hibernation in the Bone Marrow Niche. *Cell* **147**, 1146–1158 (2011).
138. Chen, J. Y. *et al.* Hoxb5 marks long-term haematopoietic stem cells and reveals a homogenous perivascular niche. *Nature* **530**, (2016).
139. Acar, M. *et al.* Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. *Nature* **526**, (2015).
140. Christodoulou, C. *et al.* Live-animal imaging of native haematopoietic stem and progenitor cells. *Nature* **578**, 278–283 (2020).

141. Kubota, Y., Takubo, K. & Suda, T. Bone marrow long label-retaining cells reside in the sinusoidal hypoxic niche. *Biochem. Biophys. Res. Commun.* **366**, 335–339 (2008).
142. Miharada, K. *et al.* Hematopoietic stem cells are regulated by Cripto, as an intermediary of HIF-1 α in the hypoxic bone marrow niche. *Ann. N. Y. Acad. Sci.* **1266**, 55–62 (2012).
143. Takubo, K. *et al.* Regulation of the HIF-1 α level is essential for hematopoietic stem cells. *Cell Stem Cell* **7**, 391–402 (2010).
144. Tikhonova, A. N. *et al.* The bone marrow microenvironment at single-cell resolution. *Nature* **569**, 222–228 (2019).
145. Addo, R. K. *et al.* Single-cell transcriptomes of murine bone marrow stromal cells reveal niche-associated heterogeneity. *Eur. J. Immunol.* **49**, 1372–1379 (2019).
146. Baccin, C. *et al.* Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. *Nat. Cell Biol.* **22**, 38–48 (2020).
147. Purton, L. E. & Scadden, D. T. Limiting Factors in Murine Hematopoietic Stem Cell Assays. *Cell Stem Cell* **1**, 263–270 (2007).
148. Dick, J. E., Magli, M. C., Huszar, D., Phillips, R. A. & Bernstein, A. Introduction of a selectable gene into primitive stem cells capable of long-term reconstitution of the hemopoietic system of W/W^v mice. *Cell* **42**, 71–79 (1985).
149. Keller, G., Paige, C., Gilboa, E. & Wagner, E. F. Expression of a foreign gene in myeloid and lymphoid cells derived from multipotent haematopoietic precursors. *Nature* **318**, 149–154 (1985).
150. Lemischka, I. R., Raulet, D. H. & Mulligan, R. C. Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell* **45**, 917–927 (1986).
151. Yamamoto, R., Wilkinson, A. C. & Nakauchi, H. Changing concepts in hematopoietic stem cells. *Science (80-)*. **362**, 895 LP – 896 (2018).
152. Szilvassy, S. J., Humphries, R. K., Lansdorp, P. M., Eaves, A. C. & Eaves, C. J. Quantitative assay for totipotent reconstituting hematopoietic stem cells by a competitive repopulation strategy. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 8736–8740 (1990).
153. Wohrer, S. *et al.* Distinct stromal cell factor combinations can separately control hematopoietic stem cell survival, proliferation, and self-renewal. *Cell Rep.* **7**, 1956–67 (2014).
154. Haas, S., Trumpp, A. & Milsom, M. D. Causes and Consequences of Hematopoietic Stem Cell Heterogeneity. *Cell Stem Cell* **22**, 627–638 (2018).
155. Benveniste, P. *et al.* Intermediate-Term Hematopoietic Stem Cells with Extended but Time-Limited Reconstitution Potential. *Cell Stem Cell* **6**, 48–58 (2010).
156. Kronstein-Wiedemann, R. & Tonn, T. Colony Formation: An Assay of Hematopoietic Progenitor Cells. in *Methods in Molecular Biology* vol. 2017 29–40 (Humana Press Inc., 2019).
157. Bock, T. A. Assay systems for hematopoietic stem and progenitor cells. *Stem Cells* **15**, 185–195 (1997).
158. Whitlock, C. A. & Witte, O. N. Long-term culture of B lymphocytes and their precursors from murine bone marrow. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 3608–3612 (1982).
159. Sutherland, H. J., Lansdorp, P. M., Henkelman, D. H., Eaves, A. C. & Eaves, C. J. Functional characterization of individual human hematopoietic stem cells cultured at limiting dilution on supportive marrow stromal layers. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 3584–3588 (1990).
160. Liu, M., Miller, C. L. & Eaves, C. J. Human Long-Term Culture Initiating Cell Assay. in *Methods in Molecular Biology* vol. 946 241–256 (Humana Press, Totowa, NJ, 2013).
161. Issaad, C., Croisille, L., Katz, A., Vainchenker, W. & Coulombel, L. A murine stromal cell line allows the proliferation of very primitive human CD34⁺/CD38⁻ progenitor cells in long-term cultures and semisolid assays. *Blood* **81**, 2916–2924 (1993).
162. Dexter, T. M., Allen, T. D. & Lajtha, L. G. Conditions controlling the proliferation of haemopoietic stem cells in vitro. *J. Cell. Physiol.* **91**, 335–344 (1977).
163. Sutherland, H., Eaves, C., Lansdorp, P., Thacker, J. & Hogge, D. Differential regulation of primitive human

- hematopoietic cells in long-term cultures maintained on genetically engineered murine stromal cells. *Blood* **78**, 666–672 (1991).
164. La Motte-Mohs, R. N., Herer, E. & Zúñiga-Pflücker, J. C. Induction of T-cell development from human cord blood hematopoietic stem cells by Delta-like 1 in vitro. *Blood* **105**, 1431–1439 (2005).
 165. Shepherd, M. S. *et al.* Single-cell approaches identify the molecular network driving malignant hematopoietic stem cell self-renewal. *Blood* **132**, 791–803 (2018).
 166. Muller-Sieburg, C. E., Whitlock, C. A. & Weissman, I. L. Isolation of Two Early B Lymphocyte Progenitors from Mouse Marrow: A Committed Pre-Pre-B Cell and a Clonogenic Thy-II0 Hematopoietic Stem Cell. *Cell* **44**, 653–662 (1986).
 167. Spangrude, G. J., Heimfeld, S., Weissman, I. L. & Heimfeld, S. Purification and Characterization of Mouse Hematopoietic Stem Cells. *Science (80-.)*. **241**, 58–62 (1988).
 168. Randall, T. D., Lund, F. E., Howard, M. C. & Weissman, I. L. Expression of Murine CD38 Defines a Population of Long-Term Reconstituting Hematopoietic Stem Cells. *Blood* **87**, 4057–4067 (1996).
 169. Chen, C.-Z. *et al.* Identification of endoglin as a functional marker that defines long-term repopulating hematopoietic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 15468–73 (2002).
 170. Christensen, J. L. & Weissman, I. L. Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14541–6 (2001).
 171. Balazs, A. B., Fabian, A. J., Esmon, C. T. & Mulligan, R. C. Endothelial protein C receptor (CD201) explicitly identifies hematopoietic stem cells in murine bone marrow. **107**, 2317–2321 (2006).
 172. Yokota, T. *et al.* The endothelial antigen ESAM marks primitive hematopoietic progenitors throughout life in mice. *Blood* **113**, 2914–23 (2009).
 173. Mayle, A., Luo, M., Jeong, M. & Goodell, M. A. Mouse Hematopoietic Stem Cell Identification And Analysis. *Cytometry. A* **83**, 27–37 (2013).
 174. Bertonecello, I., Hodgson, G. S. & Bradley, T. R. Multiparameter analysis of transplantable hemopoietic stem cells: I. The separation and enrichment of stem cells homing to marrow and spleen on the basis of rhodamine-123 fluorescence. *Exp. Hematol.* **13**, 999–1006 (1985).
 175. Wolf, N. S., Koné, A., Priestley, G. V & Bartelmez, S. H. In vivo and in vitro characterization of long-term repopulating primitive hematopoietic cells isolated by sequential Hoechst 33342-rhodamine 123 FACS selection. *Exp. Hematol.* **21**, 614–22 (1993).
 176. Doulatov, S., Notta, F., Laurenti, E. & Dick, J. E. Hematopoiesis: A Human Perspective. *Cell Stem Cell* **10**, 120–136 (2012).
 177. Huntsman, H. D. *et al.* Human hematopoietic stem cells from mobilized peripheral blood can be purified based on CD49f integrin expression. *Blood* **126**, 1631–3 (2015).
 178. Shultz, L. D. *et al.* Multiple defects in innate and adaptive immunologic function in NOD/LtSz-scid mice. *J. Immunol.* **154**, (1995).
 179. Goyama, S., Wunderlich, M. & Mulloy, J. C. Xenograft models for normal and malignant stem cells. *Blood* **125**, 2630–2640 (2015).
 180. Faiyaz Notta *et al.* Isolation of Single Human Hematopoietic Stem Cells Capable of Long-Term Multilineage Engraftment. *Science (80-.)*. **333**, 218–221 (2011).
 181. Olivier, E., Qiu, C. & Bouhassira, E. E. Novel, High-Yield Red Blood Cell Production Methods from CD34-Positive Cells Derived from Human Embryonic Stem, Yolk Sac, Fetal Liver, Cord Blood, and Peripheral Blood. *Stem Cells Transl. Med.* **1**, 604–614 (2012).
 182. Velten, L. *et al.* Human haematopoietic stem cell lineage commitment is a continuous process. *Nat. Cell Biol.* **19**, 271–281 (2017).
 183. Belluschi, S. *et al.* Myelo-lymphoid lineage restriction occurs in the human haematopoietic stem cell compartment before lymphoid-primed multipotent progenitors. *Nat. Commun.* **9**, (2018).
 184. Kent, D. G. *et al.* Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. *Blood* **113**, 6342–6350 (2009).

185. Benz, C. *et al.* Hematopoietic Stem Cell Subtypes Expand Differentially during Development and Display Distinct Lymphopoietic Programs. *Cell Stem Cell* **10**, 273–83 (2012).
186. Wilson, N. K., Kent, D. G., Theis, F. J. & Göttgens, B. Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. *Cell Stem Cell* **16**, 712–724 (2015).
187. Fatima, S., Zhou, S. & Sorrentino, B. P. Abcg2 expression marks tissue-specific stem cells in multiple organs in a mouse progeny tracking model. *Stem Cells* **30**, 210–21 (2012).
188. Gazit, R. *et al.* Fgd5 identifies hematopoietic stem cells in the murine bone marrow. *J. Exp. Med* **211**, 1315–1331 (2014).
189. Sanjuan-Pla, A. *et al.* Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* **502**, (2013).
190. Nestorowa, S. *et al.* A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood* **128**, e20–e31. (2016).
191. Cheng, C. *et al.* Endothelial cell-specific fgd5 involvement in vascular pruning defines neovessel fate in mice. *Circulation* **125**, 3142–3158 (2012).
192. Kurogane, Y. *et al.* FGD5 mediates proangiogenic action of vascular endothelial growth factor in human vascular endothelial cells. *Arterioscler. Thromb. Vasc. Biol.* **32**, 988–996 (2012).
193. Koechlein, C. S. *et al.* High-resolution imaging and computational analysis of haematopoietic cell dynamics in vivo. *Nat. Commun.* **7**, (2016).
194. Sawai, C. M. *et al.* Hematopoietic Stem Cells Are the Major Source of Multilineage Hematopoiesis in Adult Animals. *Immunity* **45**, 597–609 (2016).
195. Kataoka, K. *et al.* Evi1 is essential for hematopoietic stem cell self-renewal, and its expression marks hematopoietic cells with long-term multilineage repopulating activity. *J. Exp. Med.* **208**, 2403–2416 (2011).
196. Tajima, Y. *et al.* Continuous cell supply from Krt7-expressing hematopoietic stem cells during native hematopoiesis revealed by targeted in vivo gene transfer method. *Sci. Rep.* **7**, 1–10 (2017).
197. Ito, K. *et al.* Self-renewal of a purified Tie2 + hematopoietic stem cell population relies on mitochondrial clearance. *Science (80-.)*. **354**, 1156–1160 (2016).
198. Cabezas-Wallscheid, N. *et al.* Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. *Cell* **169**, 807-823.e19 (2017).
199. Chen, X. *et al.* Bone Marrow Myeloid Cells Regulate Myeloid-Biased Hematopoietic Stem Cells via a Histamine-Dependent Feedback Loop. *Cell Stem Cell* **21**, 747-760.e7 (2017).
200. Suzuki, N. *et al.* Combinatorial Gata2 and Sca1 expression defines hematopoietic stem cells in the bone marrow niche. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2202–2207 (2006).
201. Hills, D. *et al.* Hoxb4-YFP reporter mouse model: A novel tool for tracking HSC development and studying the role of Hoxb4 in hematopoiesis. *Blood* **117**, 3521–3528 (2011).
202. Boyer, S. W., Schroeder, A. V., Smith-Berdan, S. & Forsberg, E. C. All Hematopoietic Cells Develop from Hematopoietic Stem Cells through Flk2/Flt3-Positive Progenitor Cells. *Cell Stem Cell* **9**, 64–73 (2011).
203. TILL, J. E., MCCULLOCH, E. A. & SIMINOVITCH, L. A STOCHASTIC MODEL OF STEM CELL PROLIFERATION, BASED ON THE GROWTH OF. *Proc. Natl. Acad. Sci. United States* **51**, 29–36 (1964).
204. Humphries, R. K., Eaves, A. C. & Eaves, C. J. Self-renewal of hemopoietic stem cells during mixed colony formation in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 3629–3633 (1981).
205. Humphries, R. K., Jacky, P. B., Dill, F. J., Eaves, A. C. & Eaves, C. J. CFU-S in individual erythroid colonies derived in vitro from adult mouse marrow. *Nature* **279**, 718–720 (1979).
206. Dykstra, B. *et al.* Long-Term Propagation of Distinct Hematopoietic Differentiation Programs In Vivo. *Cell Stem Cell* **1**, 218–229 (2007).
207. Morita, Y., Ema, H. & Nakauchi, H. Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. *J. Exp. Med.* **207**, 1173–1182 (2010).
208. Muller-Sieburg, C. E., Cho, R. H., Karlsson, L., Huang, J. F. & Sieburg, H. B. Myeloid-biased hematopoietic

- stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness. *Blood* **103**, 4111–4118 (2004).
209. Naik, S. H. *et al.* Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature* **496**, 229–232 (2013).
 210. Rodriguez-Fraticelli, A. E. *et al.* Clonal analysis of lineage fate in native haematopoiesis. *Nat. Publ. Gr.* **553**, (2018).
 211. Yu, V. W. C. *et al.* Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells. *Cell* **167**, 1310–1322 (2016).
 212. Benveniste, P., Cantin, C., Hyam, D. & Iscove, N. N. Hematopoietic stem cells engraft in mice with absolute efficiency. *Nat. Immunol.* **4**, 708–713 (2003).
 213. Hamanaka, S. *et al.* Generation of transgenic mouse line expressing Kusabira Orange throughout body, including erythrocytes, by random segregation of provirus method. *Biochem. Biophys. Res. Commun.* **435**, 586–591 (2013).
 214. Carrelha, J. *et al.* Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature* **554**, 106–111 (2018).
 215. Yamamoto, R. *et al.* Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* **154**, 1112–1126 (2013).
 216. Rossi, L. *et al.* Less is more: Unveiling the functional core of hematopoietic stem cells through knockout mice. *Cell Stem Cell* **11**, 302–317 (2012).
 217. Phillips, R. L. *et al.* The Genetic Program of Hematopoietic Stem Cells. *Science (80-.)*. **288**, 1635–1640 (2000).
 218. Terskikh, A. V *et al.* From Hematopoiesis to Neurogenesis: Evidence of Overlapping Genetic Programs. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 7934–7939 (2001).
 219. Ivanova, N. B. *et al.* A stem cell molecular signature. *Science (80-.)*. **298**, 601–604 (2002).
 220. Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R. C. & Melton, D. A. ‘Stemness’: Transcriptional profiling of embryonic and adult stem cells. *Science (80-.)*. **298**, 597–600 (2002).
 221. Fortunel, N. O. Comment on ‘ “Stemness”: Transcriptional Profiling of Embryonic and Adult Stem Cells’ and ‘A Stem Cell Molecular Signature’ (I). *Science (80-.)*. **302**, 393b – 393 (2003).
 222. Venezia, T. A. *et al.* Molecular Signatures of Proliferation and Quiescence in Hematopoietic Stem Cells. *PLoS Biol.* **2**, e301 (2004).
 223. Chambers, S. M. *et al.* Hematopoietic Fingerprints: An Expression Database of Stem Cells and Their Progeny. *Cell Stem Cell* **1**, 578–591 (2007).
 224. Hu, M. *et al.* Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev.* **11**, 774–785 (1997).
 225. Brady, G., Barbara, M. & Iscove, N. N. Representative in Vitro cDNA Amplification From Individual Hemopoietic Cells and Colonies. *METHODS Mol. Cell. Biol.* **2**, 17–25 (1990).
 226. Ramos, C. A. *et al.* Evidence for Diversity in Transcriptional Profiles of Single Hematopoietic Stem Cells. *PLoS Genet.* **2**, e159 (2006).
 227. Cheng, T. *et al.* Temporal mapping of gene expression levels during the differentiation of individual primary hematopoietic cells. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 13158–13163 (1996).
 228. Miyamoto, T. *et al.* Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. *Dev. Cell* **3**, 137–147 (2002).
 229. Adolfsson, J. *et al.* Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential: A revised road map for adult blood lineage commitment. *Cell* **121**, 295–306 (2005).
 230. Månsson, R. *et al.* Molecular Evidence for Hierarchical Transcriptional Lineage Priming in Fetal and Adult Stem Cells and Multipotent Progenitors. *Immunity* **26**, 407–419 (2007).
 231. Glotzbach, J. P. *et al.* An information theoretic, microfluidic-based single cell analysis permits identification of subpopulations among putatively homogeneous stem cells. *PLoS One* **6**, 21211 (2011).

232. Moignard, V. *et al.* Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. *Nat. Cell Biol.* **15**, 363–372 (2013).
233. Nerlov, C. & Graf, T. PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors. *Genes Dev.* **12**, 2403–2412 (1998).
234. Doan, L. L. *et al.* Targeted transcriptional repression of Gfi1 by GFI1 and GFI1B in lymphoid cells. *Nucleic Acids Res.* **32**, (2004).
235. Vassen, L., Okayama, T. & Möröy, T. Gfi1b:green fluorescent protein knock-in mice reveal a dynamic expression pattern of Gfi1b during hematopoiesis that is largely complementary to Gfi1. *Blood* **109**, 2356–2364 (2007).
236. Schütte, J. *et al.* An experimentally validated network of nine haematopoietic transcription factors reveals mechanisms of cell state stability. *Elife* **5**, (2016).
237. Psaila, B. *et al.* Single-cell profiling of human megakaryocyte-erythroid progenitors identifies distinct megakaryocyte and erythroid differentiation pathways. *Genome Biol.* **17**, 83 (2016).
238. Guo, G. *et al.* Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire. *Cell Stem Cell* **13**, 492–505 (2013).
239. Petriv, O. I. *et al.* Comprehensive microRNA expression profiling of the hematopoietic hierarchy. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15443–15448 (2010).
240. Watcham, S., Kucinski, I. & Göttgens, B. New insights into hematopoietic differentiation landscapes from single-cell RNA sequencing. *Blood* **133**, 1415–1426 (2019).
241. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
242. Hashimshony, T. *et al.* CEL-Seq2: Sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol.* **17**, (2016).
243. Bagnoli, J. W. *et al.* Sensitive and powerful single-cell RNA sequencing using mcSCRIB-seq. *Nat. Commun.* **9**, (2018).
244. Paul, F. *et al.* Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell* **163**, 1663–1677 (2015).
245. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, (2017).
246. Dahlin, J. S. *et al.* A single-cell hematopoietic landscape resolves 8 lineage trajectories and defects in Kit mutant mice. *Blood* **131**, e1–e11 (2018).
247. Dong, F. *et al.* Differentiation of transplanted haematopoietic stem cells tracked by single-cell transcriptomic analysis. *Nat. Cell Biol.* **22**, 630–639 (2020).
248. Fukushima, T. *et al.* Discrimination of Dormant and Active Hematopoietic Stem Cells by G0 Marker Reveals Dormancy Regulation by Cytoplasmic Calcium. *Cell Rep.* **29**, 4144–4158.e7 (2019).
249. Olsson, A. *et al.* Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature* **537**, 698–702 (2016).
250. Grün, D. *et al.* De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome Data. *Cell Stem Cell* **19**, 266–277 (2016).
251. Notta, F. *et al.* Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science (80-.).* **351**, (2016).
252. Cabezas-Wallscheid, N. *et al.* Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. *Cell Stem Cell* **15**, 507–522 (2014).
253. Weinreb, C., Wolock, S., Tusi, B. K., Socolovsky, M. & Klein, A. M. Fundamental limits on dynamic inference from single-cell snapshots. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E2467–E2476 (2018).
254. Florian, M. C. *et al.* Aging alters the epigenetic asymmetry of HSC division. *PLOS Biol.* **16**, e2003389 (2018).
255. Lipka, D. B. *et al.* Identification of dna methylation changes at cis-regulatory elements during early steps of hsc differentiation using tagmentation-based whole genome bisulfite sequencing. *Cell Cycle* **13**, 3476–3487 (2014).
256. Zhang, X. *et al.* DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors

- in hematopoietic stem cells. *Nat. Publ. Gr.* **48**, (2016).
257. Sharma, S. & Gurudutta, G. Epigenetic Regulation of Hematopoietic Stem Cells. *Int. J. stem cells* **9**, 36–43 (2016).
 258. Roberts, A., Trapnell, C., Donaghey, J., Rinn, J. L. & Pachter, L. Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol.* **12**, R22 (2011).
 259. Dou, M. *et al.* High-Throughput Single Cell Proteomics Enabled by Multiplex Isobaric Labeling in a Nanodroplet Sample Preparation Platform. *Anal. Chem.* **91**, 13119–13127 (2019).
 260. Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* **19**, 161 (2018).
 261. Bendall, S. C. *et al.* Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science (80-.)*. **332**, 687–696 (2011).
 262. Hughes, A. J. *et al.* Single-cell western blotting. *Nat. Methods* **11**, 749–755 (2014).
 263. Knapp, D. J. H. F. *et al.* Distinct signaling programs control human hematopoietic stem cell survival and proliferation. *Blood* **129**, 307–318 (2017).
 264. Karamitros, D. *et al.* Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells article. *Nat. Immunol.* **19**, 85–97 (2018).
 265. Challen, G. A., Boles, N. C., Chambers, S. M. & Goodell, M. A. Distinct hematopoietic stem cell subtypes are differentially regulated by TGF-beta1. *Cell Stem Cell* **6**, 265–78 (2010).
 266. Sun, J. *et al.* Clonal dynamics of native haematopoiesis. *Nature* **514**, 322–327 (2014).
 267. Kumar, S. & Geiger, H. HSC Niche Biology and HSC Expansion Ex Vivo. *Trends Mol. Med.* **23**, 799–819 (2017).
 268. Meuwissen, H. J., Gatti, R. A., Terasaki, P. I., Hong, R. & Good, R. A. Treatment of Lymphopenic Hypogammaglobulinemia and Bone-Marrow Aplasia by Transplantation of Allogeneic Marrow. *N. Engl. J. Med.* **281**, 691–697 (1969).
 269. Remberger, M. *et al.* Effect of total nucleated and CD34+ cell dose on outcome after allogeneic hematopoietic stem cell transplantation. *Biol. Blood Marrow Transplant.* **21**, 889–893 (2015).
 270. Luigi Naldini. Gene therapy returns to centre stage. *Nat. Rev.* **526**, 351–60 (2015).
 271. Ok Kim, H. In-Vitro Stem Cell Derived Red Blood Cells for Transfusion: Are We There Yet? *Yonsei Med J* **55**, (2014).
 272. Aiuti, A. *et al.* Lentiviral Hematopoietic Stem Cell Gene Therapy in Patients with Wiskott-Aldrich Syndrome. *Science (80-.)*. **341**, 1233151–1233151 (2013).
 273. Biffi, A. *et al.* Lentiviral Hematopoietic Stem Cell Gene Therapy Benefits Metachromatic Leukodystrophy. *Science (80-.)*. **341**, (2013).
 274. Ribeil, J.-A. *et al.* Gene Therapy in a Patient with Sickle Cell Disease. *N. Engl. J. Med.* **376**, 848–855 (2017).
 275. Park, B., Yoo, K. H. & Kim, C. Hematopoietic stem cell expansion and generation: The ways to make a breakthrough. *Blood Res.* **50**, 194–203 (2015).
 276. Gratwohl, A. *et al.* One million haemopoietic stem-cell transplants: a retrospective observational study. *Lancet Haematol.* **2**, e91–e100 (2015).
 277. Trakarnsanga, K. *et al.* An immortalized adult human erythroid line facilitates sustainable and scalable generation of functional red cells. *Nat. Commun.* **8**, 14750 (2017).
 278. Masuda, S., Li, M. & Belmonte, J. C. I. In vitro generation of platelets through direct conversion: first report in My Knowledge (iMK). *Cell Res.* **23**, 176–178 (2013).
 279. Singh, V. K., Saini, A., Tsuji, K., Sharma, P. B. & Chandra, R. Manufacturing blood ex vivo: a futuristic approach to deal with the supply and safety concerns. *Front. cell Dev. Biol.* **2**, 26 (2014).
 280. Dick, E. P., Prince, L. R. & Sabroe, I. Ex Vivo-Expanded Bone Marrow CD34 + Derived Neutrophils Have Limited Bactericidal Ability. *Stem Cells* **26**, 2552–2563 (2008).

281. Saeki, K. *et al.* A Feeder-Free and Efficient Production of Functional Neutrophils from Human Embryonic Stem Cells. *Stem Cells* **27**, 59–67 (2009).
282. Seet, C. S. *et al.* Generation of mature T cells from human hematopoietic stem and progenitor cells in artificial thymic organoids. *Nat. Methods* **14**, 521–530 (2017).
283. Wilkinson, A. C. *et al.* Long-term ex vivo haematopoietic-stem-cell expansion allows nonconditioned transplantation. *Nature* **571**, 117–121 (2019).
284. Zhang, C. C. & Lodish, H. F. Murine hematopoietic stem cells change their surface phenotype during ex vivo expansion. *Blood* **105**, 4314–4320. (2005).
285. Fares, I. *et al.* EPCR expression marks UM171-expanded CD34+ cord blood stem cells. *Blood* **129**, 3344–3351 (2017).
286. Tomellini, E. *et al.* Integrin- α 3 Is a Functional Marker of Ex Vivo Expanded Human Long-Term Hematopoietic Stem Cells. *Cell Rep.* **28**, 1063-1073.e5 (2019).
287. Montel-Hagen, A. & Crooks, G. M. From pluripotent stem cells to T cells. *Exp. Hematol.* **71**, 24–31 (2019).
288. Bernarreggi, D., Pouyanfard, S. & Kaufman, D. S. Development of innate immune cells from human pluripotent stem cells. *Exp. Hematol.* **71**, 13–23 (2019).
289. Lis, R. *et al.* Conversion of adult endothelium to immunocompetent haematopoietic stem cells. *Nature* **545**, 439–445 (2017).
290. Ema, H. *et al.* Quantification of self-renewal capacity in single hematopoietic stem cells from normal and Lnk-deficient mice. *Dev. Cell* **8**, 907–914 (2005).
291. Iscove, N. N. & Nawa, K. Hematopoietic stem cells expand during serial transplantation in vivo without apparent exhaustion. *Curr. Biol.* **7**, 805–8 (1997).
292. Sekulovic, S. *et al.* Ontogeny stage-independent and high-level clonal expansion in vitro of mouse hematopoietic stem cells stimulated by an engineered NUP98-HOX fusion transcription factor. *Blood* **118**, 4366–4376 (2011).
293. Kent, D. G. *et al.* Isolation and Assessment of Single Long-Term Reconstituting Hematopoietic Stem Cells from Adult Mouse Bone Marrow. in *Current Protocols in Stem Cell Biology* 2A.4.1-2A.4.24 (John Wiley & Sons, Inc., 2016).
294. Zon, L. I. Intrinsic and extrinsic control of haematopoietic stem-cell self-renewal. *Nature* **453**, 306–313 (2008).
295. Reya, T. *et al.* A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature* **423**, 409–414 (2003).
296. Murphy, M. J., Wilson, A. & Trumpp, A. More than just proliferation: Myc function in stem cells. *Trends Cell Biol.* **15**, 128–137 (2005).
297. Helgason, C. D., Sauvageau, G., Lawrence, H. J., Largman, C. & Humphries, R. K. Overexpression of HOXB4 Enhances the Hematopoietic Potential of Embryonic Stem Cells Differentiated In Vitro. *Blood* **87**, 2740–2749 (1996).
298. Antonchuk, J., Guy Sauvageau & Humphries, R. K. HOXB4-Induced Expansion of Adult Hematopoietic Stem Cells Ex Vivo. *Cell* **109**, 39–45 (2002).
299. Kros, J. *et al.* In vitro expansion of hematopoietic stem cells by recombinant TAT-HOXB4 protein. *Nat. Med.* **9**, 1428–1432 (2001).
300. Rodrigues, N. P. *et al.* Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood* **106**, (2005).
301. Iwasaki, H. *et al.* Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* **106**, 1590–600 (2005).
302. Bruno, L. *et al.* Molecular signatures of self-renewal, differentiation, and lineage choice in multipotential hemopoietic progenitor cells in vitro. *Mol. Cell. Biol.* **24**, 741–56 (2004).
303. Gowney, J. D. *et al.* Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. *Blood* **106**, 494–504 (2005).

304. Passegué, E., Wagner, E. F. & Weissman, I. L. JunB Deficiency Leads to a Myeloproliferative Disorder Arising from Hematopoietic Stem Cells. *Cell* **119**, 431–443 (2004).
305. Unnisa, Z. *et al.* Meis1 preserves hematopoietic stem cells in mice by limiting oxidative stress. *Blood* **120**, 4973–81 (2012).
306. Ficara, F., Murphy, M. J., Lin, M. & Cleary, M. L. Pbx1 Regulates Self-Renewal of Long-Term Hematopoietic Stem Cells by Maintaining Their Quiescence. *Cell Stem Cell* **2**, 484–496 (2008).
307. He, S., Kim, I., Lim, M. S. & Morrison, S. J. Sox17 expression confers self-renewal potential and fetal stem cell characteristics upon adult hematopoietic progenitors. *Genes Dev.* **25**, 1613–1627 (2011).
308. Avgustinova, A. & Aznar Benitah, S. Epigenetic control of adult stem cell function. *Nat. Publ. Gr.* **17**, (2016).
309. Yuko Tadokoro, Hideo Ema, Masaki Okano, En Li, H. N. De novo DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells. *J. Exp. Med.* **204**, 715–722 (2007).
310. Challen, G. *et al.* Dnmt3a and Dnmt3b Have Overlapping and Distinct Functions in Hematopoietic Stem Cells. *Cell Stem Cell* **15**, 350–364 (2014).
311. Moran-Crusio, K. *et al.* Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**, 11–24 (2011).
312. Chaurasia, P., Gajzer, D. C., Schaniel, C., D'Souza, S. & Hoffman, R. Epigenetic reprogramming induces the expansion of cord blood stem cells. *J. Clin. Invest.* **124**, 2378–2395 (2014).
313. Lessard, J., Baban, S. & Sauvageau, G. Stage-Specific Expression of Polycomb Group Genes in Human Bone Marrow Cells. *Blood* **91**, 1216–1224 (1998).
314. Kajiume, T., Ninomiya, Y., Ishihara, H., Kanno, R. & Kanno, M. Polycomb group gene mel-18 modulates the self-renewal activity and cell cycle status of hematopoietic stem cells. *Exp. Hematol.* **32**, 571–578 (2004).
315. Klauke, K. *et al.* Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. *Nat. Cell Biol.* **15**, 353–362 (2013).
316. Iwama, A. *et al.* Enhanced Self-Renewal of Hematopoietic Stem Cells Mediated by the Polycomb Gene Product Bmi-1. *Immunity* **21**, 843–851 (2004).
317. Vasilatou, D., Papageorgiou, S., Pappa, V., Papageorgiou, E. & Dervenoulas, J. The role of microRNAs in normal and malignant hematopoiesis. *Eur. J. Haematol.* **84**, 1–16 (2010).
318. Guo, S. *et al.* MicroRNA miR-125a controls hematopoietic stem cell number. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 14229–34 (2010).
319. Yoon-Chi Han *et al.* microRNA-29a induces aberrant self-renewal capacity in hematopoietic progenitors, biased myeloid development, and acute myeloid leukemia. *J. Exp. Med.* **207**, 475–489 (2010).
320. Capitano, M. L. Toll-like receptor signaling in hematopoietic stem and progenitor cells. *Curr. Opin. Hematol.* **26**, 207–213 (2019).
321. Takizawa, H. *et al.* Pathogen-Induced TLR4-TRIF Innate Immune Signaling in Hematopoietic Stem Cells Promotes Proliferation but Reduces Competitive Fitness. *Cell Stem Cell* **21**, 225–240.e5 (2017).
322. Pietras, E. M. *et al.* Chronic interleukin-1 exposure drives haematopoietic stem cells towards precocious myeloid differentiation at the expense of self-renewal. *Nat. Cell Biol.* **18**, 607–618 (2016).
323. Mantel, C. R. *et al.* Enhancing Hematopoietic Stem Cell Transplantation Efficacy by Mitigating Oxygen Shock. *Cell* **161**, 1553–1565 (2015).
324. Povinelli, B. J. & Nemeth, M. J. Wnt5a Regulates Hematopoietic Stem Cell Proliferation and Repopulation Through the Ryk Receptor. *Stem Cells* **32**, 105–115 (2014).
325. Sauvageau, G., Iscove, N. N. & Humphries, R. K. In vitro and in vivo expansion of hematopoietic stem cells. *Oncogene* **23**, 7223–32 (2004).
326. Delaney, C. *et al.* Notch-mediated expansion of human cord blood progenitor cells capable of rapid myeloid reconstitution. *Nat. Med.* **16**, 232–236 (2010).
327. Stier, S., Cheng, T., Dombkowski, D., Carlesso, N. & Scadden, D. T. Notch1 activation increases hematopoietic stem cell self-renewal in vivo and favors lymphoid over myeloid lineage outcome. *Blood* **99**,

- 2369–78 (2002).
328. Yuki Taya,¹ Yasunori Ota,² Adam C. Wilkinson,^{1,3} Ayano Kanazawa, ¹ & Yuki Taya, Yasunori Ota, Adam C. Wilkinson, Ayano Kanazawa, Hiroshi Watarai, Masataka Kasai, Hiromitsu Nakauchi, S. Y. Depleting dietary valine permits nonmyeloablative mouse hematopoietic stem cell transplantation. *Science (80-)*. **354**, 1152–1155 (2016).
329. Cheng, C. W. *et al.* Prolonged fasting reduces IGF-1/PKA to promote hematopoietic-stem-cell- based regeneration and reverse immunosuppression. *Cell Stem Cell* **14**, 810–823 (2014).
330. Lazare, S. *et al.* Lifelong dietary intervention does not affect hematopoietic stem cell function. *Exp. Hematol.* **53**, 26–30 (2017).
331. Tang, D. *et al.* Dietary restriction improves repopulation but impairs lymphoid differentiation capacity of hematopoietic stem cells in early aging. *J. Exp. Med.* **213**, 535–553 (2016).
332. Lee, H. J., Li, N., Evans, S. M., Diaz, M. F. & Wenzel, P. L. Biomechanical force in blood development: Extrinsic physical cues drive pro-hematopoietic signaling. *Differentiation* **86**, 92–103 (2013).
333. Holst, J. *et al.* Substrate elasticity provides mechanical signals for the expansion of hemopoietic stem and progenitor cells. *Nat. Biotechnol.* **28**, 1123–1128 (2010).
334. Huang, W. *et al.* Regulatory networks in mechanotransduction reveal key genes in promoting cancer cell stemness and proliferation. *Oncogene* **38**, 6818–6834 (2019).
335. Oostendorp, R. A. J. *et al.* Long-Term Maintenance of Hematopoietic Stem Cells Does Not Require Contact with Embryo-Derived Stromal Cells in Cocultures. *Stem Cells* **23**, 842–851 (2005).
336. Metcalf, D. Hematopoietic cytokines. *Blood* **111**, 485–491 (2008).
337. Audet, J., Miller, C. L., Eaves, C. J. & Piret, J. M. Common and Distinct Features of Cytokine Effects on Hematopoietic Stem and Progenitor Cells Revealed by Dose-Response Surface Analysis. *Biotechnology Bioeng.* **80**, 393–404 (2002).
338. McCulloch, E. A., Siminovitch, L., Till, J. E., Russell, E. S. & Bernstein, S. E. The Cellular Basis of the Genetically Determined Hemopoietic Defect in Anemic Mice of Genotype SI/SF. *Blood* **26**, 399–410 (1965).
339. McCulloch, E. A., Siminovitch, L. & Till, J. E. Spleen-Colony Formation in Anemic Mice of Genotype WW *v. Repr. from Sci.* **144**, 844–846 (1964).
340. Tadokoro, Y. *et al.* Spred1 Safeguards Hematopoietic Homeostasis against Diet-Induced Systemic Stress. *Cell Stem Cell* **22**, 713-725.e8 (2018).
341. Audet, J., Miller, C. L., Rose-John, S., Piret, J. M. & Eaves, C. J. Distinct role of gp130 activation in promoting self-renewal divisions by mitogenically stimulated murine hematopoietic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 1757–62 (2001).
342. Solar, G. P. *et al.* Role of c-mpl in Early Hematopoiesis. *Blood* **92**, 4–10 (1998).
343. De Graaf, C. A. & Metcalf, D. Thrombopoietin and hematopoietic stem cells. *Cell Cycle* **10**, 1582–1589 (2011).
344. Seita, J. *et al.* Lnk negatively regulates self-renewal of hematopoietic stem cells by modifying thrombopoietin-mediated signal transduction. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2349–54 (2007).
345. Gery, S. *et al.* Lnk inhibits myeloproliferative disorder-associated JAK2 mutant, JAK2V617F. *J. Leukoc. Biol.* **85**, 957–965 (2009).
346. Yoshida, K. *et al.* Targeted disruption of gp130, a common signal transducer for the interleukin 6 family of cytokines, leads to myocardial and hematological disorders. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 407–411 (1996).
347. Bernad, A. *et al.* Interleukin-6 is required in vivo for the regulation of stem cells and committed progenitors of the hematopoietic system. *Immunity* **1**, 725–731 (1994).
348. Hawley, R. G. *et al.* Thrombopoietic potential and serial repopulating ability of murine hematopoietic stem cells constitutively expressing interleukin 11. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10297–10302 (1996).
349. Matsunaga, T., Kato, T., Miyazaki, H. & Ogawa, M. Thrombopoietin Promotes the Survival of Murine Hematopoietic Long-Term Reconstituting Cells: Comparison With the Effects of FLT3/FLK-2 Ligand and Interleukin-6. *Blood* **92**, 452–461 (1998).

350. Yamazaki, S. *et al.* TGF- β 2 as a candidate bone marrow niche signal to induce hematopoietic stem cell hibernation. *Blood* **113**, 1250–1256 (2009).
351. Yonemura, Y., Ku, H., Lyman, S. D. & Ogawa, M. In vitro expansion of hematopoietic progenitors and maintenance of stem cells: comparison between FLT3/FLK-2 ligand and KIT ligand. *Blood* **89**, 1915–1921 (1997).
352. Ueda, T. *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.* **105**, 1013–21 (2000).
353. Miller, C. L. & Eaves, C. J. Expansion in vitro of adult murine hematopoietic stem cells with transplantable lympho-myeloid reconstituting ability. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 13648–13653 (1997).
354. Petzer, A. L., Zandstra, P. W., Piret, J. M. & Eaves, C. J. Differential cytokine effects on primitive (CD34+CD38-) human hematopoietic cells: Novel responses to Flt3-ligand and thrombopoietin. *J. Exp. Med.* **183**, 2551–2558 (1996).
355. Audet, J., Zandstra, P. W., Eaves, C. J. & Piret, J. M. Advances in hematopoietic stem cell culture. *Curr. Opin. Biotechnol.* **9**, 146–151 (1998).
356. Iriuchishima, H. *et al.* Ex vivo maintenance of hematopoietic stem cells by quiescence induction through Fbxw7 α overexpression. *Blood* **117**, 2373–2377 (2011).
357. Miharada, K., Sigurdsson, V. & Karlsson, S. Dppa5 improves hematopoietic stem cell activity by reducing endoplasmic reticulum stress. *Cell Rep.* **7**, 1381–1392 (2014).
358. Kroon, E., Thorsteinsdottir, U., Mayotte, N., Nakamura, T. & Sauvageau, G. NUP98-HOXA9 expression in hemopoietic stem cells induces chronic and acute myeloid leukemias in mice. *EMBO J.* **20**, 350–361 (2001).
359. Chung, K. Y. *et al.* Enforced expression of NUP98-HOXA9 in human CD34+ cells enhances stem cell proliferation. *Cancer Res.* **66**, 11781–11791 (2006).
360. Deneault, E. *et al.* A Functional Screen to Identify Novel Effectors of Hematopoietic Stem Cell Activity. *Cell* **137**, 369–379 (2009).
361. Oostendorp, R. A. J. *et al.* Stromal cell lines from mouse aorta-gonads-mesonephros subregions are potent supporters of hematopoietic stem cell activity. *Blood* **99**, 1183–1189 (2002).
362. Zhang, C. C. & Lodish, H. F. Insulin-like growth factor 2 expressed in a novel fetal liver cell population is a growth factor for hematopoietic stem cells. *Blood* **103**, 2513–2521 (2004).
363. Moore, K. A., Ema, H. & Lemischka, I. R. In Vitro Maintenance of Highly Purified, Transplantable Hematopoietic Stem Cells. *Blood* **89**, 4337–4347 (1997).
364. Ohneda, O. *et al.* Hematopoietic Stem Cell Maintenance and Differentiation Are Supported by Embryonic Aorta-Gonad-Mesonephros Region-Derived Endothelium. *Blood* **92**, 908–919 (1998).
365. Xu, M.-J. *et al.* Stimulation of Mouse and Human Primitive Hematopoiesis by Murine Embryonic Aorta-Gonad-Mesonephros-Derived Stromal Cell Lines. *Blood* **92**, 2032–2040 (1998).
366. de Lima, M. *et al.* Cord-Blood Engraftment with Ex Vivo Mesenchymal-Cell Coculture. *N. Engl. J. Med.* **367**, 2305–2315 (2012).
367. Rebel, V. I., Dragowska, W., Eaves, C. J., Humphries, R. K. & Lansdorp, P. M. Amplification of Sca-1+ Lin-WGA+ Cells in Serum-Free Cultures Containing Steel Factor, Interleukin-6, and Erythropoietin With Maintenance of Cells With Long-Term In Vivo Reconstituting Potential. *Blood* **83**, 128–136 (1994).
368. Jingjing, X. & Chengcheng, Z. Ex vivo expansion of hematopoietic stem cells. *Sci. China Life Sci.* **58**, 839–853 (2015).
369. Fares, I., Rivest-Khan, L., Cohen, S. & Sauvageau, G. Small molecule regulation of normal and leukemic stem cells. *Curr. Opin. Hematol.* **22**, 309–316 (2015).
370. Walasek, M. A., van Os, R. & de Haan, G. Hematopoietic stem cell expansion: challenges and opportunities. *Ann. N. Y. Acad. Sci.* **1266**, 138–150 (2012).
371. Huynh, H. *et al.* Insulin-Like Growth Factor-Binding Protein 2 Secreted by a Tumorigenic Cell Line Supports Ex Vivo Expansion of Mouse Hematopoietic Stem Cells. *Stem Cells* **26**, 1628–1635 (2008).
372. Zhao, Y. *et al.* ATF4 plays a pivotal role in the development of functional hematopoietic stem cells in mouse

- fetal liver. *Blood* **126**, 2383–2391 (2015).
373. Huang, J., Nguyen-Mccarty, M., Hexner, E. O., Danet-Desnoyers, G. & Klein, P. S. Maintenance of hematopoietic stem cells through regulation of Wnt and mTOR pathways. *Nat. Med.* **18**, 1778–1785 (2012).
374. Cobas, M. *et al.* β -Catenin Is Dispensable for Hematopoiesis and Lymphopoiesis. *J. Exp. Med.* **199**, 221–229 (2004).
375. Kabiri, Z. *et al.* Wnts are dispensable for differentiation and self-renewal of adult murine hematopoietic stem cells. *Blood* **126**, 1086–1094 (2015).
376. Kirstetter, P., Anderson, K., Porse, B. T., Jacobsen, S. E. W. & Nerlov, C. Activation of the canonical Wnt pathway leads to loss of hematopoietic stem cell repopulation and multilineage differentiation block. *Nat. Immunol.* **7**, 1048–1056 (2006).
377. Koch, U. *et al.* Simultaneous loss of β - and γ -catenin does not perturb hematopoiesis or lymphopoiesis. *Blood* **111**, 160–164 (2008).
378. Purton, L. E., Bernstein, I. D. & Collins, S. J. All-trans retinoic acid enhances the long-term repopulating activity of cultured hematopoietic stem cells. *Blood* **95**, 470–7 (2000).
379. Chute, J. P. *et al.* Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 11707–11712 (2006).
380. Ghiaur, G. *et al.* Regulation of human hematopoietic stem cell self-renewal by the microenvironment's control of retinoic acid signaling. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 16121–16126 (2013).
381. Nishino, T. *et al.* Ex Vivo Expansion of Human Hematopoietic Stem Cells by Garcinol, a Potent Inhibitor of Histone Acetyltransferase. *PLoS One* **6**, e24298 (2011).
382. Seet, L.-F. *et al.* Valproic acid enhances the engraftability of human umbilical cord blood hematopoietic stem cells expanded under serum-free conditions*. *Eur. J. Haematol.* **82**, 124–132 (2009).
383. Bug, G. *et al.* Valproic Acid Stimulates Proliferation and Self-renewal of Hematopoietic Stem Cells. *Cancer Res.* **65**, 2537–2541 (2005).
384. Walasek, M. A. *et al.* The combination of valproic acid and lithium delays hematopoietic stem/progenitor cell differentiation. *Blood* **119**, 3050–3059 (2012).
385. Luchsinger, L. L. *et al.* Harnessing Hematopoietic Stem Cell Low Intracellular Calcium Improves Their Maintenance In Vitro. *Cell Stem Cell* **25**, 225-240.e7 (2019).
386. Fares, I. *et al.* Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. *Science (80-.)*. **345**, 1509–1512 (2014).
387. Cohen, S. *et al.* Hematopoietic stem cell transplantation using single UM171-expanded cord blood: a single-arm, phase 1–2 safety and feasibility study. *Lancet Haematol.* **7**, e134–e145 (2020).
388. Hoggatt, J., Mohammad, K. S., Singh, P. & Pelus, L. M. Prostaglandin E2 enhances long-term repopulation but does not permanently alter inherent stem cell competitiveness. *Blood* **122**, 2997–3000 (2013).
389. Hoggatt, J., Singh, P., Sampath, J. & Pelus, L. M. Prostaglandin E2 enhances hematopoietic stem cell homing, survival, and proliferation. *Blood* **113**, 5444–55 (2009).
390. Goessling, W. *et al.* Genetic Interaction of PGE2 and Wnt Signaling Regulates Developmental Specification of Stem Cells and Regeneration. *Cell* **136**, 1136–1147 (2009).
391. Ko, K.-H. *et al.* GSK-3 β Inhibition Promotes Engraftment of Ex Vivo-Expanded Hematopoietic Stem Cells and Modulates Gene Expression. *Stem Cells* **29**, 108–118 (2011).
392. Varnum-Finney, B., Brashem-Stein, C. & Bernstein, I. D. Combined effects of Notch signaling and cytokines induce a multiple log increase in precursors with lymphoid and myeloid reconstituting ability. *Blood* **101**, (2003).
393. Guo, B., Huang, X., Lee, M. R., Lee, S. A. & Broxmeyer, H. E. Antagonism of PPAR- γ signaling expands human hematopoietic stem and progenitor cells by enhancing glycolysis. *Nat. Med.* **24**, 360–367 (2018).
394. Tiwari, A. *et al.* Ex vivo expansion of haematopoietic stem/progenitor cells from human umbilical cord blood on acellular scaffolds prepared from MS-5 stromal cell line. *J. Tissue Eng. Regen. Med.* **7**, 871–883 (2013).
395. Bello, A. B., Park, H. & Lee, S.-H. Current approaches in biomaterial-based hematopoietic stem cell niches.

- Acta Biomater.* **72**, 1–15 (2018).
396. Lowndes, M., Junyent, S. & Habib, S. J. Constructing cellular niche properties by localized presentation of Wnt proteins on synthetic surfaces. *Nat. Protoc.* **12**, 1498–1512 (2017).
397. Bourguine, P. E. *et al.* In vitro biomimetic engineering of a human hematopoietic niche with functional properties. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E5688–E5695 (2018).
398. Bourguine, P. E., Martin, I. & Schroeder, T. Engineering Human Bone Marrow Proxies. *Cell Stem Cell* **22**, 298–301 (2018).
399. Choi, J. S., Mahadik, B. P. & Harley, B. A. C. Engineering the hematopoietic stem cell niche: Frontiers in biomaterial science. *Biotechnol. J.* **10**, 1529–1545 (2015).
400. Franke, K., Pompe, T., Bornhäuser, M. & Werner, C. Engineered matrix coatings to modulate the adhesion of CD133+ human hematopoietic progenitor cells. *Biomaterials* **28**, 836–843 (2007).
401. Ventura Ferreira, M. S. *et al.* Cord blood-hematopoietic stem cell expansion in 3D fibrin scaffolds with stromal support. *Biomaterials* **33**, 6987–6997 (2012).
402. Yeoh, J. S. G. *et al.* Fibroblast Growth Factor-1 and -2 Preserve Long-Term Repopulating Ability of Hematopoietic Stem Cells in Serum-Free Cultures. *Stem Cells* **24**, 1564–1572 (2006).
403. Zhang, C. C., Kaba, M., Iizuka, S., Huynh, H. & Lodish, H. F. Angiopoietin-like 5 and IGFBP2 stimulate ex vivo expansion of human cord blood hematopoietic stem cells as assayed by NOD/SCID transplantation. *Blood* **111**, 3415–3423 (2008).
404. Peled, T. *et al.* Pre-clinical development of cord blood-derived progenitor cell graft expanded ex vivo with cytokines and the polyamine copper chelator tetraethylenepentamine. *Cytotherapy* **6**, 344–355 (2004).
405. Young, J. C. *et al.* Inhibitors of histone deacetylases promote hematopoietic stem cell self-renewal. *Cytotherapy* **6**, 328–336 (2004).
406. Araki, H. *et al.* Chromatin-modifying agents permit human hematopoietic stem cells to undergo multiple cell divisions while retaining their repopulating potential. *Blood* **109**, (2007).
407. Holmes, T. *et al.* Glycogen Synthase Kinase-3 β Inhibition Preserves Hematopoietic Stem Cell Activity and Inhibits Leukemic Cell Growth. *Stem Cells* **26**, 1288–1297 (2008).
408. Boitano, A. E. *et al.* Aryl Hydrocarbon Receptor Antagonists Promote the Expansion of Human Hematopoietic Stem Cells. *Science (80-.)*. **329**, (2010).
409. North, T. E. *et al.* Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. *Nature* **447**, 1007–1011 (2007).
410. Nishino, T. *et al.* Ex vivo expansion of human hematopoietic stem cells by a small-molecule agonist of c-MPL. *Exp. Hematol.* **37**, 1364–1377.e4 (2009).
411. Delaney, C., Varnum-Finney, B., Aoyama, K., Brashem-Stein, C. & Bernstein, I. D. Dose-dependent effects of the Notch ligand Delta1 on ex vivo differentiation and in vivo marrow repopulating ability of cord blood cells. *Blood* **106**, 2693–9 (2005).
412. Willert, K. *et al.* Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature* **423**, 448–452 (2003).
413. Bhardwaj, G. *et al.* Sonic hedgehog induces the proliferation of primitive human hematopoietic cells via BMP regulation. *Nat. Immunol.* **2**, 172–180 (2001).
414. Bhatia, M. *et al.* Bone morphogenetic proteins regulate the developmental program of human hematopoietic stem cells. *J. Exp. Med.* **189**, 1139–48 (1999).
415. Kirouac, D. C. *et al.* Dynamic interaction networks in a hierarchically organized tissue. *Mol. Syst. Biol.* **6**, 417 (2010).
416. Csaszar, E. *et al.* Rapid Expansion of Human Hematopoietic Stem Cells by Automated Control of Inhibitory Feedback Signaling. *Cell Stem Cell* **10**, 218–229 (2012).
417. Ieyasu, A. *et al.* An All-Recombinant Protein-Based Culture System Specifically Identifies Hematopoietic Stem Cell Maintenance Factors. *Stem Cell Reports* **8**, 500–508 (2017).

418. Oedekoven, C. A. Extrinsic regulation of fate choice in mouse haematopoietic stem cells. (University of Cambridge, 2019).
419. Wilkinson, A. C., Ishida, R., Nakauchi, H. & Yamazaki, S. Long-term ex vivo expansion of mouse hematopoietic stem cells. *Nat. Protoc.* **15**, 628–648 (2020).
420. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–40 (2010).
421. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–97 (2012).
422. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
423. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).
424. Aran, D. *et al.* Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
425. Jojic, V. *et al.* Identification of transcriptional regulators in the mouse immune system. *Nat. Immunol.* **14**, 633–643 (2013).
426. Shay, T. & Kang, J. Immunological Genome Project and systems immunology. *Trends Immunol.* **34**, 602–9 (2013).
427. ImmGen at 15. *Nat. Immunol.* **21**, 700–703 (2020).
428. Engel, I. *et al.* Innate-like functions of natural killer T cell subsets result from highly divergent gene programs. *Nat. Immunol.* **17**, 728–739 (2016).
429. Yu, G. & He, Q.-Y. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst.* **12**, 477–9 (2016).
430. Mildner, A. *et al.* Genomic Characterization of Murine Monocytes Reveals C/EBP β Transcription Factor Dependence of Ly6C⁺ Cells. *Immunity* **46**, 849-862.e7 (2017).
431. Winefield, R. D., Williams, T. D. & Himes, R. H. A label-free mass spectrometry method for the quantification of protein isotopes. *Anal. Biochem.* **395**, 217–223 (2009).
432. Di Palma, S. *et al.* Highly Sensitive Proteome Analysis of FACS-Sorted Adult Colon Stem Cells. *J. Proteome Res.* **10**, 3814–3819 (2011).
433. Wang, Y. X., Wang, D. Y., Guo, Y. C. & Guo, J. Zyxin: A mechanotransducer to regulate gene expression. *Eur. Rev. Med. Pharmacol. Sci.* **23**, 413–425 (2019).
434. Pritchard, K. & Moody, C. J. Caldesmon: A calmodulin - Binding actin - Regulatory protein. *Cell Calcium* **7**, 309–327 (1986).
435. Gattazzo, F., Urciuolo, A. & Bonaldo, P. Extracellular matrix: A dynamic microenvironment for stem cell niche. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 2506–2519 (2014).
436. Himburg, H. A. *et al.* Pleiotrophin regulates the retention and self-renewal of hematopoietic stem cells in the bone marrow vascular niche. *Cell Rep.* **2**, 964–75 (2012).
437. Himburg, H. A. *et al.* Pleiotrophin mediates hematopoietic regeneration via activation of RAS. *J. Clin. Invest.* **124**, 4753–8 (2014).
438. Huynh, H. *et al.* IGF binding protein 2 supports the survival and cycling of hematopoietic stem cells. *Blood* **118**, 3236–3243 (2011).
439. Ghiaur, G. *et al.* Inhibition of RhoA GTPase activity enhances hematopoietic stem and progenitor cell proliferation and engraftment. *Blood* **108**, 2087–2094 (2006).
440. Cancelas, J. A. & Williams, D. A. Rho GTPases in hematopoietic stem cell functions. *Curr. Opin. Hematol.* **16**, 249–254 (2009).
441. Liu, C. *et al.* Lipoprotein lipase regulates hematopoietic stem progenitor cell maintenance through DHA supply. *Nat. Commun.* **9**, 1310 (2018).

442. Grigoryan, A. *et al.* LaminA/C regulates epigenetic and chromatin architecture changes upon aging of hematopoietic stem cells. *Genome Biol.* **19**, (2018).
443. Pekovic, V. & Hutchison, C. J. Adult stem cell maintenance and tissue regeneration in the ageing context: the role for A-type lamins as intrinsic modulators of ageing in adult stem cells and their niches. *J. Anat.* **213**, 5–25 (2008).
444. Ma, H. *et al.* The Sox4/Tcf7l1 axis promotes progression of BCR-ABL-positive acute lymphoblastic leukemia. *Haematologica* **99**, 1591–1598 (2014).
445. Ohlsson, E., Schuster, M. B., Hasemann, M. & Porse, B. T. The multifaceted functions of C/EBP α in normal and malignant haematopoiesis. *Leukemia* **30**, 767–775 (2016).
446. Ziegler, S. F., Ramsdell, F. & Alderson, M. R. The activation antigen CD69. *Stem Cells* **12**, 456–465 (1994).
447. Saito, Y. *et al.* Maintenance of the hematopoietic stem cell pool in bone marrow niches by EVI1-regulated GPR56. *Leukemia* **27**, 1637–1649 (2013).
448. Lehnertz, B. *et al.* HLF Expression Defines the Human Haematopoietic Stem Cell State. *bioRxiv* (2020).
449. Wang, X. *et al.* TGF- β 1 Negatively Regulates the Number and Function of Hematopoietic Stem Cells. *Stem cell reports* **11**, 274–287 (2018).
450. Blank, U. & Karlsson, S. TGF- β signaling in the control of hematopoietic stem cells. *Blood* **125**, 3542–3550 (2015).
451. Enomoto, H. *et al.* Hepatoma-derived growth factor is highly expressed in developing liver and promotes fetal hepatocyte proliferation. *Hepatology* **36**, 1519–1527 (2002).
452. Haruyama, M., Yamaichi, K., Niwa, A., Saito, M. K. & Nakahata, T. Hepatoma-Derived Growth Factor Is a Novel Factor to Promote the Proliferation of Hematopoietic Stem Cells. *Blood* **128**, 1471–1471 (2016).
453. Cruikshank, W. W., Kornfeld, H. & Center, D. M. Interleukin-16. *J. Leukoc. Biol.* **67**, 757–766 (2000).
454. Wang, L. *et al.* Myeloid-Derived Growth Factor Promotes Intestinal Glucagon-Like Peptide-1 Production in Male Mice With Type 2 Diabetes. *Endocrinology* **161**, (2020).
455. Oostendorp, R. A. J. *et al.* Embryonal subregion-derived stromal cell lines from novel temperature-sensitive SV40 T antigen transgenic mice support hematopoiesis. *J. Cell Sci.* **115**, 2099 LP – 2108 (2002).
456. Rashidi, N. M. & Lo Celso, C. Flying back to the nest: Intravital microscopy reveals how the niche can induce stemness. *Intravital* **3**, e29653 (2014).
457. Mangolini, M. *et al.* Notch2 controls non-autonomous Wnt-signalling in chronic lymphocytic leukaemia. *Nat. Commun.* **9**, 1–17 (2018).
458. Susek, K. H. *et al.* Bone marrow laminins influence hematopoietic stem and progenitor cell cycling and homing to the bone marrow. *Matrix Biol.* **67**, 47–62 (2018).
459. Huang, S.-Y. *et al.* Higher Decorin Levels in Bone Marrow Plasma Are Associated with Superior Treatment Response to Novel Agent-Based Induction in Patients with Newly Diagnosed Myeloma - A Retrospective Study. *PLoS One* **10**, e0137552 (2015).
460. Mattanovich, D. *et al.* Recombinant Protein Production in Yeasts. in *Recombinant Gene Expression* (ed. Lorence, A.) 329–358 (Humana Press, 2012).
461. Nogami, W. *et al.* A Novel, Small Non-Peptidyl Butyramide Activates Human Thrombopoietin Receptor and Promotes Megakaryopoiesis. *Blood* **110**, 2201–2201 (2007).
462. Park, S. *et al.* Fgd5 is a Rac1-specific Rho GEF that is selectively inhibited by aurintricarboxylic acid. *Small GTPases* 1–14 (2019).
463. Farhan, M. A., Azad, A. K., Touret, N. & Murray, A. G. FGD5 Regulates VEGF Receptor-2 Coupling to PI3 Kinase and Receptor Recycling. *Arterioscler. Thromb. Vasc. Biol.* **37**, 2301–2310 (2017).
464. Oguro, H., Ding, L. & Morrison, S. J. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* **13**, (2013).
465. Moignard, V. *et al.* Decoding the regulatory network of early blood development from single-cell gene expression measurements. *Nat. Biotechnol.* **33**, 269–76 (2015).

References

466. Kent, D. G. *et al.* Self-Renewal of Single Mouse Hematopoietic Stem Cells Is Reduced by JAK2V617F Without Compromising Progenitor Cell Expansion. *PLoS Biol.* **11**, e1001576 (2013).
467. Klose, M., Florian, M. C., Gerbault, A., Geiger, H. & Glauche, I. Hematopoietic Stem Cell Dynamics Are Regulated by Progenitor Demand: Lessons from a Quantitative Modeling Approach. *Stem Cells* **37**, 948–957 (2019).
468. Bai, T. *et al.* Expansion of primitive human hematopoietic stem cells by culture in a zwitterionic hydrogel. *Nat. Med.* **25**, 1566–1575 (2019).