Human fetal B-lymphopoiesis and acute lymphoblastic leukaemia: the impact of trisomy 21

Sorcha O'Byrne

Linacre College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2019

Dedication

For Dad

Abstract

Background: Children with Down Syndrome (DS), caused by an extra copy of chromosome 21 (trisomy 21, T21) are at increased risk of developing acute myeloid leukaemia (ML-DS) and B lineage acute lymphoblastic leukaemia (ALL). This points to specific effects of T21 on haematopoietic stem and progenitor cell (HSPC) biology. In ML-DS, all cases originate in fetal cells and T21, together with a *GATA1* mutation is essential for disease progression. By contrast, in DS-ALL, where the most frequent molecular abnormalities affect JAK2 signalling and *CRLF2*, clinical signs of leukaemia develop after the first year of life and there is so far no direct evidence for a fetal origin. However, previous work from the Roberts lab showed that T21 causes perturbation of haematopoiesis in second trimester fetal liver (FL) and, as well as increased erythro-megakaryopoiesis, B-lymphopoiesis was also severely impaired. Furthermore, FL HSC show a reduced ability to generate B cells *in vitro* despite increased HSC numbers. Preliminary data suggests that B-lymphopoiesis may also be reduced in T21 fetal bone marrow (BM). However, the extent of the perturbations in fetal lymphopoiesis in BM, their molecular basis and the link to DS-ALL remain unclear.

Aim: The aim of this project was to investigate the cellular and molecular mechanisms responsible for perturbed fetal B-lymphoid development in T21 and how this may contribute to ALL by addressing the following questions: 1. What is the cellular hierarchy and molecular basis for B-lymphoid development in normal fetal BM? 2. What is the impact of T21 on fetal BM B progenitor development? 3. What is the impact of T21 on the transcriptome of ALL blasts?

Results: 1. *Normal fetal B-lymphopoiesis:* Flow cytometry of FL and fetal BM showed that fetal BM is the main site of B-lymphopoiesis during the second trimester and that there are fetal specific lymphoid populations: Early Lymphoid Progenitors (ELP) and PreProB progenitors. Functional (*in vitro* cultures and *in vivo* xenografts) demonstrated that PreProB progenitors are the first committed B progenitor. Molecular analysis (RNA-sequencing, single cell RT-qPCR, ATAC-sequencing) placed PreProB progenitors downstream of ELP and upstream of ProB progenitors. PreProB progenitors are also especially proliferative and share immunophenotypic, molecular and IgH rearrangement patterns with infant ALL (iALL).

2. *Fetal B-lymphopoiesis in T21 fetal BM:* Although the frequency of immunophenotypic HSC was higher in T21 fetal BM and LMPP frequency was normal, ELP and PreProB progenitors were virtually absent, indicating an even more profound defect in B-lymphopoiesis than previously observed in FL. Functional analysis revealed decreased B-lymphoid potential from T21 HSC suggesting a cell intrinsic defect in B-lymphopoiesis at the HSC level. RNA-Sequencing analysis of normal and T21 fetal BM HSPC showed enrichment for gene signatures associated with inflammatory pathways including IFNα signalling, suggesting a potential role for extrinsic, microenvironment-driven factors. Novel co-cultures using normal fetal HSPC and normal or T21 fetal BM mesenchymal stromal cells (MSC) demonstrated that T21 MSC have reduced capacity to support fetal B-lymphopoiesis. Also, analysis of the MSC transcriptome pointed towards a role for inflammation in the perturbation in fetal B-lymphopoiesis.

3. *Transcriptomic analysis of cytogenetically matched paediatric non-DS-ALL and DS-ALL blasts:* Analysis of RNA-sequencing data from non-DS-ALL blasts, DS-ALL blasts and non-DS-ALL blasts with acquired T21 suggested that the ALL-associated cytogenetic abnormality was the main driver promoting clustering of samples by PCA. The impact of acquired T21 was distinct from that of constitutional T21, which also pointed towards a role for the trisomic microenvironment in DS-ALL because gene sets for inflammatory signalling pathways were enriched exclusively in DS-ALL blasts and HSC from DS-ALL patient BM. In addition, the increased frequency of immunophenotypic HSC observed in T21 fetal BM was found to persist in post-natal DS-ALL BM.

Conclusions: Fetal specific B-lymphoid development pathways in human BM have been identified for the first time paving the way for further studies to identify and characterise *in utero* initiation of infant and paediatric ALL. In T21, fetal BM B-lymphopoiesis is severely perturbed and specifically there appears to be a block in the normal fetal programme via ELP and PreProB progenitors. Studies of HSPC and MSC, the major cellular component of the microenvironment, indicate that both cell intrinsic and cell extrinsic mechanisms play a role, possibly through altered crosstalk between HSPC and their environment; and an increased inflammatory response that inhibits fetal B-lymphopoiesis. The inflammatory signature observed in T21 fetal life is also observed in DS-ALL and so together, these data indicate that perturbation of B-lymphopoiesis by T21 persists beyond the FL in the fetal BM and possibly in post-natal life.

Acknowledgements

Firstly, I would like to thank my co-supervisors Prof. Irene Roberts and Dr. Anindita Roy for their incredible dedication and patience in supporting and guiding me both academically and personally through this project. They are both an inspiration.

Secondly, I would like to thank all members of the extended Roberts and Roy lab past and present for their help, guidance and numerous coffees and lunches. Thank you Natalina, Nick, Siobhan, Catherine, Jana, Hannah, Ella, Lucy, Peng, Eleni, Kerol and Gemma. Also, thank you to previous members from Imperial College London and the Karadimitris Lab: Valentina, Katerina, Deena and Prof. Anastasios Karadimitris.

I am also very grateful for the support and guidance from my thesis committee: Prof. Adam Mead and Prof. Tom Milne. Thank you to to members past and present in the Mead and Milne labs for their help and expertise. A special thanks to Joe Harman and Nick Crump for their bioinformatics advice and tips. I also want to thank Dr. Mariolina Salio for her guidance at my confirmation of status and numerous helpful conversations in flow cytometry.

Many thanks to all those at the HDBR at UCL Institute of Child Health and Newcastle University. Thank you also to our collaborators, Sarah Inglott and Gary Wright at Great Ormond Street Hospital. Thank you also to childhood leukaemia Cell Bank for the ALL samples.

Many thanks to all members of the Molecular Haematology Unit at the WIMM. Specifically thank you to all those in MHU 'dirty' TC for all the chats, laughs and coordination of experiments. Thank you to the friends that I have made along the way: Laura, Julia and Vincent.

Thank you to all those in other units in the WIMM for their patience and kindness in teaching me. Thank you to Paul, Kevin, Sally and Craig in flow cytometry. Thank you to the BMS staff for their expertise with the mice. Thank you to all those in the WIMM Centre for Computational Biology. I am especially grateful for those who encouraged me to pursue and persevere with learning the bioinformatics. Thank you Steve, Nicki, Ewan, Simon, David, Charlie, Sebastian, Jelena, Hashem and the OBDS cohort of summer 2018.

I am also very grateful for all the funding received that enabled me to pursue this project and do the experiments. Thank you to the Department of Paediatrics, Bloodwise, the Alexander Thatte Fund, Linacre College, the European Haematology Association and the British Society of Immunology.

Last but by no means least, thank you to my family and friends for your unwavering support over the last 4 years. Thank you Mum, Orlaith and Aoife for supporting and believing in me; thank you Dad for inspiring me; and thank you to all my friends particularly Andreea, Dragana, Emma and Alysia for all the fun memories in Oxford and further afield.

Publications and Presentations

Publications

O'Byrne, S., Elliott, N., Buck, G., Rice, S., Knapp, D.J.H.F., Garnett, C., Godfrey, L., Crump, N., Wright, G., Inglott, S., Povinelli, B., Agraz-Doblas, A., Bueno, C., Varla, I., Fordham, N., Bennett, P., Karadimitris, A., Mead, A., Ancliff, P., Vyas, P., Menendez, P., Milne, T., Roberts, I., Roy, A., (2019). Discovery of a CD10-negative B-progenitor in human fetal life identifies unique ontogeny-related developmental programs. Blood 134, 1059–1071.

This paper is the result of my time spent working as a research assistant before starting my DPhil: Sng, C.C.T.*, **O'Byrne, S.***, Prigozhin, D.M.* Bauer, M.R.*, Harvey, J.C., Ruhle, M., Challis, B.G., Lear, S., Roberts, L.D., Workman, S., Janowitz, T., Magiera, L., Doffinger, R., Buckland, M.S., Jodrell, D.J., Semple, R.K., Wilson, T.J., Modis, Y., Thaventhiran, J.E.D., (2018). A Type III Complement Factor D Deficiency: Structural insights for inhibition of the alternative pathway. J. Allergy Clin. Immunol.

Godfrey, L., Crump, T., **O'Byrne, S.**, Lau, I., Rice, S., Harman, J., Elliott, N., Buck, G., Connor, C., Thorne, R., Knapp, D., Heidenreich, O., Vyas, P., Menendez, P., Inglott, S., Ancliff, P., Geng, H., Roberts, I., Roy, A., Milne, T., (2019). H3K79me2/3 controls enhancer promoter interactions and activation of the pan-cancer stem cell marker PROM1/CD133 in MLL-AF4 leukemia cells. Leukemia. Under review.

Louka, E., Povinelli, B., Rodrigues-Meira, A., Buck, G., Hamblin, A., Booth, C., Sousos, N., Roy, A., Elliott, N., Iskander, D., de la Fuente, J., Fordham, N., **O'Byrne, S.**, Inglott, S., Rao, A., Roberts, I., Mead, A., (2018). Heterogenous disease-propagating stem cells in juvenile myelomonocytic leukaemia. J. Ex. Med. Invited resubmission.

Roy, A., Bystry, V., Bohn, G., Goudevenou, K., Reigl, T., Papaioannou, M., Krejci, A., **O'Byrne, S.**, Chaidos, A., Grioni, A., et al. (2017). High resolution IgH repertoire analysis reveals fetal liver as the likely origin of life-long, innate B lymphopoiesis in humans. Clin. Immunol. 183, 8–16.

Kerry, J., Godfrey, L., Repapi, E., Tapia, M., Blackledge, N.P., Ma, H., Ballabio, E., **O'Byrne, S.**, Ponthan, F., Heidenreich, O., et al. (2017). MLL-AF4 Spreading Identifies Binding Sites that Are Distinct from Super-Enhancers and that Govern Sensitivity to DOT1L Inhibition in Leukemia. Cell Rep 18, 482–495.

Conferences and presentations

O'Byrne, S., Elliott, N., Buck, G., Rice, S., O'Connor, D., Oswald, J., Fuchs, H., Labbett, E-M., Roy, A., Roberts, I. (2019) Trisomy 21 Driven Pro-Inflammatory Signalling in Fetal Bone Marrow May Play a Role in Perturbed B-Lymphopoiesis and Acute Lymphoblastic Leukemia of Down Syndrome. Blood 134, 1206–1206.

O'Byrne, S. Discovery of a CD10 negative fetal B progenitor: implications for childhood leukaemia. September 2019, CLRUK Forum, Newcastle.

Godfrey, L.*, **O'Byrne, S.***, Crump, N., Lau, I., Thorne, R., Elliott, N., Buck, G., Rice, S., Connor, C., Thomas, R., Cheng, D., Ancliff, P., Inglott, S., Roberts, I., Roy, A., Milne, T. (2019). MLL-AF4 causes aberrant upregulation of PROM1 (CD133) in acute lymphoblastic leukemia by controlling enhancer-promoter interactions. EHA 2019.

Roy, A., Psaila, B., Thongjuea, S., Ashley, N., Buck, G., **O'Byrne, S.**, Elliott, N., Louka, E., Fordham, N., Hua, P., Povinelli, B., Heuston, E., Iskander, D., Karadimitris, A., Bodine, D., Mead, A., Roberts, I. (2018). Single cell profiling reveals key differences in the cellular architecture of human haematopoietic stem and progenitor cells throughout fetal and adult life. EHA 2018

Elliott, N., O'Connor, D., Roy, A., Labbett, E., **O'Byrne, S.**, Rice, S., Fordham, N., Liu, B., Cowan, G., Karadimitris, A., Vyas, P., Roberts, I. (2018). Perturbation of fetal liver stromal cell function by trisomy 21 promotes increased erythro-megakaryopoiesis by fetal HSPC. EHA 2018.

O'Byrne, S., Knapp, D., Elliott, N., Rice, S., Buck, G., Wright, G., Inglott, S., Povinelli, B., Agraz-Doblas, A., Varela, I., Mead, A., Ancliff, P., Milne, T., Menendez, P., Roberts, I., Roy, A. The fetal target cell for MLL rearranged leukaemia: implications for the origins of infant ALL. EHA-SWG Scientific meeting on ALL, April 2018, Barcelona.

Louka, E., Povinelli, B., Rodriguez-Meira, A., Buck, G., Hamblin, A., Booth, C., Roy, A., Elliott, N., Iskander, D., de la Fuente, J., Fordham, N., **O'Byrne, S.**, Inglott, S., Rao, A., Roberts, I., Mead, A. (2017). Molecular and Functional Characterization of Disease-Propagating Stem Cells in Juvenile Myelomonocytic Leukemia. Blood 130, 379–379.

O'Byrne, S., Elliott, N., Rice, S., Fordham, N., O'Connor, D., Taylor, S., Roy, A., Roberts, I. The impact of trisomy 21 on fetal B-lymphopoiesis. WIMM day 2018, Oxford.

O'Byrne, S., Elliott, N., Buck, G., Liu, B., Povinelli, B., Fordham, N., Louka, E., Bartolovic, K., Karadimitris, A., Mead, A., Roberts, I., Roy, A. (2017). Re-ordering the B cell development hierarchy in human fetal bone marrow: characterisation of a novel human fetal B progenitor. EHA 2017.

O'Byrne, S., Elliott, N., Buck, G., Liu, B., Povinelli, B., Fordham, N., Louka, E., Bartolovic, K., Karadimitris, A., Mead, A., Roberts, I., Roy, A. Reordering the B cell development hierarchy in human fetal bone marrow: characterisation of a novel human fetal B progenitor. WIMM day 2017, Oxford.

Roy, A., Bystry, V., Bohn, G., Goudevenou, K., Reigl, T., Papaioannou, M., Krejci, A., **O'Byrne, S.**, Chaidos, A., Grioni, A., et al. (2016). High Resolution Igh Repertoire Analysis Reveals the Human Fetal Liver As the Origin of Life-Long, Innate B Lymphopoiesis. Blood 128, 127–127.

List of abbreviations

α-MEM . . . . α-Minimum Essential Medium

ABM . . . . . Adult Bone Marrow

AGM . . . . . Aorta-gonad-mesenephros

ALL . . . . . Acute Lymphoblastic Leukaemia

ANOVA . . . . Analysis of Variance

aT21 . . . . . acquired Trisomy 21

ATAC-Seq . . Assay for transposase-accessible chromatin-sequencing

BCP-ALL . . B Cell Precursor Acute Lymphoblastic Leukaemia

BM . . . . . Bone Marrow

CB . . . . . Cord Blood

ChIP-Seq . . . Chromatin Immuno-precipitation Sequencing

CLP . . . . . Common Lymphoid Progenitor

CMP . . . . . Common Myeloid Progenitor

CRLF2r *CRLF2*-rearranged

CyTOF . . . . Mass Cytometry by Fluidigm

DE . . . . . Differentially Expressed

DEG . . . . . Differentially Expressed Genes

DMEM . . . . Dulbecco's Modified Eagle Medium

DS . . . . . Down Syndrome

DS-ALL . . . . Down Syndrome Acute Lymphoblastic Leukaemia

ELP . . . . . Early Lymphoid Progenitor

FACS . . . . . Fluorescence Activated Cell Sorting

FBM . . . . . Fetal Bone Marrow

FBS . . . . . Fetal Bovine Serum

FDR . . . . . False Discovery Rate

FISH . . . . . Fluorescence *in situ* hybridisation

FL . . . . . Fetal Liver

GMP . . . . . Granulocyte/Macrophage Progenitor

GO . . . . . Gene Ontologies

GSEA . . . . . Gene Set Enrichment Analysis

HDBR . . . . . Human Developmental Biology Resource

HI . . . . . Heat Inactivated

HSC . . . . . Haematopoietic Stem Cell

HSPC . . . . . Haematopoietic Stem and Progenitor Cells

iALL . . . . . infant ALL

iPSC . . . . . induced Pluripotent Stem Cell

IFN . . . . . Interferon

LMPP . . . . . Lymphoid-primed Multipotent Progenitor

MEP . . . . . Megakaryocyte/Erythroid Progenitor

miR . . . . . microRNA/miRNA

MLP . . . . . Multi-Lymphoid Progenitor

MNC . . . . . Mononuclear Cells

MPP . . . . . Multipotent Progenitor

NSG . . . . . NOD.Cg-Prkdc^{scid}Il2rg^{tmlWjl}/SzJ

ML-DS . . . . Myeloid Leukaemia of Down Syndrome

MSC . . . . . Mesenchymal Stromal Cell

NDS . . . . . Non-Down Syndrome

NES . . . . . Normalised Enrichment Score

NM . . . . . Normal (non-T21)

NPV . . . . . Nominal p-value

PAGE . . . . . Poly-acrylamide Gel Electrophoresis

PBMCs . . . . Peripheral Blood Mononuclear Cells

PBS . . . . . Phosphate Buffered Saline

PCA . . . . . Principal Component Analysis

PCR . . . . . Polymerase Chain Reaction

RISC . . . . . RNA-induced silencing complex

RMA . . . . . Robust Multi-array Average

RNA-Seq . . . RNA-Sequencing

RT-qPCR . . . Reverse Transcription quantitative-Polymerase Chain Reaction

SEM . . . . . Standard Error of Mean

SNP . . . . . Single Nucleotide Polymorphism

T21 . . . . . Trisomy 21

TAM . . . . . Transient Abnormal Myelopoiesis

TPM . . . . . Transcripts Per Million

VSN . . . . . Variance Stabilising Normalisation

List of Figures

Contents

Chapter 1

Introduction

1.1 Fetal haematopoiesis in the mouse and human

1.1.1 HSC emergence

In mammals, there are three waves of haematopoiesis that begin early in embryogenesis. Most of our understanding of these early waves of haematopoiesis comes from murine studies. In mice, the first site where blood cells develop is the yolk sac [Palis et al., 1999] [Costa et al., 2012] [Dzierzak and Bigas, 2018]. Here there are two waves of development: primitive erythropoiesis followed by a wave of erythromyeloid expansion [Palis et al., 1999] [McGrath et al., 2015a] [McGrath et al., 2015b]. In addition to differentiation of primitive erythroid and myeloid cells, there is increasing evidence to suggest that primitive B-1 cells arise in the yolk sac [Yoshimoto et al., 2011] [Yoshimoto, 2015]. "Definitive" adult type haematopoietic stem cells (HSCs) arise after embryonic day 10.5 (E10.5) from the dorsal aorta of the aorta-gonad-mesonephros (AGM) region [Müller et al., 1994] [Medvinsky and Dzierzak, 1996] [de Bruijn et al., 2000] [de Bruijn et al., 2002]. These "definitive" HSCs isolated from the E10-11 AGM share immunophenotypic and functional (long-term re-populating) properties with adult bone marrow (BM) HSCs [Müller et al., 1994] [Medvinsky and Dzierzak, 1996] [Sanchez et al., 1996] [de Bruijn et al., 2002]. After emergence, HSCs migrate to and populate the fetal liver (FL) which remains the main site of haematopoiesis in the mouse until around birth when the BM takes over [Coskun et al., 2014]. Murine HSC emergence is extensively reviewed most recently here: [Dzierzak and Bigas, 2018].

Similarly, in humans, primitive erythro-myelopoiesis occurs in the yolk sac [Huyhn et al., 1995] [Tavian et al., 1999]. Definitive HSCs arise from the AGM after 4 post-conception weeks (pcw)[Tavian et al., 1996] [Tavian et al., 1999][Tavian et al., 2001],

[Oberlin et al., 2002] [Ivanovs et al., 2011] [Easterbrook et al., 2019] and colonise the FL from 5 pcw onwards [Tavian et al., 1999]. By 10-12 pcw, haematopoietic cells are detectable in long bones but the exact timing of the onset of HSC activity in human fetal BM remains undetermined [Charbord et al., 1996] [Sinka et al., 2012] [Ivanovs et al., 2017]. The BM becomes the main site of haematopoiesis by the time of birth and this remains the case for the rest of post-natal life (figure 1.1). Human embryonic HSC emergence is reviewed extensively and most recently here: [Ivanovs et al., 2017].

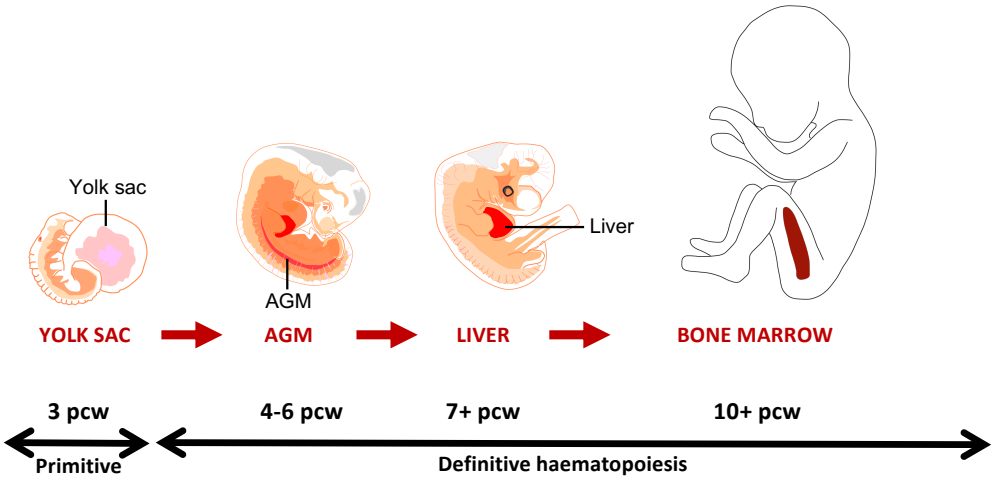


Figure 1.1: The main sites of haematopoiesis during ontogeny: primitive erythro-myelopoiesis occurs in the yolk sac and definitive HSC emerge from the AGM to colonise the FL. FL HSPC eventually migrate to the BM which will be the main site of haematopoiesis around birth. AGM: aorta-gonad-mesonephros. Illustration adapted from [Ivanovs et al., 2017] with permission from publisher (Development, the Company of Biologists Ltd.)

1.1.2 The haematopoietic hierarchy

Traditionally haematopoiesis has been modelled as a tree-like hierarchical structure with HSCs at the apex that divide and differentiate into restricted progenitors that eventually give rise to all the mature cell types of the haematopoietic and immune system [Hofer and Rodewald, 2018] [Eaves, 2015]. HSCs are defined as having multi-lineage potential, the ability to self renew and the ability to re-populate in immunodeficient mice. This model is based on decades of experiments involving the prospective immunophenotypic classification, functional and molecular characterisation of the BM in both mice and humans (figure 1.2(a)). In this active field of research the model of the haematopoietic tree is constantly being refined. Most recently the

existence of a Common Myeloid Progenitor (CMP) and Granulocyte/Macrophage progenitors (GMP) have been called into question [Drissen et al., 2016] [Drissen et al., 2019]. Many of these refinements to the haematopoietic hierarchy are as a direct result of advances in single cell technologies. In the era of single cell transcriptomics, several differing models of haematopoiesis have been proposed owing to the heterogeneity discovered in each bulk population [Notta et al., 2015] [Cvejic, 2015] [Haas et al., 2018] [Carrelha et al., 2018] [Karamitros et al., 2018] [Laurenti and Göttgens, 2018] [Jacobsen and Nerlov, 2019]. Nevertheless, the haematopoietic hierarchy tree can be used to illustrate the gradual loss of pluripotency into a mature cell type that is the basis of blood cell differentiation, providing a "snapshot" of the continuum of differentiation revealed by single cell studies. This is perhaps best illustrated in a figure from [Laurenti and Göttgens, 2018] (figure 1.2(b)).

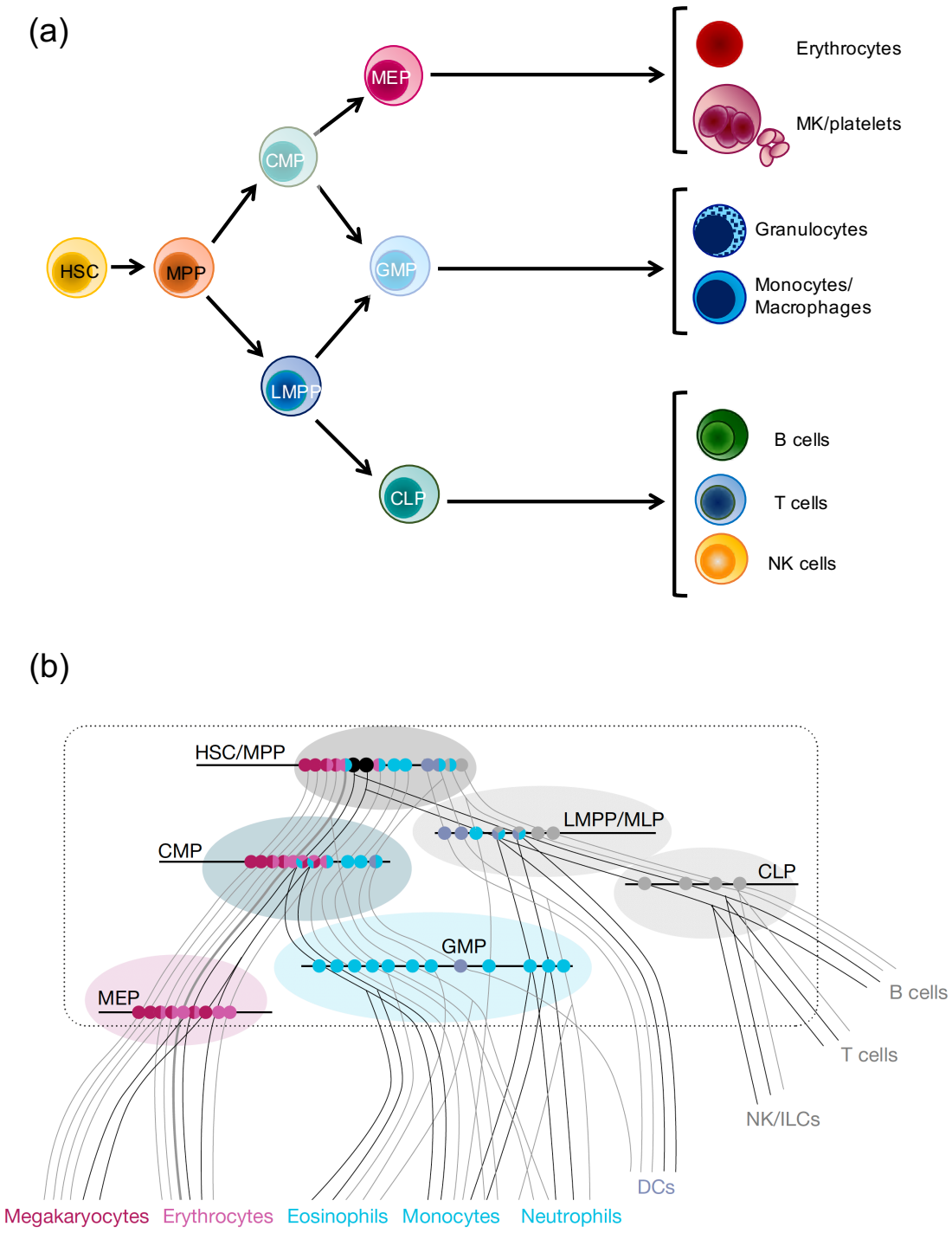


Figure 1.2: Adult haematopoietic hierarchy defined by functional and molecular analysis in murine and human studies. (a) a simplified illustration of the haematopoietic hierarchy with CMP and GMP faded to represent recent studies calling to question their existence in humans [Drissen et al., 2016] [Drissen et al., 2019]. (b) Adult haematopoietic hierarchy illustrating the heterogeneity within HSPC pools; adapted by permission from Springer Nature: Springer, Nature, From haematopoietic stem cells to complex differentiation landscapes. Laurenti and Gottgens. Copyright (2018) [Laurenti and Göttgens, 2018]. HSC, haematopoietic stem cell; MPP, multipotent progenitor; LMPP, lymphoid-primed multipotent progenitor; MLP, multi-lymphoid progenitor; CMP, common myeloid progenitor; GMP, granulocyte-macrophage progenitor; MEP, megakaryocyte-erythroid progenitor; CLP, common lymphoid progenitor.

1.1.3 Murine adult and fetal B-lymphopoiesis

In mice, adult B-lymphopoiesis starts with HSCs that differentiate into multipotent progenitors (MPP); then lymphoid-primed multipotent progenitors (LMPP); followed by common lymphoid progenitors (CLP) before finally committing to the respective lymphoid lineages: B, T and NK cell. This model of B-lymphopoiesis has been deduced after decades of studies involving prospective isolation of progenitor populations followed by molecular and functional analyses (reviewed here: [Sigvardsson, 2018]). In parallel, others have focussed on resolving the stages of B-lymphopoiesis after B-lymphoid commitment. Much of this work was lead by Richard Hardy who postulated a schema of B-lymphopoiesis that divided B progenitors into five fractions according to their cell surface markers, heavy chain rearrangement status and growth dependence on IL7 [Hardy et al., 1991]. These "Hardy" fractions are still used to decipher B progenitors in the murine adult system [Jensen et al., 2018] [Sigvardsson, 2018].

Murine adult and fetal B-lymphopoiesis is distinct [Hardy and Hayakawa, 1991]. Since Hardy and Hayakawa showed that B cells in fetal tissues are mainly CD5+ and in adult BM B cells are mainly CD5-, these CD5+ B cells have been identified as natural antibody secreting B-1 cells that play a distinct role in innate immunity. In murine fetal life, these B-1 cells arise in the yolk sac independent of HSCs in a primitive wave [Montecino-Rodriguez et al., 2016] [Ghosn et al., 2019]. Then, in fetal and adult life, in parallel with B-2 cells, there are two further waves of B-1 development that are HSC dependent [Montecino-Rodriguez et al., 2016] [Ghosn et al., 2019]. But, it is this first wave that is thought to contribute to the increased frequencies of B-1 cells reported by Hardy and Hayakawa [Hardy and Hayakawa, 1991]. In addition to the differences in B-1 cell production through ontogeny, HSC-dependent differentiation into B-2 cells develops in two waves: a fetal and adult wave [Montecino-Rodriguez et al., 2016]. While most studies of fetal and adult B-lymphopoiesis in mice concentrate on the distinctions between B-1 and B-2 developmental origins, there are others who have suggested that HSC- dependent B-2 cell differentiation is associated specifically with *Il7ra* expression [Berthault et al., 2017]

1.1.4 Human adult and fetal B-lymphopoiesis

Differences between human adult and fetal B-lymphopoiesis have also been described within the wider context of haematopoiesis as a whole [Notta et al., 2016]. Since the existence of B-1 B cells in human fetal life is debated [Griffin et al., 2011b] [Descatoire et al., 2011][Griffin et al., 2011a] [Perez-Andres et al., 2011] [Bueno et al., 2016] [Baumgarth, 2017], and given that we could not find them in FL or fetal BM [Roy et al., 2017], from herein I will use the term "B cells" to define CD34-CD19+ B cells without discriminating between B-1 and B-2 cells.

Species differences in immunophenotypic marker expression between mice and humans and the fact that human B cells are notoriously difficult to culture *in vitro* has hindered the functional dissection of human B-lymphopoiesis in the same detail as murine B-lymphopoiesis [Ichii et al., 2014]. Nevertheless, the B-lymphoid hierarchy in human post-natal tissues is relatively well characterised with many parallels in the murine system [Hystad et al., 2007] [Bendall et al., 2014] [Ichii et al., 2014] [Good et al., 2018] [Pellin et al., 2019]. It is accepted that HSCs differentiate into progenitors that first lose their megakaryocyte/erythroid potential and then myeloid potential to become a CLP that can differentiate into only the lymphoid lineages: B, T and NK. There are two immunophentypic definitions of CLP: a Lin-CD34+CD38-CD10+ progenitor described in adult BM [Galy et al., 1995] and a CD34+CD38-CD7+ progenitor described in cord blood (CB) [Hao et al., 2001]. Whether lymphoid commitment occurs prior to the expression of CD10 remains a subject of debate but differences in tissues studied (adult BM and CB) almost certainly play a role in the functional output of these immunophenotypically defined populations [Kohn et al., 2012]. Furthermore, both of these immunophenotypic CLP were later shown to be heterogeneous populations with bias towards T/NK potential in the case of CD7+ CLP and B potential in the case of CD10+ CLP [Haddad et al., 2004] [Hoebeke et al., 2007] [Ichii et al., 2010b]. More recently, there has been significant debate surrounding the progenitors that precede CLP which has been attributed to the heterogeneity within the previously described progenitor populations such as the multi-lymphoid progenitor (MLP) and LMPP [Doulatov et al., 2010] [Goardon et al., 2011] [Karamitros et al., 2018]. So whether a true functional CLP exists and what it differentiated from remains to be determined.

The developmental hierarchy of human fetal B-lymphopoiesis is not as well studied. Cells of the B-lineage have been reported in FL and fetal BM by 7-8 pcw and 12 pcw respectively however the functional relationships and molecular properties of these cells were not characterised [Asma et al., 1984] [Uckun and Ledbetter, 1988] [Grumayer et al., 1991] [Nuñez et al., 1996] [Charbord et al., 1996] [Bueno et al., 2016]. Previously the Roberts lab reported the presence of a CD127+ progenitor called an Early Lymphoid Progenitor (ELP) in FL [Roy et al., 2012] which had originally been identified in adult BM [Ryan et al., 1997]; and since then, the presence of ELP in FL has been independently confirmed by others [Alhaj Hussen et al., 2017] [Boiers et al., 2018]. While all three groups used a slightly different immunophenotypic definition of ELP, all definitions commonly used CD127+ in their definition and functional analysis of the ELP demonstrated lympho-myeloid output. A functional CLP has not been reported in the human fetus, and the point at which progenitors commit to the B-lineage remains undetermined. Expression of the cell surface marker CD19 is the hallmark of B-lineage commitment and, in contrast to adult BM where the first committed B progenitor is a CD34+CD19+CD10+ ProB progentior, the presence of two committed B progenitors that differ in their CD10 expression has been observed in FL in the Roberts lab [Roy et al., 2012]. Others have reported the simultaneous occurrence of these two progenitors in CB [Sanz et al., 2003] [Sanz et al., 2010]. However, the relationship between ELP and PreProB progenitors and their position in the fetal B-lymphoid hierarchy remains to be determined.

1.2 Down Syndrome

Down Syndrome (DS) is the most common aneuploidy in humans and accounts for 1 in 909 live births in the UK [NCARDRS, 2017]. First described in 1866 by Langdon Down [Down, 1866], the cause of DS, an extra copy of chromosome 21 (trisomy 21, T21) in all cells, was not deciphered until 1959 [Lejeune et al., 1959]. Since then, others have shown that the most common cause of T21 is the result of failed maternal non-disjunction during meiosis (reviewed here: [Antonarakis, 1998]). Live births for trisomies 18 and 13 that cause Edward's and Patau's syndrome respectively are recorded but are much rarer [NCARDRS, 2017]. With 233 protein coding genes and 424 non-protein coding genes, chromosome 21 is the smallest of all human chromosomes [Ensembl, 2019] and this may therefore explain why T21 causes less disruption to the developing embryo/fetus than other trisomies.

1.2.1 Clinical characteristics of DS

The impact of T21 differs in each individual and therefore the phenotype of DS is varied. All children born with DS display craniofacial and cognitive impairments; and they are at increased risk of being born with cardiac or gastrointestinal defects as well as developing early onset dementia later in life (reviewed here: [Liu et al., 2014]). Children with DS are also at increased risk of developing distinct haematological abnormalities [Zipursky, 2003] [Roberts et al., 2013] and malignancies despite being less cancer prone in general [Hasle et al., 2016] [Hasle, 2001] [Hasle et al., 2000].

1.2.2 Haematological abnormalities in DS

Children with DS are at increased risk of developing both acute myeloid (Myeloid Leukaemia of Down Syndrome, ML-DS) and lymphoid (DS-Acute Lymphoblastic Leukaemia, DS-ALL) leukaemias [Hasle et al., 2016], with the risk being highest for children under 5 years old [Goldacre et al., 2004] [Hasle et al., 2016]. ML-DS develops from a pre-leukamia called Transient Abnormal Myelopoiesis (TAM) which occurs in 10% of children with DS and is unique to children with DS [Zipursky, 2003] [Roberts et al., 2013]. Of these 10%, up to 20% progress to full blown ML-DS [Gamis et al., 2003] [Massey et al., 2006]. On the other hand, DS-ALL is similar to non-DS-ALL although T-ALL is extremely rare in children with DS and infant ALL (iALL) has not been reported [Whitlock et al., 2005] [Roberts and Izraeli, 2014] [Buitenkamp et al., 2014] [Lee et al., 2016] .

In addition to the increased risk of acute leukaemia, children with DS are at increased risk of infection [Cocchi et al., 2007]. Severe sepsis and sepsis-related deaths are also increased [Garrison et al., 2005]. DS is also associated with an increased incidence of several autoimmune conditions such as coeliac disease, hypothyroidism or type 1 diabetes [Goldacre et al., 2004] [Gillespie et al., 2006] [Bergholdt et al., 2006] [Kusters et al., 2009] [Marild et al., 2013] [Jorgensen et al., 2019]. This pattern of co-morbidities points towards a role for T21 in the dysregulation of the developing and maturing immune system.

Children born with DS are nearly always born with non-malignant haematological abnormalities such as polycythaemia and thrombocytopenia [Roberts et al., 2013]. Similarly, at birth lymphocyte counts are at the lower end of the normal range and neutrophil and monocyte counts tend towards the higher end of the normal range

[Roberts et al., 2013]. Multiple in-depth investigations have shown that specifically, CD4+ T cell counts are significantly lower while CD8+ T cell counts are significantly higher in children with DS compared to age matched non-DS children [Cuadrado and Barrena, 1996] [Kusters et al., 2009] [Satge and Seidel, 2018]. These studies and others have also shown that immature/naive and memory B cells are significantly reduced and there is an imbalance in IgM and IgD subtypes in the peripheral blood of children with DS [Cuadrado and Barrena, 1996] [de Hingh et al., 2005] [Kusters et al., 2009] [Verstegen et al., 2010] [Carsetti et al., 2015] [Farroni et al., 2018]. In addition to lymphocyte counts being lower, T and B lymphocytes are likely to be functionally different. There is a decreased germinal centre response to vaccination which results in fewer memory B cells [Valentini et al., 2015] [Farroni et al., 2018]. With respect to T cell function, others have shown that DS is associated with increased IFN-γ producing T cells [Franciotta et al., 2006]. Finally, there is increased expression of other inflammatory and immune related genes in the blood of DS children [Zampieri et al., 2014] [Silva et al., 2016]. However, the contribution of each immune cell type to these reported signatures remains unknown as whole blood or peripheral blood mononuclear cells (PBMCs) were analysed. While these functional differences are likely to be major drivers in the increased susceptibility to infection and autoimmune diseases in DS children, the cause(s) of these differences remain undetermined. One hypothesis is that decreased CD4+ T cell counts directly reduces the memory B cell count [Verstegen et al., 2010]. However, given that T cell independent antibody responses are also reduced, the authors of this particular study also acknowledge that there must also be an intrinsic defect in B-lymphopoiesis [Verstegen et al., 2010].

1.2.3 Fetal liver haematopoiesis in DS

Previous members of the Roberts lab have shown that the effects of T21 on haematopoiesis begin in fetal life. T21 is associated with an increase in immunophenotypic MEP [Tunstall-Pedoe et al., 2008] and HSC, and a decrease in immunophenotypic GMP and committed B progenitors (defined as CD34+CD19+) [Roy et al., 2012]. T21 FL MEP and HSC also demonstrated increased clonogenicity in colony forming assays where more erythroid and megakaryocyte-erythroid colonies were formed [Tunstall-Pedoe et al., 2008] [Roy et al., 2012]. This was confirmed by others who also showed that T21 MEP continued to expand and differentiate in an NSG xenograft model [Chou et al., 2008]. In addition to the observed skew towards megakaryopoiesis, the immunophenotypic and functional analysis of progenitor populations in T21 FL

suggested there could be a defect in the B-lymphoid compartment in DS [Roy et al., 2012].

1.2.4 Mouse models of DS

As DS is caused by trisomy of human chromosome 21 (Hsa21), generation of a mouse model that faithfully recapitulates the complex and varied human disease phenotype presents a unique challenge. With the identification of all genes (protein and non-protein coding) on Hsa21, multiple mouse models have been established and characterised. Here, I will briefly describe the most established models, but all are reviewed extensively in [Herault et al., 2017] and also put into context in [Malinge et al., 2009] and [Antonarakis, 2017].

Orthologous regions of Hsa21 are spread over three mouse chromosomes: Mmu10, 16 and 17, with the largest complete region on Mmu16. Triplication of Mmu16 in mice results in an embryonic lethal phenotype [Gropp, 1974] [Herault et al., 2017] but triplication of segments of Mmu16 (figure 1.3) has been extensively studied. The Ts65n mouse [Davisson et al., 1990] [Reeves et al., 1995] is characterised by cognitive impairment, craniofacial dysmorphy and cardiac abnormalities [Reeves et al., 1995] [Moore, 2006]. Two other models harbour smaller segments of Mmu16: Ts1Cje [Sago et al., 1998] and Ts1Rhr [Olson et al., 2004] and are also associated with cognitive differences. Both the Ts65n and Ts1Cje mouse have abnormal adult erythro- and myelopoiesis and increased HSC frequency [Kirsammer et al., 2008], however, fetal haematopoiesis is normal [Malinge et al., 2009]. While these models have been essential in determining key genes and regions on Hsa21, not all genes on Hsa21 are conserved in the mouse and also there are triplications of murine genes in each of these models that do not exist in humans [Herault et al., 2017]. Also, the role of Hsa21 miRNAs in these mice, as well as the role of genetic backgrounds on the phenotypes observed is likely to be underestimated [Klusmann et al., 2010b] [Shaham et al., 2012], [Emmrich et al., 2014] [Alexandrov et al., 2018].

In a different approach to modelling DS, O'Doherty *et al.* introduced a copy of Hsa21 to murine ES cells to make a mouse line that transmitted Hsa21 through the germline to produce the Tc1 mouse [O'Doherty et al., 2005]. While this approach was exciting, the haematopoietic abnormalities did not fully recapitulate those in the human [Alford et al., 2010] possibly because, as shown in a subsequent study, there were mutliple radiation-induced Hsa21 rearrangements [Gribble et al., 2013].

To specifically address the effects of gene balance of selected gene segments, several murine models have been developed by the Herault, Yu and Tybulewicz labs in parallel (summarised in figure 1.3), using Cre/LoxP technology to induce triplication. However, these models require more extensive study to establish their phenotype with respect to neurodevelopment, cardiac development and haematopoiesis in adult and fetal life [Antonarakis, 2017]. Furthermore, as CRISPR technology makes it quicker and easier to generate murine models, more DS mouse models are no doubt in the pipeline.

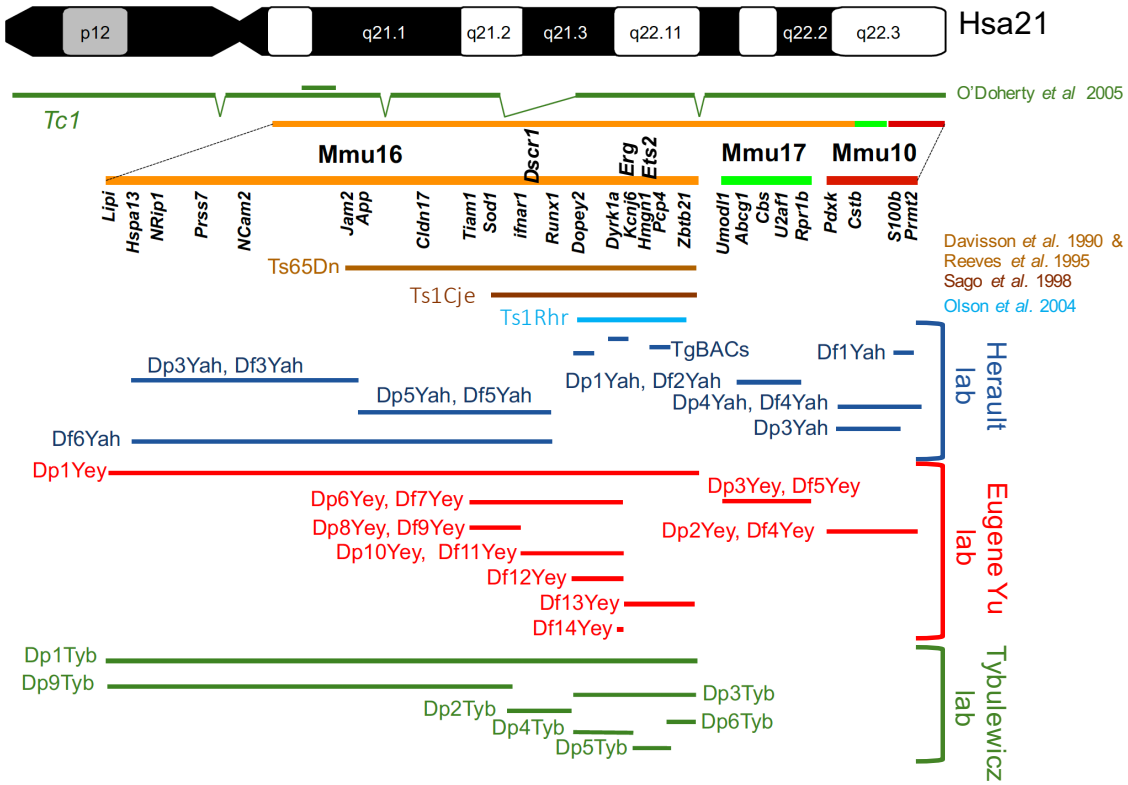


Figure 1.3: Illustration depicting transgenic mice developed to model DS. Hsa21 is shown at the top with coloured lines representing the regions that were triplicated (for Tc1) or orthologous regions in murine chromosomes (orange, bright green and red) that were triplicated in each model underneath. Several genes of interest are annotated. For the Tc1 mouse (top green line), double lines represent regions of duplications and dipped line represents a region of deletion. Light brown, dark brown and light blue lines represent regions of Mmu16 triplication in the most extensively studied mouse models (Ts65n, Ts1Cje, Ts1Rhr) with key references labelled to the right. Underneath, dark blue, red and green lines represent regions of triplication developed by labs labelled on the right. Adapted from [Herault et al., 2017] [Antonarakis, 2017] and [Malinge et al., 2009]

1.2.5 *In vitro* models of DS

According to [Antonarakis, 2017] there are twelve induced pluripotent stem cell (iPSC) lines derived from DS fibroblasts or amniotic fluid in existence. While most of the functional characterisation of these has focussed on induced differentiation into neurons, some studies have used DS iPSC to model haematopoiesis. Two studies published in tandem in *PNAS* in 2012 showed that DS iPSC differentiated into haematopoietic progenitors had increased erythroid and multi-lineage colony-forming potenital [MacLean et al., 2012] and reduced myelopoiesis [Chou et al., 2012] which is similar in some respects to the abnormalities in human FL [Tunstall-Pedoe et al., 2008] [Chou et al., 2008] [Roy et al., 2012]. However, both iPSC models failed to recapitulate the increased MEP observed in T21 FL. Faster expansion of haematopoietic lineage cells after iPSC differentiation was also observed by Murray *et al.* [Murray et al., 2015]. These iPSC models have proved useful in modelling TAM where key roles for Hsa21 genes *RUNX1*, *ETS2* and *ERG* were suggested [Banno et al., 2016]; in exploring the role of gene dosage at the chromosome level by silencing the trisomic Hsa21 using *XIST* [Chiang et al., 2018]; and also in interrogating the role of *EDNRB* (endothelin receptor) in T21 B cell differentiation [MacLean et al., 2018]. So while iPSC models are potentially valuable for investigating the role of Hsa21 genes, a comprehensive understanding of the mechanisms by which T21 perturbs fetal haematopoiesis is likely to require additional detailed studies in primary human tissue.

1.3 DS-ALL

1.3.1 Incidence and clinical characteristics of DS-ALL

Children with DS are at 33x greater risk of developing DS-ALL [Hasle et al., 2000] [Hasle, 2001] [Hasle et al., 2016] and older studies show increased rates of relapse and treatment-related mortality, particularly due to infections [Bohnstedt et al., 2013][Meyr et al., 2013] [Buitenkamp et al., 2014]. Although a recent report demonstrates that survival rates for DS-ALL are now comparable to non-DS-ALL patients [Matloub et al., 2019], infection remains the commonest cause of death in DS at all ages [Englund et al., 2013]. The incidence of DS-ALL peaks in children between the ages of 1 and 5 years old and DS-ALL has never been recorded in infants (<12 months old) [Hasle et al., 2016] [Buitenkamp et al., 2014]. DS-ALL is of a PreB immunophenotype (CD19+ CD10+/- cytoplasmic-IgM+) and T lineage ALL in DS is extremely rare [Pui et al., 1993] [Zeller et al., 2005] [Arico et al., 2008].

1.3.2 Genetics of DS-ALL

The distribution of cytogenetic abnormalities used to stratify ALL patients into subgroups differs between non-DS-ALL and DS-ALL patients. Subgroups that are common and associated with favourable outcome in non-DS-ALL (ETV6-RUNX1, high hyperploidy) are much less common in DS-ALL [Buitenkamp et al., 2014]. Similarly, high risk groups such as MLL rearranged ALLs are rare amongst DS-ALL patients [Maloney et al., 2010] [Buitenkamp et al., 2014].

Strikingly, 50-60% of DS-ALL patients (compared to 10% of non-DS-ALL patients) have a rearrangement of the *CRLF2* gene (CRLF2r) that causes its overexpression [Mullighan et al., 2009] [Russell et al., 2009] [Yoda et al., 2010] [Hertzberg et al., 2010] [Buitenkamp et al., 2014] [Schwartzman et al., 2017]. CRLF2r ALLs (non-DS and DS) are most commonly due to P2RY8-CRLF2 chromosomal rearrangements or less commonly IGH-CRLF2 chromosomal translocations [Mullighan et al., 2009] [Russell et al., 2009] and in rare cases CRLF2 overexpression is caused by a point mutation of the *CRLF2* gene itself [Hertzberg et al., 2010] [Yoda et al., 2010]. *CRLF2* encodes the thymic stromal lymphopoietin (TSLP) receptor which dimerises with IL-7Rα (*IL7R*) to transduce TSLP signalling via the JAK-STAT pathway. Interestingly, DS-ALL is also closely associated with activating mutations in the JAK/STAT signalling pathway [Bercovich et al., 2008] [Nikolaev et al., 2014], which frequently co-occur with CRLF2r [Mullighan et al., 2009] [Hertzberg et al., 2010] [Yoda et al., 2010] [Schwartzman et al., 2017]. The tonic hypersignalling of the JAK/STAT pathway induced by these aberrations was hypothesised to drive DS-ALL [Lee et al., 2016] [Bercovich et al., 2008] [Nikolaev et al., 2014] and recently Schwartzman *et al.* addressed this in a comprehensive genomic analysis of DS-ALL [Schwartzman et al., 2017]. The CRLF2r (*CRLF2* high expressing) subgroup displayed a bias towards IL7R/JAK1/JAK2 mutations that were lost in relapse, thereby demonstrating JAK/STAT signalling as an initial driver of leukaemia. In contrast, the *CRLF2* low expressing subgroup had a higher incidence of RAS pathway mutations supporting previous work demonstrating the mutual exclusivity of JAK and RAS pathway mutations in DS-ALL, [Nikolaev et al., 2014] and also another recent study on P2RY8-CRLF2 ALL [Vesely et al., 2017]. Also in agreement with Vesely *et al.*, Schwartzman *et al.* reported the increased incidence of RAS pathway mutations and persistence of loss of function *IKZF1* mutations in relapse samples. Finally, it is worth noting that while Vesely *et al.* reported that one third of CRLF2r were lost upon relapse in P2RY8-CRLF2, and others have reported similar observations where approximately 25% of P2RY8-CRLF2 DS-ALL

are subclonal [Potter et al., 2018], this was not the case in the study by Schwartzman *et al.* highlighting the need for larger studies of these leukaemias. Others have recently taken a step in this direction by performing a meta-analysis of available datasets and confirming the likely role of *IKZF1, ARID5B, GATA3* and *CDKN2A* in DS-ALL as heritable disease risk loci [Brown et al., 2019].

1.3.3 *In utero* initiation of ALL

Tracing of chromosomal translocations back to birth through analysis of CB and neonatal blood spots, as well as twin studies, suggest that several childhood leukaemias can arise *in utero* [Greaves, 2018]. Specifically, with respect to DS leukaemias, the multi-step process to ML-DS starts *in utero* and is well established [Roberts et al., 2013] [Bhatnagar et al., 2016] [Labuhn et al., 2019]. With respect to B cell precursor (BCP)-ALL, several subtypes of BCP-ALL have been traced back to birth including ETV6-RUNX1 [McHale et al., 2003] [Skorvaga et al., 2014] [Alpar et al., 2015], BCR-ABL1 [Cazzaniga et al., 2011] [Skorvaga et al., 2014], high hyperploidy [Taub et al., 2002] [Bateman et al., 2015] and MLL-rearranged [Ford et al., 1993] [Gale et al., 1997] [Fasching et al., 2000] [Yagi et al., 2000] [Greaves et al., 2003] [Skorvaga et al., 2014] . These studies have lead to the hypothesis that many paediatric BCP-ALLs begin *in utero* where a clone arises from a fetal HSC/progenitor that acquires a chromosomal translocation. Then, at some point after birth, further oncogenic mutations in this clone lead to overt leukaemia (figure 1.4). Given that P2RY8-CRLF2 rearrangements have also been backtracked to birth through neonatal blood spot analysis [Morak et al., 2012]; that CRLF2r can be early events in DS-ALL leukaemogenesis [Potter et al., 2018]; and the young age at which children with DS are most likely to develop ALL (1-5 years), it is plausible that DS-ALL also initiates *in utero*.

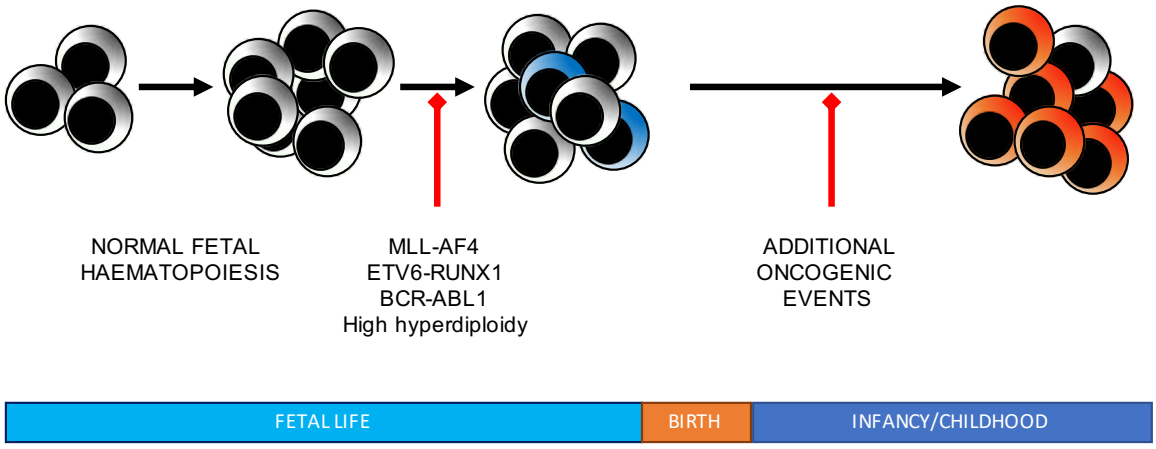


Figure 1.4: Model for *in utero* initation of ALL based on studies tracing leukaemic clones back to birth through neonatal blood spot analysis, CB analysis and twin studies.

1.3.4 Chromosome 21 genes linked to increased susceptibility to DS-ALL

***HMGN1*:** Located on chromosome 21, *HMGN1* encodes a nucleosome binding protein that has been implicated in a model of DS-ALL. Using the Ts1Rhr mouse, Lane *et al.* modelled DS-ALL by introducing a *JAK2* mutation, over-expressing *CRLF2*, knocking out one allele of *Pax5* and introducing a dominant negative isoform of *IKZF1*(Ik6) [Lane et al., 2014]. Characterisation of the leukaemia that ensued in these mice, and comparison with human leukaemia data, suggested that DS conferred a global decrease in H3K27me3 which is a repressive mark added by polycomb repressive complex 2 (PRC2). Then, through a series of *in vitro* and *in vivo* experiments when *HMGN1* only was overexpressed, the authors concluded that triplication of *HMGN1* may play a role the reduced H3K27me3 signature (inferred from gene expression analysis), leukaemic cell survival and reduced B progenitors [Lane et al., 2014]. Additional studies from the same lab have recently shown that the epigenomic changes that correlated with increased *HMGN1* expression are directly correlated with global increases in the activating histone mark H3K27ac and global increases in RNA in cells trisomic for *HMGN1* [Mowery et al., 2018]. This led to the hypothesis that over-expression of *HMGN1* causes chromatin de-compaction and enhanced transcription that improves leukaemia cell survival. However, the authors do acknowledge that the DS phenotype is "complex and polygenic" and that other trisomic genes are likely playing a role in the B-lymphoid defect and increased incidence of ALL [Mowery et al., 2018]. It is also worth noting that the direct relationship between *HMGN1* over-expression and H3K27me3 repressive marks remains unexplored.

***DYRK1A*:** In a ML-DS mouse model using the same DS mouse (Ts1Rhr) and inducing *GATA1* and *MPL* mutations, *DYRK1A* was identified as a likely mediator for the abnormal megakaryopoiesis that leads to ML-DS [Malinge et al., 2012]. Further investigations by the same lab showed that *DYRK1A* promoted quiescence via destabilisation of Cyclin D3 at the PreB transition in murine B progenitors thereby demonstrating the importance of cellular context in the function of *DYRK1A* [Thompson et al., 2015]. How, or if, trisomy of *DYRK1A* contributes to DS-ALL is unknown but recently *DYRK1A* was shown to be involved in the regulation of the DNA damage repair response [Guard et al., 2019]. These investigations point to a potential role for *DYRK1A* in the B-lymphoid defect in T21 FL.

1.4 Hypothesis and aims

The main aim of this project was to investigate the cellular and molecular mechanisms responsible for perturbed fetal B-lymphoid development in T21 and how this may contribute to ALL. I hypothesised that in DS, perturbation of fetal B-cell development by T21 promotes leukaemic transformation by altering molecular programmes that regulate normal B-lymphopoiesis. I divided the aim into three questions:

1. **What is the cellular hierarchy and molecular basis for B-lymphoid development in normal fetal BM?** I addressed this by:
   (a) characterisation of normal fetal BM by immunophenotyping, functional assays and molecular analysis.
   (b) comparison of normal fetal BM with normal FL.
2. **What is the impact of T21 on fetal BM B progenitor development?** I addressed this by:
   (a) characterisation of T21 fetal BM by immunophenotyping, functional assays and molecular analysis to compare with normal fetal BM.
   (b) investigating the role of T21 microenvironment through functional and molecular analysis.
3. **What is the impact of T21 on the transcriptome of ALL blasts?** I addressed this by:
   (a) immunophenotypic analysis of non-DS-ALL and DS-ALL patient BM.
   (b) transcriptome analysis of non-DS-ALL and DS-ALL blasts and progenitors.

Chapter 2

Methods

2.1 Samples

Fetal tissue was collected during elective termination of pregnancy after specific written informed consent was obtained for the collection and use of the fetal tissue in accordance with the Declaration of Helsinki. This fetal tissue is staged and distributed by the Human Developmental Biology Resource (HDBR, `www.hdbr.org`; details of staging in table 2.1; REC: 18/LO/0822 and 18/NE/0290). The karyotypes for all samples were confirmed by the HDBR. For samples which failed karyotype analysis, fluorescence *in situ* hybridisation (FISH) was used to confirm the lack of chromosomal abnormalities in normal samples and the presence of T21 in DS samples. Some samples were not provided by the HDBR. In these cases, all sampling, consent and experimental procedures were approved by the Hammersmith, Queen Charlotte's & Chelsea and Acton Hospitals Ethics committee prior to the commencement of the project (REC 04/Q0406/145).

Cord blood (CB) was collected from healthy newborns after elective caesarian sections and after specific written informed consent was obtained (collaboration with Prof. Suzanne Watt, REC: 04/Q0406/145).

Cryopreserved patient leukaemia samples were obtained from the Bloodwise Childhood Leukaemia CellBank (REC: 16/SW/0219)

Normal adult bone marrow (BM) was collected during elective orthopedic surgery after written informed consent and approved by INForMeD HRA approval (collaboration with Dr Beth Psaila; IRAS 199833; REC no. 16/LO/1376). Otherwise, adult BM mononuclear (MNC) were purchased from StemCell Technologies, Canada and Lonza, USA. Paediatric BM was collected during harvest from sibling donors for bone marrow transplant and approved for use in research under MREC no. 12/LO0426 (collaboration with Dr Deena Iskander, Imperial College London).

Table 2.1: HDBR Fetal staging guide [Hern, 1984]

Post conception weeks (pcw)	Foot length (mm)	Knee – heel length (mm)	Gestational age in weeks
Late 8	Above 5 and less than 6	8	10
9	6 or above and less than 8	11	11
10	8 or above and less than 10	13	12
11	10 or above and less than 13	17	13
12	13 or above and less than 17	24	14
13	17 or above and less than 20	31	15
14	20 or above and less than 23	36	16
15	23 or above and less than 25	40	17
16	25 or above and less than 27	43	18
17	27 or above and less than 31	51	19
18	31 or above and less than 33	56	20
19	33 or above and less than 36	60	21
20	36 or above and less than 40	66	22

2.2 Sample processing, cryopreservation and thawing

Fetal tissues were transported to the laboratory at 4°C and processed immediately as described previously [Roy et al., 2012]. Briefly, fetal liver was passed through a 70μm filter and Phosphate Buffered Saline (PBS) supplemented with 2% heat inactivated fetal bovine serum (HI FBS) and 1mM EDTA. Long bones were flushed with Dulbecco's Modified Eagle Medium (DMEM) and then passed through at 70μm filter to obtain a single cell suspension. Early gestation (8-10 weeks) fetal bone samples were crushed and filtered. CB was passed through a 70μm filter and diluted in PBS at a ratio of 1 part CB to 2 parts PBS. Single cell suspensions were then red cell and granulocyte depleted by density gradient separation using Ficoll-Paque PLUS (GE Healthcare) and CD34 enrichment was carried out on freshly isolated MNC using a Miltenyi CD34 MicroBead kit. Samples were then aliquoted and cryopreserved in HI FBS with 10% DMSO. Samples were thawed at 37°C for 3 minutes and then diluted in an excess of HI FBS, spun at 300xg for 5 minutes and resuspended in IMDM+10% HI FBS or PBS+2%FBS+1mM EDTA (FACS buffer) supplemented with DNase (100ng/mL). Post-enrichment CD34+ cell purity was >95% in all cases.

2.3 Flow cytometric analysis and sorting

Cells were stained for flow cytometry with fluorophore-conjugated monoclonal antibodies (mAb) and viability dye in 100uL PBS supplemented with FACS buffer for 30 minutes followed by two 1mL washes where samples were spun at 300xg for 5 minutes. Antibody details are described in table 2.2. Samples were analysed using a BD LSR II, BD Fortessa X30 and sorted using BD FACSAriaII or BD Fusion instruments. Gates were set with unstained and fluorescence minus one controls gating on viable cells. Data were analysed using FlowJo V10.0 software. Purity was always >95% (example purity plots shown in figures 2.1 and 2.2).

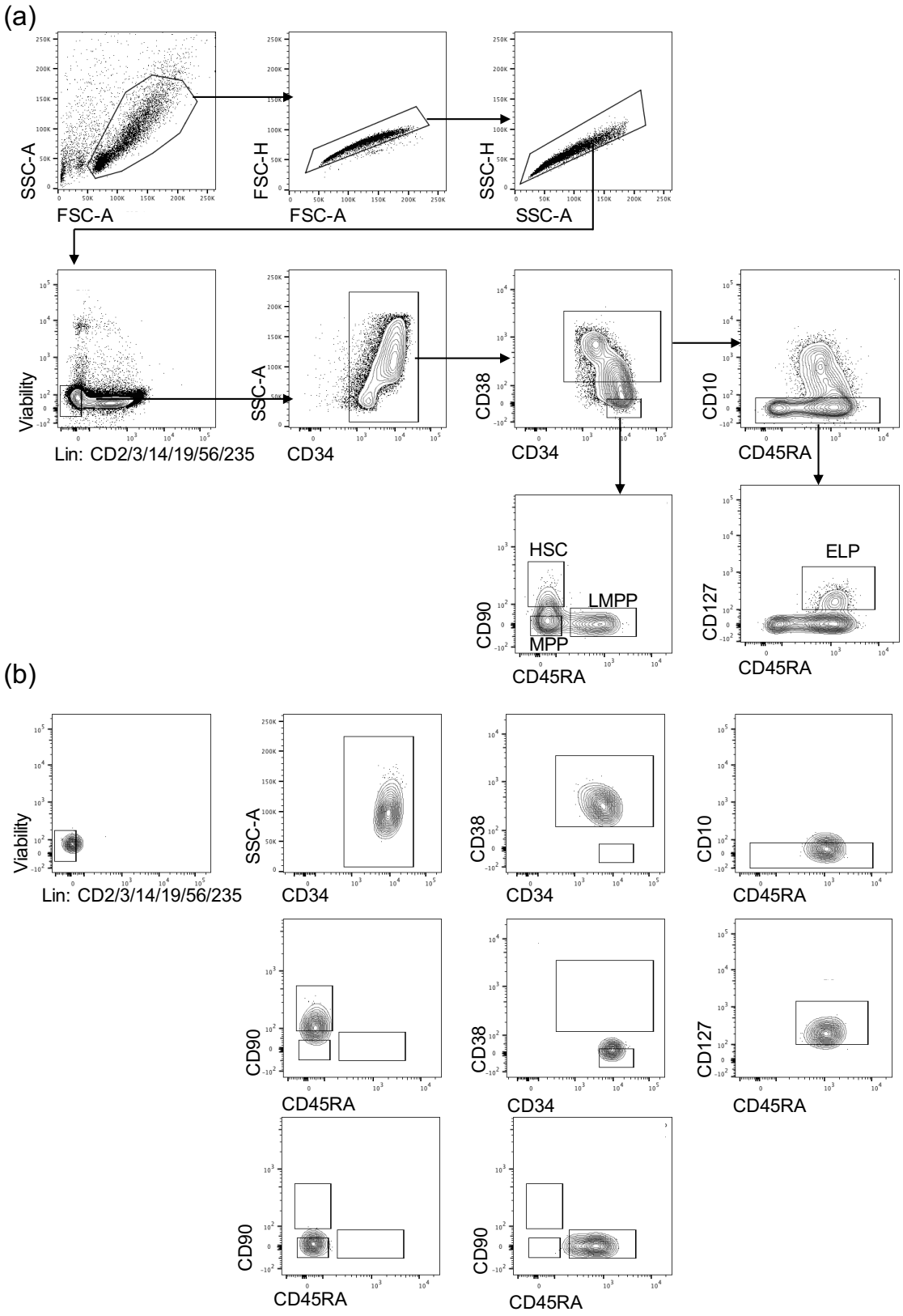


Figure 2.1: Purity plots for sorted HSC, MPP, LMPP and ELP

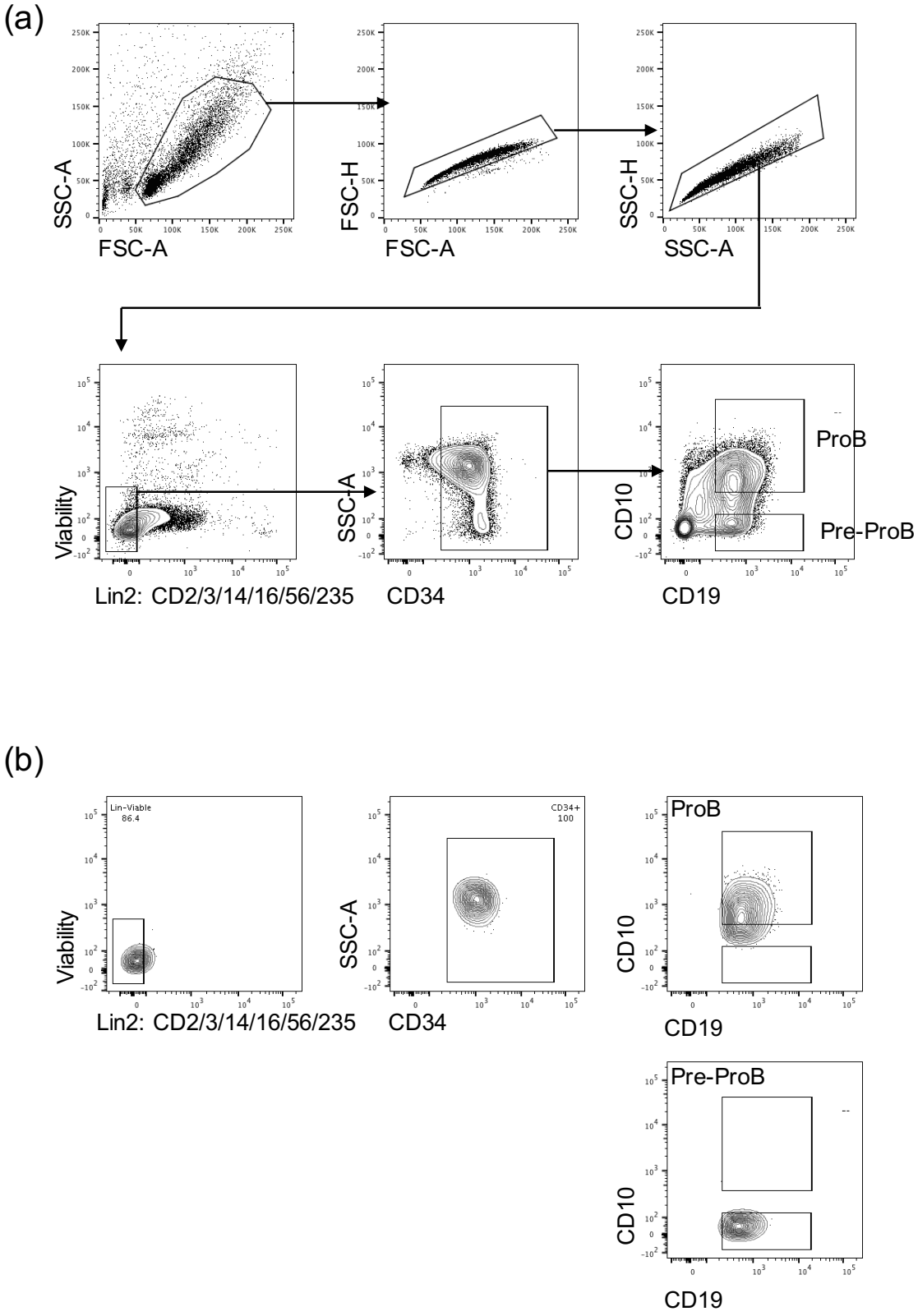


Figure 2.2: Purity plots for sorted PreProB and ProB progenitors

Table 2.2: Details of antibodies used for flow cytometry

antibody	colour	clone	company	cat no
CD2	PerCPcy5.5	RPA-2.10	Biolegend	300215
CD2	FITC	RPA-2.10	Thermofisher	11-0029-42
CD3	ef506	UCHT1	Thermofisher	4306053
CD3	PerCPcy5.5	OKT3	Biolegend	317336
CD3	FITC	OKT3	Thermofisher	11-0037-42
CD3	AF700	OKT3	Thermofisher	56-0037-42
CD3	BV711	okt3	Biolegend	317327
CD4	PE	L200	BD	550630
CD5	PE	L17F12	Biolegend	364013
CD7	BV421	MT701	BD	562635
CD8	APC	RPA-T8	Biolegend	301014
CD10	Pe_Cy7	ebioCB-CALLA	Thermofisher	25-0106-42
CD10	PE	HI10A	BD	555375
CD10	FITC	ebioCB-CALLA	Thermofisher	11-0106-42
CD14	FITC	61D3	Thermofisher	11-0149-41
CD14	PerCPcy5.5	M5E2	Biolegend	301824
CD15	PerCPcy5.5	W6D3	Biolegend	323019
CD16	PerCPcy5.5	3G8	Biolegend	302027
CD19	APC	HIB19	BD	555415
CD19	APC	HIB19	Thermofisher	17-0199-42
CD20	PE_Dazzle_594	2H7	Biolegend	302348
CD20	ef450	2H7	Thermofisher	48-0209-41
CD20	APC-Fire750	2H7	Biolegend	302357
CD27	PerCPcy5.5	M-T271	Biolegend	356407
CD27	ef450	O323	Thermofisher	48-0279
CD33	FITC	p67.6	Biolegend	366620
CD34	APC_ef780	4H11aka8G12	Thermofisher	47-0349-42
CD34	PerCP_Cy5.5	581	Biolegend	343522
CD34	PEcy7	4H11aka8G12	Thermofisher	25-0349-42
CD34	BV421	561	Biolegend	343609
CD38	BV605	HIT2	Biolegend	303532
CD38	APC	HIT2	Thermofisher	17-0389-42
CD38	af700	HIT2	Thermofisher	56-0389-42
CD38	PeDz	HIT2	Biolegend	303537
CD43	BV421	1G10	BD	562916
CD45	Alexa700	2D1	Thermofisher	56-9459-42
CD45	Alexa700	2D1	Thermofisher	56-9459-42
CD56	PE	CMSSB	Thermofisher	12-0567-41
CD56	PerCPcy5.5	HCD56	Biolegend	318321
CD56	FITC	TULY56	Thermofisher	11-0566-42
CD56	PE	CMSSB	Thermofisher	12-0567-41

Table 2.2: Details of antibodies used for flow cytometry

antibody	**colour**	**clone**	**company**	**cat no**
CD56	APC-Cy7	HCD56	Biolegend	318331
CD56	BV605	HCD56	Biolegend	318334
CD73	BV421	AD2	Biolegend	344008
CD73	APC-Cy7	AD2	Biolegend	344021
CD73	BV421	AD2	Biolegend	344008
CD90	BV421	5E10	Biolegend	328122
CD90	PE	5E10	Biolegend	328122
CD105	BV421	43A3	Biolegend	323219
CD105	APC	43A4E1	Miltenyi	130-099-125
CD123	PECy7	6H6	Thermofisher	25-1239-42
CD123	PE	9f5	BD	340545
CD123	BV650	7G3	BD	563405
CD127	FITC	RDR5	Thermofisher	11-1278-42
CD127	PE	eBioRDR5	Thermofisher	12-1278-42
CD133	PE	AC133	Miltenyi	130-098-826
CD133	APC	AC134	Miltenyi	130-098-826
CD146	PECy7	P1H12	BD	342010/562135
CD11b	FITC	ICRF44	Thermofisher	11-0118-41
CD235a	PerCPcy5.5	HIR2	Biolegend	306614
CD235a	FITC	HIR2	Thermofisher	11-9987-82
CD45.1 (mouse)	APCCy7	30-F11	Biolegend	103115
CD45RA	PE	MT4 (6B6)	BD	555904
CD45RA	BV421	HI100	Biolegend	304129
CD45RA	APC	HI100	Biolegend	304112
CD45RA	FITC	HI100	Biolegend	304106
HLA-DR	FITC	L243	Thermofisher	11-9952-41
IgD	APC_Cy7	IA6-2	Biolegend	348218
IgD	FITC	IA6-2	Biolegend	348205
IgM	FITC	MHM-88	Biolegend	314506
Ki67	AF700	KI67	Biolegend	B186549
Lineage	APC		Thermofisher	22-7776-72
Lineage	FITC		Thermofisher	22-7778-72
Viability	ef506		Thermofisher	65-0866-18
Viability	ef520		Thermofisher	65-0867-18

2.4 Cell cycle analysis

After cells were stained using standard protocol (described above) cells were fixed and permeabilised with BD Cytofix/Cytoperm solution as per the manufacturers instructions. Then intracellular staining was performed with anti-Ki67-AF700 antibody (table 2.2 for details) for 30 minutes at 4°C. Staining was followed by 2 washes and then immediately before analysis, cells were stained with 1.25μg/mL Hoechst 33341 (Sigma B2261).

2.5 MS-5 stromal co-culture system

Murine MS-5 stromal cells [Itoh et al., 1989] were a kind gift from Prof. Paresh Vyas, University of Oxford, Oxford. MS-5 stromal cells were grown to confluency and passaged as described previously [Goardon et al., 2011]. Briefly, when MS-5 cells reached 80% confluency, monolayers were gently washed with PBS and then incubated with 0.05% trypsin (Gibco) for 5 minutes at 37°C. Trypsinisation was stopped by adding an excess of media (α-MEM supplemented with 10% HI FBS, 100U/mL Penicillin, 100ug/mL Streptomycin, 2mM L-glutatmine, 50μM 2-Mercaptoethanol, 10mM HEPES) and cells were split 1 in 5 in a 75 cm^2 flask. When required for co-cultures, 0.03 x10^6 MS-5 cells were plated per well in a 24 well plate 24 hours prior to co-culture initiation. 100 sorted fetal BM CD34+ HSC and progenitors were co-cultured on confluent MS-5 stroma in α-MEM (Gibco) supplemented with 10% HI FBS, 100U/mL Penicillin, 100μg/mL Streptomycin, 2mM L-glutatmine,50μM 2-Mercaptoethanol, 10mM HEPES, SCF (20ng/mL), Flt3L (10ng/mL), IL-2 (10ng/mL) and IL-7 (5ng/mL) in 24 well plates. All cytokines were purchased from Preprotech, Rocky Hill, USA. Co-cultures were replenished with half media changes every 3-4 days. At weekly time-points B and NK-lymphoid differentiation was assessed during the re-plating of co-cultures where one fifth of the cell suspension was reserved for flow cytometry before the cells were re-plated on fresh stromal layers. In some experiments, where cell numbers permitted, multiple wells of 100-1000 cells were plated and at each time point a whole well was harvested for flow cytometry analysis. Where absolute cell numbers are quoted, CountBright beads (Invitrogen) were used to calculate the number of cells present in each well at each time point.

2.6 OP9-DL1 stromal co-culture system

OP9 stromal cells ectopically expressing notch ligand delta-like 1 (OP9-DL1) and GFP were a kind gift from Prof Adrian Thrasher (Institute of Child Health, University College, London). OP9-DL1 stromal cells were grown to confluency and passaged as described previously [Awong et al., 2007]. Stromal layers were prepared by plating 0.02x10^6 OP9-DL1 cells per well onto 0.1% gelatin coated 24-well plates. After 24 hours, 100 sorted HSC or progenitors were seeded onto the stromal layers and co-cultured with α-MEM (Gibco) supplemented with 20% HI FBS, 100U/mL Penicillin, 100μg/mL Streptomycin, 2mM L-glutatmine, Flt3L (5ng/mL) and IL-7 (5ng/mL). All cytokines were purchased from Peprotech, Rocky Hill, USA. Co-cultures were replenished with half media changes every 3-4 days. As with the MS-5 co-cultures, T-lymphoid differentiation was assessed during the re-plating of co-cultures where one fifth of the cell suspension was reserved for flow cytometry before the cells were re-plated on fresh stromal layers. Murine stromal cells were gated out of analysis according to their expression of GFP.

2.7 Clonogenic assays

Myeloid and erythroid differentiation potential of 100 sorted HSC and progenitors were assessed using colony forming assays as previously described [Roy et al., 2012]. Briefly, 100 HSC or progenitors were sorted in to 750μL MethoCult H4230 (StemCell Technologies, Vancouver, Canada) supplemented with: 100U/mL Penicillin, 100ug/mL Streptomycin, 20ng/mL IL-3, 1-ng/mL IL-6, 10ng/mL IL-11, 10ng/mL SCF, 10ng/mL Flt3L, 50ng/mL GM-CSF, 50ng/mL TPO and 4IU/mL EPO. All cytokines were purchased from Peprotech, Rocky Hill, USA with the exception of EPO (EPREX®) which was purchased from Janssen Biologics, Leiden, The Netherlands. After vigorous vortexing and allowing the methocult to settle thus removing the bubbles, cell suspensions were seeded into 24 well plates and incubated at 37°C and 5% CO2. After 12-14 days, colonies were counted and photographed. Selected colonies were picked for cytospins and staining.

2.8 Cytospins and staining

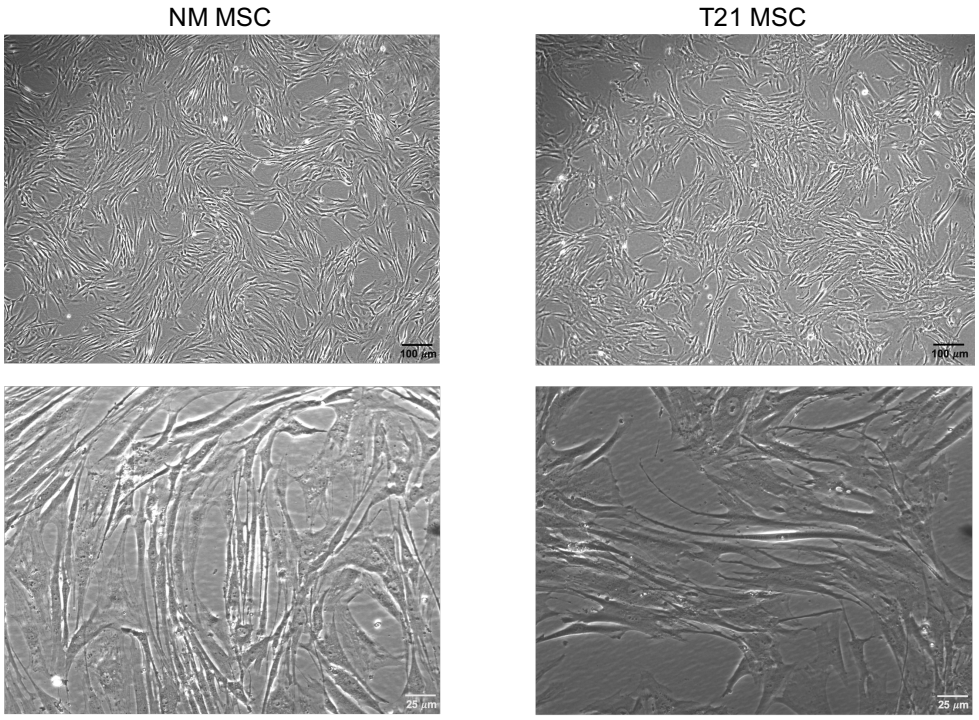
Cells were harvested from colonies or stromal co-cultures and washed by dilution in 1mL PBS followed by 300xg spin for 5 minutes. Cell pellets were re-suspended in 100uL PBS and fixed to slides using a Shandon Cytospin 4 at 400rpm for 10 minutes.

Slides were stained with Modified Wright's Stain and imaged on an EVOS AMG camera.

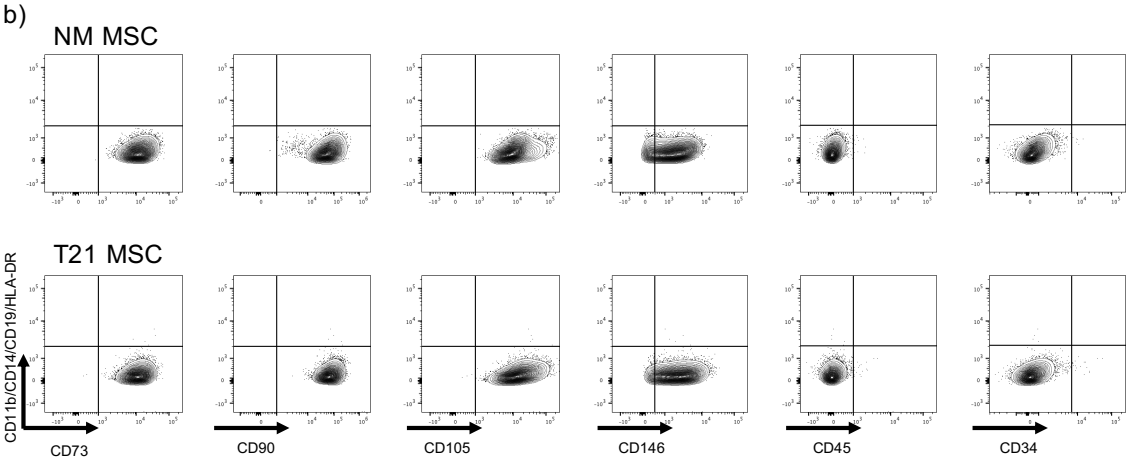
2.9 Fetal mesenchymal stromal cell co-culture system

Derivation: Fetal BM mesenchymal stromal cells (MSC) were derived by plating 1-2x10^6 freshly isolated mononuclear cells (MNC) per well in a six-well plate with αMEM (Gibco) supplemented with 10% HI FBS, 100U/mL Penicillin, 100μg/mL Streptomycin, 2mM L-glutatmine, 50μM 2-Mercaptoethanol, 10mM HEPES (this media is referred to αMEM-10 hereafter). After 48-72 hours of incubation at 37°C and 5% CO2 non-adherent cells were removed, and the media refreshed. MSC were expanded and passaged as described for MS-5 cells above but were split 1 in 3 to allow for efficient growth. MSC identity was confirmed by their morphology and internationally agreed [Dominici et al., 2006] immunophenotypic analysis (figure 2.3).

a)



b)



c)

Average expression as % viable cells ± SEM		
Marker	NM FBM	T21 FBM
CD73+	100 ± 0.00	99.97± 0.03
CD90+	99.97± 0.03	100± 0.00
CD105+	99.97± 0.03	99.67± 0.28
CD146+	69.7± 11.8	86.07± 9.59
CD45-	99.8± 0.15	99.9± 0.06
CD34-	99.77± 0.23	99.93± 0.07
CD11b/14/19/HLA-DR-	100± 0.00	99.97± 0.03

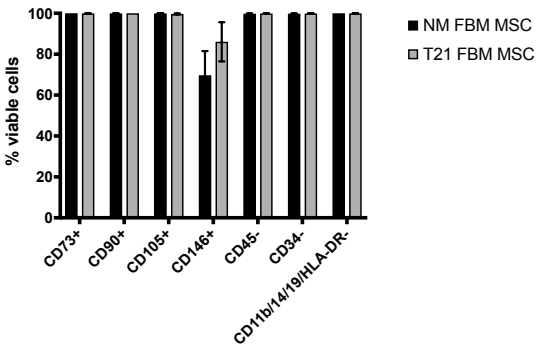


Figure 2.3: Normal (NM) and T21 MSC characterisation by morphology (a) and immunophenotyping (b) and (c)

Co-culture: 24 hours before co-culture initiation, 0.04x10^6 MSC were plated per well in a 24 well plate in 500μL of αMEM-10. After 24 hours of incubation to allow the MSC to settle and adhere, 100 sorted HSC or progenitors were seeded per well on to the monolayers in 500μL media supplemented with SCF (40ng/mL), Flt3L (20ng/mL), IL-2 (20ng/mL) and IL-7 (10ng/mL). This is illustrated in figure 2.4. Co-cultures were replenished with half media changes (supplemented with SCF, 20ng/mL; Flt3L, 10ng/mL; IL-2, 10ng/mL; and IL-7, 5ng/mL) once a week and re-plated on to fresh stromal layers once a week. During re-plating 1/5th of each well was reserved for analysis by flow cytometry. CountBright beads (Invitrogen) were used to calculate the number of cells present in each well at each time point. Where cell numbers permitted, multiple wells of 100-1000 cells were plated and upon re-plating a whole well was harvested for flow cytometry analysis. In some experiments (during optimisation), to prevent over-growth of the MSC, they were irradiated with 20Gy before being plated.

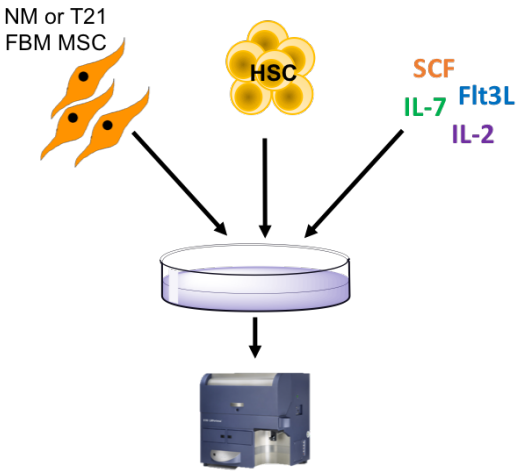


Figure 2.4: Schematic illustration of novel co-culture system

2.10 Mice and xenograft assays

NOD.Cg-Prkdc^{scid}Il2rg^{tmlWjl}/SzJ (NSG) mice were a kind gift from Prof. Terence Rabbitts group. All mice were housed in individually ventilated cages and looked after by the Biomedical Services team at the University of Oxford. All experiments were performed under a project license (PIL 3003442) approved by the UK Home Office under the Animal (Scientific Procedures) Act 1986 and in accordance with the principles of 3Rs (replacement, reduction and refinement) in animal research.

Xenograft transplants: 10,000-30,000 Lin2-CD34+CD19-CD10-, PreProB- or ProB-progenitors were sorted into α-MEM (Gibco) supplemented with 10% FBS. After sorting, cells were washed and re-suspended in PBS supplemented with 1% FBS and injected via the tail vein into sub-lethally irradiated (total dose of 2.5Gy split into two exposures of 1.25Gy each) NSG mice. At 2-3 week intervals mice were either monitored for engraftment by blood sampling or sacrificed and tissues (BM and spleen) analysed by flow cytometry. Thank you to Nicholas Fordham and Siobhan Rice for performing the mouse injections.

2.11 IgH locus rearrangement analysis

100-10,000 cells were sorted into PBS, spun at 500xg for 10 minutes and all supernatant was removed before being frozen at -20°C. Whole genome amplification was performed using the REPLI-g Single Cell Kit (Qiagen) following manufacturers instructions.

PCR of whole genome amplified DNA was performed by Gary Wright, Great Ormond Street Hospital. The BIOMED-2 protocols to detect clonality were used [Dongen et al., 2003] and PCR fragments were excised and purified using a Thermo Fisher Scientific BigDye®Terminator v3.1 Cycle Sequencing Kit for Sanger sequencing.

2.12 Bulk RNA-sequencing

For fetal BM bulk RNA-Sequencing (2015 dataset): the SMART-Seq2 protocol was followed [Picelli et al., 2014]. 100 HSC or progenitors were flow sorted directly into lysis buffer and snap frozen over 16 sort experiments. A total of 4 normal (NM) biological samples were sorted, and a total of 5 trisomy 21 (T21) biological samples were sorted (summarised in table 2.3). Not all populations could be sorted from each sample and one NM biological sample was excluded from analysis as a normal karyotype

could not be confirmed. Samples were split into batches for cDNA synthesis and paired-end library preparation. Libraries were sequenced in 12 lanes on an Illumina HiSeq 2000 instrument (Wellcome Trust Centre for Human genetics). Sorting and library preparation was performed by Natalina Elliott and Gemma Buck.

Table 2.3: Summary of libraries sequenced for fetal BM RNA-Seq experiment. Numbers in brackets indicate the number of samples used for downstream analysis as one NM sample failed FISH analysis and so was excluded from downstream analysis.

	HSC	MPP	LMPP	ELP	PreProB	ProB	B
NM fetal BM	4 (3)	4 (3)	4 (3)	4 (3)	4 (3)	3	4 (3)
T21 fetal BM	3	3	3	0	2	5	5

For ALL blast bulk RNA-Sequencing (2017 dataset): we used the library preparation protocols used for TARGET-Seq [Rodriguez-Meira et al., 2019]. 100 normal fetal BM HSC or progenitors (used for controls) or patient blast cells (and HSPC where possible) were sorted (over 7 sort experiments) directly into lysis buffer and snap frozen. In total, blasts from 20 patients were sequenced alongside three biological replicates of normal fetal BM HSC and progenitors. RNA samples were split in to 4 batches for cDNA synthesis. Where possible two technical replicates for cDNA synthesis was performed and the best quality sample (according to quantification using Agilent Technologies Tapestation High Sensitivity tapes) was used for paired-end library preparation. Libraries were pooled and sequenced on 8 lanes on an Illumina NovaSeq S4 instrument (NovoGene, China). cDNA synthesis and library preparation was performed by Gemma Buck. Samples and populations that were sorted and sequenced are summarised in table 5.1.

Table 2.4: Summary of libraries sequenced in non-DS/DS-ALL RNA-Seq experiment

	Blasts	HSC	MPP	LMPP	ELP	CLP	PreProB	ProB	B cell
non-DS ETV6-RUNX1	4	1	4	2				3	4
DS ETV6-RUNX1	2		1	2				1	2
non-DS CRLF2+	3	1	3	3		2			3
DS CRLF2+	4	3	3	4			3	4	2
non-DS OTHER	3		3	3			1	1	3
DS OTHER	1		1	1			1	1	1
non-DS acquired T21	3		2	2			2	3	2
non-DS fetal BM		3	2	3	3	3	2	3	3

2.13 RNA-sequencing analysis

Prior to my secondment in OBDS (Oxford Biomedical Data Science, formerly known as CGAT), I used bash scripts to run QC, trimming, mapping and counting for RNA-seq data (fetal BM RNA-Seq and collaborator datasets). Fastq files were checked for quality control and trimmed using TrimGalore! v0.4.1 [`https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/`]; trimmed fastq files were mapped to the Ensembl GRCh38 assembly using STAR v2.4.2a [Dobin et al., 2013] and; mapped reads were counted using featureCounts in the SubRead v1.4.5 package [Liao et al., 2014]. To establish whether removing PCR duplicates from our dataset would skew our analysis I also removed duplicates using Picard Tools v1.8 (Broad Institute). For re-mapping of collaborators and published datasets, .bam files were converted to fastq using Samtools v1.3 [Li et al., 2009], [Li, 2011] and Bedtools v2.25.0 [`https://bedtools.readthedocs.io/en/latest/index.html`]. Then, these fastq files were mapped using STAR as described above.

During my secondment in OBDS I learnt to write my own pipelines to wrap up the bash scripts detailed above. These two pipelines, one for QC and trimming and a second for mapping are summarised in figures 2.5 and 2.6.

All further downstream analysis was performed in R v3.5.1 with approaches summarised in figure 2.7 and the packages used summarised in table 2.5. For gene set enrichment analysis (GSEA) permutation type was set to 'gene set' because there were fewer than 7 biological replicates for each phenotype and the default permutation number (1000) was used.

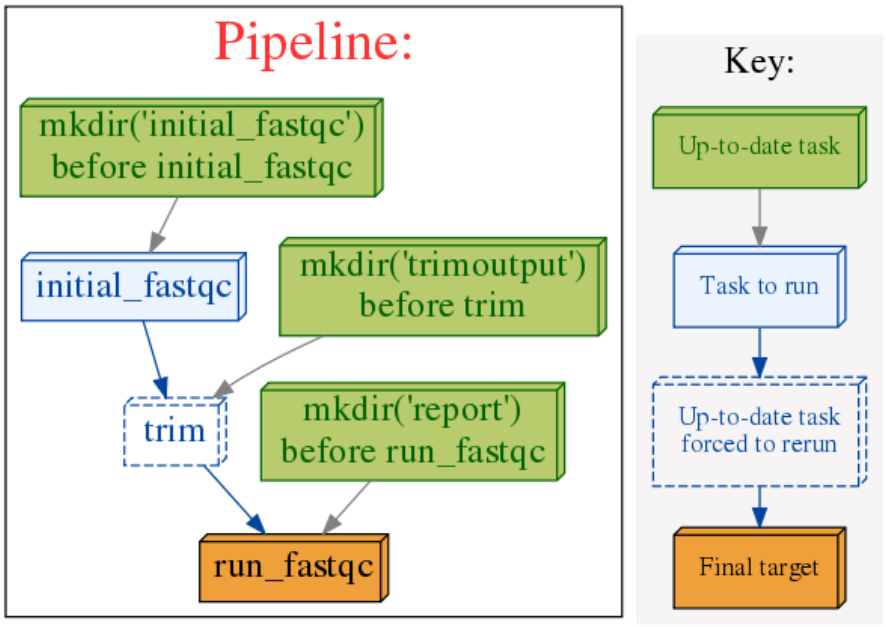


Figure 2.5: Pipeline written to QC check and trim fastq files for RNA-Sequencing data

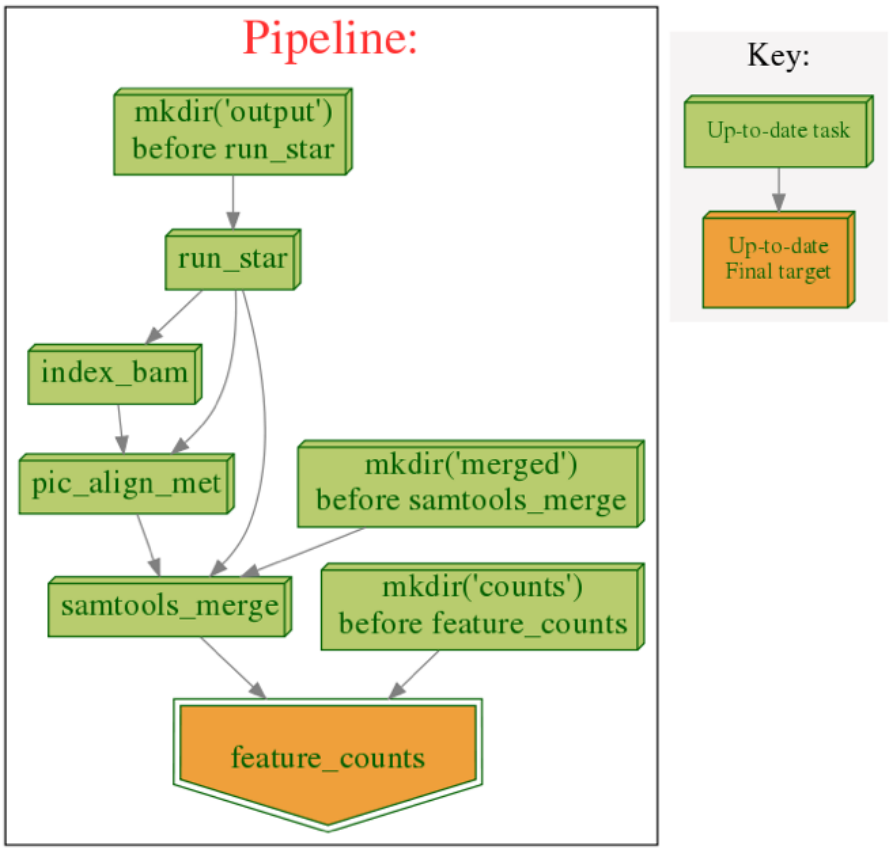


Figure 2.6: Pipeline written to map and count bulk RNA-Sequencing data

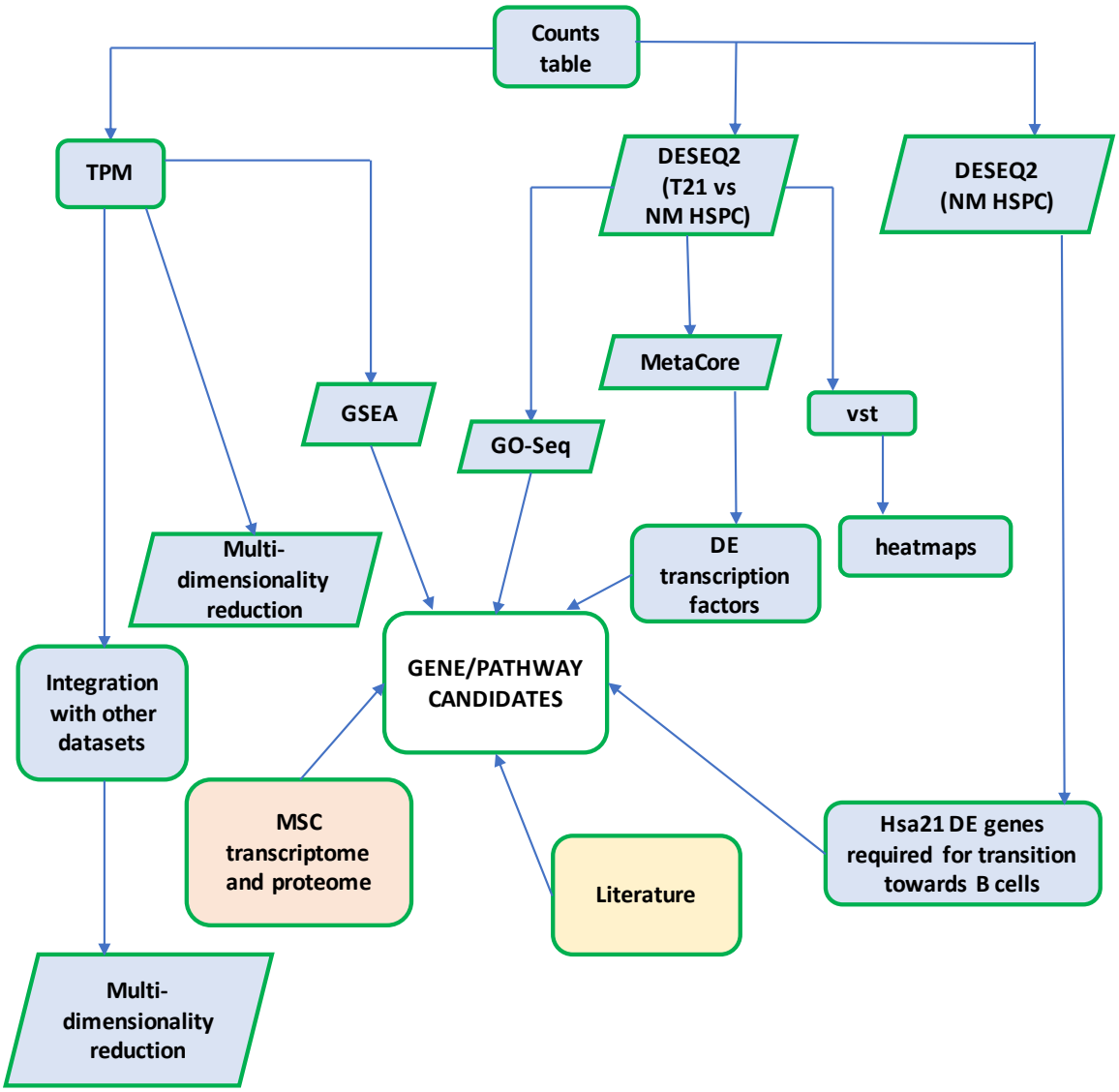


Figure 2.7: Scheme showing workflow used in RNA-Seq analysis and integration with other datasets

Table 2.5: Summary of packages used for RNA-Seq analysis

Step	CRAN/Bioconductor package or programme	Reference
Calculation of TPM	R base package	
Integration with other datasets (batch effect removal)	Limma	[Ritchie et al., 2015]
	RUVSeq	[Risso et al., 2014]
	SVA	DOI: 10.18129/B9.bioc.sva
Multi-dimensionality reduction	Principal component analysis: R base and ggplot2/plotly	
	Diffusion pseudotime: destiny	[Haghverdi et al., 2016]
	t-sne	
Differential expression analysis	DESEQ2 and visualisation in Pheatmap	[Love et al., 2014]
Gene set enrichment analysis	Broad institute graphical user interface	[Subramanian et al., 2005]
Identification of transcription factors	Metacore	Clarivate analytics
GO analysis	GO-Seq	[Young et al., 2010]
Exclusion of cell cycle genes	Seurat	[Butler et al., 2018]

Before running downstream analyses like differential expression or pathways analysis, I first performed principal component analysis (PCA) on the TPM tables to assess if there were any batch effects within the datasets. For the 2015 dataset, any "batch" effect was largely reflective of biological replicate (figure 2.8). Also, FastQC analysis suggested that PCR duplicate levels were high in the libraries sequenced. Artificially removing these duplicates is controversial as it could potentially also remove reads that reflect the high expression of certain genes ([Parekh et al., 2016]). To assess whether this would have an impact on our conclusions, I used Picard Tools v1.8 (Broad Institute) to remove PCR duplicates and performed some initial PCAs on the resulting TPM table. These PCAs (figure 2.8) and also differential analysis suggested that any effects of removing duplicates was minimal and unnecessary. Therefore, I chose to proceed in using the dataset where duplicates were not removed in my analysis.

I performed similar analysis on the 2017 DS-ALL dataset. Outliers in this considerably larger dataset necessitated using principal components 5 and 6 to look for the presence of batch effects. These PCAs showed that there was no batch effect based on sort data, sort stream, library preparation batch, or biological sample. Reassuringly the major cause for the two separate clusters was a result of immunophenotype where progenitors committed to the B-lineage and B-ALL blasts clustered separately from upstream progenitors (HSC, MPP, LMPP, ELP) (figure 2.9).

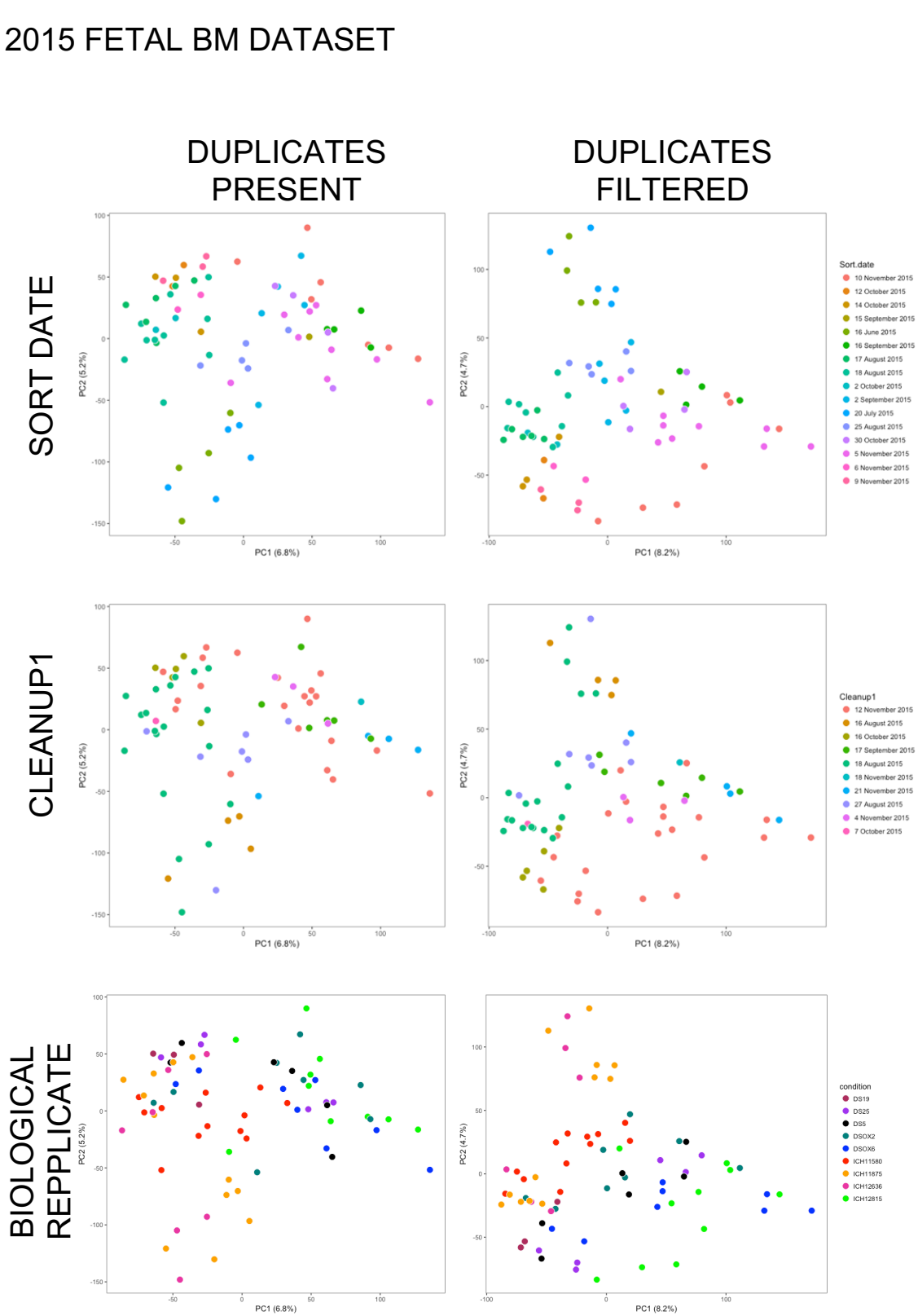


Figure 2.8: Initial analysis of 2015 Fetal BM RNA-Seq dataset to check for batch effects.

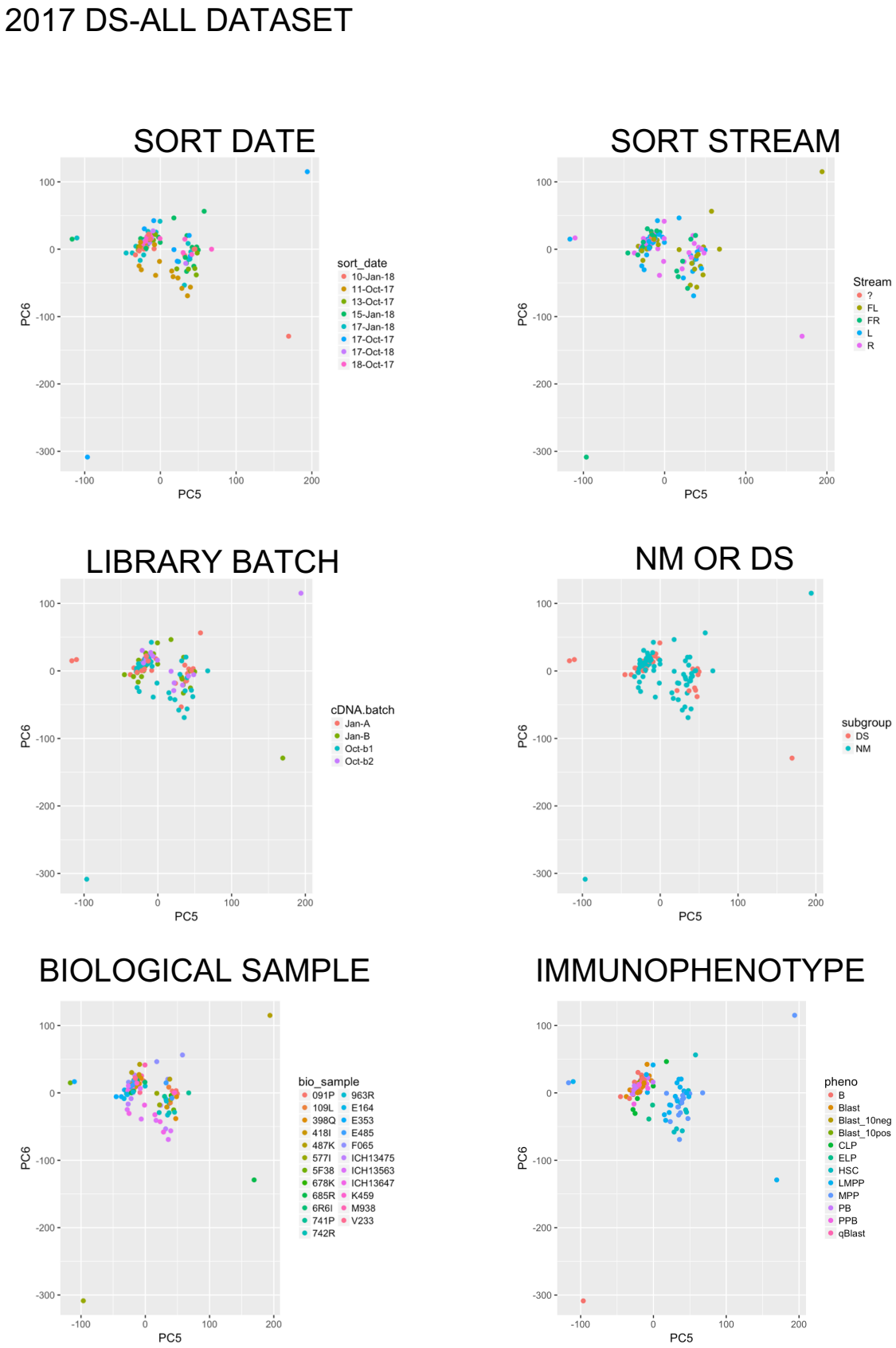


Figure 2.9: Initial analysis of 2017 DS-ALL RNA-Seq dataset to check for batch effects. NM: normal; DS: Down Syndrome.

2.14 Single cell RT-qPCR

Single cell RT-qPCR was performed using the Biomark HD microfluidics system (Fluidigm) which facilitates simultaneous quantification of expression of 96 genes from a single cell. Using this technique two approaches were taken. Briefly, HSC and progenitors were single cell index sorted into a 96 well plate containing pre-amplification mix (all 96 probes, tables 2.7 and 2.8; CellsDirect One-Step qRT-PCR reaction mix and Platinum Taq, Invitrogen; SUPERase-In RNase Inhibitor, Ambion; Low EDTA TE). Immediately following sorting, pre-amplification was performed according to table 2.6. Following pre-amplification, qPCR was performed using the Fluidigm/Biomark system with a 96.96 Dynamic Array as per manufacturers instructions. Dr Natalina Elliott operated the microfluidics system and performed the qPCR.

Two different sorting approaches were used in combination with this technique. Firstly, the whole Lin2-CD34+ compartment was index sorted to achieve an unbiased agnostic approach to interrogating the B-lymphoid hierarchy in fetal BM. In this experiment genes detailed in table 2.7 were chosen for their known involvement in specific haematopoietic pathways (lymphoid, myeloid or erythroid). To avoid under-representation of rare stem cell subsets (HSC/MPP, LMPP, ELP) these were selectively sorted at the end such that a total of 1100 cells were sorted. Secondly, in a separate set of experiments, all subsets hypothesised to be involved in the B-lymphoid hierarchy were selectively index sorted for single cell qRT-PCR with a second panel of genes chosen for their specific involvement in lymphopoiesis and childhood ALL (detailed in table 2.8).

Table 2.6: Pre-amplification PCR for single cell qRT-PCR. Set lid temperature to 105°C

Step	Temperature	Time	Cycle number
Reverse transcription	50°C	15min	1
Inactivation of RTase/ activation of Taq	95°C	2min	1
Target amplification	95°C	15sec	20
	60°C	4min	
Hold	4°C	hold	infinity

Table 2.7: Genes and TaqMan assay ID's used for agnostic approach. * Denotes failed probe.

Gene Name	TaqMan Assay ID
ACTB	Hs01060665_g1
APOC1	Hs00155790_m1
AZU1	Hs00156049_m1
B2M	Hs00984230_m1
BAALC	Hs00227249_m1
BCL11A	Hs01093197_m1
BCL2	Hs00608023_m1
BMP2	Hs00154192_m1
CCNA2	Hs00996788_m1
CCNB1	Hs01030099_m1
CCNB2	Hs01084593_g1
CCND3	Hs01017690_g1
CCNE2	Hs00180319_m1
CD19	Hs01047413_g1
CD24	Hs02379687_s1
CD27	Hs00609654_g1
CD34	Hs02576480_m1
CD38	Hs01120071_m1
CD44	Hs01075864_m1
CD52	Hs00174349_m1
CD72	Hs00960066_m1
CD79A	Hs00998119_m1
CD9	Hs01124022_m1
CD96	Hs00976975_m1
CDK6	Hs01026371_m1
CDKN2A	Hs00923894_m1
CFD	Hs00197387_m1
CNRIP1	Hs01547787_m1
CRLF2	Hs00845692_m1
CSF1R	Hs00911250_m1
CSF2RB	Hs00166144_m1
CSF3R	Hs00167918_m1
CTSG*	Hs00175195_m1
CXCR4	Hs00607978_s1
DNTT	Hs00172743_m1
DYRK1A	Hs00176369_m1
EBF1	Hs01092695_m1
ELANE	Hs00236952_m1
ELL3	Hs00228559_m1
EPOR	Hs00959427_m1

Table 2.7: Genes and TaqMan assay ID's used for agnostic approach. * Denotes failed probe.

Gene Name	TaqMan Assay ID
ERG	Hs01554629_m1
ETS2	Hs00232009_m1
FLI1	Hs00956711_m1
FLT3	Hs00174690_m1
FUT4/CD15	Hs01106466_s1
GAPDH	Hs02758991_g1
GATA1ex4/5	Hs01085823_m1
GFI1	Hs00382207_m1
GFI1B	Hs01062469_m1
HBD	Hs00426283_m1
HIST1H1B	Hs00271207_s1
HLF	Hs00171406_m1
HOXA10	Hs00172012_m1
HOXA3	Hs00601076_m1
HOXA5	Hs00430330_m1
HOXA7	Hs00600844_m1
IGF2BP3	Hs00559907_g1
IGFBP7	Hs00266026_m1
IGHM	Hs00941538_g1
IKZF1	Hs00958474_m1
IL3RA	Hs00608141_m1
IL7R	Hs00233682_m1
ITGA2B	Hs01116228_m1
JAK2	Hs01078136_m1
JCHAIN	Hs00376160_m1
KIT	Hs00174029_m1
KLF1	Hs00610592_m1
LEF1	Hs01547250_m1
LIN28B	Hs01013729_m1
LYZ	Hs00426232_m1
MEIS1	Hs01017441_m1
MKI67	Hs04260396_g1
MME	Hs00153510_m1
MPO	Hs00924296_m1
MS4A1	Hs00544819_m1
MX1	Hs00895608_m1
MX2	Hs01550814_m1
MYB	Hs00920556_m1
MYC	Hs00153408_m1
PAX5	Hs00277134_m1

Table 2.7: Genes and TaqMan assay ID's used for agnostic approach. * Denotes failed probe.

Gene Name	TaqMan Assay ID
PF4	Hs00427220_g1
PROCR	Hs00157263_m1
PROM1	Hs01009259_m1
PRTN3	Hs01553330_m1
PTPRC	Hs04189704_m1
RAG1	Hs01920694_s1
RAG2	Hs01851142_s1
RUNX1	Hs01021970_m1
RUNX2	Hs01047973_m1
SPINK2	Hs01598293_m1
STAT5A	Hs00234181_m1
TCF3	Hs01012685_m1
TCL1A	Hs00951350_m1
TFR2	Hs01056398_m1
THY1	Hs00264235_s1
VPREB1	Hs00356766_g1

Table 2.8: Genes and TaqMan probe IDs used in lymphoid single cell qRT-PCR. * Denotes failed probe

Gene Name	TaqMan Assay ID
ACTB	Hs01060665_g1
AFF1	Hs01014712_m1
B2M	Hs00984230_m1
BAALC	Hs00227249_m1
BAZ2B	Hs01002426_m1
BCAT1	Hs00398962_m1
BCL11A	Hs01093197_m1
BCL2	Hs00608023_m1
BMP2	Hs00154192_m1
CAMK2D	Hs00943538_m1
CCNA2	Hs00996788_m1
CCNB1	Hs01030099_m1
CCNB2	Hs01084593_g1
CCND3	Hs01017690_g1
CCNE2	Hs00180319_m1
CD19	Hs01047413_g1
CD24	Hs02379687_s1
CD27	Hs00609654_g1
CD34	Hs02576480_m1

Table 2.8: Genes and TaqMan probe IDs used in lymphoid single cell qRT-PCR. * Denotes failed probe

Gene Name	TaqMan Assay ID
CD38	Hs01120071_m1
CD44	Hs01075864_m1
CD52	Hs00174349_m1
CD72	Hs00960066_m1
CD79A	Hs00998119_m1
CD9	Hs01124022_m1
CD96	Hs00976975_m1
CDK6	Hs01026371_m1
CDKN2A	Hs00923894_m1
CRLF2	Hs00845692_m1
CSPG4*	Hs00361541_g1
DNTT	Hs00172743_m1
DOT1L	Hs01579928_m1
DYRK1A	Hs00176369_m1
EBF1	Hs01092695_m1
ELL3	Hs00228559_m1
EPOR	Hs00959427_m1
ERG	Hs01554629_m1
ETS2	Hs00232009_m1
FLI1	Hs00956711_m1
FLT3	Hs00174690_m1
FUT4	Hs01106466_s1
GAPDH	Hs02758991_g1
GFI1	Hs00382207_m1
GFI1B	Hs01062469_m1
HIST1H1B	Hs00271207_s1
HOXA10	Hs00172012_m1
HOXA3	Hs00601076_m1
HOXA5	Hs00430330_m1
HOXA7	Hs00600844_m1
HOXA9*	Hs00266821_m1
IGFBP7	Hs00266026_m1
IKZF1	Hs00958474_m1
IL7R	Hs00233682_m1
IRX1	Hs00411782_m1
JAK2	Hs01078136_m1
KCTD15	Hs00225337_m1
KIT	Hs00174029_m1
KLRK1	Hs00183683_m1
KMT2A	Hs00610538_m1

Table 2.8: Genes and TaqMan probe IDs used in lymphoid single cell qRT-PCR. * Denotes failed probe

Gene Name	TaqMan Assay ID
KRAS	Hs00364284_g1
LEF1	Hs01547250_m1
LIN28B	Hs01013729_m1
MEF2C	Hs00231149_m1
MEIS1	Hs01017441_m1
MKI67	Hs04260396_g1
MME	Hs00153510_m1
MPO	Hs00924296_m1
MS4A1	Hs00544819_m1
MX1	Hs00895608_m1
MX2	Hs01550814_m1
MYB	Hs00920556_m1
MYC	Hs00153408_m1
NPM1	Hs02339479_g1
NPM3	Hs00199625_m1
NRAS	Hs00180035_m1
PAX5	Hs00277134_m1
PDGFA	Hs00234994_m1
PPP1R14A	Hs00264434_m1
PRDX1	Hs00602020_mH
PRDX4	Hs01056076_m1
PROM1	Hs01009259_m1
PTPRC	Hs04189704_m1
RAF1	Hs00234119_m1
RAG1	Hs01920694_s1
RAG2	Hs01851142_s1
RFC3	Hs01082404_m1
RFC4	Hs00427469_m1
RUNX1	Hs01021970_m1
RUNX2	Hs01047973_m1
RUVBL2	Hs00272632_m1
S100A8	Hs00374264_g1
SOCS2	Hs00919620_m1
SPINK2	Hs01598293_m1
STAT5A	Hs00234181_m1
TCF3	Hs01012685_m1
TCL1A	Hs00951350_m1
VPREB1	Hs00356766_g1

2.15 Single cell RT-qPCR analysis

Ct values were normalised to the mean average Ct value of three housekeeping genes: *ACTB*, *B2M* and *GAPDH* and then transformed ($2^{-\Delta CT}$). Downstream analysis was performed in R v3.5.1. To check that biological samples were all comparable, initial PCAs were plotted and coloured by sample. In the fetal BM single cell RT-qPCR, all samples overlapped (figure 2.10 a). In the adult BM single cell RT-qPCR, initial inspection of the PCA would suggest there is some clustering according to biological sample (figure 2.10 b). With RT-qPCR such such effects are unexpected because the measurement of gene expression is targeted and upon further investigation it became clear that this separation was a result of the B cells that were sorted from one sample only (figure 2.10 c). Dot plots and violin plots were plotted using ggplot2 (v3.1.0). Principal component analysis was visualised in 2D with ggplot2 and 3D with plotly(v4.8.0). Diffusion maps and pseudotime analysis were performed using Destiny (v2.12.0) [Haghverdi et al., 2016] and Monocle (v2.10.1) [Trapnell et al., 2014].

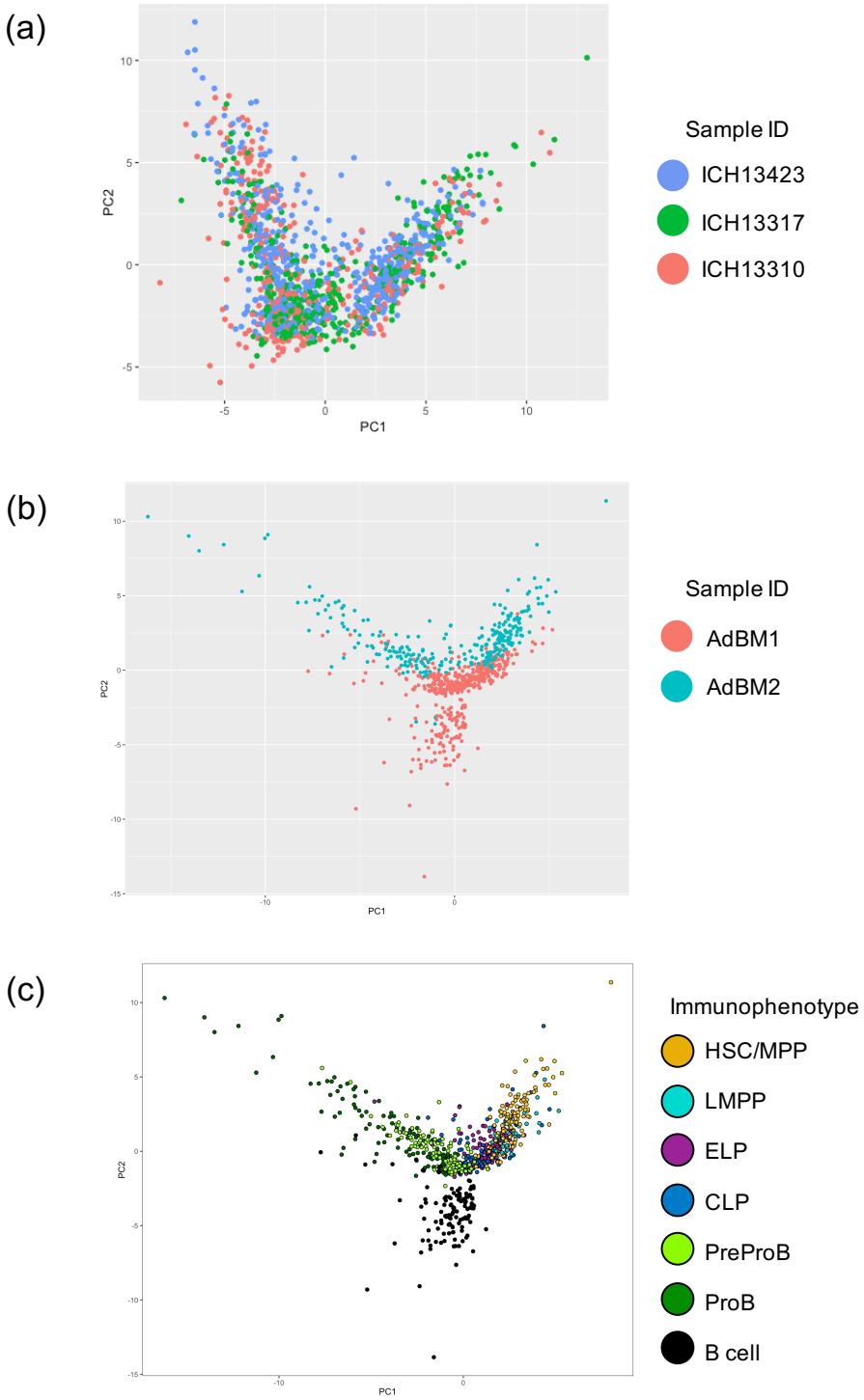


Figure 2.10: RT-qPCR representative PCA plots demonstrating that biological replicates in each single cell RT-qPCR experiment were comparable. (a) PCA of three samples used for fetal BM single cell RT-qPCR; (b) PCA of two samples used for adult BM single cell RT-qPCR; (c) PCA of adult BM single cell RT-qPCR coloured by immunophenotype demonstrating that clustering of samples is a result of B cells sorted from one sample

2.16 ATAC-Sequencing

The protocol for ATAC-Sequencing was adapted for small cell number from [Corces et al., 2016]. 2000 PreProB or ProB progenitors (n=3) were sorted directly into 25μL of transposase mix (Tn5 enzyme, 2X TD buffer and 0.01% digitonin) at 4°C. Immediately following sorting, samples were incubated at 37°C for 30 minutes at 300 rpm agitation. Transposed DNA was immediately purified using a Qiagen MinElute kit and eluted in 10μL pre-warmed EB. Transposed fragments were amplified by PCR as previously described [Buenrostro et al., 2015]. Libraries were then quantified by qPCR (NEBNext Library Quant Kit for Illumina) before being sequenced on an Illumina NextSeq instrument using 75 cycles for paired end sequencing. Many thanks to Gemma Buck and Catherine Garnett for their help and expertise in guiding me with these protocols.

2.17 ATAC-sequencing analysis

A summary of the ATAC-Sequencing analysis performed is detailed in figure 2.11. Fastq files were trimmed and mapped to the UCSC hg19 assembly using an in-house pipeline written by Jelena Telenius and described in Hay *et al.* [Hay et al 2016]. Blacklisted regions and mitochondrial reads were excluded before normalisation by scaling to 1x10^6 reads and visualisation in the UCSC genome browser [ref: Kent et al 2002]. MACS2 [Feng et al., 2012] was used to call peaks (FDR <0.05). Differential accessibility was investigated using DiffBind [Ross-Innes et al., 2012] which is a wrap-around R package for ChIP-Seq data. I performed differential analysis using DESEQ2, DESEQ and EDGER for comparison. Peaks were arbitrarily assigned genes according to distance to transcription start site using HOMER [Heinz et al., 2010]. Enrichment and over-representation analysis was performed using multiple online platforms: DAVID, CPDB and GO, powered by PANTHER [Mi et al., 2013]. Heatmaps for various gene sets were made using Deeptools v3.0.1 [Ramírez et al., 2016] using both the scale-regions and reference-point options. The scale-regions option is probably the most relevant for ATAC-sequencing data because peaks are not corresponding to binding sites of a specific protein or complex (e.g. ChIP data) and are actually representative of open chromatin where any activating or repressive complex could be binding. Overlaps of differentially accessible genes with RNA-sequencing differential analysis were calculated in R.

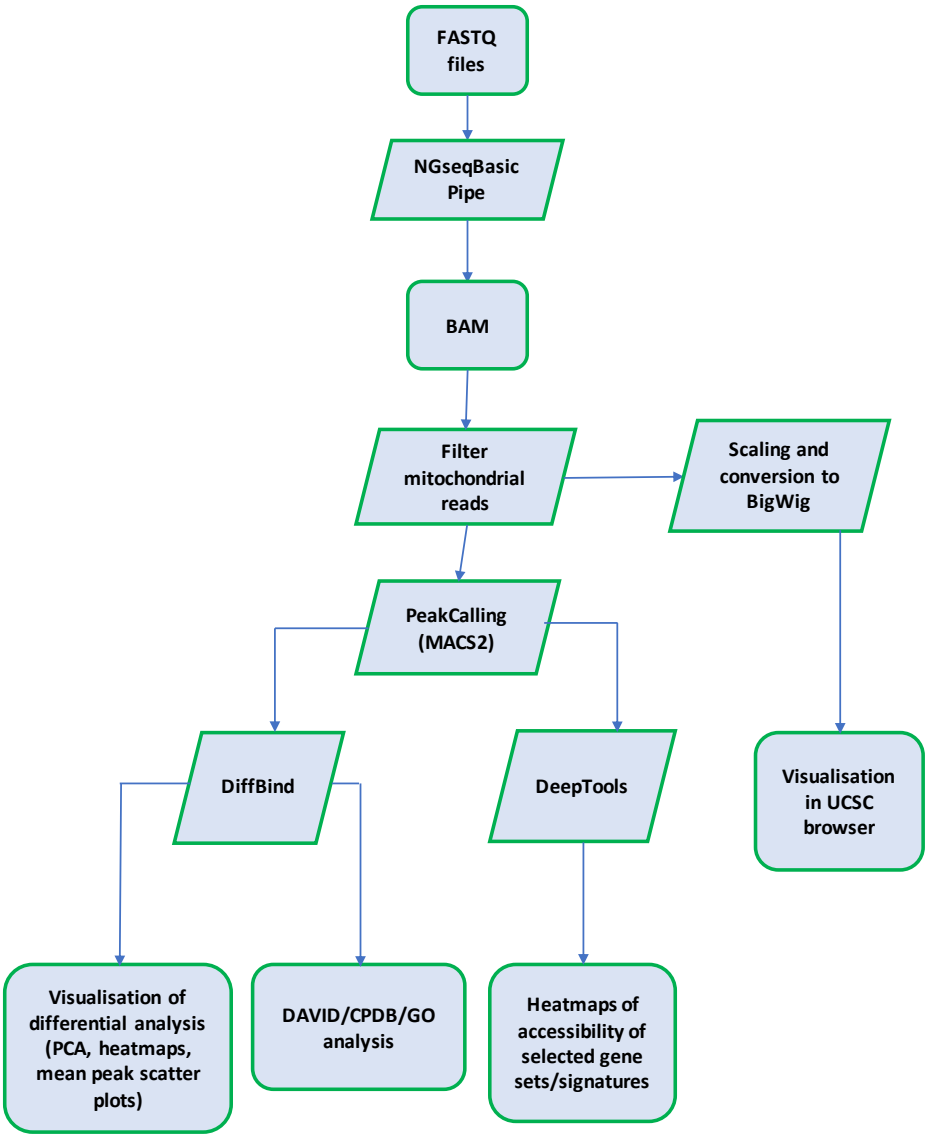


Figure 2.11: Scheme for ATAC-Seq analysis

2.18 Proteomics

Protein fractions from the membrane of MSCs and also secreted proteins were isolated for identification by mass spectrometry. At 90% confluence, cells were scraped and membrane proteins isolated using a memPER membrane extraction kit (ThermoScientific). Secreted protein was harvested from serum free media that had been conditioned for 48 hours by fully confluent MSC. All protein samples were stored at -80 °C and concentrations determined using an Invitrogen qubit kit for proteinl and by Coomassie stain after polyacrylamide gel electrophoresis (PAGE). These protein samples were then submitted to the Advanced Proteomics Facility, Target Discovery Institute, Oxford, for mass spectrometry analysis. These experiments were performed by Natalina Elliott and Ella-Mae Labbett.

2.19 Proteomics analysis

A summary of the workflow used for analysis of the MSC proteomic data is detailed in figure 2.12. Identification and quantification was performed using the Progenesis software by the Advanced Proteomics Facility, TDI Mass Spectrometry Lab, Oxford. Downstream analysis was performed in R where multiple methods of normalisation were explored (figure 2.13). Differential expression analysis was performed in Limma using both VSN normalised and median normalised matrices. Volcano plots were plotted using ggplot2.

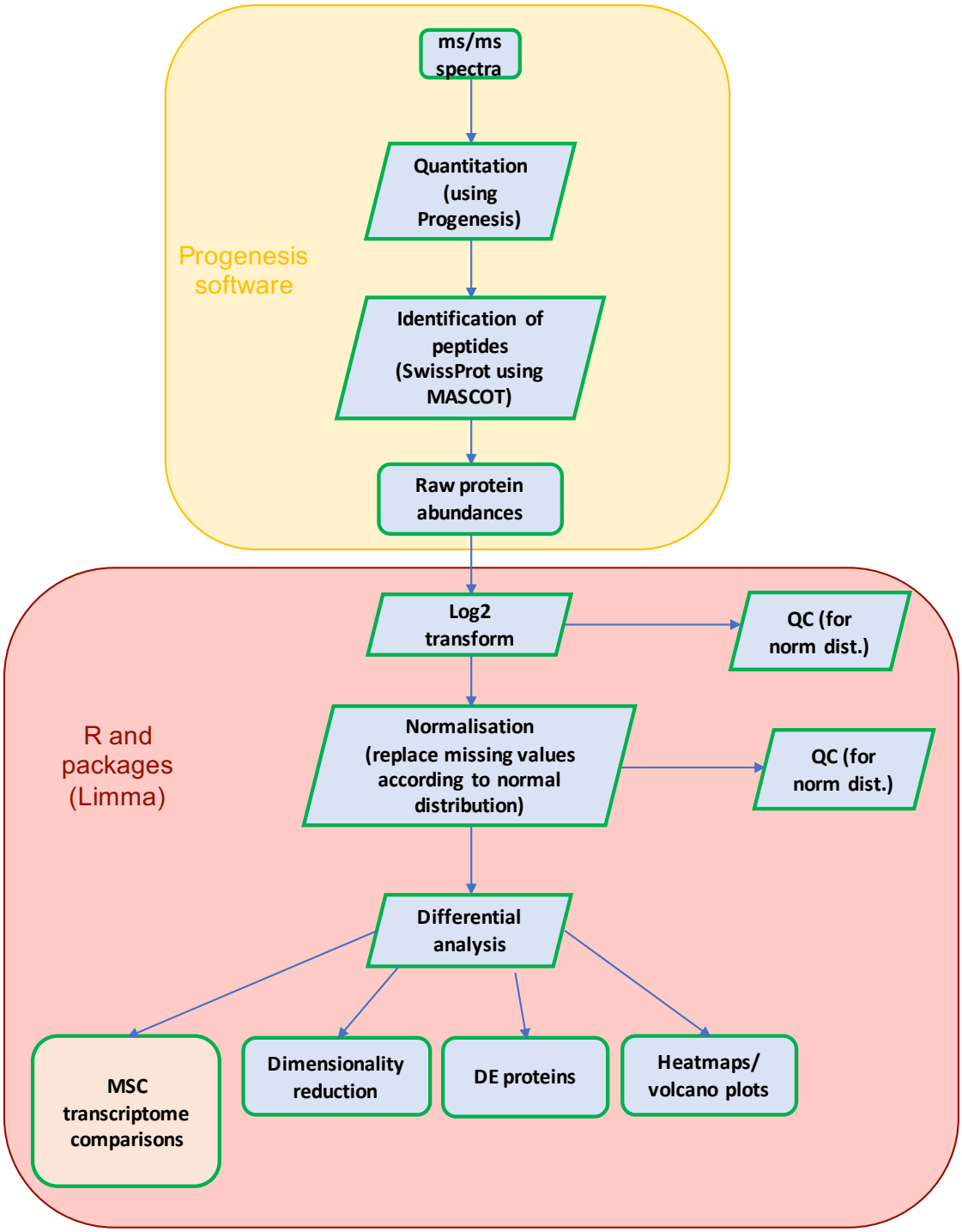


Figure 2.12: Scheme for proteomics analysis

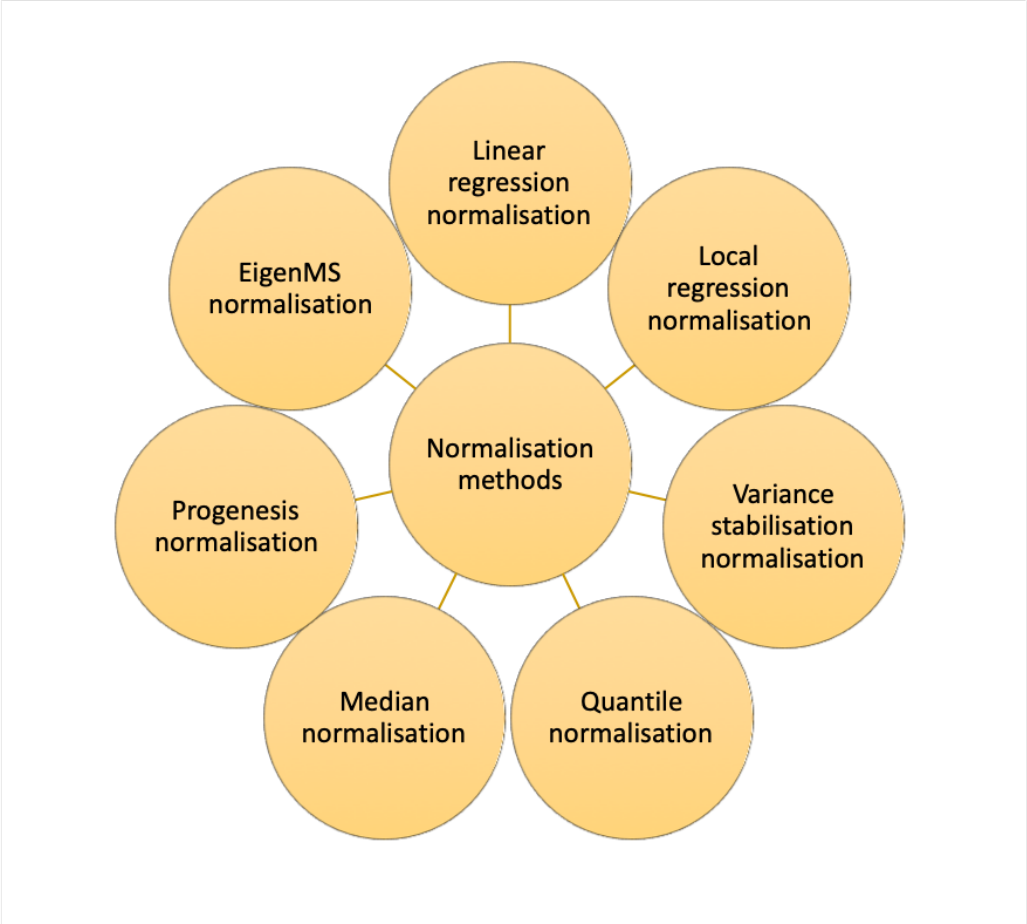


Figure 2.13: Multiple methods of normalisation as detailed in [Välikangas et al., 2016]

2.20 Microarray analysis

Thank you to Steve Taylor for helping me with this analysis by providing example code and advising me with the downstream analysis. Thank you to David O'Connor for providing the data.

A summary of the analysis performed is illustrated in figure 2.14. .CEL files were normalised by Robust Multi-array Average (RMA) using the Affymetrix power tool "apt-probeset-summarize". Downstream analysis was performed in R using Limma [Ritchie et al., 2015] where control probes and genes with no variance across samples were excluded. I tested differential expression with lowly expressed genes at three cut off points: <3, <5 and <10. After plotting raw counts across each sample as a box plot I decided a cut off <5 was most appropriate as this was roughly the value of the lower quartile (25th percentile). Differential expression analysis was performed using Limma; heatmaps were plotted using pHeatmap and; Gene Set Enrichment Analysis (GSEA) was performed using the Broad Institute's graphical user interface [Subramanian et al., 2005]. The settings for GSEA, permutation type was set to 'gene set' because there were fewer than 7 biological replicates for each phenotype and the default permutation number (1000) was used.

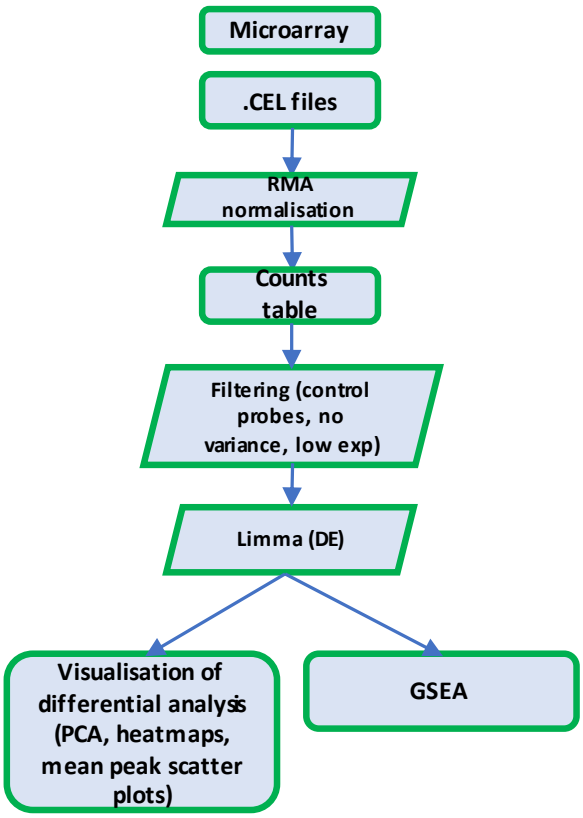


Figure 2.14: Scheme for microarray analysis

2.21 RT-qPCR

qPCR for measuring gene expression (RT-qPCR) in MSC was carried out in bulk. Cultured MSC were trypsinised and snap frozen in pellets of 1-3x10^6 cells. mRNA was extracted using the Qiashredder columns and Qiagen RNeasy kit; reverse transcription was performed using Superscript III first strand synthesis (Invitrogen); and qPCR was performed using TaqMan probes on the QuantStudio3 real-time PCR system (ThermoFisher Scientific). ΔCt values were calculated by subtracting the mean Ct value of housekeeping gene, *GAPDH*, from the test gene. ΔCt values were transformed ($2^{-\Delta CT}$) and plotted in GraphPad prism 7.

Since miRNAs are small and are not poly-adenylated, we used the miRNeasy kit from Qiagen to isolate total RNA. From the total RNA, we used the TaqMan advanced miRNA cDNA synthesis kit (ThermoFisher Scientific) to perform the poly-A tailing and reverse transcription steps. qPCR from the cDNA stage was performed as described above. The housekeeping gene used to normalise miRNA expression was miRNA-423-5p.

Details of all TaqMan probes used are described in table 2.9. With thanks to Jana Oswald (visiting student) and Hannah Fuchs (FHS student) with their help with these experiments.

2.22 Enzyme-linked immunosorbent assays (ELISAs)

0.04x10^6 MSC were plated per well in a 24 well plate in 500μL of αMEM-10. After 3 days, conditioned media was collected and snap frozen in an ethanol bath. DuoSet ELISA kits for IL-6 and IFNα2 were purchased from R&D Systems and performed according to manufacturers instructions on thawed conditioned media. Where concentrations measured exceeded the standard curve, ELISAs were repeated with diluted conditioned media and final concentration was calculated based on this dilution. Data is plotted in GraphPad prism. With thanks to Jana Oswald (visiting student) and Hannah Fuchs (FHS student) for their help with these experiments.

2.23 Statistics

Where data are shown as bar charts, they are expressed as the mean $\pm$ standard error of mean (SEM) unless otherwise stated. Where statistical significance is shown, a

paired or un-paired t-test was performed depending on whether the sample groups tested were from within biological replicates or between biological replicates. Where gene expression (RT-qPCR) across multiple cell sets was tested, ANOVA was used.

Table 2.9: Genes and TaqMan assay IDs used for bulk RT-qPCR of MSC

Gene name	TaqMan Assay ID
hsa-let-7c-3p	479365_mir
hsa-let-7c-5p	478577_mir
hsa-miR-99a-5p	478519_mir
hsa-miR-99a-3p	479224_mir
hsa-miR-125b-2-3p	478666_mir
hsa-miR-125b-5p	477885_mir
hsa-miR-155-3p	477926_mir
hsa-miR-155-5p	483064_mir
hsa-miR-423-5p	478090_mir
GAPDH	Hs99999905_m1
IL6	Hs00174131_m1
IFNAR1	Hs01066116_m1
IFNAR2	Hs01022060_m1
IFNA1	Hs00256882_s1
IFNA2	Hs00265051_s1
IFNA4	Hs01681284_sH
IFNA5	Hs04186137_sH
IFNA6	Hs00819627_s1
IFNA7	Hs01652729_s1
IFNA8	Hs00266883_s1
IFNA14	Hs00353663_s1
IFNA16	Hs03005057_sH
IFNA17	Hs00819693_sH
IL2	Hs00174114_m1
TGFB1	Hs00998133_m1
TGFB2	Hs00234244_m1
TGFB3	Hs01086000_m1

Chapter 3

The cellular and molecular characterisation of B-lymphoid development in normal fetal bone marrow

3.1 Background and aim

The developmental hierarchy for human fetal B-lymphopoiesis is poorly defined. Most of our knowledge of the dynamics of B-lymphopoiesis in the developing human fetus is inferred from murine studies and investigations on CB. Where human fetal tissues have been studied, the presence of B-lineage cells have been reported in FL and fetal BM but all of their functional and molecular properties have not been systematically characterised [Asma et al., 1984], [Nadler et al., 1984], [Uckun and Ledbetter, 1988], [Uckun, 1990], [Grumayer et al., 1991], [Nuñez et al., 1996], [Charbord et al., 1996], [Bueno et al., 2016].

Previous work investigating second trimester FL in the lab had noted the presence of a CD19-CD127+ progenitor population and a committed B progenitor that differed to adult ProB progenitors [Roy et al., 2012]. Functional analysis suggested that the CD19-CD127+ population had multi-lymphoid output with minimal myeloid potential. These progenitors were named Early Lymphoid Progenitors (ELPs) in line with a previous study identifying an immunophenotypically similar population in adult BM [Ryan et al., 1997]. More recently, other investigators have identified similar CD19-CD127+ progenitors in FL [Boiers et al., 2018], [Alhaj Hussen et al., 2017].

Previous work had also identified two types of committed B progenitors in human

FL: 'adult type' ProB progenitors that were CD34+CD19+CD10+ and a second B progenitor identified differed from adult ProB progenitors in that it was CD10-: CD34+CD19+CD10- [Roy et al., 2012]. Immunophenotypically identical CD34+CD19+ CD10- committed B progenitors had previously been described in CB; and hypothesising that these progenitors lay upstream of ProB progenitors, the CD34+CD19+CD10- progenitors were named PreProB progenitors [Sanz et al., 2010]. This group went some way towards proving this by performing *in vitro* cell division assays and some RT-qPCR of selected B cell genes. However the differentiation potential of PreProB progenitors was not comprehensively interrogated and gene expression analysis was limited to 6 genes. Furthermore, all these experiments were performed on CB PreProB progenitors and there was no comparison between CB and fetal tissues.

The aim of the experiments described in this chapter, was to systematically characterise the cellular hierarchy of B-lymphoid development in normal fetal BM using functional and molecular assays.

3.2 Immunophenotypic characterisation of fetal BM and comparison with FL

Previous published and unpublished data from the lab focussing on human FL B-lymphopoiesis lead us to postulate a hierarchy in fetal BM that differed from the established pathway in adult BM (figure 3.1). The main difference between FL and adult BM B-lymphopoiesis was the failure to identify a functional Common Lymphoid Progenitor in FL. The previously reported IL7 receptor (CD127) expressing population in adult BM [Ryan et al., 1997] appeared to be enriched in FL along with the presence of a fetal specific B progenitor which was called PreProB progenitors based on the assumption that they lay upstream of ProB progenitors in the B-lymphoid hierarchy.

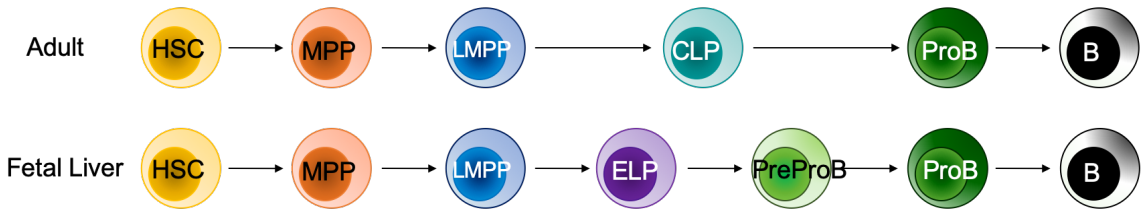


Figure 3.1: Scheme showing the known or proposed hierarchy of B-lymphopoiesis in human adult BM and FL respectively.

Table 3.1: Immunophenotypic definitions

Population	**Immunophenotype**
HSC (Haematopoietic Stem Cell)	Lin-CD34hiCD38-CD45RA-CD90+
MPP (Multipotent Progenitor)	Lin-CD34hiCD38-CD45RA-CD90-
LMPP (Lymphoid-primed Multipotent Progenitor)	Lin-CD34hiCD38-CD45RA+
ELP (Early Lymphoid Progenitor)	Lin2-CD34+CD19-CD10-CD45RA+CD127+
PreProB-progenitor	Lin2-CD34+CD19+CD10-
ProB-progenitor	Lin2-CD34+CD19+CD10+
Lin: Lineage cocktail containing: CD2/3/14/16/19/56/235a	
Lin2: Lineage cocktail containing: CD2/3/14/16/56/235a	

Investigations into human fetal haematopoiesis have largely been confined to the study of FL and while the FL is known to remain the main site of haematopoiesis until birth, in post-natal life the primary site of B-lymphopoiesis is the BM. I therefore set out to compare the emergence of B cells and their progenitors in gestation age matched FL and fetal BM. HSC, MPP and LMPP were defined by their conventional immunophenotypic markers and ELP, PreProB and ProB progenitors were defined as summarised in table 3.1. To reliably identify all progenitors by their immunophenotypes, I split each sample and ran two separate flow cytometry panels: one to identify HSC, MPP and LMPP and a second to identify ELP, PreProB, ProB and B cells (gating strategies show in figure 3.2 and figure 3.3). In later experiments I optimised these panels to include gating of ELP (by addition of anti-CD127 antibody) with the HSC, MPP and LMPP. This allowed for more efficient sorting in other experiments. Where sample cell numbers were very small (particularly in earlier gestational ages), I chose to run the latter of the two flow cytometry panels which enabled simultaneous identification of all B progenitors and HSC/MPP together (no anti-CD90 antibody in the panel) while also ruling out non-haematopoietic lineages with the addition of human CD45 to the panel. For this reason, in the analyses below, HSC and MPP are grouped together.

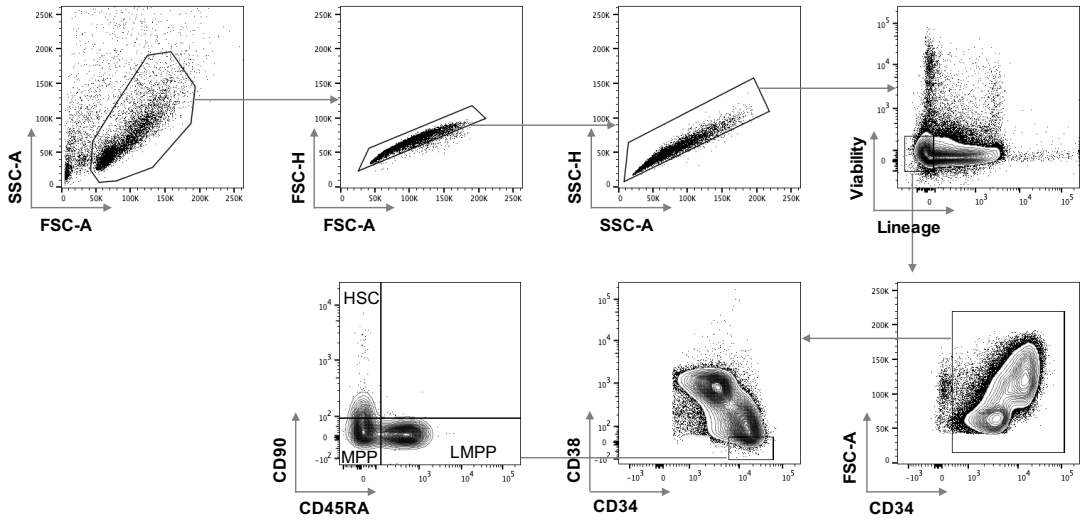


Figure 3.2: Gating strategy used for analysis of HSC, MPP and LMPP frequency in fetal BM

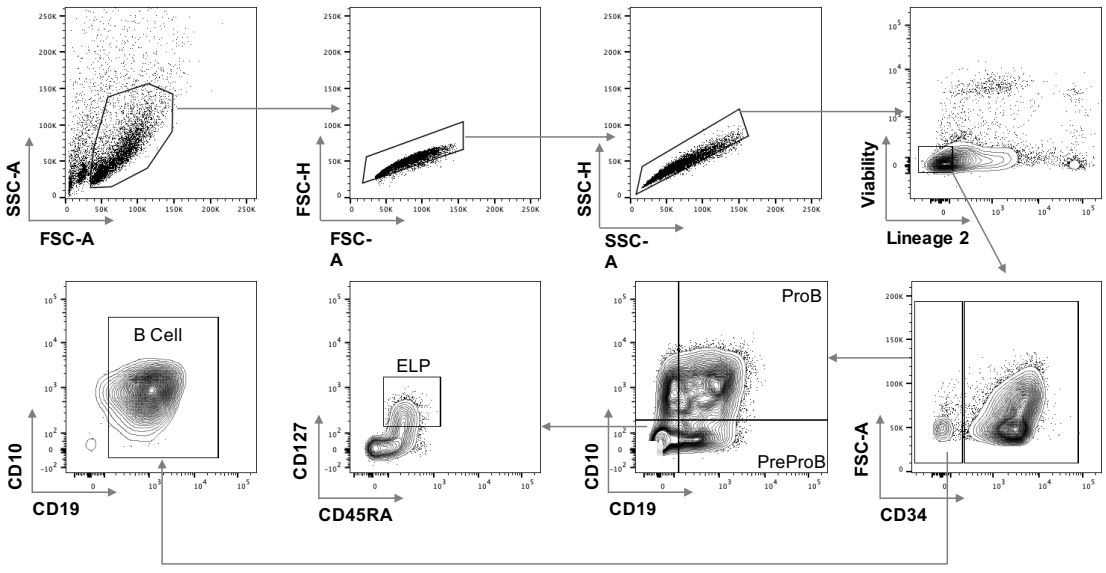


Figure 3.3: Gating strategy used for analysing the frequency of ELP, PreProB and ProB progenitors in fetal BM

In the FL, the frequencies of each population remained relatively stable throughout the second half of the first trimester and for most of second trimester (figure 3.4). The presence of haematopoietic populations in the fetal BM over this same time period was more dynamic. In line with previous immunohistochemical studies [Charbord et al., 1996], haematopoietic progenitors were not detected in the fetal BM until the beginning of second trimester (11 pcw). While HSC/MPP always remained at a higher fraction in the FL, LMPP and ELP were found at similar levels in both tissues. More striking differences were observed in the B progenitor compartment where PreProB progenitor frequencies spiked at 11 pcw in fetal BM and then dropped to a relatively stable frequency, which remained consistently higher than that in FL; ProB progenitor frequencies steadily rose from 11 pcw right through second trimester; and B cell frequencies rose sharply during mid-second trimester (figure 3.4). The emergence of HSC/MPP, LMPP, ELP and B progenitors at the same time in the fetal BM suggests that PreProB progenitors have travelled from the FL alongside other progenitors (HSC/MPP, LMPP, ProB) in an initial seeding of the BM, followed by a rapid expansion.

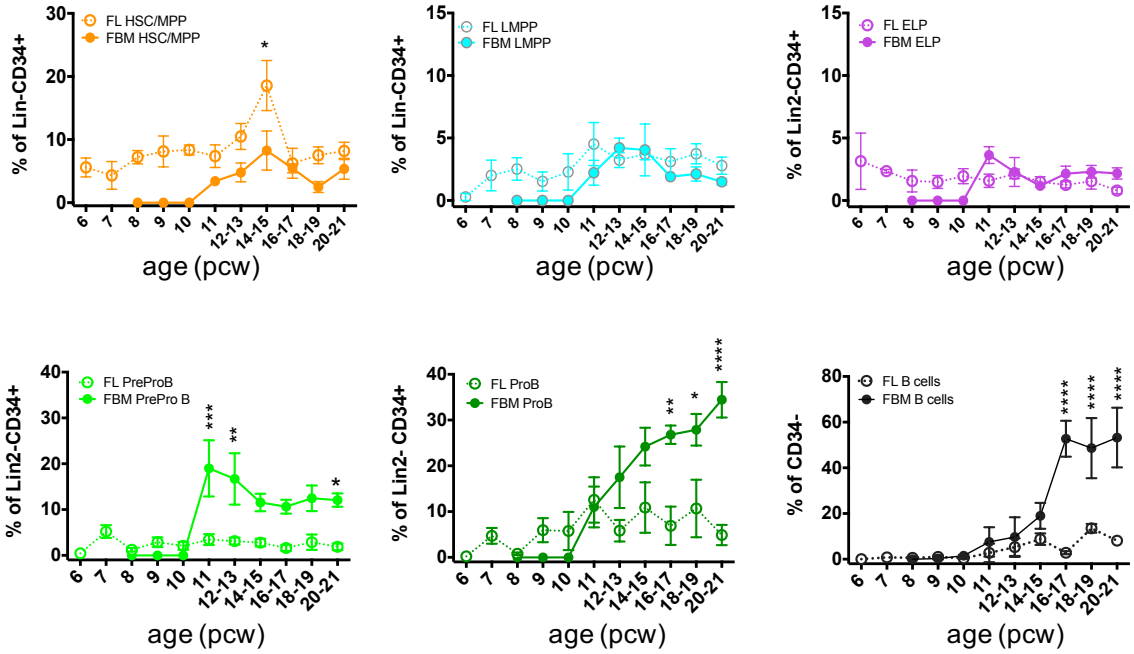


Figure 3.4: HSPC and B cell frequencies through gestation. Frequency of FL and fetal BM HSC/MPP, LMPP, ELP, PreProB progenitors, ProB progenitors and B cells from 6-21 PCW. Data expressed as a % Lin-CD34+ for HSC/MPP and LMPP (n=59 for FL an n=32-35 for fetal BM). Data expressed as a % Lin2-CD34+ for ELP, PreProB and ProB progenitors (n=39 for FL and n=32-38 for fetal BM). Data expressed as a % of CD34- for B cells (CD34-CD19+, n=37 for FL and n=32 for fetal BM). All graphs show mean ± SEM where n≥3. Statistical significance determined by two-way ANOVA. ∗p<0.05; ∗∗ p<0.01; ∗ ∗ ∗p<0.001; ∗ ∗ ∗∗p<0.0001.

To more accurately define any differences in the CD34-CD19+ B cell compartment of FL and fetal BM, I performed further flow cytometry analysis using the immunophenotypic definitions detailed in figure 3.5 on gestation-matched FL and fetal BM samples as well as CB as a control. Overall, B cells at all stages of maturation were observed in differing proportions (figures 3.6 and 3.7). While in the FL, there were approximately equal frequencies of PreB and Immature B cells, the fetal BM had a higher frequency of PreB cells and significantly fewer immature B cells when compared to FL (figures 3.6 and 3.7). These observations were in contrast to CB where PreB and immature B cells were virtually absent. Instead, in CB the majority of B cells were naive B cells (figures 3.6 and 3.7) which is to be expected given that CB is composed of circulating blood cells.

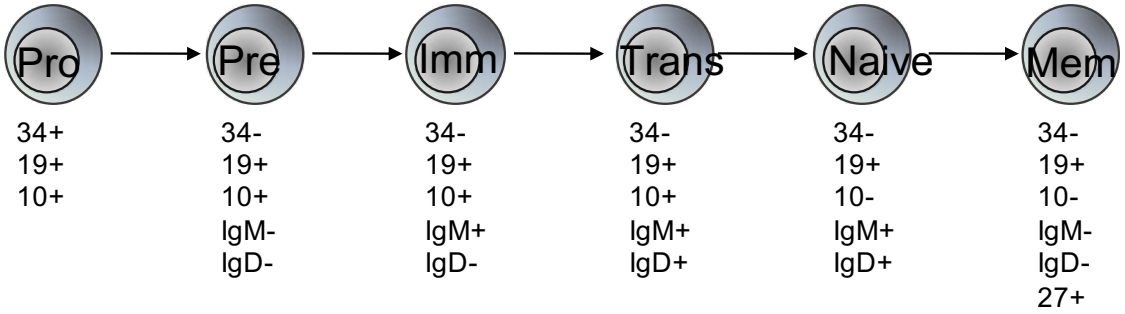


Figure 3.5: Scheme illustrating stages of B cell maturation and associated immunophenotype used to identify them. Imm: immature; Trans: transitional; Mem: memory.

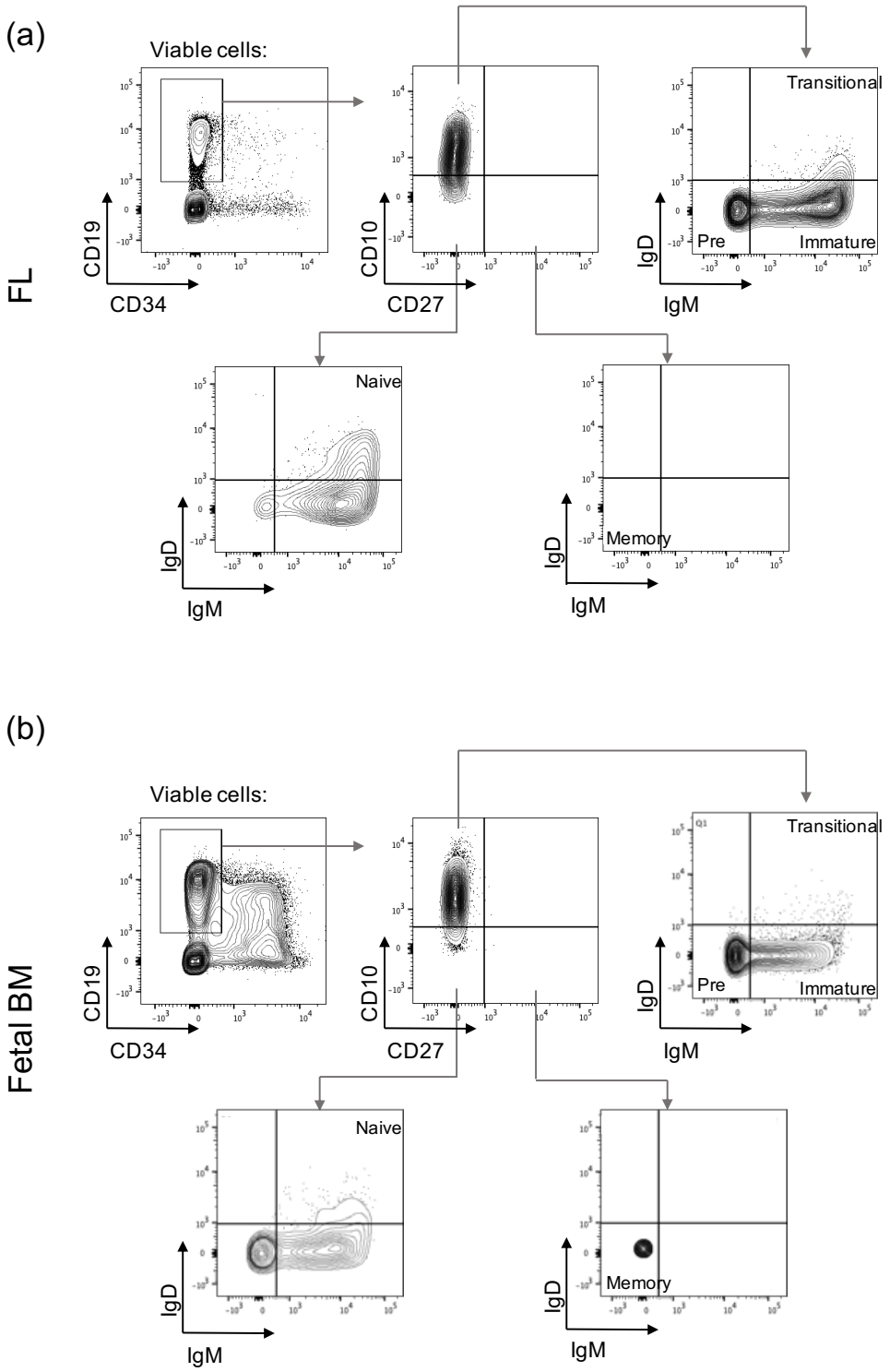


Figure 3.6: Representative flow plots from matched (a) FL (n=5) and (b) fetal BM (n=5) (16-20 pcw) showing relative proportion of mature B cells in the CD34-CD19+ compartment of each tissue.

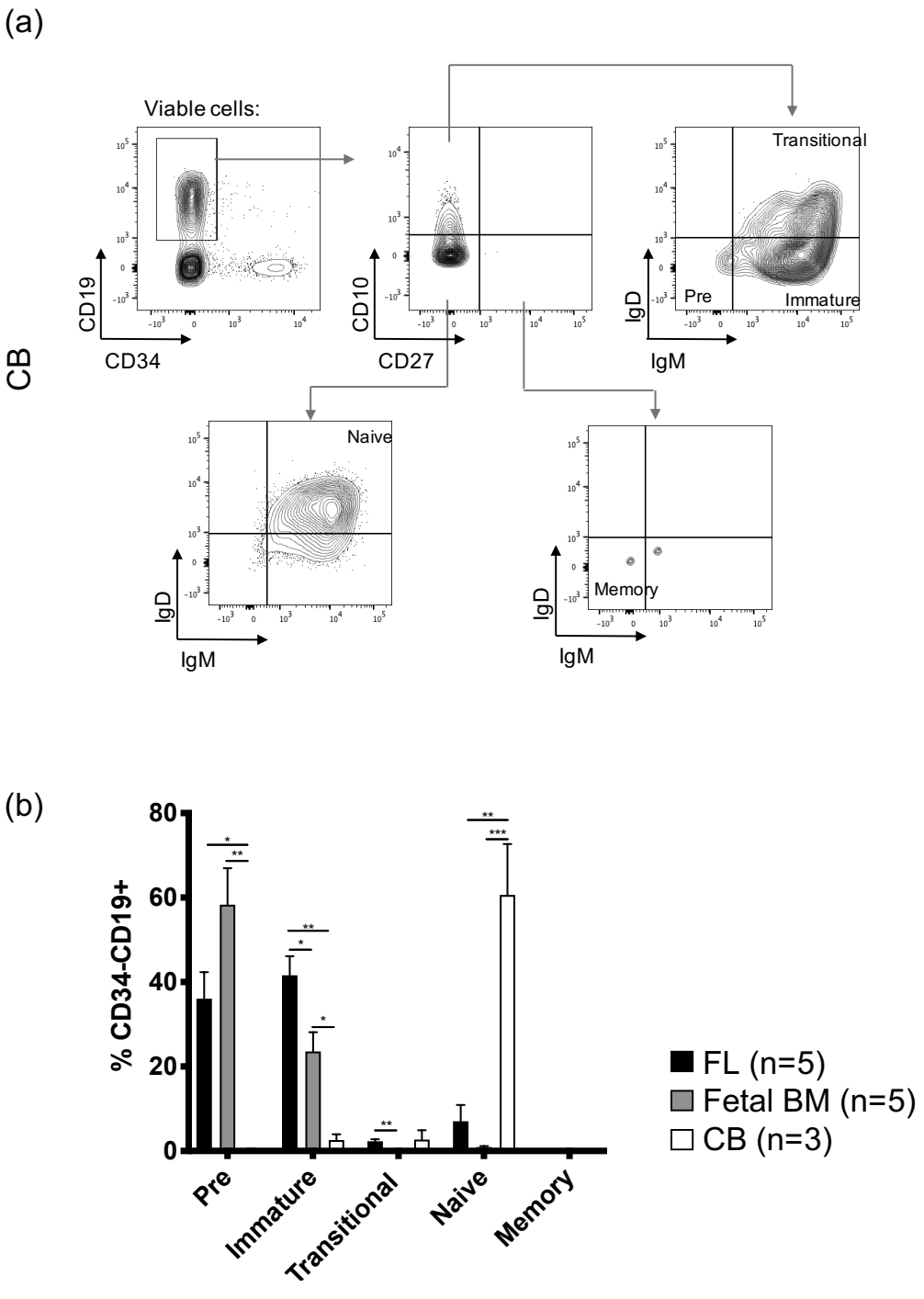


Figure 3.7: (a) representative flow plots from CB (n=3) showing relative proportion of mature B cells in the CD34-CD19+ compartment. (b) Frequencies of mature B populations. (16-20 pcw. Statistical significance determined by paired t-test. $*$ $p<0.05$, $**$ $p<0.005$, $***$ $p< 0.0005$.)

3.3 Cell cycle analysis

Given the high frequency of PreProB progenitors in the fetal BM at 11 pcw (figure 3.4). I hypothesised that a high proportion of these cells would be likely to be cycling. To address this, I performed intracellular flow cytometry analysis to determine the proportion of each cell population in each phase of the cell cycle. I designed two panels: the first to identify HSC, MPP and LMPP and the second to identify ELP, PreProB, ProB and B cells. While HSC, MPP and LMPP all showed a relatively low frequency in the S-G2-M phase of the cell cycle, ELP and PreProB progenitors had the highest frequencies in the S-G2-M phases of the cell cycle (figure 3.8). The significantly higher frequency of fetal BM PreProB progenitors in S-G2-M compared to HSC, MPP, LMPP and ProB progenitors suggests that PreProB-progenitor proliferation underpins this rapid expansion rather than differentiation of upstream progenitors.

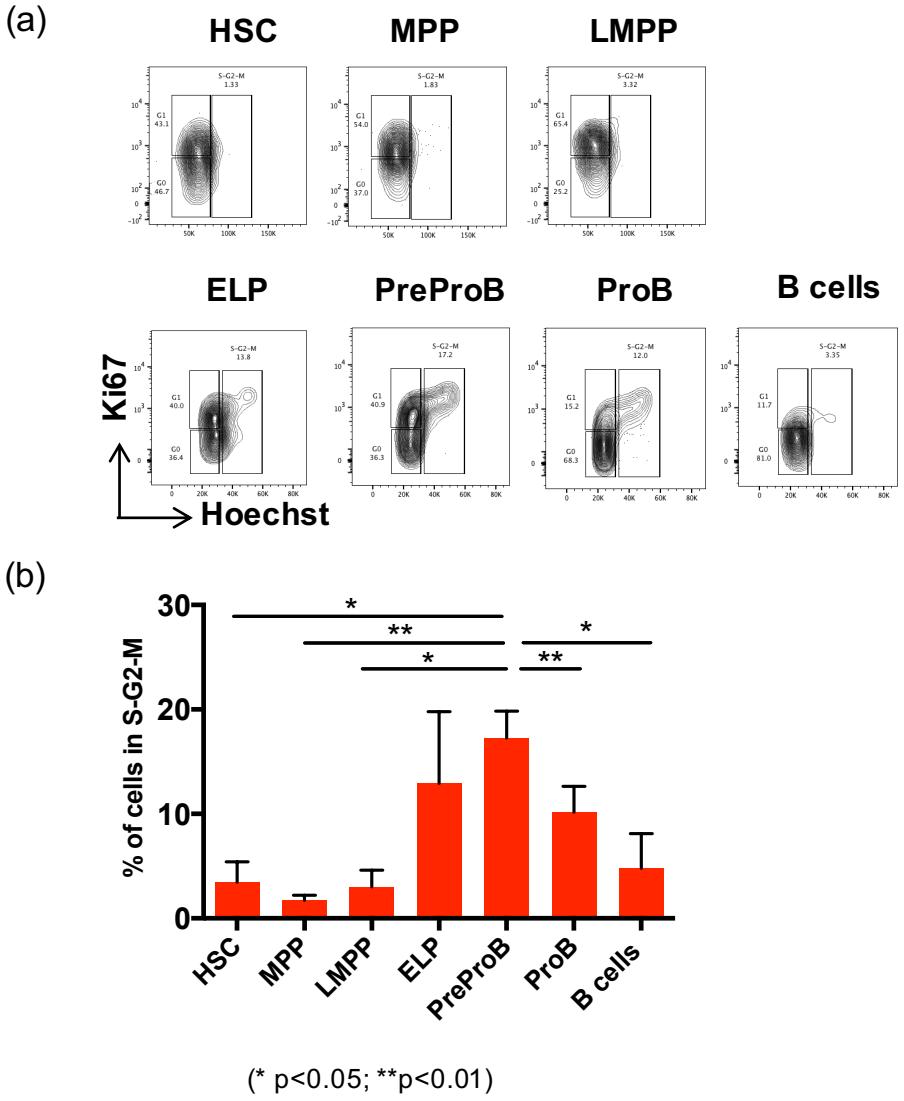


Figure 3.8: Cell cycle analysis by intra-cellular flow cytometry demonstrates that PreProB progenitors are cycling more than ProB progenitors. (a) Representative flow plots for each population (n=3 biological replicates). (b) Frequency of each population in S-G2-M phase of the cell cycle and paired t-test used to measure statistical significance between PreProB progenitors and all other progenitors.

3.4 *In vitro* differentiation potential of fetal BM HSPC

To define the functional potential of each population according to their ability to differentiate into lymphoid, myeloid and erythroid lineages I performed several *in vitro* differentiation assays using stromal cell co-cultures and clonogenic assays in methylcellulose.

3.4.1 MS-5 stromal co-cultures

The conventional way of determining B-lineage potential of human progenitors is to co-culture them with MS-5 stromal cells, a murine cell line, which also allow the output of NK cells and myeloid cells to be determined at the same time [Itoh et al., 1989]. To establish the B/NK and myeloid potential of fetal BM HSPC and B-progenitors, 100 cells of each population were sorted according to the strategy shown in figure 3.9 and plated on confluent layers of MS-5 cells. At regular time points, all cells were harvested from each well and 1/5th analysed by flow cytometry; the remaining cells were re-plated on fresh stromal layers. Using this approach, I found that by day 14, both PreProB and ProB progenitors yielded only B cells, with no NK or myeloid output thereby confirming their commitment to the B-lineage (figure 3.10). In contrast, while ELP had mainly differentiated into B cells by day 14, they also had some NK cell and myeloid potential which became more obvious by day 21. LMPP, MPP and HSC from the same samples were used as positive controls for multi-lineage output. As expected, LMPP differentiated into predominantly B cells, NK cells and some myeloid cells; and HSC and MPP co-cultures showed multi-lineage (B/NK/Myeloid) output by day 21 (figure 3.10). These data also demonstrate the hierarchical relationship between stem and progenitor cells: HSC and MPP took the longest to differentiate into mature haematopoietic cells in line with their position at the apex of the haematopoietic hierarchy.

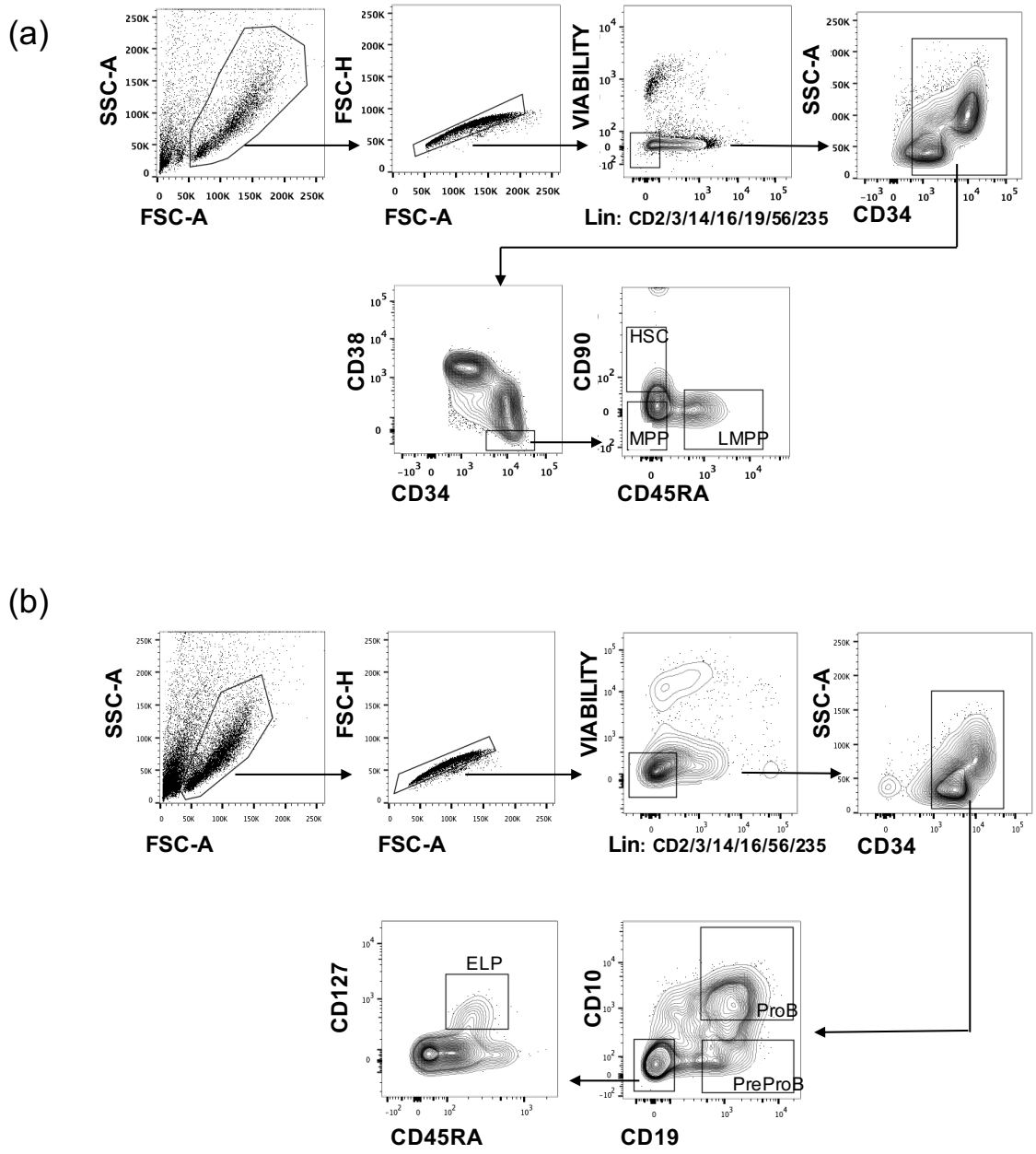


Figure 3.9: Sort gating strategy for fetal BM HSC, MPP and LMPP (a) and ELP, PreProB- and ProB progenitors (b) for establishing MS-5 stromal cell co-cultures to assess B, NK and myeloid cell potential *in vitro*.

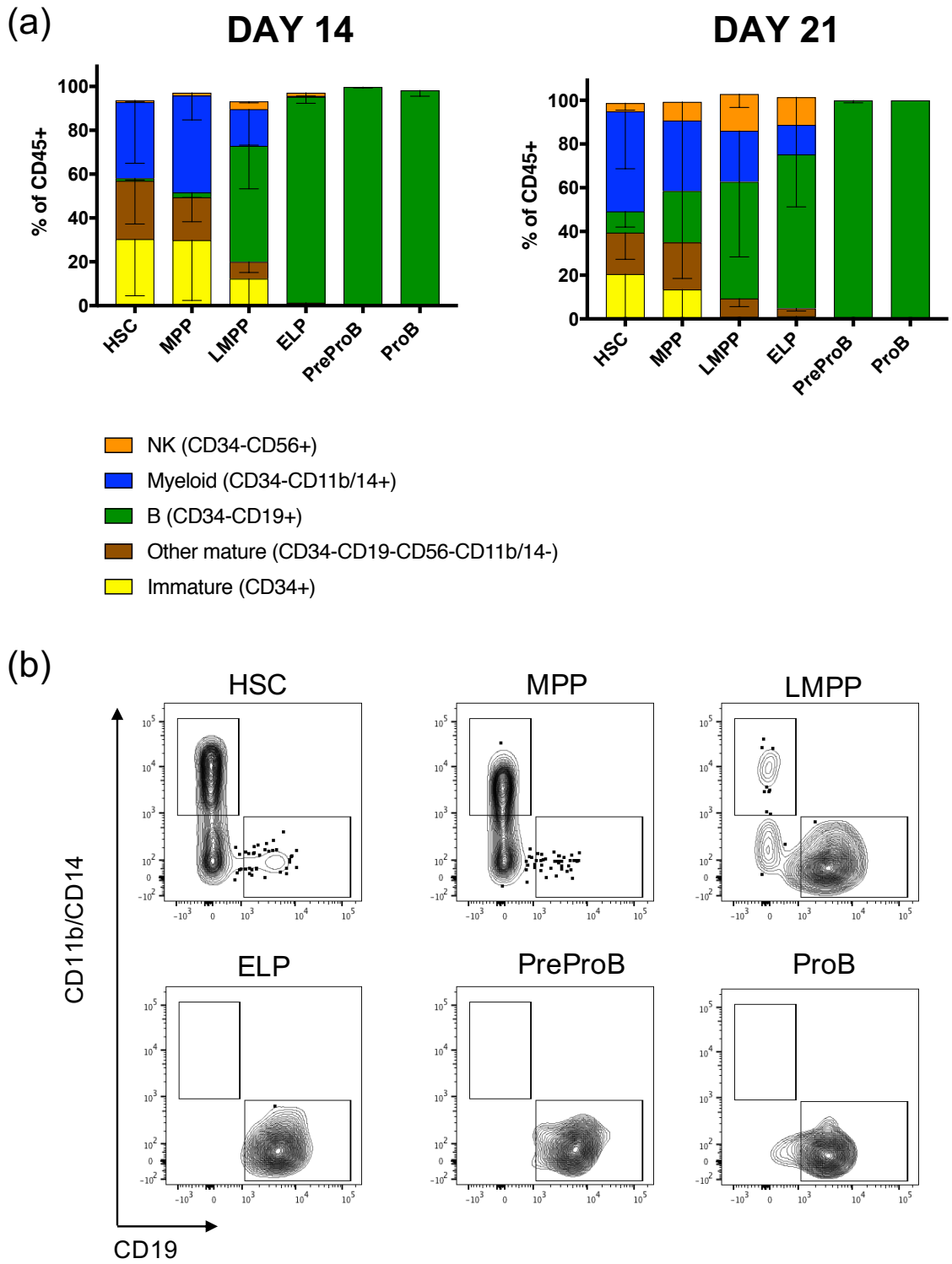


Figure 3.10: (a) B, NK and myeloid output from sorted progenitor populations at day 14 and day 21. Graphs shown as mean - SEM; n=4 (b) representative flow plots showing B/myeloid output from sorted progenitors at day 14.

Interestingly, I observed high levels of variation in the B lineage output of LMPP. Since previous work has suggested the presence of early lympho-myeloid progenitors (MLP/LMPP) in adult human BM and CB [Doulatov et al., 2010] [Goardon et al., 2011] differing by their expression of CD10, I looked at CD10 expression on fetal BM LMPP. This showed the presence of a small proportion of CD10+ LMPP in fetal BM (figure 3.11) and during the course of this project, it was confirmed that these CD10+ LMPP confer the observed B cell potential in the LMPP population in cord blood [Karamitros et al., 2018].

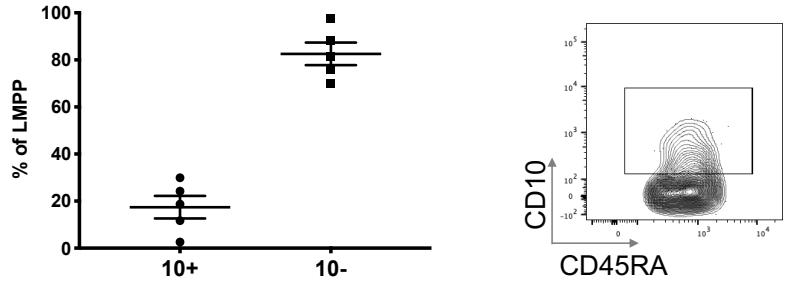


Figure 3.11: Left: The percentage of LMPP that are CD10+. Graphs shown with mean +/- SEM, n=5. Right: Representative FACS plot from fetal BM gated on Lin-CD34+CD38-D90-CD45RA+ (LMPP) showing the CD10+ gate.

In order to begin to dissect the hierarchy of B-lymphopoiesis in human fetal BM, I adapted the MS-5 co-culture assays to capture the initial changes in B lineage differentiation in ELP, PreProB and ProB progenitors. I harvested cells from the whole well (to obtain sufficient numbers for reliable analysis) and performed the analysis at day 3 and day 7 time points. Output analysis was performed using an alternative flow panel to enable the identification of ELP. Using this approach I was able to show for the first time that ELP differentiate into PreProB, ProB progenitors and eventually B cells *in vitro* thereby demonstrating that they are a source of PreProB progenitors. PreProB progenitors differentiated into ProB and B cells whereas ProB progenitors differentiated solely into B cells demonstrating that PreProB progenitors lie between ELP and ProB progenitors in the fetal B-lymphoid hierarchy (figure 3.12).

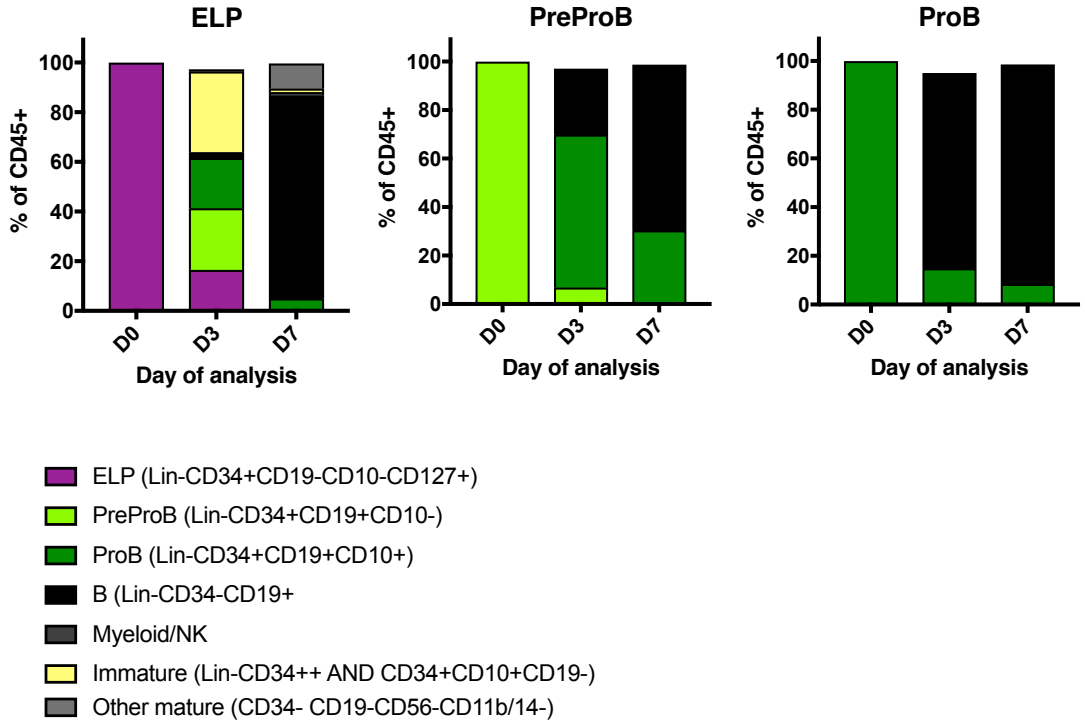


Figure 3.12: B-cell differentiation assay of flow-sorted ELP co-cultured on MS-5 stromal cells with cytokines (FLT3L, SCF, IL7 and IL2) showing output on day 3 and day 7. n=2.

Since I had previously noted the variable B lineage output of fetal BM LMPP and postulated that it might be due to variation in the level of CD10 expression on these cells, I also carried out parallel co-cultures of CD10+ and CD10- LMPP. Although these cultures produced insufficient numbers of B progenitors for reliable identification, the fetal BM CD10+ LMPP co-cultures did produce more B cells, as predicted, and in line with data from CB [Karamitros et al., 2018] (figure 3.13). In addition, what remained most striking about the short-term co-culture of LMPP was how quickly they made B cells in vitro. Within 7 days of co-culture, CD34-CD19+ B cells were easily identifiable indicating that this co-culture system efficiently induces rapid B cell differentiation in LMPP, ELP and B progenitors. Whether such rapid B cell differentiation from fetal BM LMPP happens *in vivo* is not known.

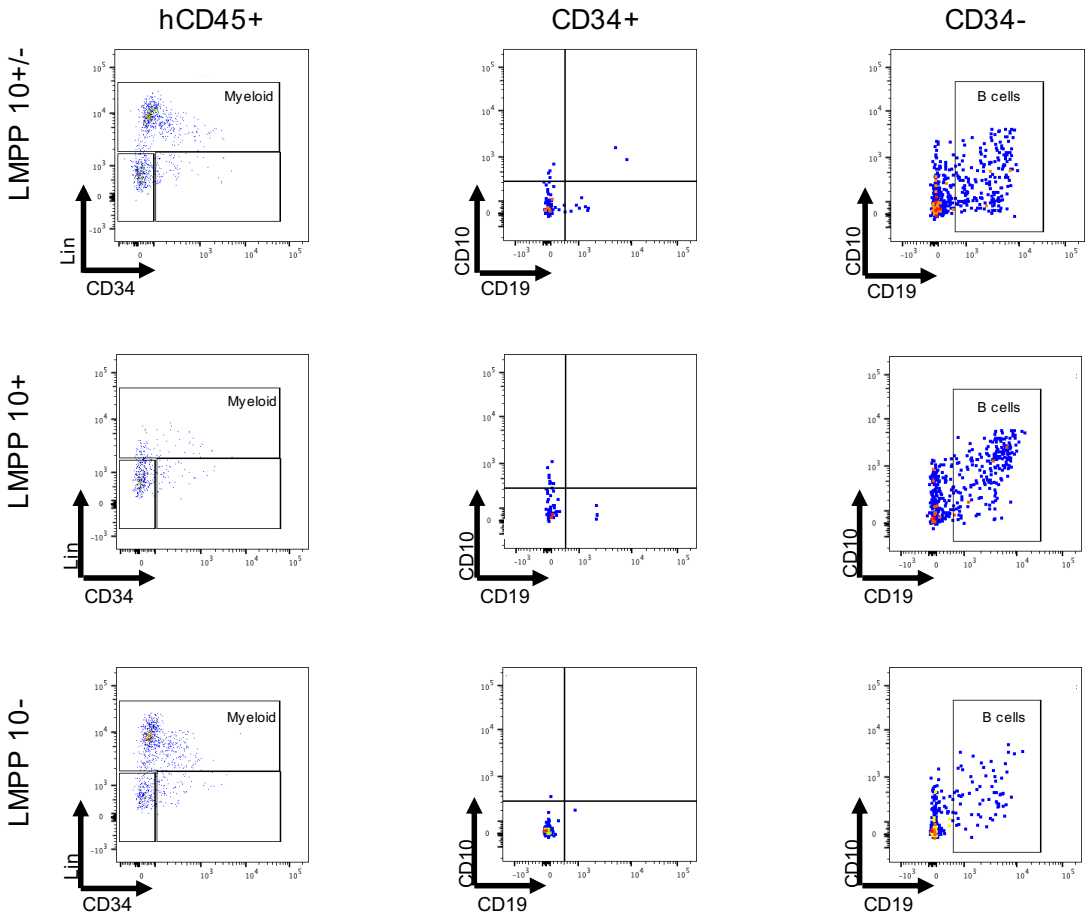


Figure 3.13: Representative flow plots from two independent experiments illustrating the B cell bias in CD10+ LMPP. Plots show the myeloid, immature and mature B output from left to right at D7 after 100 cells were sorted and plated. The top row represents output from CD10+/- LMPP, the middle row represents output from CD10+ LMPP and the bottom row represents output from CD10- LMPP.

3.4.2 MSC co-cultures

While conventional methods, such as MS-5 stromal co-cultures are excellent at establishing the B-lymphoid potential of any given HSPC population, they rely on murine stromal cells which in the case of MS-5 produce 40x the physiological levels of IL-7 [Parrish et al., 2009]. There is a danger that what we observe with these systems represents an artefact of a very artificial system and while they are the optimal system for establishing the B potential of any given population they probably do not reflect events *in vivo*. Therefore, I developed a fully human co-culture system to more closely resemble the fetal microenvironment that would enable careful dissection of the fetal B-lymphoid hierarchy.

Development and optimisation of this co-culture system for differentiating B cells was extremely challenging. Historically, human B cells are notoriously difficult to culture *in vitro* [Wolf et al., 1991] [Pribyl et al., 1995][Kurosaka et al., 1999] [Ichii et al., 2008] [Ichii et al., 2010a]. Therefore, I spent significant time optimising the derivation of the MSC (from FL and fetal BM) and their use in B cell differentiation assays.

To interrogate the hierarchy postulated as a result of the MS-5 co-cultures that suggested that PreProB progenitors lie upstream of ProB progenitors, I sorted ELP, PreProB and ProB progenitors and co-cultured them on confluent fetal BM MSC. All three progenitor populations demonstrated B cell potential on fetal BM MSC with PreProB and ProB progenitors demonstrating B cell commitment (figure 3.14). These data are consistent with the hierarchy observed in the MS-5 experiments (figure 3.12). I also observed an increased myeloid capacity of ELP in MSC co-cultures compared to MS-5 co-cultures. This, combined with the knowledge that HSC, MPP and LMPP all also produce more myeloid cells on MSC (described in Chapter 4) suggests that the MSC co-culture system is more efficient for measuring myeloid potential and therefore, it also provides further confirmation of the B-lineage commitment of PreProB progenitors.

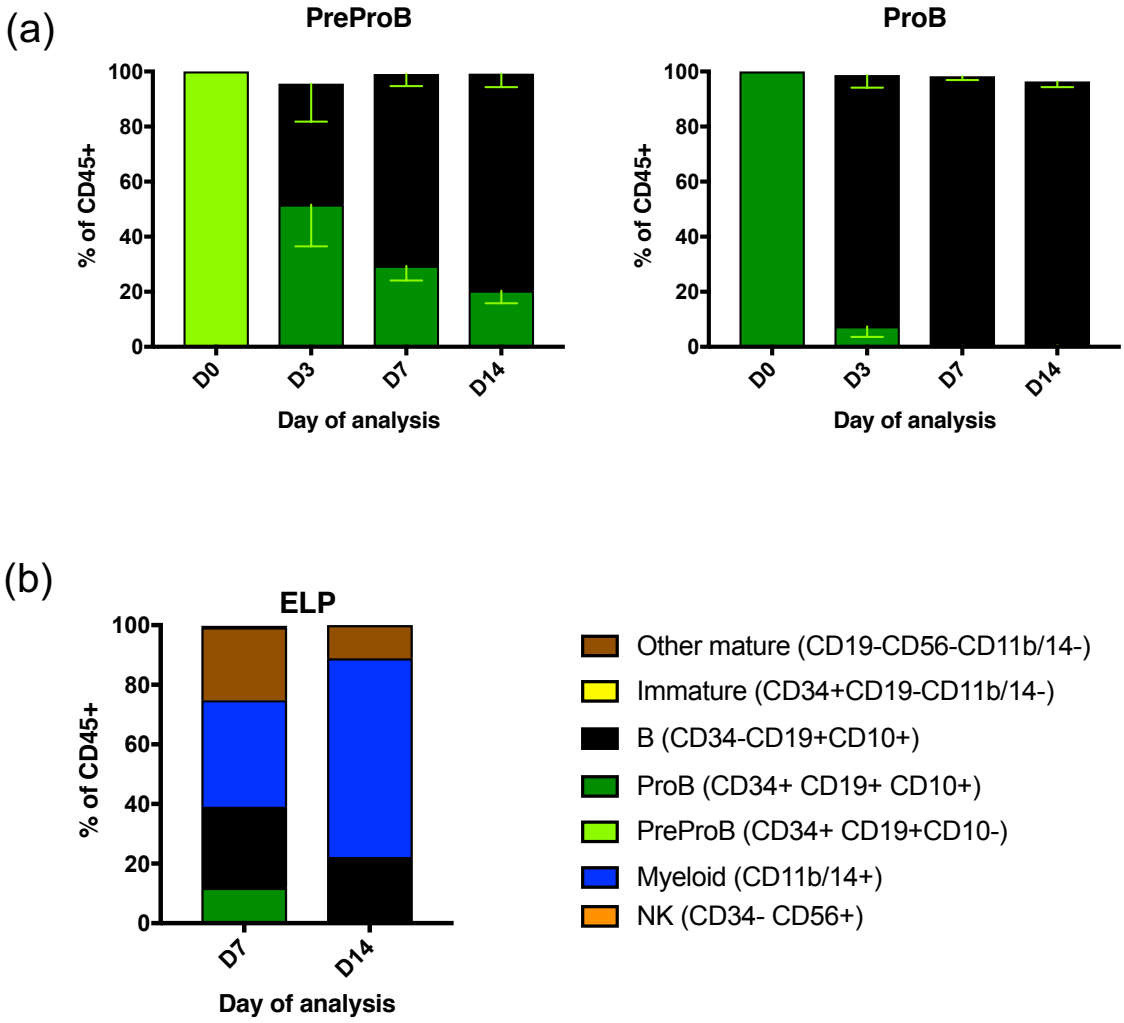


Figure 3.14: PreProB progenitors lie upstream of ProB progenitors and have no capacity to differentiate into myeloid cells: (a) PreProB progenitors differentiate into ProB progenitors and B cells; ProB progenitors only differentiate into B cells (b) ELP differentiate into B cells and myeloid cells. All co-cultures were repeated 3 times but one ELP co-culture did not grow and therefore the bar chart is indicative of two biological replicates.

3.4.3 OP9-DL1 stromal co-cultures

In order to determine whether ELP, PreProB and ProB had any T cell potential, I sorted 100 cells of each population, and HSC, MPP and LMPP as controls, for co-cultures on OP9-DL1 stromal cells. OP9-DL1 stromal cells are a murine stromal cell line that ectopically express notch ligand, delta-like 1, that promotes T-cell differentiation; and also express GFP which allows efficient elimination of stromal cells during flow cytometry analysis [Awong et al., 2007]. T cell potential was observed in HSC, MPP, LMPP and ELP OP9-DL1 co-cultures (figure 3.15). No T cells were detected in co-cultures where PreProB and ProB progenitors were seeded providing further evidence for their commitment solely to the B lineage. Analysis at 7 day time points confirmed the haematopoietic hierarchy: by day 14 ELP and LMPP T cell frequency had peaked; whereas by day 28 and 35 T cells could only be identified in the HSC and MPP co-cultures demonstrating their upstream position in the hierarchy.

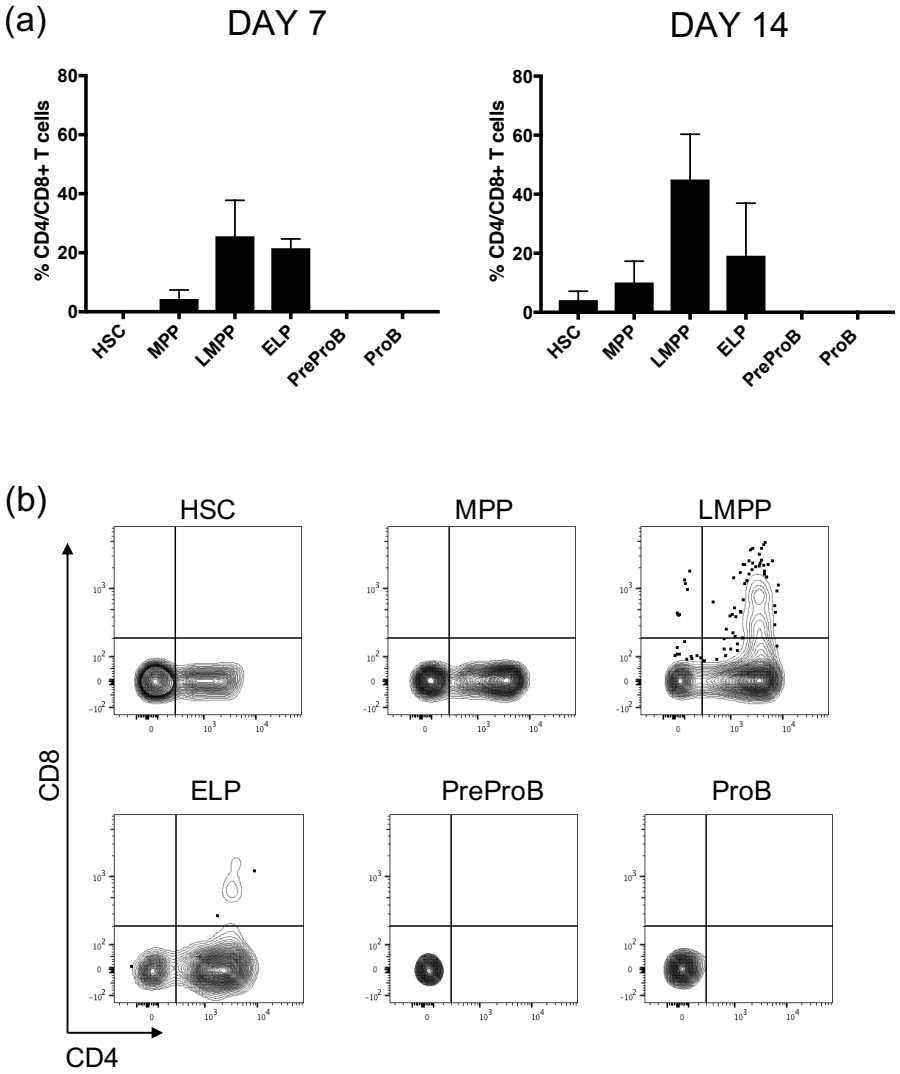


Figure 3.15: (a) T cell potential of sorted fetal BM HSC and progenitor populations. Graphs show total % of CD4+, CD8+ and CD4+CD8+ T cells as a % of hCD45+ cells at day 7 (left) or day 14 (right) of co-culture on OP9DL1 stromal cells, n=4. Graphs show mean + SEM. (b) representative flow plots showing CD4+ and CD8+ T cells produced by sorted progenitors at day 14.

3.4.4 Clonogenic assays

To assess myeloid and erythroid potential, I performed clonogenic assays in methylcelluose medium. These are the standard assays for measuring myeloid/erythroid potential in human HSPC but can be less sensitive than MS-5 co-cultures. I sorted 100 cells of each HSC and progenitor population and analysed the frequency and type of colony after 14-21 days at which point each colony is representative of a single cell that has multiplied and differentiated. Colonies were classified according to their appearance (size, morphology and haemoglobinisation) in accordance with convention (and using guides by Stem Cell Technologies; representative images in figure 3.14 (b)). Cytospins of picked colonies confirmed these classifications (figure 3.14(c)). As expected, HSC and MPP had myeloid and erythroid potential and LMPP displayed only myeloid potential. ELP, PreProB and ProB progenitors did not form any myeloid colonies *in vitro* (figure 3.16(a)).

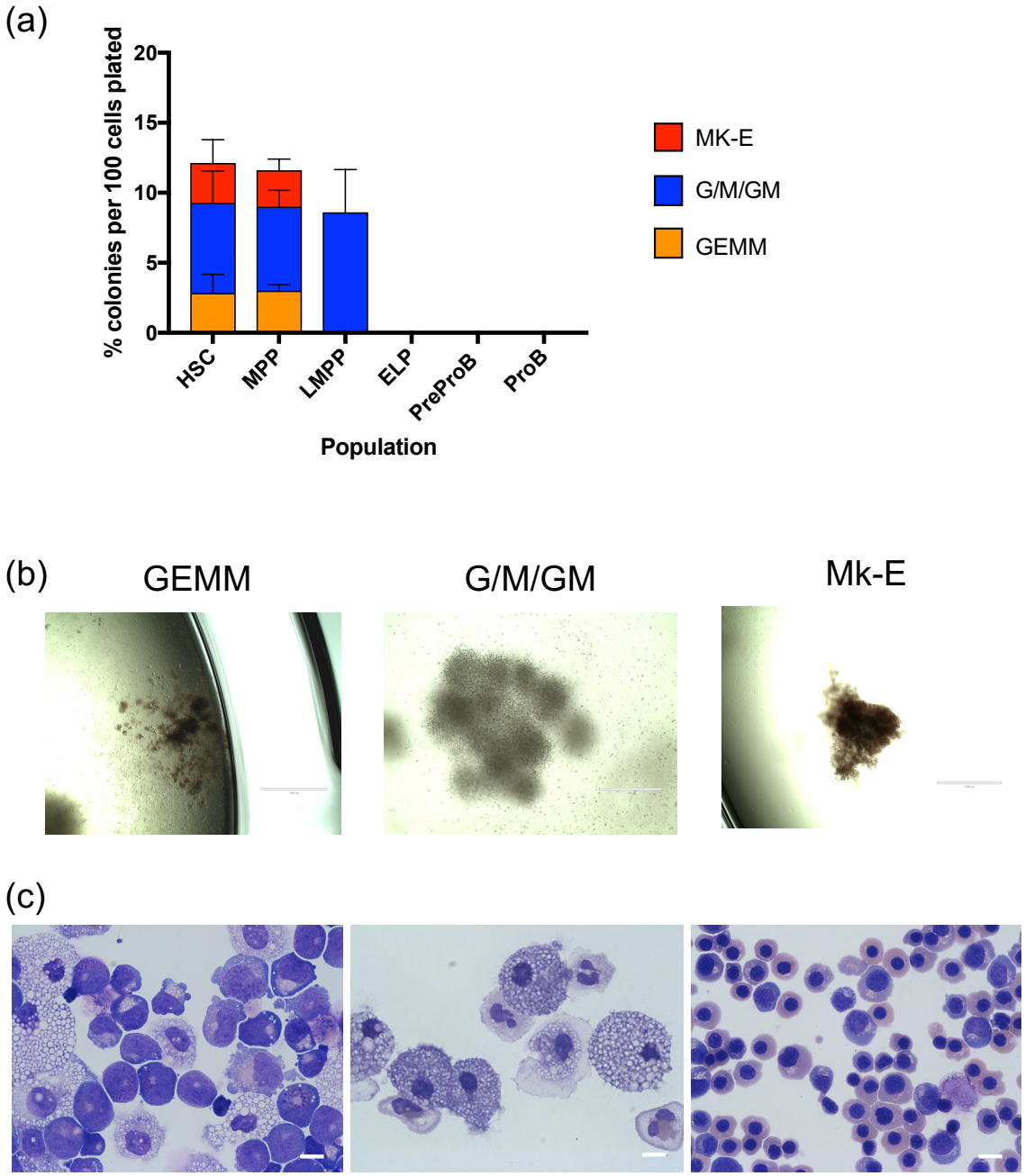


Figure 3.16: (a) Myeloid/megakaryocyte-erythroid potential of sorted fetal BM HSC and progenitor populations in clonogenic assays using MethoCult after 14-21 days (n=4). Data expressed as mean + SEM. (b) representative images of colonies identified (c) Giemsa-wright stained cytospins of picked colonies. White scale bar is 10μm.

3.5 *In vivo* functional assessment of fetal BM progenitors using a murine xenograft model

The *in vitro* data suggest that both PreProB and ProB progenitors are committed solely to the B-lineage; and that PreProB progenitors lie upstream of ProB progenitors *in vitro* proving they are the first committed B progenitor in fetal BM. To confirm this *in vivo*, xenograft transplants were performed where 10,000 flow-sorted PreProB or ProB progenitors were injected into sub-lethally irradiated NSG mice. As a positive control 10,000 Lin2-CD34+CD19-CD10-(double negative; DN) progenitors which would contain the HSC, MPP, LMPP and ELP as well as myeloid and erythroid progenitors transplanted (figure 3.17 (a)). After 14 weeks, mice were culled and their peripheral blood, BM and spleen assessed by flow cytometry. While 1/2 mice transplanted with the DN progenitors showed evidence of human CD45+ cells in the BM, none of the mice transplanted with PreProB or ProB progenitors displayed any engraftment 14 weeks post-transplant. Furthermore, the BM of one mouse from each group at weeks 3 and 6 showed no evidence of human cell engraftment in the mice transplanted with PreProB and ProB progenitors. These data are summarised in table 3.2.

Table 3.2: Summary of xenograft experiment where 10,000 progenitors were transplanted into sub-lethally irradiated NSG mice. hCD45+ cells were only observed in the BM of mice transplanted with the DN (Lin2-CD34+CD19-CD10-) cell population.

Mouse	# Cells transplanted	Population transplanted	Cull week	hCD45+ events in BM?
2.2b	10000	Lin2-34+10-19-(DN)	3	No
3.1f	10000	Lin2-34+10-19+ (PreProB)	3	No
4.1f	10000	Lin2-34+10+19+ (ProB)	3	No
4.1e	10000	Lin2-34+10-19-(DN)	6	Yes
3.1e	10000	Lin2-34+10-19+ (PreProB)	6	No
2.1f	10000	Lin2-34+10+19+ (ProB)	6	No
1.1f	10000	Lin2-34+10-19-(DN)	14	Yes
1.2b	10000	Lin2-34+10-19-(DN)	14	No
1.1e	10000	Lin2-34+10-19+ (PreProB)	14	No
1.1d	10000	Lin2-34+10+19+ (ProB)	14	No
2.2c	10000	Lin2-34+10+19+ (ProB)	14	No
2.2d	50000	Lin2-34+10+19+ (ProB)	14	No

To explore the possibility that PreProB and ProB progenitors home to the murine BM and differentiate immediately in the weeks following transplant (i.e. to engraft transiently as described in [Alhaj Hussen et al., 2017]), I also designed short-term xenograft assays. In these experiments, sub-lethally irradiated mice were transplanted

with 30,000 PreProB or ProB progenitors or DN cells. Their peripheral blood, BM and spleen were analysed by flow cytometry at 2 and 3 weeks post-transplant. This work was done with the help of Dr Nicholas Fordham and Ms Siobhan Rice (DPhil students in the lab) who kindly injected the cells that I sorted and helped in the tissue collection which I analysed.

While there were no hCD45+ events in the peripheral blood, engraftment (defined as $> 0.1\%$ of all CD45+ being hCD45+) was observed in the BM (figure 3.17 (b)) of mice transplanted with the DN and PreProB progenitor populations. In mice transplanted with PreProB progenitors, only cells committed to the B lineage (ProB and CD34-CD19+ B cells) were observed in the BM (figure 3.17 (b)). In contrast, mice transplanted with ProB progenitors engrafted at very low levels and there were too few hCD45+ events (<10 events) available for reliable identification. These data provide further evidence that PreProB progenitors lie upstream of ProB progenitors as they produce B cells *in vivo* within 2-3 weeks; while in this same time frame any transiently engrafted ProB progenitors presumably differentiated into B cells that were no longer detectable.

Mice transplanted with the DN population showed the best engraftment (figure 3.17 (b)). I observed multi-lineage output from the BM of these mice where myeloid, B, NK and T cells were observed alongside some immature CD34+ cells. Interestingly, within the hCD45+CD34+ compartment of these mice, ELP, PreProB and ProB progenitors were observed suggesting that fetal enriched ELP and PreProB progenitors generated from Lin-CD34+CD19-CD10- fetal BM cells in NSG faithfully mirror the differentiation hierarchy suggested by the flow cytometric frequency data and the *in vitro* functional data.

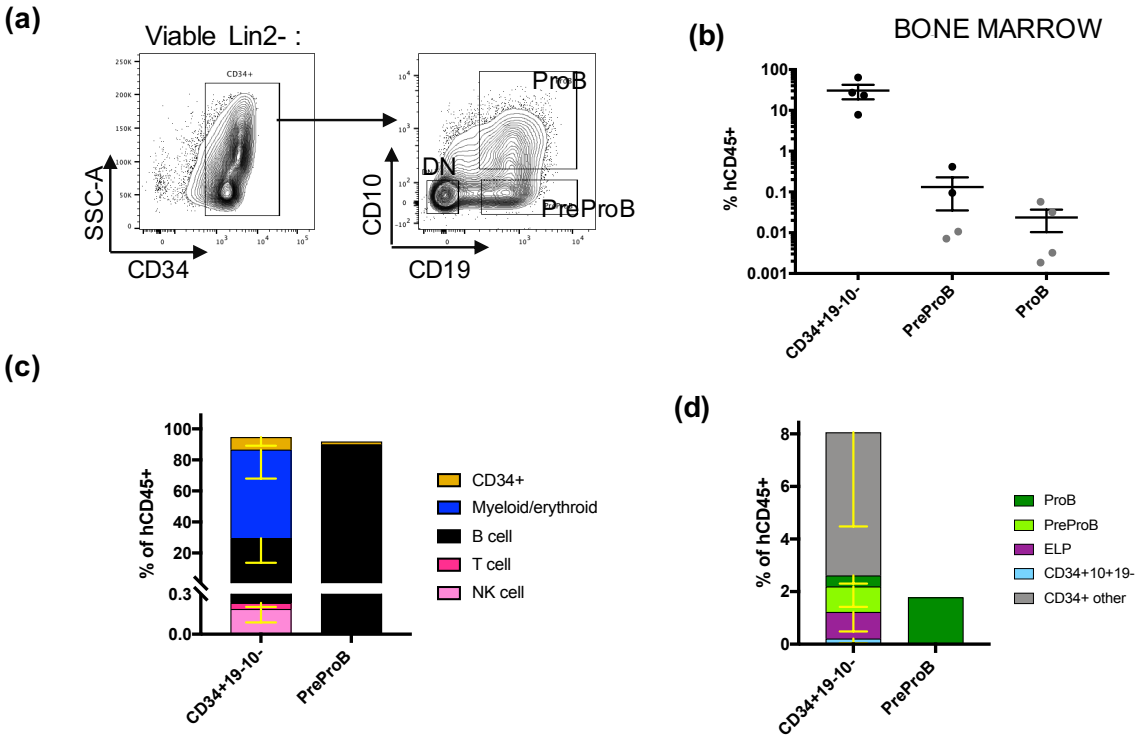


Figure 3.17: Short-term xenograft assays of fetal BM HSPC and lymphoid progenitor populations transplanted into NSG mice and analysed after 2-3 weeks. (a) sorted progenitor populations; (b) BM engraftment; mice where <10 hCD45+ cells were detected are marked in grey and not included in further analyses; (c) Lin2-CD34+CD19-CD10-(double negative; DN) cells differentiated into myeloid, and all lymphoid populations (B, T and NK), PreProB progenitors only differentiated into B cells; (d) breakdown of the hCD45+CD34+ compartment in engrafted mice.

3.6 B-lymphoid hierarchy at the molecular level

3.6.1 RNA-Sequencing

The functional analysis described above supported the B-lymphopoiesis developmental hierarchy proposed at the beginning of this work and summarised in figure 3.18. In order to determine the molecular basis for this process, 100 cells of each population were flow sorted using the cell markers detailed in the schematic, for RNA-Sequencing (RNA-Seq).

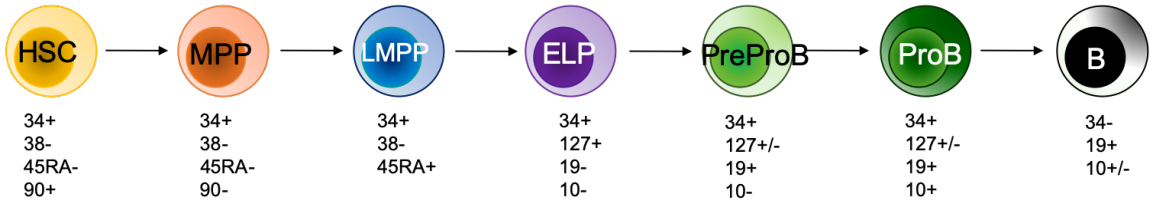


Figure 3.18: Schematic of the proposed human fetal B-lymphoid hierarchy.

RNA-Seq analysis was performed by myself as outlined in Chapter 2-Methods. First, I looked at genes known and established as being involved in B-lymphopoiesis. This showed sequential down-regulation of early lymphoid genes (*IKZF1, FLT3*); a peak in expression of early B-lymphoid specification genes (*EBF1, IL7R, DNTT, VPREB1, LEF1*) and a sequential upregulation of later, specific B-lymphoid genes (*PAX5, RAG1, RAG2, CD79B, IGHM*) along the proposed trajectory from HSC to B cells (figure 3.19). In addition, a principal component analysis (PCA) of the top 500 most variably expressed genes reflected the relationship between HSPC populations according to the proposed hierarchy (figure 3.20).

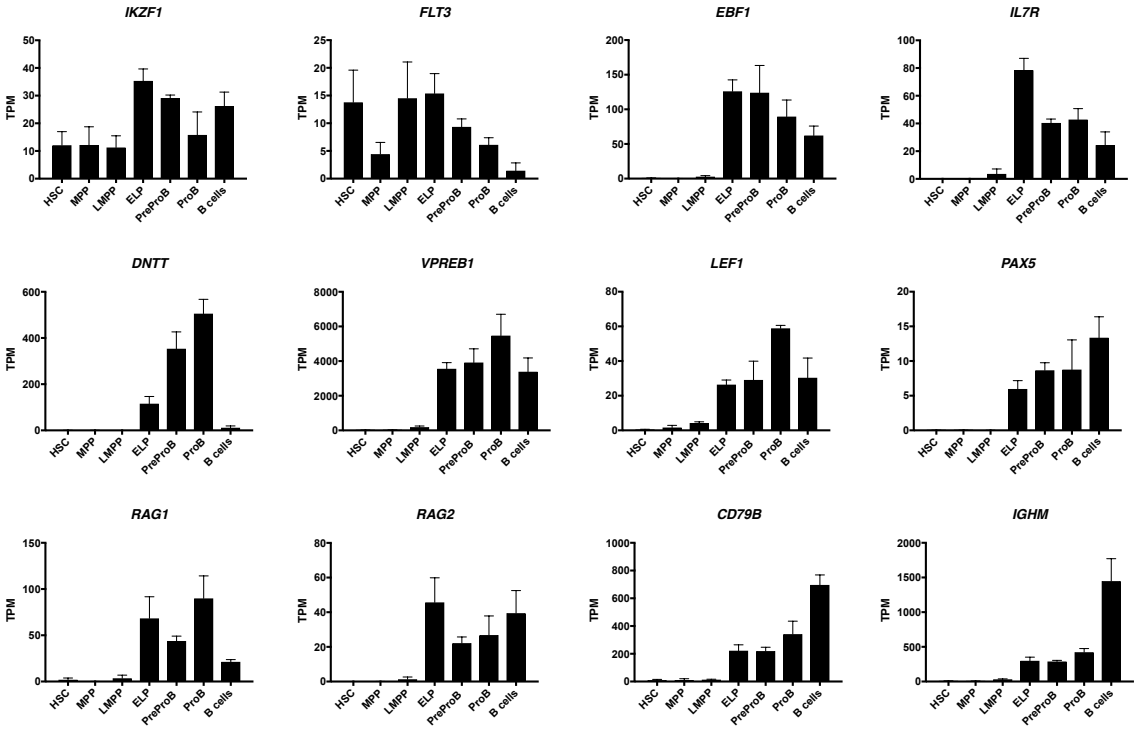


Figure 3.19: Expression of selected genes by RNA-Seq known for their role in B-lymphopoiesis. There is a sequential down-regulation of genes associated with stem cells (*IKZF1, FLT3*); peaks in expression of genes associated with early lymphoid commitment (*EBF1, IL7R*) and up-regulation of genes associated with B-lymphoid differentiation (*DNTT, VPREB1, LEF1, PAX5, RAG1, RAG2, CD79B, IGHM.*

Since both ELP and PreProB progenitors are fetal specific populations, I investigated their relationship with each other and other described populations in the B-lymphoid hierarchy (LMPP and ProB). Pairwise differential expression analysis of LMPP vs ELP, ELP vs PreProB and PreProB vs ProB showed that there were 1154, 244 and 663 differentially expressed protein coding genes (DEG; FDR <0.1) respectively (figure 3.21). In the comparison of LMPP and ELP, several genes associated with B-lymphopoiesis were significantly up-regulated in ELP (*LEF1, EBF1, PAX5, DNTT, MS4A1, CD38, CD24, BLNK, CD22*) and by contrast stem cell and early myeloid genes were down-regulated (*CD34, GATA2, GATA3, SPINK2, LMO2, ELANE, CD33, MPO, CEBPB, CSF3R*) (figure 3.22(a)). These data suggest that ELP lie downstream of LMPP and are more lymphoid "primed"/restricted than LMPP. Furthermore, when the transcriptomes of ELP were compared to PreProB progenitors, ELP over-expressed T cell (*CD2, CD3E*), stem cell (*GFI1B, HOXA7*) and myeloid (*FUT4*) (figure 3.22(b)) genes providing further evidence for their position in the B-lymphoid hierarchy upstream of PreProB progenitors and downstream of LMPP.

To confirm the position of PreProB progenitors in the fetal hierarchy, I also compared their expression profiles with ProB progenitors. There were several stem cell genes (*PROM1, SPINK2, HOXA3, HOXA5, RUNX2*); and some fetal associated (*LIN28B, CD7*) and T/NK/Myeloid genes (*CD244, CD248, CD3D, CD7, TFE3, MPO, CSF1R*) over expressed in PreProB progenitors. In contrast, genes associated with B-lymphoid maturation (*CD72, CDKN2A, MME, CD9, TCL1A, CD40*) (figure 3.22(c)) were over expressed in ProB progenitors compared with PreProB progenitors.

These data confirm the observed functional differences described earlier and provide further evidence for the proposed hierarchy. Full DE tables are available in online supplemental files.

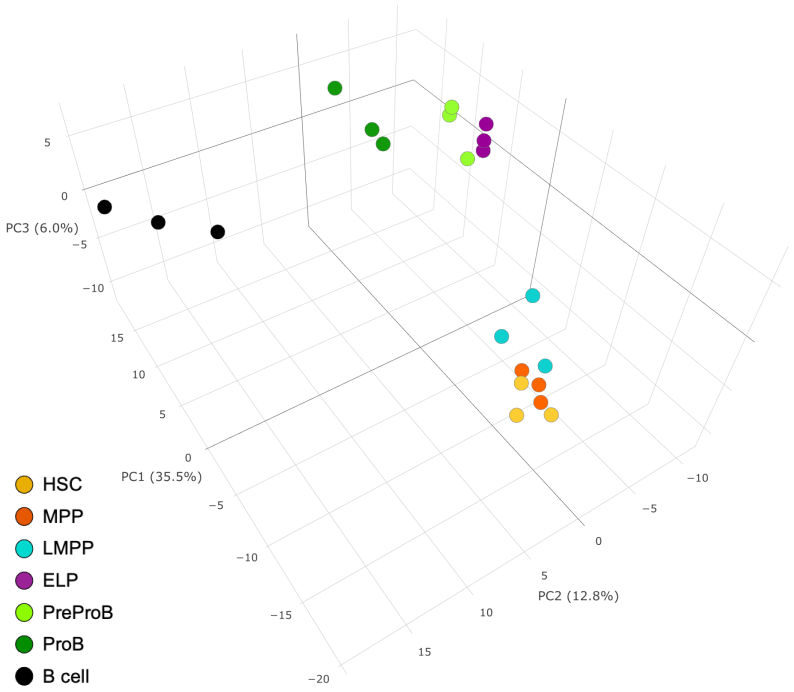


Figure 3.20: Bulk RNA-Seq of 100 cells from each flow sorted population (n=3): principal component analysis (PCA) of each HSPC population using the top 500 most variable expressed genes.

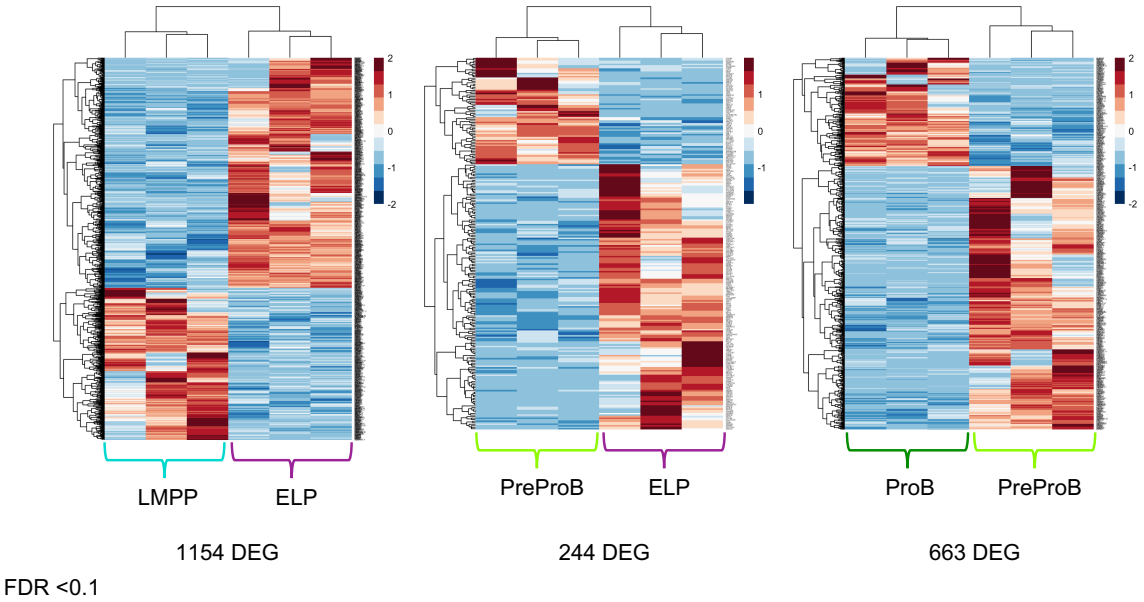


Figure 3.21: Bulk RNA-Seq of 100 cells from each flow sorted population (n=3): differential expression analysis of LMPP vs ELP, ELP vs PreProB and PreProB vs ProB progenitors performed using DESeq2, FDR < 0.1. DEG: differentially expressed genes.

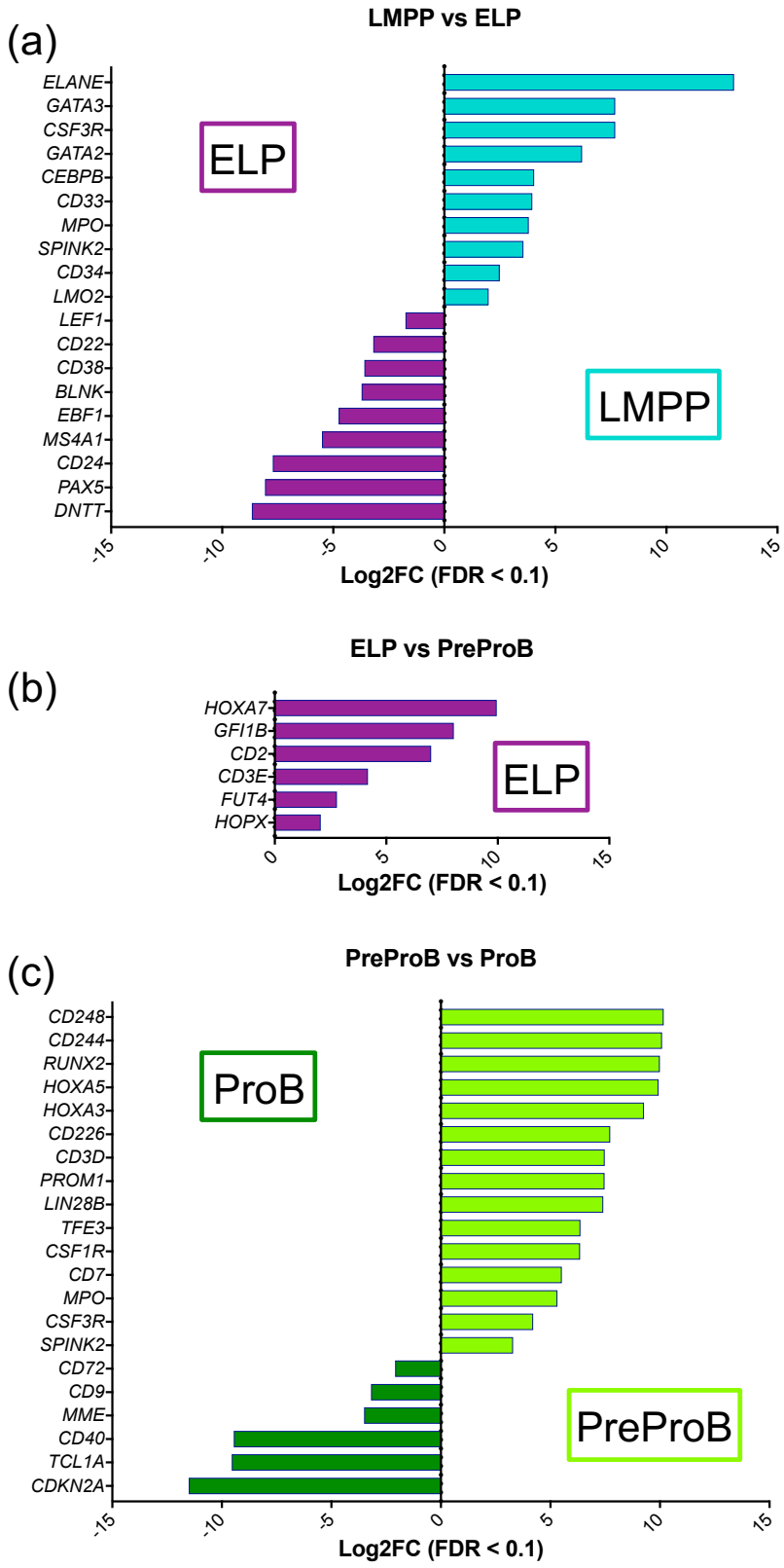


Figure 3.22: Significant DE genes by RNA-Seq between (a) LMPP and ELP; (b) ELP and PreProB; and (c) PreProB and ProB progenitors. Genes selected for their known role in haematopoietic lineage commitment from DE analysis of RNA-Seq data (n=3 per population; 100 cells/population).

3.6.2 Single cell RT-qPCR

To examine the B-lymphoid hierarchy at single cell resolution and also validate the findings from the RNA-Seq, single cell RT-qPCR was performed using the Fluidigm Biomark platform on immunophenotypically defined HSPC, B-progenitors and B cells. A panel of 96 genes were chosen based on the RNA-Seq analysis, their known involvement in B-lymphopoiesis and, given the potential of fetal BM progenitors to act as leukaemia initiating cells, a number of genes specifically known their known involvement in ALL were included (see Chapter 2-Methods for details of all probes used). A total of 1400 cells were index sorted from 3 biological samples and labelled according to their immunophenotype.

Analysis of expression of individual genes confirmed the RNA-Seq analysis: there is sequential up-regulation of B-lymphoid genes from HSC through to ProB progenitors (figure 3.23). A diffusion map of all 96 genes, placed HSC through to B cells in the differentiation trajectory proposed by the functional analysis (figures 3.24, 3.18). Colouring the diffusion map by pseudotime and also by expression of selected genes further demonstrated the hierarchy whereby ELP lay downstream of LMPP and upstream of PreProB progenitors; and PreProB progenitors lay downstream of ELP and upstream of ProB progenitors (figure 3.25).

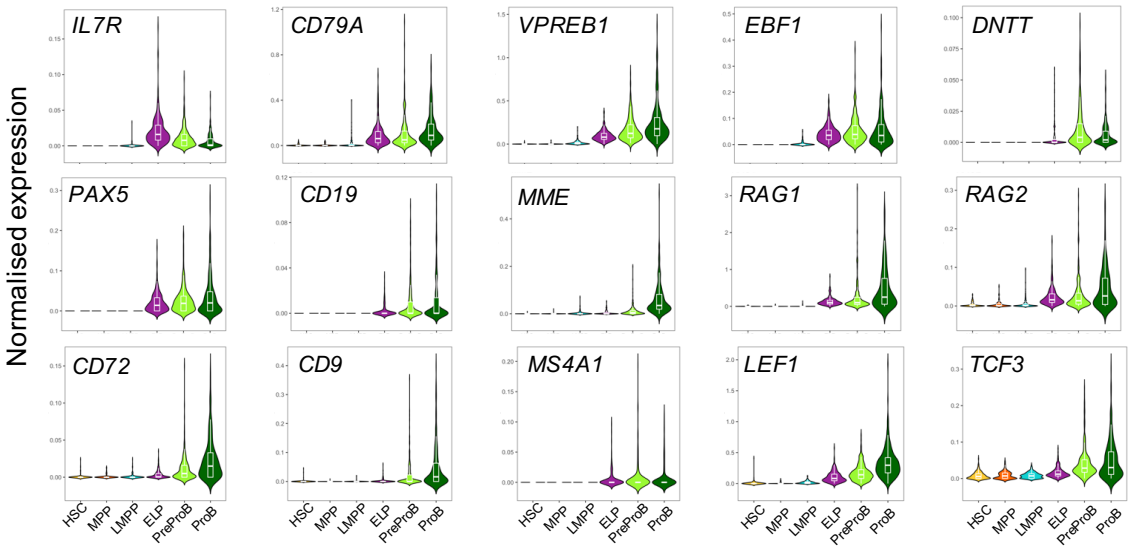


Figure 3.23: Sequential up-regulation of B-lymphoid genes determined by single cell RT-qPCR in individual flow-sorted HSC and progenitors from 3 fetal BM samples (1400 cells sorted).

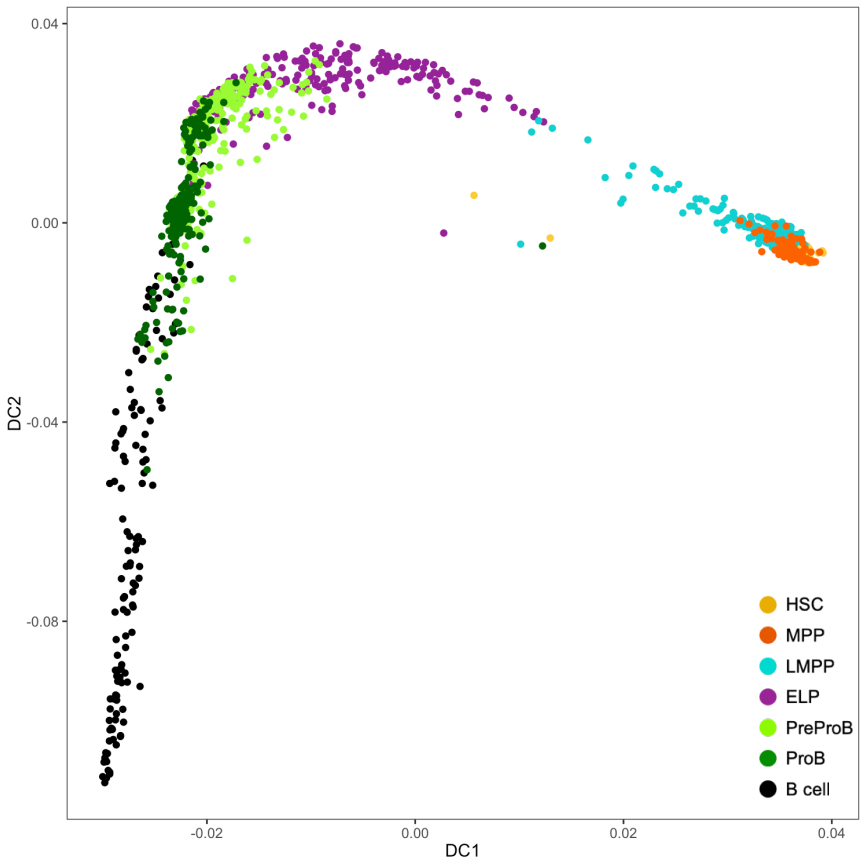


Figure 3.24: Diffusion map of single cell RT-qPCR of 96 genes expressed by fetal BM HSPC and coloured by immunophenotype showing the B-lymphoid differentiation trajectory from HSC/MPP to B cells.

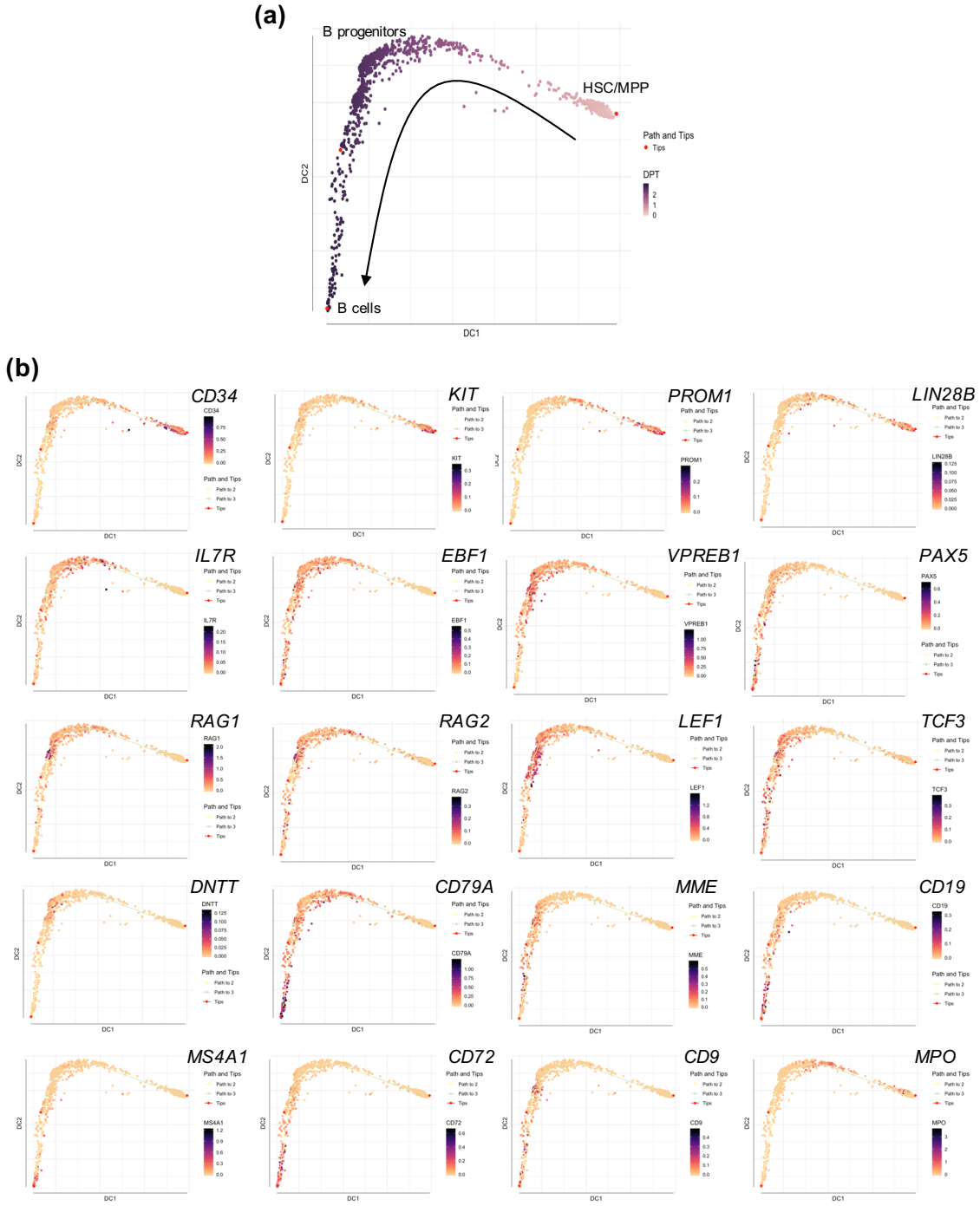


Figure 3.25: Diffusion maps of single cell RT-qPCR of 96 genes expressed by fetal BM HSPC and coloured by (a) pseudotime and (b) selected gene expression to show the B-lymphoid differentiation trajectory.

To make sure that the approach of prospectively isolating each progenitor population was not excluding any other novel B progenitor populations, I index sorted 938 progenitors (2 biological samples) from the Lin2-CD34+ fraction of fetal BM MNC and chose a panel of genes known for their involvement in lymphopoiesis, myelopoiesis and erythropoiesis for single cell RT-qPCR. Every cell that was index sorted and used for RT-qPCR was retrospectively labelled using the approach in figure 3.26. To prevent under-representation of rarer populations such as HSC and LMPP, I selectively sorted these cells towards the end of the experiment.

The diffusion map of this data (figure 3.27(a)) demonstrated three trajectories: B-lymphoid, myeloid and erythroid. A 3D diffusion map (figure 3.27 (b)) highlighted a small population that was CD123hi by flow cytometry. I hypothesise that these progenitors are dendritic cell precursors given their high expression of *JCHAIN, RUNX2 and CD44* but further investigations would be required to fully characterise these cells.

Colouring the diffusion map according to pseudotime and also by genes known for their expression in stem, B-Lymphoid, myeloid and erythroid lineages further illustrates this (figure 3.28). These analyses confirm that the prospective isolation of B progenitors closely reflects the continuum of B-lymphopoiesis from LMPP to ELP to PreProB to ProB and did not identify any additional B progenitor populations.

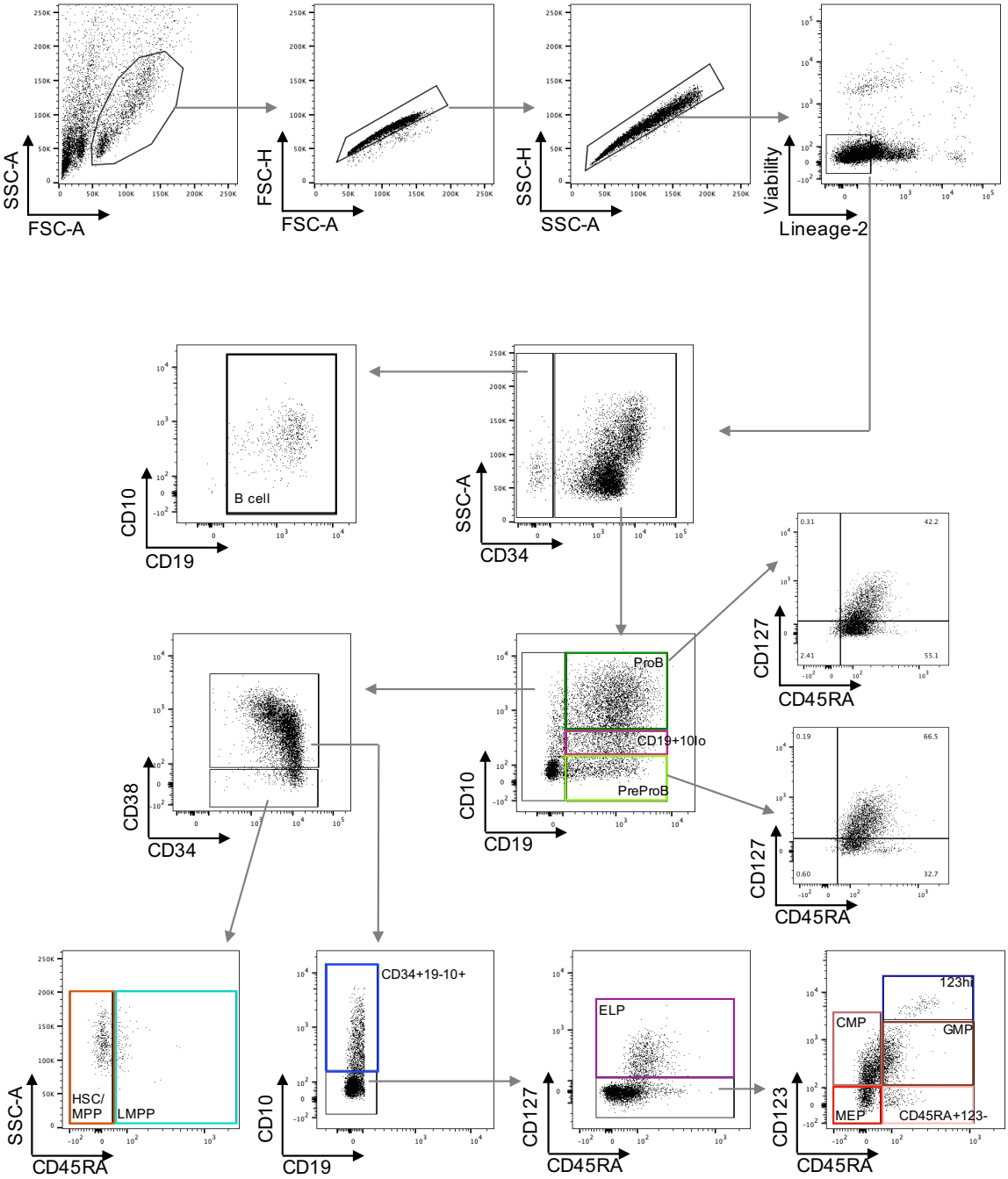


Figure 3.26: Post index sort labelling strategy for all cells in the Lin2-CD34+ compartment sorted for single cell RT-qPCR. Flow plots are representative of one biological sample where all cells were retrospectively labelled according to their immunophenotype labelled here (colours correspond to diffusion map in next figure). A total of 938 cells were sorted. CMP: Common Myeloid Progenitor, MEP: Mega-Erythroid Progenitor, GMP: Granulo-Monocyte Progenitor. Lin2: CD2, CD3, CD14, CD16, CD56 and CD235a.

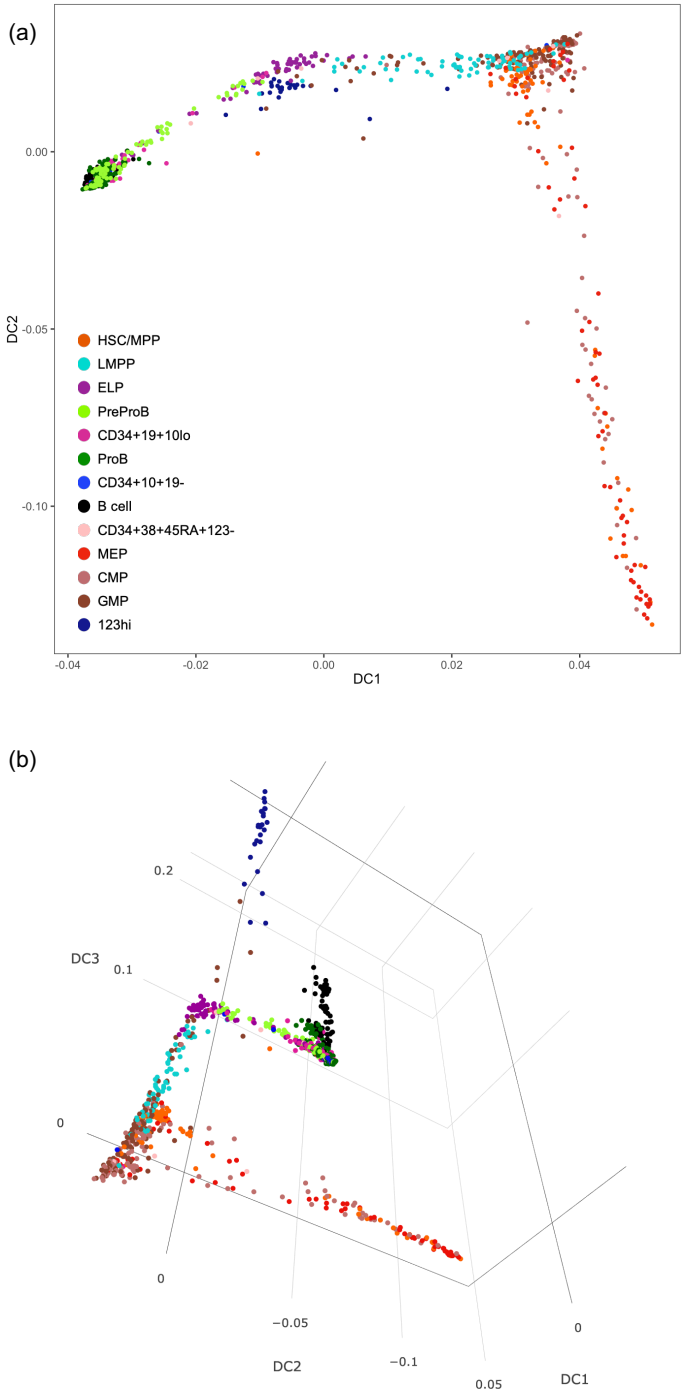


Figure 3.27: 2D (a) and 3D (b) diffusion map of single cell (n=938 from two biological samples) RT-qPCR of 96 genes expressed by fetal BM HSPC and coloured by the label they were assigned in figure 3.26. Genes were chosen for their known involvement in lymphopoiesis, erythropoiesis and myelopoiesis.

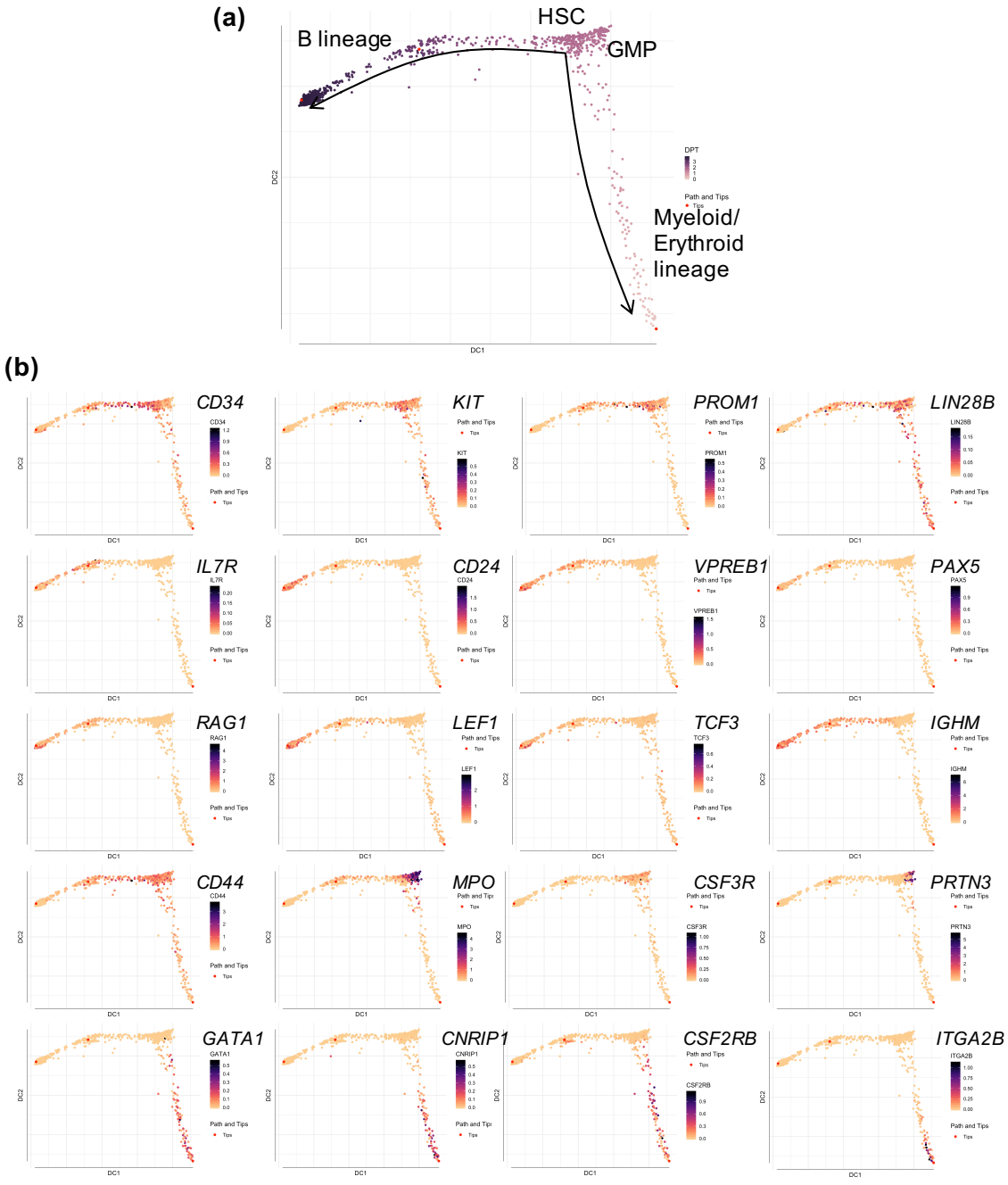


Figure 3.28: Diffusion map of single cell RT-qPCR of fetal BM HSPC shows 3 principal lineages downstream of HSC within the fetal BM Lin2-CD34+ compartment as expected (myeloid, B-lineage and myelo-erythroid (a) diffusion map coloured by pseudotime; (b) diffusion maps coloured by expression of selected genes.

To confirm these observations independently, in addition to using the R package Destiny, I also used the R package Monocle for analysis. While Destiny was written for single cell RT-qPCR data and Monocle was written for single cell RNA-Seq data, both packages can be used for both types of data. Analysis using Monocle arguably demonstrated increased resolution compared to Destiny (figure 3.29). The trajectory begins with HSC/MPP/LMPP and at the first branch point the myeloid progenitors CMP and GMP separate, this is closely followed by a second branching of megakaryocyte erythroid progenitors (MEP). Interestingly, the last branch point largely separates the PreProB progenitors from the ProB progenitors and B cells which could either reflect fetal specific programmes or multiple differentiation pathways. When the trajectory is coloured by pseudotime, the pseudotemporal ordering is comparable to the analysis with Destiny: pseudotime starts at the myeloid/erythroid lineage and ends at B cells. While on initial glance, this pseudotemporal ordering appears at odds with what we might expect i.e. HSC to be where pseudotime starts, in fact the result reflects the logic of the analytical packages used. Both Destiny and Monocle's underlying algorithms used for multi-dimentional reduction ask the question: "how related is each cell to each other?". With this in mind, it makes sense that myeloid progenitors are most different from B lineage cells and the cells that 'unite' them are their parent cells: HSC/MPP.

To establish what could be driving the pseudotime, I plotted the relative expression of each individual gene against pseudotime and coloured each cell according to its immunophenotype. The genes specified in figure 3.30 are likely to be the key drivers of this trajectory.

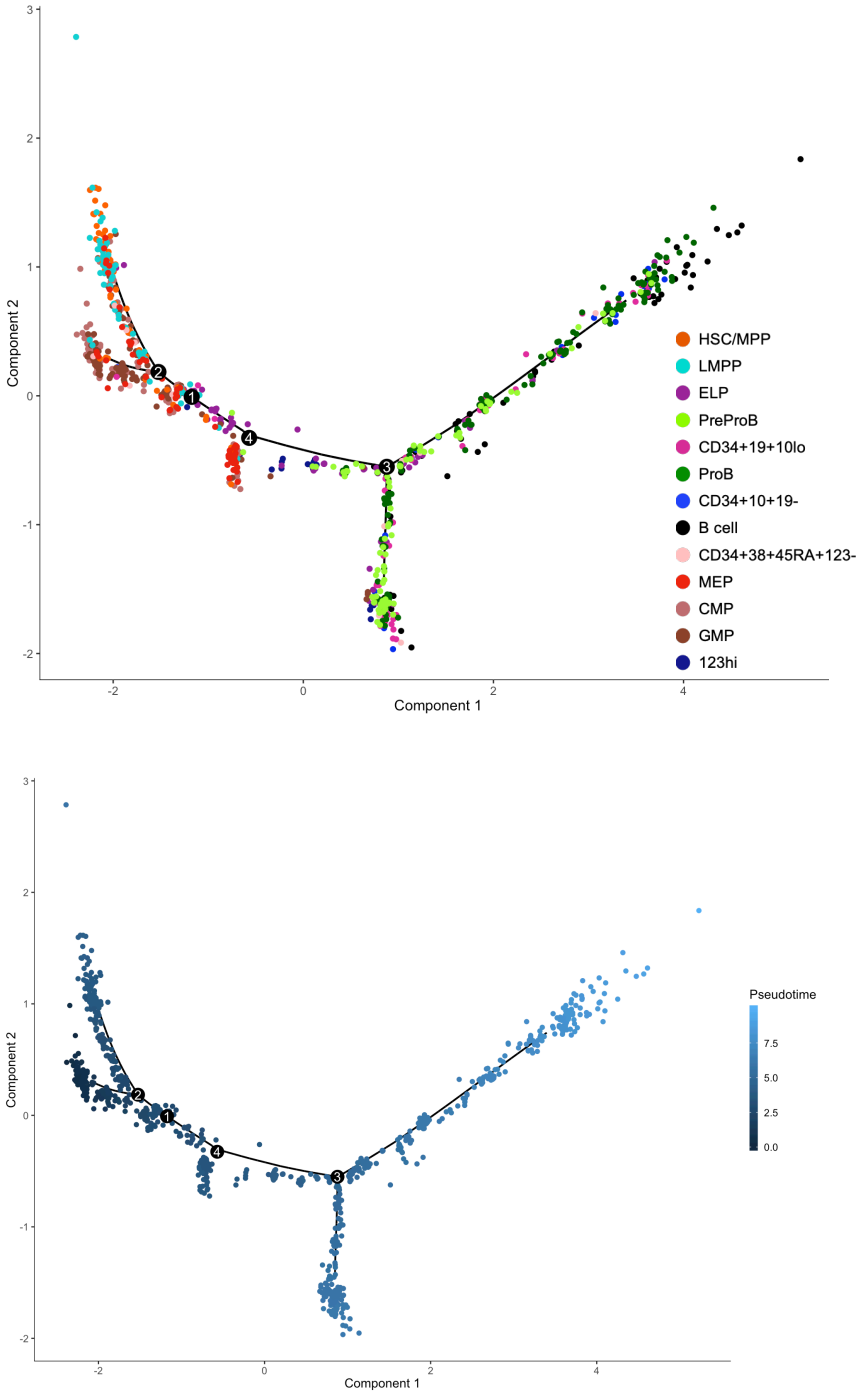


Figure 3.29: Trajectory analysis of single cell gene expression of 96 genes in index sorted single cells from fetal BM (2 biological samples, 938 cells) using Monocle. Top plot shows trajectory where each cell is coloured according to the immunophenotype label it was assigned in figure 3.26. Bottom plot shows trajectory coloured by pseudotime. Similar to analysis using Destiny, Monocle has ordered HSC and progenitors according to three principle lineages: myeloid, erythroid and B-lymphoid. Later in pseudotime there appears to be a split in the B-lymphopoiesis trajectory.

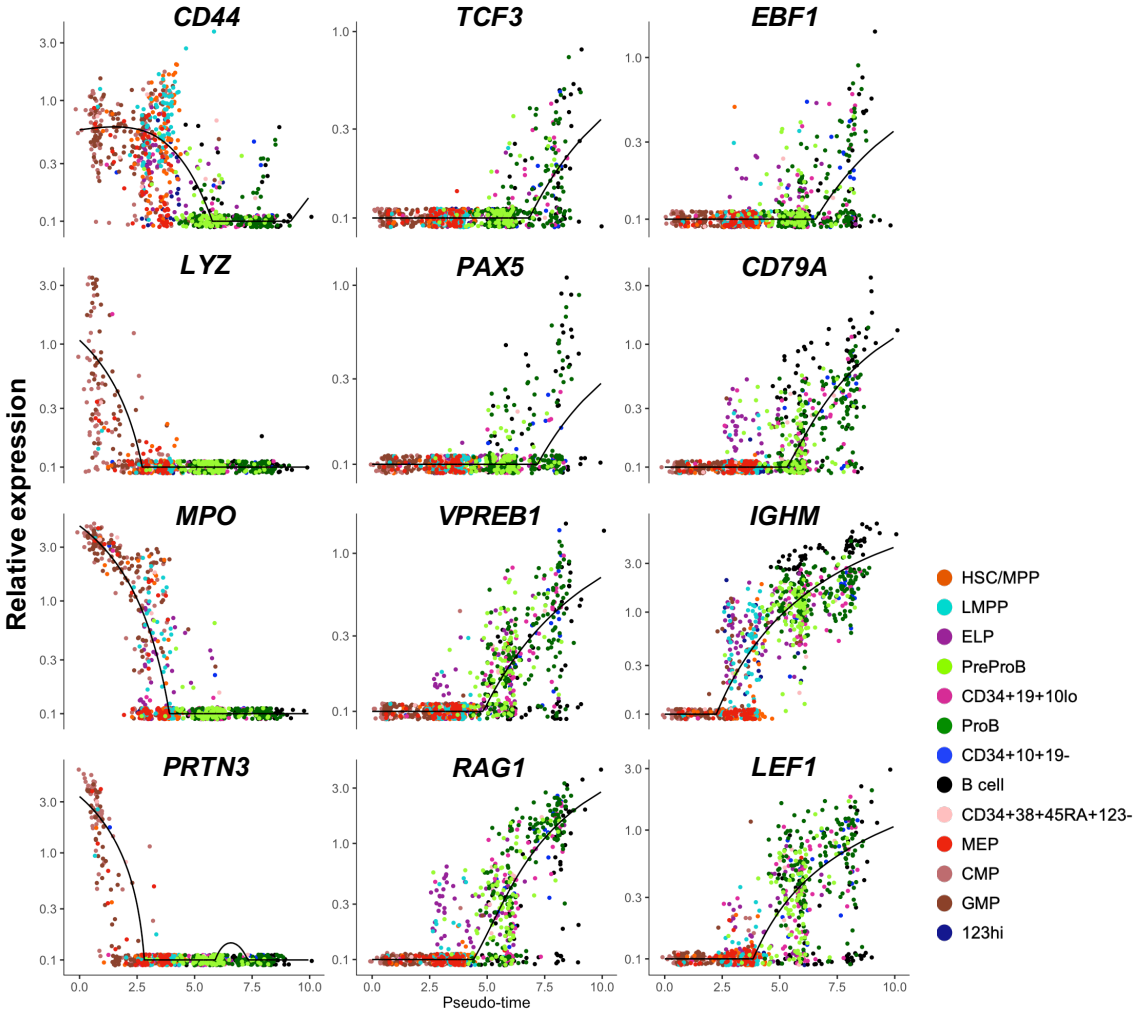


Figure 3.30: Relative expression of key genes involved in driving the pseudotemporal ordering of single cells in Monocle trajectory analysis. Gene expression determined by single cell RT-qPCR of index sorted fetal BM HSC and progenitors that were retrospectively labelled according to their immunophenotype (figure 3.26). Lin2-CD34+ cells from two biological samples were sorted to a total of 938 cells.

Finally, t-SNE is a multi-dimentional reduction algorithm commonly used to establish which cells or groups of are related to each other according to their gene expression/transcriptome. In theory such a multi-dimensional reduction technique would be useful in identifying groups of closely related cells. Therefore, I performed t-SNE analysis using the single cell RT-qPCR data of 96 genes from 938 index sorted cells to see if retrospective labelling according to immunophenotypes represented distinct populations. Since the B cells sorted for the purpose of trajectory analysis and comparison with B progenitors are CD34-, it was expected that they would cluster closely and separately from all CD34+ cells and potentially mask smaller differences between more closely related progenitor populations. Therefore I ran the t-SNE with and without B cells as shown in figure 3.31. Both of these analyses show the close relationships between each progenitor population in human fetal B-lymphopoiesis and also within the myeloid and erythroid progenitor compartments. ProB progenitors largely cluster separately but are closest to PreProB progenitors and CD34+CD19+CD10low B progenitors. This analysis reflects the continuum of B-lymphopoiesis observed in the pseudotime analysis as well as the distinction from other principle lineages: myeloid and erythroid (figure 3.31).

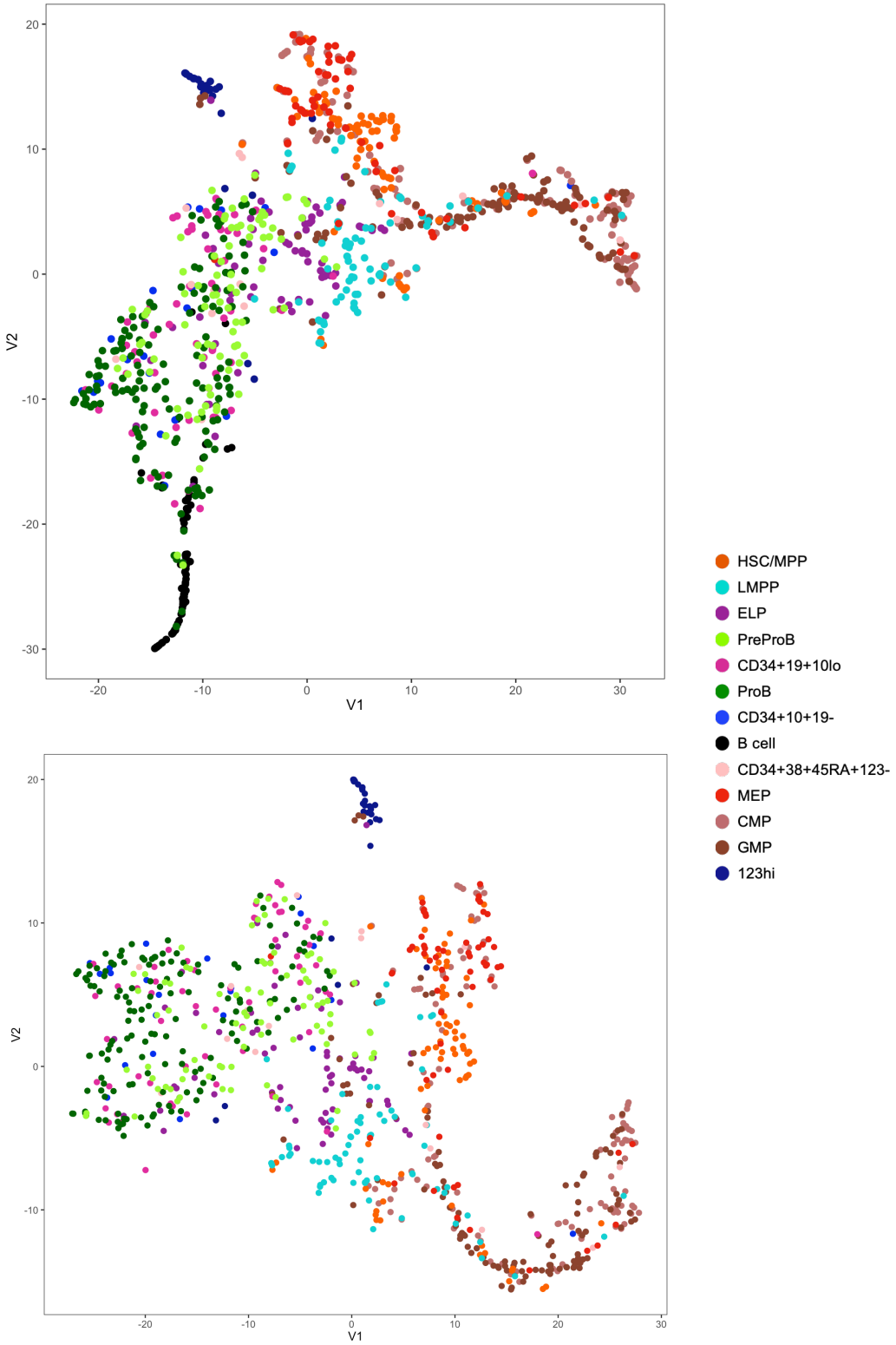


Figure 3.31: t-SNE of 938 single cells (2 biological replicates) according to the expression of 96 genes determined by single cell RT-qPCR. The top panel shows t-SNE analysis with B cells and the bottom panel shows t-SNE analysis without B cells. HSC and progenitors cluster according to their lineage (B-lymphoid, myeloid or erythroid) and reflects the continuum of B-lymphopoiesis observed in the pseudotime analysis (Destiny and Monocle analysis above).

3.6.3 IgH locus rearrangements

To make the vast repertoire of antibodies required for life, B cells start rearranging their IgH locus as they mature through a process called VDJ recombination [Roth, 2014]. This is a hallmark of B cell development and maturation. The RNAseq and single cell RT-qPCR data both confirmed that genes required for VDJ recombination and B cell receptor development (*RAG1, RAG2, DNTT, VPREB1*) begin to be expressed from ELP onwards (figure 3.32(a) and (b)). Therefore, to establish when heavy chain recombination begins in fetal B-lymphopoiesis and thereby demonstrate hierarchical relationships between the B progenitors, I investigated the re-arrangement of the IgH locus in sorted progenitor populations. To do this I whole genome amplified (WGA) DNA from sorted progenitor populations (100-10000 cells) and used PCR/heteroduplex analysis to look for IgH rearrangements. This work was done in collaboration with Gary Wright at Great Ormond Street Hospital who performed the heteroduplex analysis and sequencing on the WGA DNA I prepared from each flow sorted population.

All VDJ rearrangements detected from all populations and samples tested (n=5) are detailed in table 3.3. No rearrangements were detected in any LMPP samples. Partial IgH rearrangements in the form of DJ rearrangements were detected in 1/3 ELP, 2/4 PreProB and 2/5 ProB progenitors. Full VDJ rearrangements were detected in ProB progenitors and B cells only. These data are presented as a % of samples tested in figure 3.32(c) and demonstrate the proposed B-lymphoid hierarchy where ELP and PreProB progenitors lie upstream of ProB progenitors.

Table 3.3: VDJ rearrangements of IgH locus in whole genome amplified DNA of bulk sorted progenitor populations from 5 biological replicates. Where rearrangements were detected fragments were excised and sequenced.

Sample ID/gestation	**Population**	**Cells sorted**	**results**	**Rearrangements sequenced**
SAMPLE 1	LMPP	1118	NIL	None detected on screen
19 wk	PreProB	2562	NIL	None detected on screen
	ProB	1960	DJ	IGHD6-13*01/IGHJ4*02
	B cells	562	VDJ	IGHV1-45*02/IGHD7-27*01/IGHJ6*02, IGHV3-30*18/IGHD3-10*01/IGHJ6*02, IGHV4-59*01/IGHJ4*02
SAMPLE 2	LMPP	191	NIL	None detected on screen
22 wk	PreProB	7258	DJ	IGHD1-7*01/IGHJ5*02
	ProB	5524	DJ	IGHD7-27*01/IGHJ3*01
	B cells	74	FAILED	
SAMPLE 3	LMPP	2375	NIL	None detected on screen
17 wk	ELP	824	DJ	IGHD2-2*02/IGHJ4*02, IGHD6-6*01/IGHJ4*02
	PreProB	5692	DJ	IGHD1-26*01/IGHJ5*02, IGHD1-26*01/IGHJ3*02, IGHD4-4*01/IGHJ5*02
	ProB	26225	VDJ	IGHV3-30-3*01/IGHD1-14*01/IGHD1-26*01/IGHJ2*01, IGHD7-27*01/IGHJ4*02
	B cells	948	VDJ	IGHV2-5*01/IGHD6-13*01/IGHJ4*02, IGHV4-59*01/IGHD7-27*01/IGHJ2*01
SAMPLE 4	LMPP	2623	NIL	None detected on screen
20 wk	ELP	1167	NIL	None detected on screen
	PreProB	1313	FAILED	
	ProB	9695	VDJ	IGHV2-5*01/IGHD3-3*01/IGHJ4*02, IGHV3-9*01/IGHD6-19*01/IGHJ5*02, IGHV3-48*03/IGHD1-1*01/IGHJ5*02
SAMPLE 5	LMPP	4561	NIL	None detected on screen
20 wk	ELP	2454	NIL	None detected on screen
	PreProB	2221	NIL	None detected on screen
	ProB	4581	VDJ	IGHV3-30-3*01/IGHD2-15*01/IGHJ4*02

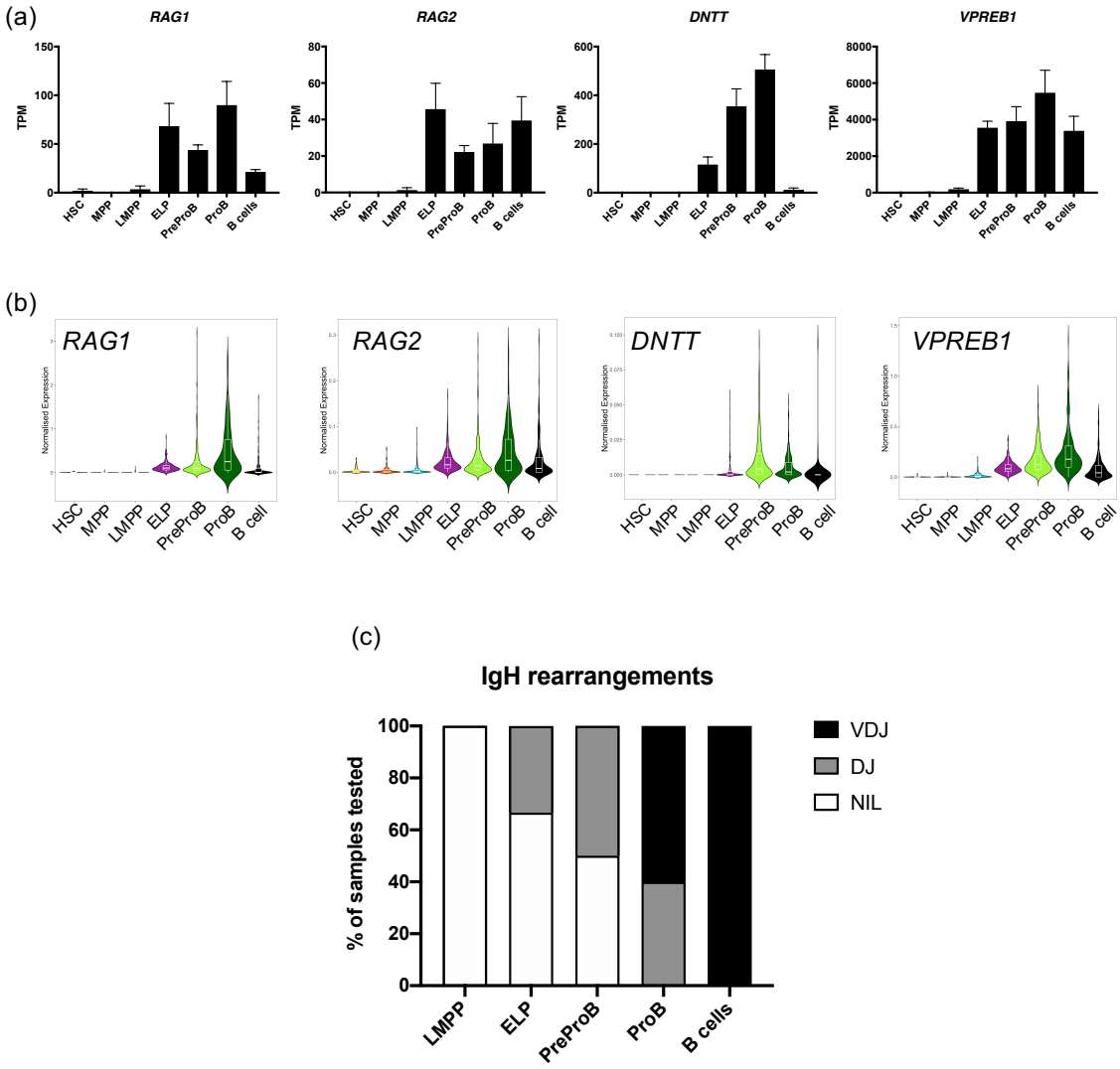


Figure 3.32: Genes associated with active VDJ recombination (*RAG1, RAG2, DNTT* and *VPREB1* are up-regulated from ELP onwards in the B-lymphopoiesis hierarchy by (a) bulk RNAseq and (b) single cell RT-qPCR analysis. (c) Percentage of DJ and VDJ rearrangements detected in bulk sorted progenitor populations n=3-5.

3.6.4 ATAC-Sequencing/Chromatin accessibility

To investigate whether the transcriptomic differences between PreProB and ProB progenitors are linked to differences in chromatin accessibility I performed ATAC-Sequencing (Assay for Transposase-Accessible Chromatin using sequencing; ATAC-Seq) adapted for small cell numbers [Corces et al., 2016] which was further adapted by a DPhil student in the lab: Dr Catherine Garnett, unpublished data). I sorted 2000 PreProB or ProB progenitors from 3 biological samples and 5000 Lin2+ fetal BM MNC from one biological sample as a control. After mapping and labelling peaks according to the closest gene, differential binding analysis using DiffBind [Ross-Innes et al., 2012] revealed that there were 351 differentially accessible elements in the vicinity of 341 genes (figure 3.33, full differential table available in online supplemental files). Manual inspection of these differentially accessible peaks suggested that despite being statistically significant, most differences in chromatin accessibility were very subtle. Nevertheless, 7 of these differentially accessible peaks were at the transcription start site of a gene (figure 3.34). Of these genes (*CUEDC1, CDH18, ST3GAL4, KCNN2, FUT9, GPR85, BRINP1*) the differential accessibility at *FUT9*, the gene that encodes key myeloid marker CD15, complements the transcription data that shows a residual myeloid gene expression programme in PreProB progenitors. Of the 341 differentially accessible genes there were also some genes known for their involvement in haematopoiesis: *GFI1, EBF1, GP9, LEF1* (figure 3.35) and the subtle differences in accessibility of these genes correspond with the relative position of PreProB and ProB progenitors in the B-lymphoid hierarchy.

Both PreProB and ProB progenitors had open chromatin upstream of or overlapping with the gene bodies of key B cell development genes including *PAX5, CD79A, RAG1, MME/CD10, DNTT* and *TCL1A* (figure 3.36) consistent with an active B cell programme and commitment to the B lymphoid lineage. Stem cell and myeloid genes expressed in PreProB progenitors (*LIN28B, CSF1R, MPO*) (figure 3.36) demonstrate qualitatively higher chromatin accessibility in PreProB progenitors compared to ProB progenitors. Finally B cell specificity is demonstrated by the lack of open chromatin at these B genes in the Lin2+(CD2, CD3, CD14, CD16, CD56 and CD235a) MNC ATAC-Seq trace (figure 3.36).

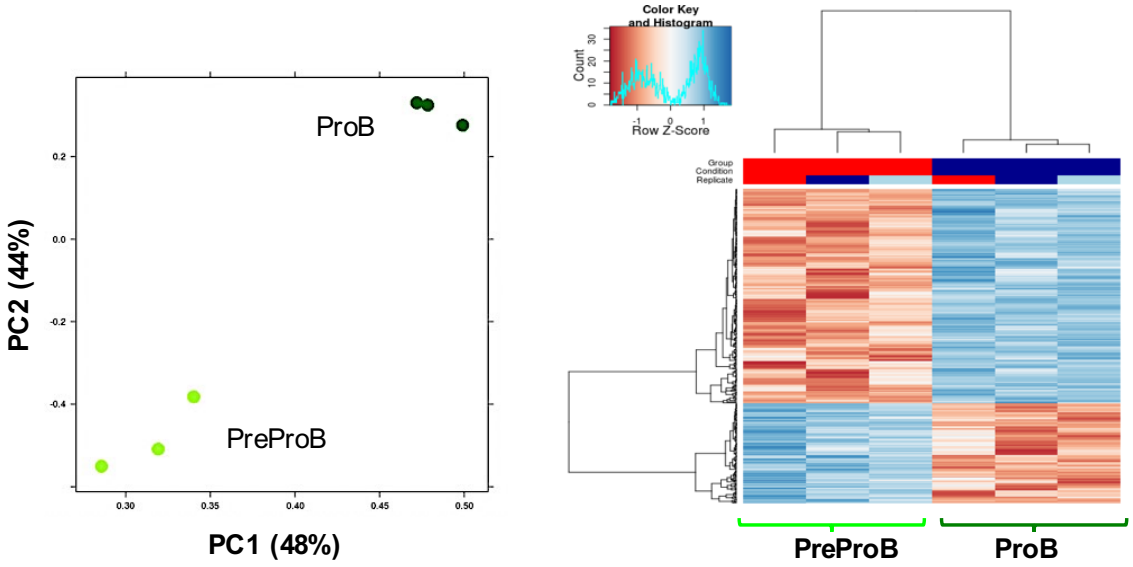


Figure 3.33: Differential chromatin accessibility analysis of ATAC-seq from PreProB and ProB progenitors summarized by PCA (left) and heatmap (right) using DiffBind. Details of differentially accessible genes available in online supplemental files. 2000 PreProB and ProB progenitors were sorted from 3 biological replicates and immediately processed for ATAC-Seq.

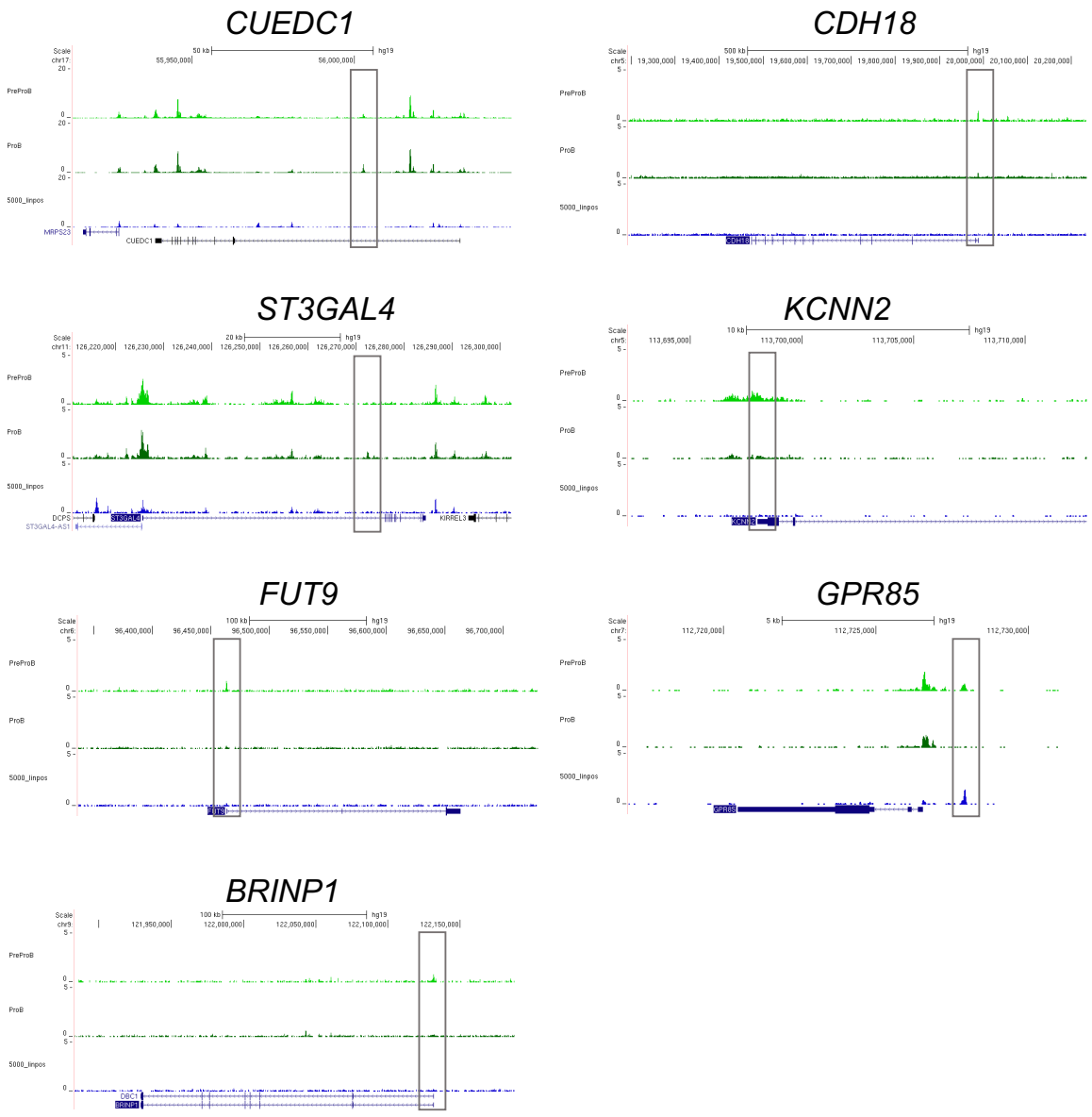


Figure 3.34: DiffBind analysis between PreProB and ProB progenitor ATAC-seq revealed 7 significant differences in chromatin accessibility at the transcription start site of a gene. These genes are shown here, with grey boxes indicating significantly differentially accessible region. FDR <0.1; data indicative of 3 biological replicates for PreProB and ProB progenitors and 1 biological sample for Lin2+MNC.

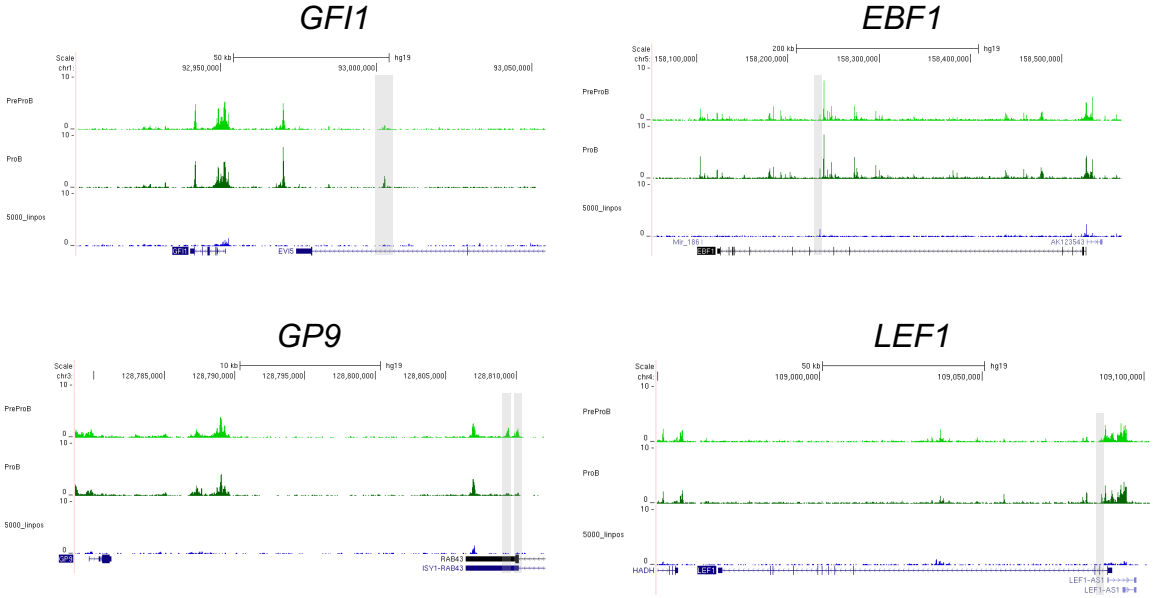


Figure 3.35: Diffbind analysis between PreProB and ProB progenitor ATAC-Seq revealed significant differences in the chromatin accessibility in the gene body of 4 genes known to have a role in haematopoiesis: *GFI1, EBF1, GP9* and *LEF1.* These genes are shown here with grey boxes to indicate differentially accessible region demonstrating the subtle but significant differences. FDR <0.1; data indicative of 3 biological replicates for PreProB and ProB progenitors and 1 biological sample for Lin2+MNC.

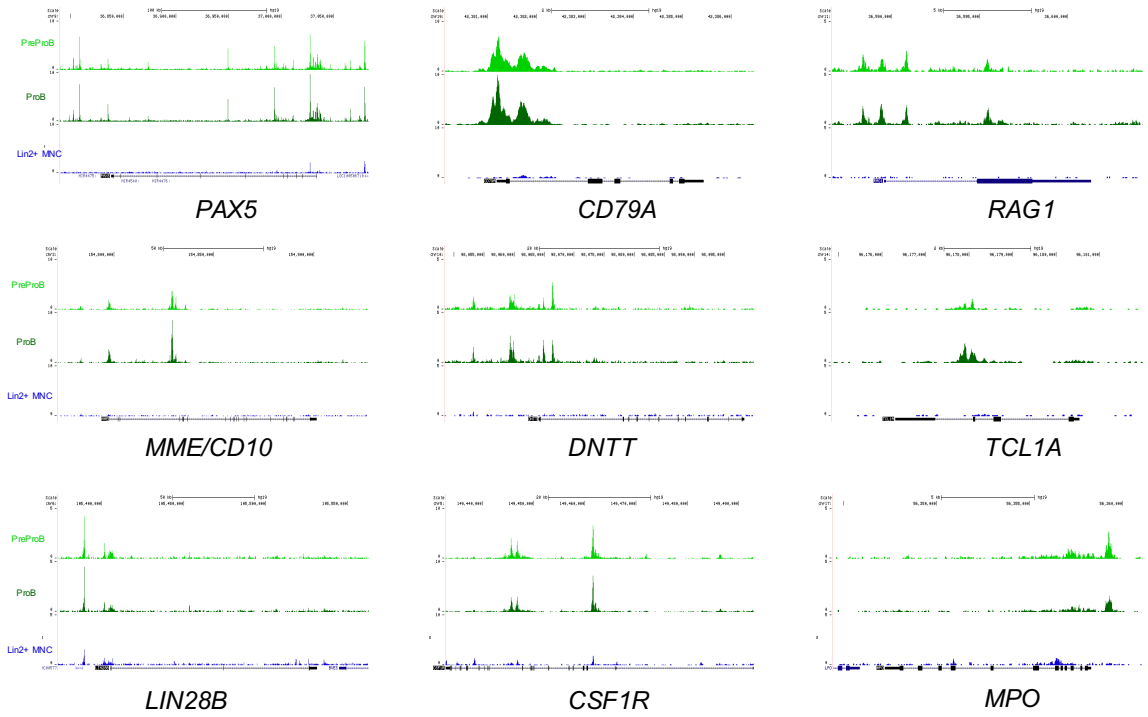


Figure 3.36: Chromatin accessibility at genes in PreProB progenitors (n=3), ProB progenitors (n=3) and Lin2+MNC (n=1) showing open chromatin at key B cell genes in committed B progenitors (PreProB progenitors and ProB progenitors) compared to Lin+MNC.

To ensure and independently confirm that any differences between chromatin accessibility in PreProB and ProB progenitors are subtle, I used deepTools to plot heatmaps of chromatin accessibility across the whole genome (figure 3.37) and HSC genes (3.38, [Chen et al., 2014]). Consistent with my analysis using DiffBind, there were no gross differences in chromatin accessibility between PreProB and ProB progenitors confirming that the differences seen were indeed subtle with the most obvious differences arising from different biological samples.

Taken together, these data show that chromatin accessibility in fetal BM progenitors reflects the transcriptomic profile of these cells and is cell context-dependent (by comparison with the Lin2+ MNC), demonstrating subtle differences even between closely related B-progenitor populations.

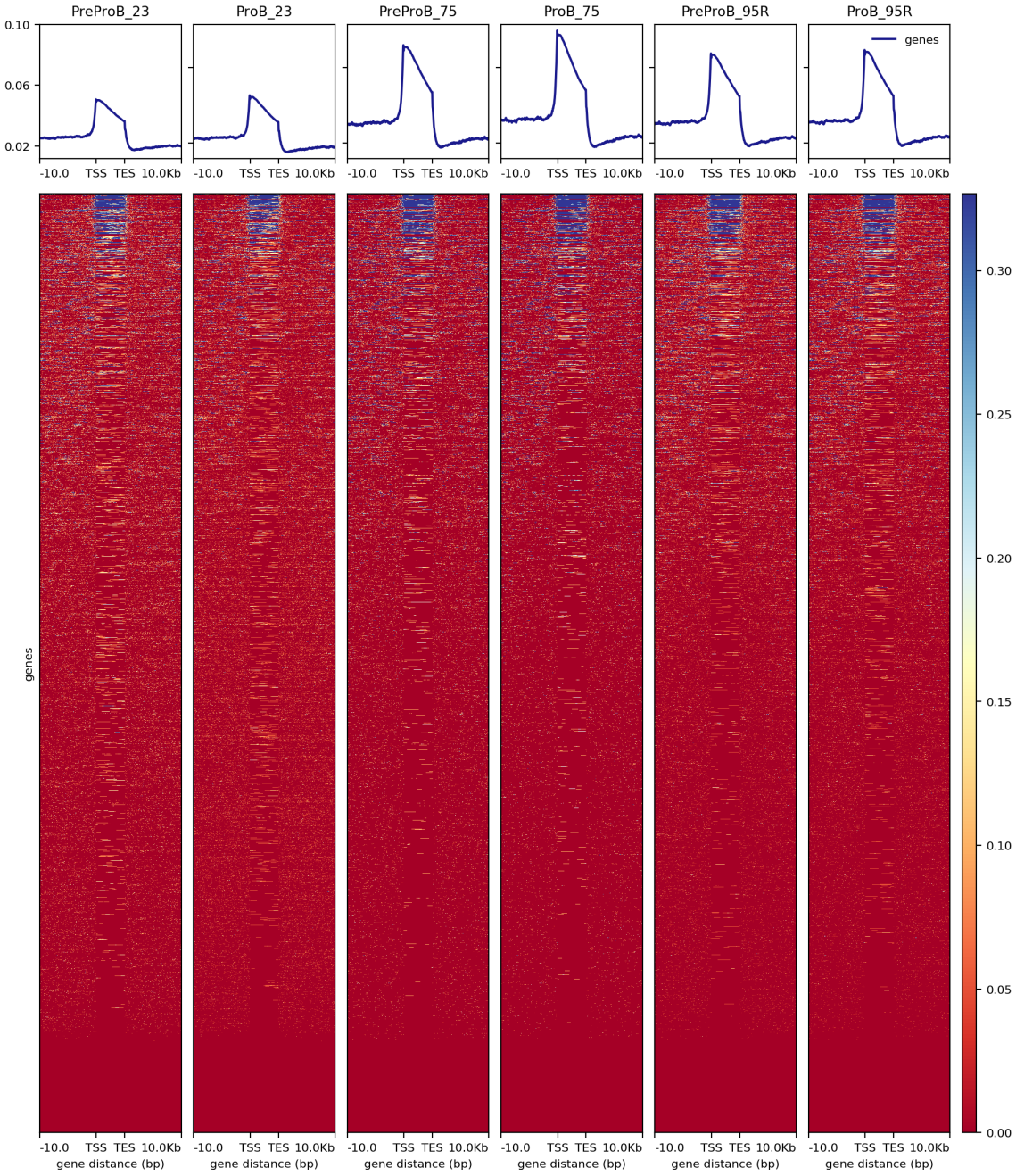


Figure 3.37: Heatmap of ATAC-Seq of PreProB and ProB progenitors produced using deepTools. The heatmap shows chromatin accessibility for each gene in the genome (one gene per row) scaled to transcription start sites (TSS) and transcription end sites (TES). There are no gross differences in chromatin accessibility across the whole genome between PreProB and ProB progenitors with the greatest difference being between biological samples.

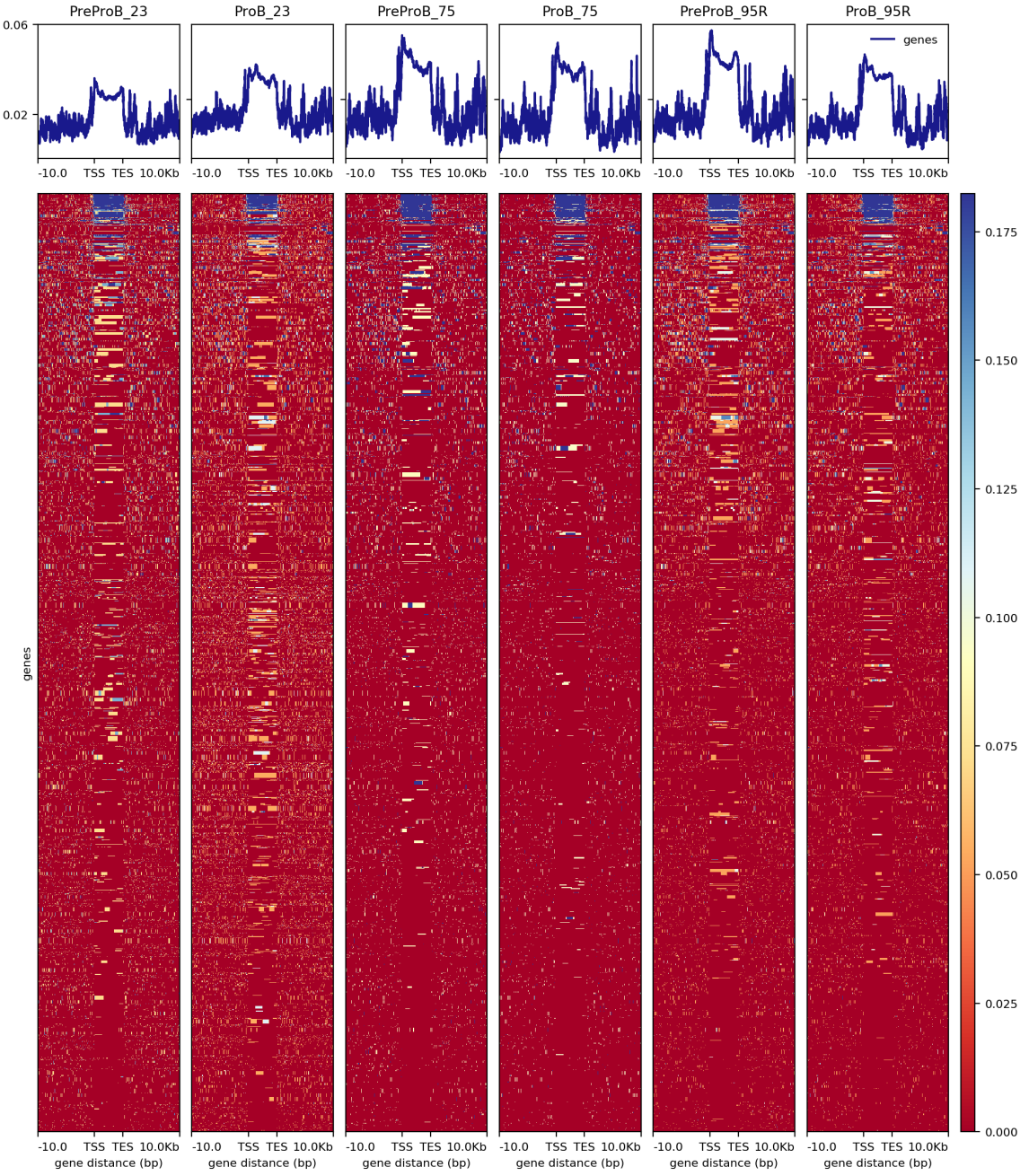


Figure 3.38: Heatmap of ATAC-Seq of PreProB and ProB progenitors produced using deepTools. The heatmap shows chromatin accessibility for HSC genes (one gene per row; HSC list from [Chen et al., 2014] scaled to transcription start sites (TSS) and transcription end sites (TES). There are no gross differences in chromatin accessibility in HSC genes between PreProB and ProB progenitors with the greatest difference being between biological samples.

3.7 Functional and molecular characterisation of adult BM PreProB progenitors: comparisons with fetal BM

To address whether fetal BM PreProB progenitors have features that are specific to fetal B-cell development, I compared their frequencies, functional and molecular characteristics with that of their very rare adult counterparts. After performing immunophenotypic analysis on paediatric and adult BM, I performed MS-5 co-culture assays to investigate B cell differentiation potential and single cell RT-qPCR to compare gene expression patterns in adult and fetal PreProB progenitors.

3.7.1 Immunophenotypic characterisation of fetal BM: comparison with post-natal BM

To compare frequencies of B progenitors, I performed flow cytometry analysis on paediatric BM (n=4) and adult BM MNC (n=7) from healthy donors using the same panels and protocols for fetal BM (n=31-35) flow cytometry analysis. Both ELP and PreProB progenitors showed a sharp and significant decline in frequency between fetal and adult BM. ELP were virtually absent in adult BM (0.33± 0.13% of Lin-CD34+ cells compared to 0.84± 0.32% of Lin-CD34+ cells in paediatric BM and 2.83± 0.38% of Lin-CD34+ cells in fetal BM). Furthermore, PreProB progenitors were present at an extremely low frequency in adult BM (1.03± 0.53%) and were also significantly reduced in paediatric BM (3.28± 0.57%; p= 0.003) compared to fetal BM. Another striking observation was the increased frequency of ProB progenitors in paediatric BM compared to both fetal and adult BM (3.39). This data has demonstrated for the first time that PreProB progenitors are highly enriched in fetal life and, while not entirely fetal-specific, might represent a previously unrecognised and functionally important population in normal human B lineage development.

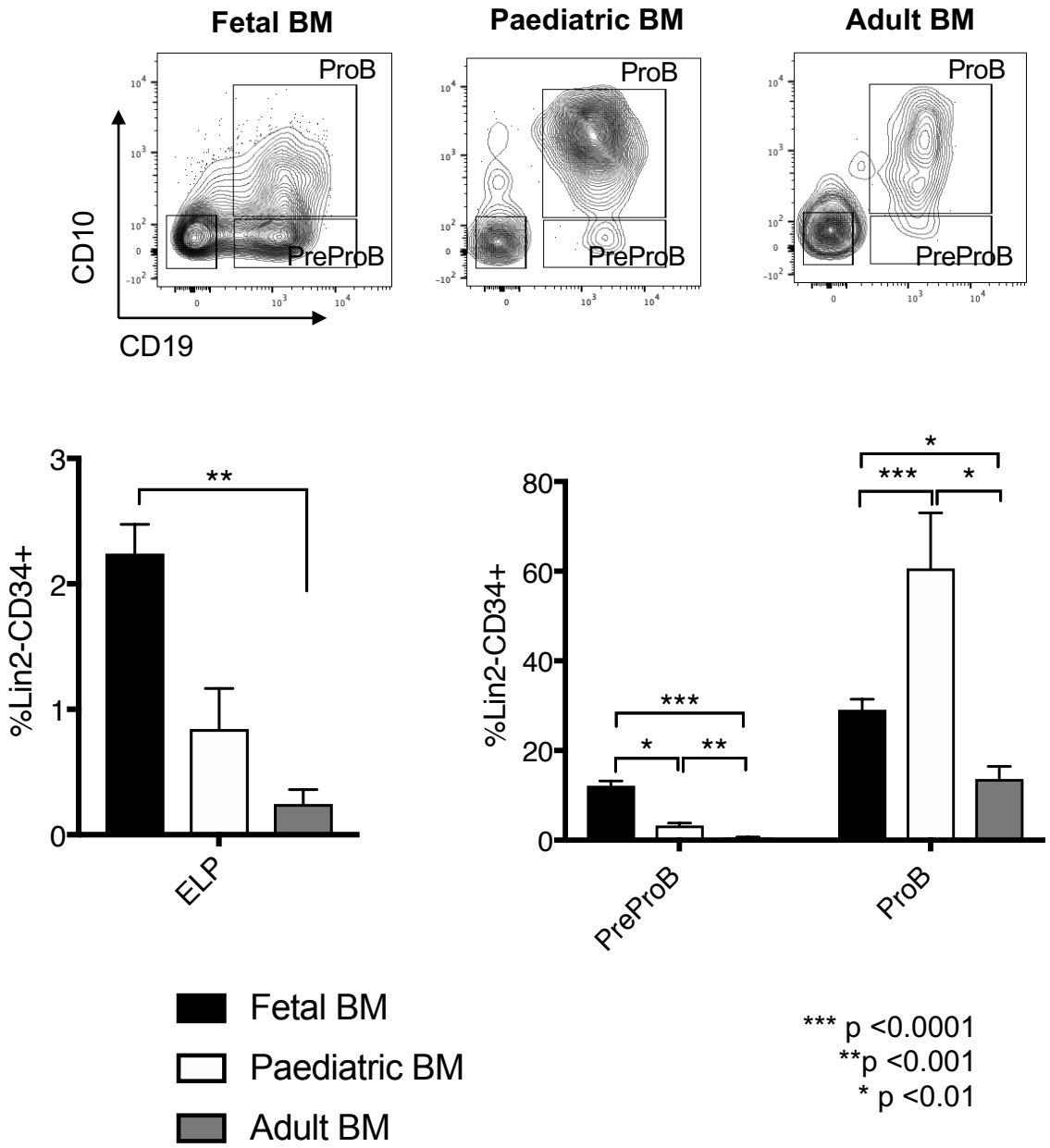


Figure 3.39: Representative flow plots (gated on Lin-CD34+) and frequencies of ELP, PreProB and ProB progenitors through ontogeny: fetal BM (n=31-35), paediatric BM (n=4) and adult BM (n=7).

3.7.2 B-lymphoid potential of adult ELP, PreProB and ProB progenitors

To establish the B/NK/myeloid potential of adult BM ELP and PreProB progenitors I sorted ELP, PreProB and ProB progenitors from 3 donors for MS-5 co-culture assays. Due to the rarity of ELP and PreProB progenitors in adult BM this proved very challenging (figure 3.40). Where possible 100 cells per well were sorted so that one whole well could be harvested for each time point and the summary of cell numbers sorted is in table 3.4. By day 7, adult BM PreProB (n=2) and ProB (n=3) progenitors differentiated into B cells only with no NK/myeloid output in a pattern consistent with fetal PreProB progenitors: PreProB progenitors differentiated into ProB and B cells and ProB progenitors differentiated into B cells (figure 3.41(a)). Using fold-change to account for the difference in number of PreProB and ProB progenitors plated, I observed only a modest expansion of adult PreProB progenitors on MS-5 compared to a fetal BM control (figure 3.41(b)). In constrast, ProB progenitors from both tissues did not exhibit any such expansion *in vitro*.

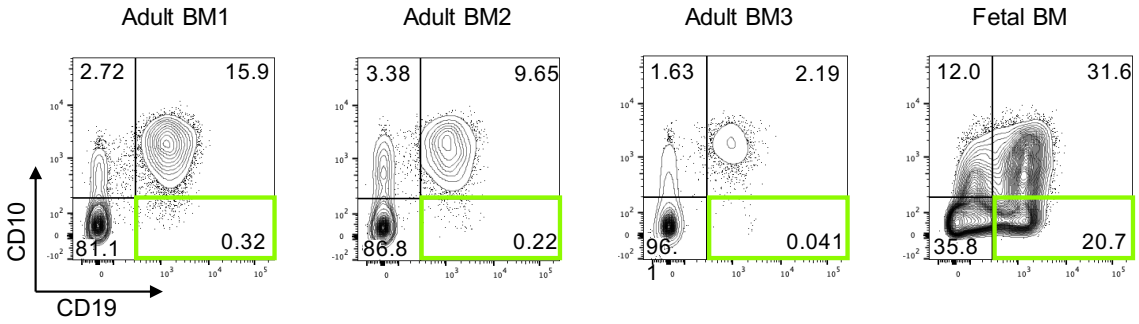


Figure 3.40: Representative flow plots of the Lin-CD34+ compartment in adult BM from three donors and one fetal BM for comparison. PreProB progenitors are indicated by the green box.

Table 3.4: Number of progenitors sorted for adult BM MS-5 co-cultures

	Lin-CD34+CD38-	ELP	PreProB	ProB
Fetal BM (ICH13195)	3 x 100	3 x 100	3 x 100	3 x 100
Adult BM1 (1509011089)	3 x 100	2 x 100	83	3 x 100
Adult BM2 (1509011090)	3 x 100	30	27	3 x 100
Adult BM3 (1509220016)	3 x 100	31	9	3 x 100

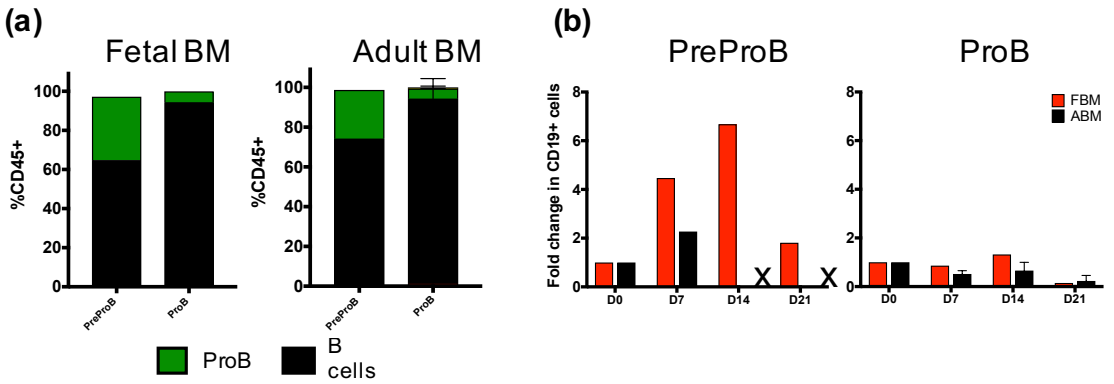


Figure 3.41: Adult BM PreProB progenitors differentiate into ProB and B cells like their fetal counterparts (a). However, their ability to proliferate on MS-5 is markedly reduced compared to their fetal counterparts (b). Adult BM, ABM; Fetal BM, FBM. X indicates "not measured" as a result of too few cells available for sorting.

In contrast to PreProB progenitors, adult BM ELP expanded at comparable levels when compared to a fetal ELP control (figure 3.42(a)). However, the identity of majority of these haematopoietic cells was undetermined at day 7, as they were not B cells (or B progenitors), NK cells or myeloid cells (figure 3.42(b)). Given that by day 14, the majority of the hCD45+ cells generated from adult BM ELP were CD19+ (1847 out of 2143 cells; figure 3.42(b)), it is possible that the early timepoint (D7) unidentifiable "other" cells detected in all three biological replicates of ELP/MS-5 co-cultures were a CD127+ population. MS-5 stromal cells are reported to produce 40x the physiological levels of IL-7 [Parrish et al., 2009] that is cross-reactive between human and murine species. Furthermore, the dependence of B-lymphopoiesis on IL-7 increases through ontogeny [Giliani et al., 2005], [Parrish et al., 2009], [Milford et al., 2016] therefore, it is possible that the MS-5 co-cultures caused an expansion of a CD127+ population (by their definition ELP express IL7 receptorα, CD127, on their surface, but the cells detected in co-culture were CD34-) which does not happen in MS-5 co-cultures using fetal ELP where IL-7 does not play such a pivotal role. Also, in these co-cultures adult ELP appeared to differentiate directly into B cells by day 14 (figure 3.42(b)). A lack of observed PreProB and ProB progenitors could be a result of inappropriate time points chosen for analysis (perhaps the dynamics of adult B-lymphopoiesis is slower) or because adult ELP are fundamentally distinct from fetal ELP and play a different role in adult lymphopoiesis.

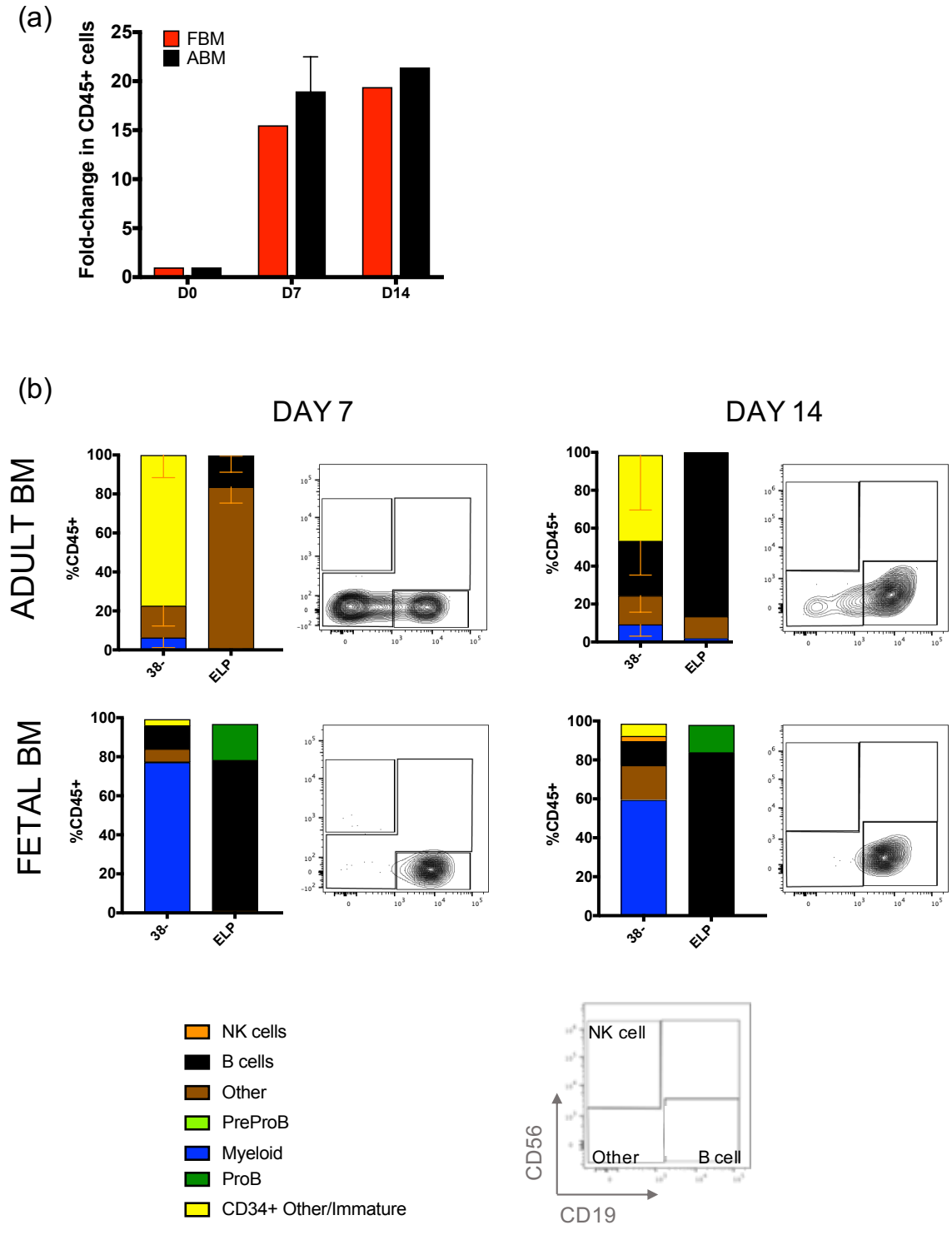


Figure 3.42: Adult ELP MS-5 co-cultures: (a) adult ELP (n=2-3) expanded *in vitro* at comparable levels to fetal ELP co-cultured in parallel (n=1). (b) Bar charts and representative flow plots showing adult ELP differentiated into an unknown CD34- population by day 7 (n=3) that disappeared day 14 (n=1). At day 14 most adult ELP had differentiated into B cells like fetal BM ELP (n=1). Due to rarity of adult ELP, I could not sort 100 ELP from each biological sample hence by day 14 n=1 for adult ELP. Key at bottom of figure denote colours used in bar charts and populations identifiable in representative flow plots.

3.7.3 Single cell RT-qPCR of adult BM B progenitors

Finally, I performed single cell RT-qPCR on flow-sorted adult BM HSPC to directly compare the B-cell developmental hierarchy between adult and fetal BM and identify any transcriptional differences. Using the sort strategy detailed in figure 3.43, I flow-sorted 758 cells from 2 biological samples for single cell RT-qPCR using the same panel of genes detailed in table 2.8 chosen for their role in B-lymphopoiesis and leukaemia. PCA demonstrated no major differences between adult and fetal HSPC, B progenitors or B cells at least by expression of the genes chosen (figure 3.44). A diffusion map suggested that adult PreProB progenitors might be more closely related to ProB progenitors (figure 3.45(a)) than their fetal counterparts; and this map coloured by pseudotime shows the differentiation trajectory expected (which is similar to fetal BM) (figure 3.45(b)).

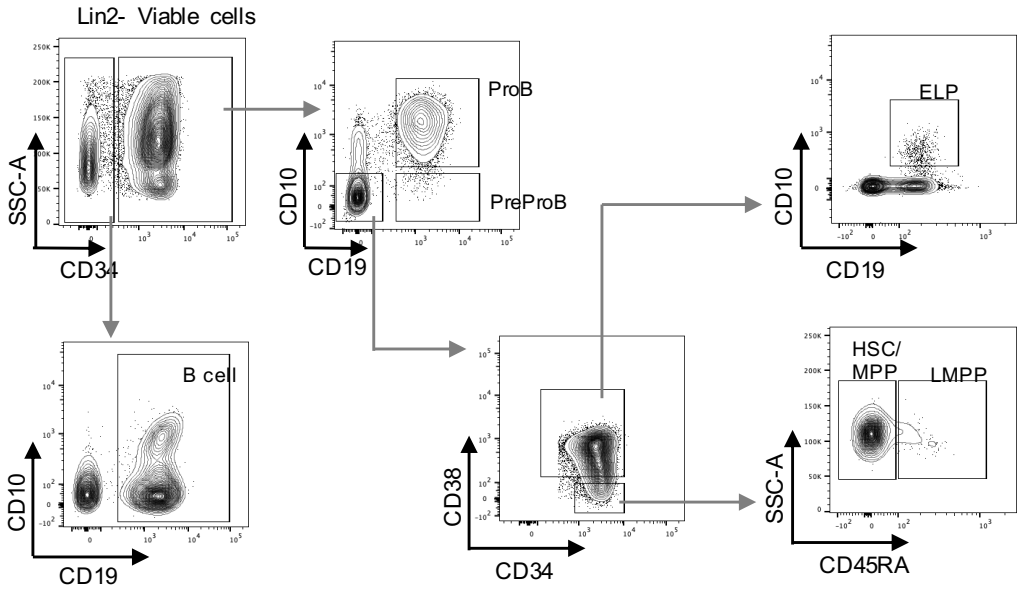


Figure 3.43: Sort gating strategy used to sort HSC/MPP, LMPP, ELP, PreProB, ProB progenitors and B cells from normal adult BM.

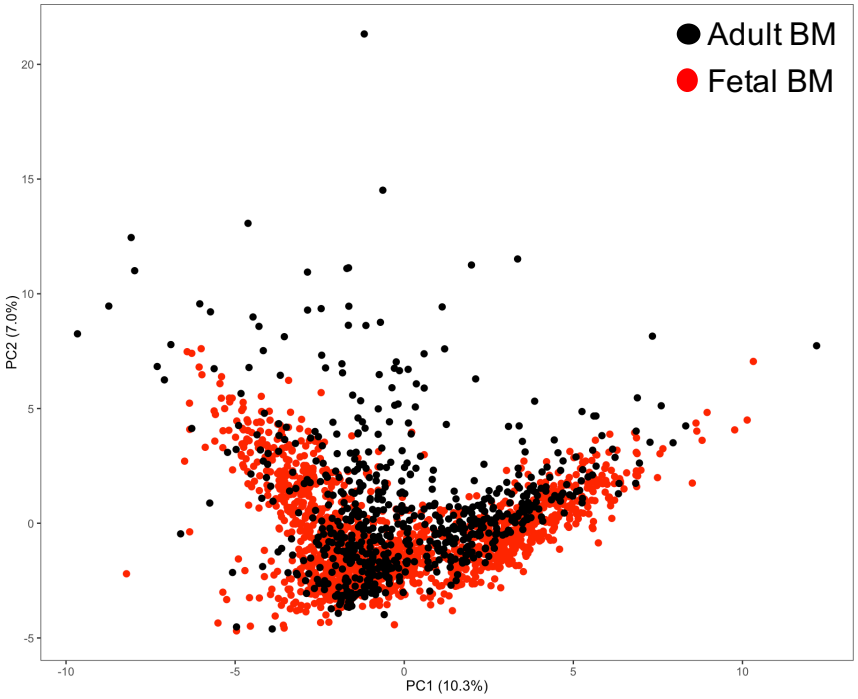


Figure 3.44: PCA shows no gross differences between gene expression measured by single cell RT-qPCR in adult BM (black) and fetal BM (red) HSPC. Each dot represents the data from a single flow sorted BM HSPC (n= 758 for adult BM cells from two donors; n= 1400 for fetal BM cells from three donors)

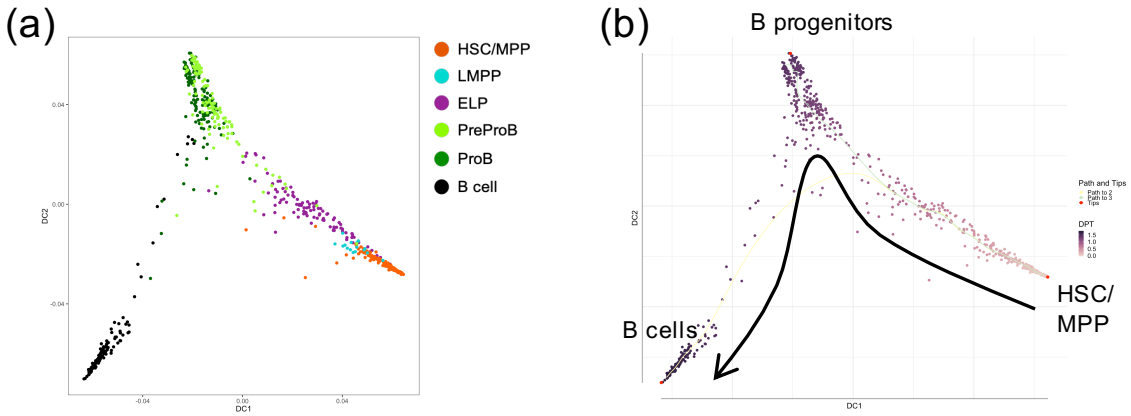


Figure 3.45: Diffusion maps of adult BM HSC and progenitor (n= 758 from two separate donors) single cell RT-qPCR coloured by (a) immunophenotype and (b) pseudotime.

Looking specifically at B lineage genes, adult and fetal PreProB progenitors express very similar levels of most B-lineage genes (*CD79A, MME, MS4A1, PAX5*; figure 3.46). However, compared to adult PreProB progenitors, fetal PreProB progenitors express higher levels of genes involved in DNA recombination (*DNTT, RAG1*) and some myeloid (*MPO*) and HSC/leukemia genes (*TCF3, LEF1, MEF2C, IL7RA*)(figure 3.46).

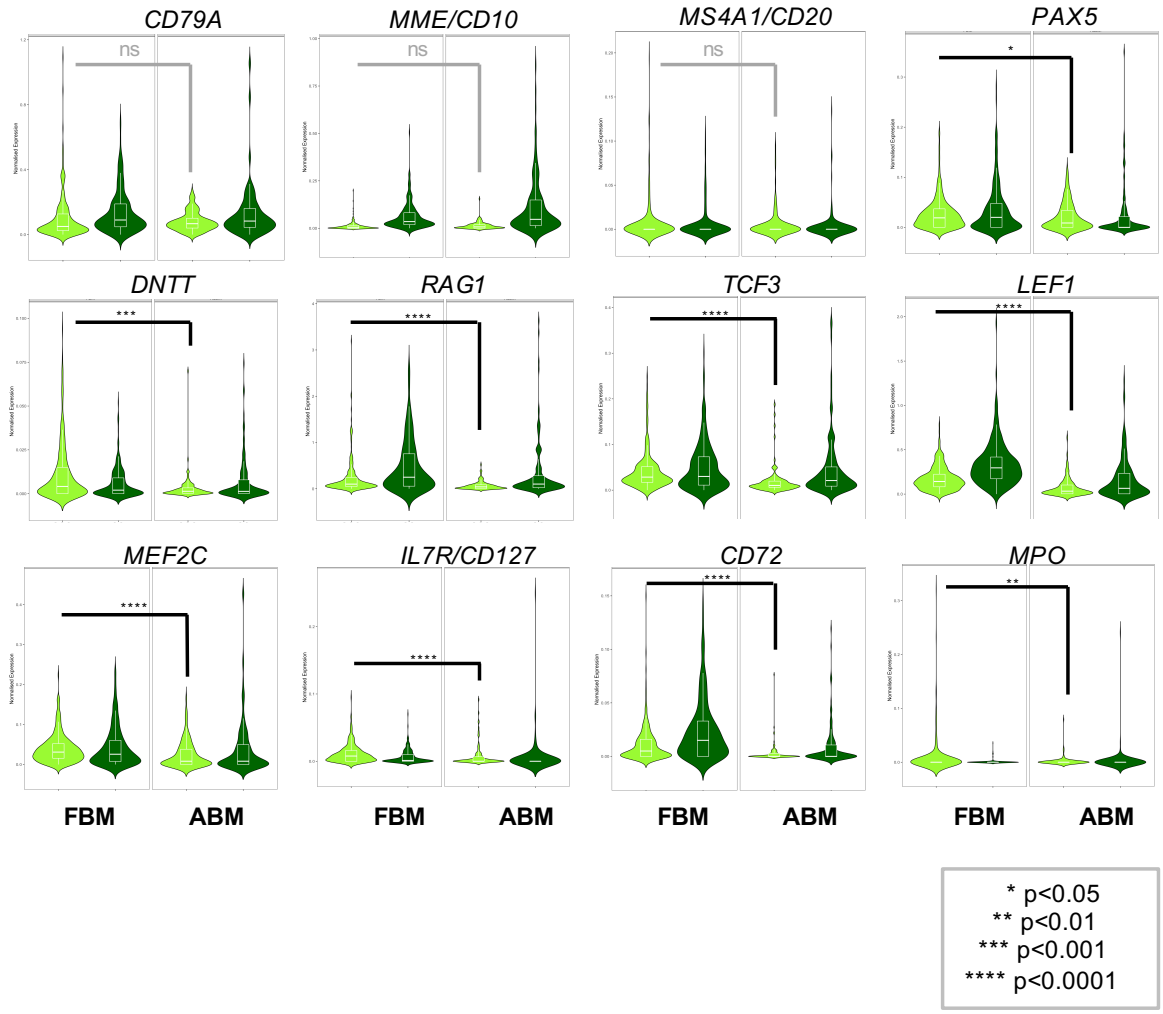


Figure 3.46: Violin plots showing gene expression of selected genes by single cell RT-qPCR. PreProB progenitors are coloured light green and ProB progenitors coloured dark green. Fetal BM, FBM, n=1400 from 3 donors; Adult BM, ABM, n=758 from 2 donors. Significance determined by non-parametric ANOVA

The overall pattern of gene expression across the B lineage hierarchy is similar in fetal BM and immunophenotypically identical adult BM progenitors, (figure 3.47). However, there are some notable differences in the level of expression between adult and fetal BM progenitors for early B genes such as *LEF1, PAX5, RAG1, RAG2, DNTT* and *TCF3* that could be a reflection of the polarised/primed B cell drive in the fetal BM microenvironment (figure 3.47).

There were also differences in the expression of genes associated with a more mature B cell phenotype (*CD72, CD79A, CD44, CD20*) which were expressed at higher levels in adult BM CD34- B cells perhaps because the CD34-CD19+ "B cells" in fetal BM mainly consist of CD10+ PreB progenitors [Roy et al., 2017] while the CD34-CD19+ "B cells" sorted from the adult BM are more mature CD10- cells just about to leave the BM for germinal centres (figure 3.48).

Finally, there were some striking differences in genes associated with fetal development (*LIN28B*), and stem cell identity (*PROM1, RUNX2*) and *CD9* which were strongly expressed in fetal BM HSPC compared to low or absent expression in adult BM HSPC (figure 3.49).

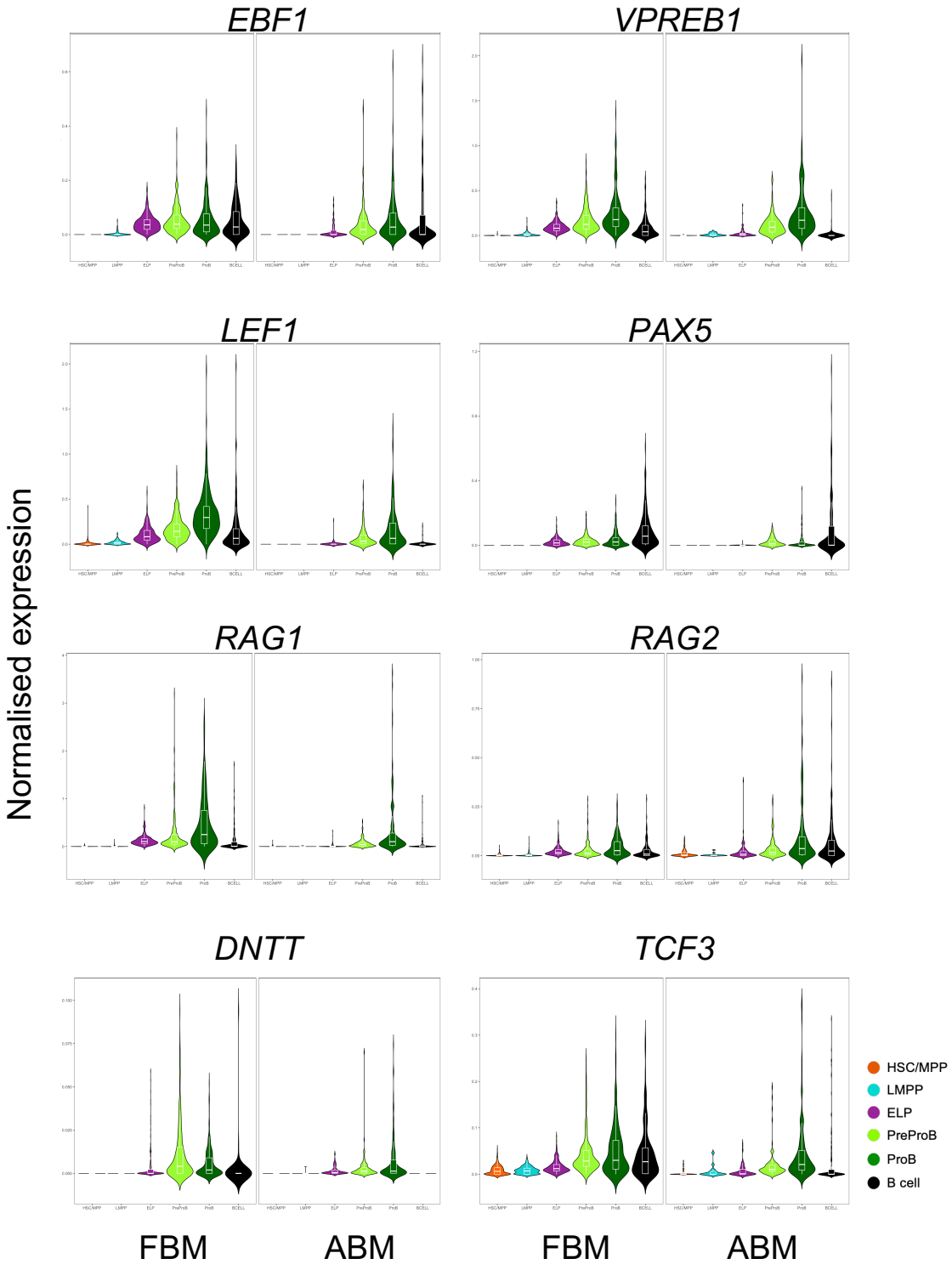


Figure 3.47: Single cell expression of genes (RT-qPCR) associated with B-lymphoid development in fetal BM (n=1400, 3 biological replicates, 17-20pcw, FBM) and adult BM (n=758, 2 biological replicates, ABM) sorted HSC and progenitors.

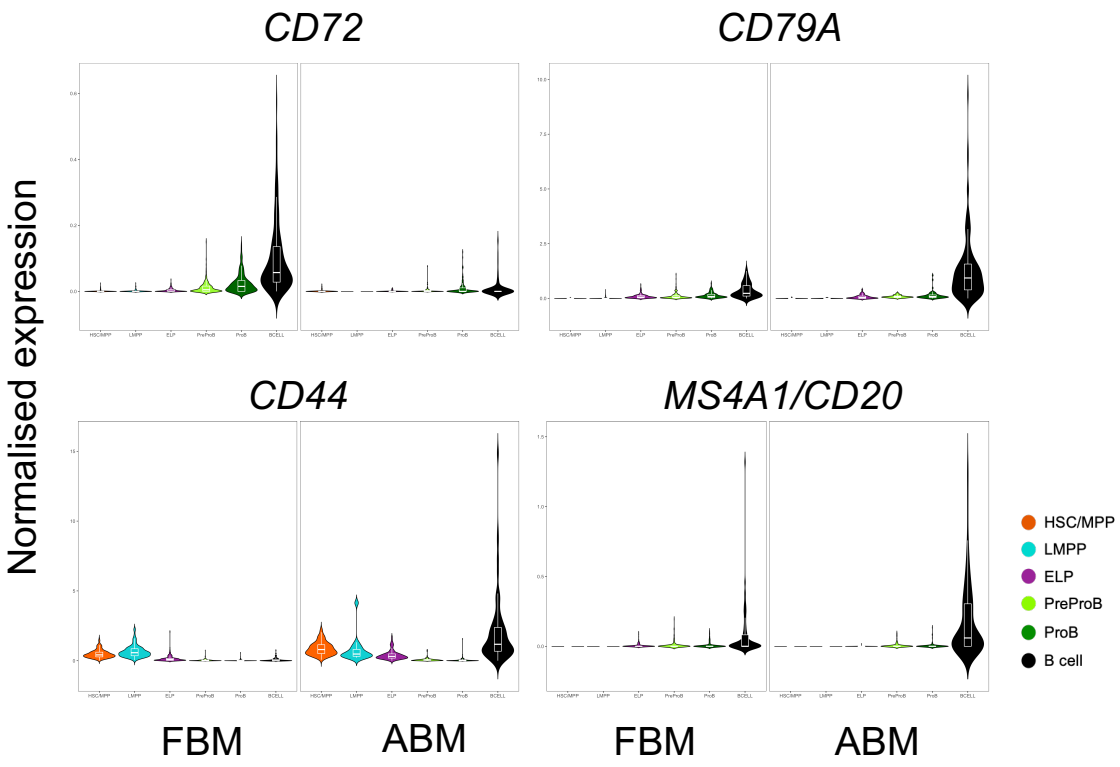


Figure 3.48: Single cell expression of genes (RT-qPCR) associated with mature B cells in fetal (n=1400, 3 biological replicates, 17-20pcw, FBM) and adult (n=758, 2 biological replicates, ABM) sorted HSC and progenitors. B cells sorted from fetal BM were mainly PreB cells and B cells sorted from adult BM were mainly immature B cells. This is reflected in the difference in expression of mature B cell genes between fetal and adult B cells.

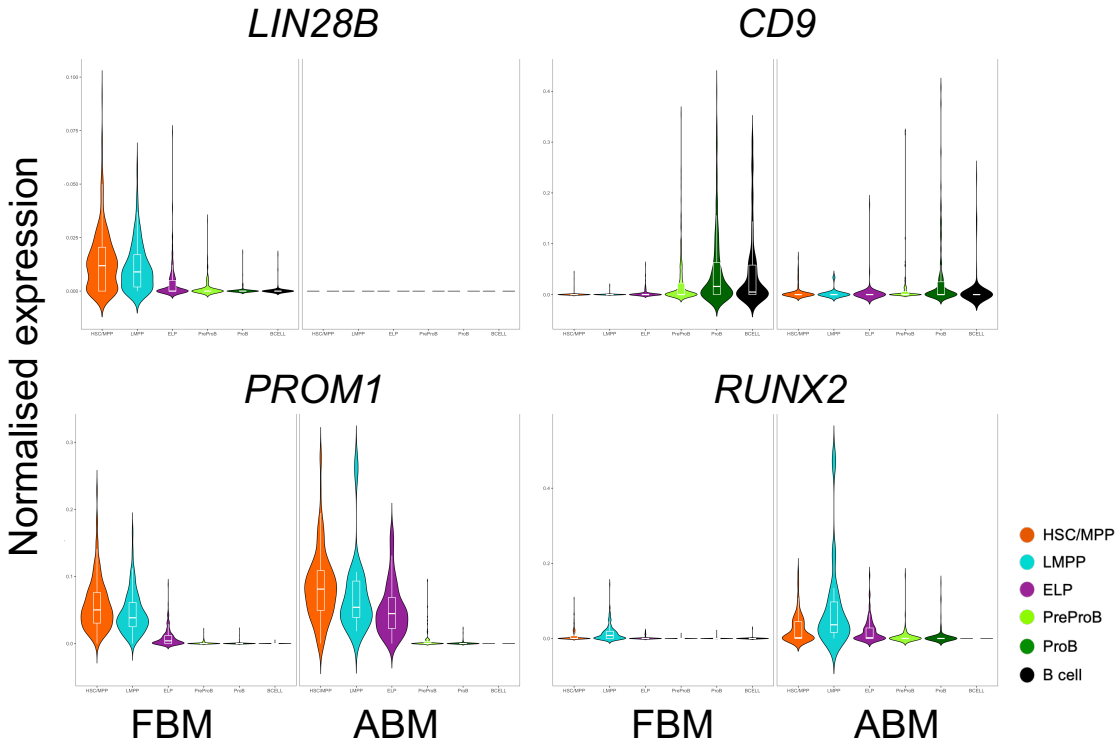


Figure 3.49: Single cell gene expression (RT-qPCR) of fetal BM (n=1400, 3 biological replicates, 17-20pcw, FBM) and adult BM HSPC (n=758, 2 biological replicates, ABM) showing selective expression of the fetal development gene (*LIN28B*) in fetal BM HSPC as well as increased expression of *CD9* in fetal BM compared to adult BM B progenitors and B cells. In contrast expression of other stem cell genes (*PROM1, RUNX2*) was higher in adult BM HSPC.

3.8 Summary and Discussion

Here I have characterised human fetal B-progenitor development in detail for the first time. The data have recently been published in the journal *Blood* [O'Byrne et al., 2019] and provide an important resource that could aid understanding of the *in utero* initiation of childhood leukaemia.

3.8.1 Fetal specific/enriched progenitors

Prior to these investigations, most of our understanding of fetal B-lymphopoiesis was based on murine studies and inference from studies of human CB [Sanz et al., 2010], [Montecino-Rodriguez and Dorshkind, 2012]. I set out to functionally and molecularly characterise B-lymphopoiesis in human fetal BM for the first time. Using flow cytometry I have shown that there are two cell populations that are enriched in fetal life: Early Lymphoid Progenitors (ELP) and PreProB progenitors. While ELP and PreProB progenitors had been previously noted in FL [Roy et al., 2012] and PreProB progenitors in CB [Sanz et al., 2010], their relationship to each other and place in the B-lymphoid hierarchy had not previously been investigated.

Using age matched FL and fetal BM samples, I show that while HSPC frequencies remain relatively constant in the FL, there is a massive expansion of B progenitors in the fetal BM at 11-12pcw. Furthermore, when compared to adult BM, the differences in frequency of ELP and PreProB progenitors are striking in the fetal BM. This is the first time such comparisons have been made and demonstrates the major differences in blood development that can occur through different stages of ontogeny.

3.8.2 Fetal B-lymphoid hierarchy

To establish where these progenitors lay in the fetal B-lymphoid hierarchy I performed extensive *in vitro* and *in vivo* differentiation assays to ascertain the B, NK, T, myeloid and erythroid potential of immunophenotypically defined HSC, MPP, LMPP, ELP, PreProB and ProB progenitors. HSC and MPP both demonstrated multi-lineage output as expected while LMPP could differentiate into lymphoid and myeloid lineages. ELP demonstrated a predominantly lymphoid differentiation potential with residual myeloid potential that was observed particularly in MSC co-cultures. PreProB and ProB progenitors only differentiated into B cells regardless of stromal feeder layer and did not grow in methylcellulose. Short term MS-5 and MSC co-cultures established that ELP can differentiate into PreProB and ProB progenitors and; PreProB progenitors

can differentiate into ProB progenitors. The reverse does not occur thereby confirming the hierarchy: ELP → PreProB → ProB → B cell.

One unresolved question is the hierarchy preceding ELP in fetal BM. During the optimisation of MSC and MS-5 co-culture assays, I also noted the presence of a population of CD10+ LMPP although the majority of fetal BM LMPP were CD10-. I also have preliminary data suggesting that most of the lymphoid potential of fetal BM LMPP lies in the CD10+ rather than the CD10-fraction. Over the course of this project this was confirmed independently in extensive single cell assays revealing the heterogeneity within the LMPP compartment in CB [Karamitros et al., 2018]. Overall, these data suggest that there may be 2 pathways of B cell development in fetal BM, one via CD10+ LMPP perhaps direct to ProB progenitors and the other via CD10- LMPP perhaps to ELP and PreProB progenitors but further work would be required to confirm this.

While xenograft assays showed that PreProB and ProB progenitors do not have long term engraftment potential (defined as the presence of >0.1% human CD45+ cells in BM 9-14 weeks after transplantation into NSG mice), short term xenograft assays demonstrated that PreProB progenitors only differentiated into B cells. Therefore, functional B cell hierarchy demonstrated by *in vitro* assays was recapitulated *in vivo*, suggesting that fetal progenitors have cell intrinsic mechanisms that dictate the fetal differentiation pathway independent of human fetal microenvironment.

The functional fetal B cell hierarchy was confirmed at a molecular level. Analysis of the transcriptome of sorted progenitors placed ELP and PreProB progenitors in a sequential order after LMPP and before ProB progenitors and this was confirmed by single cell RT-qPCR. The transcriptomes of ELP and PreProB progenitors were most similar to each other but there were important differences in the gene expression of genes associated with myeloid and T-lymphoid identity that corroborated with the functional data placing ELP upstream of PreProB progenitors. Although functionally identical, PreProB and ProB progenitors were also molecularly distinct according to their transcriptomes. Furthermore, analysis of chromatin accessibility between PreProB and ProB progenitors showed very subtle differences reflecting their commitment to B-lineage.

Functional and molecular comparisons with adult BM demonstrated that very rare

adult PreProB progenitors are also committed to the B-lymphoid lineage but probably do not have the same proliferative potential. Furthermore, these adult PreProB progenitors more closely resemble adult ProB progenitors than fetal counterparts according to single cell gene expression analysis. These data support the hypothesis that the rare PreProB progenitors sorted are more like ProB progenitors that were dim for CD10 expression rather than CD10- PreProB progenitors.

One of the other novel features of the work in this chapter was the development of a new fetal BM MSC: HSPC co-culture system which I hoped would more faithfully recapitulate the human *in vivo* BM microenvironment. Although these co-cultures were able to support fetal BM B cell development from ELP and B progenitors there were technical challenges with these co-cultures. In particular, regardless of tissue source (FL and fetal BM) or whether the MSC were irradiated to prevent overgrowth, MSC co-cultures were relatively poor at reliably supporting differentiation of HSC, MPP and LMPP to B cells which is in accordance with what Kurosaka *et al.* previously reported [Kurosaka et al., 1999]. The ability to make B cells appeared dependent on both MSC sample and HSPC sample (despite all HSPC making B cells on MS-5 in parallel): if a HSPC sample grew well on MS-5, it would also grow well on MSC but if a HSPC sample did not grow so well on MS-5, it was less likely to grow on MSC. It was also difficult to factor in the heterogeneity of MSC derived from different fetal samples. Despite this caveat, the co-culture system proved very useful in confirming the hierarchy of human fetal B-lymphopoiesis and in interrogating the role of the trisomic microenvironment (described and discussed in Chapter 4).

Having delineated B cell development in normal fetal BM in detail, this sets the scene for investigating the impact of T21 on B cell development in fetal BM. Previous work from the lab identified severe impairment of B progenitor development in FL in DS, but the work I have done in normal fetal BM clearly shows that the main site of B-lymphopoiesis in human fetal life is in the BM. In the following chapter, I go on to investigate the functional and molecular characteristics of B cell development in T21 fetal BM and discuss how this may link to the pathogenesis of DS-ALL.

3.8.3 Implications for childhood ALL

Many childhood B-ALLs are hypothesised to originate *in utero* from a fetal specific progenitor, which is transformed by in-frame chromosomal translocations and/or cooperating mutations to give rise to B-ALL in childhood. This hypothesis is supported

by twin studies [Ford et al., 1993], [Cazzaniga et al., 2011], [Alpar et al., 2015], and tracing of chromosomal translocations to birth through analysis of neonatal blood spots and cord blood[Gale et al., 1997], [McHale et al., 2003], [Mori et al., 2002] (recently comprehensively reviewed in [Greaves, 2018]).

This led me to speculate that PreProB progenitors and/or ELP could be the cell of origin for a particularly poor prognosis B-ALL that occurs in infants (iALL). iALL is a rare and hard to treat B-ALL with an immunophenotype that bears striking resemblance to PreProB progenitors: it is CD19+CD10-; is characterised by similar IgH rearrangement status; and is transcriptomically similar to fetal BM ELP and PreProB progenitors [Jansen et al., 2007] [Andersson et al., 2015]. Most cases of iALL are caused by an in-frame chromosomal translocation of the *KMT2A* gene that encodes the methyltransferase, MLL (iALL reviewed here: [Sanjuan-Pla et al., 2015]). This was of particular interest because even though chromosomal translocations are less common in DS-ALL, MLL rearrangements are extremely rare in this disease group [Buitenkamp et al., 2014] and there are no reported cases of iALL in a child with DS [Whitlock et al., 2005], [Zeller et al., 2005], [Qiao et al., 2018]. These observations, combined with our knowledge of the impact of fetal haematopoiesis on myeloid malignancy in DS [Tunstall-Pedoe et al., 2008], [Roy et al., 2012], makes the investigation of B-lymphopoiesis in T21 fetal life of particular interest (see Chapter 4).

3.8.4 Future work

The only difference in immunophenotype between PreProB and ProB progenitors is the expression of *MME*/CD10. Ideally, it should be possible to positively identify PreProB progenitors using a marker that is only expressed on PreProB progenitors. To begin to interrogate this, I filtered the RNA-Seq differential expression analysis between PreProB and ProB by genes that are expressed on the cell surface and found that 23 genes are significantly differentially expressed between PreProB and ProB progenitors (table 3.5). All of the genes that were over-expressed in PreProB progenitors are associated with stem, myeloid and T/NK lineage reflecting their position upstream of ProB progenitors in the hierarchy. Similarly, genes that were over-expressed in ProB progenitors have all been associated with B cell maturation. These genes now need to be tested for their cell surface expression by flow cytometry. One promising example is *PROM1*, which encodes the cell surface protein CD133, is over-expressed in ALL and is a direct target of the MLL-AF4 fusion protein [Kerry et al., 2017], [Godfrey, Crump *et al*, 2019, under review].

Table 3.5: Differentially expressed surface protein genes in PreProB/ProB progenitors. + log2FoldChange indicates over-expression in PreProB progenitors, - log2FoldChange indicates over-expression in ProB progenitors.

Gene	Alt gene name	log2FoldChange	padj	Expressed by
CD248		10.15861653	1.07E-05	Endothelial cells
CD244		10.0875812	1.84E-14	NK cells, T cells, dentritic cells, myeloid
ABCB1	CD243	9.672088488	0.028145788	NK cells
CR1	CD35	7.846696572	0.024186683	Haematopoietic cells
CD226		7.716869121	0.016576353	Lymphoid, stem, macrophage and platelets
CD3D		7.467829333	0.002367528	T cells
PROM1	CD133	7.462433136	2.69E-37	Stem cells
MSR1	CD204	7.327919112	0.012831048	Macrophages
CD300A		6.394380147	0.039632748	Lymphoid and monocytes
CSF1R	CD115	6.343644291	3.44E-05	Myeloid cells, platelets
ITGA3	CD49c	6.081303141	0.076488074	T cells , B cells and macrophages
GP9	CD42a	5.809350509	0.000189428	Platelets
CD7		5.509374221	0.03563293	T cells , NK cells and stem cells
IL6R	CD126	4.835371298	0.039524417	T cells, B cells and Stem cells (myeloid prog?)
CSF3R	CD114	4.204876699	0.069619883	Macrophages
SEMA4D	CD100	1.49306235	0.047503447	Lymphoid and macrophages
LILRB2	CD85D	-1.386291106	0.017033387	Myeloid and dendritic cells
CD72		-2.08332243	0.005873953	B cells, stem cells, macrophages
TNFRSF21	CD269	-2.267815451	1.96E-06	B cells
CD9		-3.183286919	4.11E-10	Platelets, myeloid, T cells and B cells
MME	CD10	-3.497277922	0.003922852	B and myeloid progenitors, NK cells
PTPRJ	CD148	-5.606124216	3.43E-09	Haematopoietic cells
CD40		-9.445158084	0.004266378	B cells, dendritic cells , stem and macrophages

Another area of future study, as mentioned above, is to clarify the relationship between LMPP and ELP, i.e. to ascertain whether all LMPP can give rise to ELP. Furthermore, I also observed the presence of a Lin-CD34+CD19-CD10+ population that was more obvious in the fetal BM of later gestation samples (>17pcw). The literature would suggest that these cells are immunophenotypic common lymphoid progenitors (CLP) but my preliminary data suggests that they are similar to ELP as they still have some residual myeloid potential. Also, RT-qPCR gene expression analysis suggests that this population also has a lymphoid bias. My investigations to date have focussed on prospective isolation of cell populations by flow cytometry and retrospective classification of these cells based on their differentiation potential largely in *in vitro* systems and at the bulk level (100 cells/population). While this has proven extremely valuable in identifying ELP and PreProB progenitors, my data shows some variation in the differentiation potential of upstream progenitors. This is a strong indicator for heterogeneity in these populations and to objectively define

the lymphoid hierarchy upstream of ELP, systematic analysis of function and gene expression would need to be carried out at the single cell level.

As mentioned above, it is possible that two different B cell differentiation pathways exist, with a predominantly CD10 negative progenitor pathway existing in fetal life whereby CD10- LMPP give rise to ELP, PreProB, ProB and B cells; and a CD10+ progenitor pathway existing in adult life whereby CD10+ LMPP give rise to CLP, ProB and B cells. This needs to be functionally and molecularly demonstrated and forms the basis of future work in the lab. The findings are likely to be relevant in our understanding of the cell of origin of specific types of infant, paediatric and adult leukaemias.

Chapter 4

The impact of Trisomy 21 on fetal B-lymphopoiesis

4.1 Background and aim

While the impact of T21 on FL megakaryopoiesis and myelopoiesis is relatively well studied, little is known about the impact of T21 on lymphopoiesis. Similarly, the link between abnormal T21 FL haematopoiesis and the multi-step model of ML-DS is well established [Bhatnagar et al., 2016] but less is known about how the perturbations to haematopoiesis mediated by T21 in fetal life underpins the increased incidence of DS-ALL. In this chapter I have characterised haematopoiesis and lymphopoiesis in T21 fetal BM and compared this to the normal fetal BM hierarchy as established in Chapter 3. How these differences potentially link to DS-ALL are explored and discussed in Chapter 5.

Previous work from the lab reported several differences in the frequency of HSPC in T21 FL compared to normal disomic (NM) FL [Tunstall-Pedoe et al., 2008], [Roy et al., 2012]. Concomitant with an apparent lack of B progenitors (described as CD34+CD19+), the frequency of immunophenotypic HSC, Megakaryocyte-Erythroid progenitors (MEP) and megakaryocytes was increased (summarised in figure 4.1). Consistent with these frequencies, functional analysis showed that T21 FL HSC and MEP had higher clonogenicity in methylcellulose colony forming assays and B cell output from T21 HSC was reduced *in vitro* [Roy et al., 2012] and *in vivo* in a xenograft model (unpublished data).

Since the data from FL suggested a defect in B-lymphopoiesis, and the main site of fetal B-lymphopoiesis is the fetal BM, I have characterised the T21 fetal BM by flow

cytometry, functional *in vitro* assays and transcriptome analysis for the first time.

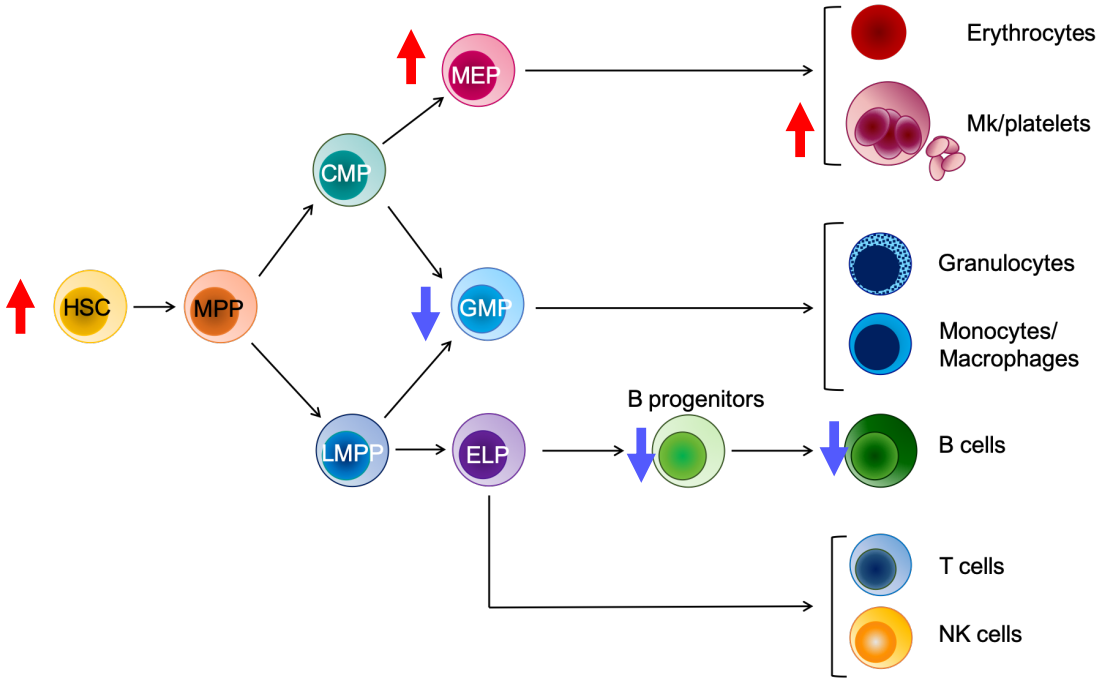


Figure 4.1: Summary of how T21 FL HSPC frequencies differ from NM HSPC frequencies: there are increased immunophenotypic HSC, MEP and megakaryocytes and decreased GMP, and committed B progenitors. HSC functional output was biased towards megakaryopoiesis and showed reduced B-lymphoid output [Roy et al., 2012]. GMP: granulocyte/monocyte progenitor; MEP: megakaryocyte/erythroid progenitor; CMP: common myeloid progenitor.

4.2 Comparison of immunophenotypic HSPC frequency in T21 fetal BM and normal fetal BM

To establish whether the apparent defect in B-lymphopoiesis was specific to the FL or was also present in the fetal BM, I performed flow cytometry analysis on second trimester T21 fetal BM and compared HSPC frequencies with NM second trimester fetal BM (data described in Chapter 3).

Consistent with what was observed in the FL, there is a significantly higher frequency of immunophenotypic HSC in the T21 fetal BM (figures 4.2, 4.3). HSC frequency is 4.7-fold higher in T21 vs normal fetal BM compared to the 3.5-fold higher frequency previously reported in T21 FL [Roy et al., 2012]. The B-progenitor defect in fetal BM is more profound than in FL. ELP and PreProB progenitors were extremely rare in fetal BM (figures 4.2 and 4.4) whereas in T21 FL, although PreProB progenitors were markedly reduced, ELP frequency was normal. In addition to this, B cell frequencies were also significantly decreased (figures 4.2, 4.4). These deficiencies are in contrast to comparable frequencies of ProB progenitors.

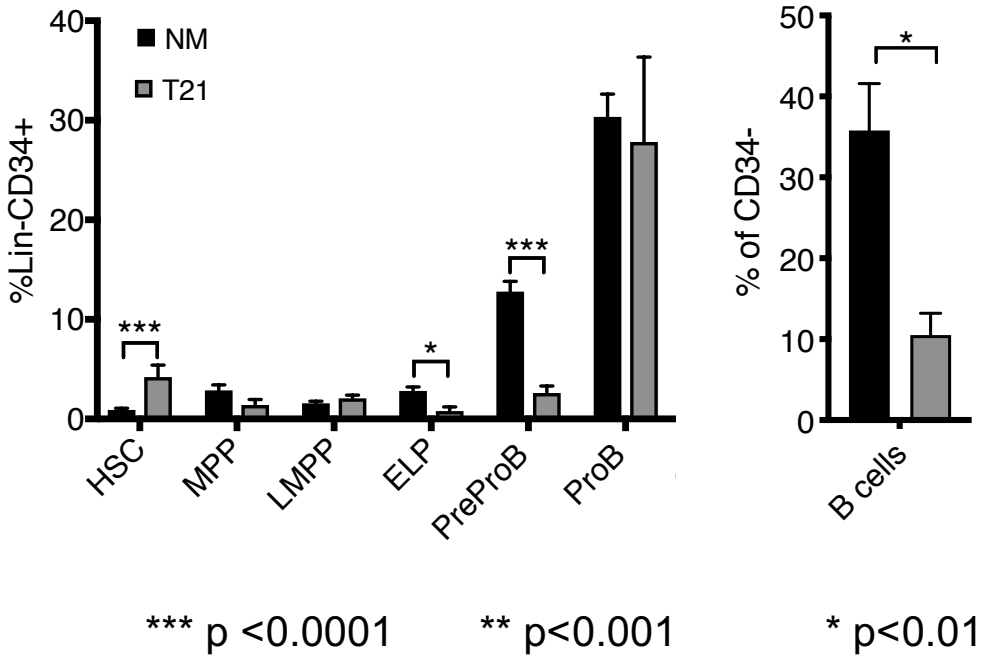


Figure 4.2: Frequencies of HSPC and B cells in T21 fetal BM (n=9-12, grey) compared to normal (NM) fetal BM(n=12-35, black). NM HSPC frequencies also shown in figure 3.4.

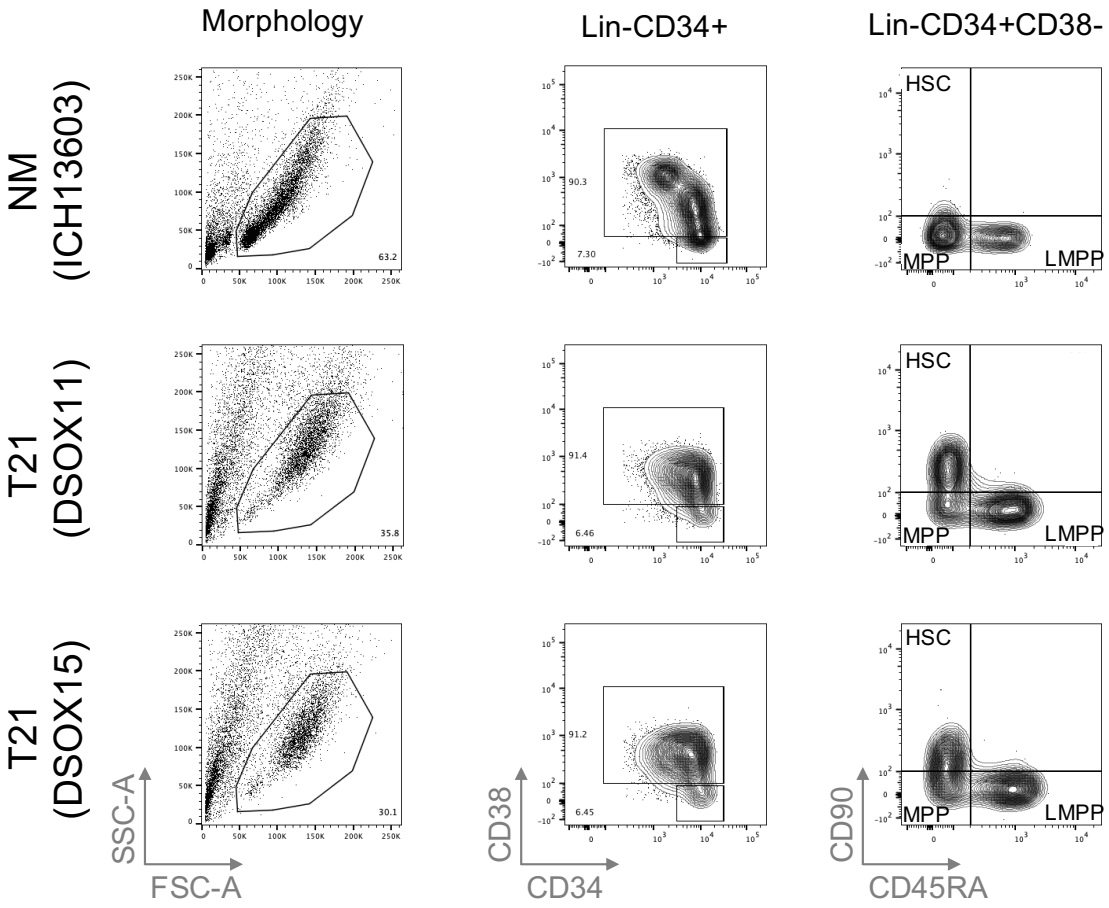


Figure 4.3: Representative flow plots showing difference in immunophenotypic HSC, MPP and LMPP frequency in two T21 fetal BM samples compared to one NM fetal BM sample. All three samples shown were CD34 enriched. Titles above each column of plots indicate the gate shown and each row is one biological sample.

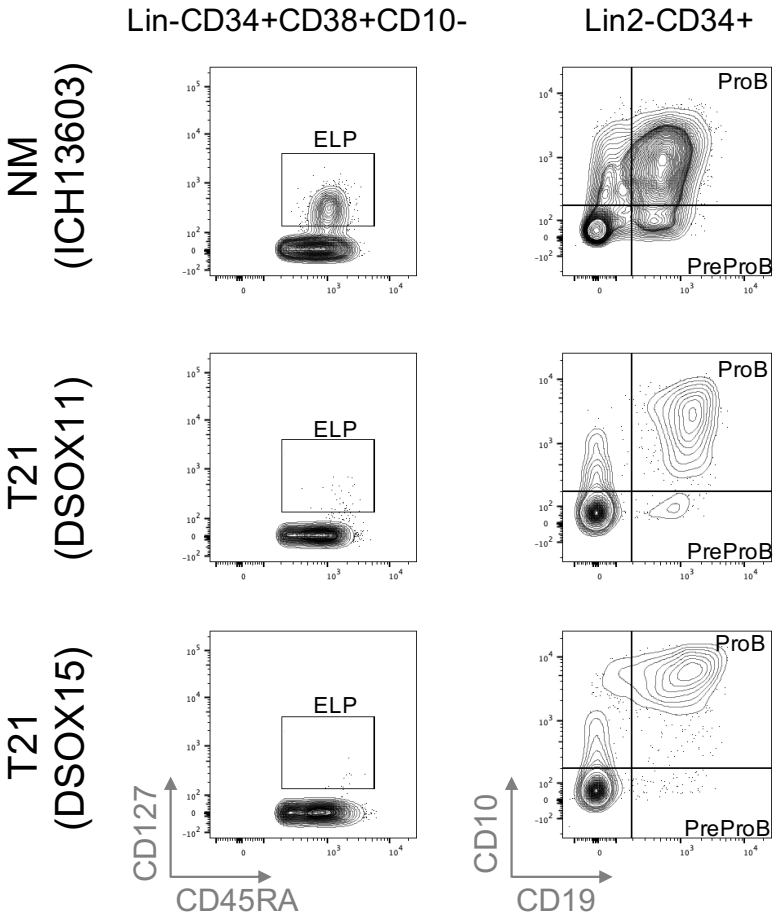


Figure 4.4: Representative flow plots showing difference in immunophenotypic ELP, PreProB, ProB- progenitor frequency in two T21 fetal BM samples compared to one NM fetal BM sample. All three samples shown were CD34 enriched.Titles above each column of plots indicate the gate shown and each row is one biological sample.

4.3 *in vitro* differentiation potential of T21 HSPC

To define the functional potential of each T21 population in the fetal B-lymphoid hierarchy, I performed MS-5 stromal co-culture and clonogenic assays.

4.3.1 MS-5 co-cultures

To establish the B/NK potential of T21 HSPC, I performed MS-5 co-culture differentiation assays. Given the lack of ELP and PreProB progenitors in T21 fetal life, I hypothesised that T21 HSC/MPP/LMPP would have reduced B-lymphoid capacity. 100 immunophenotypically defined HSPC were sorted from 1 NM biological sample and 2 T21 biological samples for co-culture on MS-5 stroma over 4 weeks in parallel. Due to their rarity, 94 and 43 ELP and 98 and 80 PreProB progenitors were sorted and plated respectively from the two T21 samples. In addition to the positive NM control plated in parallel with the T21 samples, data from previous MS-5 co-cultures of NM HSPC (Chapter 3) were collated for comparison.

By day 7, T21 PreProB and ProB progenitors had differentiated into B cells which is comparable to the dynamics observed with NM PreProB and ProB progenitors (figure 4.5). T21 ELP largely differentiated into B cells too which is consistent with the pattern observed in NM ELP/MS-5 co-cultures (figure 4.5).

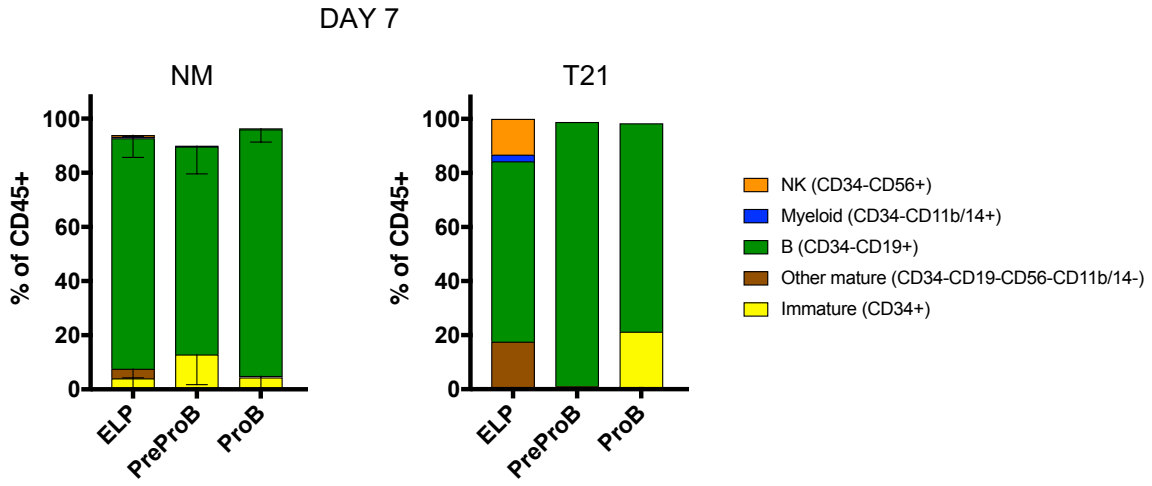


Figure 4.5: B, NK and myeloid output from MS-5 co-cultures with sorted ELP (43-94 cells for T21, 100 cells for NM), PreProB (80-98 cells for T21, 100 cells for NM) and ProB progenitors (100 cells) at day 7. NM (also shown in figure 3.12): n=4-6, T21: n=2.

In contrast to the dynamics observed downstream of T21 ELP, the frequency and dynamics of B cell differentiation in T21 HSC, MPP, LMPP MS-5 co-cultures differed from what I observed in NM fetal BM HSPC co-cultures. Overall, the B cell potential of T21 HSC, MPP and LMPP was reduced compared to NM at all time points (figure 4.6). In NM co-cultures, LMPP B cell potential peaks at day 14 and while T21 LMPP B cell numbers did "peak" at D14, this peak was reduced in magnitude and the co-culture was dominated by myeloid cells (figures 4.6 and 4.7(a)). Similarly, T21 HSC and MPP produced no or very few B cells *in vitro* by day 21/28, when peak B cell output is observed in NM MPP/HSC co-cultures (figures 4.6 and 4.7(b)). Despite the lack of B-lymphopoiesis, NK cell potential remained unaffected. In fact, mature NK cells (CD34-CD56+) were detected as early as day 14 in T21 HSPC co-cultures (figure 4.6).

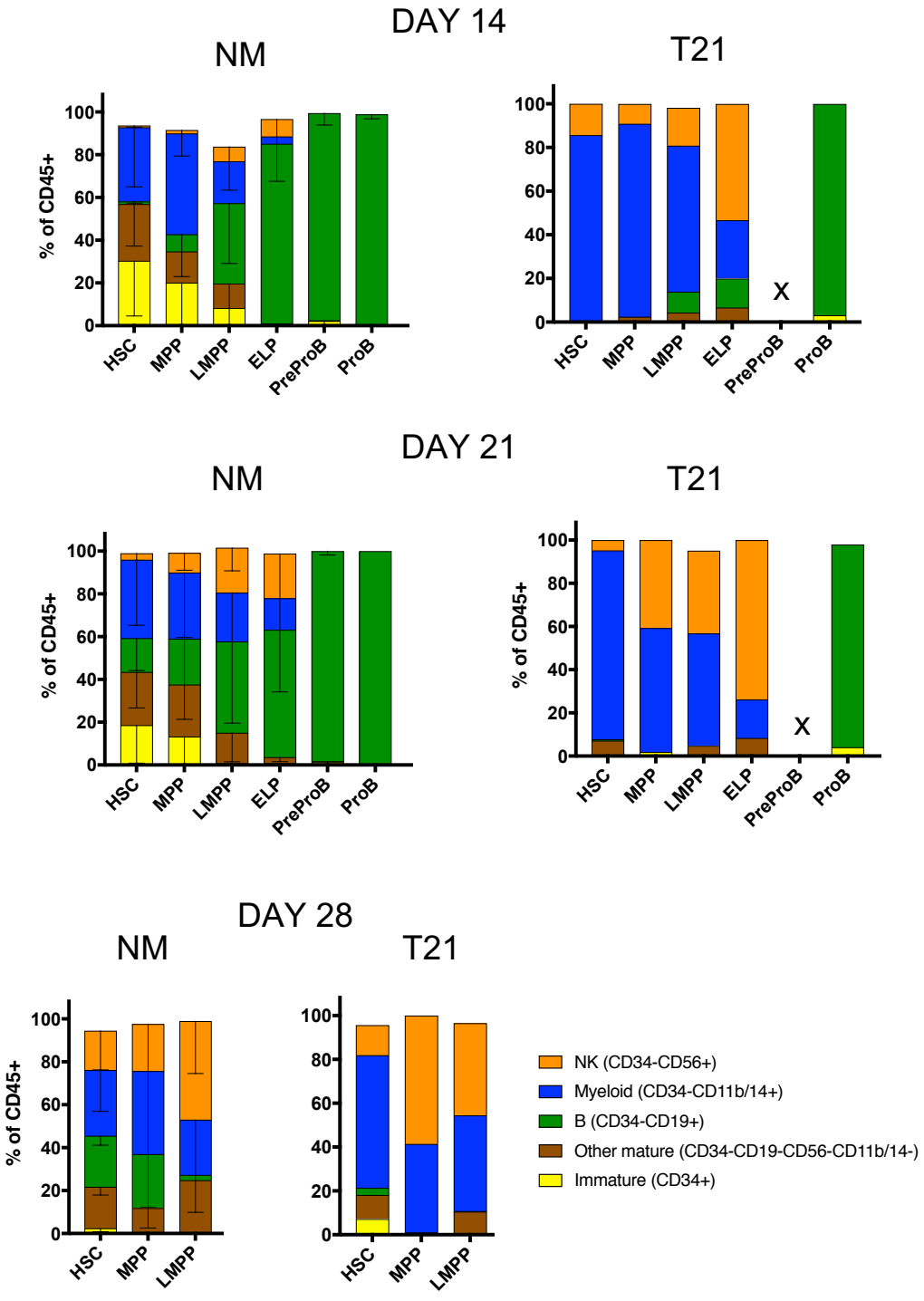


Figure 4.6: B, NK and myeloid output from MS-5 co-cultures with sorted HSPC (100 cells /population) apart from T21 ELP (43-94 sorted and plated) and T21 PreProB (80-98 sorted and plated) at days 14, 21 and 28. X indicates too few cells available for analysis. NM (also shown in figure 3.10): n=4-6, T21: n=2.

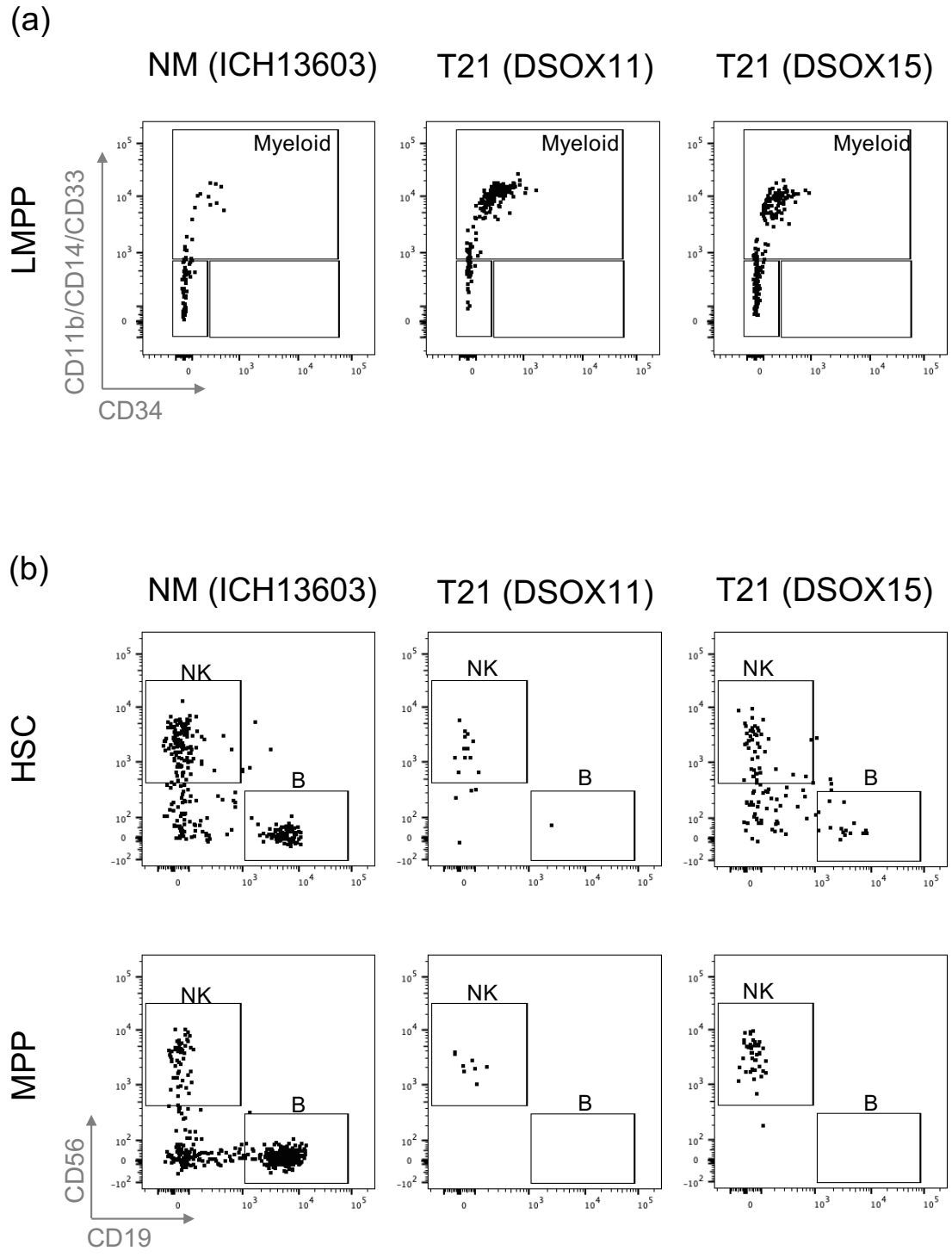


Figure 4.7: Representative flow plots gated on (a) hCD45+ cells from LMPP co-culture on MS-5 at day 14 and (b) hCD45+CD11b-CD14-CD33-CD34- from HSC and MPP co-cultures on MS-5 at day 28. One NM and two T21 samples shown.At day 0 100 sorted HSPC were plated for co-culture.

4.3.2 Clonogenic assays

The MS-5 co-cultures suggested that T21 HSPC were able to differentiate into myeloid cells. However, MS-5 co-cultures are optimised for establishing B/NK potential of a given population and are therefore not a reliable measure of myeloid potential. Therefore, to establish the erythroid and myeloid differentiation potential of T21 HSC, MPP and LMPP, I performed methylcellulose differentiation assays. I sorted 100 HSC/MPP/LMPP from 2 T21 samples and 1 NM positive control and classified the colonies that grew between days 14 and 21. While there was no marked difference in the total number of colonies formed (figure 4.8), I did notice that the GEMM colonies produced in T21 HSC and MPP cultures were larger than their normal counterparts (figure 4.9(a)). Colony identification was confirmed with selected colonies picked and stained after being fixed onto slides using a cytospin and this suggested that T21 HSPC derived colonies made more megakaryocytes (figure 4.9(b)).

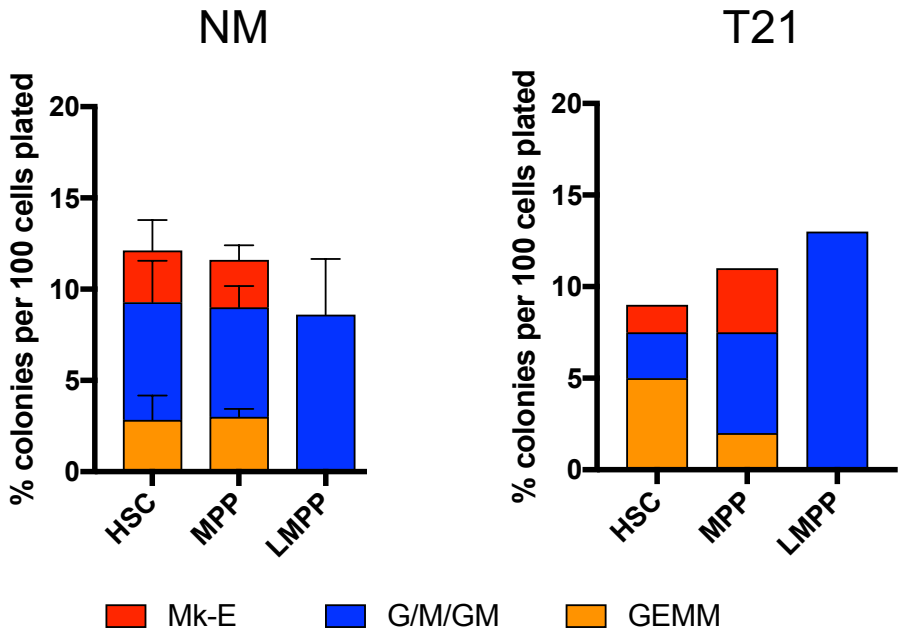


Figure 4.8: Frequency of myeloid and Mk-E colonies formed from 100 sorted progenitor populations after 14-21 days culture in methylcellulose. NM (also shown in figure 3.16): n=4, T21: n=2.

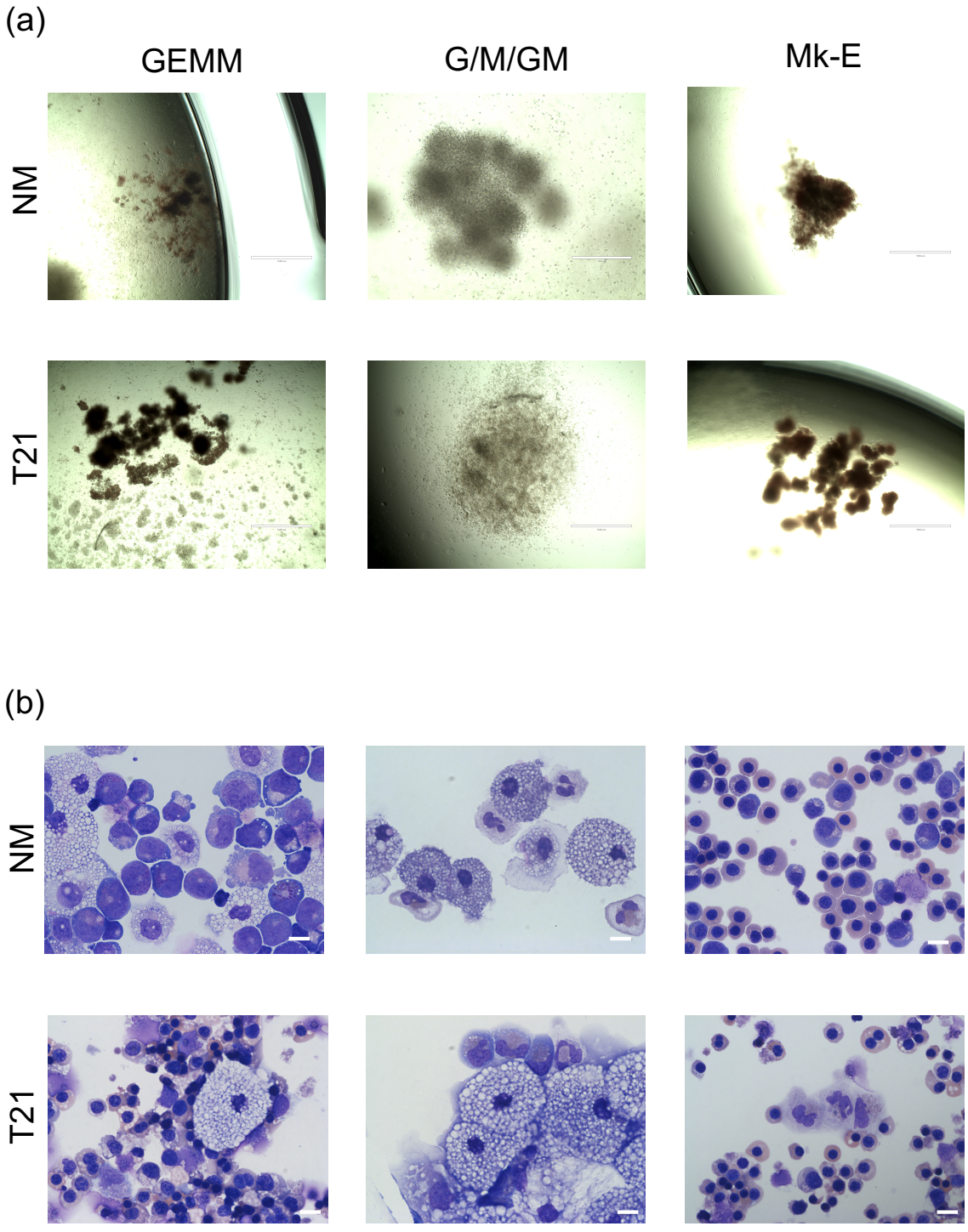


Figure 4.9: (a) Representative images of typical colonies formed after 14-21 days culture in methylcellulose where GEMM colonies from T21 progenitors were noted as being larger. b) Cytospins of picked colonies stained with Giemsa-Wright stain. T21 Mk-E colonies appeared to have more megakaryocytes than NM fetal BM. White scale bar representative of 1000 μm in (a) and 10 μm in (b) respectively. NM (also shown in figure 3.16): n=4, T21: n=2.

4.4 RNA-Sequencing of T21 fetal stem and B-lymphoid progenitor populations

The profound lack of ELP and PreProB progenitors in T21 fetal BM and the reduced ability of T21 HSPC to differentiate into B cells *in vitro* suggested a failure in T21 HSPC commitment to the B-lineage (figure 4.10). To interrogate the molecular mechanisms responsible for this, RNA-Seq on 100 sorted cells of each population (n=2-5) was performed in parallel with the RNA-Seq of NM progenitors (n=3) described in Chapter 3. While insufficient numbers of ELP were available to sort, 100 HSC, MPP and LMPP from three biological samples, 100 PreProB progenitors from two biological samples and 100 ProB and B cells from 5 biological samples were sorted for library preparation. I analysed these data myself.

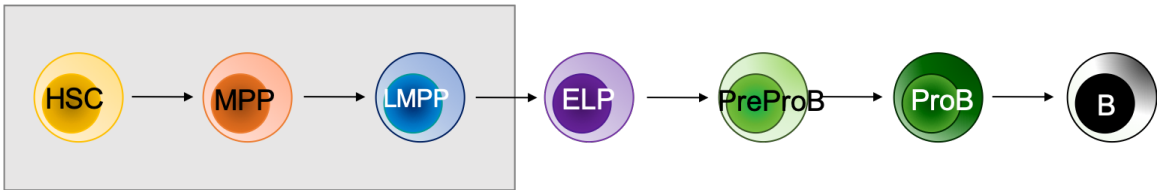


Figure 4.10: Scheme of human fetal B-lymphopoiesis hierarchy and grey box indicating proposed populations affected by T21 driven failure to commit to the B-lineage.

4.4.1 Gene expression through B-lymphopoiesis

Expression of genes known for their involvement in B-lymphopoiesis were largely comparable to their NM counterparts in each sorted T21 population (figure 4.11). Significant differences in expression were all confined to the later stages of B-lymphopoiesis and significant differences between NM and T21 PreProB progenitors should be interpreted with caution as the T21 PreProB gene expression is representative of two biological replicates due to the rarity of this cell population in T21 fetal BM.

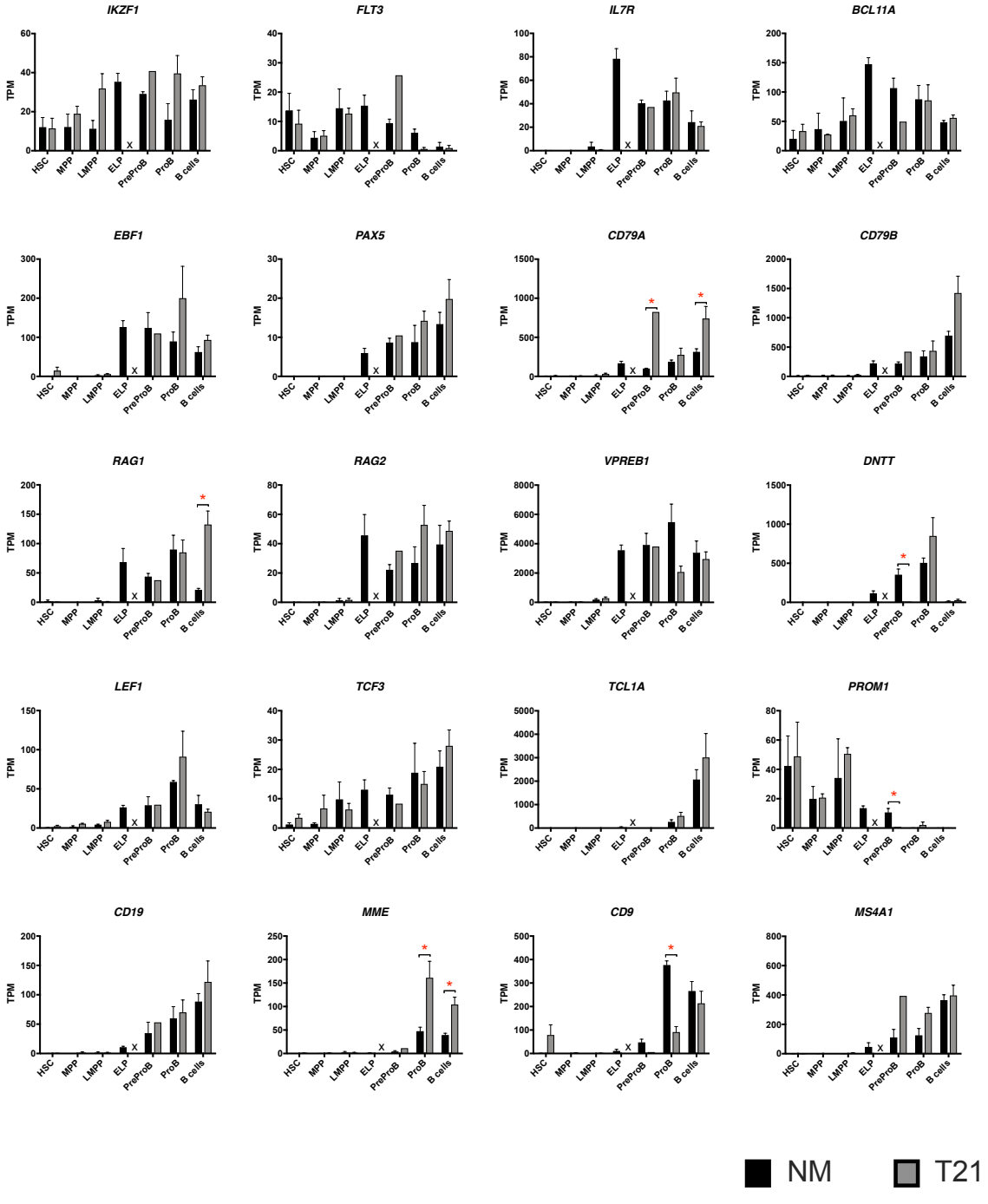


Figure 4.11: Expression of genes selected for their known role in B-lymphopoiesis in NM (n=3, black) and T21 (n=2-5, grey) sorted populations (100 cells) by RNA-Seq. $*$ denotes significance determined by DESeq2 with an FDR < 0.1.

To investigate whether the failure to commit to B-lymphopoiesis is a result of cell intrinsic bias towards other lineages, I also looked at the gene expression of key genes associated with myelopoiesis (figure 4.12), erythropoiesis (figure 4.13) and megakaryopoiesis (figure 4.14). Genes were chosen according to lists curated in our lab [Roy et al., 2012] [Psaila et al., 2016] [unpublished data; Dr. Supat Thongjuea].

There were no striking differences in the pattern of myeloid gene expression in T21 fetal BM HSPC except for the *CST7* gene (that encodes Cystatin F), which was uniquely expressed in T21 but not NM MPP and LMPP; and the *ELANE* and *CFD* genes (that encodes Complement Factor D) which were expressed in NM LMPP but at undetectable or very low levels in T21 LMPP (figure 4.12). The significance of small increases in *IL3RA* expression in PreProB progenitors is difficult to assess especially given the small number of samples (for T21 n=2; figure 4.12).

There were no striking differences in the expression of erythroid genes in T21 HSPC apart from in the *APOE* gene, expression of which was seen in T21 but not NM HSC and LMPP (figure 4.13). This is potentially interesting given the higher levels of *APOE* in T21 brain tissue [Lott and Head, 2019] [Raha-Chowdhury et al., 2019] but the significance is unclear given the very low expression overall with TPM in the range of 0-1.5 (figure 4.13). By contrast, the expression of several megakaryocyte genes was higher in T21 fetal BM HSC and/or LMPP. In particular, expression of *PF4* and *MPIG6G* genes was seen in T21 but not NM LMPP and expression of *ITGB3* was seen in T21 but not NM fetal BM HSC (figure 4.14). These findings support data previously found in T21 FL HSPC [Roy et al., 2012] and it is tempting to speculate that they might reflect an excess of megakaryocyte-primed HSC.

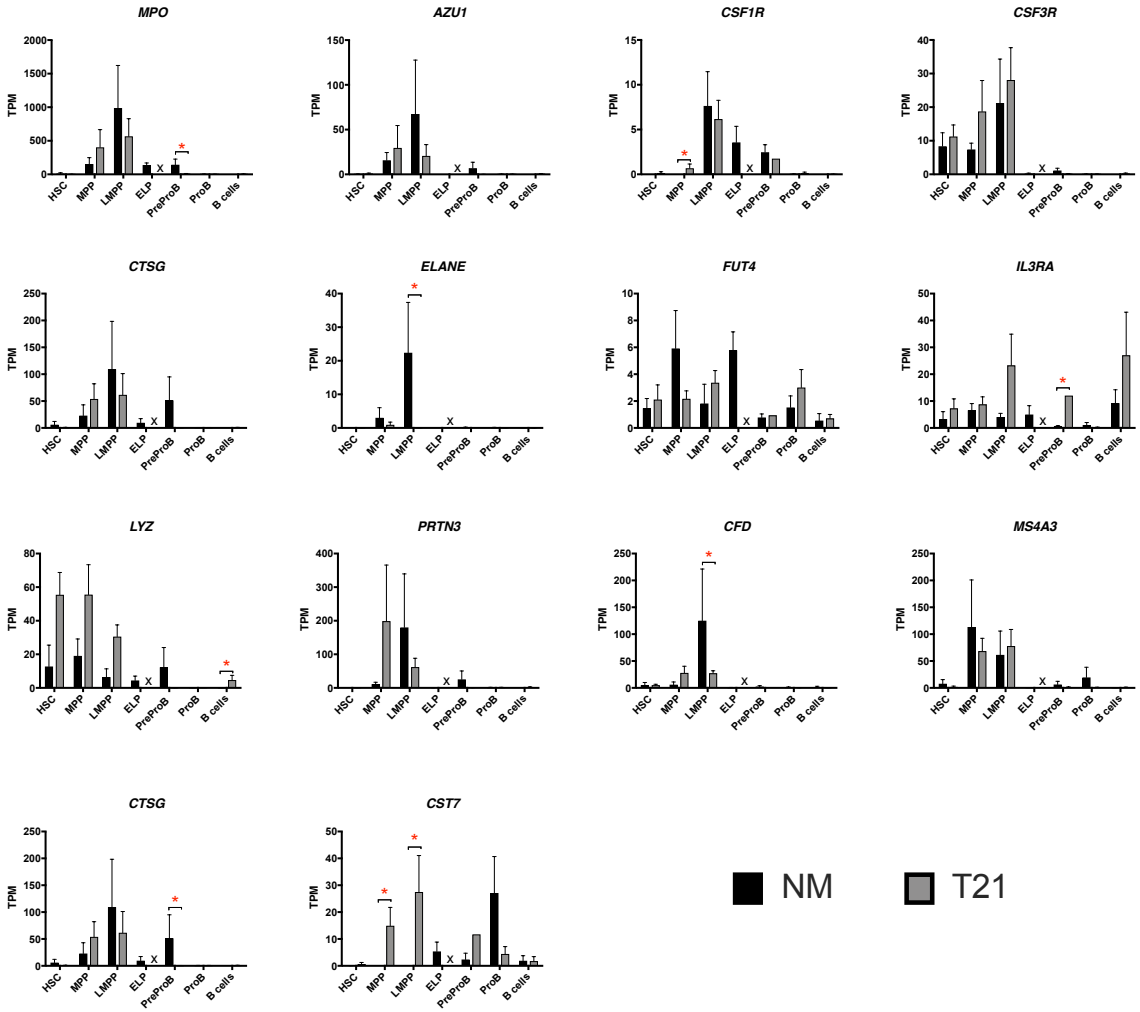


Figure 4.12: Expression of genes known for their role in myelopoiesis in NM (black) and T21 (grey) sorted populations (100 cells/population) by RNA-Sequencing. $*$ denotes significance determined by DESeq2 with an FDR < 0.1.

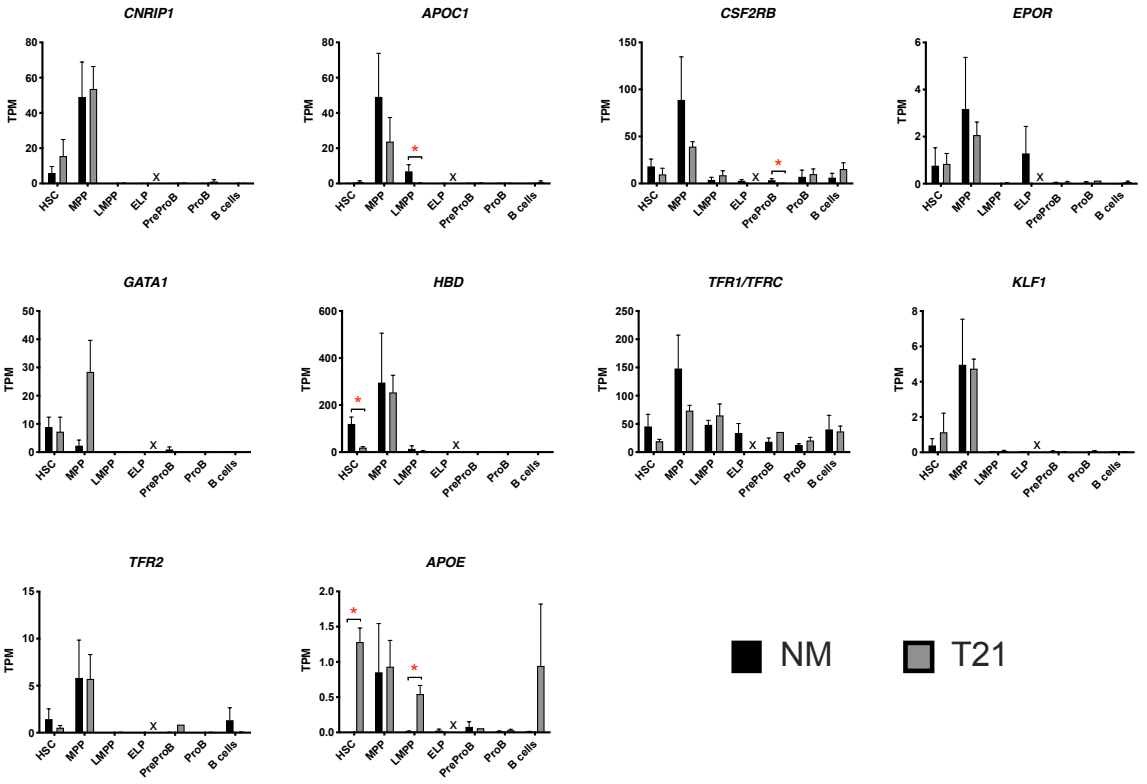


Figure 4.13: Expression of genes known for their role in erythropoiesis in NM (n=3, black) and T21 (n=2-5, grey) sorted populations (100 cells/population by RNA-Sequencing. $*$ denotes significance determined by DESeq2 with an FDR < 0.1.

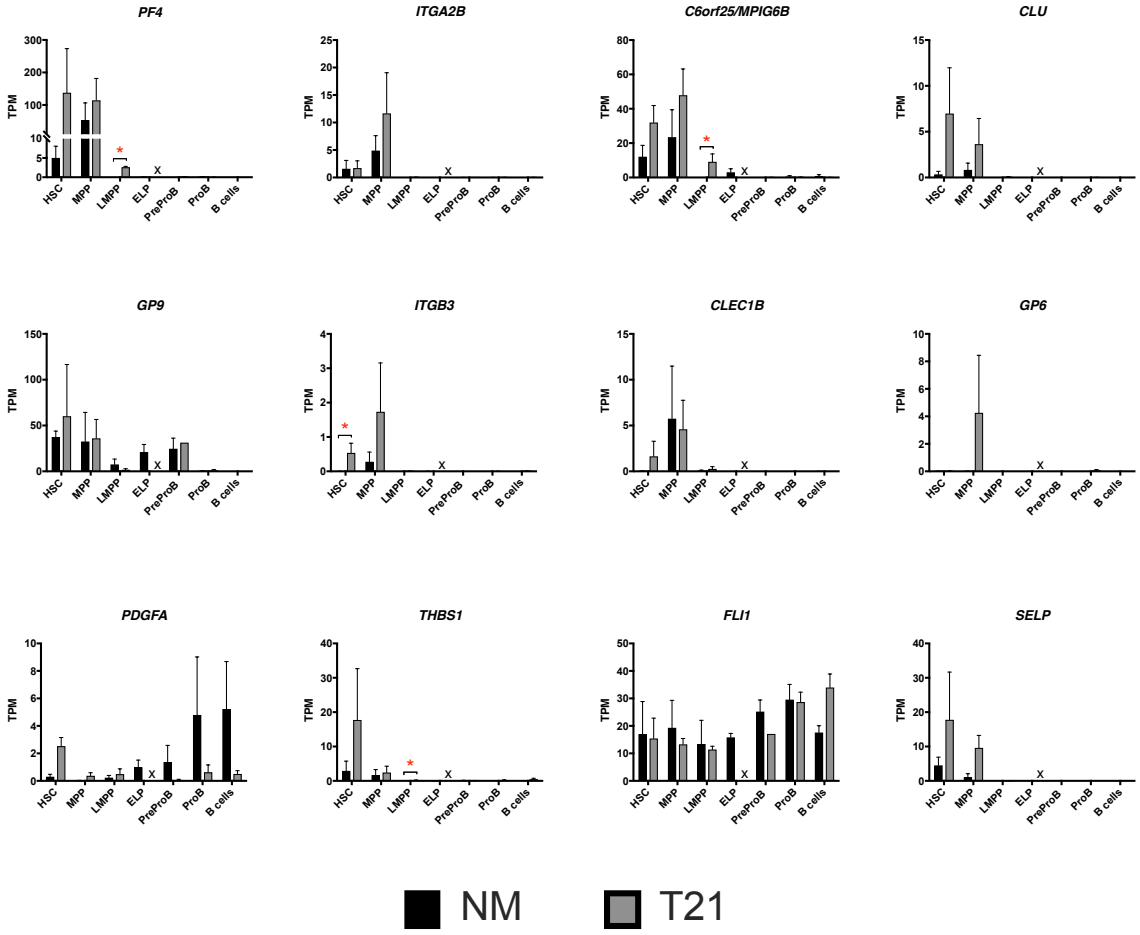


Figure 4.14: Expression of genes known for their role in megakarypoiesis in NM (n=3, black) and T21 (n=2-5, grey) sorted populations (100 cells/population) by RNA-Sequencing. $*$ denotes significance determined by DESeq2 with an FDR < 0.1.

Given that there was no dramatic difference in expression of known B-lymphoid genes in T21 fetal BM HSPC, I next looked at the expression of genes located on chromosome 21 and with a known role in haematopoiesis or leukaemia (table 4.1). These genes either showed comparable levels of expression or an increased level of expression that would be expected as a result of increased gene dosage caused by an extra copy of chromosome 21 (figure 4.15). The differences in the level of expression of the vast majority of chromosome 21 genes were not statistically significant in DE analysis (FDR < 0.1) in most of the HSPC populations. This probably reflects both the small number of biological samples analysed and also the relatively modest differences in expression level of individual genes that would be anticipated simply from the additional copy of that gene in T21 cells. Thus, the only differences that were statistically significant were increased expression of *ETS2* in ProB progenitors and *BACH1* and *USP16* in B cells (figure 4.15). What is more striking is viewing these changes in levels of expression collectively. Thus, although not reaching statistical significance for individual genes in individual populations, the concurrent 1.3-1.8-fold increases in expression level of almost all of these chromosome 21 genes in T21 fetal BM HSPC might play some role in mediating or maintaining the block in B progenitor differentiation.

Table 4.1: Genes on chromosome 21 with known roles in haematopoiesis and/or leukaemia.

Gene	**Function in haematopoiesis/leukaemia**
AIRE	Autoimmune regulator: promotes expression (in thymus) of antigens normally expressed elsewhere to facilitate negative selection of T-cells. Interacts with CREBBP [Anderson and Su, 2011]
BACH1	*BACH1* and *BACH2* promote differentiation of CLP to B-cells by repressing myeloid-related genes [Itoh-Nakadai et al., 2014].
DYRK1A	Involved in ML-DS and B cell cycle exit [Malinge et al., 2012], [Thompson et al., 2015].
ERG	Known oncogene implicated in ML-DS [Rainis et al., 2005] [Roberts and Izraeli, 2014]. Role in B-lymphopoiesis unknown.
ETS2	Related to *ERG*, implicated in ML-DS [Rainis et al., 2005], [Stankiewicz and Crispino, 2009]. Role in B-lymphopoiesis unknown.
GABPA	ETS family transcription factor. Critical regulator of myeloid and B-cell development [Yang et al., 2011].
HMGN1	H3 and H2A histone phosphorylation inhibited by HMGN1 thus maintaining open chromatin structure. Possible role in DS-ALL [Lane et al., 2014], [Mowery et al., 2018].
PAXBP1	Binds PAX3 and PAX7 and links them to methyltransferase machinery [Diao et al., 2012].
RUNX1	Essential for HSC emergence in embryonic development (Reviewed in [Gao et al., 2018]). Important regulator of expression of several B-lymphoid transcription factors including *EBF1* [Boller and Grosschedl, 2014].
SOD1	Important for haematopoietic proliferation in mice [Hadjur et al., 2001]. Role in B-lymphopoiesis unknown.
SON	Processes mRNA of genes required for maintenance of pluripotency and linked to regulation of *GATA* factors. Expressed in HSPCs and blasts [Ahn et al., 2013]. Role in B-lymphopoiesis unknown.
USP16	Required for normal lymphopoiesis and HSC function [Gu et al., 2016a]; regulates *HOX* gene expression.

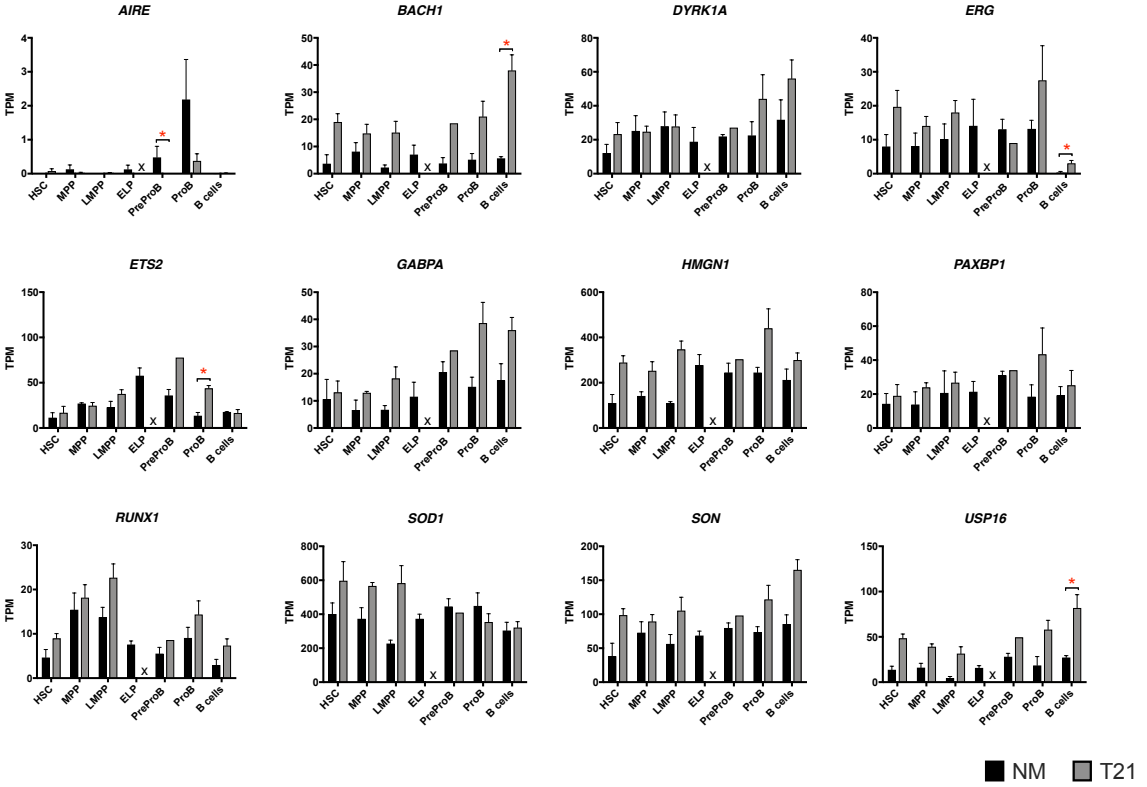


Figure 4.15: Expression of genes with known roles in haematopoiesis and/or leukaemia located on chromosome 21 in NM (n=3, black) and T21 (n=2-5, grey) sorted populations (100 cells/population). $*$ denotes significance determined by DESeq2 with an FDR < 0.1.

4.4.2 Differential analysis of RNA-Seq

I next adopted an unbiased, global approach to the analysis of the transcriptome of each progenitor population. PCA of the top 1000 most variably expressed genes showed that NM and T21 HSPC, B progenitors and B cells first separated by their lineage (PC1) and then by chromosome phenotype (PC2) (figure 4.16). Further PCA of HSPC excluding committed B progenitors (figure 4.17(a)) and with myeloid progenitors added for comparison (figure 4.17(b)) shows separation primarily by chromosome phenotype (PC1) and then according to lineage/hierarchy (PC2 or PC3)(figure 4.17). Together, these PCA show that while drive towards B-lymphoid commitment is the main source of variation between each HSPC and B cells, when only multipotent progenitors are analysed, the main source of variation is in chromosome phenotype. This suggests that the impact of T21 begins early in haematopoiesis.

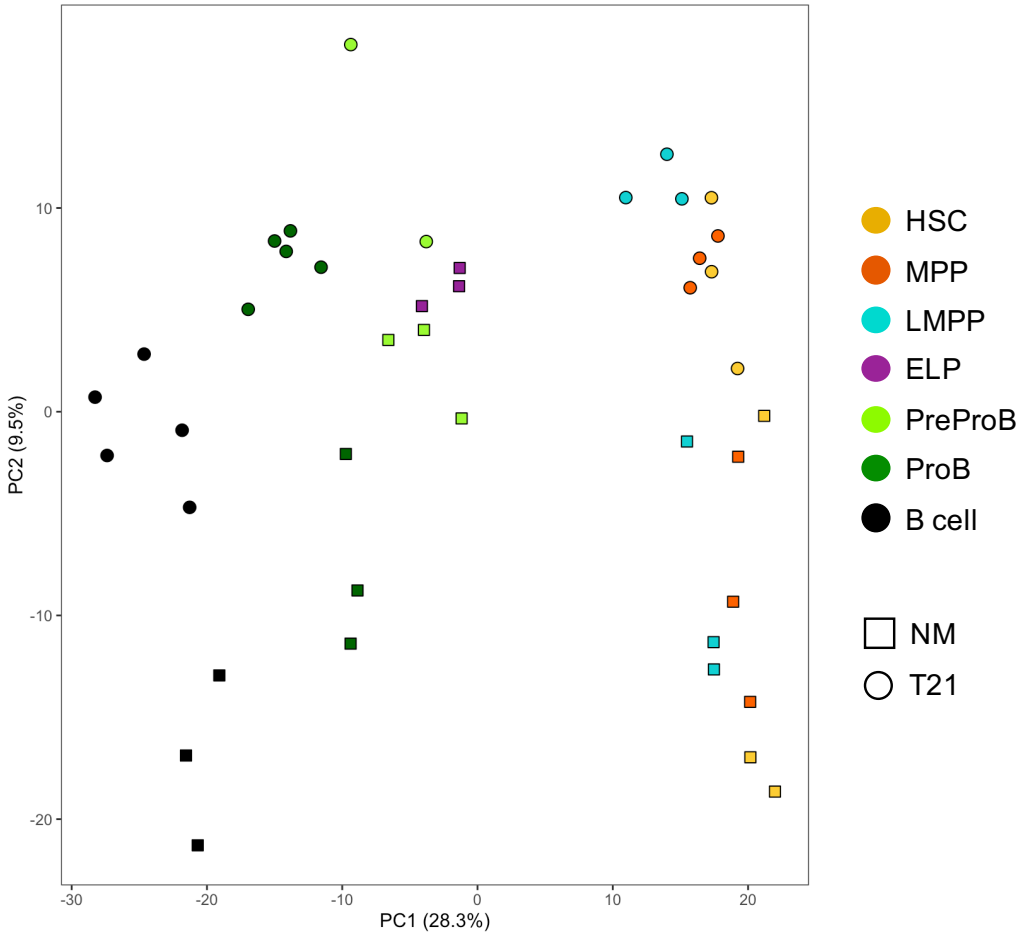


Figure 4.16: PCA of the top 1000 most variably expressed genes by RNA-Seq (100 cells/population) in NM (n=3) and T21 (n=2-5) HSPC and B cells. This shows that B-lineage commitment drives PC1 and chromosome phenotype drives PC2.

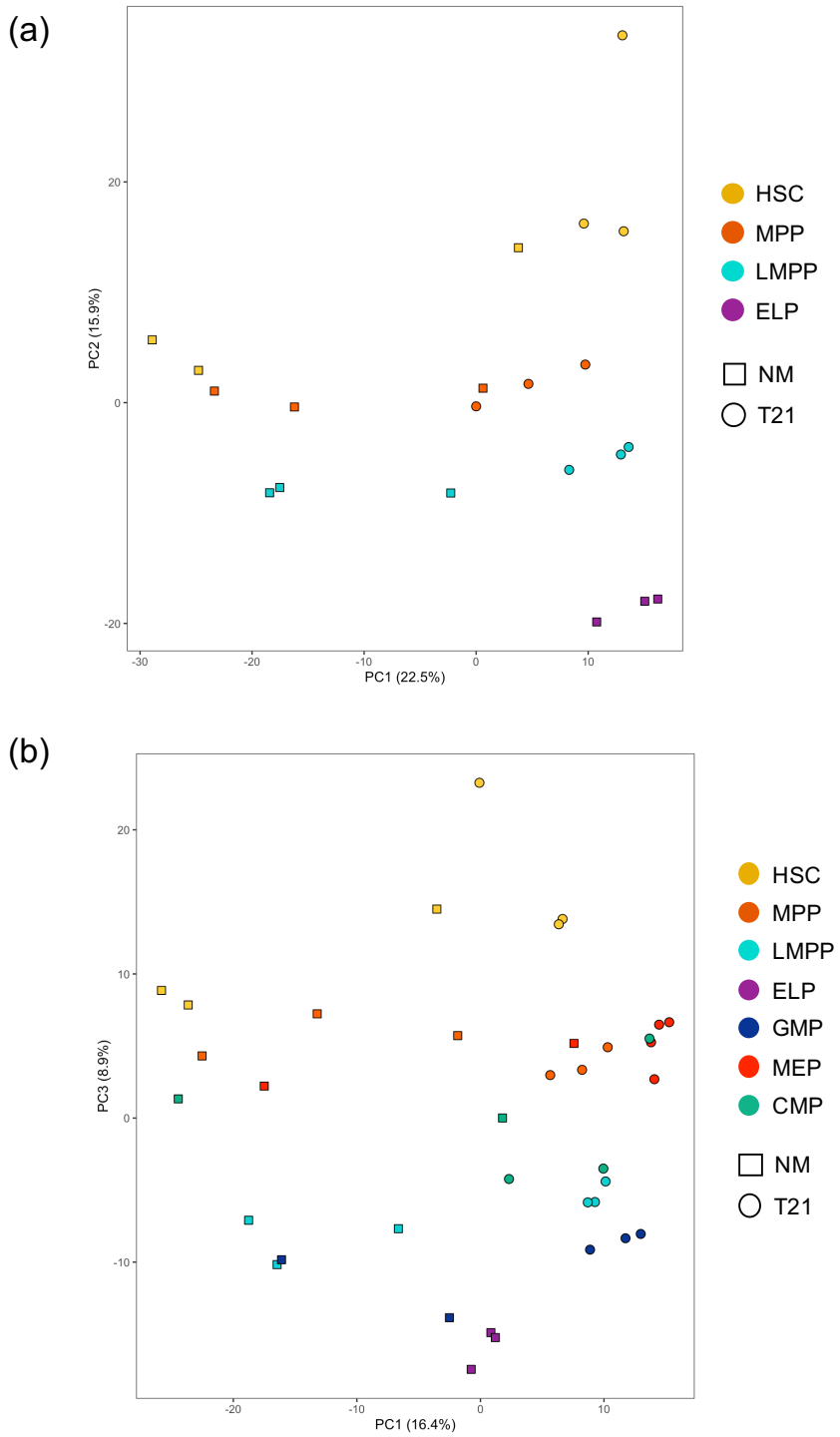


Figure 4.17: PCA of the top 1000 most variably expressed genes by RNA-seq in NM and T21 HSPC. HSPC involved in fetal B-lymphopoiesis is plotted in (a) and myeloid and erythroid progenitors (MEP, CMP and GMP) are included in (b). In both analyses, chromosome phenotype drives PC1 and population hierarchy drives PC2/3. 100 cells/population; NM: n=3; T21: n=2-3.

The PCA showing that the differences in transcriptome between NM and T21 populations was secondary to their position in the fetal B-lymphoid hierarchy (figure 4.16) was confirmed independently using diffusion map analysis (figure 4.18). Diffusion component 1 separated populations according to their position in the hierarchy and diffusion component 2 separated NM from T21 progenitor populations. Since diffusion map analysis is used to establish how related one population is to another, this analysis shows that the differentiation trajectory is not grossly affected by T21: once a T21 HSPC is committed to the B lineage, it will follow that differentiation trajectory. Nevertheless, T21 populations did separate by DC2 and so I next performed differential expression analysis.

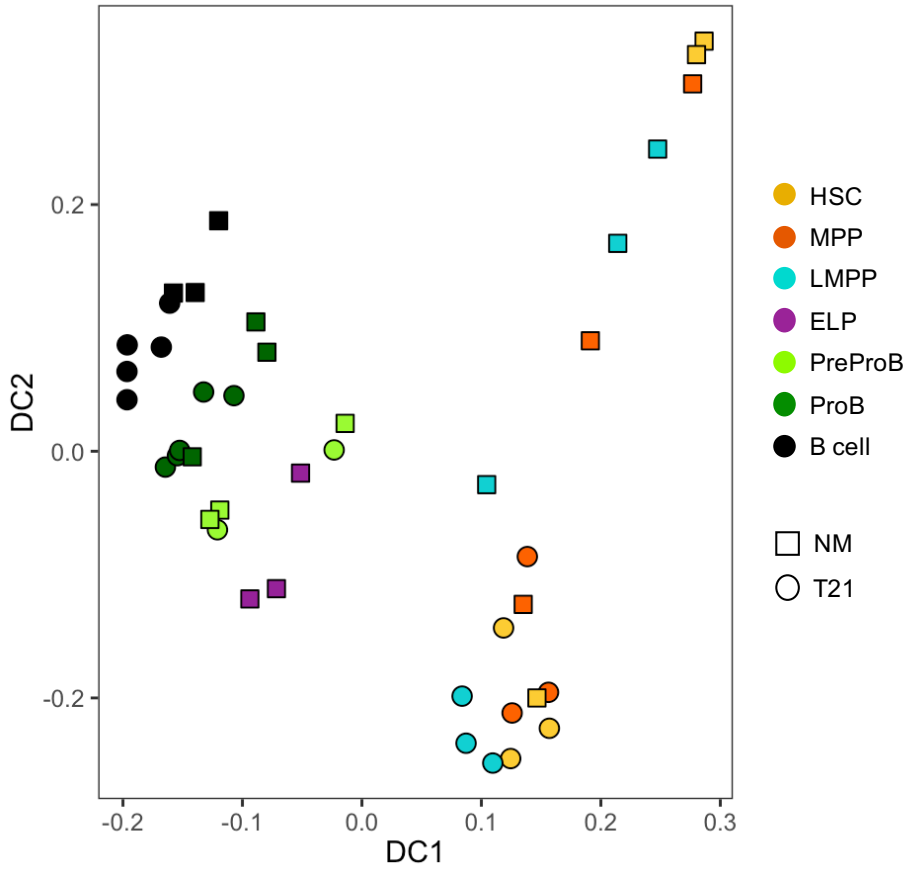


Figure 4.18: Diffusion map analysis of the whole transcriptome of NM (n=3) and T21 (n=2-4) progenitors by RNA-Seq (100 cells/population) demonstrates that the differentiation trajectory is not grossly perturbed by T21.

Pairwise differential expression analysis between each NM and T21 population revealed that there were 1052, 1204, 1232, 1116, 1031 and 1273 differentially expressed (DE) protein coding genes (FDR <0.1, complete DE tables available as online supplemental files) between NM and T21 HSC, MPP, LMPP, PreProB, ProB and B cells respectively. MA plots used to visualise all significant DE genes (protein and non-protein coding) show that in the T21 samples most DE gene are over-expressed (figure 4.19) and this is also reflected in heatmaps of DE protein coding genes for populations upstream of the defect (HSC, MPP and LMPP; figure 4.20).

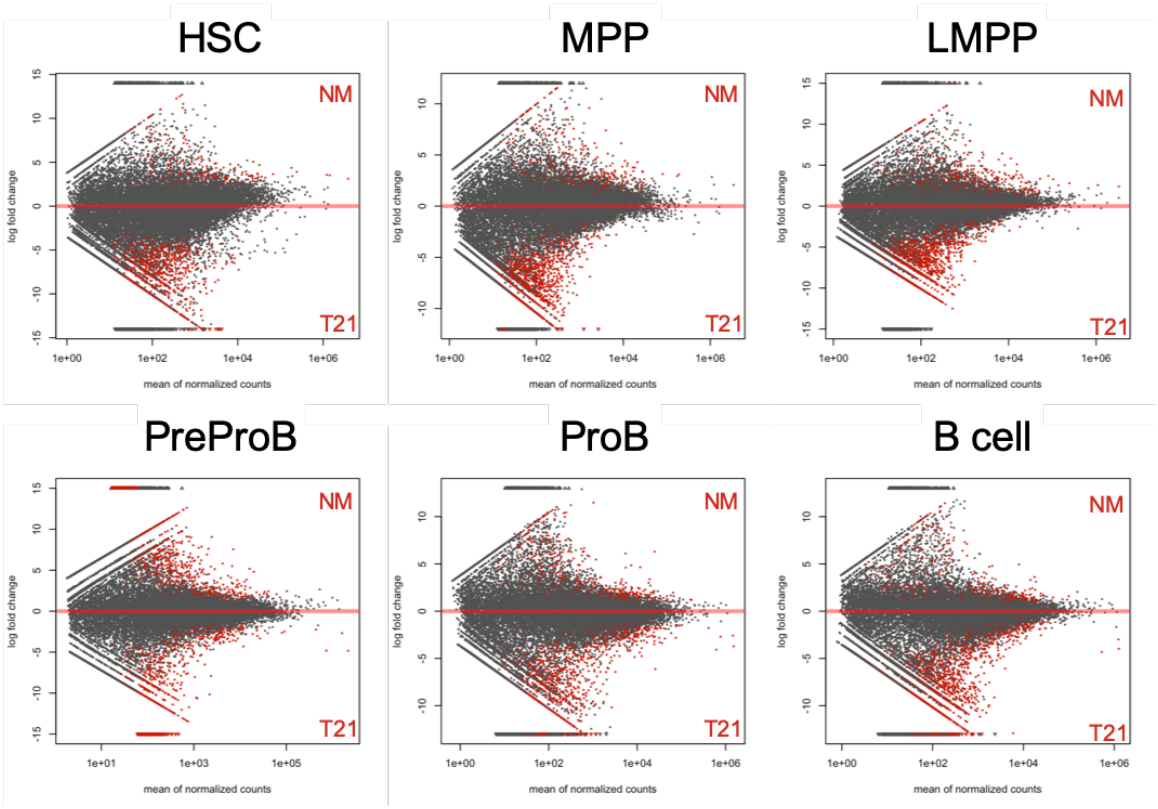


Figure 4.19: MA plots showing significant (FDR <0.1) DE genes (protein and non-protein coding; in red) by RNA-Seq (100 cells/population) determined by DESeq2 in pairwise comparisons of NM (n=3) and T21 (n=2-5) HSPC, B-progenitors and B cells.

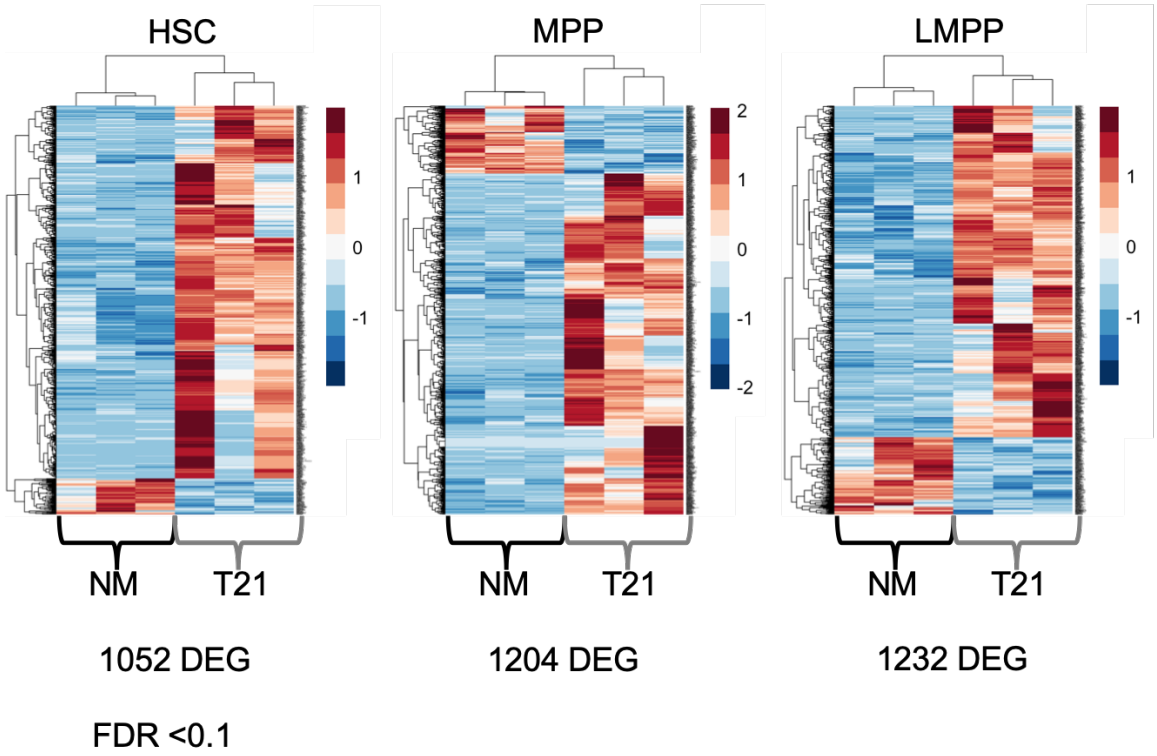


Figure 4.20: Heatmaps showing significant DE protein coding genes in HSC, MPP and LMPP by RNA-Seq. 100 cells/population; NM: n=3; T21 n=2-5; FDR <0.1.

Given that the B-lymphoid defect in T21 fetal BM is likely to be the result of molecular dysregulation in T21 HSPC upstream of B-lineage commitment, I performed GSEA on HSC and LMPP DE gene lists using hallmark and Gene Ontologies (GO) gene sets.

In the GSEA using GO gene sets, the most striking enrichment was down-regulation of gene sets associated with nucleosome organisation and chromatin silencing in HSC (shown as an enrichment in NM HSC, figure 4.21). This, combined with the DE analysis showing significant global over-expression of genes, suggests that there may be a global failure to silence gene expression programmes in T21 fetal HSC.

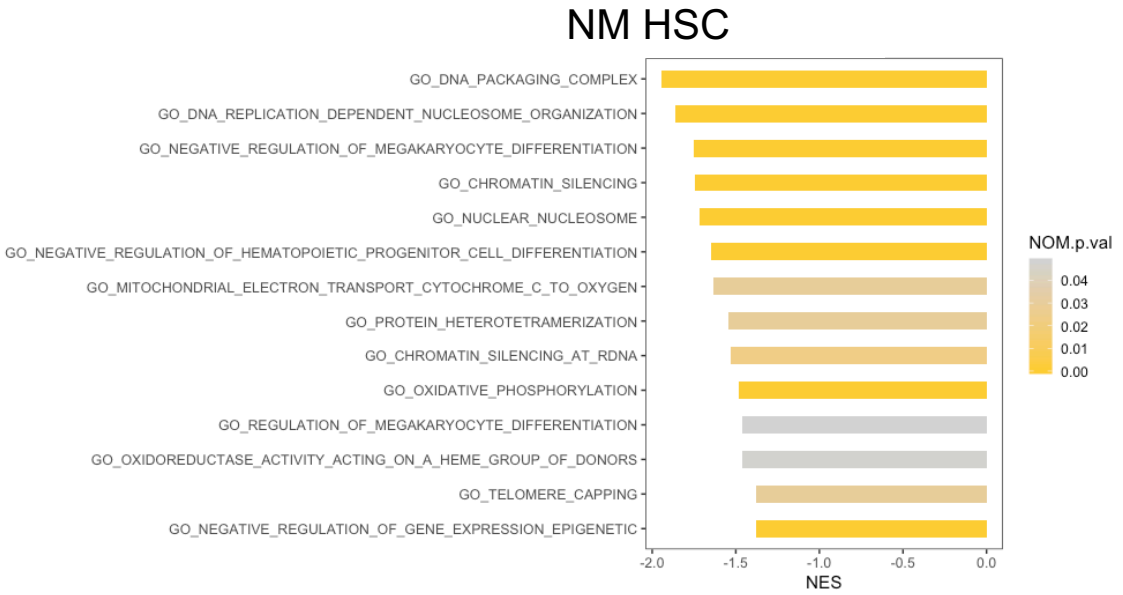


Figure 4.21: GSEA of the transcriptome (RNA-Seq; 100 cells/population) of NM (n=3) and T21 (n=3) HSC shows an enrichment of GO terms associated with nucleosome organisation and chromatin silencing in NM HSC compared to T21 HSC. NES: normalised enrichment score; FDR: false discovery rate; NPV: nominal p-value)

In the GSEA using hallmark gene sets, TGFβ signalling, IL-6 signalling, IL-2 signalling and the inflammatory response gene sets were all enriched in T21 HSC (figure 4.22). Furthermore, IFNα response was enriched in T21 LMPP (figure 4.23). These data, and the fact that the only striking difference in gene expression was increased expression of genes associated with megakaryopoiesis and not B-lymphopoiesis suggested that the defect in B-lineage commitment might be secondary and mediated by T21-driven changes in the T21 fetal BM haematopoietic microenvironment.

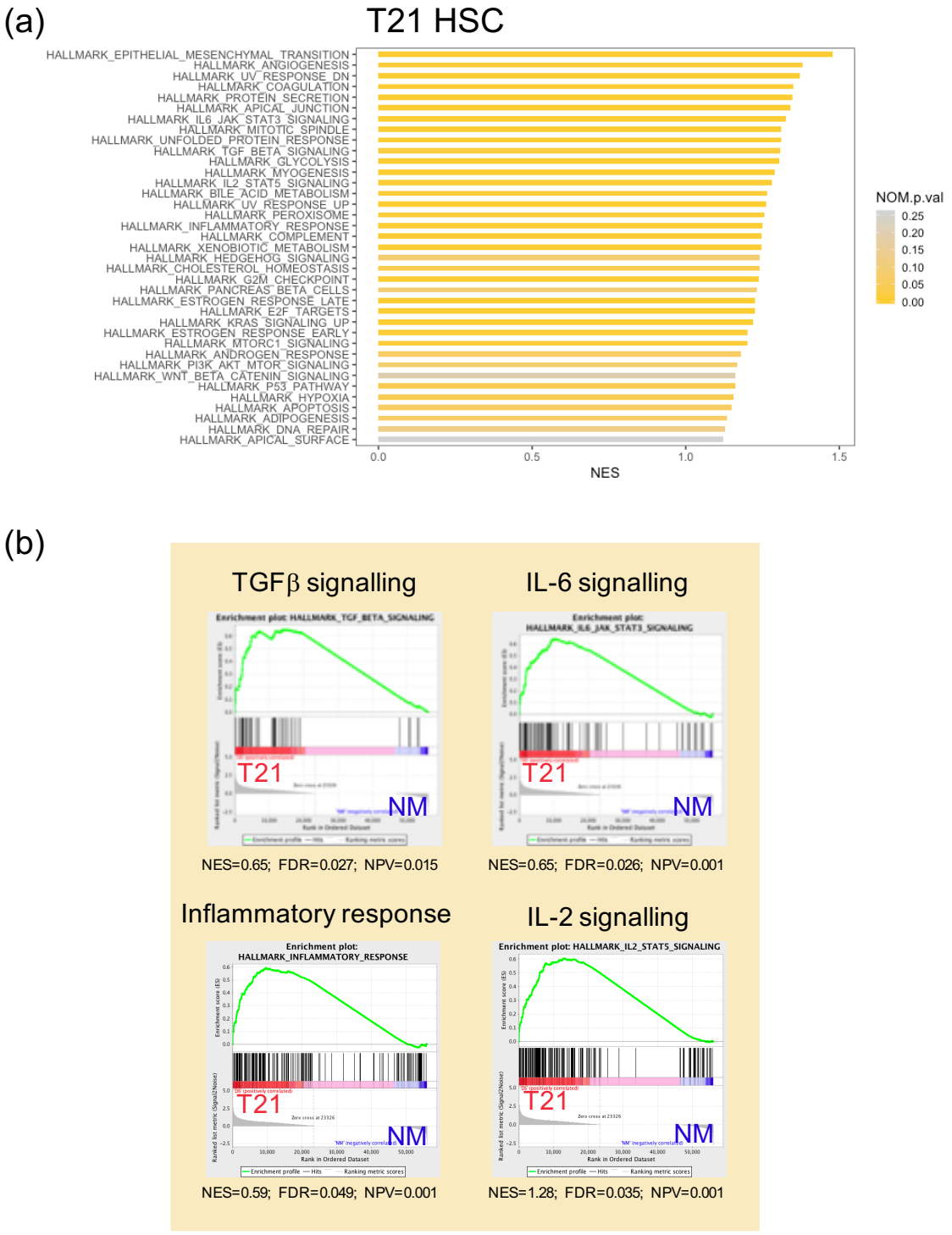


Figure 4.22: Gene set enrichment analysis of the transcriptome (by RNA-Seq; 100 cells/population) of NM (n=3) and T21 (n=3) HSC using hallmark gene sets. (a) summary of significantly enriched gene sets in T21 HSC. FDR <0.25. (b) Enrichment plots showing enrichment of hallmark gene sets associated with the haematopoietic microenvironment. NES: normalised enrichment score; FDR: false discovery rate; NPV: nominal p-value)

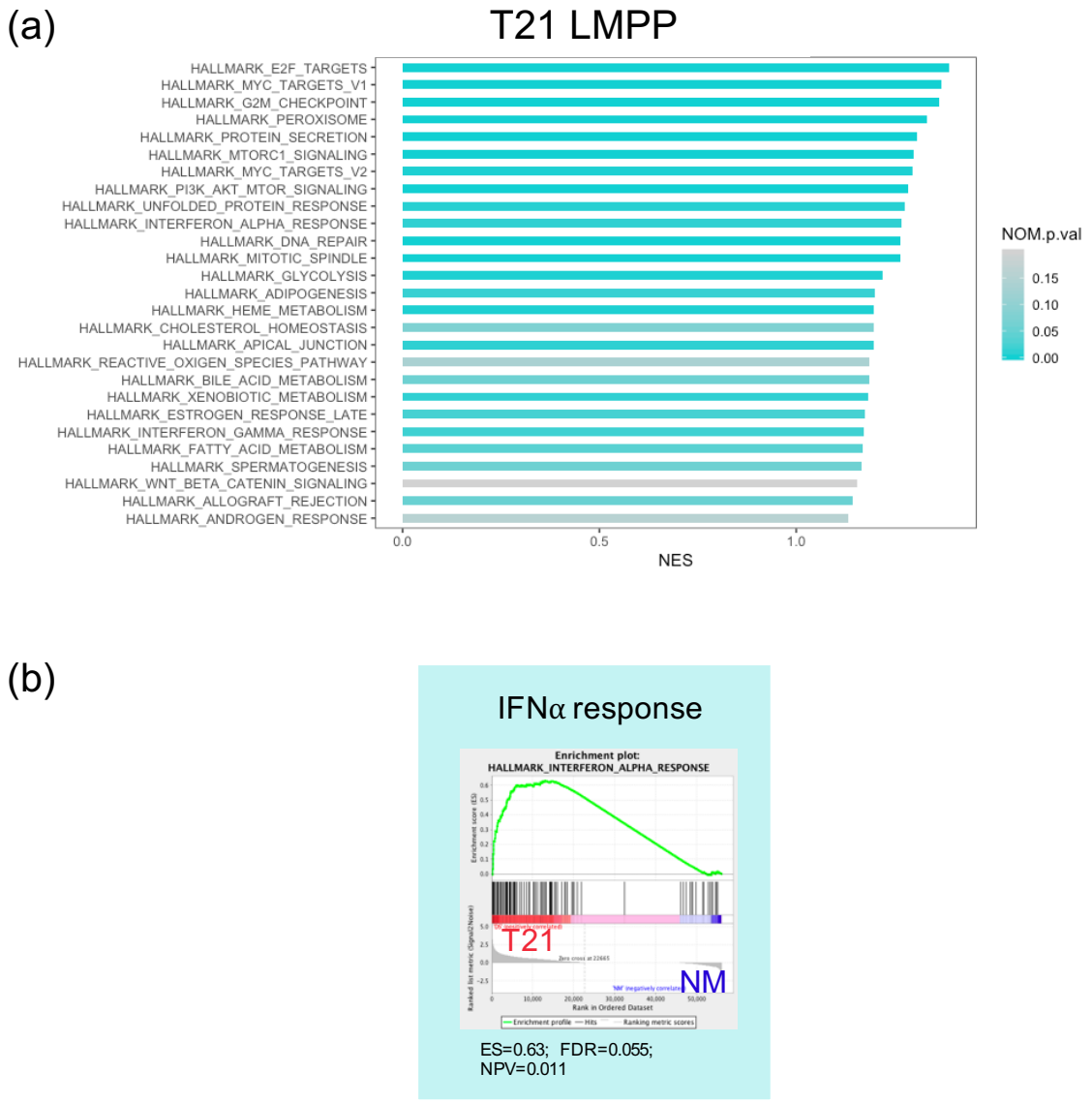


Figure 4.23: Gene set enrichment analysis of the transcriptome (by RNA-Seq; 100 cells/population) of NM (n=3) and T21 (n=3) LMPP using hallmark gene sets. (a) summary of significantly enriched gene sets in T21 LMPP. FDR <0.25. (b) Enrichment plot showing enrichment of hallmark gene set IFNα response in T21 LMPP. NES: normalised enrichment score; FDR: false discovery rate; NPV: nominal p-value)

Finally, to establish whether any chromosome 21 genes are important in driving the specification and commitment of NM human fetal B-progenitor development, I looked for genes encoded on chromosome 21 that were significantly DE in normal neighbouring progenitors from LMPP through to ProB progenitors (figure 4.24, DE analysis as part of Chapter 3). This showed that a total of 4 chromosome 21 genes are up-regulated between different stages of normal fetal B-progenitor development (*JAM2, IFNAR1, USP16* and *CFAP298*) while 7 genes are down-regulated during normal fetal B-progenitor development (*RUNX1, RRP18, PTTG1IP, ITGB2, PRDM15, ETS2* and *APP*; figure 4.24(a)).

Next, I cross-checked these 11 genes with the DE genes between NM and T21 populations. 3/11 of these genes (*PRDM15, APP* and *ETS2*) were dysregulated in T21 fetal BM and (with the exception of *PRDM15* over-expression in T21 B cells) this was confined over-expression in T21 ProB progenitors (figure 4.24(b)). The role of these genes in normal fetal B cell development or in the T21 B lineage defect is not clear as none of these genes have been shown to be important in lymphopoiesis (see table 4.1) and the TPM values are very low in the case of *PRDM15* and *APP*. However, *ETS2* is a known oncogene and is likely involved in the initiation of ML-DS [Rainis et al., 2005] [Stankiewicz and Crispino, 2009]. It is tempting to speculate that the impact that *ETS2* over-expression has on T21 MEP in the fetal liver has a similar proliferative effect on T21 ProB progenitors.

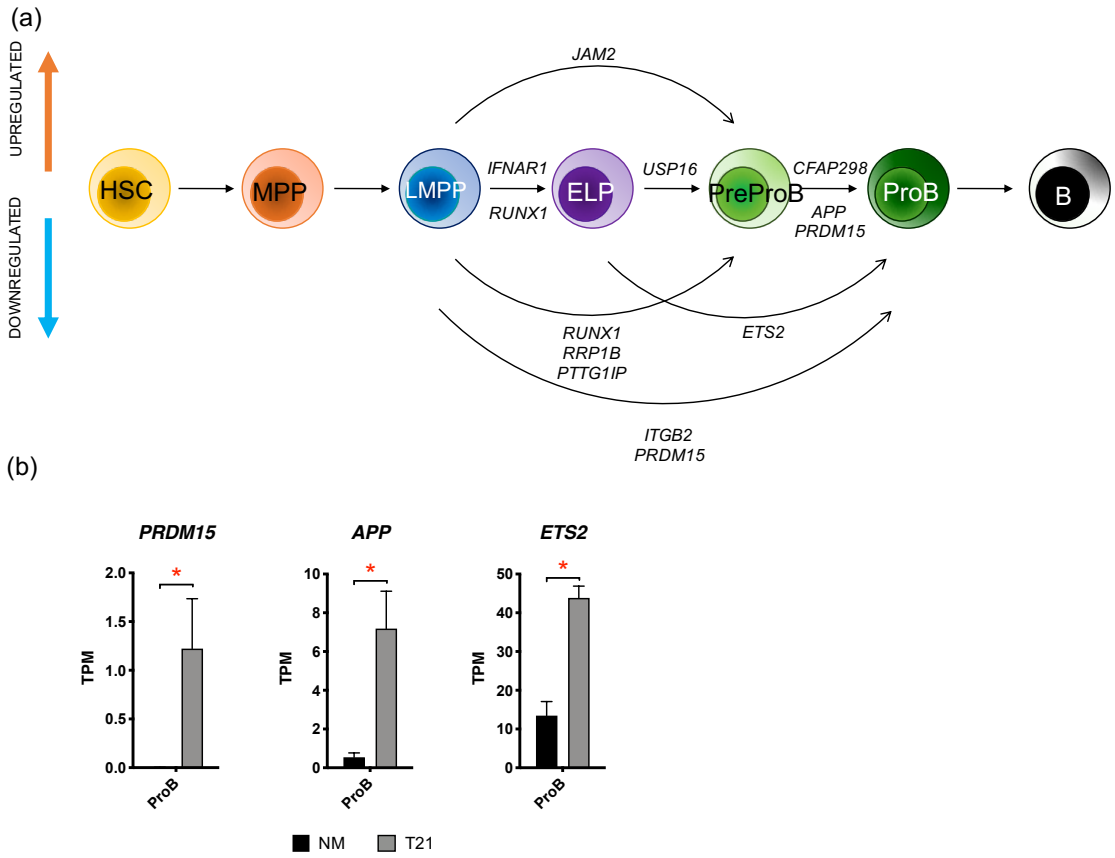


Figure 4.24: Chromosome 21 genes important in fetal B-lymphopoiesis. (a) chromosome 21 genes significantly DE between normal LMPP/ELP/PreProB/ProB (Chapter 3 by RNA-Seq, 100 cells/population, n=3). (b) expression of chromosome 21 genes differentially expressed during normal fetal BM B lymphoid progenitor development and differentially expressed between T21 and NM ProB progenitors. $*$ denotes significance determined by DESeq2 with an FDR < 0.1.

4.5 The impact of the T21 microenvironment on fetal BM haematopoiesis

Analysis of the RNA-Seq data of NM and T21 HSPC showed no striking or significant differences in the expression of genes known to have a direct role in B-lymphopoiesis. This, in conjunction with the enrichment of several gene sets involved in cytokine signalling and inflammatory responses suggested that the B-lymphoid commitment defect might be secondary and at least in part mediated by the T21 microenvironment. In order to investigate the role of the T21 microenvironment, I derived mesenchymal stromal cells (MSC) from NM and T21 FL and fetal BM for gene expression, proteomic and functional characterisation.

4.5.1 The impact of T21 MSC on normal B-lymphopoiesis

To establish whether T21 MSC alone were sufficient to perturb B-lymphopoiesis, I used the co-culture system described in Chapter 2 and Chapter 3. NM fetal BM HSPC were sorted and co-cultured on NM fetal BM MSC, T21 fetal BM MSC and MS-5 stroma in parallel at a density of 100 cells/well. At 7 day time points, a fraction of each co-culture (or where cell numbers permitted whole wells) was analysed by flow cytometry for the presence of myeloid (CD33+/CD11b+/CD14+), B cells (CD34-CD19+CD10+/-), NK cells (CD34-CD56+), immature (CD34+CD11b/14/33-) and other haematopoietic cells (CD34-CD19-CD33/CD11b/CD14-CD56-). An example of the gating strategy I used is shown in figure 4.25 where MSC were excluded according to their expression of CD45 and CD73 (CD45-CD73+). In every experiment the NM HSPC grown on MS-5 co-cultures yielded B cells. Data presented here are representative of 3 independent experiments where 3 three donors were used for NM HSPC, three donors were used for NM MSC and three donors were used for T21 MSC.

By day 14, the commitment of PreProB and ProB progenitors to the B-lineage was clear as all plated B progenitors had differentiated into B cells and there was no difference in the ability of T21 vs normal MSC to support B cell development (figure 4.26). In contrast, HSC, MPP, LMPP and ELP showed a variable decrease in B-lymphoid potential on T21 MSC that was more apparent in later time-points (figure 4.26). To make sure that any observed difference in B-lymphopoiesis was not overshadowed by the profound increase in myelopoiesis observed in the HSC/T21 MSC co-cultures (figure 4.27(a)), I also calculated the absolute number of B cells and myeloid cells produced in each HSC/MSC co-culture (figure 4.27(b)). While these

data illustrate the variability observed between each experiment, it is temping to speculate that T21 MSC induce a skew in haematopoiesis towards myelopoiesis and at the expense of B-lymphopoiesis although further experiments will be required to confirm this.

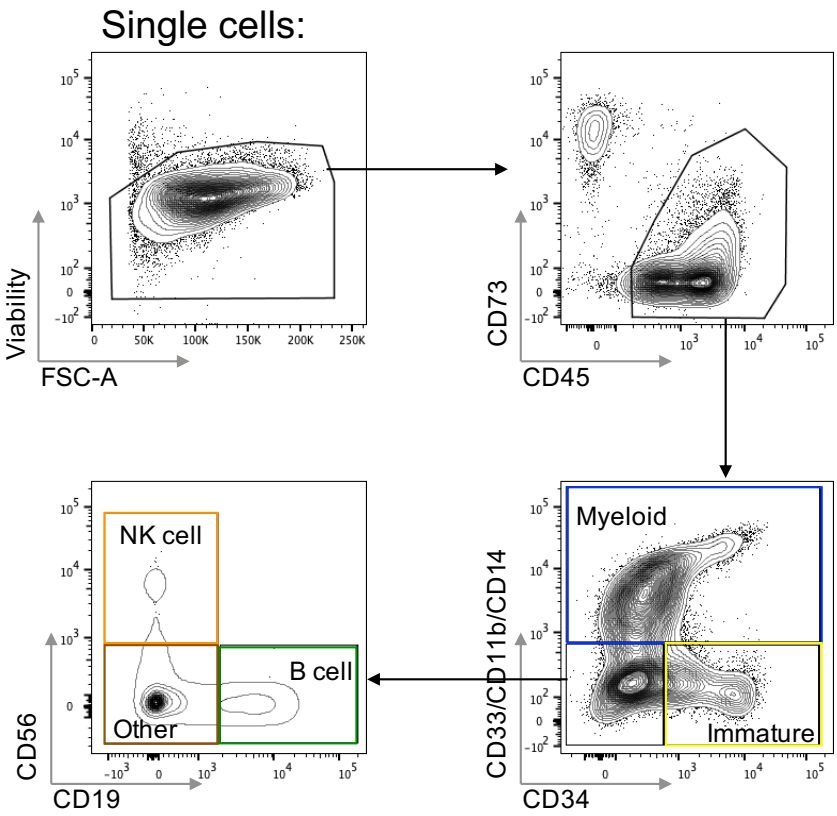


Figure 4.25: Representative plots showing gating strategy used in analysis of MSC co-culture assays. NM HSC/MSC co-culture shown.

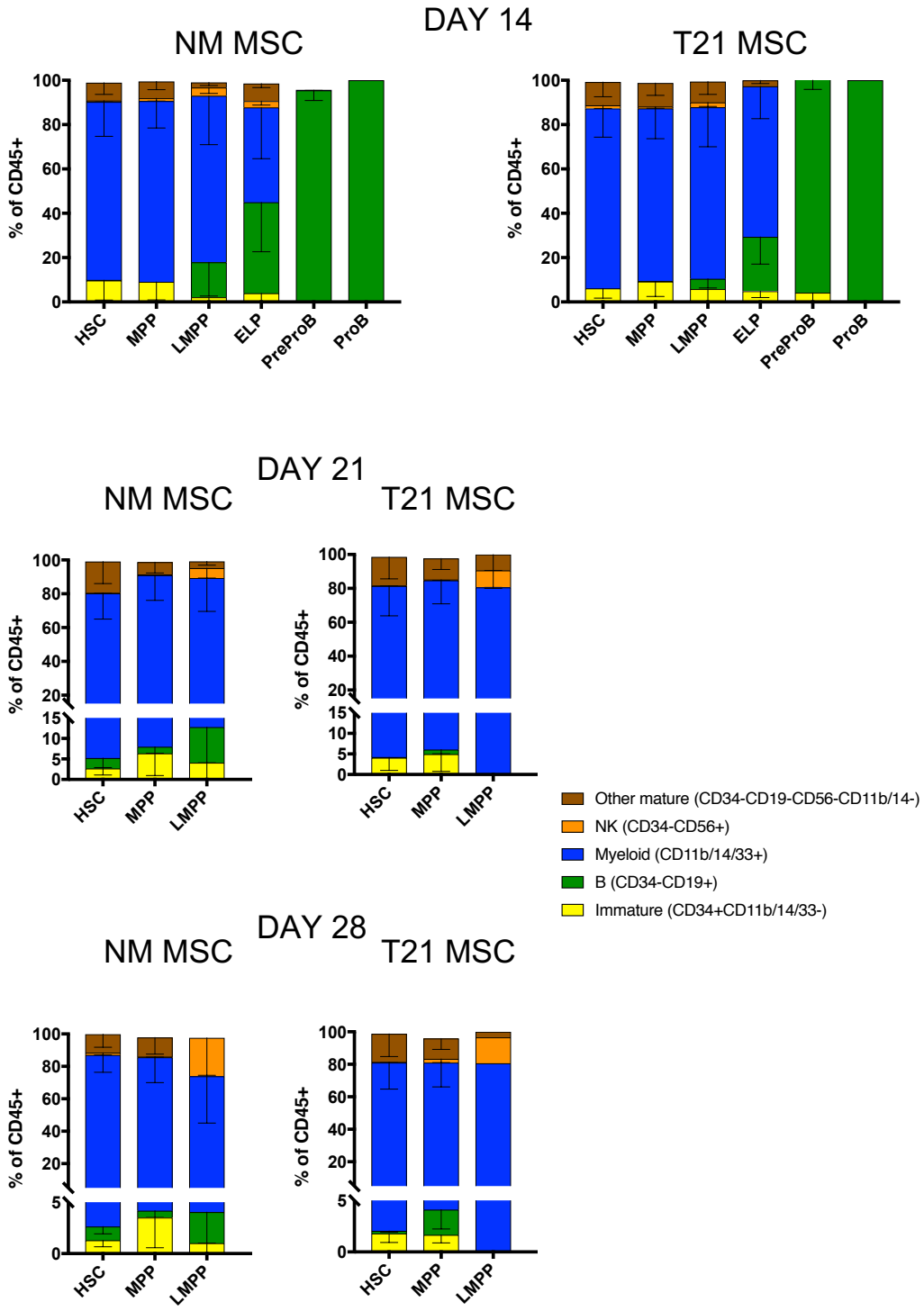


Figure 4.26: NM and T21 MSC co-culture assays: 100 HSC, MPP, LMPP, ELP, PreProB or ProB progenitors were sorted and plated on NM or T21 MSC and co-cultured for 28 days. Each bar chart shows the relative frequency of mature and immature cell types produced as a proportion of CD73-CD45+ cells. n=3.

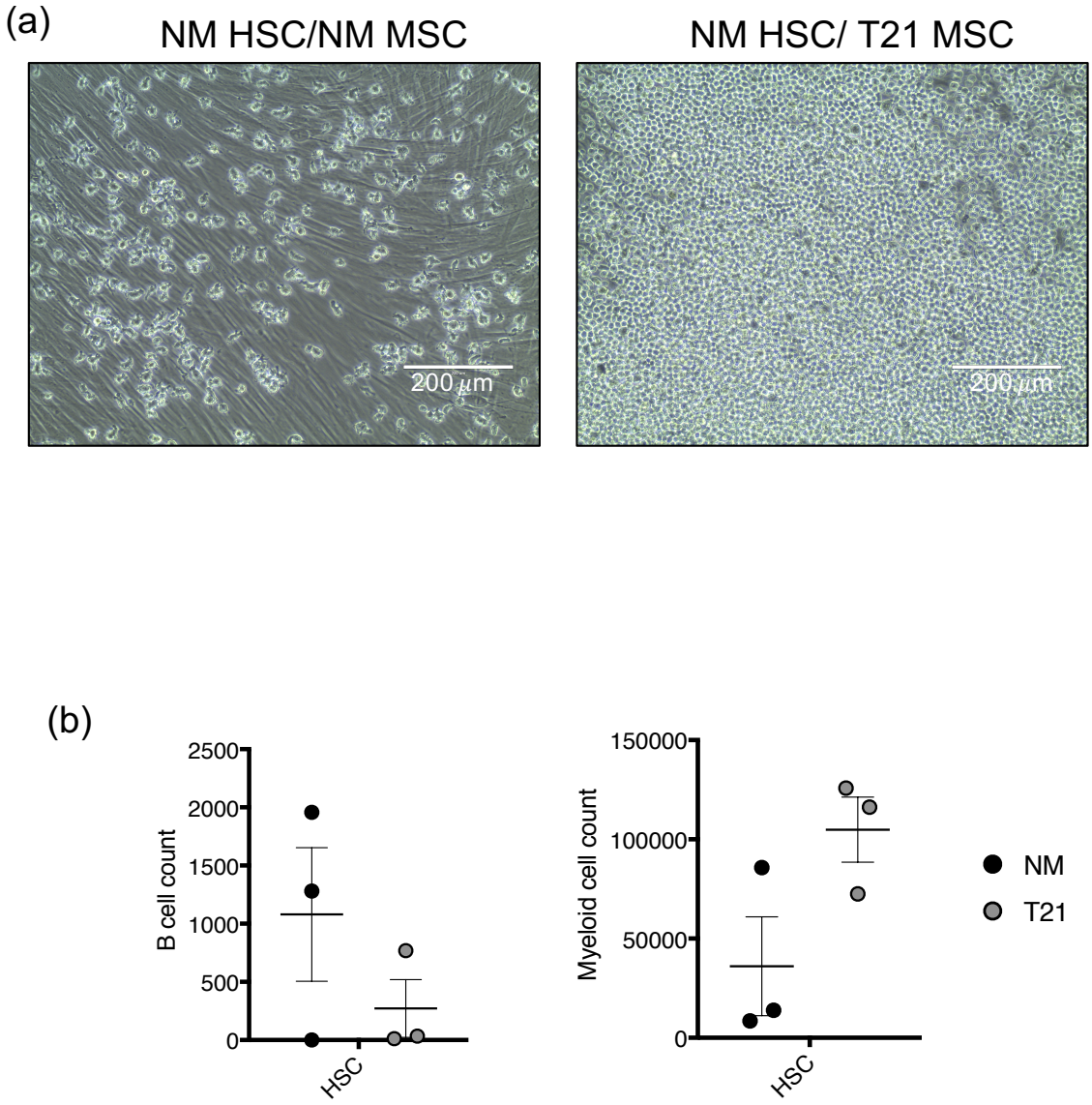


Figure 4.27: NM and T21 MSC co-culture assays: 100 HSC were sorted and plated on NM or T21 MSC and co-cultured for 28 days. (a) pictures of NM and T21 MSC co-cultures at day 28 (b) Dot plots showing the absolute count of B cells (left) and myeloid cells (right) produced from 100 sorted HSC after 21 days in culture with NM or T21 MSC. n=3. B count p-value: 0.381; Myeloid count p-value: 0.076.

4.5.2 The impact of T21 on the transcriptome of fetal MSC

The MSC co-cultures suggested that cell extrinsic factors may be playing a role in perturbing B-lymphopoiesis at the HSC and LMPP level in T21 fetal BM. To explore the putative molecular mechanisms responsible for this, I re-analysed a microarray dataset of the transcriptomes of NM and T21 MSC produced by a previous member of the lab (Dr. David O'Connor). There were a total of 5 biological replicates for NM FL and 3 biological replicates for T21 FL; and a total of 4 biological replicates for NM fetal BM and 3 biological replicates for T21 fetal BM.

Differential expression analysis comparing NM and T21 fetal BM MSC revealed 77 significant DE genes including several genes encoding secreted factors: *IL6, FGF16, GDF6, F10* and *RELN* (figures 4.28 and 4.29; complete DE tables are available in the online supplemental files). GSEA showed an enrichment for IFNα response in T21 MSC, and enrichment for TNFα and IL-6 signalling in NM MSC (figure 4.29).

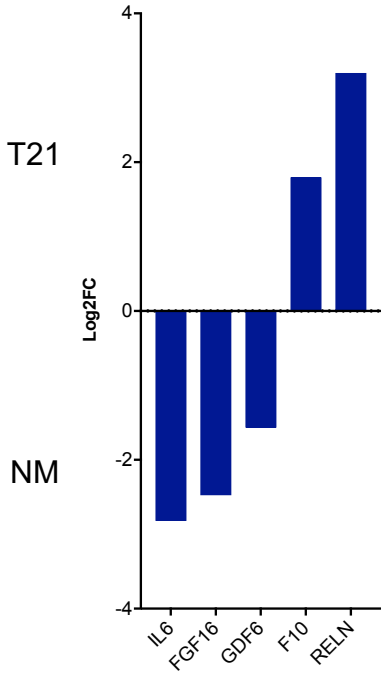


Figure 4.28: Genes encoding known secreted factors which were significantly DE ($p<0.05$) between NM fetal BM MSC (n=4) and T21 fetal BM MSC (n=3) according to microarray analysis.

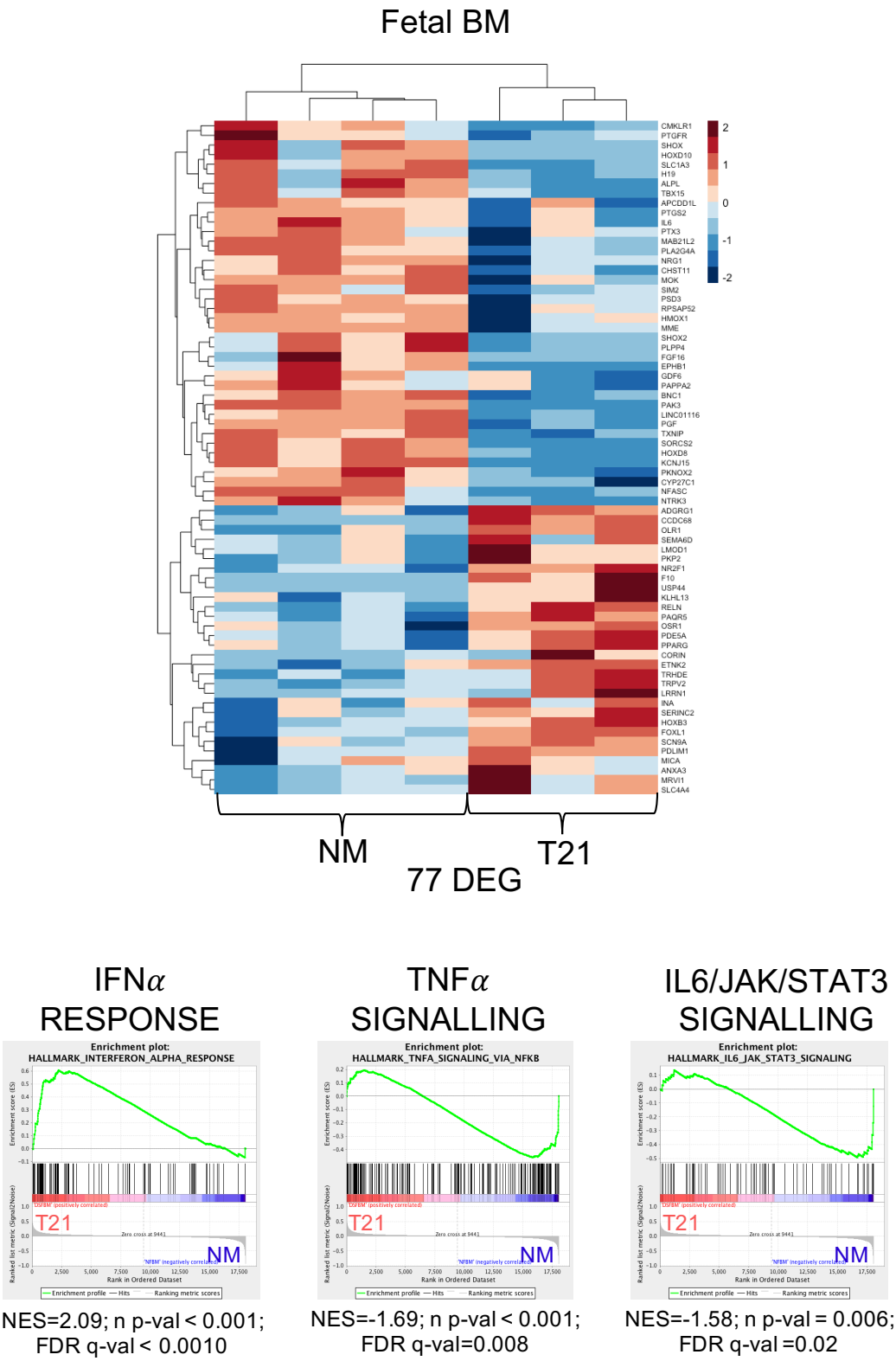


Figure 4.29: Heatmap of all significant DE ($p<0.05$) genes between NM (n=4) and T21 (n=3) fetal BM MSC. GSEA enrichment plots showing enrichments in IFNα in T21 MSC and TNFα and IL-6 in NM MSC according to gene expression by transcriptome analysis by microarray. DEG: differentially expressed genes.

Given that the main site of fetal B-lymphopoiesis is the fetal BM, I also compared gene expression of NM fetal BM MSC with NM FL MSC as a way of identifying possible BM-specific factors important for B cell development. There were 164 significant differentially expressed genes including genes encoding several known secreted factors: *ESM1, NRP2, C3, MASP1, C1S, GDF6, PDGFD, IGFBP5, RELN, FGF16, PAPPA2, BMP2, FGF7, SNCA, ANOS1* and *MMP3* (figures 4.30 and 4.31). These included increased FL MSC expression of complement protein genes *C3, C1S* and *MASP1* (liver is one of the main sites of complement protein synthesis throughout life) and IGF-signalling related genes: *IGFBP5* and *PAPPA2*.

Of the DE genes from the NM fetal BM and FL MSC comparison and encoding secreted proteins, 3 were dysregulated in T21 fetal BM MSC compared to NM fetal BM MSC. Expression of the gene *RELN* was higher in NM FL MSC compared to NM fetal BM MSC; and in T21 fetal BM MSC compared to NM fetal BM MSC (figures 4.28 and 4.30). On the other hand *GDF6* and *FGF16* expression was higher in NM fetal BM MSC than in NM FL MSC, but when compared to T21 fetal BM MSC, both of these genes were down-regulated (figures 4.28 and 4.30). Finally, GSEA also revealed that multiple inflammatory signalling and pathway gene sets were enriched in FL compared to fetal BM MSC (figure 4.31). Taken together, these data suggest that T21 fetal BM MSC share several transcriptomic characteristics with normal FL MSC which is confirmed when a PCA of the top 500 most variably expressed genes is plotted and shows 2/3 T21 fetal BM MSC clustering with FL MSC by PC1 (figure 4.32).

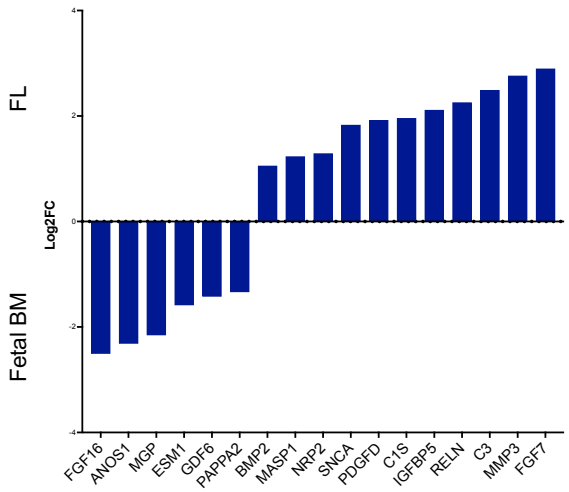


Figure 4.30: Genes encoding secreted factors significantly DE ($p<0.05$) between NM FL MSC (n=5) and NM fetal BM MSC (n=4) according to microarray analysis.

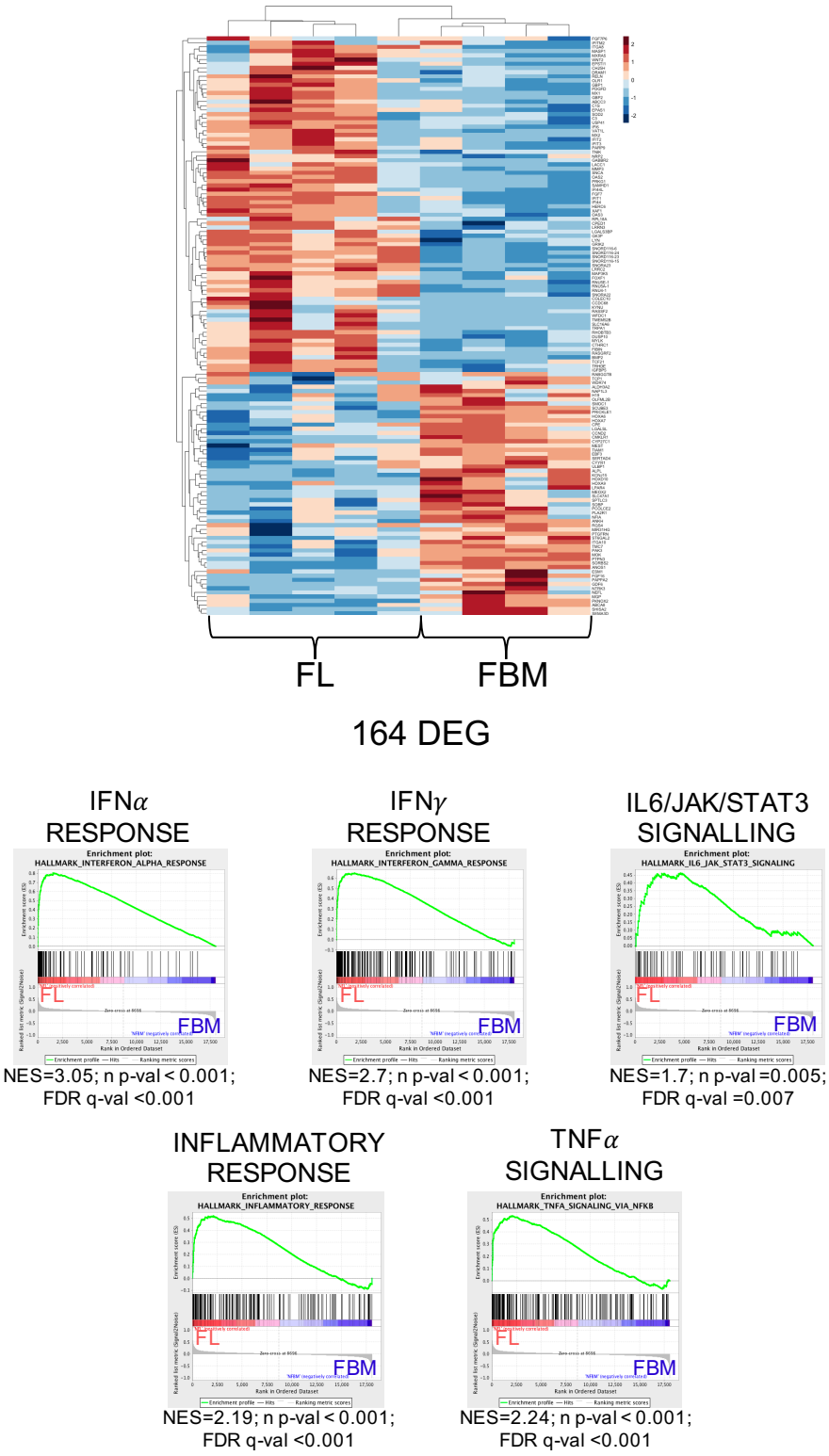


Figure 4.31: Heatmap of all significant (p<0.05) DE genes between NM FL (n=5) and NM fetal BM (n=4) MSC. GSEA enrichment plots showing enrichments in multiple inflammatory signalling pathways according to gene expression by transcriptome analysis by microarray. DEG: differentially expressed genes.

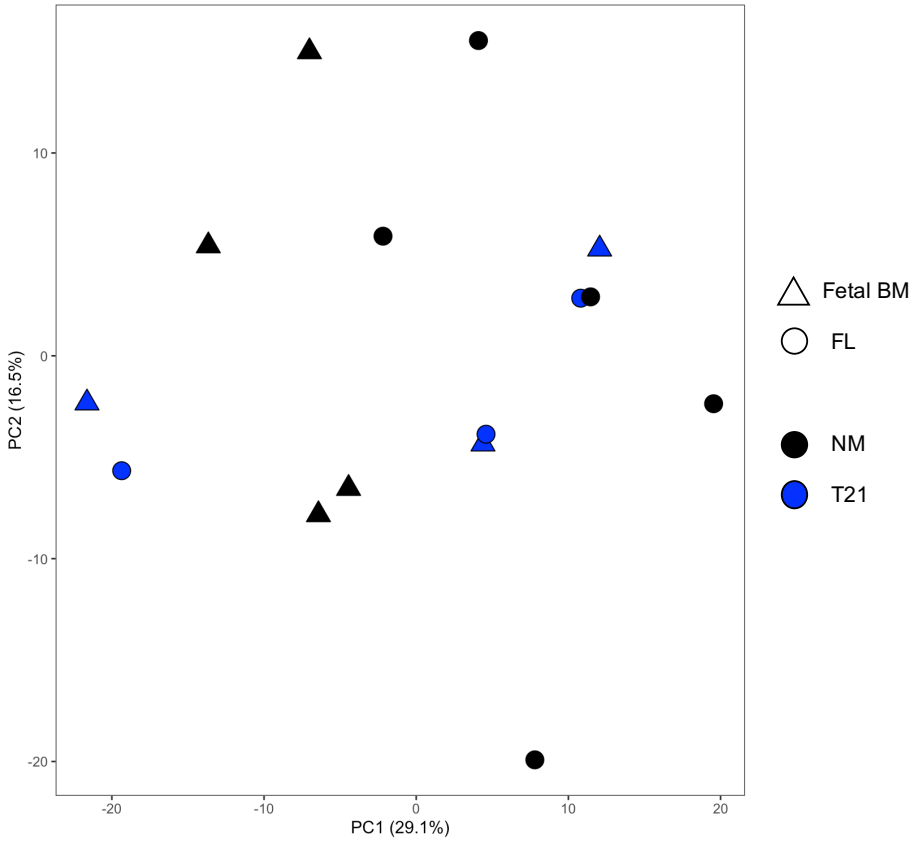


Figure 4.32: PCA of top 500 most variably expressed genes in fetal MSC by microarray. 2/3 T21 fetal BM MSC cluster with FL MSC by PC1.

4.5.3 The proteome of fetal MSC

To see if the differences observed in the transcriptome of MSC were reflected in the proteome, samples of membrane and secreted proteins from MSC were submitted for proteomic identification and quantification by mass spectrometry (Advanced Proteomics facility, TDI, Oxford). Where possible three biological replicates were submitted for each condition representative of tissue (FL or fetal BM) and disease phenotype (NM or T21)(figure 4.33). Unfortunately, only 2 biological replicates were available for T21 fetal BM proteomic analysis at the time.

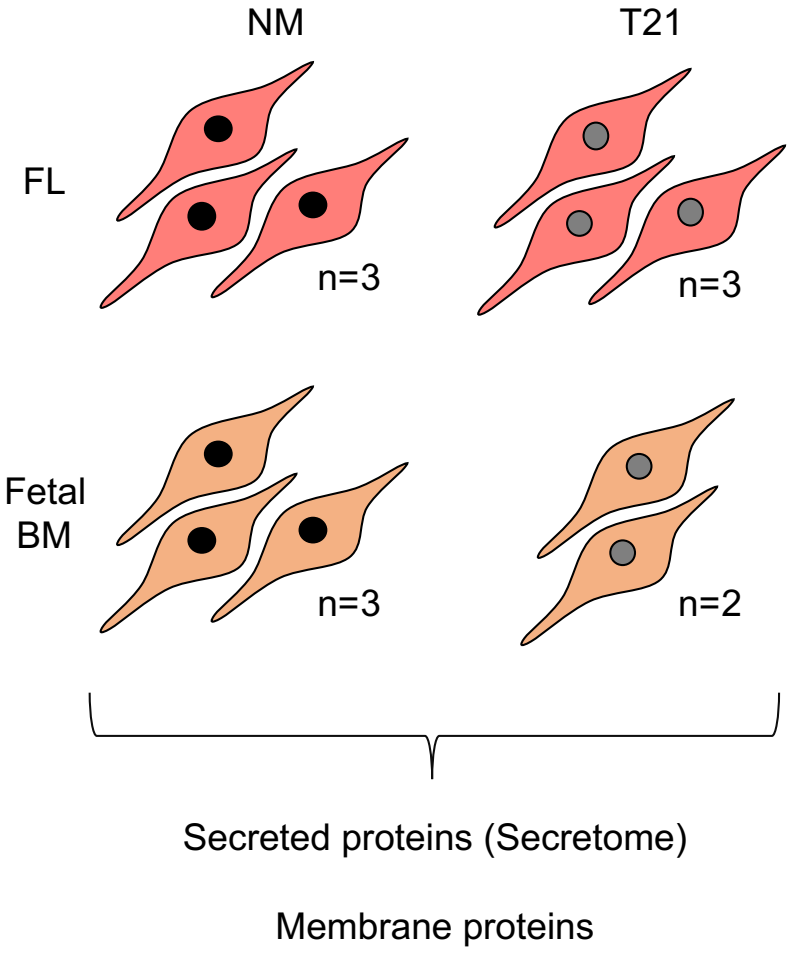


Figure 4.33: Summary of samples submitted for proteomic analysis (mass spectrometry) of membrane and secreted proteins.

Initial identification and quantification of proteins was run by Dr Roman Fisher at the Advanced Proteomics Facility, Oxford and downstream analysis was performed by myself. To establish the most appropriate method of normalisation for the dataset (and because established analysis protocols for proteomic data lag behind the vast

choice available in the transcriptomics field) I performed two kinds of normalisation (median and variance stabilising normalisation, VSN) and differential expression analysis using these two methods of normalisation which is detailed in Chapter 2. As a first step to compare normalisation methods, I ran PCA. While PCAs of both median normalised and VSN data were very similar (figures 4.34 and 4.35), I decided that the differential expression analysis using VSN dataset would provide a more accurate estimate of differentially expressed proteins as extreme values are normalised and did not skew the differential expression analysis. Also, VSN ranked well in a systematic evaluation of normalisation methods [Välikangas et al., 2016] although it was noted that resulting differential expression analysis was conservative. I think this is an appropriate approach especially when attempting to perform differential expression analysis between fetal BM MSC where there were only two biological replicates for T21 MSC.

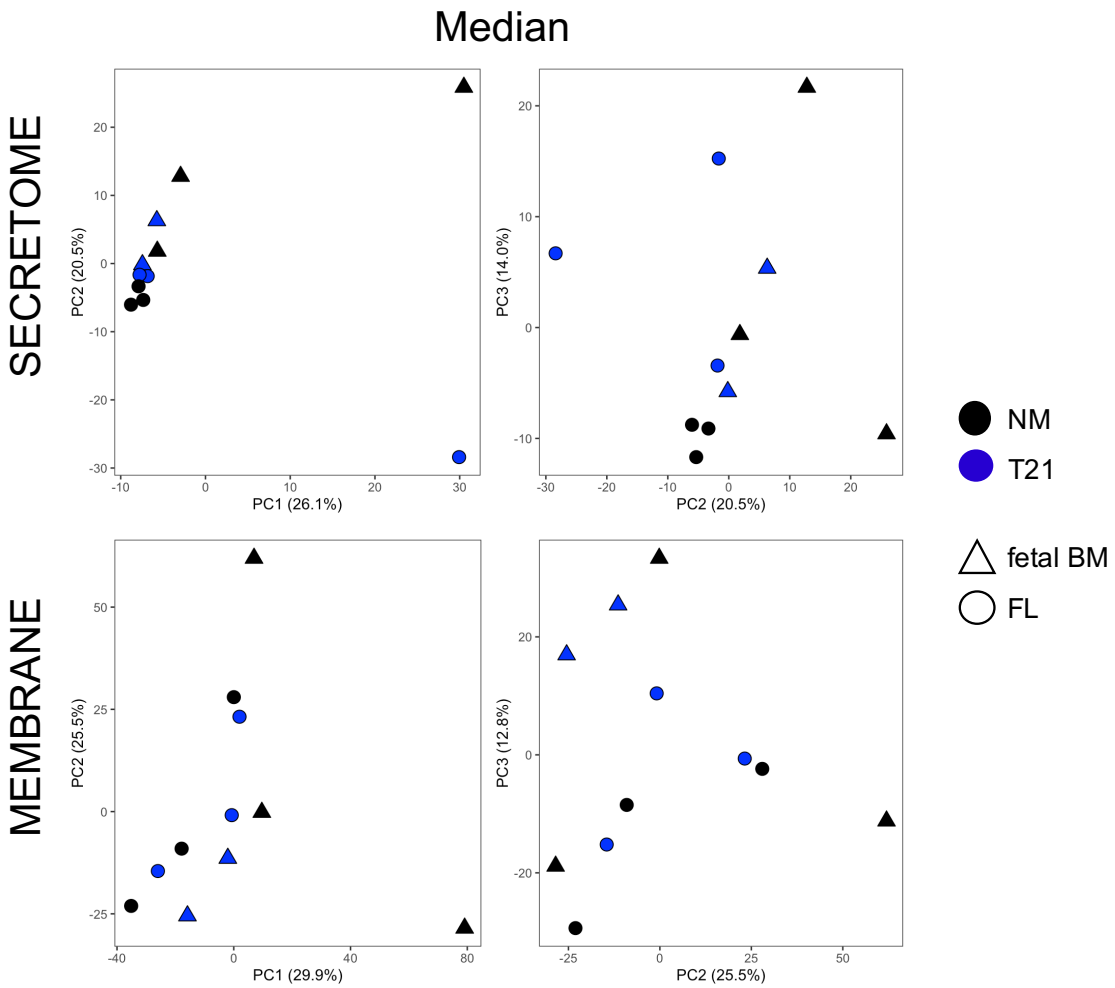


Figure 4.34: PCA of median normalised secretome and membrane proteins identified by proteomics. NM FL, NM fetal BM and T21 FL: n=3; T21 fetal BM: n=2.

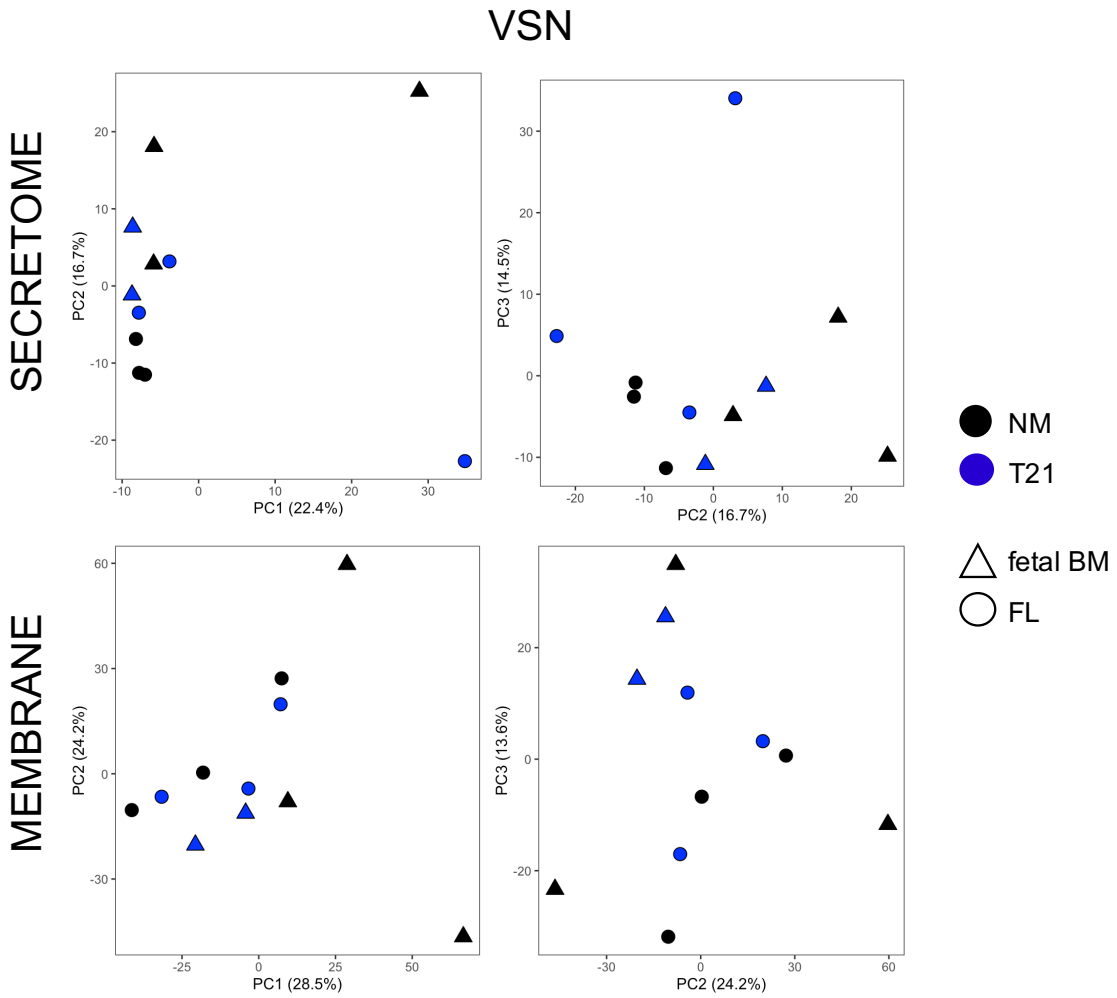


Figure 4.35: PCA of VSN normalised secretome and membrane proteins identified by proteomics. NM FL, NM fetal BM and T21 FL: n=3; T21 fetal BM: n=2.

After normalisation, I performed DE analysis comparing each tissue and chromosome phenotype in pairwise comparisons (tables available as online supplemental files). There are a couple of considerations to note when interpreting these data. Firstly, although MSC samples were fractionated prior to submission, there is contamination presumably from dead and dying MSC which has resulted in proteins that are not known/predicted to be secreted or located in the plasma membrane in each dataset. Secondly, identification of proteins using mass spectrometry in this manner will only identify and quantify the most abundant proteins. Therefore, the concentration of proteins of low relative abundance such as many cytokines cannot be determined using this method.

Starting with the comparison between NM and T21 fetal BM MSC (which should be interpreted with caution because n=2 for T21) there were 34 DE secreted proteins and 32 DE membrane proteins with a logFC > 2 and p-value <0.05 (figure 4.36). Of the secreted proteins, TRH, HAPLN1, F5, CD63, RPN1 and MFAP5 all have potential roles in haematopoiesis (figure 4.36).

As only 2 samples of T21 fetal BM were available for proteomic analysis, I also compared the FL MSC to see if there were any common DE proteins that could shed light on the the trisomic environment as a whole (figure 4.37). This analysis yielded 69 DE secreted proteins and 21 DE membrane proteins (logFC > 2 and p-value <0.05; figure 4.37). Two candidates that were DE in the secretome are of particular interest: TGFB1 and JAG1. TGFB1 is the protein TGF-β1 which is a key regulator of HSC self-renewal [Naka and Hirao, 2017] and also plays a role in MSC regulation [Abou-Ezzi et al., 2019]. JAG1 is the protein Jagged-1 which is critical for the Notch signalling pathway [Gu et al., 2016b]. In addition to these analyses, I also compared all T21 MSC with all NM MSC (figure 4.38) and found that there were 23 DE secreted proteins but no DE membrane proteins (logFC > 2 and p-value <0.05; figure 4.38). Of the 23 DE secreted proteins, 12 were over-expressed in NM MSC, including insulin like growth factor binding protein 5 (IGFBP5), C-C motif chemokine ligand 2 (CCL2) and intercellular cell adhesion molecule-1 (ICAM1); and 11 were over-expressed in T21 MSC, including several extracellular matrix components such as the proteoglycan stablising protein (HAPLN1), the tropelastin protein (ELN) and thrombospondin 3 (THBS3). None of the DE proteins increased in T21 samples were encoded by a gene on chromosome 21 and further work would be required to investigate the significance of these differences.

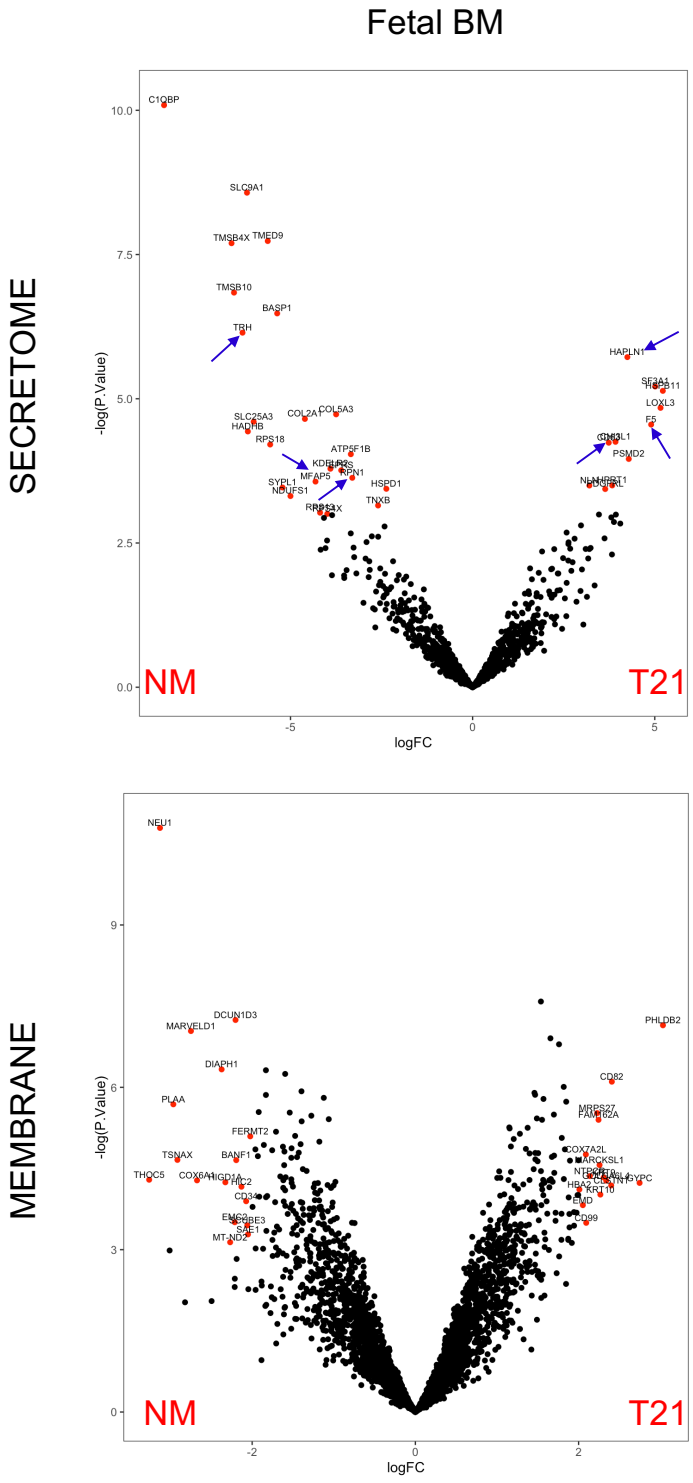


Figure 4.36: DE secretome and membrane proteins by proteomics between NM (n=3) and T21 (n=2) fetal BM MSC. p<0.05. Blue arrows highlight proteins with known roles in haematopoiesis.

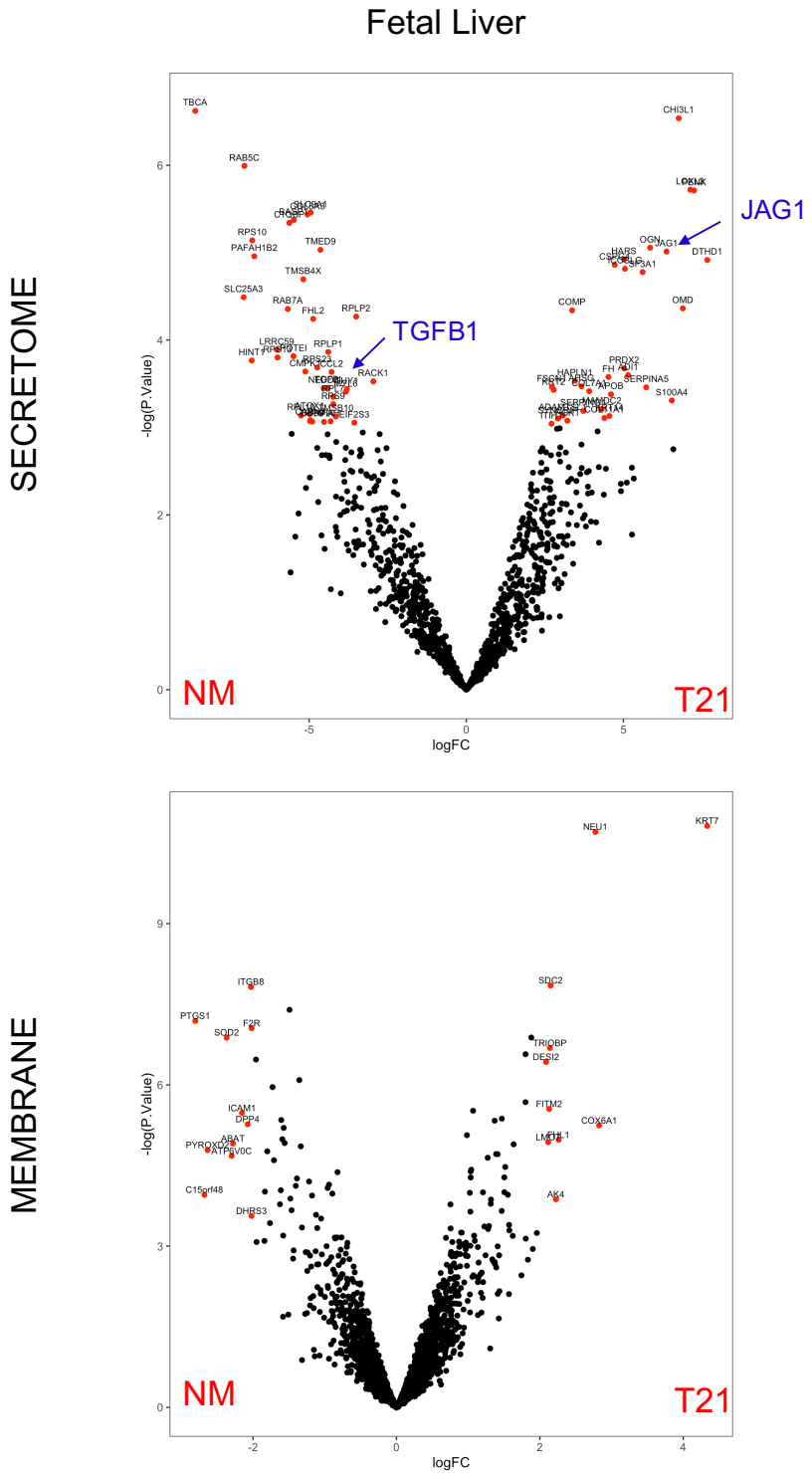


Figure 4.37: DE secretome and membrane proteins by proteomics between NM (n=3) and T21 (n=3) FL MSC. p <0.05. Blue arrows highlight TGFB1 and JAG1 proteins which are of particular interest given their known roles in haematopoiesis.

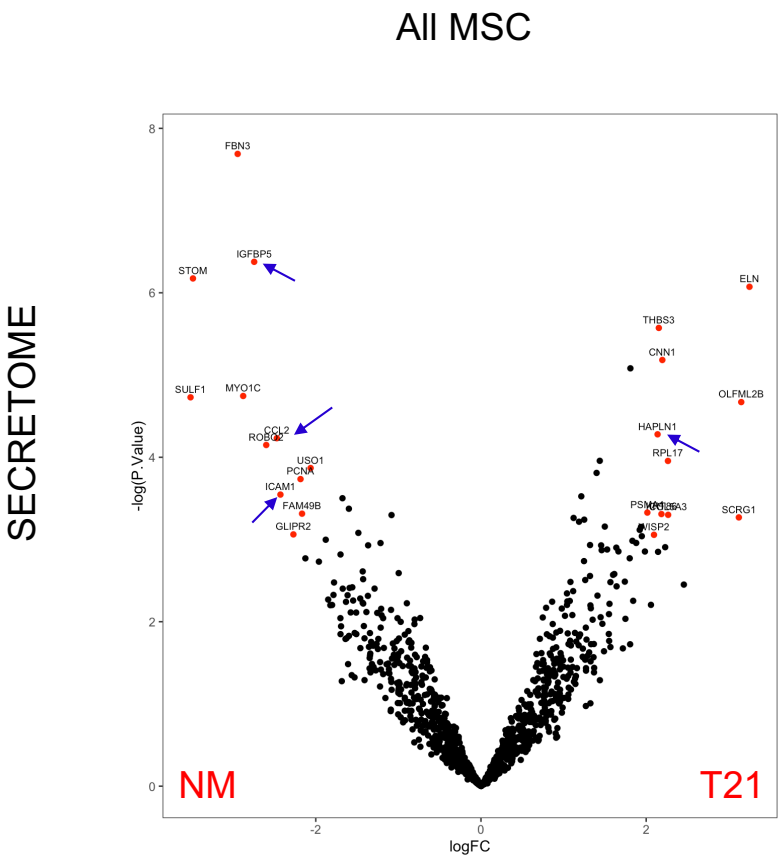


Figure 4.38: DE secretome by proteomics between NM (n=6) and T21 (n=5) MSC (FL and fetal BM combined). p <0.05. Blue arrows highlight proteins with known roles in haematopoiesis.

Finally, in an effort to see if there were other factors that I could add to the MSC co-culture system to optimise B cell output, I compared NM FL and NM fetal BM MSC. Using the same rationale as above (LogFC >2 and p <0.05) there were 129 DE secreted proteins and 48 DE membrane proteins but none which have obvious known roles in B-lymphopoiesis (figure 4.39).

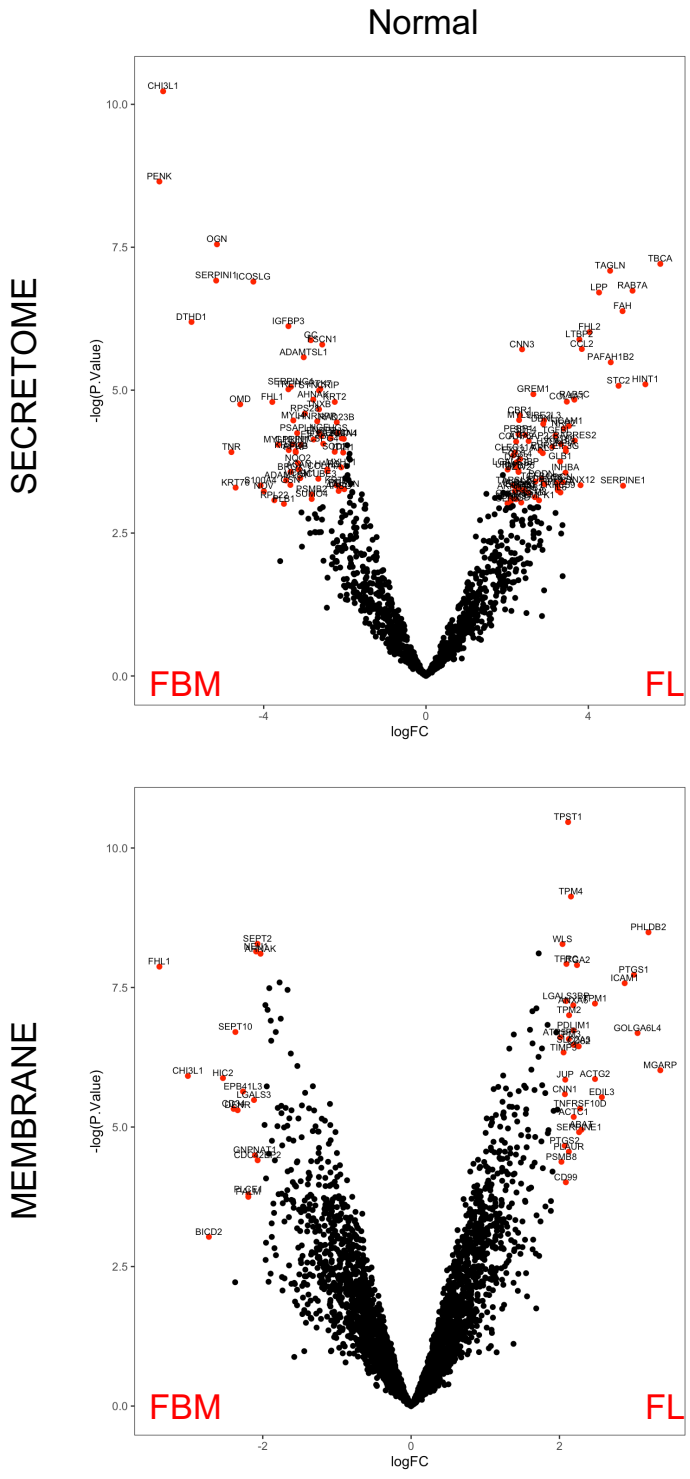


Figure 4.39: DE secretome and membrane proteins between NM FL and NM fetal BM MSC.

Next I compared the DE protein lists with the DE gene lists from the MSC transcriptome analysis (figure 4.40). Overlap between DE secreted proteins and DE genes was confined to the differences between tissues i.e. FL and fetal BM. This observation is interesting in itself because it suggests that there are no differences in secretome of NM and T21 MSC. Also, none of the genes/proteins found to be DE in the transcriptome and proteome analysis are located on chromosome 21. Strikingly, the largest overlaps were in the T21 FL vs T21 fetal BM comparison. This is probably because the microarray DE analysis revealed 2175 DE genes which is a much larger total number of DE genes when compared to the other microarray DE analyses that yielded between 86 and 184 DE genes. Whether this difference is of biological significance will require further investigation.

Of the differentially secreted proteins between NM tissues, IGFBP5 is over-expressed in NM FL compared to NM fetal BM which given its role as a negative regulator of IGF signalling [Ding et al., 2016]; the role of IGF signalling in cancer (reviewed here: [Pollak, 2012]); and specifically the pathogenesis of ML-DS [Klusmann et al., 2010a] suggests that the role of IGFBP5 in the FL microenvironment warrants further investigation.

Of the DE secreted proteins and genes in the comparison between T21 FL and T21 fetal BM MSC, the overexpression of RAB5C in T21 fetal BM MSC is of potential interest as its regulation has been linked to leukaemia progression [Tan et al., 2014].

In terms of membrane proteins, there was only one DE gene and protein in the comparison between fetal BM MSC: ALPL. However, given that the gene was over-expressed in NM and the protein over expressed in T21, this result is probably a consequence of differences in biological samples.

SECRETOME

	NM FL	T21 FL	NM Fetal BM	T21 Fetal BM
NM FL				
T21 FL	None			
NM Fetal BM	IGFBP5, LGALS3BP, NRP2, SCUBE3, SOD2	Comparison not made		
T21 Fetal BM	Comparison not made	AEBP1, ESD, PEBP1, PLBD2, RAB5C, RPS16, S100A4, TARSL2, TFRC, UFM1,	None	

MEMBRANE

	NM Fetal Liver	T21 Fetal Liver	NM Fetal BM	T21 Fetal BM
NM Fetal Liver				
T21 Fetal Liver	DHRS3, ERAP2, NRP2, PTGS2			
NM Fetal BM	IGFBP5, ITGA1, LGALS3BP, NRP2, SCUBE3, SOD2	Comparison not made		
T21 Fetal BM	Comparison not made	*	ALPL	

* ACP6, ARCN1, ATP6V0C, B2M, BANF1, COL6A2, DENR, DIABLO, DDRGK1, DDX6, DYNC1LI1, EIF4G2, EPN2, FAHD2A, FLAD1, FKBP2, FXR1, GNAI3, HNRNPK, IMPAD1, ITGA5, ITGB1, LAMTOR3, LARS2, MRPL3, MRPL21, MRPS30, MRPS18C, NAXE, NCKAP1, NDUFA4, NDUFAF7, NDUFB4, NDUFB10, NEXN, NUCB2, PABPC4, PRDX3, PRDX5, PSMD8, RDX, RPA1, RPL11, RPL23, PPP1CA, RPS8, RPS13, RPS16, ROCK1, SDHC, SEPT7, SF3B5, SPTBN1, SSBP1, TBRG4, TFRC, TIMM44, THEM6, UPF1, VMP1.

Figure 4.40: Overlaps between significantly expressed genes by microarray ($p<0.05$) and proteins differentially expressed by proteomic analysis ($p<0.05$). Overlap in gene/proteins detected in both membrane and secreted fractions is most probably a result of contamination from dead or dying cells during purification. However in some cases (e.g. NRP2) proteins are predicted to be secreted and expressed on the cell surface.

4.5.4 The role of IL-6

The cytokine IL-6 is an attractive candidate for having a role in the T21 mediated B-lymphoid defect in fetal life. It was initially described as B cell stimulating factor (BSF-2) due to its role in maturation of B cells into plasma cells [Hirano et al., 1986]. By gene expression (microarray; figure 4.41(a)) and protein expression (quantitative mass spectrometry; figure 4.41(b)), IL-6 expression is lower in T21 fetal BM MSC; and both IL-6 receptor genes (*IL6R* and *IL6ST*) are expressed (by RNA-Seq) in NM and T21 HSPC upstream of B-lineage commitment (figure 4.41(c)). The lack of *IL6R* expression in B cells (non-stimulated) is consistent with previous observations in mice [Hirata et al., 1989]. GSEA showed an enrichment for IL-6 signalling in NM fetal BM MSC compared to T21 also suggesting down-regulation of IL-6 signalling in T21 fetal BM (figure 4.29). However, GSEA of the RNA-Seq of NM vs T21 HSC showed an enrichment for IL-6 signalling in T21 HSC (figure 4.22). This enrichment would initially appear in conflict with what I observed in the MSC data. However, IL-6 is part of a family of cytokines that are all defined by their common binding to receptor subunit gp130 (encoded by the gene *IL6Rβ*). As a result of this shared receptor subunit, all members of the IL-6 cytokine family display a high degree of redundancy in the downstream signalling pathways they evoke (reviewed extensively in [Jones and Jenkins, 2018] and [Murakami et al., 2019]). Since LIF is part of the IL-6 family of cytokines and its receptor subunit gene *LIFR* is significantly over expressed in T21 HSC (figure 4.41(c)), it is reasonable to hypothesise that this enrichment is actually driven by this DE gene and is a direct result of increased LIF signalling which shares many downstream effector signalling pathways with classic IL-6 signalling responses. However, it should be noted that to date, the mechanisms that are redundant between IL-6 family members and what makes them specific and distinct remains unclear.

With evidence that T21 mediated dysregulation of IL-6 persists in post-natal life [Zampieri et al., 2014]; others showing that IL-6 affects the ability of an ELP-like cell to expand (but not commit) in mice [Maeda et al., 2005]; and IL-6 expression linked to several miRNAs located on chromosome 21 [Migita et al., 2017], [Lin et al., 2016]; I hypothesised that supplementation of IL-6 in MSC co-cultures would rescue the B-lymphoid defect caused by T21 MSC in the *in vitro* co-culture system.

MSC co-cultures: To investigate this hypothesis, I adapted the MSC co-culture system by plating 200 sorted Lin-CD34+CD38- HSPC (HSC, MPP and LMPP) on NM MSC or T21 MSC. Parallel co-cultures with MS-5 stroma were used as a positive

control. As with previous MSC and MS-5 stromal co-cultures, media was supplemented with Flt3L, SCF, IL-2 and IL-7 and then also with or without 10ng/mL IL-6. At 7-day intervals whole wells were analysed by flow cytometry using the same output panel and gating strategy described earlier in figure 4.25. These data are representative of three independent experiments.

Myelopoiesis rather than lymphopoiesis, dominated on all three stromal layers as observed previously (figure 4.42). Also, the B-lymphoid defect observed when NM HSPC were plated on T21 MSC was preserved: T21 fetal BM MSC demonstrated impaired ability to support B-lymphopoiesis compared to NM fetal BM MSC. However, the addition of IL-6 had minimal effects. While it could be argued that IL-6 may have increased B-lymphopoiesis on NM MSC, there was no discernible increase in B-lymphopoiesis on T21 MSC supplemented with IL-6 (figure 4.42).

While these data do not reveal the pattern or correlation I expected, they do not rule out the possibility that IL-6 deficiency caused by the T21 microenvironment plays a role in the T21 B-lymphoid defect. There are a number of steps that could be taken to further optimise this system: sorting pure populations of HSC, MPP and LMPP separately; or sorting other downstream progenitors that were excluded in the CD38- gate, such as ELP (Maeda *et al.* had shown that IL-6 affects ELP cycling and not commitment [Maeda et al., 2005]). However, even with these adaptations (and as with previous co-cultures where pure populations were sorted) biological variability between samples makes searching for small differences challenging and therefore perhaps a completely different approach such as knocking down IL-6 in NM MSC would be more successful. Finally, the role of other cytokines acting in a concerted manner alongside IL-6 in attenuating B-lymphopoiesis cannot be ruled out.

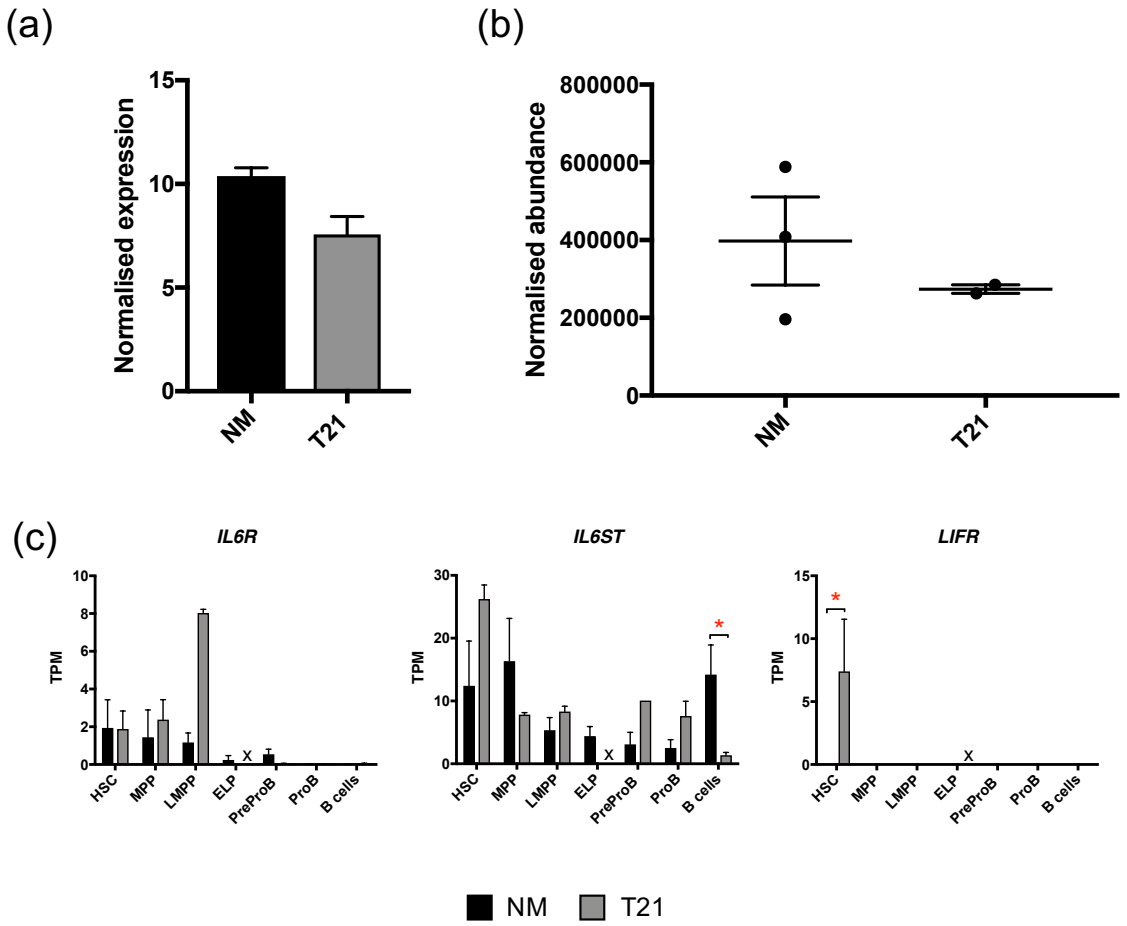


Figure 4.41: (a) *IL6* gene expression in NM (n=4) and T21 (n=3) fetal BM MSC measured by microarray. (b) IL-6 protein expression in NM (n=3) and T21 (n=2) fetal BM MSC measured by quantitative mass spectrometry (proteomics). (c) IL-6 family receptor gene expression in NM (n=3) and T21 (n=2-5) HSPC by RNA-Seq (100 cells/population) $*$ denotes significance determined by DESeq2 with an FDR < 0.1.

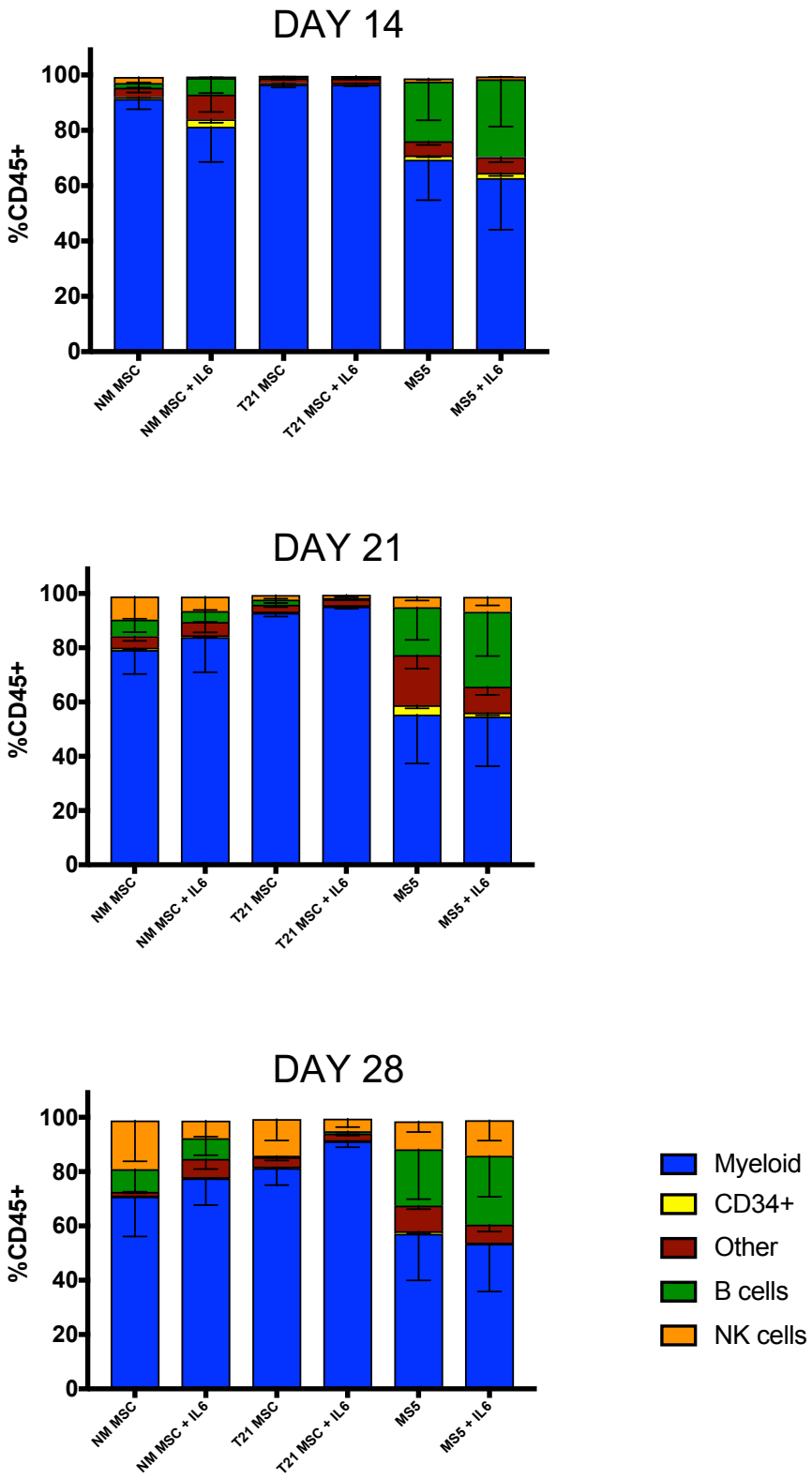


Figure 4.42: Frequency of B cells, NK cells, myeloid cells, other cells and immature cells as a proportion of all CD45+ cells after co-culture on NM or T21 MSC at indicated time points. At day 0, 200 sorted Lin-CD34+CD38- were plated for co-culture with NM MSC, T21 MSC or MS-5 stroma +/- IL-6. n=3.

Validation of IL-6 expression in MSC: Despite the lack of an obvious functional difference in the co-culture experiments after supplementation with IL-6, I sought to validate the findings of the DNA microarray and proteomics analysis showing that IL-6 was down-regulated in T21 MSC. RT-qPCR and ELISAs were used to measure *IL6* gene expression and protein secretion respectively. Surprisingly both the RT-qPCR and ELISAs showed the opposite to what had previously been observed in the trancriptomic and proteomic experiments: IL-6 expression was higher in the T21 MSC compared to NM MSC (figure 4.43). This discrepancy could in part be attributed to the inter-individual differences in biological replicates. For example, the two biological replicates (DSOX12 and DSOX15) used in RT-qPCR and ELISAs that showed highest expression of IL-6 were not available at the time for proteomic analysis. Another factor that could be important in the interpretation of these findings is that IL-6 is considered an acute phase protein and therefore its expression is very dynamic.

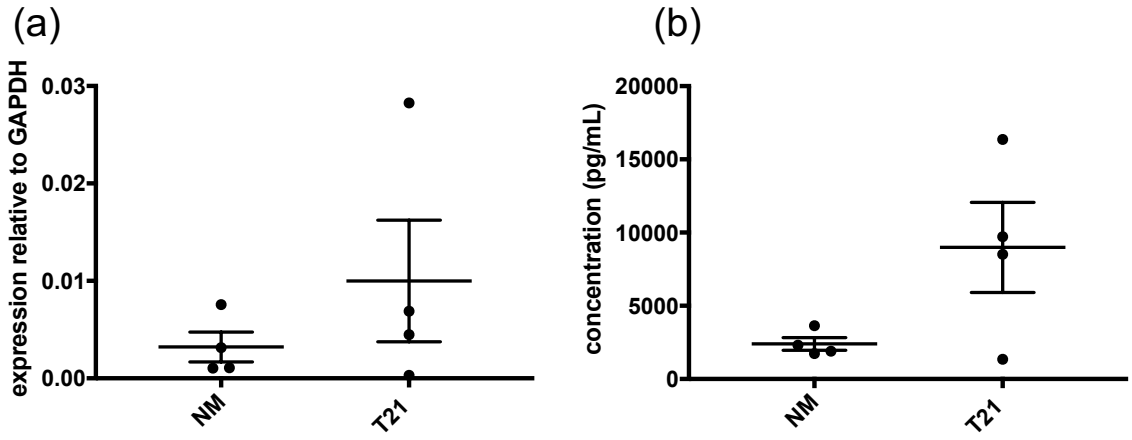


Figure 4.43: (a) RT-qPCR and (b) ELISA for IL-6 gene and protein expression in NM (n=4) and T21 (n=4) MSC.

miRNA expression in MSC: There are numerous studies that propose a role for microRNAs (miRs) in the post-transcriptional repression of IL-6 mRNA, its regulators (e.g. NF-kB) or its downstream effectors (e.g. STAT3) ([Iliopoulos et al., 2009], [Lin et al., 2016], [Jablonski et al., 2016], [Migita et al., 2017]) thereby modulating IL-6 signalling (albeit in a context dependent manner i.e. cell or cancer type). These miRs include but are not limited to a miR cluster located on chromosome 21 that encodes miR-Let7c, miR-99a, miR-125b and miR155. Since these miRs have already been implicated in haematopoiesis and/or leukaemogenesis ([Li et al., 2018], [Shaham et al., 2012], [Dorsett et al., 2008], [Emmrich et al., 2014]; all miRs associated specifically

with B-lymphopoiesis reviewed here: [Zheng et al., 2018]); and are possibly markers of inflammation, RT-qPCR for miRNA was performed on NM and T21 MSC to investigate any correlations in their expression with IL-6. Also, mature miRs are only produced after post-transcriptional processing in the cytosol where the Dicer complex cleaves the hairpin-loop pre-miRNA structure to form two complementary miRs (figure 4.44). Generally the 3' miR is targeted for degradation and the 5' is considered the dominant miR with higher relative abundance. However, there are multiple exceptions to this and so when the dominant arm is unknown, miRs are distinguished by the notation of "3p" or "5p" referring to the 3' or 5' end of the stem loop structure that they came from [Ambros et al., 2003]. As published information regarding the dominant arm for the chromosome 21 miRs was not available, RT-qPCR was performed for both 3p and 5p miRs on chromosome 21.

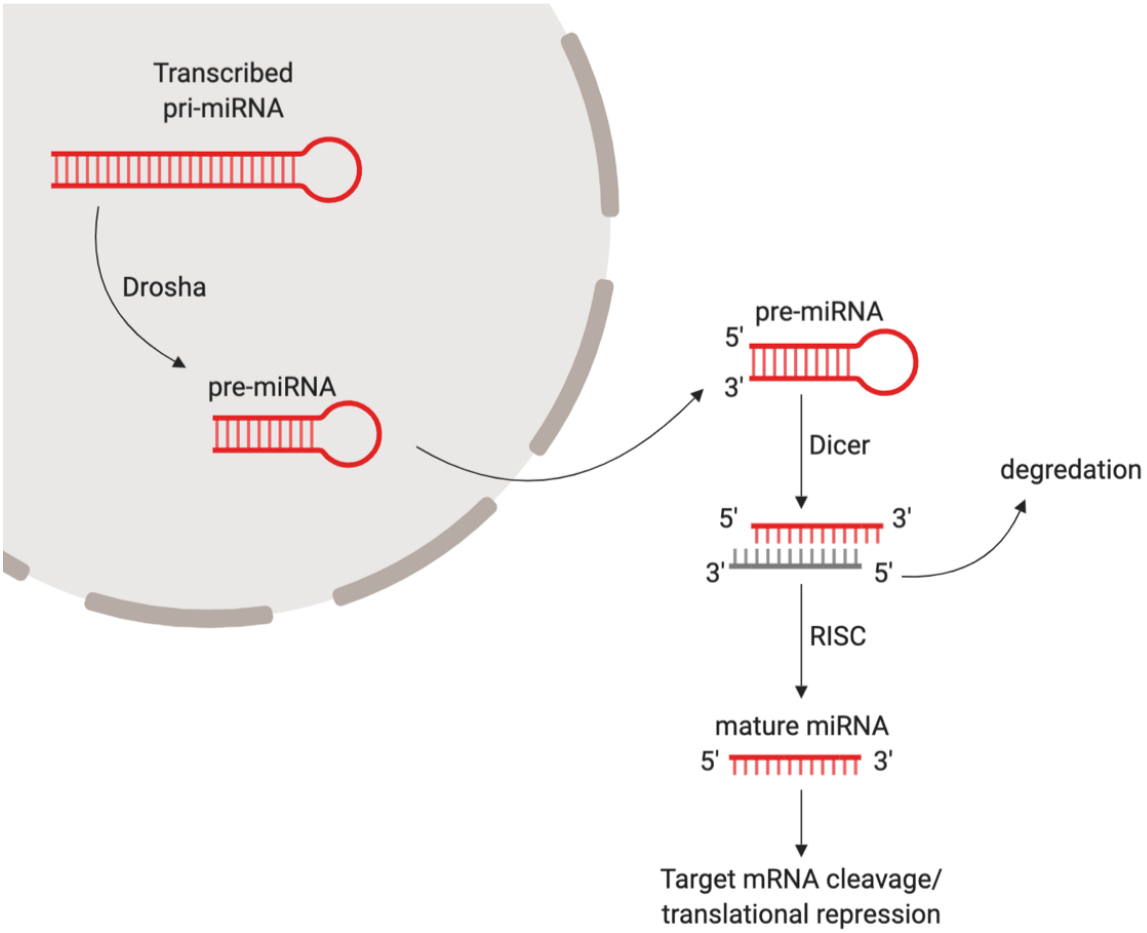


Figure 4.44: Illustration showing the main steps from transcription to mature miRNA created with BioRender.com. Usually the 3'-5' (3p) complementary strand is targeted for degradation and the mature miRNA is the remaining 5'-3' (5p) strand.

Firstly, the high relative abundance of the 5p compared to 3p forms of all miRs tested suggested that the 5p form of each miR was indeed the dominant product for that miR (figure 4.45). With respect to relative abundance of the 5p product of each miR, miR-125b, miR-Let7c and miR-99a expression was comparable between NM and T21 MSC with the exception of one outlier in T21 MSC that showed relatively high expression of miR-Let7c-5p (figure 4.45). In contrast, the relative expression of miR-155-5p in T21 MSC was higher than in NM MSC (figure 4.45) and this positively correlates with the increased IL-6 expression however, the difference was not statistically significant.

Like most of the literature regarding miR mediated attenuation of IL-6 signalling, these data are correlative and do not show a direct link between miR-155 expression and IL-6 signalling. However, a recent study where miR-155 targets were directly mapped by comparing iCLIP (a method used to find protein-RNA interactions akin to ChIP-Seq) data from wild-type and miR-155 deficient mice elegantly demonstrated that miR-155 has a cell context dependent effect on gene regulation [Hsin et al., 2018]. Specifically, miR-155/RISC only bound to *Il6ra* mRNA in macrophages and although there was a decrease in IL-6 expression in miR-155 deficient macrophages this was not significant. Therefore, much work is still required to establish the mechanism for miR-155 mediated control of IL-6 or its associated up- and downstream signalling partners in the context of T21 and fetal B-lymphopoiesis.

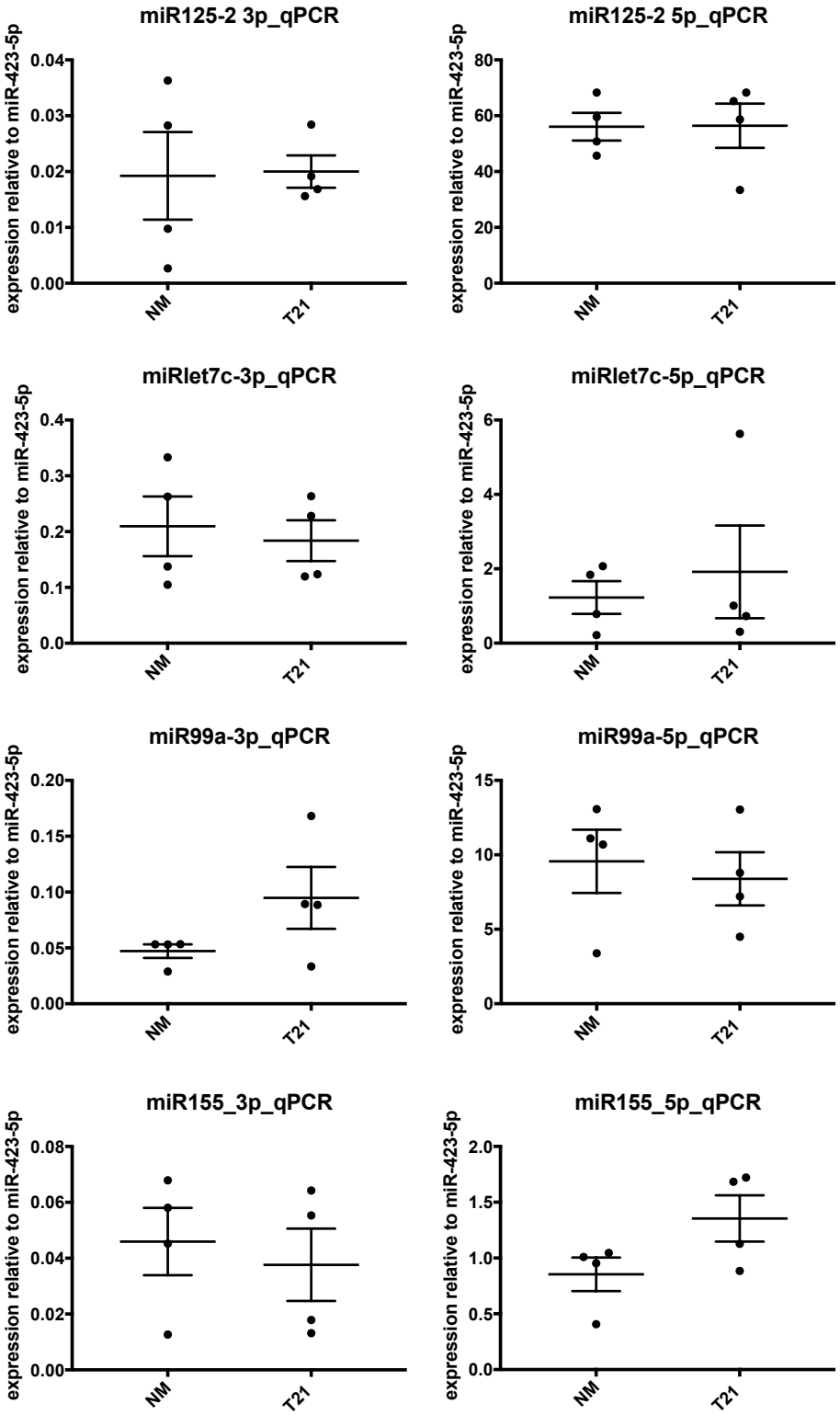


Figure 4.45: Expression of chromosome 21 miRNAs in MSC by RT-qPCR. miRNAs chosen for their known role in control of IL-6 expression or haematopoiesis. NM: n=4; T21: n=4.

4.5.5 The role of IFNα

The conflicting evidence regarding IL-6 expression in NM and T21 MSC led me to hypothesise that IL-6, being an acute phase protein and referred to as the "inflammation amplifier" [Murakami et al., 2019] is a result of a wider, systemic and dynamic inflammatory programme invoked by the T21 microenvironment. I hypothesised that it is this T21-driven inflammatory programme that leads to altered cross-talk between haematopoietic and microenvironment cells that ultimately causes the perturbations observed in T21 fetal haematopoiesis. The enrichment for the IFNα response in T21 LMPP (figure 4.23) and T21 MSC (figure 4.29); and the location of four interferon (IFN) receptor genes on chromosome 21 (*IFNAR1, IFNAR2, IFNGR2* and *IL10RB*) made IFN signalling an attractive candidate pathway worthy of further investigation.

IFNα response gene expression in HSPC: First, according to the RNA-Seq data, the expression of several genes involved in IFN signalling and response are over-expressed in T21 fetal HSPC and B cells. Many of these are significantly over-expressed in select T21 progenitor populations according to the pairwise DE analysis. For example, IFN signalling genes *IRF9, IFI27, EIF2AK2, GMPR, IFI44L, IL7, USP18* and *RSAD2* are all significantly over-expressed in T21 LMPP. In terms of receptor expression, none were significantly DE but *IFNAR1* did display a pattern of over expression consistent with what is expected due to increased gene dosage caused by an extra copy of chromosome 21.

Then, to establish whether the expression of IFN stimulated genes (ISGs) reflected the enrichment of expression of IFN signalling genes, I compared the significant DE gene lists generated by pairwise comparison of NM and T21 HSPC with published/ publicly available ISG lists (Metacore, Carvariate Analytics, [Schoggins et al., 2011] [Wu et al., 2018] [Mostafavi et al., 2016]). These lists confirm the close relationship between IFN signalling genes and IFN stimulated genes as it is very difficult to establish them in isolation because IFN signalling can act in a positive feed-forward loop to amplify an inflammatory response [Francois-Newton et al., 2011] [Michalska et al., 2018]. Nevertheless, these comparisons showed that even more ISGs are significantly over expressed in both T21 HSC and LMPP compared to NM HSC and LMPP (figures 4.47 and 4.48).

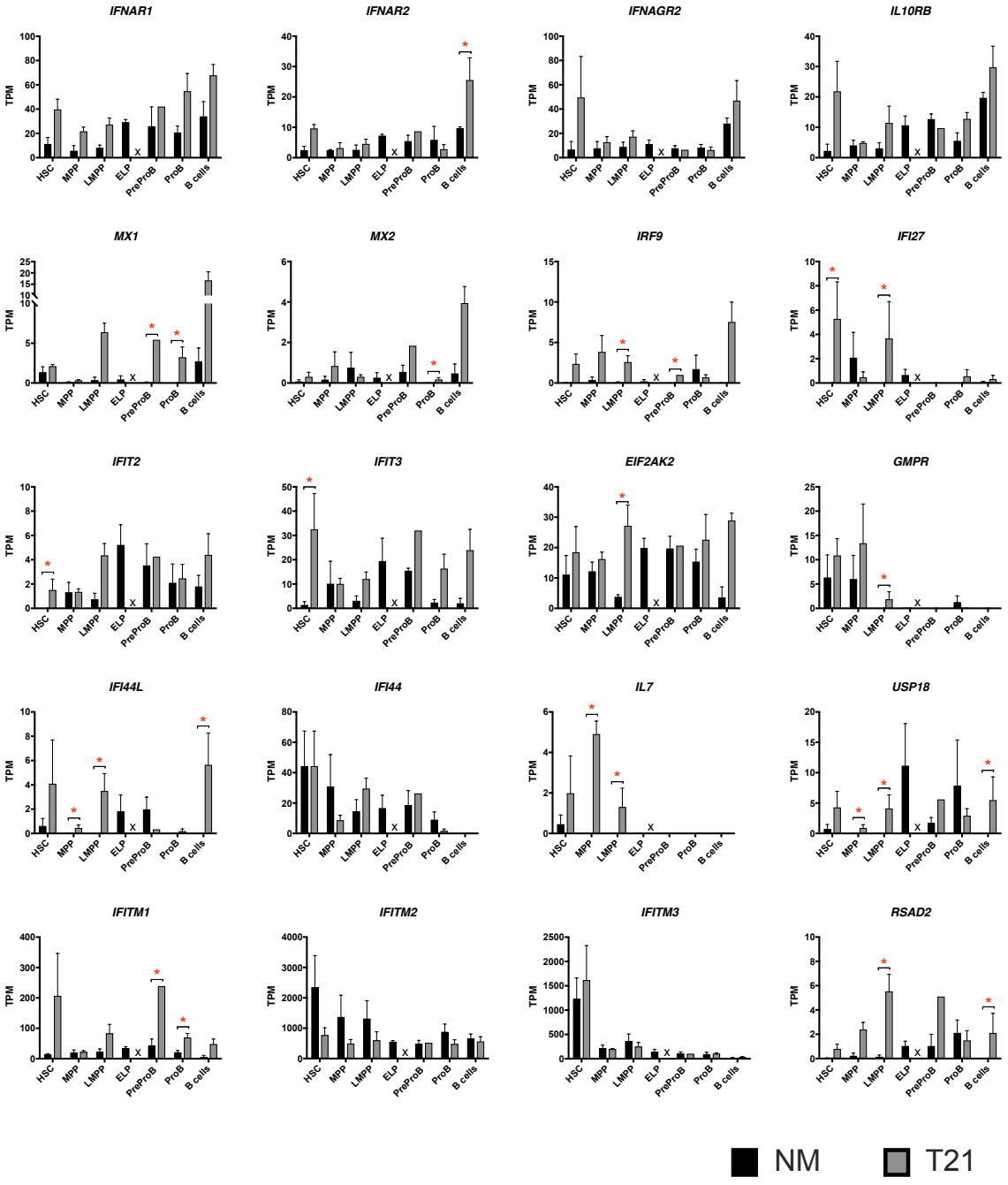


Figure 4.46: Expression of IFN receptors, signalling and response genes by RNA-Seq in NM and T21 fetal BM HSPC and B cells. NM: n=3; T21: n=2-5. $*$ denotes significance determined by DESeq2 with an FDR < 0.1.

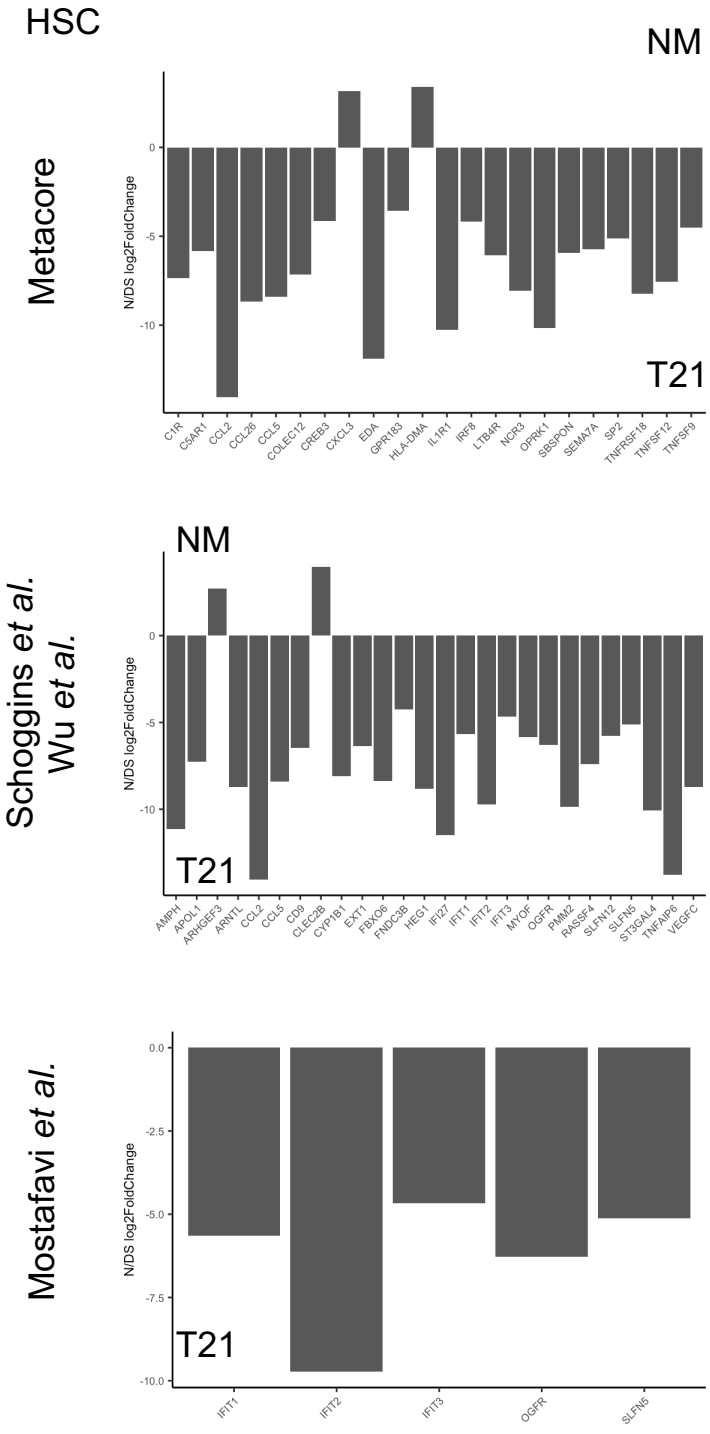


Figure 4.47: Interferon stimulated genes (ISGs) defined by Metacore, Carvariate Analytics, [Schoggins et al., 2011] [Wu et al., 2018] [Mostafavi et al., 2016]) are over-expressed in T21 HSC. A negative fold change is indicative of over-expression in T21 by RNA-Seq (100 cells/population). NM: n=3; T21: n=3.

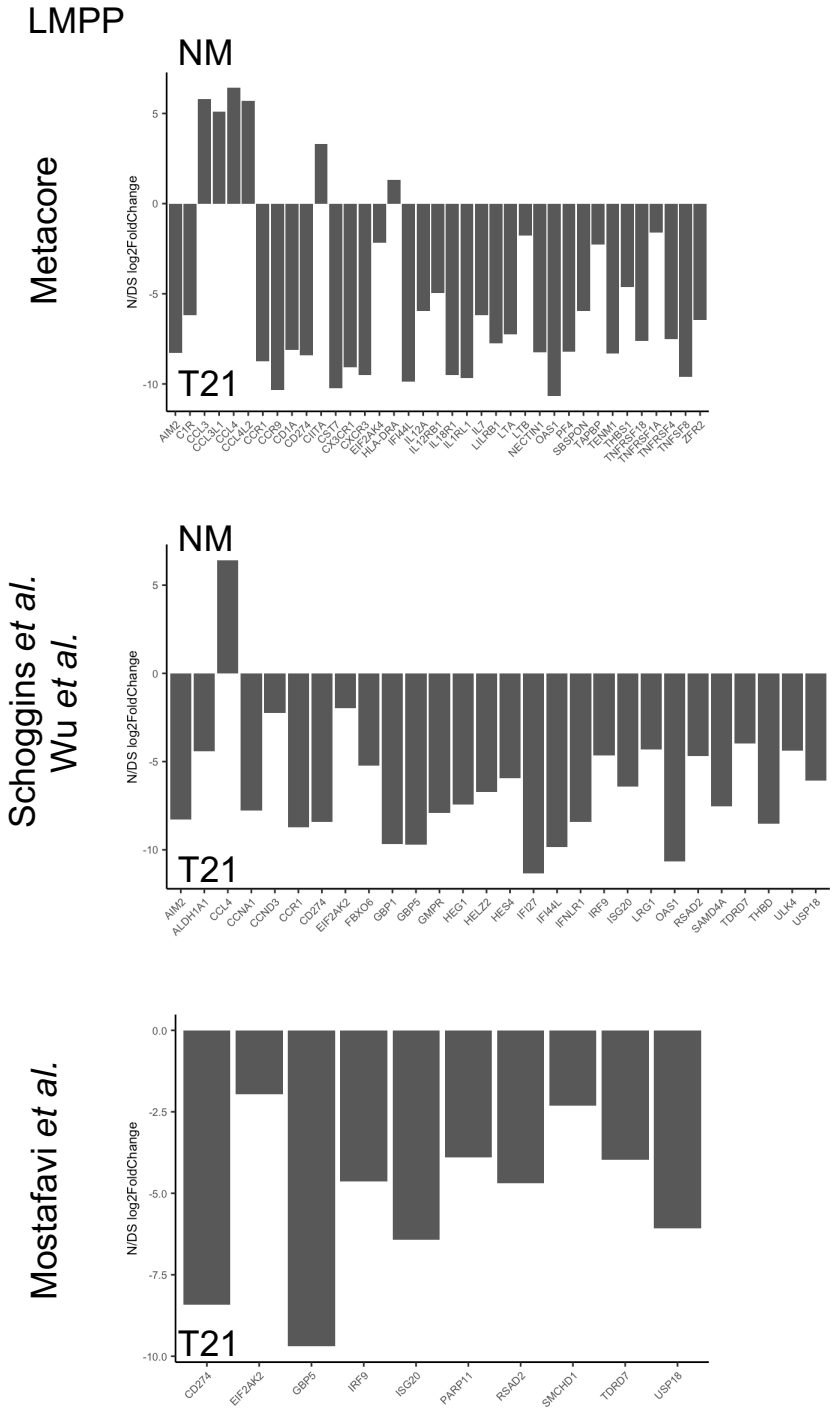


Figure 4.48: Interferon stimulated genes (ISGs) defined by Metacore, Carvariate Analytics, [Schoggins et al., 2011] [Wu et al., 2018] [Mostafavi et al., 2016]) are over-expressed in T21 LMPP. A negative fold change is indicative of over-expression in T21 by RNA-Seq (100 cells/population). NM: n=3; T21: n=3.

IFNα gene expression in MSC: To see whether the functional defect caused by T21 MSC in co-culture with NM HSC could be a result of an IFNα induced or dysregulated positive feed-forward loop where increased IFNα signalling results in the expression of more IFNα, RT-qPCR and ELISAs were performed to look for expression and secretion of IFNα. Initially I focussed on *IFNA2* only as this is the most characterised of all the 13 human subtypes of IFNα. However, since IFNα2 was undetectable by ELISA of MSC conditioned media (data not shown), I looked for the expression of *IFNA2* in parallel with *IFNA1, IFNA4, IFNA5, IFNA6, IFNA7, IFNA8, IFNA14, IFNA16* and *IFNA17* by RT-qPCR in NM and T21 MSC. Expression of all *IFNA* genes was generally very low but nevertheless 2/4 of the T21 MSC samples appeared to show higher expression of *IFNA1* when compared to NM, and expression of all other IFNA genes tested showed a similar relative level of expression with the exception of *IFNA16* and *IFNA17* which were undetectable. In parallel, I also validated what we had observed in the transcriptome analysis of the MSC: the IFNα receptor genes, *IFNAR1* and *IFNAR2*, were not significantly over-expressed although there was a trend towards higher expression in the T21 MSC (figure 4.49, bottom row).

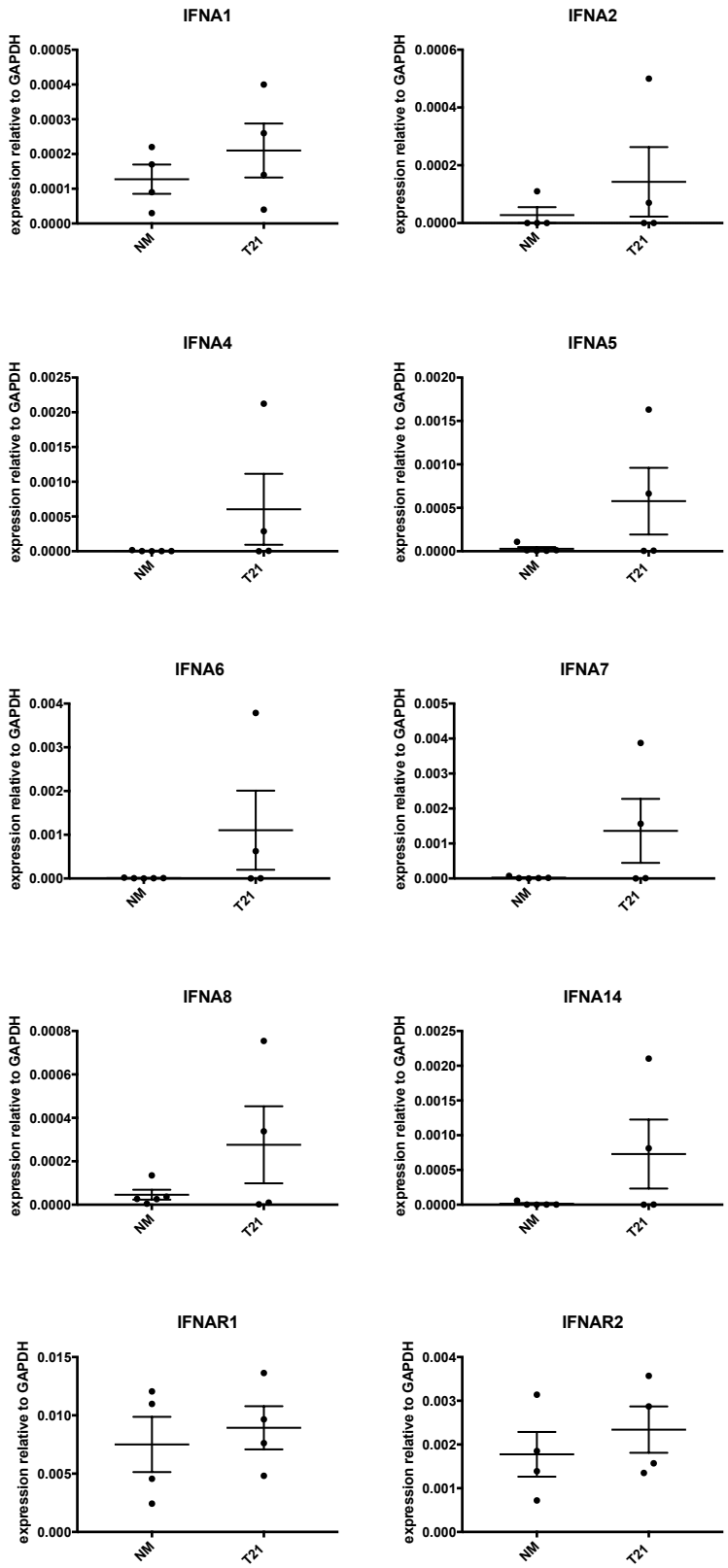


Figure 4.49: *IFNA* and *IFNAR* expression in NM (n=4) and T21 (n=4) fetal BM MSC by RT-qPCR

4.5.6 The role of other inflammatory signalling pathways

TGFβ signalling has a well documented and complex role in haematopoiesis and leukaemogenesis (reviewed here: [Blank and Karlsson, 2015] and [Naka and Hirao, 2017]). Since the TGFβ superfamily of cytokines consists of 33 different functional genes that are all related in homology but can elicit overlapping and distinct responses via different receptors [Derynck and Budi, 2019], deducing the role of each family member in isolation is complex. The role of TGFβ1 signalling is one of the most studied and is linked to the maintenance of HSC quiescence (when expressed at high levels); promoting myelopoiesis while inhibiting lymphopoiesis from HSC (when expressed at low levels) [Challen et al., 2010]; and plays a role in ageing HSC [Naka and Hirao, 2017]. In contrast, BMP-signalling (also part of the TGFβ superfamily) is associated with promotion of lymphopoiesis from HSCs [Naka and Hirao, 2017]. Since TGFβ signalling was also enriched in T21 HSC (figure 4.22), I looked at the expression of key genes involved in TGFβ1 signalling and BMP signalling in the RNA-Seq data. Then, to see if the functional difference observed in NM and T21 MSC was correlated with TGFβ expression, I used RT-qPCR to measure *TGFB1, TGFB2* and *TGFB3* expression in MSC.

Expression by RNA-Seq of *TGFB1, TGFB2* and *TGFB3* genes was increased in T21 HSC compared to NM with the difference being significant with respect to *TGFB2* and *TGFB3* (FDR <0.1). The two most common receptors responsible for TGFβ signal transduction (*TGFBR1* and *TGFBR2*) appeared to be expressed at comparable levels in NM and T21 HSPC and B cells with no significant differences (figure 4.50(a)). Expression of the genes that drove the TGFβ signalling enrichment showed that they were all expressed at low levels. Further to this, these four genes (*BCAR3, LTBP2, SKI* and *PMEPA1*) are all negative regulators of TGFβ signalling (figure 4.50(b)).

The fifth gene that was driving the TGFβ signalling enrichment was *BMP-2* (also significantly DE between NM and T21 HSC; figure 4.50(c)). Gene expression of the receptors (*BMPR1A, BMPR1B* and *BMPR2*) that transduce BMP-2 showed a pattern of over-expression in T21 HSC that was significant in the case of *BMPR1B*. This pattern continued in the expression of SMAD proteins known as common effectors (*SMAD1, SMAD5* and *SMAD9*) for BMP signalling (4.51(a)). As a comparison, the expression of SMAD genes associated with TGFβ signal transduction did not show such consistent patterns of over-expression in the T21 HSPC (figure 4.51(b)). Finally, expression of SMAD proteins that are common to both pathways (*SMAD4*

and *SMAD7*) showed that *SMAD4*, the role of which is to translocate to the nucleus and activate transcription, was relatively similarly expressed; and *SMAD7* known as a downstream repressor of TGFβ signalling was over-expressed in T21 HSC albeit to a not significant extent (figure 4.51(c)). While these data suggest a role for BMP signalling over TGFβ signalling in T21 B-lymphopoiesis, these gene expression patterns do not provide any indication as to whether these gene products are phosphorylated and therefore actively participating in the proposed signalling pathways. Extensive work is still required to investigate the relative roles of these pathways in T21 fetal B-lymphopoiesis.

Nevertheless, to see if the functional differences in NM and T21 MSC co-cultures could be a direct result of TGFβ stimulation, RT-qPCR was also performed for *TGFB1, TGFB2* and *TGFB3* on NM and T21 fetal BM MSC. While no differences in expression were significant, again probably due to insufficient biological replicates, *TGFB1* expression was lower in T21 MSC, *TGFB2* expression was marginally higher in T21 MSC and there was a spread in expression of *TGFB3* albeit at low levels (figure 4.52). Although the difference in pattern of expression of TGFβ is potentially interesting with respect to the defects observed in T21 fetal life, in the adult BM microenvironment, the main sources of TGFβ are megakaryocytes [Zhao et al., 2014] and non-myelinating Schwann cells [Yamazaki et al., 2011]. Given the bias towards megakaryo-erythropoiesis observed in T21 FL, it is tempting to speculate that the TGFβ signalling enrichment observed in T21 HSC (which is driven by negative regulators of TGFβ signalling) is in fact a response to the increased TGFβ signalling from increased numbers of megakaryocytes in the environment; and the contribution of TGFβ in the B-lymphoid defect being to push HSC towards myelopoiesis thus inhibiting B-lymphopoiesis as shown in the mouse model reported by Challen *et. al* [Challen et al., 2010].

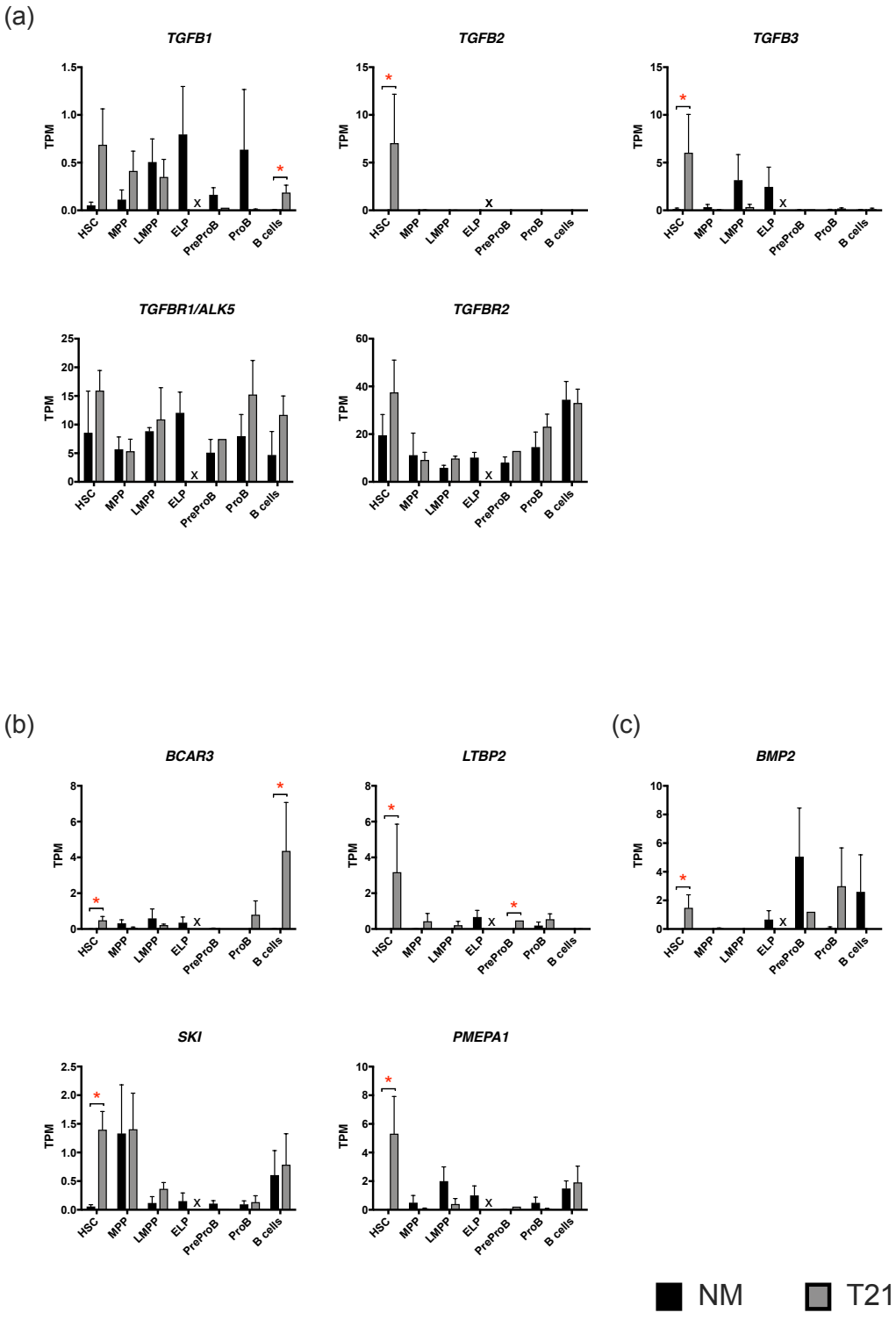


Figure 4.50: Expression of genes involved in TGFβ1 signalling by RNA-Seq of fetal BM HSPC and B cells. (a) *TGFB1/2/3* cytokine expression and commonly used TGFβ1 receptors (*TGFBR1* and *TGFBR2*). (b) negative regulators of TGFβ-signalling (*BCAR3, LTBP2, SKI* and *PMEPA1*) and (c) expression of *BMP2* that drove the enrichment observed in T21 HSC. $*$ denotes significance determined by DESeq2 with an FDR < 0.1. NM, n=3; T21 n=2-5.

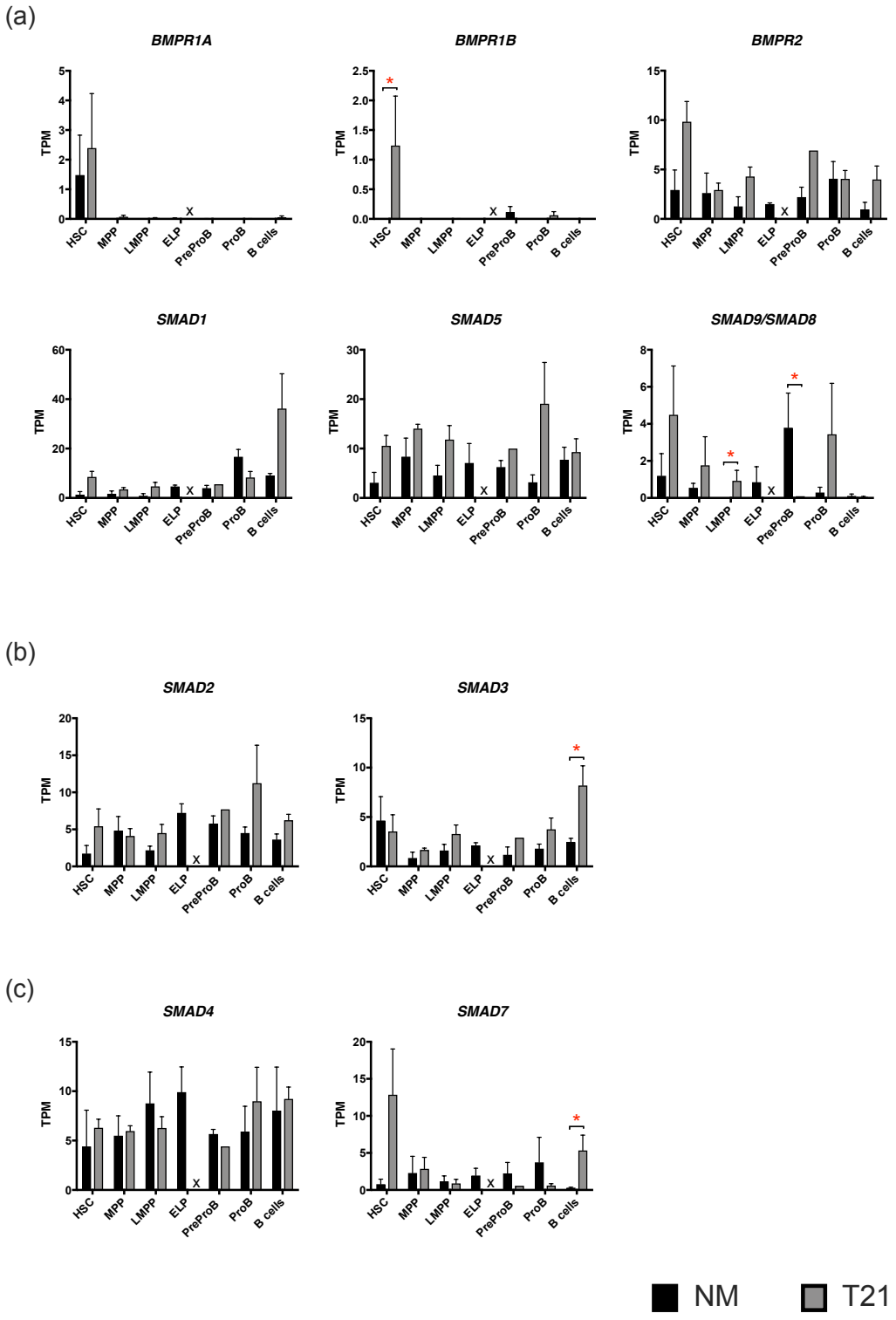


Figure 4.51: Expression of genes involved in BMP-signalling by RNA-Seq of fetal BM HSPC and B cells. (a) BMP receptor (*BMPR1A, BMPR1B* and *BMPR2*) and signalling effector (*SMAD1, SMAD5* and *SMAD9/8*) expression. (b) TGFβ1 signalling effector expression (*SMAD2* and *SMAD3*). (c) Shared BMP and TGFβ1 signalling effector expression (*SMAD4* and *SMAD7*. These data suggest BMP rather than TGFβ signalling is perturbed in T21 fetal B-lymphopoiesis. $*$ denotes significance determined by DESeq2 with an FDR < 0.1.NM, n=3; T21 n=2-5.

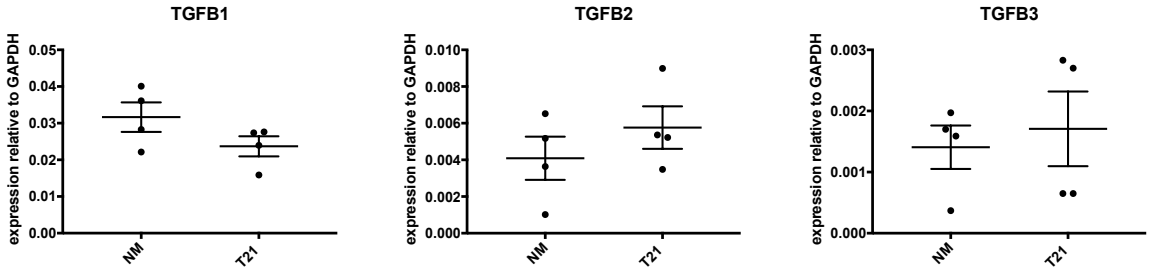


Figure 4.52: Expression of 3 of the TGFβ superfamily of cytokines measured by RT-qPCR in NM (n=4) and T21 (n=4) fetal BM MSC.

4.6 Summary and discussion

The results from this chapter show that T21 fetal B-lymphopoiesis is profoundly perturbed from the HSC stage to the ProB stage and these perturbations have a profound impact on B cell frequency. These data suggest a possible role for T21-driven pro-inflammatory changes in both the HSPC and associated stromal cells that could underpin or contribute to a bias in haematopoiesis that results in the defect in B-lymphopoiesis. Given the recent data about the role of delayed infection/chronic inflammation in the pathogenesis of childhood ALL [Greaves, 2018], a T21-driven over-active inflammatory haematopoietic environment could also play a role in the increased incidence of ALL in DS children (explored in Chapter 5).

4.6.1 T21 mediated perturbation of fetal B-lymphopoiesis

Prior to these experiments, little was known about extent of the B-lymphoid defect caused by T21 in human fetal life. Previous work on the FL concentrated on the impact of T21 on megakaryopoiesis, erythropoiesis and myelopoiesis [Tunstall-Pedoe et al., 2008] [Roy et al., 2012] [Chou et al., 2008] while providing preliminary evidence of the B-lymphoid defect [Roy et al., 2012]. Since data in Chapter 3 clearly demonstrate that the main site of B-lymphopoiesis is the fetal BM, I have begun to characterise T21 fetal BM for the first time. I have shown that concomitant with a significant increase in immunophenotypic HSC, there are significantly fewer (and in some cases no) ELP or PreProB progenitors in T21 fetal BM when compared to NM fetal BM. At the functional level, HSC, MPP and LMPP sorted from T21 fetal BM have compromised B-lymphoid output compared to normal HSPC in MS-5 co-cultures. This is in direct contrast with the comparable B-lymphoid output of rare T21 ELP and PreProB and abundant ProB progenitors co-cultured on MS-5 stroma.

Whole transcriptome analysis of sorted progenitor populations and comparison with normal counterparts showed no gross differences in key B-lymphopoiesis genes or overall B-lymphoid trajectories (according to PCA and diffusion maps). Instead, analysis of expression of megakaryocyte genes points towards a bias in priming of progenitors which is a concept increasingly employed to describe haematopoiesis[Enver and Greaves, 1998] [Notta et al., 2016][Carrelha et al., 2018] [Karamitros et al., 2018] [Laurenti and Göttgens, 2018]. In addition to this possible bias, thousands of protein coding genes were differentially expressed throughout the fetal B-lymphopoiesis hierarchy in T21 progenitors compared to NM, with genes being mainly over-expressed

in T21 populations. This, combined with GSEA showing a down regulation of gene sets associated with chromatin organisation suggests that there may be a global failure to silence gene expression in T21 HSPC. This is supported by murine studies showing that global amplification of gene expression is associated with T21 [Mowery et al., 2018]. Specifically, when the chromosome 21 gene *HMGN1* is over-expressed in this system, the B cell phenotypes (increase in H3K27 acetylation) observed in DS mouse models, including DS-ALL, is recapitulated [Lane et al., 2014] [Mowery et al., 2018]. Since the data presented here also show an increase in *HMGN1* expression in T21 HSPC (figure 4.15), this suggests that *HMGN1* plays a potentially important role in DS-ALL leukaemogenesis.

The whole transcriptome analysis by RNA-Seq of flow sorted fetal BM populations (100 cells/population) also pointed towards a role for the T21 microenvironment and specifically multiple inflammatory pathways: TGFβ signalling, IL-6 signalling, Inflammatory response, IL-2 signalling and IFNα response. This was supported by novel co-culture experiments which showed that T21 MSC could negatively affect the B-lymphoid potential of NM HSC. Attempting to deduce the roles that individual (or even families of) cytokines and the pathways they elicit in isolation is especially challenging given the similar and sometimes overlapping signalling pathways involved. Furthermore, dissecting the role of cell intrinsic mechanisms responsible for the reduced B-lymphoid output of T21 HSPC co-cultured on MS-5 stromal cells, cannot be separated from any autocrine or paracrine signalling that occurred *in vivo* during HSPC genesis (endothelial to haematopoietic transition) or differentiation prior to sorting and co-culture *in vitro*. Nevertheless, I attempted to rescue the B-lymphoid defect caused by T21 MSC in NM HSPC/T21 MSC co-cultures by supplementing co-cultures with IL-6. Unfortunately, this was insufficient to rescue the defect and while IL-6 initially seemed like the most promising candidate, its expression appeared to be largely sample dependent. Nevertheless, the difference in IL-6 expression between NM and T21 MSC suggested that perhaps the IL-6 signalling was part of a varied and complex response to other pro-inflammatory stimuli.

The IFNα response pathway is a particularly attractive candidate pathway because several genes involved in IFNα signalling including both IFNα receptors are encoded on chromosome 21 (*INFAR1, IFNAR2, IFNGR1, IL10RB, MX1* and *MX2*). Furthermore, there is the possibility that IL-6 expression is controlled by interferon regulatory factors (IRFs; [Luo and Zheng, 2016]). Others have also suggested a role for IFNα signalling in

DS disease pathology in post-natal life [Woo et al., 2013] [Sullivan et al., 2016] [Sullivan et al., 2017] with some suggesting that DS should be interpreted as an interferonopathy [Sullivan et al., 2017]. However, the role of IFNα signalling in T21 haematopoiesis and specifically in human fetal life remained unclear. The data presented here would suggest that increased expression of IFNα receptors *IFNAR1* and *IFNAR2* on T21 HSPC and MSC may confer increased sensitivity to IFNα signalling that leads to a pro-inflammatory positive feed-forward signalling cascades [Michalska et al., 2018]. Expression of several IFNα signalling and response genes display a pattern of over-expression particularly in T21 HSC and LMPP. However, it is not clear where IFNα itself is coming from. MSC express IFNα subtypes at very low levels (according to RT-qPCR) and there is no evidence of expression in NM or T21 HSPC (according to RNA-Seq). These findings do not preclude the possibility that interactions between HSPC and MSC result in the up-regulation of IFNα genes. This could in turn be responsible for the functional bias observed in NM HSPC/T21 MSC co-cultures, that resulted in excess myelopoiesis at the expense of B-lymphopoiesis.

In addition to the lack of B-progenitors, I also observed a significantly increased frequency of immunophenotypic HSC in T21 fetal BM. Increased inflammatory signalling and specifically increased IFNα signalling could be responsible for this difference as IFNα signalling is also associated with HSC exit from quiescence [Essers et al., 2009]. Although chronic exposure of IFNα signalling can cause HSC to re-enter quiescence to prevent exhaustion [Pietras et al., 2016], the combinatorial effects with other chromosome 21 genes, such as *DYRK1A* [Thompson et al., 2015], could play a key role in increased cycling and therefore frequency of HSC in T21 fetal BM. Alternatively, this observation could be explained by earlier events in fetal development as HSC emergence from the AGM also relies on sterile inflammation [Li et al., 2014][Ottersbach, 2019]. This work was initially described in mice where IFN signalling increased HSPC production in the absence of infection. Since then, this phenomenon has been observed by other groups in zebrafish [He et al., 2015] [Espin-Palazon et al., 2014] and recently the source of IFNα was shown to be primitive macrophages [Mariani et al., 2019].

There is also some evidence of anti-inflammatory control mechanisms. The negative regulator for IFNα signalling, *USP18* is up-regulated in T21 HSC, MPP and LMPP but this combined with only negative regulators of the anti-inflammatory TGFβ signalling pathway suggests that the pro-inflammatory signals are overwhelmingly predominant.

4.6.2 Inflammation and leukaemia

Recently Mel Greaves and colleagues postulated a role for delayed infection/inflammation in leukaemogenesis by causing DNA damage [Greaves, 2018]. Since observations showing an enrichment for inflammatory signalling in the leukaemia setting have also been reported by others [Hemminki et al., 2013] [Giustacchini et al., 2017] [Cuartero et al., 2018] [Forte et al., 2019], I hypothesise that T21-driven pro-inflammatory pathways are a key event in the initiation of DS-ALL that contributes to the increased incidence in children with DS. This is supported by others' work showing that pro-inflammatory pathways can persist in post-natal life in people with DS [Zampieri et al., 2014] [Sullivan et al., 2016] [Sullivan et al., 2017] and this forms the basis of the next chapter. Therefore, these data also potentially provide mechanistic insight into DS-ALL pathogenesis as well as immunodeficiency and autoimmune conditions also observed in children with DS [Goldacre et al., 2004] [Kusters et al., 2009]. Finally, the role of pro-inflammatory signals in a tumour microenvironment is likely to differ [Greten and Grivennikov, 2019] and so the role of T21-mediated pro-inflammatory signals could operate in a time and context dependent manner.

4.6.3 Future work

While these data show that B-lymphopoiesis is compromised in T21 fetal life, I have not established how ProB progenitor frequencies recover. As discussed in Chapter 3 it is possible that there are two different B-lymphoid pathways working in parallel in fetal BM where both eventually give rise to ProB progenitors. It is therefore tempting to speculate that in the absence of the fetal specific ELP pathway, the "adult type" pathway makes up for the deficit in T21 fetal BM. However, the explanation for the B cell defect in DS is likely to be more complex than absence of the ELP/PreProB progenitor pathway as the data show that T21 HSC/MPP/LMPP have reduced B-lymphoid output.

Overall, these data provide preliminary but promising evidence for the mechanism(s) responsible for T21 driven perturbations in fetal haematopoiesis. The data presented here suggest a role for T21 driven chromatin dysregulation and pro-inflammatory signalling in T21 fetal B-lymphopoiesis. However, much of the mechanistic insight thus far relies heavily on transcriptomic data and therefore extensive work is still required to establish the functional relevance. For example, increased expression of a gene does not always correlate with increased presence or activation of a protein in a particular

signalling cascade. Therefore experiments firstly validating the gene expression (single cell RT-qPCR) and then interrogating the protein activation (CyTOF/mass cytometry) of specific key effectors will be vital in establishing the role of the aforementioned pathways. Furthermore, the effects of these pathways requires further functional investigation which is a challenging prospect as it is likely that multiple pathways are acting in concert to perturb B-lymphopoiesis. As a starting point, cross-over co-culture assays to see if T21 B-lymphopoiesis can be rescued by NM MSC could be established.

Another area of future study is the MSC themselves. MSC are very loosely defined by their morphology and immunophenotype according to internationally agreed convention [Dominici et al., 2006]. This loose definition could explain the functional variation I observed in my co-cultures and is also a subject of recent debate in the field [Sipp et al., 2018]. Therefore, I think it would be really interesting to characterise these cells at a single cell level immediately after derivation, and then in later passage numbers to identify potential sub-populations and gauge how these cells change during cell culture. In addition to characterisation of MSC, it could also be useful to characterise the other cellular components of the BM stromal microenvironment such as chondrocytes, fibroblasts and osteoblasts which are all possibly contributing to the cytokine milieu *in vivo*, which undoubtedly has an impact on haematopoiesis as is evidenced by recent studies in mice [Baryawno et al., 2019] [Tikhonova et al., 2019].

Having established the extent of the B-lymphoid defect in T21 fetal BM and the possible role of inflammatory pathways, this sets the scene for exploring the potential causes of increased incidence of ALL in DS children. In the following chapter, I go on to investigate the immunophenotypic and molecular characteristics of ALL blasts from children with and without DS, using data from T21 and NM fetal ProB progenitors for comparison as a putative leukaemia initiating cell in these children.

Chapter 5

DS-ALL biology and implications for fetal B-lymphopoiesis

5.1 Background and aim

The evolution of ML-DS from a fetal progenitor in DS children is well characterised (reviewed in [Roberts and Izraeli, 2014] and [Bhatnagar et al., 2016]), but less is known about the origins of DS-ALL. DS-ALL is characterised by a paucity of cytogenetic aberrations such as ETV6-RUNX1 and high hyperdiploidy that are common in non-DS-ALL (also more commonly referred to as PreB-ALL or BCP-ALL) [Forestier et al., 2008] [Buitenkamp et al., 2014]. Instead, the most common genomic alteration in DS-ALL is a rearrangement affecting the *CRLF2* gene (CRLF2r) which is found in 60% of all DS-ALLs [Mullighan et al., 2009] [Russell et al., 2009] [Yoda et al., 2010] [Hertzberg et al., 2010] [Buitenkamp et al., 2014] [Schwartzman et al., 2017] and causes the over-expression of *CRLF2*. This contrasts with the relative rarity of these aberrations in non-DS-ALL where 10% of cases have CLRF2r [Mullighan et al., 2009] [Hertzberg et al., 2010] [Buitenkamp et al., 2014] [Schwartzman et al., 2017]. In addition to these studies, a very recent meta-analysis identified 4 heritable risk loci in DS-ALL in the *IKZF1, ARID5B, GATA3* and *CDKN2A* genes [Brown et al., 2019] which confirms earlier investigations implicating point mutations/SNPs in these genes [Schwartzman et al., 2017] [Vesely et al., 2017].

Despite these extensive investigations into the genetic aberrations common in DS-ALL; the *in utero* origins of multiple other subtypes of non-DS-ALL (ETV6-RUNX1, high hyperdiploidy, BCR-ABL1 and MLL-AF4; reviewed here: [Greaves, 2018]); and the high incidence of DS-ALL in young children (27-fold; 1-4yrs) [Hasle et al., 2016], the link between the B-lymphoid defect in T21 fetal life and DS-ALL leukaemogenesis

remains elusive. Further to this, acquired T21 (aT21) is the most common cytogenetic aberration in non-DS-ALL [Viguie, 2001] [Izraeli et al., 2007] [Huret et al., 2013] pointing towards a direct role for chromosome 21 in leukaemogenesis. Therefore, I set out to characterise the transcriptomes of cytogenetically matched non-DS-ALL and DS-ALL blasts for comparison with fetal progenitors. Specifically, I hypothesised that transcriptomic signatures conferred by an extra copy of chromosome 21 in blast cells would provide mechanistic insight into leukaemogenesis of DS-ALL. In order to investigate this, in addition to performing RNA-Seq on sorted cytogenetically matched non-DS and DS-ALL blasts, I performed RNA-Seq on non-DS-ALL blasts with aT21 and no other detectable cytogenetic abnormalities. Then, where possible, HSPC from non-DS and DS-ALL patient BM were sorted for RNA-Seq. Since it is clear that many non-DS-ALLs arise *in utero*; that DS-ALL occurs in young children; and that the immmunophenotype of DS-ALL blasts can closely resemble CD19+CD10+ ProB progenitors and B cells (figure 5.1), it is possible that T21 fetal BM ProB progenitors or B cells may be the fetal cell of origin for DS-ALL. Therefore, I began these investigations by analysing the RNA-Seq from NM and T21 fetal ProB progenitors. I also initially included B cells in these analyses as in fetal BM the majority of B cells are CD34-CD19+CD10+CD20+/- PreB cells as described in Chapter 3 and [Roy et al., 2017].

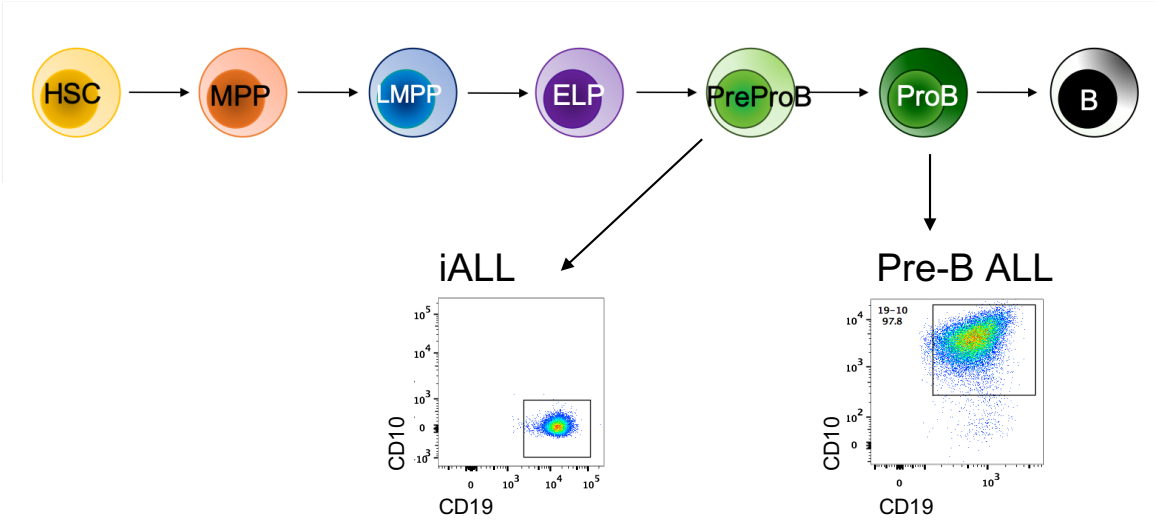


Figure 5.1: Schematic showing the relationship between putative fetal cell populations of origin for infant ALL (iALL) discussed in Chapter 3; and paediatric non-DS/DS-ALL (PreB-ALL).

5.2 Transcriptomic analysis of NM and T21 fetal BM ProB progenitors

Differential analysis of the RNA-Seq data from NM and T21 ProB progenitors and B cells (100 cells/population, n=3-5) by DESeq2 revealed that there were 1031 and 1073 significant (FDR <0.1) DE protein coding genes respectively (figure 5.2(a); complete DE tables available as online supplemental files). Since ProB progenitors are a putative cell of origin for paediatric ALL, I performed further analyses on the ProB transcriptomic data. Manual inspection of the significant DE genes showed a striking pattern of over-expression of genes associated with the cell cycle (*MKI67, ETS2, CENPF, CENPE, TOP2A, TTK, CDK6, DONSON, RAD21, CEP170, CENPA* and *BRCA2*) in T21 ProB progenitors compared to NM (figure 5.2(b)). GSEA for GO and Hallmark gene sets also demonstrated a striking enrichment for gene sets associated with the cell cycle (figure 5.3). These data could explain the relative sparing of ProB progenitors in the T21 fetal B-lymphopoiesis hierarchy where frequencies between NM and T21 fetal BM are not significantly different. Furthermore, if ProB progenitors are the cell of origin for paediatric ALL, the increased cell cycling inferred by these data and presumably as a result of recovery from the fetal B-lymphoid defect, could provide the permissive cellular context for an oncogenic hit in the initiation of leukaemia. To further investigate the putative link between fetal B-lymphopoiesis, the B-lymphoid defect and increased incidence of ALL in DS children, I performed immunophenotypic and transcriptomic analysis on ALL samples from the BM of non-DS and DS-ALL patients.

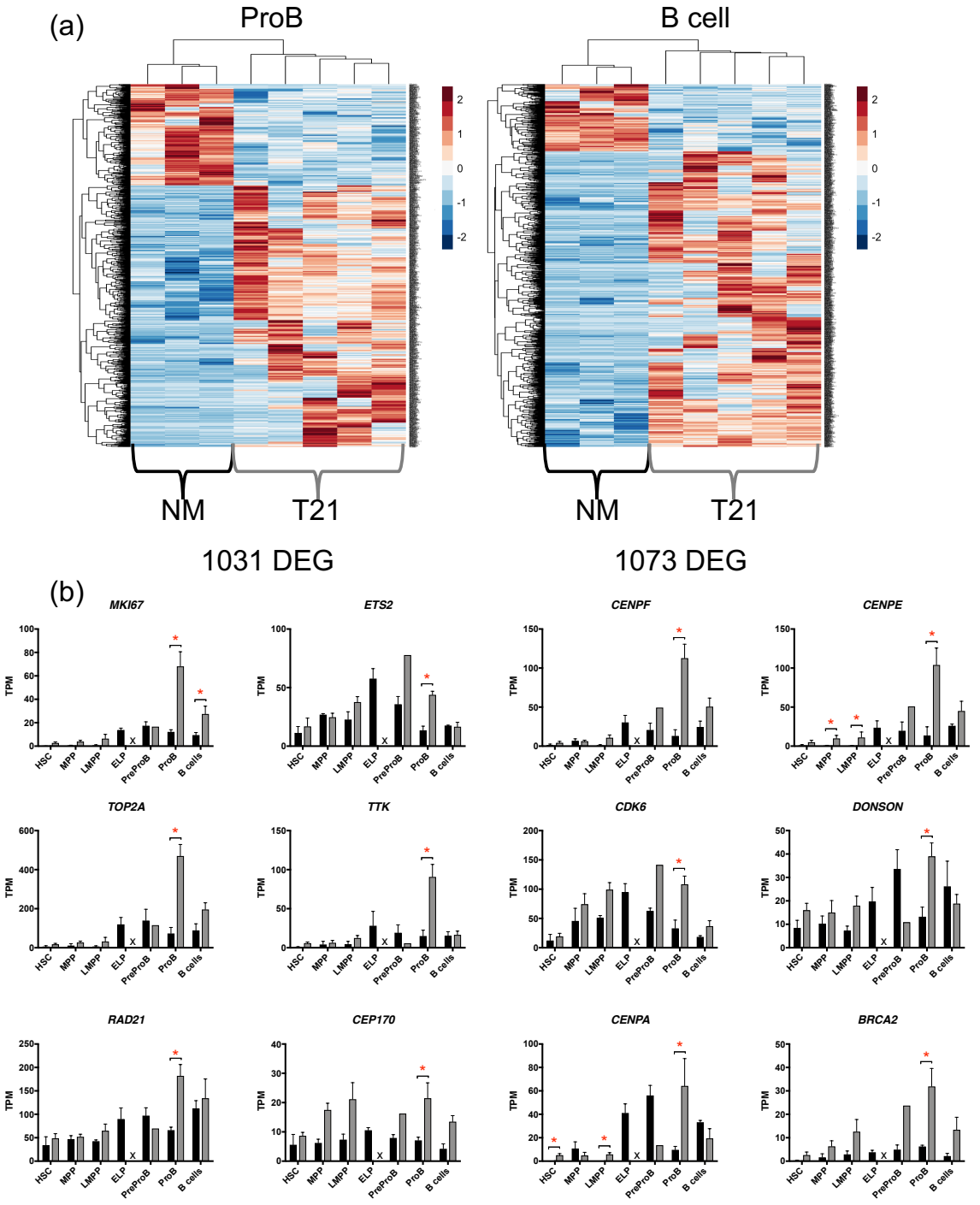


Figure 5.2: (a) Heatmaps showing 1031 and 1073 differentially expressed protein coding genes between NM and T21 fetal BM ProB and B cells respectively by RNA-Seq (100 cells/population). (b) Gene expression (by RNA-Seq; 100 cells/population) of genes selected for their involvement in the cell cycle and significant differential expression between NM and T21 ProB progenitors. For NM n=3 and for T21 n=5. DEG: differentially expressed gene. $*$ denotes significance determined by DESeq2; FDR <0.1.

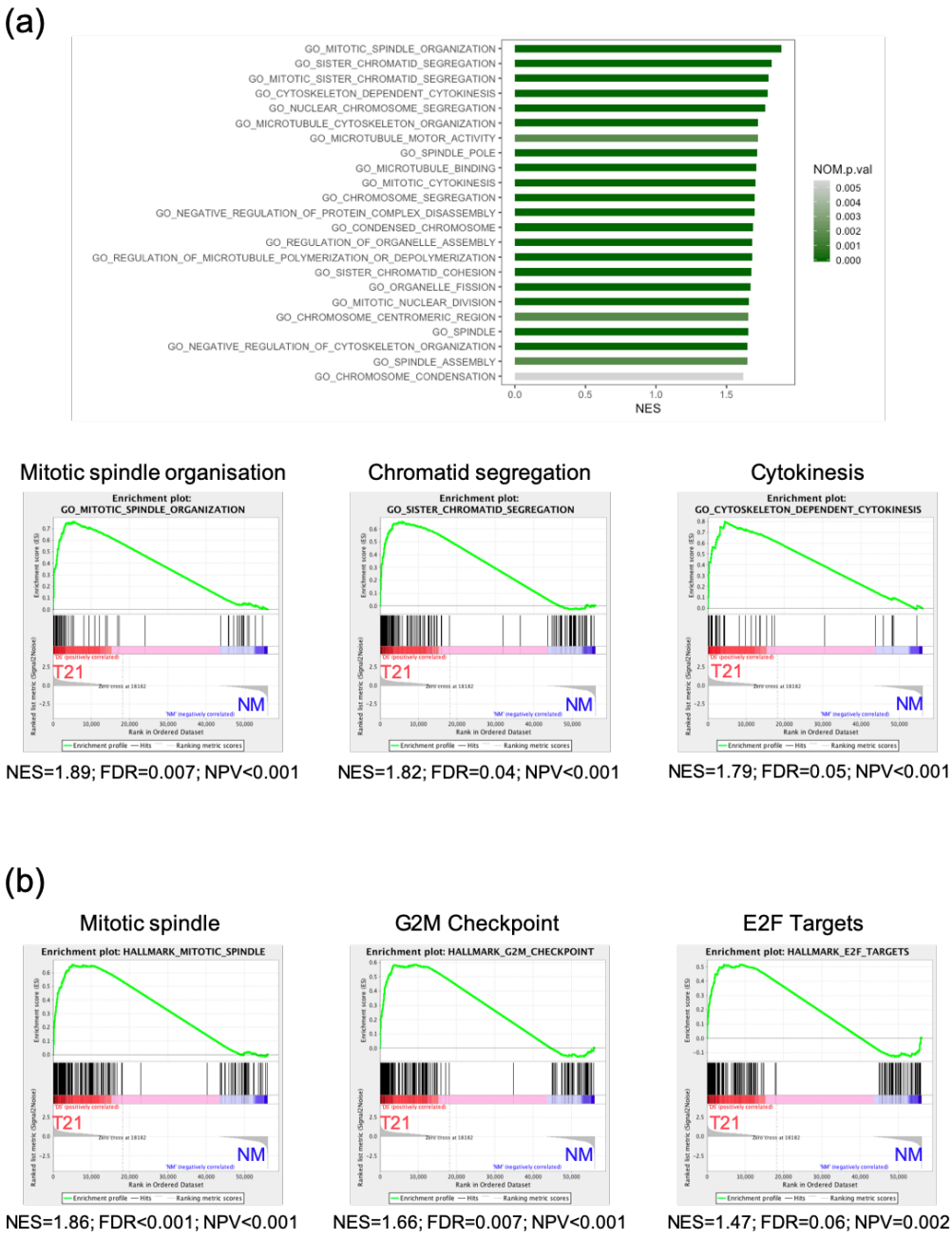


Figure 5.3: GSEA using (a) GO and (b) Hallmark gene sets show a significant enrichment for gene sets associated with cell cycle progression in T21 fetal BM ProB progenitors according to their transcriptome (by RNA-Seq; 100 cells/population. NM: n=3; T21: n=5). NES: normalised enrichment score; FDR: false discovery rate; NPV: nominal p-value)

5.3 Immunophenotypic characterisation of non-DS-ALL and DS-ALL BM

Cytogenetically-matched non-DS-ALL and DS-ALL blasts were stained and sorted for bulk RNA-Seq (100 cells/population). Cytogenetic groups available were ETV6-RUNX1 (good prognosis; common in non-DS-ALL), CRLF2r (commonest cytogenetic abnormality in DS-ALL), 'Other' (not ETV6-RUNX1, *IKZF1* deletion, CRLF2r or high hyperploidy) and acquired trisomy 21 (aT21; commonest cytogenetic abnormality in non-DS-ALL where an extra copy of chromosome 21 is found in blast cells only). A total of 20 patient BM samples were stained and sorted. At the same time, to enable comparison between datasets, 3 NM fetal BM samples were stained and immunophenotypic HSC, MPP, LMPP, ELP, CLP, PreProB, ProB and B cells were sorted for RNA-Seq. For these experiments, I designed and optimised new flow panels to allow the parallel sorting of ELP alongside HSC, MPP and LMPP and added further markers to a new leukaemia blast panel: CD20, CD133 and CRLF2/TSLPR. In collaboration with Sarah Inglott, Great Ormond Street Hospital, a sort strategy was designed to accurately sort patient blasts from cryopreserved BM (figure 5.4). Blasts, regardless of cytogenetic sub-group differed in their detailed immunophenotype from sample to sample. While all blasts were CD19+ and generally CD38lo and CD45lo, the expression of CD10, CD20 and CD34 was variable from sample to sample. In figure 5.5 I show the flow plots of each sample sorted in the CRLF2r subgroup as an example.

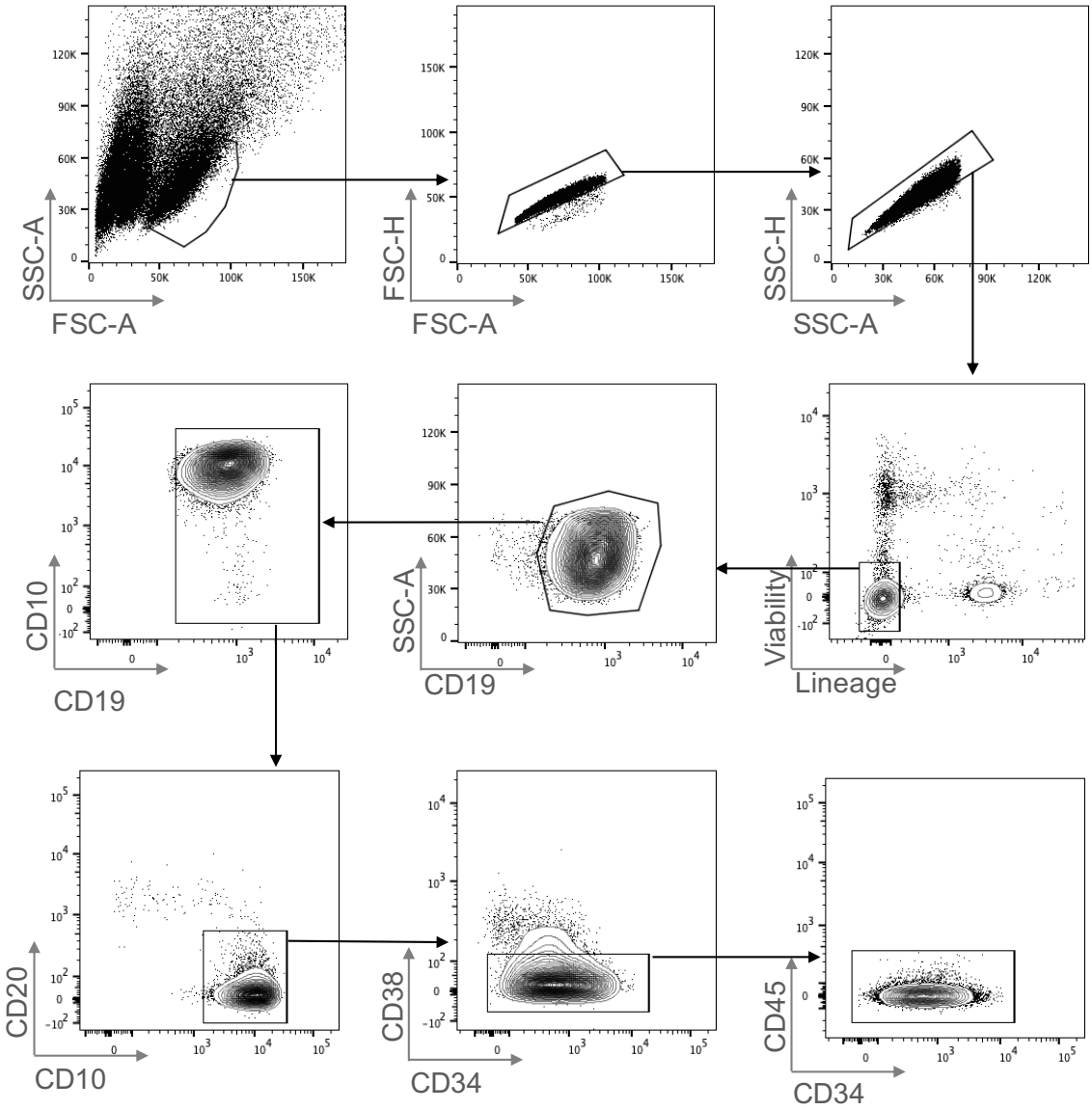


Figure 5.4: ALL blast flow gating strategy. Sample shown is a non-DS-ALL with CRLF2r. Lineage (Lin2): CD2, CD3, CD14, CD16, CD56 and CD235a.

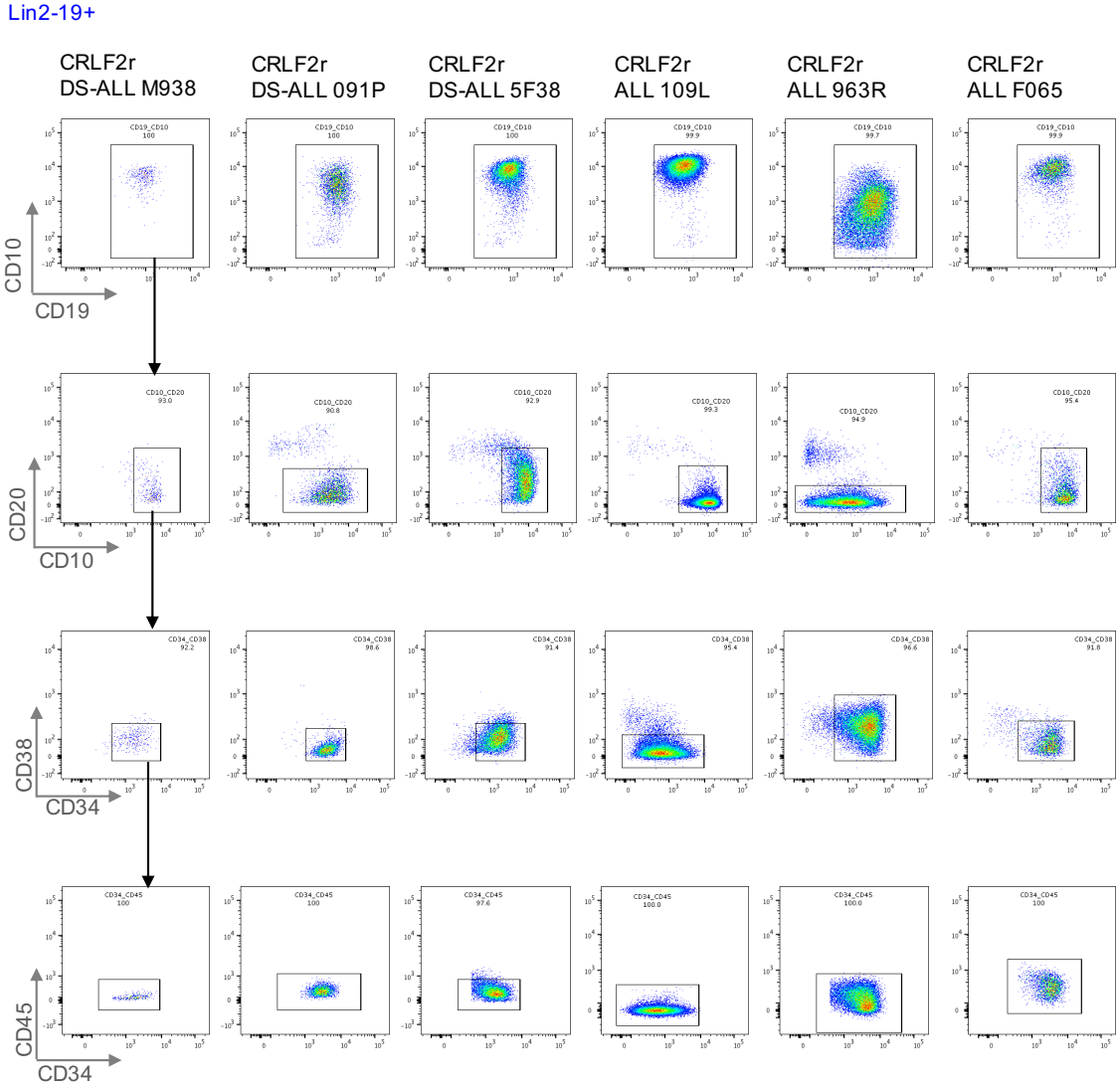


Figure 5.5: Flow plots from CRLF2r ALLs sorted for RNA-Seq. Columns are representative of one sample (labelled at the top) and rows are representative of the same gate. The top row is gated on Lin2-CD19+ cells. Lin2: CD2, CD3, CD14, CD16, CD56 and CD235a.

In addition to sorting patient blasts, I also performed flow cytometry analysis to look for HSPC in patient BM. To my surprise, the increased frequency of HSC in T21 fetal BM is preserved in DS-ALL patient BM and this is irrespective of cytogenetic subgroup (figure 5.6).

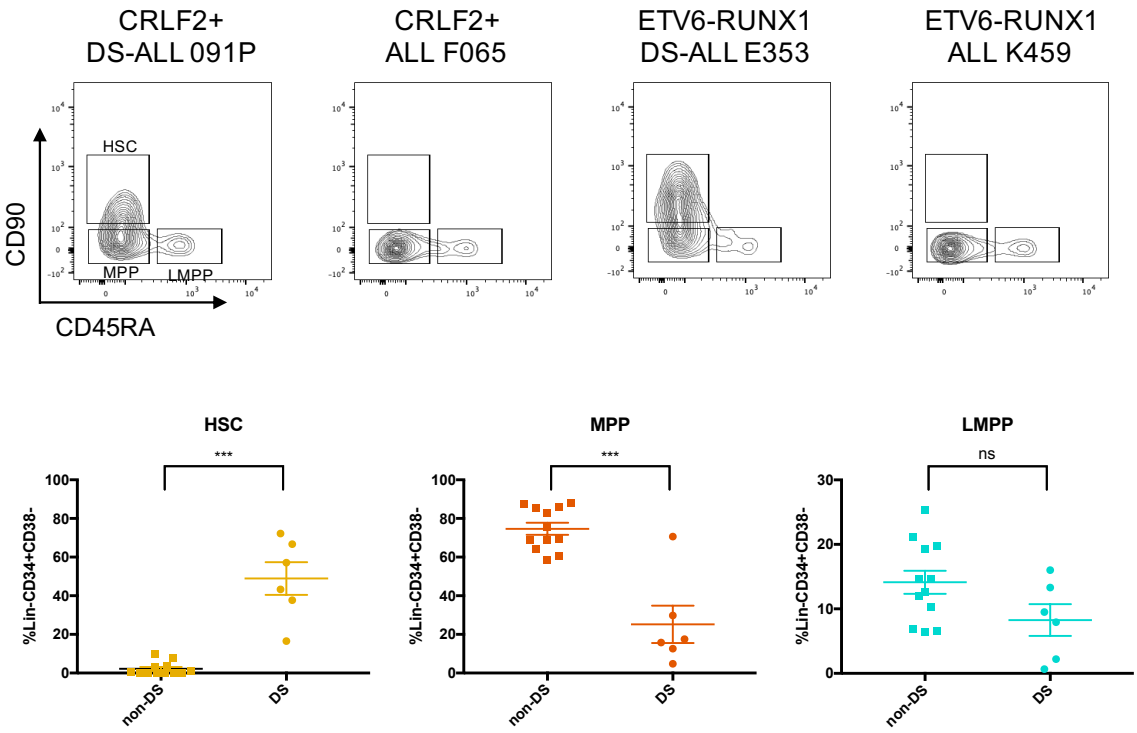


Figure 5.6: HSC, MPP and LMPP frequencies in non-DS-ALL and DS-ALL paediatric BM. The increased immunophenotypic HSC frequency observed in T21 fetal BM is preserved in the post-natal leukaemic setting. Top: representative flow plots showing that the significantly increased HSC frequencies are irrespective of cytogenetic status (gated on Lin-CD34+CD38-). Bottom: Frequencies of HSC, MPP and LMPP in non-DS-ALL and DS-ALL paediatric BM. Non-DS-ALL: n=12; DS-ALL: n=6. $***$ $p<0.0001$.

Since *ex vivo* expansion of patient blasts is notoriously challenging to establish [Pal et al., 2016], I performed a preliminary experiment to test the possibility that fetal BM MSC could support blast expansion. Blasts were sorted (when spare sample was available) for co-culture on NM fetal BM MSC with and without cytokines (Flt3L, SCF, IL7 and IL2) at a density of 500 or 1000 blasts/well. Co-cultures were monitored by flow cytometry at 7 day intervals. Out of a total of 9 biological samples, 3 samples survived in co-culture to day 14. While no expansion was observed (generally approximately a maximum of 100 cells were observed in wells originally plated with 500 blasts), there were some interesting observations in changes in the immunophenotype of the blasts.

Sample 398Q: sample 398Q (a DS-ALL of 'other' cytogenetic status) had the immunophenotype CD19+CD10+/-CD20-CD34+CD45lo. Since there appeared to be CD10+ and CD10- populations of blasts in this sample, CD10+ and CD10- blasts were co-cultured separately and in parallel to a co-culture containing both CD10+ and CD10- blasts that I called the bulk population. After 14 days in culture, all blasts had lost the expression of CD34 on their surface and while CD10- blasts remained CD10-, the CD10+ blasts appeared to lose their expression of CD10 (figure 5.7).

Sample 742R: sample 742R (a DS-ALL of CRLF2r cytogenetic status) had the immunophenotype CD19+CD10+CD20+/-CD34+/-CD45lo when sorted for co-culture. By day 14, in contrast to sample 398Q, the blasts generally maintained this CD19+CD10+CD34- immunophenotype (figure 5.7).

Sample 418I: sample 418I (a non-DS-ALL of 'other' cytogenetic status) had the immunophenotype CD19+CD10-CD20-CD34+CD45lo. By D14, the blast population was small but had maintained the same immunophenotype: CD19+CD10-CD34+. In addition, there was also a pronounced myeloid expansion that could either be from residual healthy BM cells with myeloid potential or, less likely, from de-differentiation of ALL blasts (figure 5.7).

Taken together, these data, although preliminary, suggest that fetal BM MSC might be a suitable alternative to murine stromal cell lines like MS-5 that are optimal for forcing B-lymphopoiesis rather than maintaining populations. It is particularly interesting that in these co-cultures 2 of the 3 samples that survived showed that the CD10- immunophenotype of blasts was maintained because in my experience of MS-5 co-cultures, expression of CD10 can be detected very soon after culture initiation regardless of population plated.

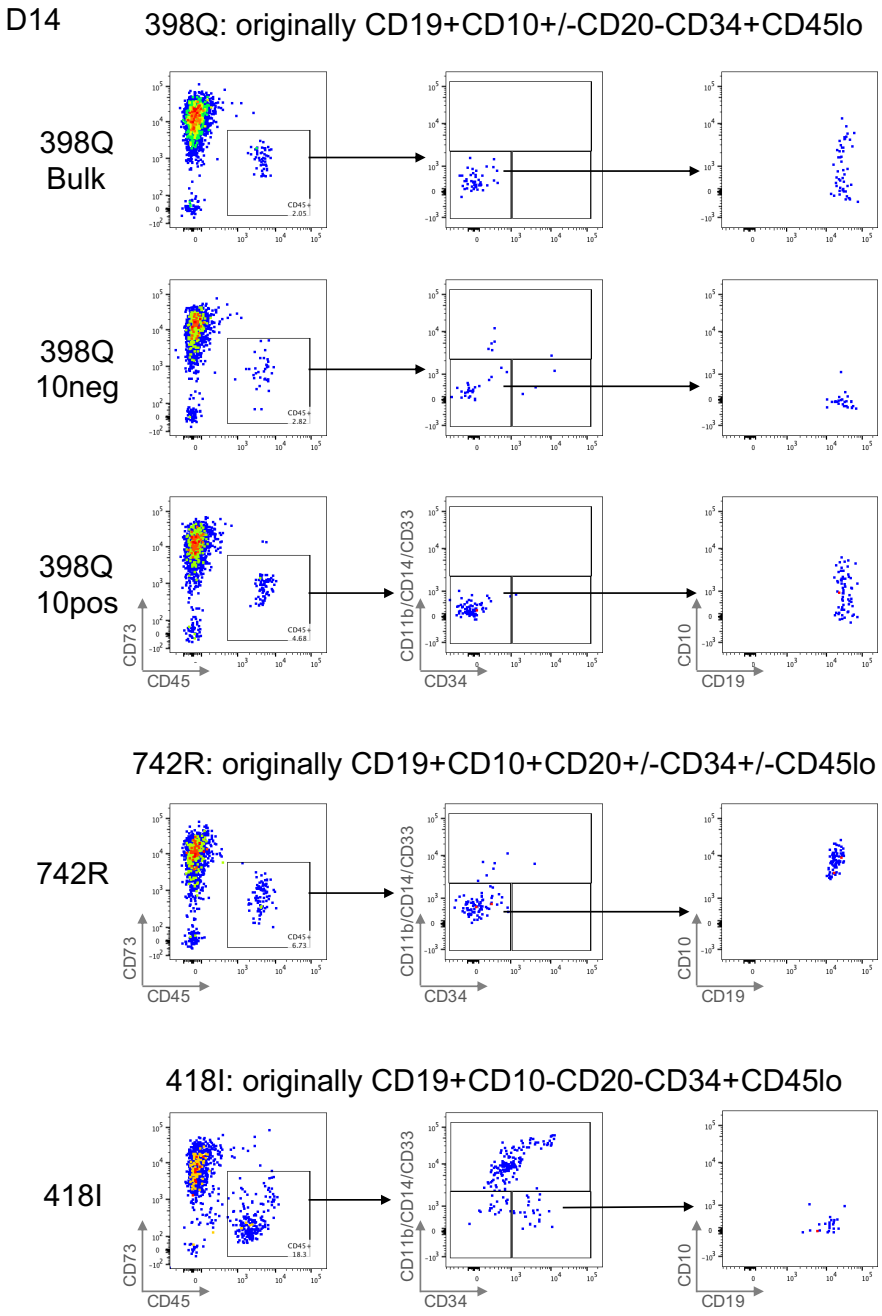


Figure 5.7: Flow plots from day 14 analysis of co-culture of 500 sorted ALL blasts. 3/9 samples plated on NM fetal BM MSC had live CD45+ cells *in vitro* by day 14 of culture with Flt3L, SCF, IL7 and IL2. Samples 398Q and 742R are DS-ALL blasts and sample 481I is a non-DS-ALL sample. Note the persistence of the CD10- blast cells in samples 398Q and 481I.

5.4 Transcriptomic analysis of cytogenetically-matched ALL blast cells by RNA-Sequencing

Bulk RNA-Seq (100 cells/population) was performed on cytogenetically matched flow-sorted blasts and where possible HSPC. A summary of the populations sorted and sequenced successfully is detailed in table 5.1 (Chapter 2-Methods). Figure 5.8 illustrates the comparisons made and discussed in the following sections of this chapter.

Table 5.1: Summary of libraries sequenced for RNA-Seq of non-DS/DS-ALL RNA-Seq experiment

	Blasts	HSC	MPP	LMPP	ELP	CLP	PreProB	ProB	B cell
non-DS ETV6-RUNX1	4	1	4	2				3	4
DS ETV6-RUNX1	2		1	2				1	2
non-DS CRLF2+	3	1	3	3		2			3
DS CRLF2+	4	3	3	4			3	4	2
non-DS OTHER	3		3	3			1	1	3
DS OTHER	1		1	1			1	1	1
non-DS acquired T21	3		2	2			2	3	2
non-DS fetal BM		3	2	3	3	3	2	3	3

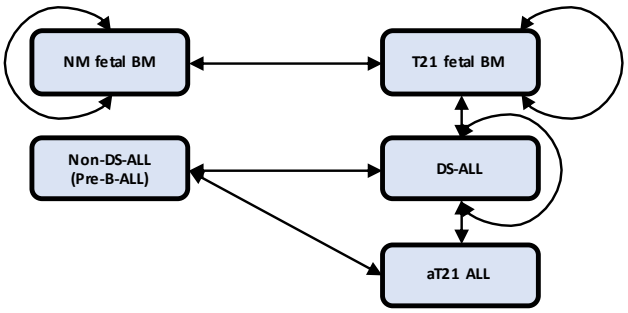


Figure 5.8: Comparisons made between the dataset discussed here (DS-ALL, non-DS-ALL and aT21-ALL) and the previous RNA-Seq dataset (NM fetal BM and T21 fetal BM; Chapter 3 and Chapter 4). NM fetal BM n=3; T21 fetal BM n=2-5; non-DS-ALL n=10; DS-ALL n=7; aT21 ALL n=3.

5.4.1 Differential analysis of ALL blast populations

As a first step in the comparison of the transcriptomes of non-DS-ALL and DS-ALL blasts, I performed PCA of the top 1000 and top 500 most variably expressed genes (figure 5.9). These analyses show that the gene expression profile of blast populations is largely shaped by cytogenetic subgroup with ETV6-RUNX1+ blasts separating by PC1 and CRLF2r/aT21 blasts separating by PC2 (figure 5.9). Since the 'Other' cytogenetic classification is likely to be a very heterogenous group, I also ran the analysis with these data excluded but this did not drastically change the pattern of clustering (figure 5.9 bottom row).

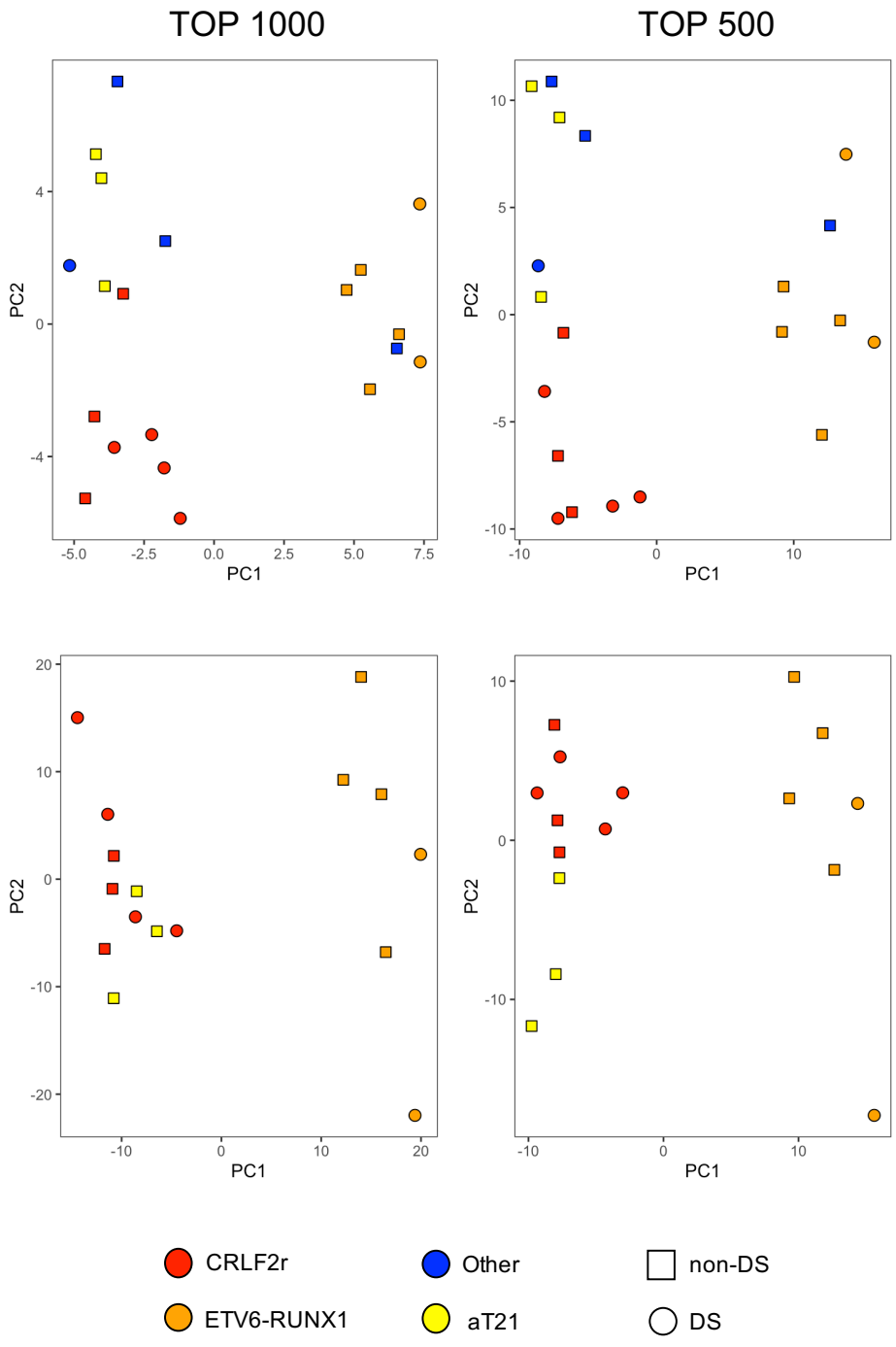
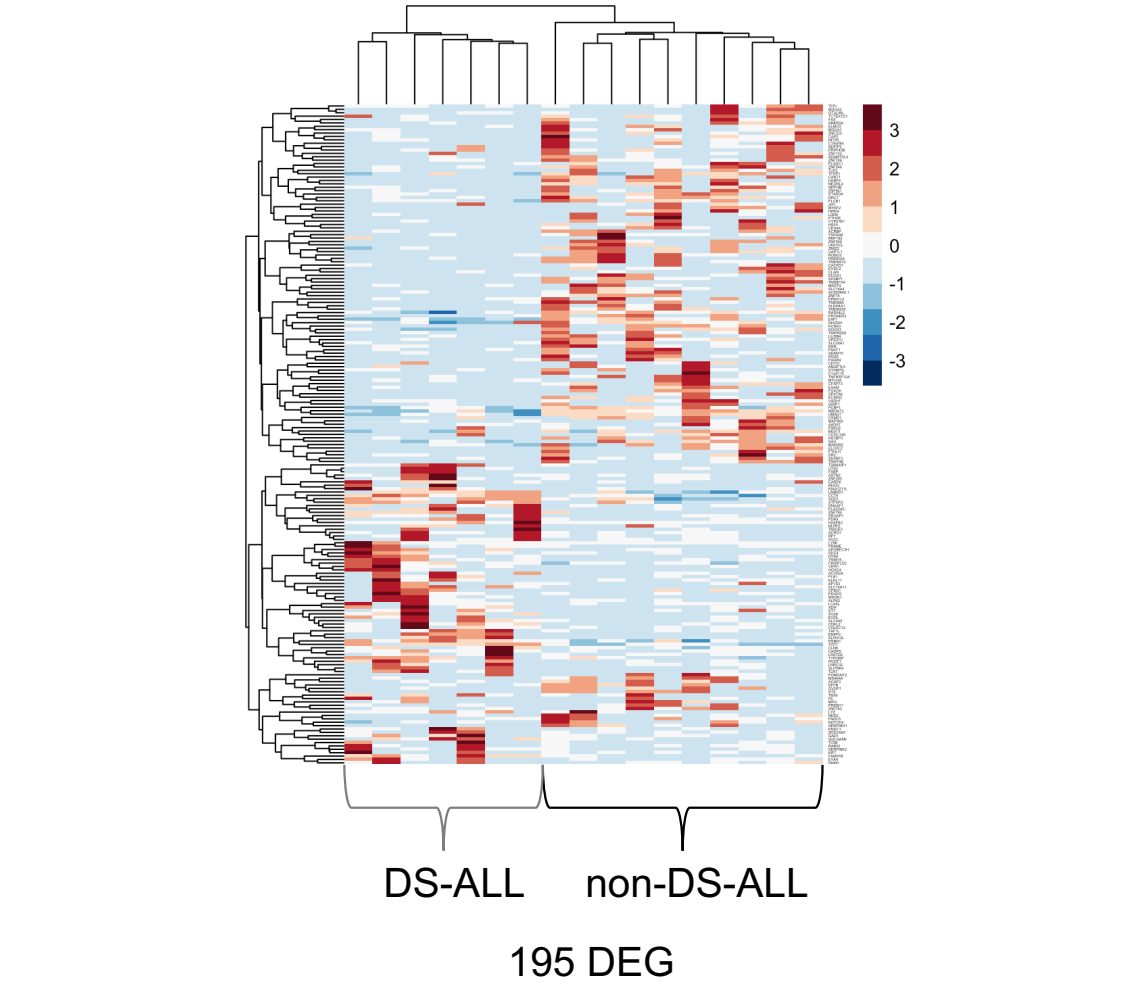


Figure 5.9: PCA of the top 1000 (left) or top 500 (right) most variably expressed genes by RNA-Seq (100 cells/population) of non-DS-ALL and DS-ALL blasts. Samples are coloured according to cytogenetic subgroup with the 'Other' subgroup included in top plots and excluded in bottom plots. Non-DS-ALL: n=13; DS-ALL: n=7.

Next, to look for differences in non-DS-ALL and DS-ALL, I performed differential expression analysis. Since the PCA had already demonstrated that most differences are between cytogenetic subgroups it perhaps is not surprising that there were relatively few DE protein coding genes between non-DS-ALL and DS-ALL blasts (195 DE protein coding genes, FDR <0.1, full differential table available as online supplemental file)(figure 5.10). These analyses also demonstrate the previously observed heterogeneity in DS-ALL [Hertzberg et al., 2010] [Buitenkamp et al., 2014]. Nevertheless, GSEA showed an enrichment for inflammatory pathways in DS-ALL blasts compared to non-DS-ALL blasts (figure 5.10). This has not previously been described and is interesting given that I observed similar gene set enrichments in fetal BM T21 HSPC compared to NM suggesting that such gene expression programmes are either preserved in postnatal life or re-activated in the DS leukaemic setting.

To ascertain whether differences caused by T21 were being masked by the transcriptional programme dictated by cytogenetic group, I then ran differential expression analysis on the most common cytogenetic subgroup in DS-ALL, CRLF2r. No hallmark gene sets were significantly enriched in CRLF2r DS-ALL compared to non-DS-ALL figure 5.11) and there were no significant DE chromosome 21 genes with a known role in haematopoiesis.



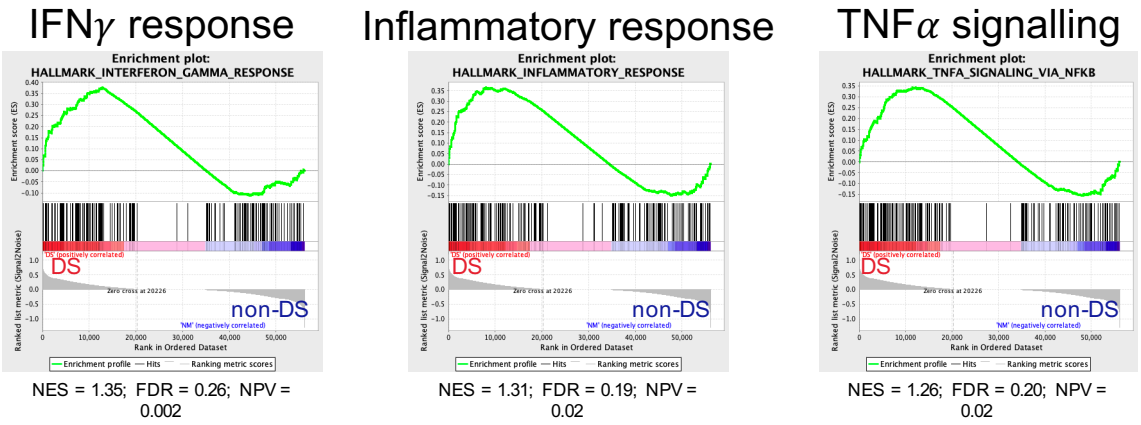


Figure 5.10: Top: Heatmap showing all significant (FDR <0.1 determined by DESeq2) DE protein coding genes by RNA-Seq between non-DS-ALL and DS-ALL blasts. Bottom: GSEA using hallmark gene sets showing enrichment for pro-inflammatory IFNγ, Inflammatory response and TNFα signalling pathways. Non-DS-ALL: n=10; DS-ALL: n=7. DEG: differentially expressed genes; NES: normalised enrichment score; FDR: false discovery rate; NPV: nominal p-value.

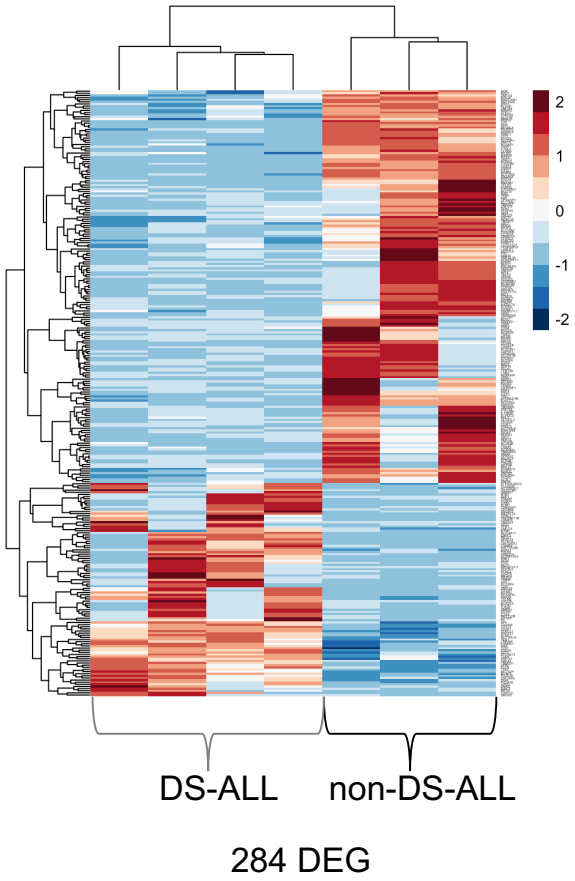


Figure 5.11: Heatmap showing all significant (FDR <0.1 determined by DESeq2) DE protein coding genes by RNA-Seq between CLRF2r non-DS-ALL and DS-ALL blasts. Non-DS-ALL: n=3; DS-ALL: n=4. DEG; differentially expressed genes.

I hypothesised that the transcriptional signatures conferred by an extra copy of chromosome 21 in blast cells would provide mechanistic insight into the leukaemogenesis of DS-ALL. Since the acquisition of an extra copy of chromosome 21 in blast cells is the most common genetic aberration in non-DS-ALL, I compared the transcriptomes of aT21 blasts with all other non-DS-ALL blasts and DS-ALL blasts. When aT21 blasts were compared to non-DS-ALL blasts and DS-ALL blasts there were 847 and 780 DE protein coding genes respectively (figure 5.12(a); full differential tables available as online supplemental files). GSEA between aT21 and DS-ALL blasts showed a strong enrichment for interferon signalling pathways (figure 5.12(b)) which is reflected in the increased pattern of expression of selected IFN genes in DS-ALL blasts (figure 5.13). This is very interesting because it shows that the common inflammatory signature seen irrespective of cytogenetic subgroup in DS-ALL and in T21 fetal BM haematopoiesis, is absent in aT21 ALL; pointing towards a role for the T21 microenvironment in leukaemogenesis and/or maintenance of leukaemia blasts. In addition to this observation I also noted that there were very few (64 different genes in each case) DE genes that were commonly DE in non-DS-ALL vs DS-ALL and the comparisons with aT21 (figure 5.14)

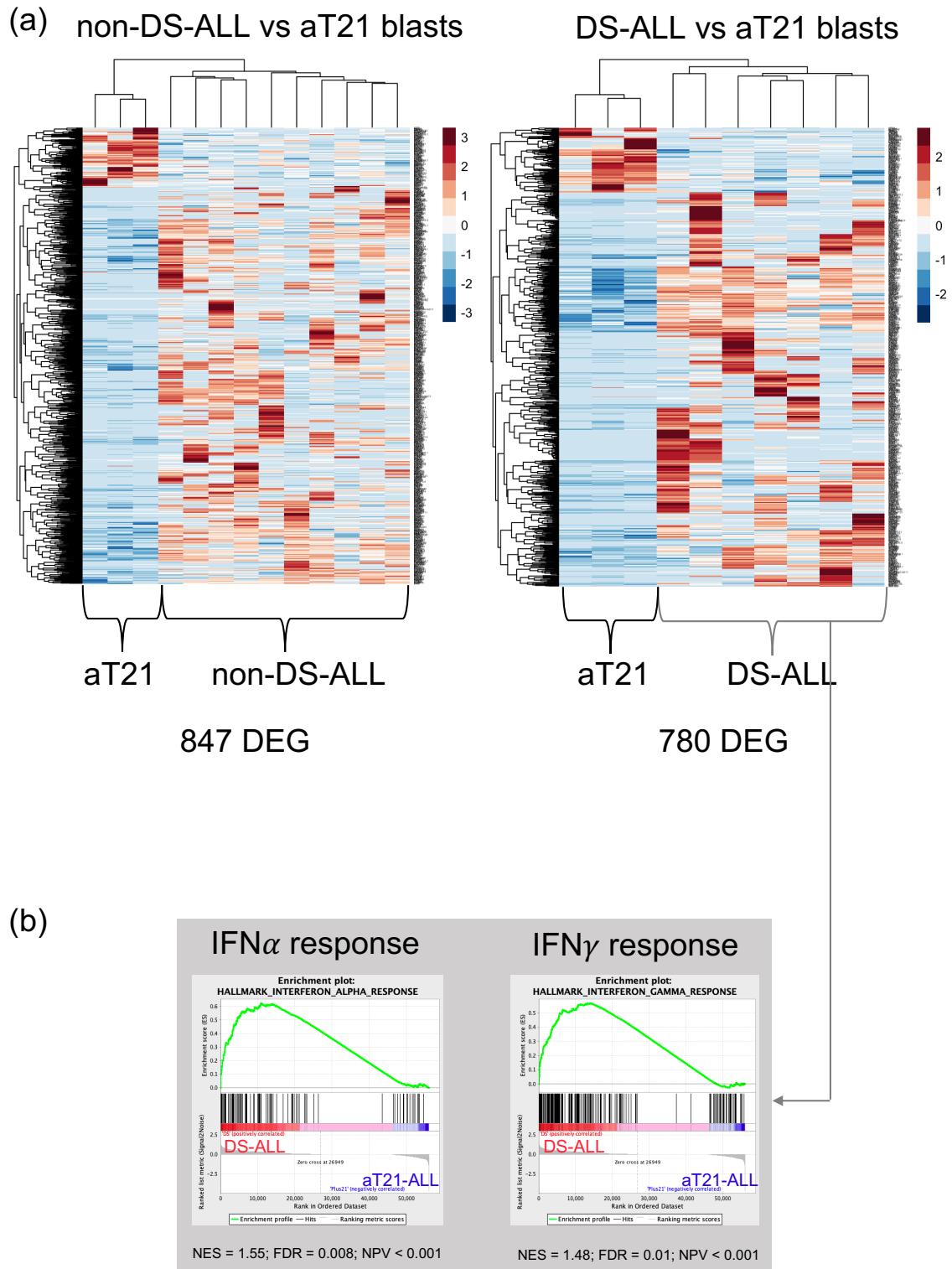


Figure 5.12: Heatmaps showing all significant (FDR <0.1 determined by DESeq2) DE protein coding genes by RNA-Seq between non-DS-ALL (left) or DS-ALL (right) blasts and blasts with aT21. Non-DS-ALL: n=10; DS-ALL: n=7; aT21-ALL n=3. DEG: differentially expressed genes; NES: normalised enrichment score; FDR: false discovery rate; NPV: nominal p-value.

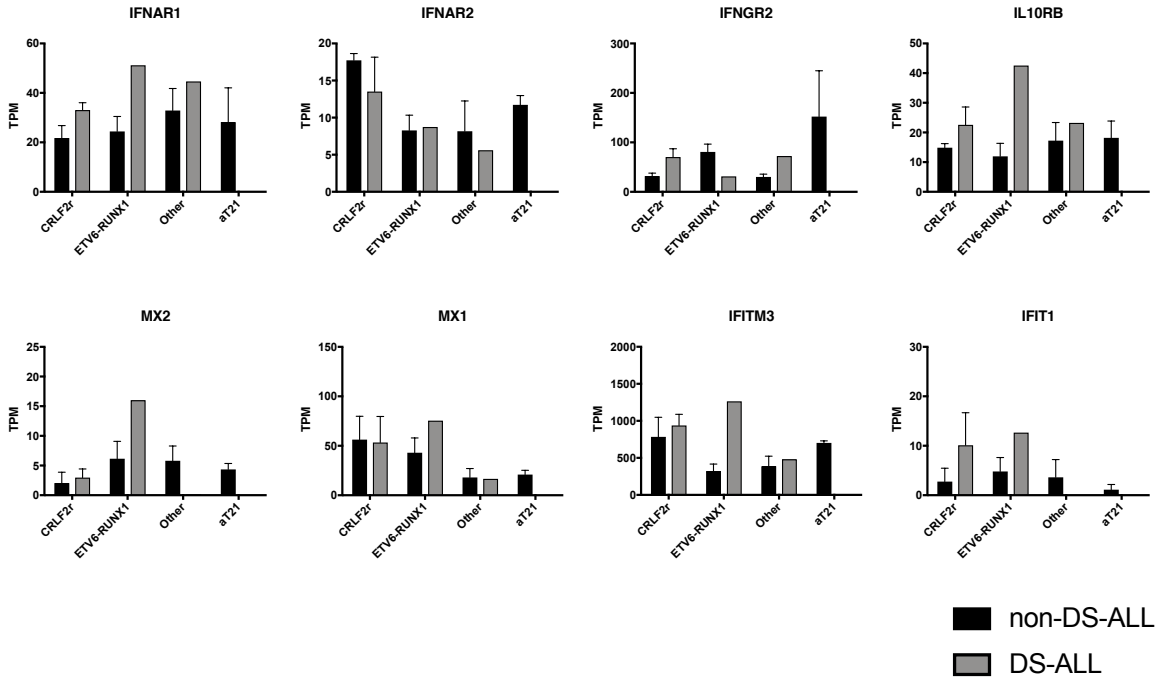


Figure 5.13: IFN signalling gene expression in non-DS and DS-ALL blasts by RNA-Seq (100 cells/population). non-DS-ALL: n=10; DS-ALL: n=7; aT21-ALL n=3.

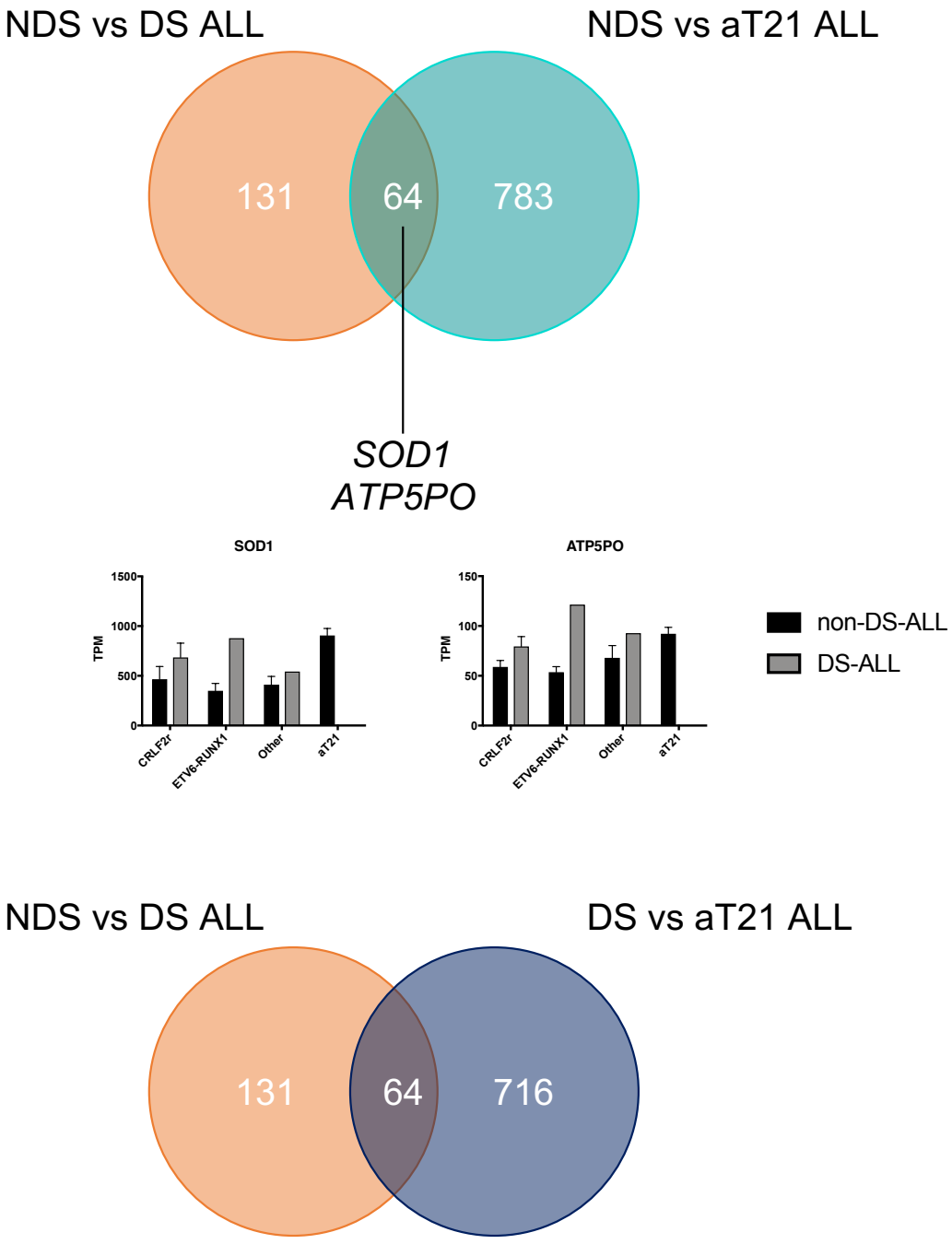


Figure 5.14: Common significant (FDR <0.1) DE protein coding genes between non-DS-ALL (NDS)/DS-ALL and aT21/NDS-ALL and aT21/DS-ALL comparisons. Only two genes located on chromosome 21 are significantly DE in DS-ALL and aT21-ALL blasts compared to non-DS-ALL: *SOD1* and *ATP5PO*.

5.4.2 Differential analysis of NM and T21 leukaemic HSC

Finally, differential expression analysis between HSC from non-DS-ALL (n=2) and DS-ALL (n=3) patients also showed an enrichment for pro-inflammatory signalling pathways (figure 5.15). These analyses were limited because of the inherent difficulty in obtaining adequate numbers of immunophenotypic HSC from leukaemic BM. Therefore, findings can only be considered preliminary as there are potential confounding factors including the small sample size (n=2 for non-DS-ALL HSC) and all the DS-ALL HSC being from CRLF2r DS-ALLs. However, in combination with the multiple other enrichments in fetal HSPC and DS-ALL, these data suggest a role for T21 driven pro-inflammatory signalling not only in the perturbation of T21 fetal haematopoiesis but also in facilitating leukaemogenesis in post-natal life.

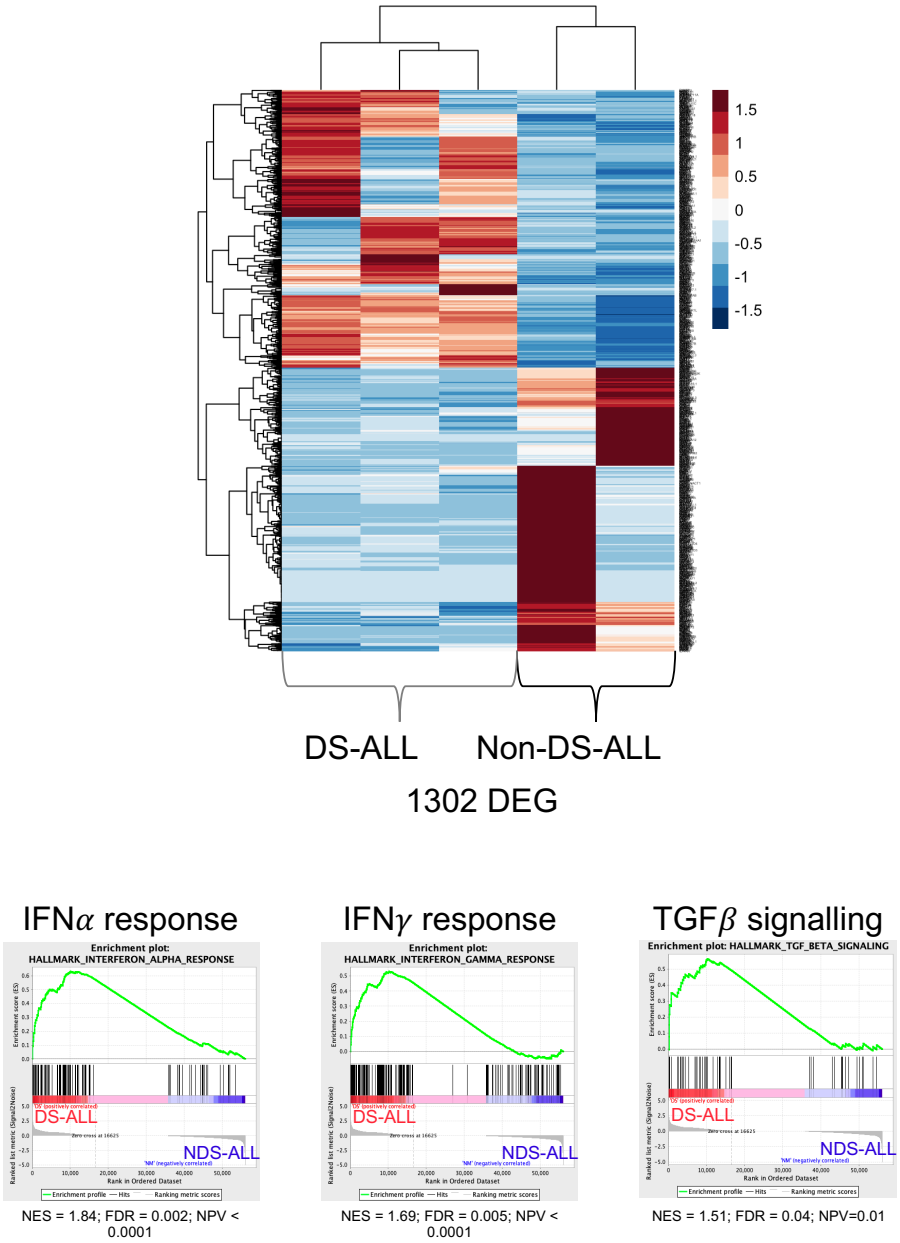


Figure 5.15: Top: Heatmap showing all significantly (FDR <0.1) differentially expressed protein coding genes between non-DS (NDS)-ALL (n=2) and DS-ALL (n=3) HSC by RNA-Seq (100 cells/population). Bottom: key enrichment plots for hallmark gene sets by GSEA. DEG: differentially expressed genes; NES: normalised enrichment score; FDR: false discovery rate; NPV: nominal p-value.

5.4.3 Comparison with normal fetal B progenitors

The comparison of the transcriptomes of fetal B progenitors (both NM and T21) with non-DS-ALL and DS-ALL blasts presents a challenge because these data are from two separate datasets sorted, processed and sequenced at different dates. To address this, I sorted HSPC and B cells from three NM fetal BM samples for RNA-Seq at the same time as the ALL samples as having the "same" cell populations (different biological samples) in each dataset would provide an internal control that would allow cross-dataset comparisons. After extensively testing three independent methods of batch removal (Limma, RUVSeq and SVA), I concluded that RUVSeq was the most appropriate as by PCA it appeared to effectively remove batch effects while also maintaining the variation in the datasets. This was achieved by specifying which transcriptomes should be similar i.e. the NM fetal BM HSC that were present in both datasets so that RUVSeq could remove variation from the whole dataset that could thereby be attributed to differences in between datasets.

However, before looking for similarities and differences between NM and T21 progenitors and ALL blasts, I performed a PCA on the ALL dataset by itself as an extra way to make sure RUVSeq wasn't causing artificial manipulations in the dataset. This PCA of the top 1000 most variable genes showed the B-lymphoid hierarchy was driving PC1 and this is exactly what I had observed in the previous dataset discussed in Chapter 3. All ALL blasts regardless of cytogenetic sub-group clustered away from NM progenitors. (figure 5.16).

After batch effect removal using RUVSeq, PCA of top 1000 and top 100 most variably expressed genes across both datasets clearly showed a separation of HSPC from B progenitors, B cells and blasts (figures 5.17 and 5.18). This is to be expected given that all blasts are of the B-lineage. Interestingly, the PCA of the top 100 most variably expressed genes caused blasts to cluster closely with committed B progenitors (PreProB and ProB progenitors) (figure 5.18).

Since PCA is a multi-dimensional reduction technique used to ask "how different is each transcriptome?", it is more appropriate to ask "how similar is each transcriptome?" when studying cell differentiation in any biological system. Therefore in addition to PCA, I also plotted diffusion maps. A similar pattern was observed with HSPC generally separating from committed B progenitors, B cells and blasts by DC1 and ELP bridging the gap (figure 5.19). Interestingly, by DC3 blasts clustered closely with

ProB progenitors (figure 5.19) supporting the hypothesis that ProB progenitors are a putative cell of origin in non-DS-ALL and DS-ALL.

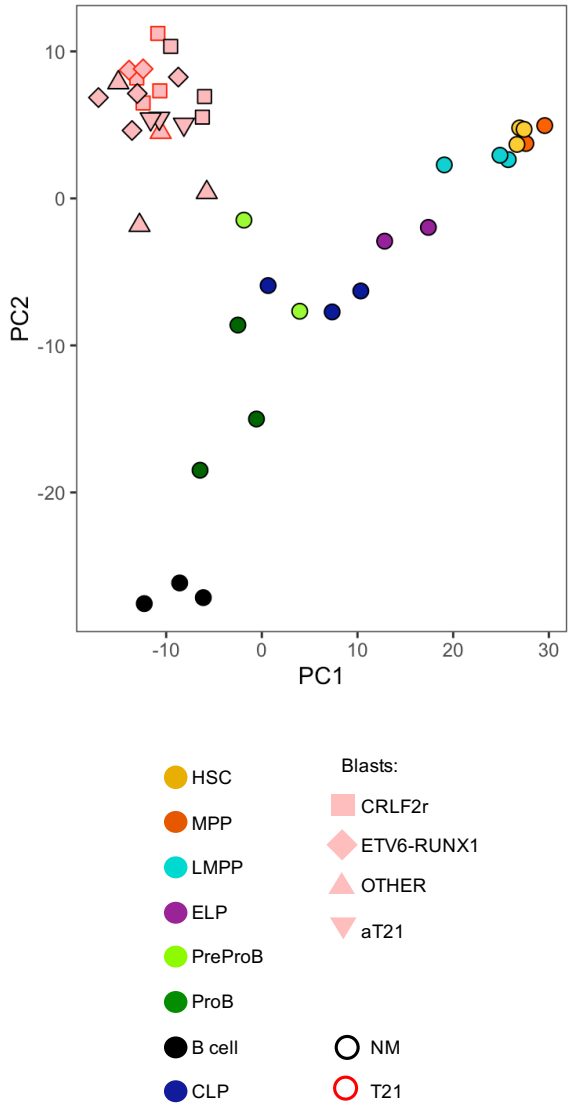


Figure 5.16: PCA of top 1000 most variably expressed genes by RNA-Seq (100 cells/population) of NM fetal BM HSPC and B cells and non-DS-ALL and DS-ALL blasts. Blasts regardless of cytogenetic subtype cluster separately from NM fetal BM progenitors and B cells. NM fetal BM: n=2-3; non-DS-ALL: n=13; DS-ALL: n=7.

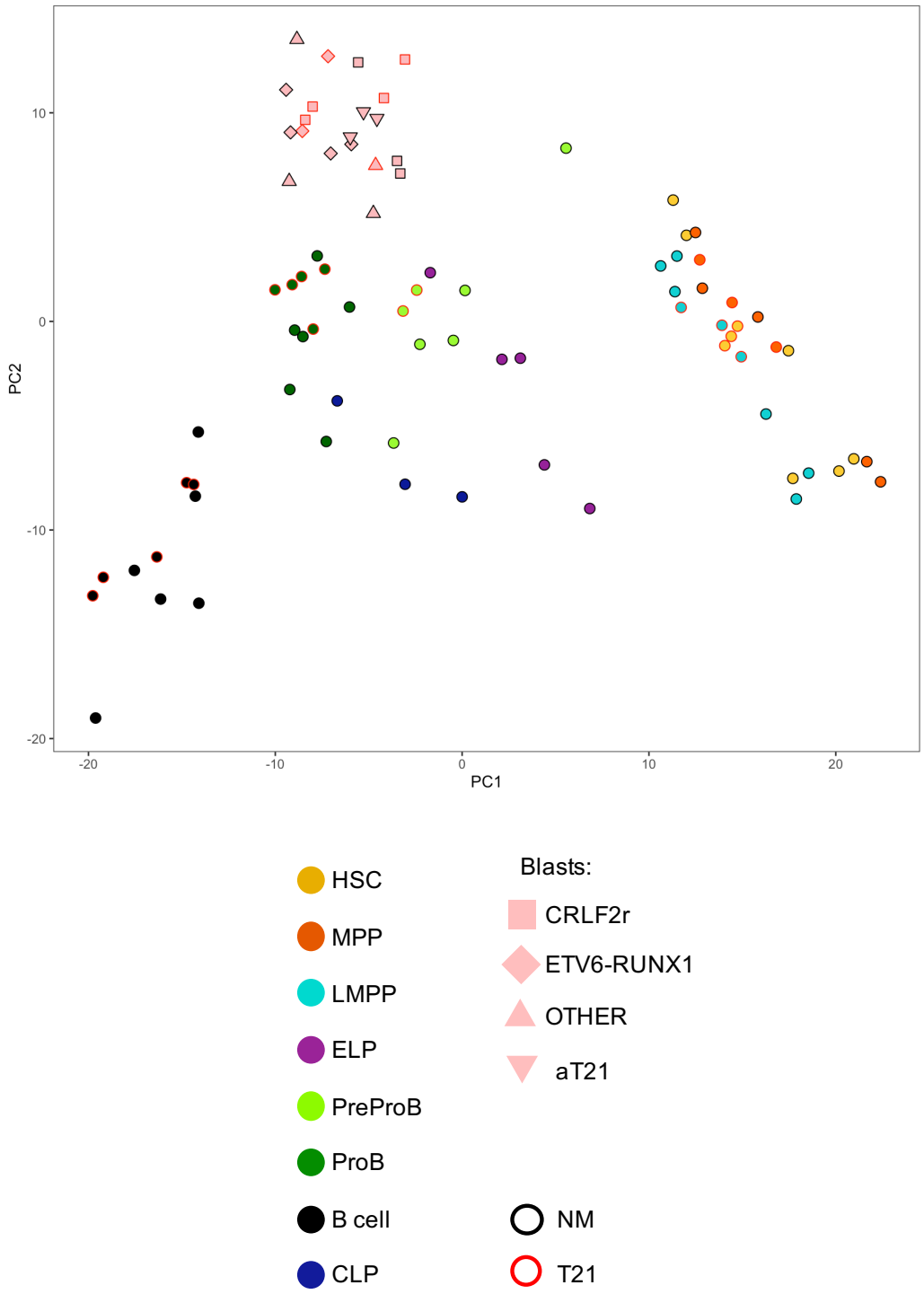


Figure 5.17: PCA of top 1000 most variably expressed genes by RNA-Seq (100 cells/population) of NM/T21 fetal BM HSPC and B cells and non-DS-ALL and DS-ALL blasts. HSC, MPP and LMPP cluster away from B-lineage progenitors, B cells and ALL blasts. Blasts cluster separately from B progenitors and B cells but are closest to the B progenitors. NM fetal BM: n=3-6; T21 fetal BM: n=2-5; non-DS-ALL: n=13; DS-ALL: n=7.

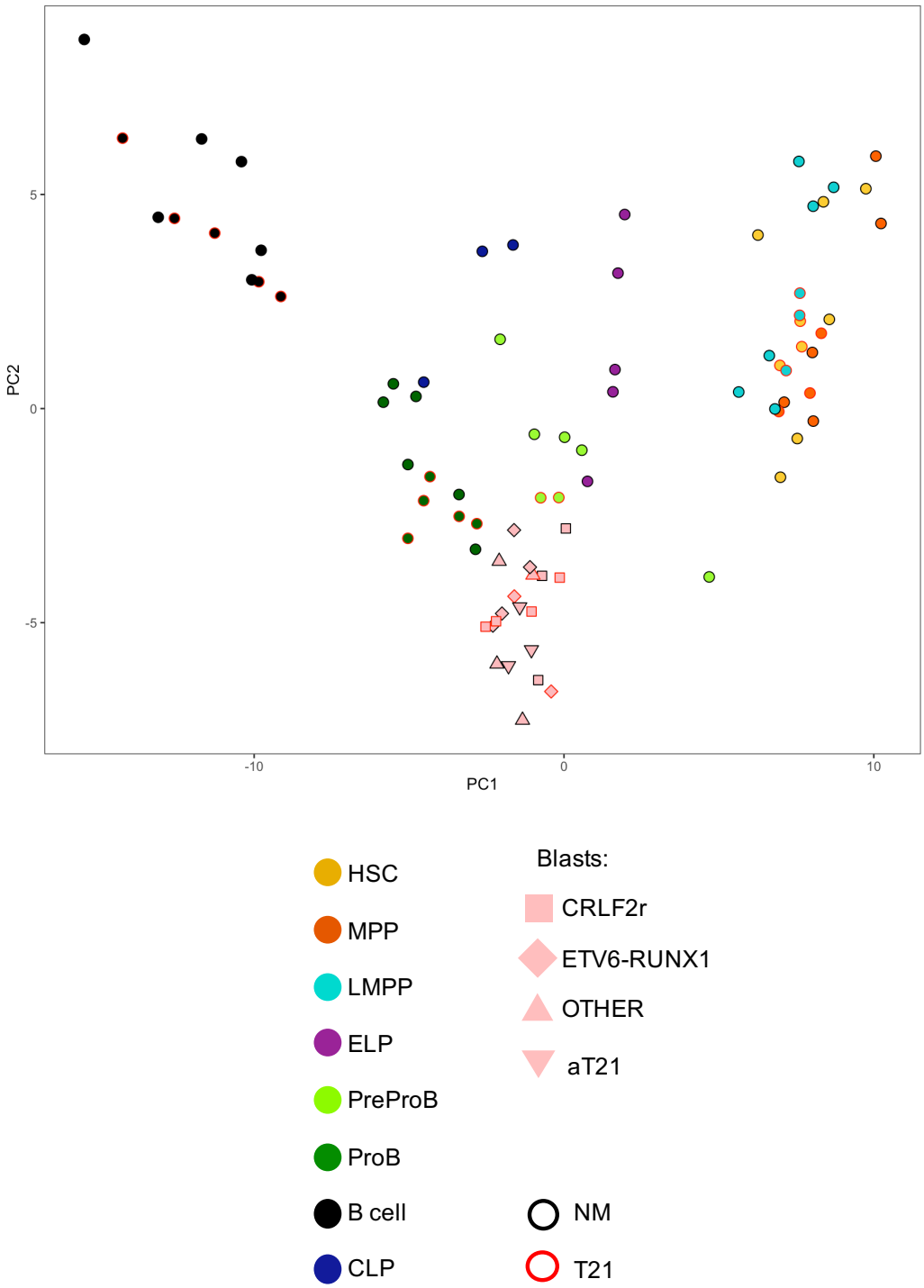


Figure 5.18: PCA of top 100 most variably expressed genes by RNA-Seq (100 cells/population) of NM/T21 fetal BM HSPC and B cells and non-DS-ALL and DS-ALL blasts. HSC, MPP and LMPP cluster away from B-lineage progenitors, B cells and ALL blasts. Blasts cluster most closely with committed B progenitors (PreProB and ProB progenitors). NM fetal BM: n=3-6; T21 fetal BM: n=2-5; non-DS-ALL: n=13; DS-ALL: n=7.

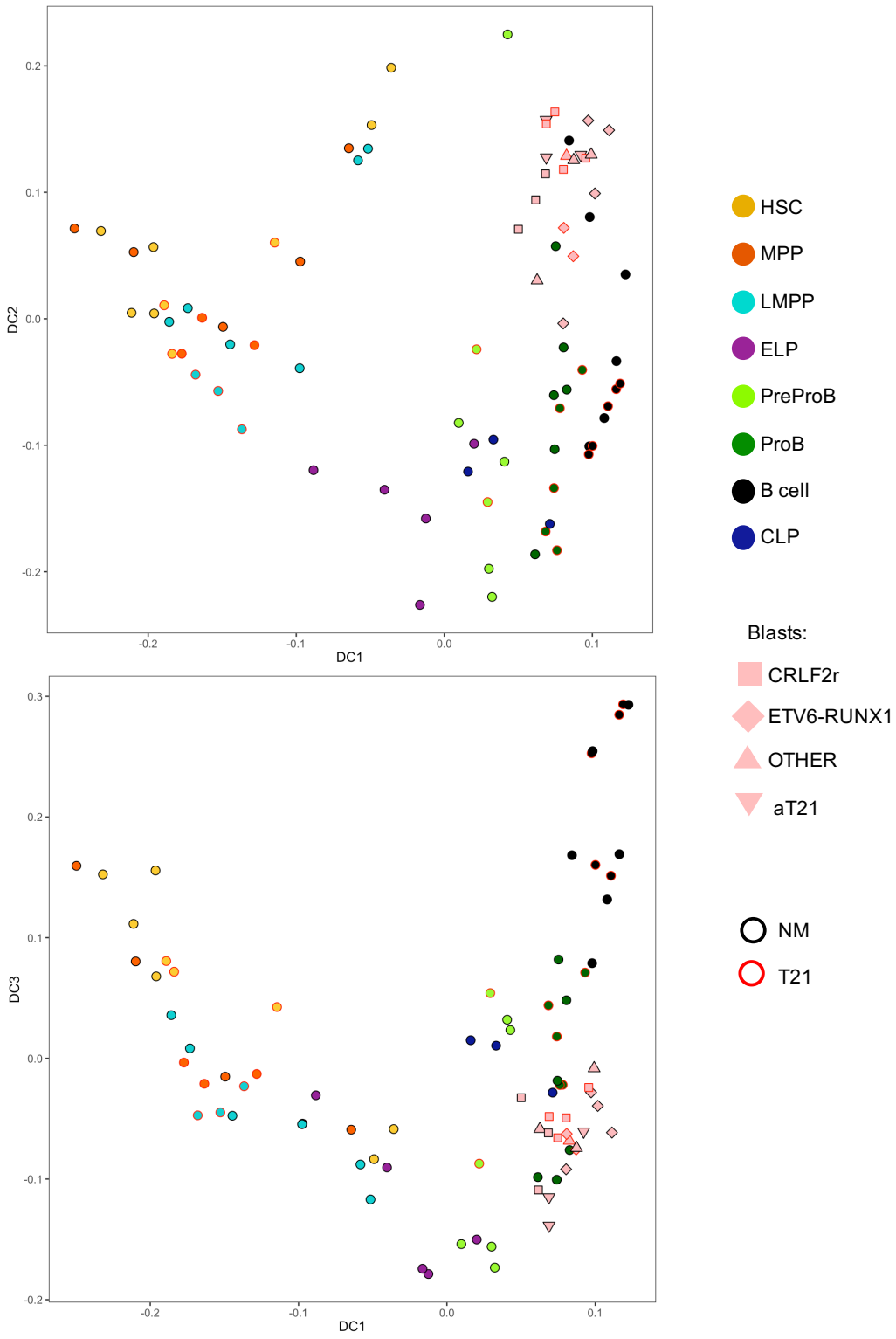


Figure 5.19: Diffusion maps of transcriptomes by RNA-Seq (100 cells/population) of NM/T21 fetal BM HSPC and B cells and non-DS-ALL and DS-ALL blasts. These diffusion maps suggest that ALL blasts share transcriptional similarities with committed B progenitors and specifically ProB progenitors irrespective of whether the population is NM or T21. NM fetal BM: n=3-6; T21 fetal BM: n=2-5; non-DS-ALL: n=13; DS-ALL: n=7.

5.5 Summary and discussion

The preliminary data and results from this chapter point towards a role for T21 driven pro-inflammatory signalling in the biology of DS-ALL. This, in combination with data and results discussed in Chapter 3 and Chapter 4 suggests that the inflammatory signatures observed in T21 fetal BM persist in the leukaemic setting. These data provide an important starting point for elucidating the molecular mechanism(s) involved in leukaemogenesis in children with DS.

5.5.1 T21 ProB progenitors

Comparison of NM and T21 fetal ProB progenitors showed a striking enrichment for gene sets associated with the cell cycle, suggesting that T21 ProB progenitors are cycling more. While this remains to be confirmed by other experiments, such as cell cycle analysis by flow cytometry and proliferation assays, this difference in cell cycling/proliferative capacity could explain the relative recovery of B-lymphopoiesis in children with DS, although several studies show that B lymphopenia is common in these children [de Hingh et al., 2005] [Verstegen et al., 2010]. Such cycling in the T21 fetal BM is likely to be the consequence of an as yet unknown compensatory mechanism at play in the T21 fetal BM, in response to the absence of normal fetal ELP and PreProB progenitors. It is not clear whether this is cell-intrinsic or mediated by extrinsic factors in the BM microenvironment.

The increased cycling of T21 ProB progenitors, in combination with their immunophenotype, also makes them an attractive candidate as the cell of origin for DS-ALL. Indeed, comparisons made by PCA between the two separate RNA-Seq transcriptome datasets show that ProB progenitors and ALL blasts cluster closely together. However, it should be noted that the similarities that cause such clustering could be a result of similarities in cell cycle gene expression themselves and on their own are insufficient to identify the cell of origin.

5.5.2 non-DS-ALL and DS-ALL: similarities and differences

Immunophenotypic and transcriptome analysis of DS-ALL blasts and comparison with non-DS-ALL blasts demonstrated the heterogeneity within the blast population described previously by others [Hertzberg et al., 2010], [Buitenkamp et al., 2014], [Lee et al., 2016]. When CRLF2r blast transcriptomes from non-DS-ALL and DS-ALL patients were compared, few genes were DE and there were no hallmark gene sets

enriched in the DS-ALL blasts. This, combined with PCA showing blasts clustering according to their cytogenetic group, suggests that the contribution of T21 to these leukaemias is most likely to lie with the early steps towards leukaemogenesis, as opposed to the interaction with the oncogenic hit of a chromosomal translocation.

Multiple comparisons between non-DS-ALL, DS-ALL and aT21-ALL showed an enrichment for pro-inflammatory signalling pathways in DS-ALL blasts. Additional comparisons between HSC isolated from non-DS-ALL and DS-ALL patients also showed enrichments for pro-inflammatory signalling pathways. This hints towards a distinct role for the trisomic microenvironment in the DS-ALL setting, that could contribute to leukaemogenesis and/or leukaemia maintenance.

Very recently, a larger DS-ALL RNA-Seq study [Kubota et al., 2019] confirmed what I have observed here: DS-ALL blast transcriptomes are shaped by their cytogenetic sub-group. Although the distribution of certain cytogenetic abnormalities is different between non-DS-ALL and DS-ALL, there is as much heterogeneity in DS-ALL as there is in non-DS-ALL. Kubota *et al.* also speculate how differences in global methylation profile of DS-ALL blasts could contribute to the increased incidence of ALL in DS children. While these findings will need to be independently confirmed by bisulfite sequencing/ChIP-Seq (the authors performed a DNA methylation array), this is interesting because I also noted (in Chapter 4) gene set enrichments suggesting a role for dysregulation of epigenetically controlled gene expression in T21 HSC. Taken together my data suggest that multiple mechanisms are likely to contribute towards DS-ALL pathogenesis.

5.5.3 Future work

These data are preliminary and will form the basis of a new project in the lab. Here I have performed immunophenotypic and transcriptomic analysis of DS-ALL and non-DS-ALL that has led to the generation of a hypothesis regarding the molecular pathogenesis of DS-ALL, that could explain the increased incidence of ALL in children with DS. I hypothesise that increased T21-driven pro-inflammatory signalling in the BM microenvironment increases the likelihood of certain oncogenic events, that combined with the recovery/compensatory mechanisms employed in the wake of the fetal B-lymphoid defect leads to the increased incidence of DS-ALL. This is an attractive hypothesis given that other independent groups have also suggested a role for inflammation in leukaemogenesis [Hemminki et al., 2013] [Greaves, 2018] [Cuartero

et al., 2018] [Forte et al., 2019] and very recently in a commentary others have also hypothesised a role for IFN signalling in DS-ALL leukaemogenesis specifically [Birger et al., 2019]. However, the role of inflammatory pathways in DS-ALL is likely to be complex and multi-factorial since inflammation has distinct roles in maintenance of homeostasis in the HSC niche, [Pinho and Frenette, 2019] and also in BM niche dysregulation in leukaemia [Baryawno et al., 2019] [Ahsberg et al., 2019]. Therefore, testing this hypothesis will require extensive further work. Firstly, I think a more exhaustive analysis of these datasets could identify a core gene signature that could be attributed to T21 leukaemia. These signatures have the potential to aid future analyses of single cell RNA-Seq datasets that are currently being generated in the lab. For example, single cell RNA-Seq of whole fetal BM/leukaemic BM has the potential to provide insight into the transcriptomes of haematopoietic and niche cells in parallel. In addition to further RNA-Seq experiments to dissect the heterogeneity and multiple cell types involved, it is also essential that findings are validated by alternative methods such as single cell RT-qPCR. Then following on from this, I think it would also be incredibly informative to look for active inflammatory signalling pathways using CyTOF/mass cytometry. Ultimately, elucidating the active signalling pathways involved before and during leukaemia could identify new therapeutic targets, which is very much needed in the case of DS-ALL where treatment related infection is a significant cause of death in this patient group [Bohnstedt et al., 2013] [Meyr et al., 2013] [Buitenkamp et al., 2014] [Athale et al., 2018] [Matloub et al., 2019].

Chapter 6

Discussion

6.1 Background

As discussed during this thesis, studies in mouse models as well as in humans suggest that chromosome 21 confers a unique susceptibility to acute leukaemia. Previous work [Tunstall-Pedoe et al., 2008] [Chou et al., 2008] [Roy et al., 2012] has shown that T21 causes widespread perturbation of haematopoiesis in FL. In addition to a marked expansion in immunophenotypic HSC and MEP, both GMP and B progenitors are decreased in T21 FL. This disruption, in combination with acquired N-terminal GATA1 mutations, appears to play a pivotal role in the development of ML-DS since this leukaemia is unique to DS and GATA1 mutations are not leukaemogenic in the absence of T21 [Roberts et al., 2013] [Bhatnagar et al., 2016]. In contrast to ML-DS, which arises in FL, DS-ALL is characterised by BM infiltration and yet the effects of T21 on fetal B-lymphopoiesis in fetal BM, and how this may relate to the pathogenesis of DS-ALL remains relatively unexplored. In this thesis I have addressed this by investigating the cellular hierarchy and molecular basis for B-lymphoid development in normal human fetal BM and T21 human fetal BM for the first time; and directly comparing the T21 fetal transcriptome with those of ALL blast cells.

6.2 Defining normal fetal BM B-lymphopoiesis

While adult BM is relatively well characterised, there was very little known about normal B cell development in human fetal BM when I started my project. In order to understand the impact of T21 on fetal B-lymphopoiesis, I first had to characterise normal fetal BM B-lymphopoiesis in detail.

I first determined the frequency of immunophenotypic HSPC and B cells in normal

fetal BM and compared this with normal FL through the first and second trimesters using flow cytometry. This showed that haematopoiesis in fetal BM begins at the end of the first trimester with the simultaneous detection of CD10- PreProB progenitors, ProB progenitors and other HSPC at approximately 11 pcw. In contrast to HSC, MPP, LMPP and ELP, both PreProB and ProB progenitors expanded rapidly at the beginning of first trimester in the fetal BM with the CD10- PreProB progenitors proving to be the most proliferative of these HSPC populations by cell cycle analysis.

To investigate the B-lymphoid hierarchy I performed extensive *in vitro* differentiation assays to look for B/NK cell potential (MS-5 co-cultures), T cell potential (OP9-DL1 co-cultures) and myeloid/erythroid potential (colony forming assays). As expected, fetal BM HSC and MPP showed multi-lineage potential in that they differentiated into B, NK, T, myeloid and erythroid cells while LMPP had lost their erythro-megakaryocytic potential and generated both lymphoid (B, NK and T) and myeloid lineages. ELP, as defined in my studies (Lin2-CD34+CD19-CD10-CD45RA+CD127+), mainly generated lymphoid cells but were able to generate small numbers of myeloid cells in MS-5 and fetal BM MSC co-cultures. This supported other work, published during my project, that showed that ELP in FL also had residual myeloid potential [Alhaj Hussen et al., 2017][Boiers et al., 2018]. By contrast, PreProB and ProB progenitors only produced B cells both *in vitro* and *in vivo*. Furthermore, I was able to demonstrate both *in vitro* and *in vivo* that PreProB progenitors lie upstream of ProB progenitors showing for the first time that, at a functional level, PreProB progenitors are the earliest progenitor in fetal BM committed to the B lineage.

Although PreProB and ProB progenitors are functionally similar in that they both differentiate into B cells, my data show that they are molecularly distinct. Transcriptome profiling by RNA-Seq showed multiple differences in the PreProB progenitor and ProB progenitor transcriptomes. Single cell RT-qPCR for genes associated with B-lineage commitment and leukaemia confirmed this and also placed PreProB progenitors between ELP and ProB progenitors in a differentiation trajectory. This was further supported by analysis of VDJ rearrangement of the IgH locus by PCR. I also performed ATAC-Seq on PreProB and ProB progenitors, hypothesising that chromatin accessibility would be altered in more mature B progenitors, supporting the differences in gene expression I had already observed. Indeed, there were subtle differences in the chromatin accessibility between PreProB and ProB progenitors.

Finally comparisons with adult PreProB progenitors suggested that fetal BM PreProB progenitors have unique gene expression programmes. More extensive studies of the individual HSPC populations, and comparison with their post-natal counterparts will be needed to better understand the significance of these differences. However, this first part of my project has recently been published [O'Byrne et al., 2019] and is important because it demonstrates that CD10- PreProB progenitors are unique to fetal life. In addition, as many childhood ALLs arise *in utero* and CD10- PreProB progenitors share several immunophenotypic and transcriptomic similarities with infant ALL, these data raise the possibility that PreProB progenitors may be a putative cell of origin for iALL.

6.3 The impact of Trisomy 21 on fetal B-lymphopoiesis

Having demonstrated the fetal B-lymphoid hierarchy in normal fetal BM, and with evidence that T21 fetal B-lymphopoiesis is perturbed in the FL [Roy et al., 2012], I measured the frequency of T21 fetal BM progenitors involved in the B-lymphoid hierarchy. Interestingly, the B-lymphoid defect was both different and more profound in T21 fetal BM than previously found in FL. In contrast to FL, the frequency of ELP, as well as PreProB progenitors and B cells, was significantly reduced in T21 fetal BM compared to normal fetal BM. On the other hand, the frequency of ProB progenitors in T21 fetal BM was not significantly reduced. As noted previously in FL, T21 also caused perturbation of B cell development upstream of B-lineage commitment. Although, immunophenotypic HSC frequency was increased in T21 fetal BM, functional analysis of their differentiation potential demonstrated that T21 HSC as well as other T21 HSPC (MPP, LMPP and ELP) all had reduced B-lymphoid potential in MS-5 co-culture assays.

Transcriptome analysis by RNA-Sequencing of sorted T21 HSPC populations and B cells suggested that the B-lymphoid defect is likely to be complex and not solely due to a direct effect of T21 on HSPC. There were few significant differences in the expression of key B-lymphoid genes or Hsa21 genes with a known role in haematopoiesis; and instead key individual gene expression differences appeared to be in the expression of megakaryocyte genes. Indeed, my analysis found no evidence for a T21-associated change in any single Hsa21 gene which would explain the B lymphoid defect, suggesting that mouse models may be of limited value in unpicking the molecular basis of impaired

B-lymphopoiesis in DS. Further to this, the trajectory inferred by diffusion map analysis was similar in T21 HSPC compared to normal HSPC indicating that once the B-lymphoid programme is switched on, B cells can be produced although the reduced numbers of B cells indicates that differentiation is abnormal. Of particular note, my transcriptomic analysis revealed a marked overall increase in gene expression in T21 HSPC compared to their normal counterparts. This was evident in all of the populations which I studied and suggests that T21 may perturb some of the normal mechanisms employed by differentiating cells to silence alternative lineage-associated programmes. Further interrogation of the epigenetic differences between T21 and normal fetal HSPC will be needed to investigate this further.

Given these patterns of perturbation of B cell development, I decided to focus on comparing the transcriptomes of HSPC upstream of the block in B-lymphopoiesis (HSC and LMPP). This revealed enrichment for gene sets associated with cytokine signalling/inflammatory pathways pointing towards a role for the trisomic micro-environment. To investigate this further, I used the MSC co-culture system to investigate the impact of T21 MSC on B-lymphopoiesis by co-culturing normal fetal BM HSPC with normal or T21 MSC. Although normal fetal BM progenitors committed to the B-lineage (PreProB progenitors and ProB progenitors) were unaffected, all upstream normal fetal BM HSPC showed reduced B-lymphoid potential on T21 MSC and this was particularly profound in normal HSC/T21 MSC co-cultures.

Transcriptome analysis of the T21 MSC also showed enrichments in inflammatory gene sets. Therefore, I sought to further investigate the roles of IL-6, IFNα and TGFβ. While co-cultures with IL-6 proved inconclusive, this may reflect differences between biological samples, but also possibly the absence of other coordinated inflammatory responses, that together act in concert to disrupt the B-lymphoid programme. Initial analysis of other candidates such as IFNα and TGFβ are also promising but will require further validation.

While others have suggested a role for increased inflammatory signalling, namely IFNα in DS-associated immunodeficiency [Sullivan et al., 2017], this is the first time such signalling pathways have been linked to abnormal fetal haematopoiesis. It is tempting to speculate that T21 driven pro-inflammatory signalling could not only explain the haematopoietic abnormalities observed in T21 fetal BM, but also the increased incidence of DS-ALL. In light of the increasing evidence from multiple labs

proposing a role for inflammation in leukaemogenesis (and cancer in general), whereby chronic inflammation induces DNA damage [Krawczyk et al., 2014] [Greaves, 2018] [Jones and Jenkins, 2018] [Greten and Grivennikov, 2019] and reported differences in gene expression the blood cells of individuals with DS [Zampieri et al., 2014] [Sullivan et al., 2016] [Sullivan et al., 2017], it seems likely that this inflammatory signalling persists into post-natal life contributing to increased incidence of leukaemia.

6.4 The role of Trisomy 21 in DS-ALL

To directly interrogate the role of T21 in DS-ALL, I performed transcriptome analysis on cytogenetically matched paediatric non-DS-ALL and DS-ALL blasts. Non-DS-ALLs with acquired T21 were included in this analysis in order to probe the effects of the additional copy of Hsa21 in the absence of the T21 microenvironment. Interestingly, comparisons between non-DS-ALL and DS-ALL blasts showed enrichments for inflammatory signalling pathway gene sets (IFNγ response, inflammatory response and TNFα) in the DS-ALL blasts. Furthermore, when compared to non-DS-ALL blasts with acquired T21, DS-ALL blasts showed a significant enrichment for IFNα and IFNγ response much like the T21 HSPC in fetal BM. This either suggests an ongoing role for T21 microenvironment-driven pro-inflammatory signalling after birth in T21 BM or, alternatively, that these pro-inflammatory changes are the result of HPSC programming during fetal life, which subsequently undergo transformation to DS-ALL. Both of these possibilities are supported by similar enrichments (IFNα response, IFNγ response and TGFβ signalling) in immunophenotypic HSC from DS-ALL patients when compared to HSC from non-DS-ALL patients. My data analysis also shows that the cytogenetic abnormalities present in DS-ALL have the biggest influence over the transcriptome whether compared with non-DS-ALL or within the DS-ALL patient group. My results are confirmed by 3 recent studies published during the course of my project which also show that the same recurrent cytogenetic drivers of leukaemia in non-DS-ALL also drive DS-ALL [Schwartzman et al., 2017] [Vesely et al., 2017] [Kubota et al., 2019].

Uniquely, these data have been compared to fetal B-lymphoid cells and represent the first evidence linking perturbation of fetal lymphopoiesis by T21 to the pathogenesis of DS-ALL. Specifically, these preliminary data point towards a role for T21-driven pro-inflammatory signalling in haematopoiesis and leukaemia.

6.5 Technical limitations

A major limitation to working with primary human cells is the size of the samples and hence the number of cells available for analysis. In particular, the numbers of HSPC both in normal and T21 samples is usually considerably lower in fetal BM compared to FL, which makes isolation of the HSC, MPP and LMPP populations from fetal BM especially challenging. This means that although I endeavoured to plan experiments such that several assays could be performed on the same biological sample, this was not always possible and some of my experiments used lower than optimal cell numbers, or had a small number of biological replicates. Another limitation was in cell sorting and laser availability. In optimising my sort panels, I could confidently sort up to 4 populations from each panel using 10-12 colours. It is exciting to hear of a new sorter with five lasers and 6-way sorting being developed by BD Biosciences that would enable further optimisation of my sort strategies and potentially reduce 'wasted' cells.

Finally, although biological variability certainly has a role to play in the variable output from co-cultures, it is very likely that in the case of the MSC co-cultures this variability was also the result of a heterogeneous population of MSC. At the start of this project, I used the accepted immunophenotypic and morphological definition of MSC to define the MSC used in co-culture [Dominici et al., 2006]. However, the very recent single cell analysis in mice demonstrating heterogeneity in the murine stromal BM microenvironment suggests that human MSC are also a mixed population [Baryawno et al., 2019] [Tikhonova et al., 2019]. This is being addressed in the lab as single cell transcriptomic studies on whole fetal BM (normal and T21) which includes stromal cells is currently under way.

6.6 Future work

The work in my project has led to a number of interesting questions. Firstly, following on from the identification of the normal fetal BM B-lymphopoiesis hierarchy, the relationship between this hierarchy and the adult B-lymphopoiesis hierarchy has not been fully addressed. I hypothesise that these pathways develop in parallel in the fetal BM, and that this facilitates the relative recovery of ProB progenitors observed in T21 fetal BM. Flow cytometry data from fetal BM suggests that an adult type immunophenotypic CLP (Lin-CD34+CD38+CD10+) is present but its functional and molecular signatures remain uncharacterised. To investigate the hypothesis of

a parallel fetal and adult B-lymphoid pathway, further functional experiments at the single cell level are required, as well as comprehensive transcriptomic analysis of post-natal tissues in parallel with fetal tissues.

Secondly, to establish the role of inflammation in the B-lymphoid defect in T21 fetal BM and the role it plays in leukaemogenesis of DS-ALL, extensive further work is required and this is currently underway. The expression of key genes identified by the RNA-Seq analysis and involved in inflammation are currently being tested by single cell RT-qPCR in normal and T21 fetal BM HSPC. As is evidenced by the transcriptomic data presented in this thesis, chronic inflammation is unlikely to exclusively affect haematopoiesis directly, and inflammation induced changes to the microenvironment could also ultimately impact on haematopoiesis [Goedhart et al., 2018] [Greten and Grivennikov, 2019] [Forte et al., 2019]. To address this, single cell RNA-Seq of whole BM is being performed and it is hoped that these data may be used to tease apart the altered cross-talk between microenvironment and HSPC in the same biological sample. Following on from this it would be interesting to see what signalling pathways are active using CyTOF/mass cytometry; and finally functional analyses either by attempting to rescue or replicate the B-lymphoid defect are essential.

Another interesting finding from the transcriptomic data analysis of T21 fetal BM was the enrichment for chromatin silencing and nucleosome organisation gene sets in normal HSC, suggesting a dysregulation of epigenetic silencing in T21 HSC. These findings are interesting given others' research linking chromosome 21 gene, *HMGN1*, to epigenomic signatures associated with a DS-ALL mouse model [Lane et al., 2014] [Mowery et al., 2018]. Since there is an apparent global failure to silence in gene expression in T21 HSPC, it would be very interesting to look at activating (i.e. H3K27Ac) and repressive (H3K27Me3) histone marks by ChIP-Seq.

6.7 Conclusions

This project has characterised human fetal B-lymphopoiesis in detail for the first time [O'Byrne et al., 2019]. These data have important implications which may aid our understanding of childhood leukaemia. Following on from this, I have shown that fetal B-lymphopoiesis is significantly altered in T21 fetal BM. Much of the functional and transcriptomic analysis points towards the role of the trisomic microenvironment in these differences and particularly T21 driven pro-inflammatory pathways that are

preserved in the DS-ALL setting. While these findings are unlikely to be the sole contributor that explains the B-lymphoid defect in T21 fetal life, the link between inflammatory signalling and leukaemogenesis points towards a significant role for these pathways in T21 fetal B-lymphopoiesis and leukaemogenesis.

Bibliography

[Abou-Ezzi et al., 2019] Abou-Ezzi, G., Supakorndej, T., Zhang, J., Anthony, B., Krambs, J., Celik, H., Karpova, D., Craft, C. S., and Link, D. C. (2019). TGF-B Signaling Plays an Essential Role in the Lineage Specification of Mesenchymal Stem/Progenitor Cells in Fetal Bone Marrow. *Stem Cell Reports*, 13(1):48–60.

[Ahn et al., 2013] Ahn, E. E.-Y., Higashi, T., Yan, M., Matsuura, S., Hickey, C. J., Lo, M.-C., Shia, W.-J., DeKelver, R. C., and Zhang, D.-E. (2013). SON Protein Regulates GATA-2 through Transcriptional Control of the MicroRNA 23a-27a-24-2 Cluster. *The Journal of Biological Chemistry*, 288(8):5381–5388.

[Ahsberg et al., 2019] Ahsberg, J., Xiao, P., Okuyama, K., Somasundaram, R., Strid, T., Qian, H., and Sigvardsson, M. (2019). Progression of progenitor B cell leukemia is associated with alterations of the bone marrow micro-environment. *Haematologica*, page haematol.2018.214031.

[Alexandrov et al., 2018] Alexandrov, P. N., Percy, M. E., and Lukiw, W. J. (2018). Chromosome 21-Encoded microRNAs (mRNAs): Impact on Down's Syndrome and Trisomy-21 Linked Disease. *Cellular and Molecular Neurobiology*, 38(3):769–774.

[Alford et al., 2010] Alford, K. A., Slender, A., Vanes, L., Li, Z., Fisher, E. M. C., Nizetic, D., Orkin, S. H., Roberts, I., and Tybulewicz, V. L. J. (2010). Perturbed hematopoiesis in the Tc1 mouse model of Down syndrome. *Blood*, 115(14):2928–2937.

[Alhaj Hussen et al., 2017] Alhaj Hussen, K., Vu Manh, T.-P., Guimiot, F., Nelson, E., Chabaane, E., Delord, M., Barbier, M., Berthault, C., Dulphy, N., Alberdi, A. J., Burlen-Defranoux, O., Socié, G., Bories, J. C., Larghero, J., Vanneaux, V., Verhoeyen, E., Wirth, T., Dalod, M., Gluckman, J. C., Cumano, A., and Canque, B. (2017). Molecular and Functional Characterization of Lymphoid Progenitor Subsets Reveals a Bipartite Architecture of Human Lymphopoiesis. *Immunity*, 47(4):680–696.e8.

[Alpar et al., 2015] Alpar, D., Wren, D., Ermini, L., Mansur, M. B., Delft, F. W. v., Bateman, C. M., Titley, I., Kearney, L., Szczepanski, T., Gonzalez, D., Ford, A. M., Potter, N. E., and Greaves, M. (2015). Clonal origins of ETV6 RUNX1+ acute lymphoblastic leukemia: studies in monozygotic twins. *Leukemia*, 29(4):839–846.

[Ambros et al., 2003] Ambros, V., Bartel, B., Bartel, D. P., Burge, C. B., Carrington, J. C., Chen, X., Dreyfuss, G., Eddy, S. R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G., and Tuschl, T. (2003). A uniform system for microRNA annotation. *RNA*, 9(3):277–279.

[Anderson and Su, 2011] Anderson, M. S. and Su, M. A. (2011). Aire and T cell Development. *Current opinion in immunology*, 23(2):198–206.

[Andersson et al., 2015] Andersson, A. K., Ma, J., Wang, J., Chen, X., Gedman, A. L., Dang, J., Nakitandwe, J., Holmfeldt, L., Parker, M., Easton, J., Huether, R., Kriwacki, R., Rusch, M., Wu, G., Li, Y., Mulder, H., Raimondi, S., Pounds, S., Kang, G., Shi, L., Becksfort, J., Gupta, P., Payne-Turner, D., Vadodaria, B., Boggs, K., Yergeau, D., Manne, J., Song, G., Edmonson, M., Nagahawatte, P., Wei, L., Cheng, C., Pei, D., Sutton, R., Venn, N. C., Chetcuti, A., Rush, A., Catchpoole, D., Heldrup, J., Fioretos, T., Lu, C., Ding, L., Pui, C.-H., Shurtleff, S., Mullighan, C. G., Mardis, E. R., Wilson, R. K., Gruber, T. A., Zhang, J., Downing, J. R., and Project, f. T. S. J. C. R. H.-W. U. P. C. G. (2015). The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nature Genetics*, 47(4):330–337.

[Antonarakis, 1998] Antonarakis, S. E. (1998). 10 Years ofGenomics,Chromosome 21, and Down Syndrome. *Genomics*, 51(1):1–16.

[Antonarakis, 2017] Antonarakis, S. E. (2017). Down syndrome and the complexity of genome dosage imbalance. *Nature Reviews Genetics*, 18(3):147–163.

[Arico et al., 2008] Arico, M., Ziino, O., Valsecchi, M. G., Cazzaniga, G., Baronci, C., Messina, C., Pession, A., Santoro, N., Basso, G., and Conter, V. (2008). Acute lymphoblastic leukemia and Down syndrome. *Cancer*, 113(3):515–521.

[Asma et al., 1984] Asma, G. E., Langlois van den Bergh, R., and Vossen, J. M. (1984). Development of pre-B and B lymphocytes in the human fetus. *Clinical and Experimental Immunology*, 56(2):407–414.

[Athale et al., 2018] Athale, U. H., Puligandla, M., Stevenson, K. E., Asselin, B., Clavell, L. A., Cole, P. D., Kelly, K. M., Laverdiere, C., Leclerc, J.-M., Michon, B., Schorin, M. A., Sulis, M. L., Welch, J. J. G., Harris, M. H., Neuberg, D. S., Sallan, S. E., and Silverman, L. B. (2018). Outcome of children and adolescents with Down syndrome treated on Dana-Farber Cancer Institute Acute Lymphoblastic Leukemia Consortium protocols 00–001 and 05-001. *Pediatric Blood & Cancer*, 65(10):e27256.

[Awong et al., 2007] Awong, G., La Motte-Mohs, R. N., and Zúñiga-Pflücker, J. C. (2007). Generation of pro-T cells in vitro: potential for immune reconstitution. *Seminars in Immunology*, 19(5):341–349.

[Banno et al., 2016] Banno, K., Omori, S., Hirata, K., Nawa, N., Nakagawa, N., Nishimura, K., Ohtaka, M., Nakanishi, M., Sakuma, T., Yamamoto, T., Toki, T., Ito, E., Yamamoto, T., Kokubu, C., Takeda, J., Taniguchi, H., Arahori, H., Wada, K., Kitabatake, Y., and Ozono, K. (2016). Systematic Cellular Disease Models Reveal Synergistic Interaction of Trisomy 21 and GATA1 Mutations in Hematopoietic Abnormalities. *Cell Reports*, 15(6):1228–1241.

[Baryawno et al., 2019] Baryawno, N., Przybylski, D., Kowalczyk, M. S., Kfoury, Y., Severe, N., Gustafsson, K., Kokkaliaris, K. D., Mercier, F., Tabaka, M., Hofree, M., Dionne, D., Papazian, A., Lee, D., Ashenberg, O., Subramanian, A., Vaishnav, E. D., Rozenblatt-Rosen, O., Regev, A., and Scadden, D. T. (2019). A Cellular Taxonomy of the Bone Marrow Stroma in Homeostasis and Leukemia. *Cell*, 177(7):1915–1932.e16.

[Bateman et al., 2015] Bateman, C. M., Alpar, D., Ford, A. M., Colman, S. M., Wren, D., Morgan, M., Kearney, L., and Greaves, M. (2015). Evolutionary trajectories of hyperdiploid ALL in monozygotic twins. *Leukemia*, 29(1):58–65.

[Baumgarth, 2017] Baumgarth, N. (2017). A Hard(y) Look at B-1 Cell Development and Function. *The Journal of Immunology*, 199(10):3387–3394.

[Bendall et al., 2014] Bendall, S. C., Davis, K. L., Amir, E.-a. D., Tadmor, M. D., Simonds, E. F., Chen, T. J., Shenfeld, D. K., Nolan, G. P., and Pe'er, D. (2014). Single-Cell Trajectory Detection Uncovers Progression and Regulatory Coordination in Human B Cell Development. *Cell*, 157(3):714–725.

[Bercovich et al., 2008] Bercovich, D., Ganmore, I., Scott, L. M., Wainreb, G., Birger, Y., Elimelech, A., Shochat, C., Cazzaniga, G., Biondi, A., Basso, G., Cario, G.,

Schrappe, M., Stanulla, M., Strehl, S., Haas, O. A., Mann, G., Binder, V., Borkhardt, A., Kempski, H., Trka, J., Bielorei, B., Avigad, S., Stark, B., Smith, O., Dastugue, N., Bourquin, J.-P., Tal, N. B., Green, A. R., and Izraeli, S. (2008). Mutations of JAK2 in acute lymphoblastic leukaemias associated with Down's syndrome. *The Lancet*, 372(9648):1484–1492.

[Bergholdt et al., 2006] Bergholdt, R., Eising, S., Nerup, J., and Pociot, F. (2006). Increased prevalence of Down's syndrome in individuals with type 1 diabetes in Denmark: a nationwide population-based study. *Diabetologia*, 49(6):1179.

[Berthault et al., 2017] Berthault, C., Ramond, C., Burlen-Defranoux, O., Soubigou, G., Chea, S., Golub, R., Pereira, P., Vieira, P., and Cumano, A. (2017). Asynchronous lineage priming determines commitment to T cell and B cell lineages in fetal liver. *Nature Immunology*, 18(10):1139–1149.

[Bhatnagar et al., 2016] Bhatnagar, N., Nizery, L., Tunstall, O., Vyas, P., and Roberts, I. (2016). Transient Abnormal Myelopoiesis and AML in Down Syndrome: an Update. *Current Hematologic Malignancy Reports*, 11(5):333–341.

[Birger et al., 2019] Birger, Y., Shiloh, R., and Izraeli, S. (2019). Mechanisms of Leukemia Evolution: Lessons from a Congenital Syndrome. *Cancer Cell*, 36(2):115–117.

[Blank and Karlsson, 2015] Blank, U. and Karlsson, S. (2015). TGF-B signaling in the control of hematopoietic stem cells. *Blood*, 125(23):3542–3550.

[Bohnstedt et al., 2013] Bohnstedt, C., Levinsen, M., Rosthøj, S., Zeller, B., Taskinen, M., Hafsteinsdottir, S., Björgvinsdóttir, H., Heyman, M., and Schmiegelow, K. (2013). Physicians compliance during maintenance therapy in children with Down syndrome and acute lymphoblastic leukemia. *Leukemia*, 27(4):866–870.

[Boiers et al., 2018] Boiers, C., Richardson, S. E., Laycock, E., Zriwil, A., Turati, V. A., Brown, J., Wray, J. P., Wang, D., James, C., Herrero, J., Sitnicka, E., Karlsson, S., Smith, A. J. H., Jacobsen, S. E. W., and Enver, T. (2018). A Human IPS Model Implicates Embryonic B-Myeloid Fate Restriction as Developmental Susceptibility to B Acute Lymphoblastic Leukemia-Associated ETV6-RUNX1. *Developmental Cell*, 44(3):362–377.e7.

[Boller and Grosschedl, 2014] Boller, S. and Grosschedl, R. (2014). The regulatory network of B-cell differentiation: a focused view of early B-cell factor 1 function. *Immunological Reviews*, 261(1):102–115.

[Brown et al., 2019] Brown, A. L., Smith, A. J. d., Gant, V. U., Yang, W., Scheurer, M. E., Walsh, K. M., Chernus, J. M., Kallsen, N. A., Peyton, S. A., Davies, G. E., Ehli, E. A., Winick, N., Heerema, N. A., Carroll, A. J., Borowitz, M. J., Wood, B. L., Carroll, W. L., Raetz, E. A., Feingold, E., Devidas, M., Barcellos, L. F., Hansen, H. M., Morimoto, L., Kang, A. Y., Smirnov, I., Healy, J., Laverdière, C., Sinnett, D., Taub, J. W., Birch, J. M., Thompson, P., Spector, L. G., Pombo-de Oliveira, M. S., DeWan, A. T., Mullighan, C. G., Hunger, S. P., Pui, C.-H., Loh, M. L., Zwick, M. E., Metayer, C., Ma, X., Mueller, B. A., Sherman, S. L., Wiemels, J. L., Relling, M. V., Yang, J. J., Lupo, P. J., and Rabin, K. R. (2019). Inherited genetic susceptibility of acute lymphoblastic leukemia in Down syndrome. *Blood*, page blood.2018890764.

[Bueno et al., 2016] Bueno, C., van Roon, E. H. J., Muñoz-López, A., Sanjuan-Pla, A., Juan, M., Navarro, A., Stam, R. W., and Menendez, P. (2016). Immunophenotypic analysis and quantification of B-1 and B-2 B cells during human fetal hematopoietic development. *Leukemia*.

[Buenrostro et al., 2015] Buenrostro, J., Wu, B., Chang, H., and Greenleaf, W. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]*, 109:21.29.1–21.29.9.

[Buitenkamp et al., 2014] Buitenkamp, T. D., Izraeli, S., Zimmermann, M., Forestier, E., Heerema, N. A., Heuvel-Eibrink, M. M. v. d., Pieters, R., Korbijn, C. M., Silverman, L. B., Schmiegelow, K., Liang, D.-C., Horibe, K., Arico, M., Biondi, A., Basso, G., Rabin, K. R., Schrappe, M., Cario, G., Mann, G., Morak, M., Panzer-Grümayer, R., Mondelaers, V., Lammens, T., Cavé, H., Stark, B., Ganmore, I., Moorman, A. V., Vora, A., Hunger, S. P., Pui, C.-H., Mullighan, C. G., Manabe, A., Escherich, G., Kowalczyk, J. R., Whitlock, J. A., and Zwaan, C. M. (2014). Acute lymphoblastic leukemia in children with Down syndrome: a retrospective analysis from the Ponte di Legno study group. *Blood*, 123(1):70–77.

[Butler et al., 2018] Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*, 36(5):411–420.

[Carrelha et al., 2018] Carrelha, J., Meng, Y., Kettyle, L. M., Luis, T. C., Norfo, R., Alcolea, V., Boukarabila, H., Grasso, F., Gambardella, A., Grover, A., Högstrand, K., Lord, A. M., Sanjuan-Pla, A., Woll, P. S., Nerlov, C., and Jacobsen, S. E. W. (2018). Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature*, 554(7690):106–111.

[Carsetti et al., 2015] Carsetti, R., Valentini, D., Marcellini, V., Scarsella, M., Marasco, E., Giustini, F., Bartuli, A., Villani, A., and Ugazio, A. G. (2015). Reduced numbers of switched memory B cells with high terminal differentiation potential in Down syndrome. *European Journal of Immunology*, 45(3):903–914.

[Cazzaniga et al., 2011] Cazzaniga, G., Delft, F. W. v., Nigro, L. L., Ford, A. M., Score, J., Iacobucci, I., Mirabile, E., Taj, M., Colman, S. M., Biondi, A., and Greaves, M. (2011). Developmental origins and impact of BCR-ABL1 fusion and IKZF1 deletions in monozygotic twins with Ph+ acute lymphoblastic leukemia. *Blood*, 118(20):5559–5564.

[Challen et al., 2010] Challen, G. A., Boles, N. C., Chambers, S. M., and Goodell, M. A. (2010). Distinct Hematopoietic Stem Cell Subtypes Are Differentially Regulated by TGF-B1. *Cell Stem Cell*, 6(3):265–278.

[Charbord et al., 1996] Charbord, P., Tavian, M., Humeau, L., and Peault, B. (1996). Early ontogeny of the human marrow from long bones: an immunohistochemical study of hematopoiesis and its microenvironment [see comments]. *Blood*, 87(10):4109–4119.

[Chen et al., 2014] Chen, L., Kostadima, M., Martens, J. H. A., Canu, G., Garcia, S. P., Turro, E., Downes, K., Macaulay, I. C., Bielczyk-Maczynska, E., Coe, S., Farrow, S., Poudel, P., Burden, F., Jansen, S. B. G., Astle, W. J., Attwood, A., Bariana, T., Bono, B. d., Breschi, A., Chambers, J. C., Consortium, B., Choudry, F. A., Clarke, L., Coupland, P., Ent, M. v. d., Erber, W. N., Jansen, J. H., Favier, R., Fenech, M. E., Foad, N., Freson, K., Geet, C. v., Gomez, K., Guigo, R., Hampshire, D., Kelly, A. M., Kerstens, H. H. D., Kooner, J. S., Laffan, M., Lentaigne, C., Labalette, C., Martin, T., Meacham, S., Mumford, A., Nürnberg, S. T., Palumbo, E., Reijden, B. A. v. d., Richardson, D., Sammut, S. J., Slodkowicz, G., Tamuri, A. U., Vasquez, L., Voss, K., Watt, S., Westbury, S., Flicek, P., Loos, R., Goldman, N., Bertone, P., Read, R. J., Richardson, S., Cvejic, A., Soranzo, N., Ouwehand, W. H., Stunnenberg, H. G., Frontini, M., and Rendon, A. (2014).

Transcriptional diversity during lineage commitment of human blood progenitors. *Science*, 345(6204):1251033.

[Chiang et al., 2018] Chiang, J.-C., Jiang, J., Newburger, P. E., and Lawrence, J. B. (2018). Trisomy silencing by XIST normalizes Down syndrome cell pathogenesis demonstrated for hematopoietic defects in vitro. *Nature Communications*, 9(1):5180.

[Chou et al., 2012] Chou, S. T., Byrska-Bishop, M., Tober, J. M., Yao, Y., VanDorn, D., Opalinska, J. B., Mills, J. A., Choi, J. K., Speck, N. A., Gadue, P., Hardison, R. C., Nemiroff, R. L., French, D. L., and Weiss, M. J. (2012). Trisomy 21-associated defects in human primitive hematopoiesis revealed through induced pluripotent stem cells. *Proceedings of the National Academy of Sciences*, 109(43):17573–17578.

[Chou et al., 2008] Chou, S. T., Opalinska, J. B., Yao, Y., Fernandes, M. A., Kalota, A., Brooks, J. S. J., Choi, J. K., Gewirtz, A. M., Danet-Desnoyers, G.-a., Nemiroff, R. L., and Weiss, M. J. (2008). Trisomy 21 enhances human fetal erythro-megakaryocytic development. *Blood*, 112(12):4503–4506.

[Cocchi et al., 2007] Cocchi, G., Mastrocola, M., Capelli, M., Bastelli, A., Vitali, F., and Corvaglia, L. (2007). Immunological patterns in young children with Down syndrome: is there a temporal trend? *Acta Paediatrica*, 96(10):1479–1482.

[Corces et al., 2016] Corces, M. R., Buenrostro, J. D., Wu, B., Greenside, P. G., Chan, S. M., Koenig, J. L., Snyder, M. P., Pritchard, J. K., Kundaje, A., Greenleaf, W. J., Majeti, R., and Chang, H. Y. (2016). Lineage-specific and single cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nature genetics*, 48(10):1193–1203.

[Coskun et al., 2014] Coskun, S., Chao, H., Vasavada, H., Heydari, K., Gonzales, N., Zhou, X., de Crombrugghe, B., and Hirschi, K. K. (2014). Development of the fetal bone marrow niche and regulation of HSC quiescence and homing ability by emerging osteolineage cells. *Cell reports*, 9(2):581–590.

[Costa et al., 2012] Costa, G., Kouskoff, V., and Lacaud, G. (2012). Origin of blood cells and HSC production in the embryo. *Trends in Immunology*, 33(5):215–223.

[Cuadrado and Barrena, 1996] Cuadrado, E. and Barrena, M. J. (1996). Immune Dysfunction in Down's Syndrome: Primary Immune Deficiency or Early Senescence of the Immune System? *Clinical Immunology and Immunopathology*, 78(3):209–214.

[Cuartero et al., 2018] Cuartero, S., Weiss, F. D., Dharmalingam, G., Guo, Y., Ing-Simmons, E., Masella, S., Robles-Rebollo, I., Xiao, X., Wang, Y.-F., Barozzi, I., Djeghloul, D., Amano, M. T., Niskanen, H., Petretto, E., Dowell, R. D., Tachibana, K., Kaikkonen, M. U., Nasmyth, K. A., Lenhard, B., Natoli, G., Fisher, A. G., and Merkenschlager, M. (2018). Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nature Immunology*, 19(9):932–941.

[Cvejic, 2015] Cvejic, A. (2015). Mechanisms of fate decision and lineage commitment during haematopoiesis. *Immunology and Cell Biology*.

[Davisson et al., 1990] Davisson, M. T., Schmidt, C., and Akeson, E. C. (1990). Segmental trisomy of murine chromosome 16: a new model system for studying Down syndrome. *Progress in Clinical and Biological Research*, 360:263–280.

[de Bruijn et al., 2000] de Bruijn, M. F., Speck, N. A., Peeters, M. C., and Dzierzak, E. (2000). Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *The EMBO journal*, 19(11):2465–2474.

[de Bruijn et al., 2002] de Bruijn, M. F. T. R., Ma, X., Robin, C., Ottersbach, K., Sanchez, M.-J., and Dzierzak, E. (2002). Hematopoietic Stem Cells Localize to the Endothelial Cell Layer in the Midgestation Mouse Aorta. *Immunity*, 16(5):673–683.

[de Hingh et al., 2005] de Hingh, Y. C. M., van der Vossen, P. W., Gemen, E. F. A., Mulder, A. B., Hop, W. C. J., Brus, F., and de Vries, E. (2005). Intrinsic Abnormalities of Lymphocyte Counts in Children with Down Syndrome. *The Journal of Pediatrics*, 147(6):744–747.

[Derynck and Budi, 2019] Derynck, R. and Budi, E. H. (2019). Specificity, versatility, and control of TGF-B family signaling. *Sci. Signal.*, 12(570):eaav5183.

[Descatoire et al., 2011] Descatoire, M., Weill, J.-C., Reynaud, C.-A., and Weller, S. (2011). A human equivalent of mouse B-1 cells? *The Journal of Experimental Medicine*, 208(13):2563–2564.

[Diao et al., 2012] Diao, Y., Guo, X., Li, Y., Sun, K., Lu, L., Jiang, L., Fu, X., Zhu, H., Sun, H., Wang, H., and Wu, Z. (2012). Pax3/7bp Is a Pax7- and Pax3-Binding Protein that Regulates the Proliferation of Muscle Precursor Cells by an Epigenetic Mechanism. *Cell Stem Cell*, 11(2):231–241.

[Ding et al., 2016] Ding, M., Bruick, R. K., and Yu, Y. (2016). Secreted IGFBP5 mediates mTORC1-dependent feedback inhibition of IGF-1 signalling. *Nature Cell Biology*, 18(3):319–327.

[Dobin et al., 2013] Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1):15–21.

[Dominici et al., 2006] Dominici, M., Le Blanc, K., Mueller, I., Slaper-Cortenbach, I., Marini, F., Krause, D., Deans, R., Keating, A., Prockop, D., and Horwitz, E. (2006). Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy*, 8(4):315–317.

[Dongen et al., 2003] Dongen, J. J. M. v., Langerak, A. W., Brüggemann, M., Evans, P. a. S., Hummel, M., Lavender, F. L., Delabesse, E., Davi, F., Schuuring, E., García-Sanz, R., Krieken, J. H. J. M. v., Droese, J., González, D., Bastard, C., White, H. E., Spaargaren, M., González, M., Parreira, A., Smith, J. L., Morgan, G. J., Kneba, M., and Macintyre, E. A. (2003). Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: Report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*, 17(12):2257.

[Dorsett et al., 2008] Dorsett, Y., McBride, K. M., Jankovic, M., Gazumyan, A., Thai, T.-H., Robbiani, D. F., Di Virgilio, M., San-Martin, B. R., Heidkamp, G., Schwickert, T. A., Eisenreich, T., Rajewsky, K., and Nussenzweig, M. C. (2008). MicroRNA-155 Suppresses Activation-Induced Cytidine Deaminase-Mediated Myc-Igh Translocation. *Immunity*, 28(5):630–638.

[Doulatov et al., 2010] Doulatov, S., Notta, F., Eppert, K., Nguyen, L. T., Ohashi, P. S., and Dick, J. E. (2010). Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. *Nature Immunology*, 11(7):585–593.

[Down, 1866] Down, J. L. H. (1866). Observations on an ethnic classification of idiots. *Heredity*, 21(4):695–697.

[Drissen et al., 2016] Drissen, R., Buza-Vidas, N., Woll, P., Thongjuea, S., Gambardella, A., Giustacchini, A., Mancini, E., Zriwil, A., Lutteropp, M., Grover,

A., Mead, A., Sitnicka, E., Jacobsen, S. E. W., and Nerlov, C. (2016). Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nature Immunology*, 17(6):666–676.

[Drissen et al., 2019] Drissen, R., Thongjuea, S., Theilgaard-Mönch, K., and Nerlov, C. (2019). Identification of two distinct pathways of human myelopoiesis. *Science Immunology*, 4(35):eaau7148.

[Dzierzak and Bigas, 2018] Dzierzak, E. and Bigas, A. (2018). Blood Development: Hematopoietic Stem Cell Dependence and Independence. *Cell Stem Cell*, 22(5):639–651.

[Easterbrook et al., 2019] Easterbrook, J., Rybtsov, S., Gordon-Keylock, S., Ivanovs, A., Taoudi, S., Anderson, R. A., and Medvinsky, A. (2019). Analysis of the Spatiotemporal Development of Hematopoietic Stem and Progenitor Cells in the Early Human Embryo. *Stem Cell Reports*, 12(5):1056–1068.

[Eaves, 2015] Eaves, C. J. (2015). Hematopoietic stem cells: concepts, definitions and the new reality. *Blood*, pages blood–2014–12–570200.

[Emmrich et al., 2014] Emmrich, S., Rasche, M., Schöning, J., Reimer, C., Keihani, S., Maroz, A., Xie, Y., Li, Z., Schambach, A., Reinhardt, D., and Klusmann, J.-H. (2014). miR-99a/100-125b tricistrons regulate hematopoietic stem and progenitor cell homeostasis by shifting the balance between TGFB and Wnt signaling. *Genes & Development*, 28(8):858–874.

[Englund et al., 2013] Englund, A., Jonsson, B., Zander, C. S., Gustafsson, J., and Annerén, G. (2013). Changes in mortality and causes of death in the Swedish Down syndrome population. *American Journal of Medical Genetics Part A*, 161(4):642–649.

[Ensembl, 2019] Ensembl (2019). Chromosome 21: 1-46,709,983 - Chromosome summary - Homo sapiens - Ensembl genome browser 97. Assembly GRCh38.P12.

[Enver and Greaves, 1998] Enver, T. and Greaves, M. (1998). Loops, Lineage, and Leukemia. *Cell*, 94(1):9–12.

[Espin-Palazon et al., 2014] Espin-Palazon, R., Stachura, D. L., Campbell, C. A., Garcia-Moreno, D., Del Cid, N., Kim, A. D., Candel, S., Meseguer, J., Mulero, V., and Traver, D. (2014). Proinflammatory signaling regulates hematopoietic stem cell emergence. *Cell*, 159(5):1070–1085.

[Essers et al., 2009] Essers, M. A. G., Offner, S., Blanco-Bose, W. E., Waibler, Z., Kalinke, U., Duchosal, M. A., and Trumpp, A. (2009). IFNa activates dormant haematopoietic stem cells in vivo. *Nature*, 458(7240):904–908.

[Farroni et al., 2018] Farroni, C., Marasco, E., Marcellini, V., Giorda, E., Valentini, D., Petrini, S., D'Oria, V., Pezzullo, M., Cascioli, S., Scarsella, M., Ugazio, A. G., De Vincentiis, G. C., Grimsholm, O., and Carsetti, R. (2018). Dysregulated miR-155 and miR-125b Are Related to Impaired B-cell Responses in Down Syndrome. *Frontiers in Immunology*, 9.

[Fasching et al., 2000] Fasching, K., Panzer, S., Haas, O. A., Marschalek, R., Gadner, H., and Panzer-Grümayer, E. R. (2000). Presence of clone-specific antigen receptor gene rearrangements at birth indicates an in utero origin of diverse types of early childhood acute lymphoblastic leukemia. *Blood*, 95(8):2722–2724.

[Feng et al., 2012] Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X. S. (2012). Identifying ChIP-seq enrichment using MACS. *Nature protocols*, 7(9).

[Ford et al., 1993] Ford, A. M., Ridge, S. A., Cabrera, M. E., Mahmoud, H., Steel, C. M., Chan, L. C., and Greaves, M. (1993). In utero rearrangements in the trithorax-related oncogene in infant leukaemias. *Nature*, 363(6427):358–360.

[Forestier et al., 2008] Forestier, E., Izraeli, S., Beverloo, B., Haas, O., Pession, A., Michalová, K., Stark, B., Harrison, C. J., Teigler-Schlegel, A., and Johansson, B. (2008). Cytogenetic features of acute lymphoblastic and myeloid leukemias in pediatric patients with Down syndrome: an iBFM-SG study. *Blood*, 111(3):1575–1583.

[Forte et al., 2019] Forte, D., Krause, D. S., Andreeff, M., Bonnet, D., and Méndez-Ferrer, S. (2019). Updates on the hematological tumor microenvironment and its therapeutic targeting. *Haematologica*, page haematol.2018.195396.

[Franciotta et al., 2006] Franciotta, D., Verri, A., Zardini, E., Andreoni, L., De Amici, M., Moratti, R., and Nespoli, L. (2006). Interferon-gamma- and interleukin-4-producing T cells in Down's syndrome. *Neuroscience Letters*, 395(1):67–70.

[Francois-Newton et al., 2011] Francois-Newton, V., Almeida, G. M. d. F., Payelle-Brogard, B., Monneron, D., Pichard-Garcia, L., Piehler, J., Pellegrini, S., and Uze, G. (2011). USP18-Based Negative Feedback Control Is Induced by Type I and

Type III Interferons and Specifically Inactivates Interferon a Response. *PLOS ONE*, 6(7):e22200.

[Gale et al., 1997] Gale, K. B., Ford, A. M., Repp, R., Borkhardt, A., Keller, C., Eden, O. B., and Greaves, M. F. (1997). Backtracking leukemia to birth: Identification of clonotypic gene fusion sequences in neonatal blood spots. *Proceedings of the National Academy of Sciences*, 94(25):13950–13954.

[Galy et al., 1995] Galy, A., Travis, M., Cen, D., and Chen, B. (1995). Human T, B, natural killer, and dendritic cells arise from a common bone marrow progenitor cell subset. *Immunity*, 3(4):459–473.

[Gamis et al., 2003] Gamis, A. S., Woods, W. G., Alonzo, T. A., Buxton, A., Lange, B., Barnard, D. R., Gold, S., Smith, F. O., and Children's Cancer Group Study 2891 (2003). Increased age at diagnosis has a significantly negative effect on outcome in children with Down syndrome and acute myeloid leukemia: a report from the Children's Cancer Group Study 2891. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 21(18):3415–3422.

[Gao et al., 2018] Gao, L., Tober, J., Gao, P., Chen, C., Tan, K., and Speck, N. A. (2018). RUNX1 and the endothelial origin of blood. *Experimental Hematology*, 68:2–9.

[Garrison et al., 2005] Garrison, M. M., Jeffries, H., and Christakis, D. A. (2005). Risk of Death for Children with Down Syndrome and Sepsis. *The Journal of Pediatrics*, 147(6):748–752.

[Ghosn et al., 2019] Ghosn, E., Yoshimoto, M., Nakauchi, H., Weissman, I. L., and Herzenberg, L. A. (2019). Hematopoietic stem cell-independent hematopoiesis and the origins of innate-like B lymphocytes. *Development*, 146(15):dev170571.

[Giliani et al., 2005] Giliani, S., Mori, L., Basile, G. D. S., Deist, F. L., Rodriguez-Perez, C., Forino, C., Mazzolari, E., Dupuis, S., Elhasid, R., Kessel, A., Galambrun, C., Gil, J., Fischer, A., Etzioni, A., and Notarangelo, L. D. (2005). Interleukin-7 receptor a (IL-7ra) deficiency: cellular and molecular bases. Analysis of clinical, immunological, and molecular features in 16 novel patients. *Immunological Reviews*, 203(1):110–126.

[Gillespie et al., 2006] Gillespie, K. M., Dix, R. J., Williams, A. J. K., Newton, R., Robinson, Z. F., Bingley, P. J., Gale, E. A. M., and Shield, J. P. H. (2006). Islet Autoimmunity in Children With Down's Syndrome. *Diabetes*, 55(11):3185–3188.

[Giustacchini et al., 2017] Giustacchini, A., Thongjuea, S., Barkas, N., Woll, P. S., Povinelli, B. J., Booth, C. A. G., Sopp, P., Norfo, R., Rodriguez-Meira, A., Ashley, N., Jamieson, L., Vyas, P., Anderson, K., Segerstolpe, A., Qian, H., Olsson-Strömberg, U., Mustjoki, S., Sandberg, R., Jacobsen, S. E. W., and Mead, A. J. (2017). Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nature Medicine*, 23(6):692–702.

[Goardon et al., 2011] Goardon, N., Marchi, E., Atzberger, A., Quek, L., Schuh, A., Soneji, S., Woll, P., Mead, A., Alford, K. A., Rout, R., Chaudhury, S., Gilkes, A., Knapper, S., Beldjord, K., Begum, S., Rose, S., Geddes, N., Griffiths, M., Standen, G., Sternberg, A., Cavenagh, J., Hunter, H., Bowen, D., Killick, S., Robinson, L., Price, A., Macintyre, E., Virgo, P., Burnett, A., Craddock, C., Enver, T., Jacobsen, S. E. W., Porcher, C., and Vyas, P. (2011). Coexistence of LMPP-like and GMP-like Leukemia Stem Cells in Acute Myeloid Leukemia. *Cancer Cell*, 19(1):138–152.

[Goedhart et al., 2018] Goedhart, M., Cornelissen, A., Kuijk, C., Geerman, S., Kleijer, M., van Buul, J. D., Huveneers, S., Raaijmakers, M., Young, H. A., Wolkers, M. C., Voermans, C., and Nolte, M. A. (2018). Interferon-gamma Impairs Maintenance and Alters Hematopoietic Support of Bone Marrow Mesenchymal Stromal Cells. *Stem Cells and Development*.

[Goldacre et al., 2004] Goldacre, M. J., Wotton, C. J., Seagroatt, V., and Yeates, D. (2004). Cancers and immune related diseases associated with Down's syndrome: a record linkage study. *Archives of Disease in Childhood*, 89(11):1014–1017.

[Good et al., 2018] Good, Z., Sarno, J., Jager, A., Samusik, N., Aghaeepour, N., Simonds, E. F., White, L., Lacayo, N. J., Fantl, W. J., Fazio, G., Gaipa, G., Biondi, A., Tibshirani, R., Bendall, S. C., Nolan, G. P., and Davis, K. L. (2018). Single-cell developmental classification of B cell precursor acute lymphoblastic leukemia at diagnosis reveals predictors of relapse. *Nature Medicine*, 24(4):474–483.

[Greaves, 2018] Greaves, M. (2018). A causal mechanism for childhood acute lymphoblastic leukaemia. *Nature Reviews Cancer*, page 1.

[Greaves et al., 2003] Greaves, M. F., Maia, A. T., Wiemels, J. L., and Ford, A. M. (2003). Leukemia in twins: lessons in natural history. *Blood*, 102(7):2321–2333.

[Greten and Grivennikov, 2019] Greten, F. R. and Grivennikov, S. I. (2019). Inflammation and Cancer: Triggers, Mechanisms, and Consequences. *Immunity*, 51(1):27–41.

[Gribble et al., 2013] Gribble, S. M., Wiseman, F. K., Clayton, S., Prigmore, E., Langley, E., Yang, F., Maguire, S., Fu, B., Rajan, D., Sheppard, O., Scott, C., Hauser, H., Stephens, P. J., Stebbings, L. A., Ng, B. L., Fitzgerald, T., Quail, M. A., Banerjee, R., Rothkamm, K., Tybulewicz, V. L. J., Fisher, E. M. C., and Carter, N. P. (2013). Massively Parallel Sequencing Reveals the Complex Structure of an Irradiated Human Chromosome on a Mouse Background in the Tc1 Model of Down Syndrome. *PLOS ONE*, 8(4):e60482.

[Griffin et al., 2011a] Griffin, D. O., Holodick, N. E., and Rothstein, T. L. (2011a). Human B1 cells are CD3-: A reply to "A human equivalent of mouse B-1 cells\?" and "The nature of circulating CD27+CD43+ B cells". *The Journal of Experimental Medicine*, 208(13):2566–2569.

[Griffin et al., 2011b] Griffin, D. O., Holodick, N. E., and Rothstein, T. L. (2011b). Human B1 cells in umbilical cord and adult peripheral blood express the novel phenotype CD20+CD27+CD43+CD70-. *The Journal of Experimental Medicine*, 208(1):67–80.

[Gropp, 1974] Gropp, A. (1974). Animal model of human disease. Autosomal trisomy, developmental impairment and fetal death. *The American Journal of Pathology*, 77(3):539–542.

[Grumayer et al., 1991] Grumayer, E. R., Griesinger, F., Hummell, D. S., Brunning, R. D., and Kersey, J. H. (1991). Identification of novel B-lineage cells in human fetal bone marrow that coexpress CD7 [see comments]. *Blood*, 77(1):64–68.

[Gu et al., 2016a] Gu, Y., Jones, A. E., Yang, W., Liu, S., Dai, Q., Liu, Y., Swindle, C. S., Zhou, D., Zhang, Z., Ryan, T. M., Townes, T. M., Klug, C. A., Chen, D., and Wang, H. (2016a). The histone H2a deubiquitinase Usp16 regulates hematopoiesis and hematopoietic stem cell function. *Proceedings of the National Academy of Sciences*, 113(1):E51–E60.

[Gu et al., 2016b] Gu, Y., Masiero, M., and Banham, A. H. (2016b). Notch signaling: its roles and therapeutic potential in hematological malignancies. *Oncotarget*, 7(20):29804–29823.

[Guard et al., 2019] Guard, S. E., Poss, Z. C., Ebmeier, C. C., Pagratis, M., Simpson, H., Taatjes, D. J., and Old, W. M. (2019). The nuclear interactome of DYRK1a reveals a functional role in DNA damage repair. *Scientific Reports*, 9(1):6539.

[Haas et al., 2018] Haas, S., Trumpp, A., and Milsom, M. D. (2018). Causes and Consequences of Hematopoietic Stem Cell Heterogeneity. *Cell Stem Cell*, 22(5):627–638.

[Haddad et al., 2004] Haddad, R., Guardiola, P., Izac, B., Thibault, C., Radich, J., Delezoide, A.-L., Baillou, C., Lemoine, F. M., Gluckman, J. C., Pflumio, F., and Canque, B. (2004). Molecular characterization of early human T/NK and B-lymphoid progenitor cells in umbilical cord blood. *Blood*, 104(13):3918–3926.

[Hadjur et al., 2001] Hadjur, S., Ung, K., Wadsworth, L., Dimmick, J., Rajcan-Separovic, E., Scott, R. W., Buchwald, M., and Jirik, F. R. (2001). Defective hematopoiesis and hepatic steatosis in mice with combined deficiencies of the genes encoding Fancc and Cu/Zn superoxide dismutase. *Blood*, 98(4):1003–1011.

[Haghverdi et al., 2016] Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F., and Theis, F. J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nature Methods*, 13(10):845–848.

[Hao et al., 2001] Hao, Q.-L., Zhu, J., Price, M. A., Payne, K. J., Barsky, L. W., and Crooks, G. M. (2001). Identification of a novel, human multilymphoid progenitor in cord blood. *Blood*, 97(12):3683–3690.

[Hardy et al., 1991] Hardy, R. R., Carmack, C. E., Shinton, S. A., Kemp, J. D., and Hayakawa, K. (1991). Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *Journal of Experimental Medicine*, 173(5):1213–1225.

[Hardy and Hayakawa, 1991] Hardy, R. R. and Hayakawa, K. (1991). A developmental switch in B lymphopoiesis. *Proceedings of the National Academy of Sciences of the United States of America*, 88(24):11550–11554.

[Hasle, 2001] Hasle, H. (2001). Pattern of malignant disorders in individuals with Down's syndrome. *The Lancet Oncology*, 2(7):429–436.

[Hasle et al., 2000] Hasle, H., Clemmensen, I. H., and Mikkelsen, M. (2000). Risks of leukaemia and solid tumours in individuals with Down's syndrome. *The Lancet*, 355(9199):165–169.

[Hasle et al., 2016] Hasle, H., Friedman, J. M., Olsen, J. H., and Rasmussen, S. A. (2016). Low risk of solid tumors in persons with Down syndrome. *Genetics in Medicine*, 18(11):1151–1157.

[He et al., 2015] He, Q., Zhang, C., Wang, L., Zhang, P., Ma, D., Lv, J., and Liu, F. (2015). Inflammatory signaling regulates hematopoietic stem and progenitor cell emergence in vertebrates. *Blood*, 125(7):1098–1106.

[Heinz et al., 2010] Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell*, 38(4):576–589.

[Hemminki et al., 2013] Hemminki, K., Liu, X., Försti, A., Ji, J., Sundquist, J., and Sundquist, K. (2013). Subsequent leukaemia in autoimmune disease patients. *British Journal of Haematology*, 161(5):677–687.

[Herault et al., 2017] Herault, Y., Delabar, J. M., Fisher, E. M. C., Tybulewicz, V. L. J., Yu, E., and Brault, V. (2017). Rodent models in Down syndrome research: impact and future opportunities. *Disease Models & Mechanisms*, 10(10):1165–1186.

[Hern, 1984] Hern, W. M. (1984). Correlation of fetal age and measurements between 10 and 26 weeks of gestation. *Obstetrics and Gynecology*, 63(1):26–32.

[Hertzberg et al., 2010] Hertzberg, L., Vendramini, E., Ganmore, I., Cazzaniga, G., Schmitz, M., Chalker, J., Shiloh, R., Iacobucci, I., Shochat, C., Zeligson, S., Cario, G., Stanulla, M., Strehl, S., Russell, L. J., Harrison, C. J., Bornhauser, B., Yoda, A., Rechavi, G., Bercovich, D., Borkhardt, A., Kempski, H., Kronnie, G. t., Bourquin, J.-P., Domany, E., and Izraeli, S. (2010). Down syndrome acute lymphoblastic leukemia, a highly heterogeneous disease in which aberrant expression of CRLF2 is associated with mutated JAK2: a report from the International BFM Study Group. *Blood*, 115(5):1006–1017.

[Hirano et al., 1986] Hirano, T., Yasukawa, K., Harada, H., Taga, T., Watanabe, Y., Matsuda, T., Kashiwamura, S.-i., Nakajima, K., Koyama, K., Iwamatsu, A., Tsunasawa, S., Sakiyama, F., Matsui, H., Takahara, Y., Taniguchi, T., and Kishimoto, T. (1986). Complementary DNA for a novel human interleukin (BSF-2) that induces B lymphocytes to produce immunoglobulin. *Nature*, 324(6092):73.

[Hirata et al., 1989] Hirata, Y., Taga, T., Hibi, M., Nakano, N., Hirano, T., and Kishimoto, T. (1989). Characterization of IL-6 receptor expression by monoclonal and polyclonal antibodies. *The Journal of Immunology*, 143(9):2900–2906.

[Hoebeke et al., 2007] Hoebeke, I., De Smedt, M., Stolz, F., Pike-Overzet, K., Staal, F. J. T., Plum, J., and Leclercq, G. (2007). T-, B- and NK-lymphoid, but not myeloid cells arise from human CD34+CD38-CD7+ common lymphoid progenitors expressing lymphoid-specific genes. *Leukemia*, 21(2):311–319.

[Hofer and Rodewald, 2018] Hofer, T. and Rodewald, H.-R. (2018). Differentiation-based model of hematopoietic stem cell functions and lineage pathways. *Blood*, 132(11):1106–1113.

[Hsin et al., 2018] Hsin, J.-P., Lu, Y., Loeb, G. B., Leslie, C. S., and Rudensky, A. Y. (2018). The effect of cellular context on miR-155-mediated gene regulation in four major immune cell types. *Nature Immunology*, 19(10):1137–1145.

[Huret et al., 2013] Huret, J.-L., Ahmad, M., Arsaban, M., Bernheim, A., Cigna, J., Desangles, F., Guignard, J.-C., Jacquemot-Perbal, M.-C., Labarussias, M., Leberre, V., Malo, A., Morel-Pair, C., Mossafa, H., Potier, J.-C., Texier, G., Viguie, F., Yau Chun Wan-Senon, S., Zasadzinski, A., and Dessen, P. (2013). Atlas of Genetics and Cytogenetics in Oncology and Haematology in 2013. *Nucleic Acids Research*, 41(D1):D920–D924.

[Huyhn et al., 1995] Huyhn, A., Dommergues, M., Izac, B., Croisille, L., Katz, A., Vainchenker, W., and Coulombel, L. (1995). Characterization of hematopoietic progenitors from human yolk sacs and embryos. *Blood*, 86(12):4474–4485.

[Hystad et al., 2007] Hystad, M. E., Myklebust, J. H., Bø, T. H., Sivertsen, E. A., Rian, E., Forfang, L., Munthe, E., Rosenwald, A., Chiorazzi, M., Jonassen, I., Staudt, L. M., and Smeland, E. B. (2007). Characterization of Early Stages of Human B Cell Development by Gene Expression Profiling. *The Journal of Immunology*, 179(6):3662–3671.

[Ichii et al., 2014] Ichii, M., Oritani, K., and Kanakura, Y. (2014). Early B lymphocyte development: Similarities and differences in human and mouse. *World Journal of Stem Cells*, 6(4):421–431.

[Ichii et al., 2010a] Ichii, M., Oritani, K., Yokota, T., Kanakura, Y., and Kincade, P. W. (2010a). Stromal Cell-Free Conditions Favorable for Human B Lymphopoiesis in Culture. *Journal of immunological methods*, 359(1-2):47–55.

[Ichii et al., 2008] Ichii, M., Oritani, K., Yokota, T., Nishida, M., Takahashi, I., Shirogane, T., Ezoe, S., Saitoh, N., Tanigawa, R., Kincade, P. W., and Kanakura, Y. (2008). Regulation of human B lymphopoiesis by the transforming growth factor-B superfamily in a newly established coculture system using human mesenchymal stem cells as a supportive microenvironment. *Experimental Hematology*, 36(5):587–597.

[Ichii et al., 2010b] Ichii, M., Oritani, K., Yokota, T., Zhang, Q., Garrett, K. P., Kanakura, Y., and Kincade, P. W. (2010b). The Density of CD10 Corresponds to Commitment and Progression in the Human B Lymphoid Lineage. *PLOS ONE*, 5(9):e12954.

[Iliopoulos et al., 2009] Iliopoulos, D., Hirsch, H. A., and Struhl, K. (2009). An Epigenetic Switch Involving NF-kB, Lin28, Let-7 MicroRNA, and IL6 Links Inflammation to Cell Transformation. *Cell*, 139(4):693–706.

[Itoh et al., 1989] Itoh, K., Tezuka, H., Sakoda, H., Konno, M., Nagata, K., Uchiyama, T., Uchino, H., and Mori, K. J. (1989). Reproducible establishment of hemopoietic supportive stromal cell lines from murine bone marrow. *Experimental hematology*, 17(2):145–153.

[Itoh-Nakadai et al., 2014] Itoh-Nakadai, A., Hikota, R., Muto, A., Kometani, K., Watanabe-Matsui, M., Sato, Y., Kobayashi, M., Nakamura, A., Miura, Y., Yano, Y., Tashiro, S., Sun, J., Ikawa, T., Ochiai, K., Kurosaki, T., and Igarashi, K. (2014). The transcription repressors Bach2 and Bach1 promote B cell development by repressing the myeloid program. *Nature Immunology*, 15(12):1171–1180.

[Ivanovs et al., 2017] Ivanovs, A., Rybtsov, S., Ng, E. S., Stanley, E. G., Elefanty, A. G., and Medvinsky, A. (2017). Human haematopoietic stem cell development: from the embryo to the dish. *Development*, 144(13):2323–2337.

[Ivanovs et al., 2011] Ivanovs, A., Rybtsov, S., Welch, L., Anderson, R. A., Turner, M. L., and Medvinsky, A. (2011). Highly potent human hematopoietic stem cells first emerge in the intraembryonic aorta-gonad-mesonephros region. *The Journal of Experimental Medicine*, 208(12):2417–2427.

[Izraeli et al., 2007] Izraeli, S., Rainis, L., Hertzberg, L., Smooha, G., and Birger, Y. (2007). Trisomy of chromosome 21 in leukemogenesis. *Blood Cells, Molecules, and Diseases*, 39(2):156–159.

[Jablonski et al., 2016] Jablonski, K. A., Gaudet, A. D., Amici, S. A., Popovich, P. G., and Guerau-de Arellano, M. (2016). Control of the Inflammatory Macrophage Transcriptional Signature by miR-155. *PLOS ONE*, 11(7):e0159724.

[Jacobsen and Nerlov, 2019] Jacobsen, S. E. W. and Nerlov, C. (2019). Haematopoiesis in the era of advanced single-cell technologies. *Nature Cell Biology*, 21(1):2.

[Jansen et al., 2007] Jansen, M. W. J. C., Corral, L., van der Velden, V. H. J., Panzer-Grümayer, R., Schrappe, M., Schrauder, A., Marschalek, R., Meyer, C., den Boer, M. L., Hop, W. J. C., Valsecchi, M. G., Basso, G., Biondi, A., Pieters, R., and van Dongen, J. J. M. (2007). Immunobiological diversity in infant acute lymphoblastic leukemia is related to the occurrence and type of MLL gene rearrangement. *Leukemia*, 21(4):633–641.

[Jensen et al., 2018] Jensen, C. T., Åhsberg, J., Sommarin, M. N. E., Strid, T., Somasundaram, R., Okuyama, K., Ungerbäck, J., Kupari, J., Airaksinen, M. S., Lang, S., Bryder, D., Soneji, S., Karlsson, G., and Sigvardsson, M. (2018). Dissection of progenitor compartments resolves developmental trajectories in B-lymphopoiesis. *Journal of Experimental Medicine*, page jem.20171384.

[Jones and Jenkins, 2018] Jones, S. A. and Jenkins, B. J. (2018). Recent insights into targeting the IL-6 cytokine family in inflammatory diseases and cancer. *Nature Reviews Immunology*, 18(12):773.

[Jorgensen et al., 2019] Jorgensen, I. F., Russo, F., Jensen, A. B., Westergaard, D., Lademann, M., Hu, J. X., Brunak, S., and Belling, K. (2019). Comorbidity landscape of the Danish patient population affected by chromosome abnormalities. *Genetics in Medicine*, page 1.

[Karamitros et al., 2018] Karamitros, D., Stoilova, B., Aboukhalil, Z., Hamey, F., Reinisch, A., Samitsch, M., Quek, L., Otto, G., Repapi, E., Doondeea, J., Usukhbayar, B., Calvo, J., Taylor, S., Goardon, N., Six, E., Pflumio, F., Porcher, C., Majeti, R., Göttgens, B., and Vyas, P. (2018). Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. *Nature Immunology*, 19(1):85.

[Kerry et al., 2017] Kerry, J., Godfrey, L., Repapi, E., Tapia, M., Blackledge, N. P., Ma, H., Ballabio, E., O'Byrne, S., Ponthan, F., Heidenreich, O., Roy, A., Roberts, I., Konopleva, M., Klose, R. J., Geng, H., and Milne, T. A. (2017). MLL-AF4 Spreading Identifies Binding Sites that Are Distinct from Super-Enhancers and that Govern Sensitivity to DOT1l Inhibition in Leukemia. *Cell Reports*, 18(2):482–495.

[Kirsammer et al., 2008] Kirsammer, G., Jilani, S., Liu, H., Davis, E., Gurbuxani, S., Beau, M. M. L., and Crispino, J. D. (2008). Highly penetrant myeloproliferative disease in the Ts65dn mouse model of Down syndrome. *Blood*, 111(2):767–775.

[Klusmann et al., 2010a] Klusmann, J.-H., Godinho, F. J., Heitmann, K., Maroz, A., Koch, M. L., Reinhardt, D., Orkin, S. H., and Li, Z. (2010a). Developmental stage-specific interplay of GATA1 and IGF signaling in fetal megakaryopoiesis and leukemogenesis. *Genes & Development*, 24(15):1659–1672.

[Klusmann et al., 2010b] Klusmann, J.-H., Li, Z., Böhmer, K., Maroz, A., Koch, M. L., Emmrich, S., Godinho, F. J., Orkin, S. H., and Reinhardt, D. (2010b). miR-125b-2 is a potential oncomiR on human chromosome 21 in megakaryoblastic leukemia. *Genes & Development*, 24(5):478–490.

[Kohn et al., 2012] Kohn, L. A., Hao, Q.-L., Sasidharan, R., Parekh, C., Ge, S., Zhu, Y., Mikkola, H. K. A., and Crooks, G. M. (2012). Lymphoid priming in human bone marrow begins before expression of CD10 with upregulation of L-selectin. *Nature Immunology*, 13(10):963–971.

[Krawczyk et al., 2014] Krawczyk, J., O'Dwyer, M., Swords, R., Freeman, C., and Giles, F. J. (2014). The Role of Inflammation in Leukaemia. In Aggarwal, B. B., Sung, B., and Gupta, S. C., editors, *Inflammation and Cancer*, Advances in Experimental Medicine and Biology, pages 335–360. Springer Basel, Basel.

[Kubota et al., 2019] Kubota, Y., Uryu, K., Ito, T., Seki, M., Kawai, T., Isobe, T., Kumagai, T., Toki, T., Yoshida, K., Suzuki, H., Kataoka, K., Shiraishi, Y., Chiba, K., Tanaka, H., Ohki, K., Kiyokawa, N., Kagawa, J., Miyano, S., Oka, A., Hayashi, Y., Ogawa, S., Terui, K., Sato, A., Hata, K., Ito, E., and Takita, J. (2019). Integrated genetic and epigenetic analysis revealed heterogeneity of acute lymphoblastic leukemia in Down syndrome. *Cancer Science*, 0(0).

[Kurosaka et al., 1999] Kurosaka, D., LeBien, T. W., and Pribyl, J. A. R. (1999). Comparative studies of different stromal cell microenvironments in support of human B-cell development. *Experimental Hematology*, 27(8):1271–1281.

[Kusters et al., 2009] Kusters, M. A. A., Verstegen, R. H. J., Gemen, E. F. A., and de Vries, E. (2009). Intrinsic defect of the immune system in children with Down syndrome: a review. *Clinical and Experimental Immunology*, 156(2):189–193.

[Labuhn et al., 2019] Labuhn, M., Perkins, K., Matzk, S., Varghese, L., Garnett, C., Papaemmanuil, E., Metzner, M., Kennedy, A., Amstislavskiy, V., Risch, T., Bhayadia, R., Samulowski, D., Hernandez, D. C., Stoilova, B., Iotchkova, V., Oppermann, U., Scheer, C., Yoshida, K., Schwarzer, A., Taub, J., Crispino, J. D., Weiss, M. J., Hayashi, A., Taga, T., Ito, E., Ogawa, S., Reinhardt, D., Yaspo, M.-L., Campbell, P. J., Roberts, I., Constantinescu, S., Vyas, P., Heckl, D., and Klusmann, J.-H. (2019). Mechanisms of Progression of Myeloid Preleukemia to Transformed Myeloid Leukemia in Children with Down Syndrome. *Cancer Cell*.

[Lane et al., 2014] Lane, A. A., Chapuy, B., Lin, C. Y., Tivey, T., Li, H., Townsend, E. C., Bodegom, D. v., Day, T. A., Wu, S.-C., Liu, H., Yoda, A., Alexe, G., Schinzel, A. C., Sullivan, T. J., Malinge, S., Taylor, J. E., Stegmaier, K., Jaffe, J. D., Bustin, M., Kronnie, G. t., Izraeli, S., Harris, M. H., Stevenson, K. E., Neuberg, D., Silverman, L. B., Sallan, S. E., Bradner, J. E., Hahn, W. C., Crispino, J. D., Pellman, D., and Weinstock, D. M. (2014). Triplication of a 21q22 region contributes to B cell transformation through HMGN1 overexpression and loss of histone H3 Lys27 trimethylation. *Nature Genetics*, 46(6):618–623.

[Laurenti and Göttgens, 2018] Laurenti, E. and Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. *Nature*, 553(7689):418–426.

[Lee et al., 2016] Lee, P., Bhansali, R., Izraeli, S., Hijiya, N., and Crispino, J. D. (2016). The biology, pathogenesis and clinical aspects of acute lymphoblastic leukemia in children with Down syndrome. *Leukemia*, 30(9):1816–1823.

[Lejeune et al., 1959] Lejeune, J., Gautier, M., and Turpin, R. (1959). [Study of somatic chromosomes from 9 mongoloid children]. *Comptes Rendus Hebdomadaires Des Seances De l'Academie Des Sciences*, 248(11):1721–1722.

[Li et al., 2018] Li, G., So, A. Y.-L., Sookram, R., Wong, S., Wang, J. K., Ouyang, Y., He, P., Su, Y., Casellas, R., and Baltimore, D. (2018). Epigenetic silencing of miR-125b is required for normal B-cell development. *Blood*, 131(17):1920–1930.

[Li, 2011] Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21):2987–2993.

[Li et al., 2009] Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16):2078–2079.

[Li et al., 2014] Li, Y., Esain, V., Teng, L., Xu, J., Kwan, W., Frost, I. M., Yzaguirre, A. D., Cai, X., Cortes, M., Maijenburg, M. W., Tober, J., Dzierzak, E., Orkin, S. H., Tan, K., North, T. E., and Speck, N. A. (2014). Inflammatory signaling regulates embryonic hematopoietic stem and progenitor cell production. *Genes & Development*, 28(23):2597–2612.

[Liao et al., 2014] Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7):923–930.

[Lin et al., 2016] Lin, K.-Y., Ye, H., Han, B.-W., Wang, W.-T., Wei, P.-P., He, B., Li, X.-J., and Chen, Y.-Q. (2016). Genome-wide screen identified let-7c/miR-99a/miR-125b regulating tumor progression and stem-like properties in cholangiocarcinoma. *Oncogene*, 35(26):3376–3386.

[Liu et al., 2014] Liu, B., Filippi, S., Roy, A., and Roberts, I. (2014). Stem and progenitor cell dysfunction in human trisomies. *EMBO reports*, page e201439583.

[Lott and Head, 2019] Lott, I. T. and Head, E. (2019). Dementia in Down syndrome: unique insights for Alzheimer disease research. *Nature Reviews Neurology*, 15(3):135.

[Love et al., 2014] Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15:550.

[Luo and Zheng, 2016] Luo, Y. and Zheng, S. G. (2016). Hall of Fame among Pro-inflammatory Cytokines: Interleukin-6 Gene and Its Transcriptional Regulation Mechanisms. *Frontiers in Immunology*, 7.

[MacLean et al., 2018] MacLean, G. A., McEldoon, J., Huang, J., Allred, J., Canver, M. C., and Orkin, S. H. (2018). Downregulation of Endothelin Receptor B Contributes to Defective B Cell Lymphopoiesis in Trisomy 21 Pluripotent Stem Cells. *Scientific Reports*, 8(1):8001.

[MacLean et al., 2012] MacLean, G. A., Menne, T. F., Guo, G., Sanchez, D. J., Park, I.-H., Daley, G. Q., and Orkin, S. H. (2012). Altered hematopoiesis in trisomy 21 as revealed through in vitro differentiation of isogenic human pluripotent cells. *Proceedings of the National Academy of Sciences*, 109(43):17567–17572.

[Maeda et al., 2005] Maeda, K., Baba, Y., Nagai, Y., Miyazaki, K., Malykhin, A., Nakamura, K., Kincade, P. W., Sakaguchi, N., and Coggeshall, K. M. (2005). IL-6 blocks a discrete early step in lymphopoiesis. *Blood*, 106(3):879–885.

[Malinge et al., 2012] Malinge, S., Bliss-Moreau, M., Kirsammer, G., Diebold, L., Chlon, T., Gurbuxani, S., and Crispino, J. D. (2012). Increased dosage of the chromosome 21 ortholog Dyrk1a promotes megakaryoblastic leukemia in a murine model of Down syndrome. *The Journal of Clinical Investigation*, 122(3):948–962.

[Malinge et al., 2009] Malinge, S., Izraeli, S., and Crispino, J. D. (2009). Insights into the manifestations, outcomes, and mechanisms of leukemogenesis in Down syndrome. *Blood*, 113(12):2619–2628.

[Maloney et al., 2010] Maloney, K. W., Carroll, W. L., Carroll, A. J., Devidas, M., Borowitz, M. J., Martin, P. L., Pullen, J., Whitlock, J. A., Willman, C. L., Winick, N. J., Camitta, B. M., and Hunger, S. P. (2010). Down syndrome childhood acute lymphoblastic leukemia has a unique spectrum of sentinel cytogenetic lesions that influences treatment outcome: a report from the Children's Oncology Group. *Blood*, 116(7):1045–1050.

[Mariani et al., 2019] Mariani, S. A., Li, Z., Rice, S., Krieg, C., Fragkogianni, S., Robinson, M., Vink, C. S., Pollard, J. W., and Dzierzak, E. (2019). Pro-inflammatory Aorta-Associated Macrophages Are Involved in Embryonic Development of Hematopoietic Stem Cells. *Immunity*, 50(6):1439–1452.e5.

[Marild et al., 2013] Marild, K., Stephansson, O., Grahnquist, L., Cnattingius, S., Soderman, G., and Ludvigsson, J. F. (2013). Down Syndrome Is Associated with Elevated Risk of Celiac Disease: A Nationwide Case-Control Study. *The Journal of Pediatrics*, 163(1):237–242.

[Massey et al., 2006] Massey, G. V., Zipursky, A., Chang, M. N., Doyle, J. J., Nasim, S., Taub, J. W., Ravindranath, Y., Dahl, G., and Weinstein, H. J. (2006). A prospective study of the natural history of transient leukemia (TL) in neonates with Down syndrome (DS): Children's Oncology Group (COG) study POG-9481. *Blood*, 107(12):4606–4613.

[Matloub et al., 2019] Matloub, Y., Rabin, K. R., Ji, L., Devidas, M., Hitzler, J., Xu, X., Bostrom, B. C., Stork, L. C., Winick, N., Gastier-Foster, J. M., Heerema, N. A., Stonerock, E., Carroll, W. L., Hunger, S. P., and Gaynon, P. S. (2019). Excellent long-term survival of children with Down syndrome and standard-risk ALL: a report from the Children's Oncology Group. *Blood Advances*, 3(11):1647–1656.

[McGrath et al., 2015a] McGrath, K. E., Frame, J. M., Fegan, K. H., Bowen, J. R., Conway, S. J., Catherman, S. C., Kingsley, P. D., Koniski, A. D., and Palis, J. (2015a). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. *Cell Reports*, 11(12):1892–1904.

[McGrath et al., 2015b] McGrath, K. E., Frame, J. M., and Palis, J. (2015b). Early hematopoiesis and macrophage development. *Seminars in Immunology*, 27(6):379–387.

[McHale et al., 2003] McHale, C. M., Wiemels, J. L., Zhang, L., Ma, X., Buffler, P. A., Guo, W., Loh, M. L., and Smith, M. T. (2003). Prenatal origin of TEL-AML1 positive acute lymphoblastic leukemia in children born in California. *Genes, Chromosomes and Cancer*, 37(1):36–43.

[Medvinsky and Dzierzak, 1996] Medvinsky, A. and Dzierzak, E. (1996). Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell*, 86(6):897–906.

[Meyr et al., 2013] Meyr, F., Escherich, G., Mann, G., Klingebiel, T., Kulozik, A., Rossig, C., Schrappe, M., Henze, G., Stackelberg, A. v., and Hitzler, J. (2013). Outcomes of treatment for relapsed acute lymphoblastic leukaemia in children with Down syndrome. *British Journal of Haematology*, 162(1):98–106.

[Mi et al., 2013] Mi, H., Muruganujan, A., and Thomas, P. D. (2013). PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Research*, 41(Database issue):D377–386.

[Michalska et al., 2018] Michalska, A., Blaszczyk, K., Wesoly, J., and Bluyssen, H. A. R. (2018). A Positive Feedback Amplifier Circuit That Regulates Interferon (IFN)-Stimulated Gene Expression and Controls Type I and Type II IFN Responses. *Frontiers in Immunology*, 9.

[Migita et al., 2017] Migita, K., Iwanaga, N., Izumi, Y., Kawahara, C., Kumagai, K., Nakamura, T., Koga, T., and Kawakami, A. (2017). TNF-a-induced miR-155 regulates IL-6 signaling in rheumatoid synovial fibroblasts. *BMC Research Notes*, 10.

[Milford et al., 2016] Milford, T.-A. M., Su, R. J., Francis, O. L., Baez, I., Martinez, S. R., Coats, J. S., Weldon, A. J., Calderon, M. N., Nwosu, M. C., Botimer, A. R., Suterwala, B. T., Zhang, X.-B., Morris, C. L., Weldon, D. J., Dovat, S., and Payne, K. J. (2016). TSLP or IL-7 provide an IL-7ra signal that is critical for human B lymphopoiesis. *European Journal of Immunology*, pages n/a–n/a.

[Montecino-Rodriguez and Dorshkind, 2012] Montecino-Rodriguez, E. and Dorshkind, K. (2012). B-1 B Cell Development in the Fetus and Adult. *Immunity*, 36(1):13–21.

[Montecino-Rodriguez et al., 2016] Montecino-Rodriguez, E., Fice, M., Casero, D., Berent-Maoz, B., Barber, C. L., and Dorshkind, K. (2016). Distinct Genetic Networks Orchestrate the Emergence of Specific Waves of Fetal and Adult B-1 and B-2 Development. *Immunity*, 45(3):527–539.

[Moore, 2006] Moore, C. S. (2006). Postnatal lethality and cardiac anomalies in the Ts65dn Down syndrome mouse model. *Mammalian Genome: Official Journal of the International Mammalian Genome Society*, 17(10):1005–1012.

[Morak et al., 2012] Morak, M., Attarbaschi, A., Fischer, S., Nassimbeni, C., Grausenburger, R., Bastelberger, S., Krentz, S., Cario, G., Kasper, D., Schmitt, K., Russell, L. J., Pötschger, U., Stanulla, M., Eckert, C., Mann, G., Haas, O. A., and Panzer-Grümayer, R. (2012). Small sizes and indolent evolutionary dynamics challenge the potential role of P2ry8-CRLF2–harboring clones as main relapse-driving force in childhood ALL. *Blood*, 120(26):5134–5142.

[Mori et al., 2002] Mori, H., Colman, S. M., Xiao, Z., Ford, A. M., Healy, L. E., Donaldson, C., Hows, J. M., Navarrete, C., and Greaves, M. (2002). Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proceedings of the National Academy of Sciences*, 99(12):8242–8247.

[Mostafavi et al., 2016] Mostafavi, S., Yoshida, H., Moodley, D., LeBoité, H., Rothamel, K., Raj, T., Ye, C., Chevrier, N., Zhang, S.-Y., Feng, T., Lee, M., Casanova, J.-L., Clark, J., Hegen, M., Telliez, J.-B., Hacohen, N., De Jager, P.,

Regev, A., Mathis, D., and Benoist, C. (2016). Parsing the Interferon Transcriptional Network and Its Disease Associations. *Cell*, 164(3):564–578.

[Mowery et al., 2018] Mowery, C. T., Reyes, J. M., Cabal-Hierro, L., Higby, K. J., Karlin, K. L., Wang, J. H., Kimmerling, R. J., Cejas, P., Lim, K., Li, H., Furusawa, T., Long, H. W., Pellman, D., Chapuy, B., Bustin, M., Manalis, S. R., Westbrook, T. F., Lin, C. Y., and Lane, A. A. (2018). Trisomy of a Down Syndrome Critical Region Globally Amplifies Transcription via HMGN1 Overexpression. *Cell Reports*, 25(7):1898–1911.e5.

[Mullighan et al., 2009] Mullighan, C. G., Collins-Underwood, J. R., Phillips, L. A. A., Loudin, M. G., Liu, W., Zhang, J., Ma, J., Coustan-Smith, E., Harvey, R. C., Willman, C. L., Mikhail, F. M., Meyer, J., Carroll, A. J., Williams, R. T., Cheng, J., Heerema, N. A., Basso, G., Pession, A., Pui, C.-H., Raimondi, S. C., Hunger, S. P., Downing, J. R., Carroll, W. L., and Rabin, K. R. (2009). Rearrangement of CRLF2 in B-progenitor– and Down syndrome–associated acute lymphoblastic leukemia. *Nature Genetics*, 41(11):1243–1246.

[Murakami et al., 2019] Murakami, M., Kamimura, D., and Hirano, T. (2019). Pleiotropy and Specificity: Insights from the Interleukin 6 Family of Cytokines. *Immunity*, 50(4):812–831.

[Murray et al., 2015] Murray, A., Letourneau, A., Canzonetta, C., Stathaki, E., Gimelli, S., Sloan-Bena, F., Abrehart, R., Goh, P., Lim, S., Baldo, C., Dagna-Bricarelli, F., Hannan, S., Mortensen, M., Ballard, D., Syndercombe Court, D., Fusaki, N., Hasegawa, M., Smart, T. G., Bishop, C., Antonarakis, S. E., Groet, J., and Nizetic, D. (2015). Brief Report: Isogenic Induced Pluripotent Stem Cell Lines From an Adult With Mosaic Down Syndrome Model Accelerated Neuronal Ageing and Neurodegeneration. *STEM CELLS*, 33(6):2077–2084.

[Müller et al., 1994] Müller, A. M., Medvinsky, A., Strouboulis, J., Grosveld, F., and Dzierzakt, E. (1994). Development of hematopoietic stem cell activity in the mouse embryo. *Immunity*, 1(4):291–301.

[Nadler et al., 1984] Nadler, L. M., Korsmeyer, S. J., Anderson, K. C., Boyd, A. W., Slaughenhoupt, B., Park, E., Jensen, J., Coral, F., Mayer, R. J., and Sallan, S. E. (1984). B cell origin of non-T cell acute lymphoblastic leukemia. A model for discrete stages of neoplastic and normal pre-B cell differentiation. *Journal of Clinical Investigation*, 74(2):332–340.

[Naka and Hirao, 2017] Naka, K. and Hirao, A. (2017). Regulation of Hematopoiesis and Hematological Disease by TGF-B Family Signaling Molecules. *Cold Spring Harbor Perspectives in Biology*, 9(9):a027987.

[NCARDRS, 2017] NCARDRS, N. C. A. a. R. D. R. S. (2017). National Congenital Anomaly and Rare Disease Registration Service: Congenital anomaly statistics 2017. *Public Health England*.

[Nikolaev et al., 2014] Nikolaev, S. I., Garieri, M., Santoni, F., Falconnet, E., Ribaux, P., Guipponi, M., Murray, A., Groet, J., Giarin, E., Basso, G., Nizetic, D., and Antonarakis, S. E. (2014). Frequent cases of RAS-mutated Down syndrome acute lymphoblastic leukaemia lack JAK2 mutations. *Nature Communications*, 5:4654.

[Notta et al., 2015] Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O. I., Wilson, G., Kaufmann, K. B., McLeod, J., Laurenti, E., Dunant, C. F., McPherson, J. D., Stein, L. D., Dror, Y., and Dick, J. E. (2015). Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*, page aab2116.

[Notta et al., 2016] Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O. I., Wilson, G., Kaufmann, K. B., McLeod, J., Laurenti, E., Dunant, C. F., McPherson, J. D., Stein, L. D., Dror, Y., and Dick, J. E. (2016). Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*, 351(6269):aab2116.

[Nuñez et al., 1996] Nuñez, C., Nishimoto, N., Gartland, G. L., Billips, L. G., Burrows, P. D., Kubagawa, H., and Cooper, M. D. (1996). B cells are generated throughout life in humans. *The Journal of Immunology*, 156(2):866–872.

[Oberlin et al., 2002] Oberlin, E., Tavian, M., Blazsek, I., and Péault, B. (2002). Blood-forming potential of vascular endothelium in the human embryo. *Development*, 129(17):4147–4157.

[O'Byrne et al., 2019] O'Byrne, S., Elliott, N., Rice, S., Buck, G., Fordham, N., Garnett, C., Godfrey, L., Crump, N. T., Wright, G., Inglott, S., Hua, P., Psaila, B., Povinelli, B., Knapp, D. J. H. F., Agraz-Doblas, A., Bueno, C., Varela, I., Bennett, P., Koohy, H., Watt, S. M., Karadimitris, A., Mead, A. J., Ancliff, P., Vyas, P., Menendez, P., Milne, T. A., Roberts, I., and Roy, A. (2019). Discovery of a CD10-negative B-progenitor in human fetal life identifies unique ontogeny-related developmental programs. *Blood*, 134(13):1059–1071.

[O'Doherty et al., 2005] O'Doherty, A., Ruf, S., Mulligan, C., Hildreth, V., Errington, M. L., Cooke, S., Sesay, A., Modino, S., Vanes, L., Hernandez, D., Linehan, J. M., Sharpe, P. T., Brandner, S., Bliss, T. V. P., Henderson, D. J., Nizetic, D., Tybulewicz, V. L. J., and Fisher, E. M. C. (2005). An Aneuploid Mouse Strain Carrying Human Chromosome 21 with Down Syndrome Phenotypes. *Science*, 309(5743):2033–2037.

[Olson et al., 2004] Olson, L. E., Richtsmeier, J. T., Leszl, J., and Reeves, R. H. (2004). A Chromosome 21 Critical Region Does Not Cause Specific Down Syndrome Phenotypes. *Science*, 306(5696):687–690.

[Ottersbach, 2019] Ottersbach, K. (2019). Endothelial-to-haematopoietic transition: an update on the process of making blood. *Biochemical Society Transactions*, 47(2):591–601.

[Pal et al., 2016] Pal, D., Blair, H. J., Elder, A., Dormon, K., Rennie, K. J., Coleman, D. J. L., Weiland, J., Rankin, K. S., Filby, A., Heidenreich, O., and Vormoor, J. (2016). Long-term in vitro maintenance of clonal abundance and leukaemia-initiating potential in acute lymphoblastic leukaemia. *Leukemia*.

[Palis et al., 1999] Palis, J., Robertson, S., Kennedy, M., Wall, C., and Keller, G. (1999). Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development*, 126(22):5073–5084.

[Parekh et al., 2016] Parekh, S., Ziegenhain, C., Vieth, B., Enard, W., and Hellmann, I. (2016). The impact of amplification on differential expression analyses by RNA-seq. *Scientific Reports*, 6:25533.

[Parrish et al., 2009] Parrish, Y. K., Baez, I., Milford, T.-A., Benitez, A., Galloway, N., Rogerio, J. W., Sahakian, E., Kagoda, M., Huang, G., Hao, Q.-L., Sevilla, Y., Barsky, L. W., Zielinska, E., Price, M. A., Wall, N. R., Dovat, S., and Payne, K. J. (2009). IL-7 Dependence in Human B Lymphopoiesis Increases during Progression of Ontogeny from Cord Blood to Bone Marrow. *The Journal of Immunology*, 182(7):4255–4266.

[Pellin et al., 2019] Pellin, D., Loperfido, M., Baricordi, C., Wolock, S. L., Montepeloso, A., Weinberg, O. K., Biffi, A., Klein, A. M., and Biasco, L. (2019). A comprehensive single cell transcriptional landscape of human hematopoietic progenitors. *Nature Communications*, 10(1):2395.

[Perez-Andres et al., 2011] Perez-Andres, M., Grosserichter-Wagener, C., Teodosio, C., Dongen, J. J. M. v., Orfao, A., and Zelm, M. C. v. (2011). The nature of circulating CD27+CD43+ B cells. *Journal of Experimental Medicine*, 208(13):2565–2566.

[Picelli et al., 2014] Picelli, S., Faridani, O. R., Bjorklund, A. K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols*, 9(1):171–181.

[Pietras et al., 2016] Pietras, E. M., Mirantes-Barbeito, C., Fong, S., Loeffler, D., Kovtonyuk, L. V., Zhang, S., Lakshminarasimhan, R., Chin, C. P., Techner, J.-M., Will, B., Nerlov, C., Steidl, U., Manz, M. G., Schroeder, T., and Passegué, E. (2016). Chronic interleukin-1 exposure drives haematopoietic stem cells towards precocious myeloid differentiation at the expense of self-renewal. *Nature Cell Biology*, 18(6):607–618.

[Pinho and Frenette, 2019] Pinho, S. and Frenette, P. S. (2019). Haematopoietic stem cell activity and interactions with the niche. *Nature Reviews Molecular Cell Biology*, 20(5):303.

[Pollak, 2012] Pollak, M. (2012). The insulin and insulin-like growth factor receptor family in neoplasia: an update. *Nature Reviews Cancer*, 12(3):159–169.

[Potter et al., 2018] Potter, N., Jones, L., Blair, H., Strehl, S., Harrison, C. J., Greaves, M., Kearney, L., and Russell, L. J. (2018). Single-cell analysis identifies CRLF2 rearrangements as both early and late events in Down syndrome and non-Down syndrome acute lymphoblastic leukaemia. *Leukemia*, page 1.

[Pribyl et al., 1995] Pribyl, J. R., Dittel, B. N., and Lebien, T. W. (1995). In Vitro Studies of Human B Lymphopoiesisa. *Annals of the New York Academy of Sciences*, 764(1):9–18.

[Psaila et al., 2016] Psaila, B., Barkas, N., Iskander, D., Roy, A., Anderson, S., Ashley, N., Caputo, V. S., Lichtenberg, J., Loaiza, S., Bodine, D. M., Karadimitris, A., Mead, A. J., and Roberts, I. (2016). Single-cell profiling of human megakaryocyte-erythroid progenitors identifies distinct megakaryocyte and erythroid differentiation pathways. *Genome Biology*, 17(1):83.

[Pui et al., 1993] Pui, C. H., Raimondi, S. C., Borowitz, M. J., Land, V. J., Behm, F. G., Pullen, D. J., Hancock, M. L., Shuster, J. J., Steuber, C. P., and Crist, W. M. (1993). Immunophenotypes and karyotypes of leukemic cells in children with Down syndrome and acute lymphoblastic leukemia. *Journal of Clinical Oncology*, 11(7):1361–1367.

[Qiao et al., 2018] Qiao, B., Austin, A. A., Schymura, M. J., and Browne, M. L. (2018). Characteristics and survival of children with acute leukemia with Down syndrome or other birth defects in New York State. *Cancer Epidemiology*, 57:68–73.

[Raha-Chowdhury et al., 2019] Raha-Chowdhury, R., Henderson, J. W., Raha, A. A., Vuono, R., Bickerton, A., Jones, E., Fincham, R., Allinson, K., Holland, A., and Zaman, S. H. (2019). Choroid Plexus Acts as Gatekeeper for TREM2, Abnormal Accumulation of ApoE, and Fibrillary Tau in Alzheimer's Disease and in Down Syndrome Dementia. *Journal of Alzheimer's Disease*, 69(1):91–109.

[Rainis et al., 2005] Rainis, L., Toki, T., Pimanda, J. E., Rosenthal, E., Machol, K., Strehl, S., Göttgens, B., Ito, E., and Izraeli, S. (2005). The Proto-Oncogene ERG in Megakaryoblastic Leukemias. *Cancer Research*, 65(17):7596–7602.

[Ramírez et al., 2016] Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, 44(W1):W160–W165.

[Reeves et al., 1995] Reeves, R. H., Irving, N. G., Moran, T. H., Wohn, A., Kitt, C., Sisodia, S. S., Schmidt, C., Bronson, R. T., and Davisson, M. T. (1995). A mouse model for Down syndrome exhibits learning and behaviour deficits. *Nature Genetics*, 11(2):177–184.

[Risso et al., 2014] Risso, D., Ngai, J., Speed, T. P., and Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. *Nature Biotechnology*, 32(9):896–902.

[Ritchie et al., 2015] Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., and Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7):e47.

[Roberts et al., 2013] Roberts, I., Alford, K., Hall, G., Juban, G., Richmond, H., Norton, A., Vallance, G., Perkins, K., Marchi, E., McGowan, S., Roy, A., Cowan, G., Anthony, M., Gupta, A., Ho, J., Uthaya, S., Curley, A., Rasiah, S. V., Watts, T., Nicholl, R., Bedford-Russell, A., Blumberg, R., Thomas, A., Gibson, B., Halsey, C., Lee, P.-W., Godambe, S., Sweeney, C., Bhatnagar, N., Goriely, A., Campbell, P., and Vyas, P. (2013). GATA1-mutant clones are frequent and often unsuspected in babies with Down syndrome: identification of a population at risk of leukemia. *Blood*, 122(24):3908–3917.

[Roberts and Izraeli, 2014] Roberts, I. and Izraeli, S. (2014). Haematopoietic development and leukaemia in Down syndrome. *British Journal of Haematology*, 167(5):587–599.

[Rodriguez-Meira et al., 2019] Rodriguez-Meira, A., Buck, G., Clark, S.-A., Povinelli, B. J., Alcolea, V., Louka, E., McGowan, S., Hamblin, A., Sousos, N., Barkas, N., Giustacchini, A., Psaila, B., Jacobsen, S. E. W., Thongjuea, S., and Mead, A. J. (2019). Unravelling Intratumoral Heterogeneity through High-Sensitivity Single-Cell Mutational Analysis and Parallel RNA Sequencing. *Molecular Cell*, 73(6):1292–1305.e8.

[Ross-Innes et al., 2012] Ross-Innes, C. S., Stark, R., Teschendorff, A. E., Holmes, K. A., Ali, H. R., Dunning, M. J., Brown, G. D., Gojis, O., Ellis, I. O., Green, A. R., Ali, S., Chin, S.-F., Palmieri, C., Caldas, C., and Carroll, J. S. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*, 481(7381):389–393.

[Roth, 2014] Roth, D. B. (2014). V(D)J Recombination: Mechanism, Errors, and Fidelity. *Microbiology Spectrum*, 2(6).

[Roy et al., 2017] Roy, A., Bystry, V., Bohn, G., Goudevenou, K., Reigl, T., Papaioannou, M., Krejci, A., O'Byrne, S., Chaidos, A., Grioni, A., Darzentas, N., Roberts, I. A. G., and Karadimitris, A. (2017). High resolution IgH repertoire analysis reveals fetal liver as the likely origin of life-long, innate B lymphopoiesis in humans. *Clinical Immunology*, 183:8–16.

[Roy et al., 2012] Roy, A., Cowan, G., Mead, A. J., Filippi, S., Bohn, G., Chaidos, A., Tunstall, O., Chan, J. K. Y., Choolani, M., Bennett, P., Kumar, S., Atkinson, D., Wyatt-Ashmead, J., Hu, M., Stumpf, M. P. H., Goudevenou, K., O'Connor, D., Chou, S. T., Weiss, M. J., Karadimitris, A., Jacobsen, S. E., Vyas, P., and

Roberts, I. (2012). Perturbation of fetal liver hematopoietic stem and progenitor cell development by trisomy 21. *Proceedings of the National Academy of Sciences*, 109(43):17579–17584.

[Russell et al., 2009] Russell, L. J., Capasso, M., Vater, I., Akasaka, T., Bernard, O. A., Calasanz, M. J., Chandrasekaran, T., Chapiro, E., Gesk, S., Griffiths, M., Guttery, D. S., Haferlach, C., Harder, L., Heidenreich, O., Irving, J., Kearney, L., Nguyen-Khac, F., Machado, L., Minto, L., Majid, A., Moorman, A. V., Morrison, H., Rand, V., Strefford, J. C., Schwab, C., Tönnies, H., Dyer, M. J. S., Siebert, R., and Harrison, C. J. (2009). Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. *Blood*, 114(13):2688–2698.

[Ryan et al., 1997] Ryan, D. H., Nuccie, B. L., Ritterman, I., Liesveld, J. L., Abboud, C. N., and Insel, R. A. (1997). Expression of Interleukin-7 Receptor by Lineage-Negative Human Bone Marrow Progenitors With Enhanced Lymphoid Proliferative Potential and B-Lineage Differentiation Capacity. *Blood*, 89(3):929–940.

[Sago et al., 1998] Sago, H., Carlson, E. J., Smith, D. J., Kilbridge, J., Rubin, E. M., Mobley, W. C., Epstein, C. J., and Huang, T.-T. (1998). Ts1cje, a partial trisomy 16 mouse model for Down syndrome, exhibits learning and behavioral abnormalities. *Proceedings of the National Academy of Sciences*, 95(11):6256–6261.

[Sanchez et al., 1996] Sanchez, M.-J., Holmes, A., Miles, C., and Dzierzak, E. (1996). Characterization of the First Definitive Hematopoietic Stem Cells in the AGM and Liver of the Mouse Embryo. *Immunity*, 5(6):513–525.

[Sanjuan-Pla et al., 2015] Sanjuan-Pla, A., Bueno, C., Prieto, C., Acha, P., Stam, R. W., Marschalek, R., and Menéndez, P. (2015). Revisiting the biology of infant t(4;11)/MLL-AF4+ B-cell acute lymphoblastic leukemia. *Blood*, 126(25):2676–2685.

[Sanz et al., 2003] Sanz, E., Alvarez-Mon, M., Martıinez-A, C., and Hera, A. d. l. (2003). Human cord blood CD34+Pax-5+ B-cell progenitors: single-cell analyses of their gene expression profiles. *Blood*, 101(9):3424–3430.

[Sanz et al., 2010] Sanz, E., Muñoz-A, N., Monserrat, J., Van-Den-Rym, A., Escoll, P., Ranz, I., Álvarez Mon, M., and de-la Hera, A. (2010). Ordering human CD34+CD10-CD19+ pre/pro B-cell and CD19- common lymphoid progenitor stages in two pro-B-cell development pathways. *Proceedings of the National Academy of Sciences*, 107(13):5925–5930.

[Satge and Seidel, 2018] Satge, D. and Seidel, M. G. (2018). The Pattern of Malignancies in Down Syndrome and Its Potential Context With the Immune System. *Frontiers in Immunology*, 9.

[Schoggins et al., 2011] Schoggins, J. W., Wilson, S. J., Panis, M., Murphy, M. Y., Jones, C. T., Bieniasz, P., and Rice, C. M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature*, 472(7344):481–485.

[Schwartzman et al., 2017] Schwartzman, O., Savino, A. M., Gombert, M., Palmi, C., Cario, G., Schrappe, M., Eckert, C., Stackelberg, A. v., Huang, J.-Y., Hameiri-Grossman, M., Avigad, S., Kronnie, G. t., Geron, I., Birger, Y., Rein, A., Zarfati, G., Fischer, U., Mukamel, Z., Stanulla, M., Biondi, A., Cazzaniga, G., Vetere, A., Wagner, B. K., Chen, Z., Chen, S.-J., Tanay, A., Borkhardt, A., and Izraeli, S. (2017). Suppressors and activators of JAK-STAT signaling at diagnosis and relapse of acute lymphoblastic leukemia in Down syndrome. *Proceedings of the National Academy of Sciences*, page 201702489.

[Shaham et al., 2012] Shaham, L., Binder, V., Gefen, N., Borkhardt, A., and Izraeli, S. (2012). MiR-125 in normal and malignant hematopoiesis. *Leukemia*, 26(9):2011–2018.

[Sigvardsson, 2018] Sigvardsson, M. (2018). Molecular Regulation of Differentiation in Early B-Lymphocyte Development. *International Journal of Molecular Sciences*, 19(7).

[Silva et al., 2016] Silva, C. R. S., Biselli-Perico, J. M., Zampieri, B. L., Silva, W. A., de Souza, J. E. S., Burger, M. C., Goloni-Bertollo, E. M., and Pavarino, E. C. (2016). Differential Expression of Inflammation-Related Genes in Children with Down Syndrome.

[Sinka et al., 2012] Sinka, L., Biasch, K., Khazaal, I., Péault, B., and Tavian, M. (2012). Angiotensin-converting enzyme (CD143) specifies emerging lympho-hematopoietic progenitors in the human embryo. *Blood*, 119(16):3712–3723.

[Sipp et al., 2018] Sipp, D., Robey, P. G., and Turner, L. (2018). Clear up this stem-cell mess. *Nature*, 561(7724):455.

[Skorvaga et al., 2014] Skorvaga, M., Nikitina, E., Kubes, M., Kosik, P., Gajdosechova, B., Leitnerova, M., Copakova, L., and Belyaev, I. (2014). Incidence of Common

Preleukemic Gene Fusions in Umbilical Cord Blood in Slovak Population. *PLOS ONE*, 9(3):e91116.

[Stankiewicz and Crispino, 2009] Stankiewicz, M. J. and Crispino, J. D. (2009). ETS2 and ERG promote megakaryopoiesis and synergize with alterations in GATA-1 to immortalize hematopoietic progenitor cells. *Blood*, 113(14):3337–3347.

[Subramanian et al., 2005] Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550.

[Sullivan et al., 2017] Sullivan, K. D., Evans, D., Pandey, A., Hraha, T. H., Smith, K. P., Markham, N., Rachubinski, A. L., Wolter-Warmerdam, K., Hickey, F., Espinosa, J. M., and Blumenthal, T. (2017). Trisomy 21 causes changes in the circulating proteome indicative of chronic autoinflammation. *Scientific Reports*, 7(1):14818.

[Sullivan et al., 2016] Sullivan, K. D., Lewis, H. C., Hill, A. A., Pandey, A., Jackson, L. P., Cabral, J. M., Smith, K. P., Liggett, L. A., Gomez, E. B., Galbraith, M. D., DeGregori, J., and Espinosa, J. M. (2016). Trisomy 21 consistently activates the interferon response. *eLife*, 5:e16220.

[Tan et al., 2014] Tan, Y. S., Kim, M., Kingsbury, T. J., Civin, C. I., and Cheng, W.-C. (2014). Regulation of RAB5c is important for the growth inhibitory effects of MiR-509 in human precursor-B acute lymphoblastic leukemia. *PloS One*, 9(11):e111777.

[Taub et al., 2002] Taub, J. W., Konrad, M. A., Ge, Y., Naber, J. M., Scott, J. S., Matherly, L. H., and Ravindranath, Y. (2002). High frequency of leukemic clones in newborn screening blood samples of children with B-precursor acute lymphoblastic leukemia. *Blood*, 99(8):2992–2996.

[Tavian et al., 1996] Tavian, M., Coulombel, L., Luton, D., Clemente, H., Dieterlen-Lièvre, F., and Péault, B. (1996). Aorta-associated CD34+ hematopoietic cells in the early human embryo. *Blood*, 87(1):67–72.

[Tavian et al., 1999] Tavian, M., Hallais, M. F., and Peault, B. (1999). Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo. *Development*, 126(4):793–803.

[Tavian et al., 2001] Tavian, M., Robin, C., Coulombel, L., and Péault, B. (2001). The Human Embryo, but Not Its Yolk Sac, Generates Lympho-Myeloid Stem Cells. *Immunity*, 15(3):487–495.

[Thompson et al., 2015] Thompson, B. J., Bhansali, R., Diebold, L., Cook, D. E., Stolzenburg, L., Casagrande, A.-S., Besson, T., Leblond, B., Désiré, L., Malinge, S., and Crispino, J. D. (2015). DYRK1a controls the transition from proliferation to quiescence during lymphoid development by destabilizing Cyclin D3. *Journal of Experimental Medicine*, 212(6):953–970.

[Tikhonova et al., 2019] Tikhonova, A. N., Dolgalev, I., Hu, H., Sivaraj, K. K., Hoxha, E., Cuesta-Dominguez, A., Pinho, S., Akhmetzyanova, I., Gao, J., Witkowski, M., Guillamot, M., Gutkin, M. C., Zhang, Y., Marier, C., Diefenbach, C., Kousteni, S., Heguy, A., Zhong, H., Fooksman, D. R., Butler, J. M., Economides, A., Frenette, P. S., Adams, R. H., Satija, R., Tsirigos, A., and Aifantis, I. (2019). The bone marrow microenvironment at single-cell resolution. *Nature*, page 1.

[Trapnell et al., 2014] Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N. J., Livak, K. J., Mikkelsen, T. S., and Rinn, J. L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology*, 32(4):381–386.

[Tunstall-Pedoe et al., 2008] Tunstall-Pedoe, O., Roy, A., Karadimitris, A., Fuente, J. d. l., Fisk, N. M., Bennett, P., Norton, A., Vyas, P., and Roberts, I. (2008). Abnormalities in the myeloid progenitor compartment in Down syndrome fetal liver precede acquisition of GATA1 mutations. *Blood*, 112(12):4507–4511.

[Uckun, 1990] Uckun, F. M. (1990). Regulation of human B-cell ontogeny. *Blood*, 76(10):1908–1923.

[Uckun and Ledbetter, 1988] Uckun, F. M. and Ledbetter, J. A. (1988). Immunobiologic differences between normal and leukemic human B-cell precursors. *Proceedings of the National Academy of Sciences*, 85(22):8603–8607.

[Valentini et al., 2015] Valentini, D., Marcellini, V., Bianchi, S., Villani, A., Facchini, M., Donatelli, I., Castrucci, M. R., Marasco, E., Farroni, C., and Carsetti, R. (2015). Generation of switched memory B cells in response to vaccination in Down syndrome children and their siblings. *Vaccine*, 33(48):6689–6696.

[Verstegen et al., 2010] Verstegen, R. H. J., Kusters, M. A. A., Gemen, E. F. A., and De Vries, E. (2010). Down Syndrome B-Lymphocyte Subpopulations, Intrinsic Defect or Decreased T-Lymphocyte Help. *Pediatric Research*, 67(5):563–569.

[Vesely et al., 2017] Vesely, C., Frech, C., Eckert, C., Cario, G., Mecklenbräuker, A., Stadt, U. z., Nebral, K., Kraler, F., Fischer, S., Attarbaschi, A., Schuster, M., Bock, C., Cavé, H., Stackelberg, A. v., Schrappe, M., Horstmann, M. A., Mann, G., Haas, O. A., and Panzer-Grümayer, R. (2017). Genomic and transcriptional landscape of P2ry8-CRLF2-positive childhood acute lymphoblastic leukemia. *Leukemia*, 31(7):1491–1501.

[Viguie, 2001] Viguie, F. (2001). +21 or trisomy 21; http://atlasgeneticsoncology.org/Anomalies/tri21id1041.html; accessed 20 Sep 2019.

[Välikangas et al., 2016] Välikangas, T., Suomi, T., and Elo, L. L. (2016). A systematic evaluation of normalization methods in quantitative label-free proteomics. *Briefings in Bioinformatics*, page bbw095.

[Whitlock et al., 2005] Whitlock, J. A., Sather, H. N., Gaynon, P., Robison, L. L., Wells, R. J., Trigg, M., Heerema, N. A., and Bhatia, S. (2005). Clinical characteristics and outcome of children with Down syndrome and acute lymphoblastic leukemia: a Children's Cancer Group study. *Blood*, 106(13):4043–4049.

[Wolf et al., 1991] Wolf, M. L., Buckley, J. A., Goldfarb, A., Law, C. L., and LeBien, T. W. (1991). Development of a bone marrow culture for maintenance and growth of normal human B cell precursors. *The Journal of Immunology*, 147(10):3324–3330.

[Woo et al., 2013] Woo, A. J., Wieland, K., Huang, H., Akie, T. E., Piers, T., Kim, J., and Cantor, A. B. (2013). Developmental differences in IFN signaling affect GATA1s-induced megakaryocyte hyperproliferation. *The Journal of Clinical Investigation*, 123(8):3292–3304.

[Wu et al., 2018] Wu, X., Thi, V. L. D., Huang, Y., Billerbeck, E., Saha, D., Hoffmann, H.-H., Wang, Y., Silva, L. A. V., Sarbanes, S., Sun, T., Andrus, L., Yu, Y., Quirk, C., Li, M., MacDonald, M. R., Schneider, W. M., An, X., Rosenberg, B. R., and Rice, C. M. (2018). Intrinsic Immunity Shapes Viral Resistance of Stem Cells. *Cell*, 172(3):423–438.e25.

[Yagi et al., 2000] Yagi, T., Hibi, S., Tabata, Y., Kuriyama, K., Teramura, T., Hashida, T., Shimizu, Y., Takimoto, T., Todo, S., Sawada, T., and Imashuku, S. (2000). Detection of clonotypic IGH and TCR rearrangements in the neonatal blood spots of infants and children with B-cell precursor acute lymphoblastic leukemia. *Blood*, 96(1):264–268.

[Yamazaki et al., 2011] Yamazaki, S., Ema, H., Karlsson, G., Yamaguchi, T., Miyoshi, H., Shioda, S., Taketo, M., Karlsson, S., Iwama, A., and Nakauchi, H. (2011). Nonmyelinating Schwann Cells Maintain Hematopoietic Stem Cell Hibernation in the Bone Marrow Niche. *Cell*, 147(5):1146–1158.

[Yang et al., 2011] Yang, Z.-F., Drumea, K., Cormier, J., Wang, J., Zhu, X., and Rosmarin, A. G. (2011). GABP transcription factor is required for myeloid differentiation, in part, through its control of Gfi-1 expression. *Blood*, 118(8):2243–2253.

[Yoda et al., 2010] Yoda, A., Yoda, Y., Chiaretti, S., Bar-Natan, M., Mani, K., Rodig, S. J., West, N., Xiao, Y., Brown, J. R., Mitsiades, C., Sattler, M., Kutok, J. L., DeAngelo, D. J., Wadleigh, M., Piciocchi, A., Dal Cin, P., Bradner, J. E., Griffin, J. D., Anderson, K. C., Stone, R. M., Ritz, J., Foà, R., Aster, J. C., Frank, D. A., and Weinstock, D. M. (2010). Functional screening identifies CRLF2 in precursor B-cell acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, 107(1):252–257.

[Yoshimoto, 2015] Yoshimoto, M. (2015). The first wave of B lymphopoiesis develops independently of stem cells in the murine embryo. *Annals of the New York Academy of Sciences*, 1362(1):16–22.

[Yoshimoto et al., 2011] Yoshimoto, M., Montecino-Rodriguez, E., Ferkowicz, M. J., Porayette, P., Shelley, W. C., Conway, S. J., Dorshkind, K., and Yoder, M. C. (2011). Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone progenitor lacking B-2 potential. *Proceedings of the National Academy of Sciences of the United States of America*, 108(4):1468–1473.

[Young et al., 2010] Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology*, 11(2):R14.

[Zampieri et al., 2014] Zampieri, B. L., Biselli-Perico, J. M., Souza, J. E. S. d., Burger, M. C., Junior, W. A. S., Goloni-Bertollo, E. M., and Pavarino, E. C. (2014). Altered Expression of Immune-Related Genes in Children with Down Syndrome. *PLOS ONE*, 9(9):e107218.

[Zeller et al., 2005] Zeller, B., Gustafsson, G., Forestier, E., Abrahamsson, J., Clausen, N., Heldrup, J., Hovi, L., Jonmundsson, G., Lie, S. O., Glomstein, A., Hasle, H., and the Nordic Society of Paediatric Haematology and Oncology (NOPHO) (2005). Acute leukaemia in children with Down syndrome: a population-based Nordic study. *British Journal of Haematology*, 128(6):797–804.

[Zhao et al., 2014] Zhao, M., Perry, J. M., Marshall, H., Venkatraman, A., Qian, P., He, X. C., Ahamed, J., and Li, L. (2014). Megakaryocytes maintain homeostatic quiescence and promote post-injury regeneration of hematopoietic stem cells. *Nature Medicine*, 20(11):1321–1326.

[Zheng et al., 2018] Zheng, B., Xi, Z., Liu, R., Yin, W., Sui, Z., Ren, B., Miller, H., Gong, Q., and Liu, C. (2018). The Function of MicroRNAs in B-Cell Development, Lymphoma, and Their Potential in Clinical Practice. *Frontiers in Immunology*, 9.

[Zipursky, 2003] Zipursky, A. (2003). Transient leukaemia – a benign form of leukaemia in newborn infants with trisomy 21. *British Journal of Haematology*, 120(6):930–938.