Investigating gene expression patterns and amino acid metabolism in CML LSCs using transcriptomics and *ex-vivo* expansion

Kiron Roy, BSc., MSc.

Thesis submitted to the University of Glasgow in accordance with the requirements for the degree of Doctor of Philosophy

School of Cancer Sciences

College of Medical, Veterinary and Life Sciences

University of Glasgow September 2025

Abstract

Chronic myeloid leukaemia (CML) is a myeloproliferative disease where a normal haematopoietic stem cell (HSC) undergoes chromosomal translocation, leading to the formation of the Philadelphia chromosome (Ph). The Ph contains the fusion oncogene breakpoint cluster region abelson murine Leukemia 1 (BCR::ABL1), which encodes a constitutively active tyrosine kinase. BCR::ABL1 signalling transforms a HSC into a leukaemic stem cell (LSC), which is responsible for propagating CML. Tyrosine kinase inhibitors (TKIs) have vastly improved patient outcomes in the long and short term, however BCR::ABL1 inhibition alone is insufficient to eradicate all disease causing LSCs.

Bulk transcriptome profiling is a method which has been used to study CML LSCs for over 20 years, and has revealed much regarding the phenotype and molecular features of the LSC. Recent advances such as single cell transcriptomics have again increased the granularity at which we understand LSCs, now providing the capacity to phenotypically identify BCR::ABL1$^+$ LSCs, predict treatment responses, and identify features of treatment resistant LSCs. Metabolomics and transcriptomics have also shown that LSCs possess enhanced metabolism compared to their HSC counterparts, with metabolic adaption also being a mechanism by which LSCs are capable of surviving TKI treatment. Therefore, we aimed to conduct a meta-analysis on CML LSC transcriptomic datasets to determine if we could identify novel BCR::ABL1$^+$ specific signals and metabolic genes across phenotypes, ranging from CD34$^+$ to CD34$^+$CD38$^-$CD90$^+$CD26$^+$ cells. Additionally, we aimed to use recent advances in *ex-vivo* HSC culture and apply them to LSCs, to identify the metabolic weaknesses of murine CML LSCs, and validate these findings *in-vivo*.

In our meta-analysis, we show numerous uncharacterised genes differentially expressed in LSCs compared to bystander BCR::ABL1$^-$ or normal HSCs across datasets and phenotypes, and a core of commonly expressed genes that would likely represent meaningful therapeutic targets. We also observed widespread changes in all areas of metabolism, particularly OXPHOS, amino acid (AA) and fatty acid metabolism in Lin$^-$CD34$^+$CD38$^-$ LSCs. While we detected fewer changes within the Lin$^-$CD34$^+$CD38$^-$CD90$^+$ compartment, we were still able to detect metabolic changes, suggesting this weakness is present even in a highly primitive phenotype.

To assess the utility of *ex-vivo* culture methods, we conducted hibernation culture and poly-vinyl alcohol (PVA) based expansion culture on CML LSCs. We found that the low number of LSCs that could be purified from leukaemic mice made hibernation culture a challenge when considering hypothesis testing, even though it may be more physiologically relevant. Using PVA based *ex-vivo* expansion, we were capable of expanding LSCs from a c-kit enriched or sorted population under normoxia or hypoxia. We then used expanded LSCs to conduct an AA dropout screen, wherein LSCs were grown in media individually missing each of the AAs. Among the top hits were valine, leucine and isoleucine (the branched chain amino acids (BCAAs)), which we further characterised using a combination of expanded LSCs and cell lines. We found BCAA restriction *in-vitro* to be highly efficacious, with a 75% reduction in BCAA concentration having a strong anti-leukaemic effect in both cell lines and expanded cells. We also show that BCAA restriction reduced *ex-vivo* expanded HSC growth substantially. Lastly, we used dietary restriction of BCAAs to treat murine CML *in-vivo*, however, the combination of BCAA restriction and CML was poorly tolerated, and so we could not detect any anti-leukaemic effect.

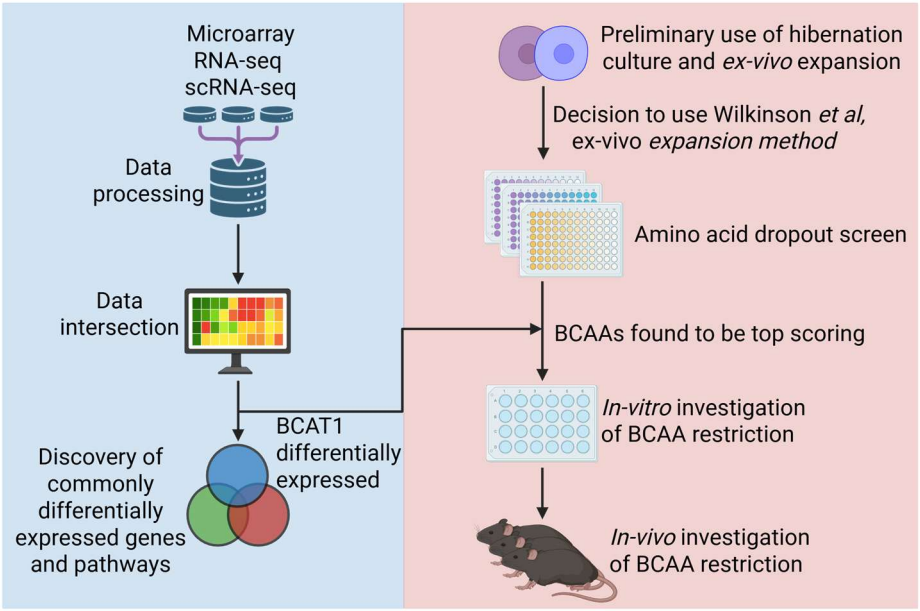


Figure 1: Graphical abstract. Transcriptomic datasets were identified and analysed, finding GAS2, CACNA1D, RXFP1, ZNF385D and FRY to be differentially expressed across datasets and phenotypes. BCAT1, the first enzyme in the branched chain amino acid (BCAA) metabolism pathway, was upregulated in predicted primitive LSCs. Utilising *ex-vivo* LSC culture methods an amino acid dropout screen was conducted, finding BCAA restriction to have a strong anti-leukaemic effect. BCAA restriction was further characterized *in-vitro* using LC-MS and Seahorse, however *in-vivo* BCAA restriction caused rapid weight loss, preventing any anti-leukaemic effect from being detected.

Table of Contents

List of figures

List of tables

Acknowledgements

Foremost, I would like to thank Professor Vignir Helgason for the opportunity to conduct a PhD in his laboratory, and especially for his patience and belief in me. His creation of a research environment that values open and honest communication and collaboration has allowed me to grow as a person and as a scientist. Special thanks to Kevin for the time and effort put into both my project and thesis, and also for his “good” jokes and shocking late-night appearance at the Sony.

I would like to express my gratitude to past and present members of M08 for their insight, support, and for the great moments we shared. You have all been instrumental to this project and I am grateful.

Other individuals out with M08 I would like to thank are the flow cytometry core of Tom, Yi-Hsia, and especially Jen for project specific insight. Additionally, I would like to acknowledge the BRU/BSU staff, the metabolomics team at the Beatson, and Karen Dunn, who was instrumental to this project. Thank you to all members of L3 for creating an enjoyable and collaborative working environment, particularly Lijun for sharing animal work responsibilities.

I would like to thank my partner Rose for her unconditional love, support and patience throughout this process. Lastly, I am eternally grateful for the encouragement and support provided by my family over the years, who have wholeheartedly supported my endeavours.

Author’s Declaration

I declare that all work presented is my own, unless otherwise stated. The contents of this thesis have not been submitted previously for any other degree.

Kiron Roy

Abbreviations

Abbreviation	Full name
1C	1-carbon
AAs	Amino acids
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
AMPK	AMP-activated protein kinase
ARG2	Arginase-2
ATG13	Autophagy-related protein 13
ATM	Ataxia-telangiectasia mutated
ATP	Adenosine triphosphate
AV	annexin V
Baf	Bafilomycin
BCAA	Branched chain amino acid
BCAT	Branched-chain aminotransferase
BCKA	Branched-chain α-keto acid
BCKDH	Branched-chain α-keto acid dehydrogenase
BCKDK	Branched-chain keto acid dehydrogenase kinase
BCR::ABL1	Breakpoint Cluster Region Abelson murine Leukemia 1
BCR::ABL1^{T315I}	T315I mutant BCR::ABL1
Bl6	Black 6
BM	Bone marrow
BP	Acute blast phase
C1orf21	chromosome 1 open reading frame 21
CACNA1D	calcium voltage-gated channel subunit alpha1 D
cAMP	Cyclic AMP
CCyR	Complete cytogenetic response
CD90	THY1
CITE-seq	Cellular Indexing of Transcriptomes and Epitopes by Sequencing
c-kit	CD117
CLL	chronic lymphocytic leukaemia
CLP	Common lymphoid progenitors
CML	Chronic myeloid leukaemia
CMP	Common myeloid progenitors
CoA	Coenzyme A
CP	Chronic phase
CRHBP	Corticotropin releasing hormone binding protein
CRIM1	Cysteine-rich Motor neuron 1
CXCL12	C-X-C motif chemokine 12
CXCL4	C-X-C motif chemokine 4
CXCR4	C-X-C motif chemokine receptor 4
DE	Differentially expressed
DEPDC1	DEP domain containing 1
dFBS	Dialysed FBS
DHCR24	4-dehydrocholesterol reductase
DMR	Deep molecular remission
DMSO	Dimethyl sulfoxide
DOX	Doxycycline
DTG	Double transgenic

EBSS	Earle's Balanced Salt Solution
EC	Endothelial cell
ECAR	extracellular acidification rate
EED	Embryonic ectoderm development
EGA	EGAS00001005509
ELN	European LeukemiaNet
EPCR	CD201
ER	Endoplasmic reticulum
E-SLAM	$Lineage^{-}Sca1^{+}ckit^{+}CD150^{+}CD48^{-}EPCR^{+}$
ETC	Electron transport chain
ELTS	EUTOS Long-Term Survival
EZH2	Enhancer of zeste homolog 2
FACS	Flourescance activated cell sorting
FAO	Fatty acid oxidation
FBS	Fetal bovine serum
FH	fumarate hydratase
FISH	Fluorescent in situ hybridization
FITC	Fluorescein isothiocyanate
FKBP4	FK506-binding protein 4
FoxO	Forkhead box O
FOXO3	Forkhead box O3
GAS2	Growth Arrest-Specific 2 protein
G-CSF	Granulocyte-colony stimulating factor
GEO	Gene Expression Omnibus
GMP	Granulocyte/Macrophage progenitor
GRB2	Growth factor receptor-bound protein 2
GRPEL1	GrpE-like 1, mitochondrial
GSE184	GSE218184
GSE233	GSE236233
GSE312	GSE76312
GSEA	Gene set enrichment analysis
HBSS	Hanks' Balanced Salt Solution
HCL	Hydrochloric acid
HCQ	Hydroxy chloroquine
HIF	Hypoxia inducible factor
HPLC	High pressure liquid chromatography
HREs	Hypoxia responsive elements
HSC	Haematopoietic stem cell
HSD17B	hydroxysteroid 17-beta dehydrogenase
IDH	Isocitrate-dehydrogenase
IGF-1	Insulin-like growth factor 1
IL-11	Interluekin-11
IL-6	Interleukin-6
IM	Imatinib mesylate
ITS -X	Insulin-Transferrin-Selenium-Ethanolamine
KCNK5	potassium two pore domain channel subfamily K member 5
KIC	2-keto-isocaproate/4-methyl-2-oxopentanoic acid
KIV	2-keto-isovalerate/3-methyl-2-oxobutanoic acid
KMV	α-keto-ß-methylvaleric acid/3-methyl-2-oxopentanoate
KOH	Potassium Hydroxide
KW	Kruskal-Wallis

KYNU	Kynureninase
LAT	L-type amino acid transporter
LGALS1	Galectin-1
Lhx2	LIM homeobox 2
Lin-	Lineage negative
Log2FC	Log 2 fold change
LSC	Leukaemic stem cell
LSK	$Lineage^{-}Sca1^{+}ckit^{+}$
LSK-SLAM	$Lineage^{-}Sca1^{+}ckit^{+}CD150^{+}CD48^{-}$
LT-HSC	Long-term haematopoietic stem cell
LY6E	lymphocyte antigen 6 family member E
MDH1	Malate dehydrogenase 1
MDH2	Malate dehydrogenase 2
MEP	Megakaryocytes/erythrocytes progenitor
MKs	Megakaryocytes
MMP	Multipotent progenitor
MMR	Major molecular response
MPC	Mitochondrial pyruvate carrier 1
MPP	Multipotent progenitors
MS	Musashi RNA binding protein
MSCs	Mesenchymal Stem Cells
mTOR	Mechanistic Target of Rapamycin kinase protein
MW	Mann-Whitney U test
NET1	neuroepithelial cell transforming gene 1
OCR	Oxygen consumption rate
ORA	Over representation analysis
OXPHOS	Oxidative phosphorylation
padj	Adjusted p value
PB	Peripheral blood
PBS	Dulbecco's phosphate buffered saline
PDK4	Pyruvate dehydrogenase kinase
PDX	Patient derived xenograft
PEG	Poly ethylene glycol
Pen/Strep	Penicillin-streptomycin
PGAM1	Phosphoglycerate Mutase 1
Ph	Philadelphia chromosome
PI	Propidium iodide
PI3K	Phosphoinositide 3-kinase
PIASy	Protein Inhibitor ff Activated STAT 4
PPIF	Cyclophilin D
PPM1K	Mg^{2+}/Mn^{2+}-dependent 1K protein phosphatase
PRC2	Polycomb repressive complex 2
PRDX3	Peroxiredoxin 3
PVA	Polyvinyl alcohol
ROS	Reactive oxygen species
RPMI	Roswell Park Memorial Institute
RXFP1	RXFP1 Relaxin family peptide receptor 1
SCENITH	Energetic metabolism by profiling Translation inhibition
SCF	Stem cell factor
scRNA-seq	Single-cell RNA sequencing
SEM	Standard error of the mean

SH2	Src homology 2
SH3BP4	SH3 Domain Binding Protein 4
SLC	Solute carrier
STAT5	Signal transducer and activator of transcription 5
STG	Single transgenic
ST-HSC	Short-term haematopoietic stem cell
TCA	Tricarboxylic acid cycle
TET	Ten-eleven translocation
TGF-β	Transforming growth factor beta
TKI	Tyrosine kinase inhibitor
TMEM97	Transmembrane protein 97
TNF-α	Tumour necrosis factor-α
TPO	Thrombopoietin
TXN	Thioredoxin
ULK1	Unc-51-like kinase 1
VCAM1	Vascular cell adhesion protein 1
VDAC	Voltage Dependent Anion Channel 1
VEGFA	Vascular endothelial growth factor A
VWF	Von Willebrand factor
ZNF385D	Zinc Finger Protein 385D
αKG	α-ketoglutarate

Chapter 1 Introduction

1.1 Introduction to haematopoietic stem cells

Haematopoietic stem cells (HSCs) are a rare cell type present in the bone marrow (BM) of adult mammals which possess extensive self-renewal potential alongside multipotency. This means that over the lifetime of a mammal, HSCs have the capacity to maintain their stemness while producing daughter cells, with the potential to differentiate into any adult haematopoietic-lineage cell[1]. HSCs were first discovered through studies aiming to discern aspects of radiosensitivity in 1961, where donor haematopoietic colonies were transplanted into lethally irradiated mice, rescuing them through full reconstitution of the haematopoietic system[2]. Transplantation and measurement of engraftment still remains the gold standard of quantifying HSC activity to this day. Although the most widely accepted definition of a HSC is a cell that can fully reconstitute the haematopoietic system of an irradiated/immunocompromised recipient, instead a HSC could be described as a cell that has the potential to do so at some point in its lifespan. For example, when HSCs were mobilized by granulocyte-colony stimulating factor (G-CSF), their engraftment potential was reduced without loss of quiescence. However, upon treatment cessation, the functional potential of the treated BM was markedly greater than that of the untreated BM[3,4]. Or, that cycling HSCs with high mitochondrial membrane potential (MMP) had massively increased engraftment potential (70-fold) when treated with an inhibitor of glycolysis post-transplant[5]. This indicates that HSC stemness can reversibly change during its lifetime, and that even though the cell cycle is a strong indicator of whether a phenotypic HSC can engraft or not, this is not always the case when steady state haematopoiesis is disrupted.

1.1.1 The haematopoietic hierarchy

HSCs sit atop the haematopoietic hierarchy, where this hierarchy represents the stages of differentiation a cell can undergo in the process of losing stemness and becoming a functional progenitor. When a HSC divides one of three outcomes are possible; symmetric self-renewal which produces an additional long-term HSC (LT-HSC), symmetric differentiation which yields two progenitor cells, or asymmetric self-renewal, where the parent cell remain stem while producing a daughter

progenitor cell[6]. These daughter cells produced after the first division are defined as short-term HSCs (ST-HSCs), which in turn can self-renew or divide again into multipotent progenitors (MPPs).

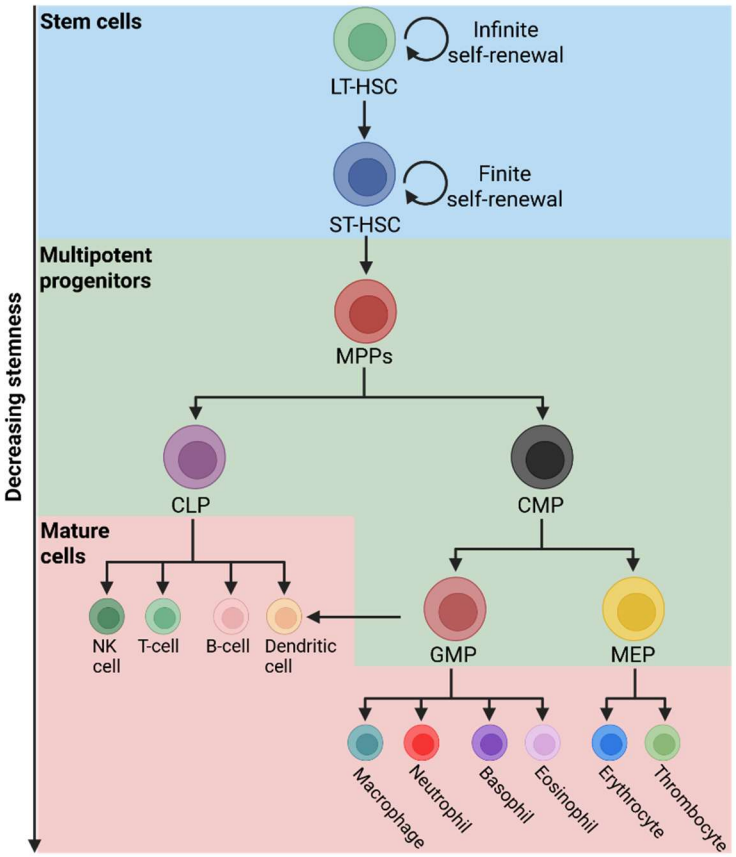


Figure 1-1 The haematopoietic hierarchy. Long-term haematopoietic stem cells (LT-HSCs) are the most primitive cell type with the potential to self-renew and/or differentiate into short-term haematopoietic stem cells (ST-HSCs) which do not have infinite self-renewal capacity. ST-HSCs can then give rise to multipotent progenitors (MPPs), wherein cells make a choice whether to become common myeloid progenitors (CMPs) or common lymphoid progenitors (CLPs). These common progenitors can then undergo numerous additional differentiation steps to become functional mature cells.

There are two differentiation fates for MPPs, they can become common myeloid progenitors (CMPs) or common lymphoid progenitors (CLPs), representing cells which can only differentiate into these specific lineages. CLPs will eventually become lymphocytes (T-cells, B-cells, natural killer cells, dendritic cells), while CMPs will either divide into a granulocyte/macrophage progenitor (GMPs) or megakaryocyte/erythrocyte progenitors (MEPs) which can further differentiate into numerous other cell types, such as erythrocytes and thrombocytes[6] (Figure 1-1).

This model is useful for understanding haematopoiesis, but recent insights provided by single-cell RNA sequencing (scRNA-seq) and lineage tracing have unravelled the true complexity of the haematopoietic hierarchy and HSC heterogeneity. For example, some phenotypic LT-HSCs show strong biases toward specific lineages, muddying the waters regarding whether these cells are MPPs with infinite self-renewal capacity or oligopotent LT-HSCs[7,8]. Utilising scRNA-seq and transplantation, *Rodriguez-Fraticelli et al*. show that a substantial number of LT-HSCs are destined for megakaryocytic differentiation, and that this fate occurs before other types of lineage commitment[9]. It has also been shown that megakaryocyte/platelet primed HSCs marked by Von Willebrand Factor (VWF) were capable of producing $VWF^{-/+}$ cells, while VWF^{-} HSCs could only produce more VWF^{-} progeny, indicating a hierarchical relationship[10]. Additionally, VWF^{-} HSCs were found to be lymphoid biased, while VWF^{+} HSCs were platelet/myeloid biased. If HSCs with differing lineage biases do have a hierarchical relationship, better characterization would be worthwhile when considering the efficiency of transplantation and gene therapy, particularly for human HSCs. Furthermore, whether these hierarchies still exist when steady state haematopoiesis is disturbed could provide greater insight into the propagation of haematopoietic malignancies, and exactly which cell types should be targeted. Causing further complication, lineage restriction appears to be pliable, with studies showing platelet restricted HSCs in a primary transplant were then capable of both myeloid and lymphoid differentiation *in-vitro*[7]. Aside from phenotype not being a definitive measure of cell fate, sister cell experiments have shown that HSCs with near indistinguishable gene expression profiles can be pre-committed to different lineages[11]. Fully understanding haematopoietic cell fate through transplantation remains a challenge, as lineage primed and lineage committed cells cannot easily be distinguished, as the BM environment itself may only be permitting certain lineages to be produced by that cell[12].

1.1.2 Human and murine HSC markers

After the discovery of the HSC, developing methods of accurately demarcating these cells from other BM or blood cells represented the next step into further characterisation. HSC markers have been studied extensively in both humans and mice (Table 1-1), and have been analysed and validated using a combination of

flow cytometry/fluorescence activated cell sorting (FACS) followed by transplantation, and in more recent years gene signatures from scRNA-seq.

In humans, the lineage negative (Lin^-)$CD45RA^-CD34^+CD38^-$ compartment is highly enriched for HSC activity, with the frequency of functional HSCs in the BM being approximately 1:30,000[13]. A number of other markers are capable of providing more specificity within this compartment such as, THY1 (CD90), CD201 (EPCR), CD117 (c-kit), CD49f and CD133[14]. From these markers CD90 has the greatest power to delineate functional HSCs from progenitors, with $Lin^-CD45RA^-CD34^+CD38^-CD90^+$ cells showing 9x greater chimerism than their $CD90^-$ counterparts when engrafted into immunocompromised mice[15]. Moreover, when examining cell types produced by these different fractions, $CD90^+$ cells were capable of giving rise to all phenotypes within the $CD34^+$ compartment, while $CD90^-$ cells could not produce $CD90^+$ cells. This indicates that under steady state haematopoiesis, CD90 is capable of delineating the largest population of functional HSCs from a heterogenous compartment. However, while $CD90^-$ HSCs lack the functional potential of their $CD90^+$ counterparts, loss of expression is not necessarily permanent. For example, Takeda *et al*. showed that after 6 days of *ex-vivo* culture, $Lin^-CD34^+CD38^-FLT3^-$ $CD90^+$ cells became $CD90^-$ while cycling, and then regained CD90 expression for another two days[16]. Additionally, Hua *et al*. took $Lin^-CD45RA^-CD34^+CD38^-CD90^-$ cells and treated them with a histone deacetylase, and at 2 days, almost all $CD90^-$ cells became $CD90^+$ for atleast 5 days[17]. Taken together, this indicates that while stemness does appear to be hierarchical, the potential for a progenitor to regain its functional potential does exist, if a cell for some reason decides to do so. Within the $CD90^+$ fraction, CD49f can again delineate a more stem population, the LT-HSC, with the use of EPCR and CD133 also becoming popularised to identify HSCs. However, unless specifically studying HSC hierarchy at the most primitive level, this specificity may not be required, as a $CD34^+$ population is sufficient to reconstitute the BM of an irradiated recipient.

Human and murine HSC phenotypes are not identical, with research groups using varied markers to delineate slightly different cell populations in mice. For example, $Lin^-Sca1^+c\text{-}kit^+$ (LSK) $CD34^-CD135^-$ or LSK $CD150^+CD48^-$ (LSK-SLAM) or $EPCR^+$LSK-SLAM (E-SLAM) cells[18]. Within the LSK compartment, CD150 expression has some capacity to separate the lineage bias and stemness of HSCs; $CD150^{hi}$ cells

show myeloid bias and contribute more to chimerism over time, CD150^{med} cells show lymphoid bias but more stable chimerism, while CD150^{neg} cells show poor reconstitution with a lymphoid bias[19]. This axis of CD150 HSC activity was further elucidated via transplantation studies utilizing 10 HSCs, showing that CD150^{hi} cells contained the greatest reconstitution capacity in secondary transplants, whilst CD150^{neg} cells showed the poorest. It has also been shown that the c-kit expression levels within the LSK compartment can separate cells based on cell cycle and metabolic activity. c-kit^{intermediate} cells were found to be quiescent in the BM but highly proliferative post-transplant, while c-kit^{Hi} cells were found to be more metabolically active, less capable of expansion and haematopoietic reconstitution, and lost engraftment potential during secondary and tertiary transplant[20]. A study by Wilson *et al*. identified markers giving the highest enrichment for functional HSCs to date, where they generated molecular signatures from scRNA-seq data. From this, they found 28 genes, including higher Sca-1 and CD150 alongside lower CD48 expression, to correspond to more primitive cells. Subsequent transplantation showed the engraftment potential of LS^{hi}K-SLAM cells outperformed LS^{low}K-SLAM cells, which was further refined to include EPCR^{+} cells. Single cell transplantation using this phenotype showed engraftment in 26/39 mice (67%), representing an extraordinarily pure population[21]. Another study combining sc-RNAseq and DNA barcoding found that the gene Tcf15 was highly upregulated in HSCs with long-term repopulating capacity within the LSK-SLAM compartment, which also correlated with EPCR expression[22]. Despite these advances, within the more primitive phenotypes (LSK-SLAM and its derivatives) true HSC activity is restricted to 40-67% of cells[21]. It could be that phenotype alone may not be capable of discriminating LT-HSCs with complete accuracy, or that the technical and stochastic aspects of sorting and transplanting a single cell in its current form only permit successful transplant up to a maximal limit, representing the current HSC detection threshold.

Table 1-1 Human and murine HSC markers. Positive markers coloured green, negative markers in black.

HSC markers	
Human	**Murine**
CD34	Lineage markers
CD38	Sca-1
Lineage markers	c-kit/CD117
CD90	CD150
CD45RA	CD48
CD201[23,24]	CD201/EPCR
c-kit/CD117	CD45
CD133[25,26]	CD34
CD49f[27]	CD135
	CD41

1.2 HSC regulation

1.2.1 Introduction to HSC fates

The capacity of a HSC to retain stemness, divide or differentiate, is linked with both cell intrinsic and extrinsic factors. For example, its cycling status, genetic/metabolic programme and phenotype, interaction with niche cells, exposure to secreted factors and physical location in the BM[6].

1.2.2 Localization of adult haematopoiesis

In adult humans, haematopoiesis occurs in the red marrow which is localized to the axial skeleton (cranial bones, vertebrae, sternum and other bones responsible for the protection of important organs), while all murine bones have the capacity for haematopoiesis[28,29]. The long bones are most commonly studied in mice with regards to haematopoiesis, given their ease of access and abundance of BM cells. These are composed of three segments, the epiphysis (poles of the bone), the diaphysis (shaft) and the metaphysis, a smaller segment which separates the epiphysis and the diaphysis (Figure 1-2)[30]. Within these regions of the bone, the BM environment can be divided into two regions, the endosteal niche and the perivascular niche. The endosteal niche is more rigid and closest to the bone while the perivascular niche is less rigid, closest to vasculature and more oxygenated[31]. HSCs reside in both of these locations, although they are more abundant near sinusoids, with 85% of HSCs being within 10μm of a sinusoidal blood vessel[32].

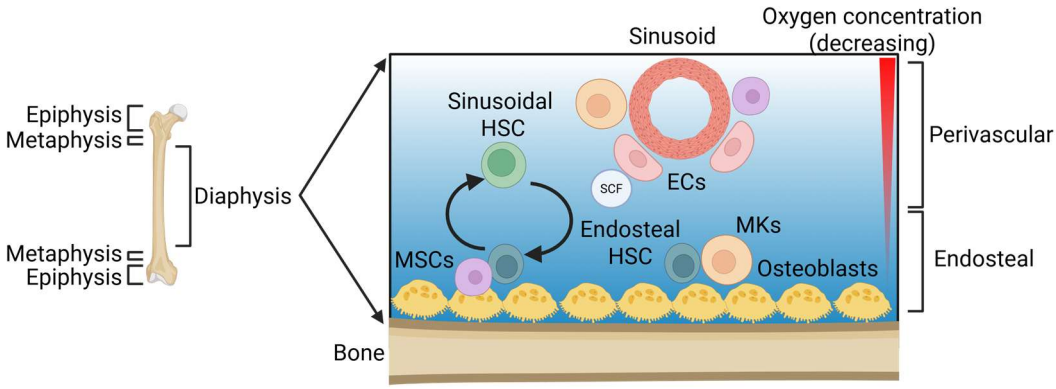


Figure 1-2 The long bone structure. The long bones are composed of the epiphysis (ends), diaphysis (shaft) and metaphysis which separates the epiphysis and diaphysis. Numerous cell types are present in these niches which support HSC activity and maintenance, with HSCs being more abundant in sinusoidal regions.

1.2.3 BM niche cells and HSCs

Aside from intrinsic HSC factors such as their current genetic programme and phenotype, there are numerous cell types capable of regulating haematopoiesis through direct interaction and secretion of growth factors from both haematopoietic and non-haematopoietic lineages.

Mesenchymal stem cells (MSCs) are a rare cell type defined as being $CD105^{+}$ $CD73^{+}CD90^{+}MHC\text{-}II^{Lo}CD11b^{-}CD14^{-}CD34^{-}CD45^{-}CD31^{+}$ and capable of self-renewal and differentiation into numerous cell types such as osteoblasts, endothelial cells (ECs), fibroblasts and adipocytes[33]. MSCs secrete numerous cytokines such as stem cell factor (SCF), C-X-C motif chemokine 12 (CXCL12), C-X-C motif chemokine 4 (CXCR4), interluekin-11 (IL-11), thrombopoietin (TPO) and vascular cell adhesion protein 1 (VCAM1), with CXCL12 and SCF being critical to HSC retention and proliferation respectively[34]. It has been shown that direct communication between MSCs and HSCs regulates the stemness of both cell types, with numerous studies showing strongly deleterious effects of MSC depletion on HSC stemness and retention[35,36]

In adult mammals ECs originate from localised endothelial progenitor cells, and can be defined as $CD45^{-}Ter119^{-}Sca1^{+}PDPN^{+}$ or $Ter119^{-}Sca1^{high}PDPN^{-}$ (sinusoidal and arterial), they line blood vessels and are highly abundant in the BM[37,38]. Similar to MSCs, ECs secrete SCF alongside CXCL12, with arterial ECs having a much greater contribution comparatively[39]. Additionally, ECs can manipulate HSCs through the modification of vascular permeability. For example, disruption of ECs causes increased reactive oxygen species (ROS) levels in HSCs, leading to increased cycling and depletion[40]. This effect was found to be rescuable through treatment with ROS scavengers.

Megakaryocytes (MKs) are of the haematopoietic lineage, and are the largest cell type in the BM. They can be localized in the sinusoidal or perivascular niche, with the former being preferred. MKs are most commonly known for their production of platelets, however they have also been proven to have the capacity to alter HSC fate[40,41]. Much like MSCs and ECs, MKs produce cytokines important for HSC localization and quiescence such as transforming growth factor beta 1 and CXCL4, but when under stress (e.g. injury) can induce proliferation through the

production of fibroblast growth factor 1 and insulin-like growth factor 1[42,43]. MKs and HSCs have also been shown to be closely associated spatially in the BM, in CXCL4$^{-/-}$ mice HSC cycling was enhanced, with megakaryocytic deletion causing aberrant haematopoiesis[44] .

1.2.4 HSC energy metabolism

HSCs in mature mammals under steady state conditions are almost always quiescent, this and their requirement for infinite self-renewal mean that HSCs maintain a particular metabolic profile to preserve these features (Figure 1-3).

Quiescent HSCs utilize anaerobic respiration and are more glycolytic than oxidative in order to reduced ROS output, with a large contribution to this effect being the hypoxic nature of the BM. For example, HSCs have reduced mitochondrial activity and a lower oxygen consumption rate (OCR) compared to their more differentiated progeny in the blood, which is also reflected in metabolomic analyses[45]. It was also shown that within the HSC compartment (LSK-SLAM) cells could be defined as either high (cycling primed) or low (quiescent) for MMP, and that only MMP high cells were sensitive to the inhibition of glycolysis *in-vitro*, either through inhibition of the mitochondrial pyruvate carrier (MPC) or the glucose transporter solute carrier (SLC)2A1[5]. Aside from reliance on glycolysis, it has been shown that fatty acid oxidation (FAO) is indispensable to the function of HSCs taken from older mice (24 months), and that this could be due to decreased metabolic plasticity in aged HSCs[46]. A high fat diet did not reduce the number of phenotypic HSCs, but did impair their functional potential, showing substantially reduced engraftment potential (10-40%) into primary and secondary recipients.

Unlike HSPCs, HSCs heavily utilize amino acids (AAs) to fulfil energetic needs. This was shown through the metabolic profiling of HSPCs (HSCs, LSKs, Lin$^-$, Lin$^+$) where AA concentrations unanimously increased up this hierarchy[47]. HSCs had the highest uptake of AAs out of these cell types and through tracing experiments[*] were

[*]A method used to study the metabolic content of a cell or biological suspension is liquid chromatography-mass spectrometry (LC-MS). This can be used be applied to measure the absolute or relative abundance of metabolites, but also to track the contribution of certain metabolites across metabolic pathways, known as tracing. In tracing experiments the metabolite to be traced can be exogenously supplied with heavy labelled carbons (^{13}C), the abundance of these within other metabolites can then be used as an indication of the flux of the labelled metabolite through a metabolic pathway, as well as its contribution to the formation of metabolites.

shown to have higher rates of AA catabolism into tricarboxylic acid cycle (TCA) intermediates. It was found that supplementing HSCs with α-ketoglutarate (αKG) and pyruvate caused greatly increased glucose uptake, suggesting that AA breakdown into TCA intermediates may be another factor contributing to the more glycolytic nature of HSCs[5]. Aside from adenosine triphosphate (ATP) generation the purpose of this enhanced AA metabolism is thought to be to limit protein synthesis to avoid proteotoxic stress. Knockout of the AA sensing gene GCN2 resulted in the accumulation of protein aggregates and apoptosis in fluorouracil treated mice, but had no effect in untreated mice with quiescent HSCs. To investigate the AA dependencies of HSCs, Taya *et al*. cultured murine HSCs in AA deficient media, showing that cysteine and valine dropouts had the strongest effects[48]. This was further investigated by Wilkinson *et al*. who showed that branched chain amino acid (BCAA) deficient media poorly supported but did not completely block HSC maintenance, a topic that will be covered in greater detail in section 1.5[49]

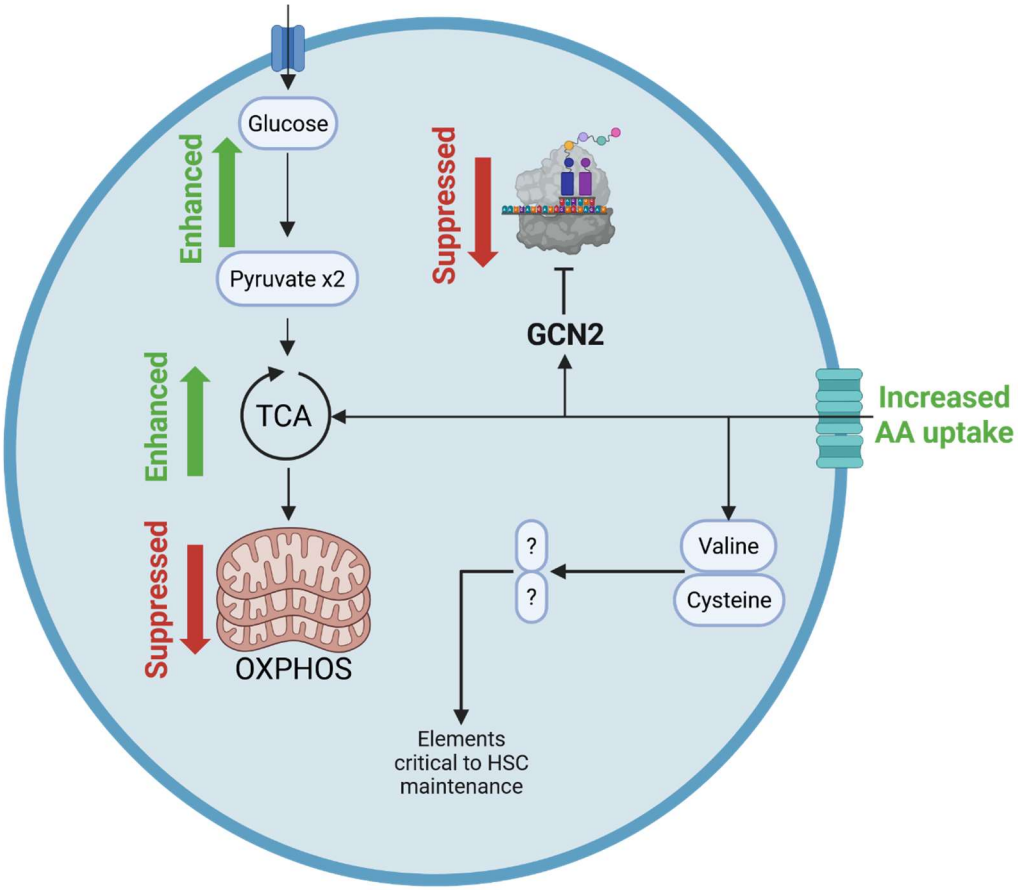


Figure 1-3 HSCs metabolism preserves stemness. Suppressed OXPHOS and enhanced glycolysis reduce ROS accumulation, while enhanced AA uptake and metabolism fuels TCA intermediate production and reduces proteotoxic effects. Through undefined mechanisms valine and cysteine are indispensable to HSCs.

1.2.5 Hypoxia, ROS and autophagy

Intimately linked to the metabolic state of a HSC are the effects of hypoxia, ROS and autophagy (Figure 1-4). As previously mentioned the BM is hypoxic, with oxygen concentrations ranging between 1-6%, with the distribution of HSPCs along this gradient being non-uniform[50,51]. As oxygen is the final electron acceptor in the electron transport chain (ETC), hypoxia reduces the capacity of a HSC to utilize OXPHOS. Part of a cell's response to hypoxia is regulated by the hypoxia inducible factor (HIF) genes, which are stably expressed under hypoxic conditions but hydroxylated and subject to proteasomal degradation under atmospheric conditions[52]. HIFs are transcription factors which bind to the hypoxia responsive elements (HREs) of genes involved in key hypoxia responses resulting in the modulation of metabolism, quiescence and stemness[53]. For example, HSCs in HIF1α deficient mice lost quiescence and had decreased functional potential and frequency[54]. Additionally, when the gene vascular endothelial growth factor A (VEGFA) was mutated at its HRE site, the frequency of phenotypic HSCs increased, but their functional potential was greatly decreased, (25-75% between 4-15 weeks)[55]. Transcriptionally, HIF signalling has been shown to upregulate glycolytic enzymes such as LUT1, LDHA and PDK1, helping maintain ATP production without reliance on OXPHOS[56]. Historically HIFs have been considered to be key contributors to HSC self-renewal, however in recent years this paradigm has been challenged, with HIF signalling possibly being redundant for HSC self-renewal. For example, *Vukovic et al*. show that HIF1α deficient cells retained the potential for serial reconstitution, while Guitart et al, showed that cells from both HIF1α and HIF2α deficient mice retained serial reconstitution capacity under steady state and disturbed haematopoiesis[57,58].

At physiological levels ROS signalling can control HSC fate, for example, whether symmetric or asymmetric division should occur, and where a HSC should localize. However, higher levels of ROS impair HSC function[59]. Deletion of genes involved in hypoxia/ROS responses such as ataxia-telangiectasia mutated and Forkhead box O3 (FOXO3) results in HSC exhaustion, increased senescence, apoptosis and even BM failure at 24 weeks in mouse models[60,61]. To further avoid accumulation of ROS-damaged mitochondria in HSCs, it has been shown that HSCs are capable of passing these mitochondria on in a targeted manner to daughter cells[62]. This oxidative phenotype was shown to be rescuable using ROS scavengers, proving that

the restriction of HSC metabolism to a glycolytic state is crucial to maintaining stemness. A key gene in the maintenance of both the metabolic state and ROS levels of HSCs is the mammalian target of rapamycin (mTOR), which is a master regulator of metabolism, protein synthesis, autophagy and nutrient sensing. It has been shown that deletion of negative regulators of mTOR resulted in ROS accumulation, increased mitochondrial biogenesis, and through these mechanisms, loss of stemness and exit from quiescence. Additionally, mTOR signalling is required to prevent aberrant haematopoiesis. For example, deletion of the mTORC1 component Raptor caused splenomegaly and the accumulation of monocytoid cells, while deletion of a negative regulator of mTORC2, PTEN, resulted in the development of leukaemia in adult mice[63,64].

Autophagy is the catabolic process of the degradation of intracellular components such as organelles, proteins and lipids through delivery into autophagosomes, functioning to regulate metabolic activity, aid in stress regulation and to generate additional metabolites[65]. In short, an initiation complex forms, beginning a cascade of events resulting in the formation of an autophagosome containing components to be recycled. The autophagosome then fuses with a lysosome, resulting in the degradation of these components. Autophagy initiation begins with mTOR, which in a phosphorylated state, directly associates with and prevents the activity of Unc-51-like kinase 1 (ULK1) and autophagy-related protein 13 (ATG13), which are two proteins that can initiate the autophagophore formation[66]. Under stressful conditions or nutrient depletion, ULK1 can be dephosphorylated leading to dissociation from the mTOR complex and phosphorylation of ATG13 and other ATG proteins, eventually leading to autophagosome production and cargo degradation. AMP-activated protein kinase (AMPK) also regulates autophagy through direct inhibition of mTOR and activation of ULK1[67]. Primitive cells, especially HSCs, have high levels of autophagy which have been shown to be critical for HSC regulation. HSCs deficient in ATG7 could not reconstitute the BM after lethal irradiation, showed deficient colony forming capacity, accumulation of damaged mitochondria, and increased ROS and DNA damage[68]. This suggests that removal of damaged mitochondria via autophagy (mitophagy) is a key mechanism HSCs utilise to limit OXPHOS and prevent the build-up of damaged mitochondria capable of causing HSCs genetic harm.

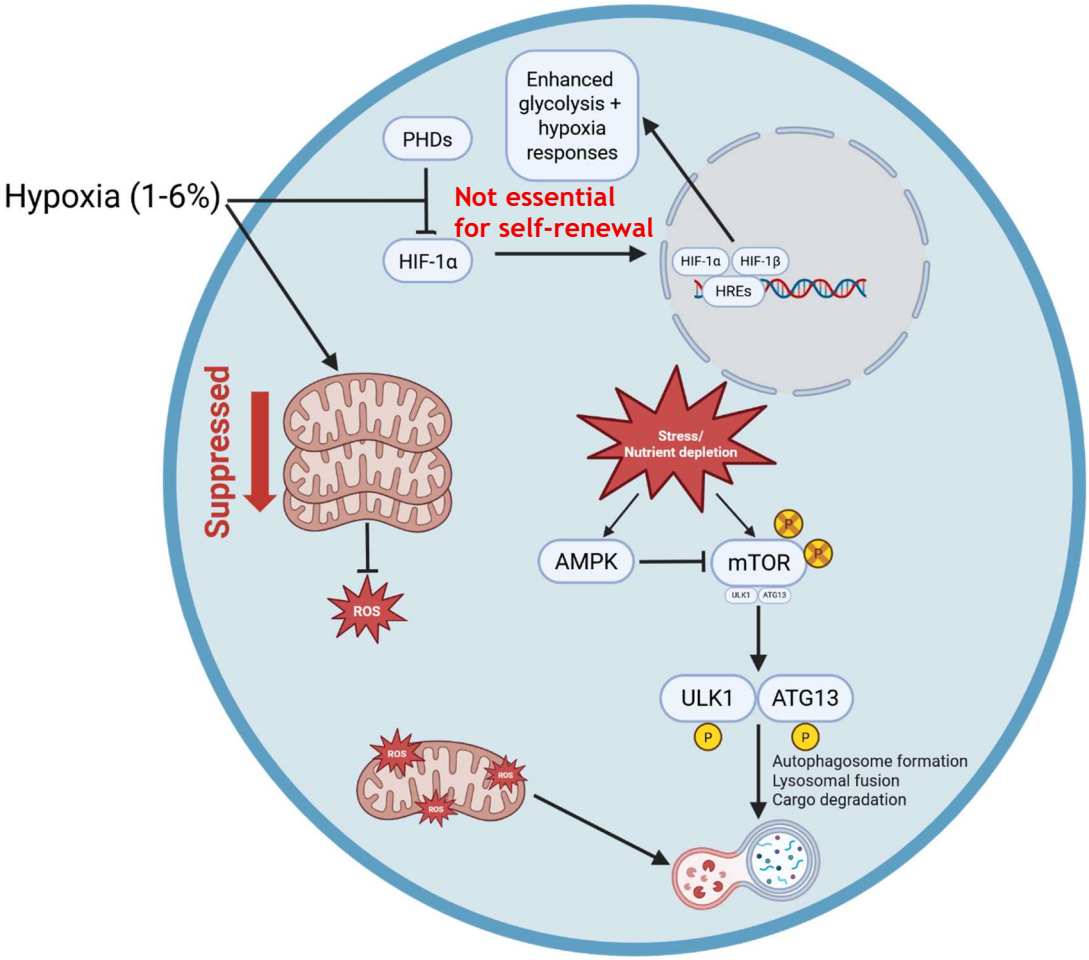


Figure 1-4 Effects of ROS, hypoxia and autophagy on HSC stemness. Low oxygen concentrations reduce OXPHOS capacity and stabilize HIFs, which translocate to the nucleus and cause the transcription of hypoxia response genes. Autophagy can be initiated by the dephosphorylation of mTOR, allowing ULK1 and ATG13 to dissociate and begin autophagosome formation. Damaged mitochondria can be recycled to prevent the build-up of ROS.

1.3 Chronic myeloid leukaemia and the leukaemic stem cell.

1.3.1 Introduction to CML pathology

Chronic myeloid leukaemia (CML) is a myeloproliferative neoplasm driven by haematopoietic stem cells (HSCs) which acquire a genetic aberration in the form of a reciprocal translocation between chromosomes 9 and 22. This results in the formation of the Philadelphia chromosome (Ph), which contains the fusion oncogene breakpoint cluster region abelson murine leukaemia 1 (BCR::ABL1), a constitutively active tyrosine kinase[69]. The mechanism behind Ph formation which results in the transformation of a HSC into a leukaemic stem cell (LSC) is still unknown. Unlike many other cancers, CML is not hereditary and there are no known risk factors aside from exposure to radiation, age and sex, with females being at slightly greater risk (45% male and 55% female)[70,71].

CML has a global incidence of 2 people per 100,000 and an approximate worldwide prevalence of 5 million, accounting for 15% of adult leukaemia and 1% of all childhood leukaemia with a median age of diagnosis being 60[69]. Pathologically, CML can be divided into two phases, chronic phase (CP) and the acute blast phase (BP), with most individuals being diagnosed in the CP.

CML treatment was revolutionised by the development of tyrosine kinase inhibitors (TKIs), which improved the 5-year survival rate from 20% to 95%, with 10-year survival only being slightly reduced to 85%[72]. The first TKI to be developed was Imatinib mesylate (IM), and it, like the majority of TKIs developed thereafter, functions through competitive inhibition of the ATP binding domain of BCR::ABL1, attenuating its oncogenic and proliferative signalling. Despite being highly efficacious at killing leukaemic progenitors, TKIs are incapable of targeting LSCs, which possess BCR::ABL1-independent survival mechanisms, meaning that individuals diagnosed with CML require lifelong TKI treatment, with 70% of patients being unable to enter treatment free remission[73]. Aside from the intrinsic TKI resistance of LSCs, BCR::ABL1 can develop mutations which further reduce TKI efficacy, notably the point mutation T315I (BCR::ABL1^{T315I}), which reduces the probability of survival by 60% over the course of 5 years[74].

Since longevity is increasing and current therapies cannot ensure the eradication of minimal residual disease, CML prevalence is increasing over time. Socioeconomic, morbidity, and mortality burdens will continue to increase if a curative approach cannot be developed. In the US in 2010 there were an estimated 70,000 CML patients, rising to 112,000 in 2020, with incidence predicted to rise until 2050[75]. Extrapolating this to a world population of 8 billion, global CML prevalence could potentially increase to >10 million cases[76]. Additionally, access to TKI does not appear to be equitable worldwide, with North America and Europe having the best outcomes and access, with central Africa having the poorest[77].

1.3.2 BCR::ABL1 signalling

The fusion gene BCR::ABL1 is the only known requirement for the development of CML, possessing immense oncogenic potential through the binding of Src homology 2 (SH2) domains of effector proteins to its autophosphorylated tyrosine residues, particularly Tyr177 at the n-terminal of the BCR domain[78]. Under normal circumstances ABL1 can be localized to the cytoplasm or nucleus, with the outcome of its signalling being compartment dependant. For example, ABL1 is required for cell cycle arrest and apoptotic responses to DNA damage[79]. However, when fused to BCR this nuclear translocation is impaired, potentially allowing cells to resist apoptosis and causing Abl1 to be exposed to cytoplasmic effectors for a greater duration[80]. Additionally, BCR::ABL1 (compared to ABL1) is capable of

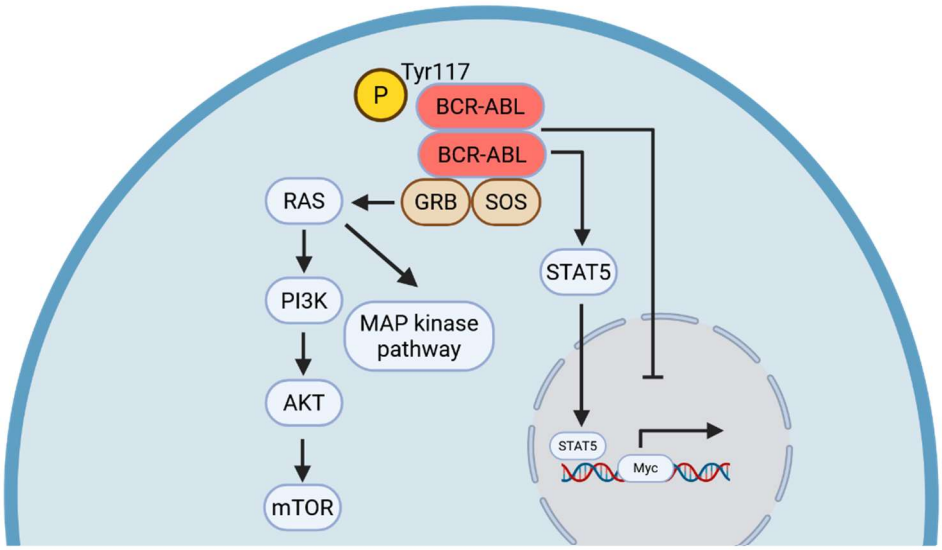


Figure 1-5 Key pathways activated by BCR::ABL1. Tyr177 at the n-terminal of BCR autophosphorylates, resulting in the recruitment of effector proteins, notably GRB2/SOS complex. This leads to the activation of the RAS signalling pathway. BCR::ABL1 signalling activates STAT5, which causes increased transcription, importantly of MYC. BCR::ABL fusion prevents nuclear translocation and apoptosis.

dimerization, which further enhances tyrosine-kinase activity by increasing the number of phosphotyrosine residues for effector binding[78]. Effectors activated by BCR::ABL1 include master regulators such as the GRB2/SOS complex which activates RAS and consequently phosphatidylinositol 3-kinase (PI3K)/mTOR, Signal transducer and activator of transcription 5 (STAT5) and indirectly MYC (Figure 1-5)[78].

Even though BCR::ABL1 is the only common genetic aberration present in CML, its presence alone does not always cause leukaemia. For example, the Ph^+ is detectable in individuals who never develop CML, and in a mouse model it was found that BCR::ABL1 acquisition alone was not enough to drive disease[81,82]. This suggests that perhaps a certain type of pre-leukaemic HSC, or a certain immunological landscape may be required for CML LSC to develop, however these factors remain undefined.

1.3.3 The CML LSC

The existence of the CML LSC was first discovered in 1999 by Holyoake *et al.*, where they were initially characterised as Ph^+ engraftable quiescent $CD34^+$ cells. Current research indicates that the CML LSC resides somewhere within $Lin^-CD34^+CD38^-CD45RA^-CD117^-CD26^+CD90^+$ compartment, highly similar to that of HSCs[83]. Definitively, an LSC from any leukaemia could be defined as a cell which is capable of infinite self-renewal, engraftment into immunocompromised mice, and the capacity to initiate leukaemia[84]. Contrary to this definition of an LSC, CML LSCs from the CP show extremely poor engraftment and disease-causing potential compared to LSCs from other haematological cancers (e.g. acute myeloid leukaemia (AML)), given that BCR::ABL1 drives proliferation which is in direct opposition with HSC activity. The damage BCR::ABL1 expression causes to primitive phenotypes has been characterised well in murine models, such as the double transgenic (DTG) Koschmieder *et al* model[85]. A healthy mouse might possess between 3000-5000 phenotypic HSCs, while a leukaemic mouse has between 400-800 HSCs/LSCs. From this model, the frequency of a long-term engrafting and disease causing LSC was found to be 1/80 cells in the LSK-SLAM-Flt3- phenotype. This means that of the 400-800 phenotypic LSCs that may be purified from a leukaemic mouse, at maximum there are 10 LSCs capable of both engraftment and leukaemogenesis. This depressed self-renewal capacity of CML

LSCs has made their characterisation particularly challenging, as the functional and disease-causing capacity of an LSC would typically be measured by engraftment.

1.3.4 LSC markers

In recent years, substantial effort has been made to improve the phenotypic recognition of LSCs, as the capacity to purify LSC subpopulations leads to the capability to target them. Through a combination of qPCR, transcriptomics and FACS, a number of markers have been proposed and validated to variable degrees, notably IL-1RAP, CD25, DPP4 (CD26), CD35, CD33 and CD93. Of these markers, the most widely adopted (and so far most useful) has been CD26. CD26 as an LSC marker was first identified by Herrmann *et al*. in 2014[86,87]. They showed that Lin^- $CD26^+$ LSCs were capable of engrafting into immunocompromised mice and retained the capacity to produce multilineage BCR::$ABL1^+$ progenitors, while corresponding $CD26^-$ LSCs only produced BCR::$ABL1^-$ progenitors. In 2022, Ebian *et al*. showed that $CD26^+$ LSC levels correlated with BCR::ABL1 transcript abundance in the peripheral blood (PB) at diagnosis and after 3 months of TKI treatment, with $CD26^+$ abundance corresponding to higher risk[88]. Further characterisation by Warfvinge *et al* in 2024 showed that the combination of $CD26^+/CD35^-$ could delineate bystander HSCs from BCR::$ABL1^+$ LSCs, and that poor responders had a higher burden of these cells at diagnosis[89]. Furthermore, they showed that Lin^- $CD34^+CD38^-CD45RA^-CD26^+CD35^-$ LSCs were as quiescent as their HSC counterparts, exhibiting stem cell like behaviour. Additionally, Camacho *et al* showed that 14/73 individuals in deep molecular remission (DMR) possessed $CD26^+$ LSCs where BCR::ABL1 transcripts were not detectable, indicating that BCR::ABL1 transcript detection from the PB may not necessarily reflect complete eradication of LSCs[90]. However, the abundance of these LSCs do decrease as DMR is achieved, and these LSCs were only present in 1 of 3 patients who were profiled that relapsed after TKI discontinuation. Therefore, there is no strong evidence that $CD26^+$ LSCs are useful for predicting outcome past diagnosis, or are solely responsible for driving relapse after TKI treatment.

Another marker of interest is CD93, where Kinstrie *et al*. show that the $CD93^+$ fraction represents a quiescent population. They also showed that $Lin^-CD34^+CD93^+$ LSCs could engraft, while $CD93^-$ LSCs could not, with output cells post engraftment

being almost all BCR:ABL^{+}[91]. Additionally, there was an increase in CD93 expressing LSCs in patients with molecular recurrence compared to those who responded well. There is certainly an overlap in the phenotypes demarcated by CD26 and CD93, as shown by Warfvinge *et al.*, however the capacity for a combination of these markers to predict treatment response and their engraftment potential is yet to be discerned. While the Lin^{-}CD34^{+}CD38^{-}CD45RA^{-} CD117^{-}CD26^{+}CD90^{+} population is thought to contain all leukaemogenic activity at diagnosis, the possibility of rare cell types within the CD26^{-}/CD90^{-} fraction also having some leukaemogenic potential remains a possibility[83].

1.3.5 LSC energy metabolism

Metabolic adaption is one of the hallmarks of cancer, with much evidence suggesting that CML LSCs metabolically rewire due to BCR::ABL1, and that these changes could represent a targetable weakness[92].

Kuntz *et al*. showed that CD34^{+} LSCs have higher rates of basal and maximal respiration compared to their healthy CD34^{+} counterparts. These LSCs also had increased mitochondrial mass and membrane potential and could be selectively targeted using Tigecycline, a mitochondrial protein translation inhibitor[93]. Rattigan *et al*. utilized ^{13}C glucose (in CD34^{+}/CD34^{+}CD38^{-} cells) and glutamine (in CD34^{+} cells) tracing in normal and CML cells, showing that both CD34^{+}/CD34^{+}CD38^{-} cells from CML patients had greater glucose catabolism, and that CD34^{+} LSPCs had greater glutamine catabolism and greater levels of metabolites compared to healthy counterparts[94]. Additionally, IM treatment partially reversed this phenotype but did not affect pyruvate anaplerosis, which was proven to be therapeutically exploitable using an MPC inhibitor *in-vitro* and in patient derived xenograft (PDX) models. In additional works, Rattigan *et al*. went on to show that CML CD34^{+} cells were arginine auxotrophs, lacking a functioning urea cycle. This was also therapeutically exploitable *in-vitro* and *in-vivo* using BCT-100, a recombinant enzyme capable of degrading arginine. One carbon (1C) metabolism and serine/glycine synthesis has been shown to be dysregulated in numerous cancers, recently including CML. Utilizing a combination of microarray, tracing and xenograft models, Zarou *et al*. showed that inhibition of *de-novo* purine synthesis was highly efficacious at reducing leukaemic burden[95]. Using glucose tracing, Shaowei *et al*. showed that in murine CML, LSCs treated with TKI for 2

days had depressed metabolism, however after two weeks an increase in the use of glucose for AA synthesis was observed[96]. Through scRNA-seq a similar effect was seen, initially primitive LSCs had reduced metabolic capacity, which was partially restored after two weeks of TKI treatment, with OXPHOS signatures recovering to levels similar to vehicle controls. Aberrant FAO is a less well characterised feature of LSCs, with a recent study showing that when CML cells were co-cultured with adipocytes they stimulated these cells to secrete fatty acids, and that under hypoxic conditions and low glucose concentrations, there was an accumulation of lipid droplets[97]. Moreover it has been shown that CD36, a fatty acid translocase, demarcates a subpopulation of $CD34^{+}38^{-}$ LSCs that are less sensitive to TKI and that these resistant cells could be targeted using an antibody therapy against CD36[98]. Branched chain amino acids (BCAAs) have been shown to be dysregulated in numerous myeloid malignancies, and are discussed in section 1.5.

1.3.6 The role of autophagy and hypoxia in CML

It was initially hypothesized that LSCs had reduced autophagy compared to HSCs given that BCR::ABL1 activates mTOR[99]. This could have explained the reduced engraftment potential of LSCs, however it has also been shown that LSCs from the CP have increased expression of ATG4B, BECN1 and ATG5. Contrary to this, Altman *et al*. showed that while BCR::$ABL1^{+}$ cells have lower levels of autophagy, they are highly dependent upon it[100–102]. What is certain is that TKI treatment of $CD34^{+}38^{-}$ LSCs and BCR::$ABL1^{+}$ cell lines enhances autophagy, and that this increase is a mechanism by which treatment resistant LSCs could persist[101]. Numerous studies have shown inhibition of autophagy to synergise with TKI treatment; for example, it was found that knockout of ATG5/7 genes sensitized both cell lines and primary cells to TKI-induced death, and that ATG7 knockdown resulted in enhanced OXPHOS and ROS, strongly suggesting that increased mitophagy is a mechanism LSCs use to reduce TKI induced stress[103,104]. Given these promising findings the autophagy inhibitor hydroxychloroquine (HCQ) was assessed in a phase II clinical trial in conjunction with TKI. However, reductions in BCR:ABL1 levels in the blood were minimal, due to inconsistent HCQ levels in the blood and circulating concentrations not reaching quantities required for complete inhibition of autophagy[105]. Recent work by the Helgason lab has utilized a new autophagy inhibitor MRT403, which showed greater specificity compared to HCQ in CML

mouse models, suggesting that development or utilization of more specific autophagy inhibitors could be a successful strategy[106].

Aside from the metabolic regulation hypoxia imposes on HSCs, it appears to play an additional role in LSCs. For example, it has been shown that when $CD34^+CD38^+$ and $CD34^+CD38^-$ LSCs were cultured with and without IM, under hypoxic or normoxic conditions, hypoxia cultured cells showed greater viability compared to their normoxic counterparts[107]. Unsurprisingly, the engraftment potential of hypoxia cultured $CD34^+$ LSCs outperformed normoxia cultured LSCs, with freshly harvested samples showing the greatest engraftment potential. It was also shown that HIF-2α was upregulated in human and murine LSCs after TKI treatment, and that inhibition reduced LSC number and increased TKI sensitivity whether through pharmacological or genetic manipulation. Mechanistically, this occurred due to the accumulation of ROS and decreased vascularization of the BM due to reduced VEGF expression[108]. Even though this has been successful in mouse models, HIF-2α is required to prevent endoplasmic reticulum (ER) stress in healthy HSCs/HSPCs. Therefore it remains to be determined how dependence on hypoxia could be exploited to selectively target LSCs in humans[109].

1.3.7 The role of transcriptomics to unravel LSC heterogeneity.

The advent of -omics technologies, first bulk sequencing methods on sorted populations and in recent years single cell multiomics, has allowed for the functional and phenotypic profiling of LSCs to be conducted with incredible resolution.

The first high plex transcriptome microarray study on CML LSCs was conducted in 2001 by Ohmine *et al.*, utilizing healthy and leukaemic $CD34^+CD38$ cells. They found the gene PIASy to be downregulated during CML progression, and that its inhibition in KCL22s induced apoptosis[110]. Subsequently *Zheng et al.* in 2006 (18,400) genes, and then Blanco *et al.* in 2007 (8746 genes), investigated $CD34^+$ cells from healthy/CP individuals also using microarrays[111,112]. These experiments were followed up by multitudes of RNA-seq and microarray experiments from human and murine CML models, asking more specific questions in more specific phenotypes. For example, investigating the differences between healthy/leukaemic progenitors and stem cells, the effects of different generations

of TKI on LSCs, and the differences in transcriptional response between responders and non-responders[113].

The first scRNA-seq dataset that focused on CML LSCs was published by Giustacchini *et al.*, where >2000 $Lin^{-}CD34^{+}CD38^{-}$ cells were sequenced[114]. This dataset included numerous different cells of interest, for example, cells from normal HSCs, and from patients in the CP, pre-BP, BP, and after TKI treatment for months to years, and from good and poor responders. Importantly, they also multiplexed cells with BCR::ABL1 specific primers during sequencing, allowing $BCR::ABL1^{+}$ LSCs to be distinguished from $BCR::ABL1^{-}$ bystanders. Numerous discoveries were made from this dataset, which are still relevant to today's research landscape. For example, constructing a gene signature capable of discriminating bystanders from LSCs, characterizing LSC treatment response heterogeneity, and creating gene signatures associated with poor and good responses. Additional scRNA-seq experiments built upon this foundation, with Warfvinge *et al*. showing that CD26 and CD35 could distinguish LSCs from bystanders using Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-seq) and flow cytometry, and that optimal TKI responders possessed a greater MK burden at diagnosis compared to suboptimal responders[115]. Scott *et al*. showed that quiescent LSCs possess an embryonic-like gene signature which proved to be a targetable vulnerability in PDX models[116]. Another important advance in diagnostics was made by Krishnan *et al.,* who utilized scRNA-seq and mass cytometry from the BM of normal, optimal, sub-optimal and non-responders[117]. From this, they show a set of gene signatures for LSCs and natural killer cells allowing for prediction scores with >80% accuracy to be made regarding treatment responses.

1.3.8 TKI response groups and disease categorisation

To determine disease burden and the efficacy of TKI treatment for an individual, qPCR is used to measure the relative quantity of BCR::ABL1 to ABL1 transcripts in the blood. These values are reported as a percentage of BCR::ABL1 to ABL1 on a logarithmic scale, with log reductions of 2 (1% BCR::ABL to ABL1), 3 (0.1%), 4 (0.01%), 4.5 (0.0032%) and 5 (0.001%) respectively corresponding to complete cytogenic remission (CCyR), and major molecular responses (MMRs) 3-5. To assign risk categories, relative BCR::ABL1 abundances are measured at 3, 6 and 12 months, with the 2020 European LeukemiaNet (ELN) guidelines stating that from these values, individuals can be categorised as being treatment Optimal, Warning, or Failure (Table **1-2**)[118]. Additional parameters can also be factored in to predict risk, such as BCR::ABL1 mutation status, and risk scores such as EUTOS Long-Term Survival (ELTS) and Sokal, which use parameters such as age, spleen size, platelet count, percentage of myeloblasts, blasts, and eosinophils.

When considering how patients will respond to TKI, LSC persistence and resistance are separate issues. Persistence refers to LSCs which survive TKI treatment due to extrinsic/intrinsic factors where BCR::ABL1 is effectively inhibited, for example metabolic adaption, enhanced autophagy or microenvironmental factors. Resistance refers to when BCR::ABL1 is not sufficiently inhibited, preventing signal attenuation, for example BCR::ABL1 mutations, increased expression of BCR::ABL1, or enhanced activity of TKI efflux transporters. Contextualising this, out of 100 individuals diagnosed in the CP, after 5 years of TKI treatment, 12 individuals will have achieved treatment free remission, 62 will remain on TKI due to LSC persistence, while 26 will have failed treatment (resistance), despite TKI switching[72]. Considering initial response rates, within 12 months of taking IM or Dasatinib *Kantarjian et al*. found that CCyR had been achieved in 66% and 77% of individuals respectively, with 28% and 46% of these individuals also achieving =>MMR4 (where >CCyR is an Optimal response at 12 months), while 3.5% and 1.9% individuals progressed to the BP (Failure)[119]. A study by *Saglio et al*. showed that at 12 months of taking either IM or Nilotinib, atleast CCyR had been achieved in 65% and 78% of individuals respectively, while =>MMR4 had been achieved by 22% and 44% of individuals, with rates of progression to BP being 4% and <1%[120]. Taken together, this indicates disease persistence (which Optimal and some Warning individuals, depending on other risk factors, fall into) should be the main focus,

as it affects more individuals. However, when instead considering mortality, investigating mechanisms of resistance (treatment Failure) would be preferable. Having the option of effective treatments for the BP would be ideal, however, to avoid the associated heterogeneity and genetic instability it would be more practical to identify individuals who are more likely to progress at diagnosis, and to provide more specific treatments while in the CP. In recent years there have been numerous methods developed to predict TKI responses and to understand the features of treatment resistant LSCs, however these are yet to be clinically impactful[115,117].

Table 1-2 2020 ELN guidelines. TKI response categories, based on the table 4 from the 2020 ELN recommendations[118].

Timepoint	Optimal	Warning	Failure
Baseline	NA	High-risk ACA, high-risk ELTS score	NA
3 months	≤10%	>10% **(no CCyR)**	>10% if confirmed within 1–3 months **(no CCyR)**
6 months	≤1% **(=>CCyR)**	>1–10% **(no CCyR)**	>10% **(no CCyR)**
12 months	≤0.1% **(=>MMR3)**	>0.1–1% **(CCyR only)**	>1% **(no CCyR)**
Any time	≤0.1% **(=>MMR3)**	>0.1–1%, loss of ≤0.1% (**CCyR only, or move from =>MMR3 to CCyR)**	>1%, resistance mutations, high-risk ACA (**no CCyR**)

1.4 *Ex-vivo* HSC culture

1.4.1 Introduction to *ex-vivo* HSC culture

The capacity to grow HSCs *ex-vivo* has been long sought-after, given the near limitless therapeutic value this would bring. Potential applications include stem cell transplant for haematopoietic and immunological malignancies, synthetic blood for transfusion and anti-ageing therapies. HSC culture can be divided into two broad categories, where HSCs are to be expanded, and where HSCs are to be maintained in a quiescent state similar to their *in-situ* behaviour. Additionally, there are methods where HSCs/HSPCs are cultured on their own (2D) in conjunction with other cells (2.5D), or on a scaffold or matrix-like substance to mimic the BM environment (3D) (Figure 1-6). Despite these advances, major challenges in HSC culture are still unresolved, such as increased stress levels resulting in disturbed proteostasis, accumulation of ROS, and issues arising from replicative stress and DNA damage.

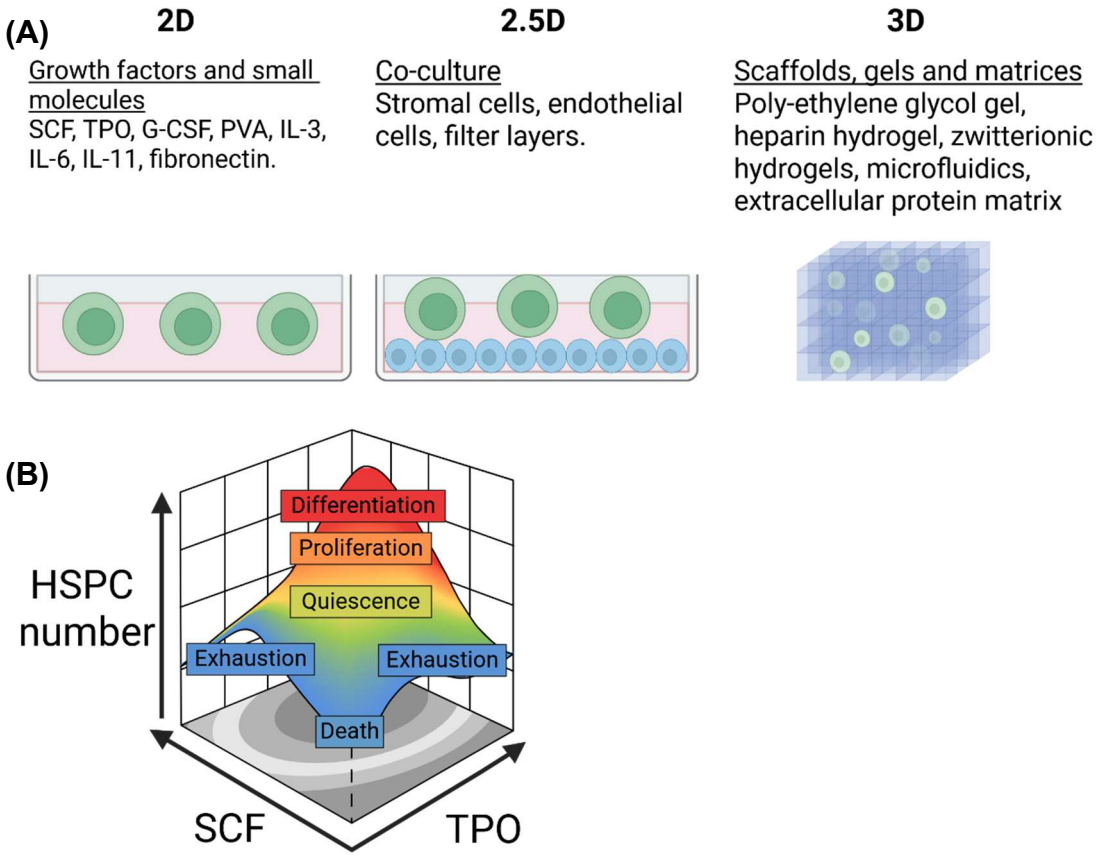


Figure 1-6 Methods of HSC expansion and cytokine requirements. (A) Different components of 2D, 2.5D and 3D HSC culture. (B) Cytokine requirements for TPO/SCF based HSC culture. Too little cytokine results in cell death, too much results in differentiation, while quantities in between result in maintenance of quiescence or proliferation. Figure (B) based upon the graphical abstract of Kobayashi *et al.*[107].

1.4.2 2D culture

2D culture aims to provide a minimal cytokine milieu, usually replacing Fetal bovine serum (FBS) with BSA, or a chemical substitute to allow for a fully defined media. The most widely utilized and successful 2D murine HSC expansion method was published by Wilkinson *et al.*, utilizing SCF (10ng/ml), TPO (100ng/ml) and polyvinyl alcohol (PVA) to expand HSCs, which was improved upon by culturing under physoxic conditions (5% oxygen)[121,122]. PVA based expansion facilitated a 236-899 fold change of functional, transplantable HSCs in both irradiated and non-irradiated hosts, and has provided the means by which to measure HSC phenotype and clonal haematopoiesis more easily. Utilising this method, *Che et al.* found that single LT-HSC clones and their progeny had heterogenous functional potential, and that EPCR and Fgd5, or a gene signature, could better demarcate functional from non-functional HSCs *ex-vivo*[123,124]. Furthermore, this knowledge allowed for *Becker et al.* to expand HSCs and conduct gene editing with much greater effectiveness on $CD48^-/EPCR^+$ single clones, as their colonies had a greater likelihood of transplanting[125]. A similar system was pioneered by Kobayashi *et al.*, who used low concentrations of SCF (3ng/ml) and TPO (0.1ng/ml) under hypoxic (1% oxygen) conditions, alongside high levels of fatty acids. This allowed for the culture of quiescent and long-term engraftable HSCs for up to 1 month in culture[126]. Deviating from the use of SCF and TPO, Oedekoven *et al.* more minimally showed that quiescent HSCs could be cultured only using IL-11 (20ng/ml), without a strict requirement for serum or a serum replacement[127]. An alternative method capable of reducing reliance upon mice for the supply of transplantable cells is the genetically modified haematopoietic progenitor cell line, the HPC^{LSK}. HPC^{LSKs} can be cultured using interleukin-6 (IL-6) and SCF at quantities titrated from cytokine-producing cells[128]. This stem and progenitor like cell line is generated from LSKs transduced with LIM homeobox 2 (Lhx2), which are capable of long-term engraftment alongside culture for many months before losing stem activity. Lhx2 is normally expressed in embryonic HSCs, and is also overexpressed in CML, having the capacity to downregulate genes involved in differentiation, and upregulate those involved in stemness. Aside from using these cells to investigate haematopoiesis, Doma *et al.* also further modified these cells by knocking in p210, the main BCR::ABL1 isoform, and other oncogenes involved in haematopoietic malignancies. While this method has some upsides compared to

the Wilkinson *et al*. method, the HPC-LSKs are not phenotypically true HSCs (LSK $CD48^{+}CD150^{+}/CD150^{-}$) and require FBS, however it is possible PVA could work as a suitable replacement in this platform. Welner *et al*., utilized the Wilkinson *et al* method of HSC expansion on murine LSCs showing that they could be expanded *ex-vivo*. However, these cells could only be expanded for 14 days before exhaustion, and could not consistently reconstitute the BM and cause CML in the recipient[129]. A different set of cytokines may be required for LT-LSCs to be expanded, which was further investigated by Shah *et al*. They found that long term engrafting and disease causing murine LSCs had low c-kit expression within the LSK compartment, and that unlike HSCs, SCF expanded multipotent LSCs and not LT-LSCs[130]. It is possible that engraftable and disease causing LSCs cannot be expanded *ex-vivo* without other elements of the BM to support, or without better understanding the nutrient requirements and cytokine usage of LSCs.

Despite research into human HSC culture, it is unfortunately less advanced than murine expansion. Recently, Sakurai *et al*. tested the Wilkinson *et al*. method of HSC expansion using cord blood, finding expansion of $CD34^{+}CD38^{-}$ cells to only be 3-4 fold over 7 days, whereas murine expansion of HSPCs was around 18-fold[131]. They found that compared to mouse cells, cytokine stimulation of the human cells was less effective, with reduced levels of phosphorylated PI3K, JAK–STAT and MAPK. Instead, they screened chemical agonists for these pathways as a replacement for cytokines and for an alternate to PVA, allowing for an improved expansion rate of $CD34^{+}$ cells of approximately 50-fold over 1 month. Another recent advancement has shown that if a HSC is to lose its functional potential *ex-vivo*, this occurs before the first division between 6 and 24 hours, during initial adaption to culture conditions. This loss of stemness could be counteracted using a JAK inhibitor[132]. This could enhance previous models of HSC expansion by allowing for the retention of a larger fraction of starting HSCs.

When considering initiating cell numbers, phenotype and lineage bias in *ex-vivo* culture or transplantation, it was found that when mouse BM was collected and processed under 3% oxygen, up to a 5-times greater number of HSCs ($LSK\text{-}CD48^{-}CD34^{-}$) could be recovered[133]. Exposure to normoxic conditions for as little as 30 minutes majorly impacted stem cell frequency as well as functional potential as proven by primary and secondary transplantation. When repeated using human

cord blood a 3-fold increase in HSCs was observed. Even though this method massively increases HSC yield it has not been adopted by the wider community, likely due to practicality and access to a hypoxia chamber. However, it does invite the question of whether all experiments regarding HSC phenotype and function that do not begin with a hypoxic harvest are already introducing bias toward which HSCs are being studied.

1.4.3 2.5D culture

Co-culture methods for both murine and human HSCs most commonly use MSCs or ECs to recapitulate the BM environment, as they have been proven to directly interact with HSCs, and produce factors critical for HSC maintenance and growth such as SCF and CXCL12. HSCs can be grown directly upon this layer or separated by filter layers with the former being more efficient.

To determine which cells might be capable of supporting HSC maintenance and growth, >100 stromal cell lines and clones from these lines were used for co-culture to measure LTC-CFC and transplantation assays[134]. This showed that numerous foetal stromal cells independent of starting tissue of origin were capable of supporting transplantable HSCs, with stromal cells from urogenital ridges showing the greatest supporting potential. Utilizing stromal cells in co-culture, *Jing et al*. further characterised their interactions with HSCs, through kinetic measurements of non-adherent HSCs, those adhering to the surface of the stromal layer (phase-bright), and those beneath the stromal layers (phase-dim)[135]. They showed that phase-bright cells had the greatest cycling rates, while phase-dim cells cycled more slowly but retained a more immature phenotype. They also showed that phase-bright and non-adherent cells lost stemness more quickly than phase-dim cells, and at an indistinguishable rate to non-adherent cells. Additionally, migration of cells below the stromal layer could be disrupted by inhibiting integrin ß1 or CXCR4, indicating that the homing of HSCs to specific locations enriched for stromal cells is critical to retention and stemness. More recently, the Essers group published an abstract where they developed a workflow for the purification of a subtype of BM stromal cell known as reinvigorating Mesenchymal Stromal Cells, alongside co-culture with these cells. They state that this method supports both bulk and single cell expansion of functional murine HSCs which outperforms all current methods regardless of the system/method used.

When published, this could represent another leap forward in our understanding of HSC/LSC biology[136].

1.4.4 3D culture

The BM environment itself is soft with a wide range of stiffnesses, dependant on the location. Hydrogels aim to mimic this environment from both a structural and chemical perspective to provide HSCs their required niche without supporting cells, and are comprised of water, proteins and polysaccharides, alongside cytokines which can be supplied in an attached or soluble state[137].

It was shown that $CD34^+$ cells cultured in a zwitterionic hydrogel could be expanded 73-fold, and that these cells could engraft for at least 24 weeks in immunocompromised mice[138]. Mechanistically, the hydrogel reduced ROS levels most likely due to its suppression of oxidative metabolism, similar to the findings of Igarashi *et al*. who treated expanded murine HSCs with an OXPHOS inhibitor IACS, which greatly improved phenotypic stem cell frequency in culture[122]. Another study utilized SCF and VCAM-1 immobilized in a lipid bilayer to mimic stromal cell cytokine signalling. They found that in conjunction these cytokines modulated HSC morphology through FOXO3a activation, and that soluble SCF disrupted this activity, suggesting the method of cytokine presentation altered downstream signalling in HSCs[139]. Another study showed that polyethylene glycol (PEG) gels could be used to simulate the extracellular matrix by attaching SCF and interferon-γ resulting in an expansion of murine LSKs. However the functional potential of these cells was never measured, leaving the effectiveness of this method unclear[140]. Gvaramia *et al*. utilized human HSCs and starPEG–heparin hydrogel with attached cytokine and interestingly found variations in gel stiffness modulated cell cycle states, with greater stiffness corresponding to increased quiescence[141]. A recent study by Donnelly *et al*. explored the effects of numerous mechanical properties on the expansion of $CD34^+CD38^-$ cells and MSCs. Using this knowledge they developed a method allowing the expansion of HSCs without additional media supplements, allowing for CRISPR gene editing and engraftment potential similar to that of fresh cells[142].

1.5 Branched chain amino acids

1.5.1 Introduction to BCAA and metabolism

The BCAAs (valine, leucine and isoleucine) represent 3 of the 9 essential AAs and are hydrophobic in nature, playing critical roles in metabolism and protein synthesis. As there are no *de-novo* synthesis pathways for the BCAAs in mammals, they must be transported into cells. This is facilitated by the L-type amino acid transporter (LAT) family. The LAT family, also referred to as solute carrier (SLC) proteins, are NA^+-independent transporters, with LAT1 being the most researched of the family[143]. After transport into the cell, BCAAs undergo two reactions which are common, followed by numerous additional reactions which are mostly unique to each of the BCAAs (Figure 1-7). Firstly, BCAAs undergo reversible transamination, catalysed by branched-chain aminotransferases 1 or 2 (BCAT1/2). BCAT1 is cytoplasmic while BCAT2 is mitochondrial, with each enzyme having tissue-specific localization. While this reaction is reversible, it more commonly occurs in the direction as described and in the case of mitochondrial BCAA catabolism, the transporter SLC25A44 is responsible for further transportation[144]. Transamination results in the formation of a corresponding branched-chain *α*-keto acid (BCKA) which are 2-keto-isocaproate/4-methyl-2-oxopentanoic acid (KIC), α-keto-ß-methylvaleric acid/3-methyl-2-oxopentanoate (KMV), and 2-keto-isovalerate/3-methyl-2-oxobutanoic acid (KIV), generated from leucine, isoleucine and valine respectively. Additionally, this reaction converts *α*KG into glutamate. The second and rate limiting step is the irreversible oxidative decarboxylation of the BCKA into a corresponding coenzyme A (CoA) compound catalysed by the branched-chain *α*-keto acid dehydrogenase (BCKDH) complex, which occurs in the mitochondria (isovaleryl-CoA from KIC, 2-methylbutyryl-CoA from KMV, and isobutyryl-CoA from KIV). The BCKDH complex is negatively regulated through phosphorylation which is mediated by branched-chain keto acid dehydrogenase kinase (BCKDK) and Mg^{2+}/Mn^{2+}-dependent 1K protein phosphatase (PPM1K)[143]. From here, the acyl-CoAs are separately metabolised, with isoleucine ultimately producing propionyl-CoA and acetyl-CoA, leucine producing acetoacetate and acetyl-CoA and valine becoming propionyl-CoA[145]. It has also

been shown that leucine can bind to the protein *Sestrin2* to activate mTORC1, inhibiting autophagy alongside activating protein and lipid synthesis[144].

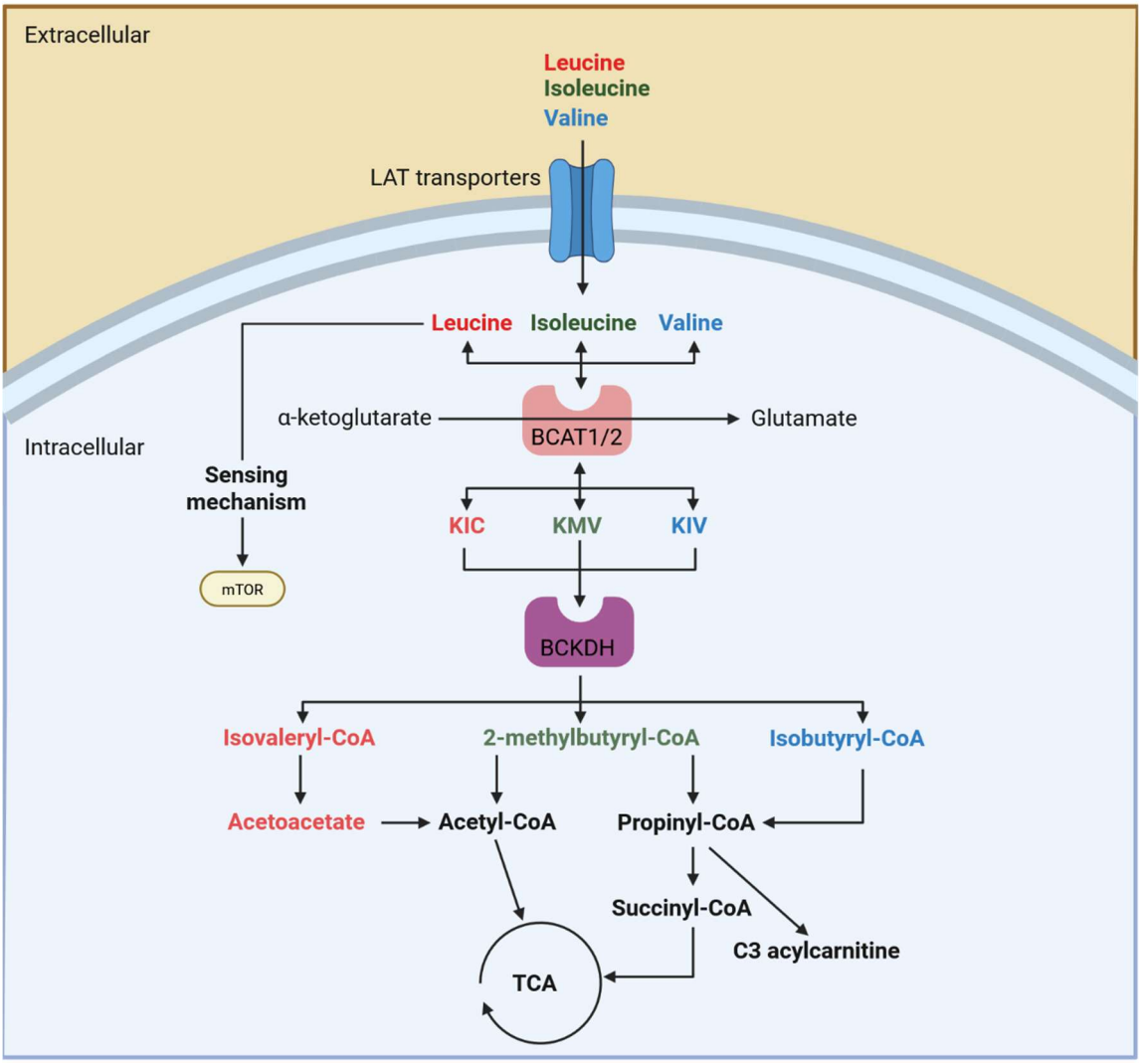


Figure 1-7 The BCAA metabolism pathway. Branched chain amino acids are transported into the cell by the LAT transporters. Leucine levels can be detected by mTOR leading to its activation and protein synthesis upon an increase in its concentration. BCAAs undergo transamination catalysed by branched-chain aminotransferase (BCAT) enzymes forming a corresponding branched-chain *a*-keto acid (BCKA) (2-keto-isocaproate/4-methyl-2-oxopentanoic acid (KIC), α-keto-ß-methylvaleric acid/3-methyl-2-oxopentanoate, (KMV), and 2-keto-isovalerate/3-methyl-2-oxobutanoic acid (KIV). This reaction can also convert αKG to glutamate. BCKAs are converted into a co-enzyme A (CoA) compound by the branched-chain *a*-keto acid dehydrogenase (BCKDH) complex (isovaleryl-CoA from KIC, 2-methylbutyryl-CoA from KMV, and isobutyryl-CoA from KIV). Acyl-CoAs are separately metabolised, with isoleucine ultimately producing propionyl-CoA and acetyl-CoA, leucine producing acetoacetate and acetyl-CoA and valine becoming propionyl-CoA.

1.5.2 The role of BCAAs in myeloid malignancies

BCAA dysregulation has been shown to be involved in numerous myeloid malignancies such as AML, acute lymphoblastic leukaemia (ALL) and CML, alongside metabolic diseases such as type-2 diabetes and metabolic-associated fatty liver disease.

A recent publication by *Jiang et al*. investigated the difference between BCR::ABL1/BCR::ABL1^{T315I} cells, finding that BCR::ABL1 cells had far greater expression of BCAT1 and PPM1K compared to BCR::ABL1^{+} cells[146]. Furthermore, knockdown of BCAT1/2 in BCR::ABL1^{T315I} cells greatly increased survival and reduced leukaemic burden, with higher levels of BCAT1 expression/activity corresponding to increased CREB phosphorylation. Hattori *et al*. showed that BP CML mice had significantly higher levels of BCAAs compared to CP CML mice, and that inhibition of BCAT1 resulted in impaired growth of BC cells *in-vitro* and *in-vivo*[147]. Mechanistically they showed that the leukaemogenic role of the BCAAs was occurring due to the interaction between Musashi RNA binding protein (MS)-12 and BCAT-1 transcripts, enhancing its translation and causing enhanced mTOR activation from the conversion of BCKAs back into BCAAs. In BC mice compared to CP there was no significant difference in BCKA concentrations which would be expected, however a large increase in αKG concentrations was detected, which could be attributed to this reaction from the conversion of glutamate to αKG.

It was also shown that CD34^{+} AML (and ALL to a lesser extent) cells had higher concentrations of BCAAs compared to healthy CD34^{+} cells, but had lower levels of BCAA in serum compared to healthy controls[148]. AML/ALL cells also had greater expression of BCAT1 causing enhanced levels of BCAAs utilizing αKG. Using dietary restriction of BCAAs and PDX models it was shown that this intervention reduced engraftment and self-renewal potential of primary AML/ALL cells, with this reduced stemness being attributed to the inactivation of polycomb repressive complex 2 (PRC2). PRC2 is an epigenetic regulator of stemness, specifically the signalling of its enzymatically active subunit enhancer of zeste homolog 2 (EZH2) and embryonic ectoderm development (EED) was attenuated, causing loss of stemness. PRC2 and EZH2 have also been shown to be dysregulated in CML and to be critical to the maintenance of CML LSCs. For example, it has been shown that PRC2 components were upregulated in CD34^{+}CD38^{-} LSCs compared to their HSC

counterparts, with a large downregulation of EZH1 and a large upregulation of EZH2[113]. TKI treatment *in-vitro* caused a downregulation of EZH2 but left EZH1 expression unchanged suggesting a difference in TKI sensitivity between PRC2 components. Using PDX models it was shown that EZH2 inhibition further sensitized LSCs to TKI treatment. More recently it has been shown that the oncogenic function of EZH2 in CML is partially exerted through altered mRNA splicing, possibly through CELF2[149]. Another study showed that a BCAT1 knockdown in AML cell lines resulted in a reduction in viability and colony forming capacity through αKG accumulation, ultimately causing HIF degradation[150]. When investigating BCAAs as a risk factor for cancers, BCAA plasma concentrations have both positive (pancreatic cancer) and negative (colorectal cancer) risk associations, however for CML, AML and chronic lymphocytic leukaemia (CLL) no relationship could be established, with BCAA levels actually being inversely correlated to ALL risk[151–153].

As previously mentioned, the capacity of a cell to engraft and reconstitute a hosts haematopoietic system is the gold standard of HSC activity, however this requires irradiation to create space in the BM or very large numbers of donor cells. Irradiation destroys most BM cells and can be problematic because even at sublethal doses can cause mice to suffer from severe weight loss. Additionally, as some HSCs are dependent upon the presence of other BM cells which may not be present in the transplant material, cells which engraft may not be truly representative of the initial HSC content. Taya *et al*. showed that dietary restriction of valine could be used to deplete HSC content over the course of 1-4 weeks (~50% decrease), with a maximal decrease occurring at week 4[48]. After two weeks 5000 LSKs were transplanted into valine$^-$ or control diet mice which were subsequently switched to a complete diet. While control mice showed no engraftment, valine$^-$ mice showed between 5 and 25% PB engraftment, symptomatically these mice were healthier than if irradiation had been used, for example 15/15 mice lost fertility post irradiation, while 12/13 mice on a valine-diet retained fertility. Human CD34$^+$38$^-$ cells were also shown to be sensitive to valine depletion, indicating that a similar approach could be used in humans for transplantation, avoiding the damaging effects of irradiation. This method was improved upon by Wilkinson *et al,* who showed that restrictions of all BCAAs was better tolerated *in-vivo* due to the avoidance of amino acid imbalance-associated

toxicity, but with highly similar effect sizes[154]. The mechanism behind HSC depletion was not elucidated, with the authors of these papers showing that it wasn't caused by ROS accumulation or the inhibition of autophagy. Depletion was also unlikely to be occurring through a EZH2 dependant manner given that it has been shown that *EZH2*^{KO} murine embryos had no defects and that there was no difference between HSC frequencies at 9 and 15 days. It was also shown that engraftment of *EZH2*^{KO} HSCs did not impair engraftment potential in adult HSCs, strongly suggesting that the effect was EZH2 independent in HSCs, but not in myeloid malignancies[155]. Conversely, *Eed*^{KO} pups were pale runts, becoming fatally infected within 2-weeks of birth. When the BM of *Eed*^{KO} mice was further investigated it was found that within the lineage- fraction no c-kit expression was detected, however a CD48^{-}CD150^{+} population was present within the Lin^{-}Sca-1^{+}Kit^{-} phenotype, whereas this population was negligible in WT BM within the same phenotype. Unsurprisingly, competitive transplantation of *Eed*^{KO} vs WT BM cells showed greatly diminished *Eed*^{KO} engraftment potential with no meaningful levels of engraftment present by 4 weeks[155]. Given what is known about BCAAs in myeloid neoplasms and PRC2 it is possible, but unproven, that BCAA induced depletion of HSCs is EED mediated.

As previously outlined there have been numerous studies investigating BCAA restriction in murine models, however applying dietary restriction to humans is a more complex matter given the complex relationships humans have with food. While there have been no studies involving BCAA restriction for myeloid malignancies in humans, they have been utilized in trials to treat other metabolic diseases, for example, circulating BCAA levels have been shown to be associated with the pathogenesis of type 2 diabetes[156]. *Karusheva et al*. aimed to determine whether reducing BCAA intake could improve insulin sensitivity, and after 4 weeks found no effect on whole body and hepatic insulin sensitivity. However, meal derived insulin secretion was 28% lower and adipose tissue was more metabolically active[157]. More recently this study was improved upon through the utilization of an isocaloric and isonitrogenous diet, finding that over 7 days circulating BCAA levels could be reduced by around 50%, which mildly reduced insulin insensitivity ($p=0.096$). Even though the effect sizes in these studies were modest, that these interventions could be conducted and did not cause any adverse effects suggests

that dietary restriction of BCAAs could be a viable option when considering novel treatments for myeloid malignancies.

1.6 Aims

This research aimed to address two questions; Firstly, whether we could identify highly differentially expressed genes and metabolic genes across CML transcriptomic datasets through a meta-analysis. And secondly, whether we could use our meta-analysis results, alongside next generation HSC culture methods to identify the metabolic weaknesses of LSCs.

Chapter 2 Materials and methods

2.1 Materials

2.1.1 General reagents

Table 2-1 Commonly used reagents.

Product	Manufacturer	Catalogue
^{13}C-Glucose	Sigma-Aldrich	389374
2-Mercarptoethanol 55mM	Gibco	21985-023
500 mL filter unit	Thermo Fisher	FB12566504
Acetonitrile	VWR International Ltd	83639.32
Acridine Orange	Thermo Fisher	L13159-14
AMD 3100 octahydrochloride	Bio-Techne	3299
Ammonium acetate (NH4Ac)	Sigma Aldrich	A1542-500G
Anhydrous DMSO	Sigma Aldrich	276855-100ML
Bafilomycin A1	Stratech Scientific	A8627-APE
CCCP	Sigma Aldrich	C2759
CD117 MicroBeads	Miltenyi Biotech	130-091224
Compensation beads	Thermo Fisher	01-2222-42
Compensation beads	Thermo Fisher	01-2222-42
CountBright Absolute Counting Beads	Thermo Scientific	C36950
Dialysed FBS	Merck	F0392-500ML
DNase I	Sigma Aldrich	11284923001
Dulbecco's Phosphate Buffered Saline (PBS)	Gibco	14190-094
Earle's Balanced Salt Solution (EBSS)	Thermo Fisher	24010043
EasyStrainer 40µm, for 50mL tubes	Greiner	542040
EDTA	Acros Organics	118432500
Fc block	BioLegend	553142
Foetal Bovine Serum (FBS)	Gibco	10500064
FxCycle™ PI/RNase Staining Solution	Thermo Fisher	F10797
G-CSF	PeproTech	300-23
Glycine	Sigma-Aldrich	50050-1G-KC
Ham's F-12 Nutrient Mix	Thermo Fisher	11765054
Hank's Balanced Salt Solution (HBSS)	Gibco	14065049
HEPES	Sigma Aldrich	H3375-100G
Hydrochloric acid	Thermo Fisher	7647-01-0
Hydroxychloroquine	Fisher Scientific	10646271
Imatinib Mesylate	LC Laboratories	I-5508
Insulin-Transferrin-Selenium-	Thermo Fisher	51500056
Iscove's Modified Dulbecco's Medium (IMDM)	Gibco	21980
L-Alanine	Sigma-Aldrich	A7627-100G
L-Arginine	Sigma-Aldrich	11040-1G-KC
L-Asparagine	Sigma-Aldrich	11150-1G-KC
L-Aspartic acid	Sigma-Aldrich	11189-1G-KC
L-Cystine 2HCl	Sigma-Aldrich	C6727-25G
L-Glutamic Acid	Sigma-Aldrich	49450-1K-KC
L-Glutamine	Life Technologies	25030-024

L-Glutamine	Life Technologies	25030-024
L-Histidine	Sigma-Aldrich	53370-1G-KC
L-Hydroxyproline	Sigma-Aldrich	81709-1G-KC
L-Isoleucine	Sigma-Aldrich	17403-25G
L-Leucine	Sigma-Aldrich	L8912-25G
L-Lysine HCL	Sigma-Aldrich	62930-1G-KC-
L-Methionine	Sigma-Aldrich	64320-1G-KC-
L-Phenylalanine	Sigma-Aldrich	P2126-1G-KC
L-Proline	Sigma-Aldrich	81709-1G-KC
LS Separation columns	Miltenyi Biotec	130-042-401
L-Serine	Sigma-Aldrich	84960-1G-KC
L-Threonine	Sigma-Aldrich	89180-1G-KC
L-Tryptophan	Sigma-Aldrich	T0254-1G
L-Tyrosine disodium	Sigma-Aldrich	T3754-1G-KC
L-Valine	Sigma-Aldrich	94620-1G-KC
LysoTracker Blue DND-22	Invitrogen	L7525
Magnesium chloride	Sigma Aldrich	208337
Methanol	Fisher Scientific	M/4056/17
Mod TestDiet® 5WA1 (control diet)	IPS Product Supplies	5WA1
Mod TestDiet® 9G3X (BCAA- diet)	IPS Product Supplies	9G3X
Mouse P-Selectin/	Biotechne	737-PS-200
Murine IL-11	Thermo Fisher	220-11-10UG
Murine SCF	Thermo Fisher	250-03-10UG
Murine TPO	Thermo Fisher	315-14-100UG
Nilotinib	LC Laboratories	N-8207
Oligomycin	Stratech	C3007-APE
Penicillin-Streptomycin	Life Technologies	15140-122
Polyvinyl alcohol, 98-99% hydrolysed,	Thermo Fisher	9002-89-5
Potassium Bicarbonate ($KHCO_3$)	Sigma Aldrich	60339
Potassium hydroxide	Sigma-Aldrich	P5958-500G
Propidium iodide	Sigma-Aldric	P4864-10ML
Propidium iodide	Sigma-Aldrich	P4864-10ML
Resazurin	Sigma-Aldrich	R7017-5G
Roswell Park Memorial Institute media	Gibco	31870-025
Rotenone	Sigma Aldrich	R8875
RPMI Medium 1640 without amino	Thermo Fisher	ME24313P1
RPMI Medium 1640 without amino	United States	R8999-02A-
Sodium bicarbonate	Sigma Aldrich	S5761
StemSpan SFEM	StemCell	09600

Table 2-2 Flow cytometry antibodies (mouse)

Product	Colour	Manufacturer	Catalogue
7-AAD	-	BD Pharmingen	51-6898IE
Annexin V	APC	BioLegend	640941
CD150 (SLAM)	APC	BioLegend	115910
CD45	FITC	BD biosciences	553079
CD45.1	FITC	BioLegend	110706
CD45.2	Alexa fluor 700	BioLegend	109822
CD45.2	PerCP/Cy5.5	BD biosciences	552950
CD48	BV421	BioLegend	103414
CD48	PE	BioLegend	103406
c-kit (CD117)	APC-Cy7	BioLegend	105826
EPCR	PE	Thermo Fisher	12-2012-82
Gr1	PE-Cy7	BioLegend	108416
Lineage cocktail	Pacific blue	BioLegend	133310
Lineage cocktail	Af700	BioLegend	133313
Mac1	APC	BioLegend	101212
Sca-1	PE-Cy7	BioLegend	108114
Zombie Green™ Fixable	-	BioLegend	423111
Zombie NIR™ Fixable	-	BioLegend	423105

2.1.2 Equipment

Table 2-3 Key equipment.

Name	Manufacturer
Attune NxT Flow Cytometer	Thermo Fisher Scientific
CASY Cell Counter and Analyser	Roche
Celldrop	Denovix
FACS Fusion	BD Biosciences
FACS Verse	BD Biosciences
Incucyte	Sartorius
Infinite M200 Pro Plate Reader	Tecan
Seahorse XFe96 Analyser	Agilent
Sony ID7000	Sony
Ultimate 3000 HPLC System	Thermo Fisher Scientific
Whitley H35 hypoxystation	Don Whitley Scientific

2.1.3 Commercial media and buffers

Table 2-4 *Ex-vivo* expansion media.

Component	Concentration
Hams F-12 nutrient mix	-
HEPES	10mM
ITSX	1x
PVA	1mg/ml
Penicillin-Streptomycin (Pen/Strep)	1x
Glutamine	1x
TPO	100ng/ml
SCF	10ng/ml

Table 2-5 Cell line culture media.

Component	Concentration
Commercial/formulated RPMI 1640	-
dFBS/FBS	10%
L-Glutamine	2mM
Pen/Strep	100 IU/mL

Table 2-6 Plasmax for cell culture.

Component	Concentration
Formulated Plasmax	-
dFBS	10%
L-Glutamine	2mM
Pen/Strep	100 IU/mL

Table 2-7 DAMP recovery media.

Component	Concentration
PBS	-
Magnesium Chloride	1M
Trisodium Citrate	0.155M
Human Serum Albumin	0.2
DNase I	2500U/ml

Table 2-8 Overnight recovery media

Component	Concentration
IMDM	-
FBS	20%
L-Glutamine	2mM
Pen/Strep	100 IU/mL

Table 2-9 Hibernation culture media.

Component	Concentration
StemSpan SFEM	-
FBS (optional)	10%
Pen/Strep	1%
L-Glutamine	1%
2-Mercaptoethanol	0.2%
IL-11	20ng/mL

Table 2-10 Red blood cell lysis buffer.

Component	For 1L
Milli-Q water	-
Ammonium acetate (NH_4Ac)	8.02g
Potassium Bicarbonate ($KHCO_3$)	1.00g
EDTA	0.02g

Table 2-11 Freezing media.

Component	Concentration
FBS	90%
DMSO	10%

Table 2-12 Metabolite extraction buffer.

Component	Concentration
Methanol	50%
Acetonitrile	30%
Milli-Q water	20%

2.1.4 Custom RPMI

To create Roswell Park Memorial Institute media (RPMI) with the capacity to drop out specific AAs, first a concentrated stock solution was made, which could then be supplemented with powdered RPMI components.

Table 2-13 Concentrations and solvents required to make AA stocks for custom RPMI.

Amino Acid	Solvent	Conc (mM)	Mass (g)
L-Aspartic acid	0.1M KOH	7.5	0.53
L-Cystine 2HCl	0.1M KOH	10.4	0.68
L-Histidine	0.1M KOH	4.8	0.98
L-Isoleucine	0.1M KOH	19.0	0.85
L-Leucine	0.1M KOH	19.1	0.95
L-Lysine HCL	0.1M KOH	13.7	1.58
L-Methionine	0.1M KOH	5.0	1.73
L-Phenylalanine	0.1M KOH	4.5	1.88
L-Tryptophan	0.1M KOH	1.2	2.48
L-Tyrosine disodium	0.1M KOH	6.4	2.63
Glycine	Water	6.7	0.05
L-Arginine	Water	57.5	0.23
L-Asparagine	Water	18.9	0.38
L-Glutamic Acid	Water	6.8	0.83
L-Hydroxyproline	Water	7.6	1.13
L-Proline	Water	8.7	2.03
L-Serine	Water	14.3	1.45
L-Threonine	Water	8.4	2.33
L-Valine	Water	8.5	1.85

As specified by Table 2-13, AAs were weighed, and put in either of 2x50ml falcons, depending on whether the AA was potassium hydroxide (KOH) or water soluble. Subsequently, 25ml of either KOH or milli-Q water was added to respective falcons, alongside 2ml 37% hydrochloric acid (HCL) to the KOH falcon. Depending on whether $BCAA^-$ or full media was being made, BCAAs were not added. Falcons were then vortexed thoroughly and sonicated in a warm water bath. After AAs were dissolved, they were combined into one Duran bottle and supplemented further with 2ml 37% HCL. This 50x stock was then stored at 4°C. To make 500ml of custom RPMI, 4.7g of powdered RPMI was added to 400ml of milli-Q water in a Duran bottle, and supplemented with 1g sodium bicarbonate. After these components dissolved, 10ml of the 50x AA master mix was added, and adjusted to pH 7.2 with 1M KOH (approximately 8ml). Subsequently media was passed through a 500ml filter unit, ready for use or storage (4°C). This formulation was created

by Dr Stephen Wilkie. To create RPMI with variable levels of BCAAs, full AA RPMI and BCAA$^-$ RPMI were combined using the appropriate volumes. For BCAA deficient Seahorse media, the same set of steps were taken, aside from the omission of sodium bicarbonate. For BCAA deficient media used for tracing, AA and glucose free powdered RPMI was used in place of the AA free powdered RPMI, which was additionally supplemented with 11.1mM of ^{13}C glucose.

2.1.5 Custom Plasmax

To create Plasmax missing each of the AAs, first stock solutions of each of the individual AAs were made, which were then combined in an n-1 manner to create a master mix AA stock solution.

Table 2-14 Concentrations and volumes required to make AA stocks for custom Plasmax.

Amino acid	Mass(g)	Conc (mM)	Volume (µl)
Glycine	0.0601	80	206.25
L-Alanine	0.0891	100	255
L-Asparagine	0.0656	50	41
L-Aspartic acid	0.0333	25	12
L-Glutamic acid	0.0368	25	196
L-Histidine	0.2096	100	60
L-Isoleucine	0.1312	100	70
L-Leucine	0.1312	100	85
L-Lysine	0.1827	100	110
L-Methionine	0.1492	100	15
L-Phenylalanine	0.1652	100	34
L-Proline	0.1151	100	180
L-Serine	0.1051	100	70
L-Threonine	0.1191	100	120
L-Tryptophan	0.1021	50	78
L-Tyrosine	0.0906	50	74
L-Valine	0.1172	100	115
L-Cystine	0.1202	50	65
L-Glutamine	NA	200	162.5
L-Arginine	0.2107	100	32
Taurine	0.1252	100	65
L-Citrulline	0.1752	100	27.5
L-Ornithine	0.1686	100	40

As specified by Table 2-14, AA masses were added to 10ml of sterile PBS, then filter sterilized, aside from cysteine and tyrosine which were insoluble. From

these stocks, the specified volumes of AA mixtures were added to a 5ml bijoux tube in an n-1 manner, which was then made up to 2ml with PBS. These AA dropout tubes and single AA stocks were stored at 4°C. To create final medias, components shown in Table 2-15 were made up in Earle's Balanced Salt Solution (EBSS) at the specified concentrations. Following this, media was supplemented with 37% HCL to set the pH to 7.4, followed by filter sterilization. Fully made-up AA deficient Plasmax was then stored at 4°C. This formulation was created by Dr Kevin Rattigan.

Table 2-15 Formulation of custom Plasmax from AA stocks.

Component	Dilution
EBSS	-
Vitamin mix	1:100
Extra nutrient mix	1:100
Salts mix	1:1000
Uric acid	1:500
Amino acid stock mix	1:22.7

2.1.6 Computational resources

Table 2-16 Key packages used for computational and statistical analyses.

Name	Version
ArrayExpress	1.56.0
arrayQualityMetrics	3.52.0
Biobase	2.54.0
biomaRt	2.52.0
clusterProfiler	4.4.4
ComplexHeatmap	2.12.1
dplyr	1.0.9
EnsDb.Hsapiens.v86	2.99.0
EnsDb.Mmusculus.v79	2.99.0
GEOquery	2.64.2
ggplot2	3.4.2
ggprism	1.0.5
hgu133a.db	3.13.0
hgu133plus2.db	3.13.0
hugene10sttranscriptcluster.db	8.8.0
limma	3.50.0
MetaboAnalystR	3.2.0
msigdbr	7.5.1
oligo	1.60.0
oligoClasses	1.58.0
org.Hs.eg.db	3.15.0
pd.hugene.1.0.st.v1	3.14.1
rstatix	0.7.0
scales	1.2.0
Seurat	5.0.3
SeuratDisk	0.0.0.9020
stringr	1.4.0
tidyr	1.2.0
vctrs	0.6.5
VennDetail	1.12.0

Table 2-17 Programmes and webtools.

Name	Version
Featurecounts	2.0.7
FlowJo	10.10.0
Metaboanalyst	6.0
Metabolite autoplotter	2.6
Mutliqc	*v1.26*
R	4.5.1
Skyline	25.1
SRA toolkit	*3.0.0.*
STAR	2.7.11b
Trimgalore	0.6.6
Wave	2.6.4.24

2.2 Transcriptomics and statistics

2.2.1 Microarray processing

Where raw intensity files were available, they were obtained from either Gene Expression Omnibus (GEO) or arrayExpress, read into R and RLE normalised. If raw data was not available, normalized data was instead read into R. Probes with low expression levels were filtered based on the intensity distribution, while PCA was used to visualise sample variability. Probes were annotated with gene symbols using the corresponding probe/annotation sets, which was followed by differential testing using limma.

2.2.2 RNA-seq processing

Where raw fastq files were available, reads were obtained from GEO using fasterqdump, then subject to QC using trimgalore, fastqc and multiqc. Reads were aligned to either GRCm38 or GRCh38.105 using STAR, and counts were quantified using featurecounts. Count matrices were read into R, where genes with fewer than 10 counts across all conditions were dropped. Normalization and differential expression were conducted using DESeq2. From the normalized data, PCA plots were produced, and where datasets had a clear batch or patient effect, batch correction was applied by supplying the corresponding variable as a co-variate to DESeq2. If raw data was not available, the normalized matrix was read into R and limma was used for differential expression.

2.2.3 scRNA-seq processing

To process scRNA-seq data, where matrix, feature barcode and metadata files or other CellRanger outputs were available, these were read into R. If instead an RDS object was available, this was read in. These files were then used to create Seurat objects, and where QC had not previously been applied, cells with <200 features and genes present in <50 cells were filtered. Cells with high mitochondrial counts were removed, based on the threshold set in the corresponding publications. For dimensionality reduction, the first 20 principal components were used, and Seurat clusters were generated with default perplexity and resolution.

SCtransform was used for normalisation and visualisation on UMAPS, while log-normalised values were used to visualise mean expression of a gene on a per-patient basis. To visualise GSE236233 and GSE218184 on their own, they were integrated based upon patient. For differential expression, cells in the clusters of interest were pseudobulked patient wise, where DESeq2 was used to conduct differential expression from the pseudobulked raw counts. Differential expression for GSE76312 was conducted using FindMarkers with MAST, as there were many biological replicates with few cells. Cell type identification as described by section 3.2.

2.2.4 Statistical analysis

Statistical analysis was conducted in RStudio (4.5.1), with error bars representing the standard error of the mean (SEM). Base R, rstatix and ggplot2 packages were used to create plots. Normality was tested for using the Shapiro-Wilk test. A p or padj value of < 0.05 was typically considered to be statistically significant.

2.3 Flow cytometry

2.3.1 Flow cytometer settings

When working from 96 well plates, after staining, cells were made up into a final volume of 200µl/well, 180µl of which was acquired. When working from tubes, final volumes that allowed for acquisition at no greater than 8000 events per second were used. All experiments on the FacsVerse or Attune NxT were compensated using the corresponding fluorescent colours with compensation beads, or cells where required. Experiments on the SonyID7000 were unmixed using appropriate spectral references. Fluorescence minus one controls were used during sorting, and for markers where gating was challenging, such as the SLAM gate and for EPCR. FCS files were exported from the FacsVerse, Attune NxT software or SonyID7000 software to be analysed using FlowJo.

2.3.2 Apoptosis

Stages of cell death can be measured using the translocation of phosphatidylserine to the cell membrane as a proxy. Apoptosis was measured using an annexin V (AV) antibody which has a high affinity of phosphatidylserine alongside 7AAD, a fluorescent dye which binds to double stranded DNA. Utilizing these dyes, cells can be classified as live ($7AAD^{-}AV^{-}$), in late apoptosis ($7AAD^{+}AV^{+}$) or in early apoptosis ($7AAD^{+}AV^{-}$). Cell lines and primary cells were washed in PBS at 300 RCF for 5 minutes, then stained at room temperature in the dark for 20 minutes in a total volume of 100µl of HBSS, with 5µl of 7AAD and 5µl AV (FITC). After incubation. cells were further diluted with HBSS to 200µl, wherein fluorescence was measured via flow cytometry.

2.3.3 Cell cycle

Cell cycle phases were identified using propidium iodide (PI), a fluorescent dye capable of binding DNA. Fluorescence in a cell is proportional to the quantity of dye bound to DNA, G1 phase cells have low DNA content whereas S phase will have more, with G2 phase having twice as much DNA compared to G1. Cell lines and primary cells were washed in PBS at 300 RCF for 5 minutes, then fixed through the drop wise addition of 70% ice-cold ethanol while tubes were vortexed, followed by 3 hours incubation at -20°C. After this, cells were stained with 500µl of PI

staining solution for 20 minutes in the dark at room temperature. Subsequently, cells were diluted with PBS to 200µl and ran on the Attune NxT. Fluorescence was measured in linear mode with the voltage modified so that the G1 peak was set to 200. Analysis was conducted using FlowJo's cell cycle analysis tool.

2.3.4 Lysosomal measurement

Lysosomal content of *ex-vivo* expanded LSCs cultured in $BCAA^-$ media +/- Bafilomycin (Baf) was measured using the fluorescent dye lysotracker blue. After 24 hours of treatment, wells were supplemented with 50nM of lysotracker, followed by 2 hours of incubation at 20% O_2 and 5% CO_2 at 37°C. Cells were then washed, stained with expanded LSC markers, then subject to flow cytometry. TKI treated wells were also run as a control for the activation of autophagy.

2.3.5 Cell type identification

To sort for LT-HSCs or quantify cell frequencies from BM or *ex-vivo* cultured cells, flow cytometry was used. Fresh BM or hibernation cultured HSCs were stained with CD45.1 and/or CD45.2 (if required), Sca-1, c-kit, CD150, CD48 (all at 1:100), a lineage cocktail (1:10), and optionally a zombie dye (1:100) in 100ul total volume of PBS (Figure 2-1A). For *ex-vivo* expanded cells a similar panel was used, aside from the omission of CD45 antibody and the addition of EPCR, if required (Figure 2-1B). To detect PB and spleen chimerism, the same gating strategy as seen in the first three panels of Figure 2-1A was used.

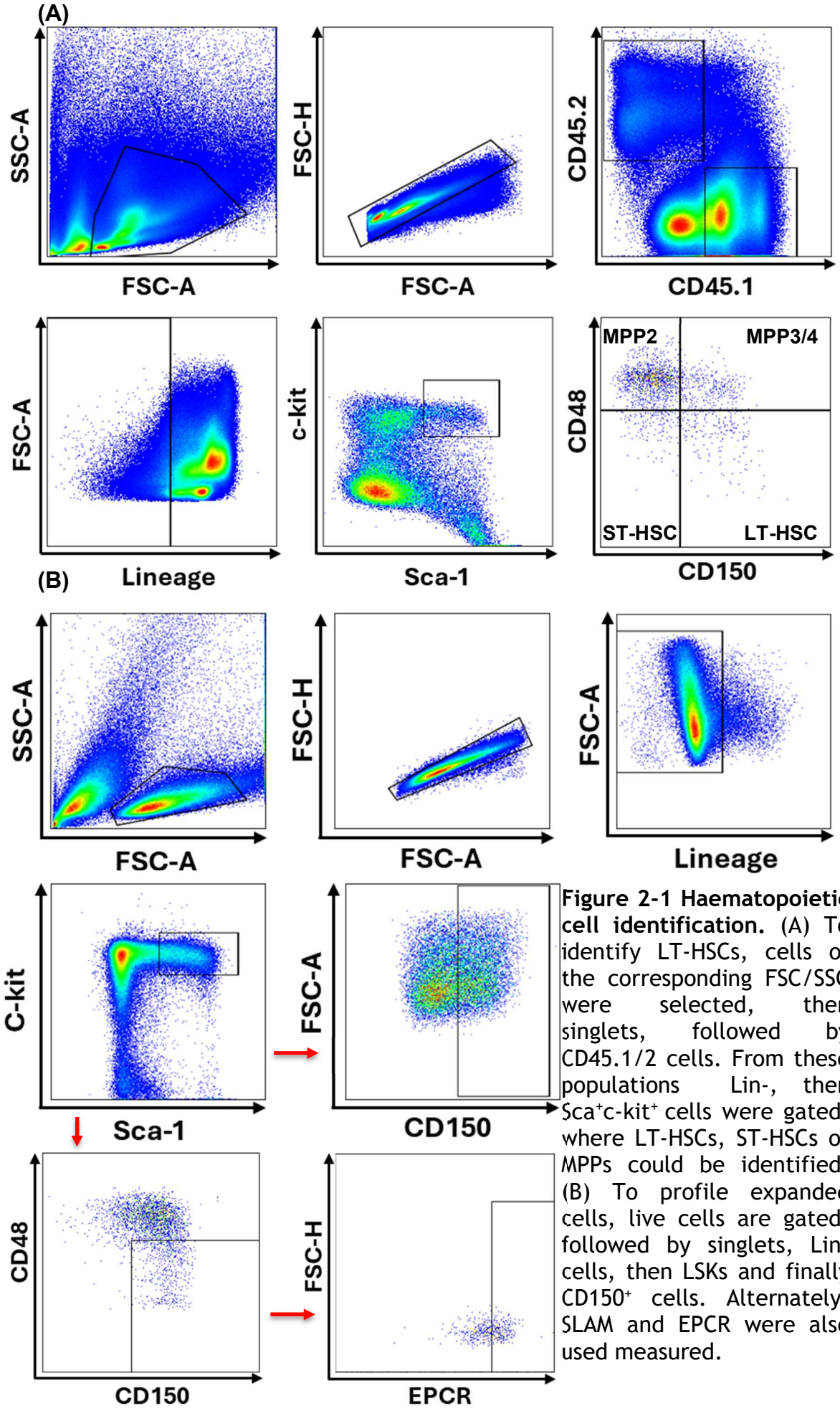


Figure 2-1 Haematopoietic cell identification. (A) To identify LT-HSCs, cells of the corresponding FSC/SSC were selected, then singlets, followed by CD45.1/2 cells. From these populations Lin-, then $Sca^{+}c\text{-}kit^{+}$ cells were gated, where LT-HSCs, ST-HSCs or MPPs could be identified. (B) To profile expanded cells, live cells are gated, followed by singlets, Lin- cells, then LSKs and finally $CD150^{+}$ cells. Alternately, SLAM and EPCR were also used measured.

2.4 Animal work

2.4.1 Licensing

All animal work was conducted on Professor Vignir Helgasons project licence.

2.4.2 Murine model of CML

The inducible model of murine CML known as the double transgenic (DTG) model was created by Koschmieder *et al.*, wherein p210 BCR::ABL1 is expressed in HSPC populations under the control of a tetracycline sensitive element (SCL-Tta). Provision of tetracycline in the drinking water suppresses the expression of BCR::ABL1, and when removed, BCR::ABL1 transcription is initiated, leading to a disease similar to CP CML. Typically tetracycline withdrawal results in symptomatic disease between two and four weeks. Single transgenic (STG) mice are result of the breeding of DTG or STG mice, where only the SCL-tTa promoter or BCR::ABL1 are inherited, meaning that leukaemia cannot be induced.

2.4.3 Non-leukaemic mouse strains

B6.SJL-*Ptprc*α *Pepcb*/BoyJ (Bl6) (JAX Stock #002014) or STG mice were used for *ex-vivo* expansion of HSCs, as transplant recipients, and for dietary intervention. STG/DTG tomato^{+} (GWTM) mice were used for *ex-vivo* expansion, transplant and dietary intervention where specified. These mice were generated by crossing DTG mice with the mTmG mouse (Gt(ROSA)26Sor$^{tm4(ACTB-tdTomato,-EGFP)Luo}$ (JAX Stock #007576), resulting in healthy or leukaemic mice where haematopoietic lineage cells possess cell membrane bound tdTomato, a fluorescent marker protein. NOD.Cg-Rag^{1tm1Mom}Kit^{W-41J}Il2rgt^{m1Wjl}/EavJ (NRGW) mice (JAX Stock #026014) are an immunodeficient strain, and were used to test the effects of dietary intervention.

2.4.4 Blood sampling and processing

To analyse PB chimerism, cell types and metabolites, blood was harvested from the tail vein at volumes between 10-80µl, or as described by 2.4.5. Prior to flow cytometric analysis, red blood cells were incubated with red blood cell lysis buffer at a ratio of 1:100 for 10-15 minutes at room temperature, depending on opacity. Cells were then washed for 5 minutes at 300 RCF, then stained for 20 minutes at

room temperature with CD45.1 and CD45.2, and Gr1 and Mac1 if appropriate, at concentrations of 1:100. To obtain plasma for LC-MS, whole blood was centrifuged at 1000 RCF for 10 minutes, where we aimed to obtain atleast 5µl of supernatant.

2.4.5 Tissue processing

Mice were culled using a CO_2 chamber, and a secondary method of cutting the femoral artery was used. Post cull, blood was obtained from the vena cava, and spleens were harvested and weighed. To obtain BM cells, the hind leg bones were harvested from mice. The ends of these bones were cut, then centrifuged into 20µl of PBS in nested Eppendorfs at maximum speed for 15 seconds. After this, cells were passed through 40µm EasyStrainers into 50ml falcon tubes, with the filters washed with 10ml of PBS. Cells were then counted from this volume. To obtain BM supernatant, after centrifugation (1000 RCF for 10 minutes) into the nested Eppendorfs, 5µl of supernatant was carefully removed.

2.4.6 c-kit enrichment

To initiate culture from BM cells, or to prepare BM cells to sort for LT-HSCs, murine cells were subject to c-kit enrichment. After BM cell harvest, cells were washed in PBS at 300 RCF for 5mins, resuspended in 80µl of PBS and 20µl/mouse of c-kit enrichment beads, incubated for 20mins at 4°C, then washed with 5ml of PBS at 300 RCF for 5 minutes. During this time, enrichment columns with filters were set up on the magnetic stand and washed with 5ml of PBS, and 15ml falcon tubes for the c-kit$^-$ fraction was set up underneath the column to catch this flow through. After the incubation period, cells were resuspended in 500µl PBS and passed through the filter atop the columns. When this volume had flowed through the column, 3x3ml of PBS was used to wash the filter and column into the c-kit$^-$ fraction tube. After this, the column was removed and flushed using the plunger into the c-kit$^+$ fraction tube with 5ml of PBS, twice. Cells from the c-kit fractions were then counted and subject to flow cytometry to evaluate the successfulness of enrichment.

2.4.7 Chimeric models and *ex-vivo* culture transplant

To generate a transplantable model of CML, combinations of DTG, STG, Bl6 and *ex-vivo* cultured cells were used (section 2.5.11). To achieve this, mice between

the ages of 8-12 weeks were irradiated with two doses of 4.25Gy 3 hours apart, then transplanted the next day. Due to irradiation toxicity, this dosage was lowered to 2x3.5Gy, with transplant on the same day. Cells were injected via the tail vein at a volume of 200µl. For transplants conducted prior to October 2024, mice were provided with antibiotics in the drinking water, while prophylactic usage of antibiotics was forbidden after this date. Transplant efficiency was measured through weekly or fortnightly blood sampling, where PB chimerism was measured using CD45.1 and CD45.2 markers. After culling, blood, BM, and spleen were harvested and analysed as previously described (section 2.3.5) .

2.4.8 Generating cells for *ex-vivo* culture

To generate cells for *ex-vivo* culture, DTG mice were induced until symptomatic disease was detected through body condition scoring, weight loss, or increases in Gr1/Mac1 levels. For *ex-vivo* culture from normal HSCs, cells from Bl6 and STG background mice were used.

2.4.9 Dietary intervention

To modify mouse diets, chow provided by the BRU was replaced with either control or $BCAA^-$ diets. All mice were weighed every day for the first 4 days after dietary modification, while we aimed to weigh $BCAA^-$ diet mice daily for the duration of the experiments. Where a singular mouse was required to be returned to control diet and not a cage mate, it was split from its cage. To measure chow consumption, cages were supplied with 110g of chow, this was weighed daily and topped up to 110g, if less than 50g remained. The food consumption metric was calculated by dividing the mass of food consumed since the previous measurement, by the total mass of the mice in that cage.

2.4.10 *In-vivo* drug treatment

To test the efficacy of HSC mobilizing agents, G-CSF, AMD3100 and P-selectin were utilized. Bl6 Mice were subcutaneously treated with either vehicle, one dose of AMD3100 (1mg/kg 30mins before culling) on its own, or alongside P-selectin (1mg/kg twice per day for 1 day), or G-CSF at 125 µg/kg twice a day for two days, or with G-CSF at 250µg/day for 5 days (STG mouse used, conducted at a separate time to the other mice). To sort and transplant cells from the 5-day AMD3100/G-

CSF treated mouse, blood was harvested postmortem via cardiac puncture. From this, 100µl of blood was washed with PBS at 300 RCF for 5mins, then resuspended in 200µl PBS alongside $1x10^6$ Bl6 carrier cells. Cells were filtered, then transplanted into an irradiated Bl6 mouse. The remaining volume was taken for cell sorting.

To treat CML alongside dietary intervention, mice were provided with Imatinib (IM) twice per day via oral gavage at a concentration of 100mg/kg.

2.4.11 IDEXX

To profile the abundance of blood cells, at minimum 50µl of whole blood was used for IDEXX sampling. Clots were removed by pipetting prior to acquisition.

2.5 Cell culture

2.5.1 Cell lines

K562s and KCL22s were available in house.

2.5.2 Culture conditions

CML cell lines were cultured in flasks in RPMI, in an incubator set to 20% O_2 and 5% CO_2 at 37°C. If an experiment using BCAA deficient RPMI was being conducted, dialysed FBS (dFBS) rather than normal FBS was used, alongside the corresponding BCAA deficient media. *Ex-vivo* grown cells were cultured in the relevant growth medias (hibernation or expansion media) in either 96 or 24 well plates, in an incubator set to 20% O_2 and 5% CO_2 at 37°C. For hypoxic culture, *ex-vivo* expanded LSCs were cultured in an incubator set at 1%O_2 and 5% CO_2 at 37°C, while for physoxic culture, cells were cultured in a humidified H35 Hypoxystation (Don Whitley Scientific) at 5% O_2 and 5% CO_2 at 37°C.

2.5.3 Cryopreservation and recovery

To maintain a stock of cell lines K562s or KCL22s were counted, washed in PBS at 300 RCF for 5minutes and resuspended in freezing media at a concentration of 20M cells/ml in 2mL cryopreservation vials. Vials were then place into Mr. Frosty freezing containers, and subsequently transferred to liquid nitrogen the next day.

To recover cryopreserved K562s or KCL22s, cryovials were thawed in the water bath at 37°C for a short period (1-2 minutes). Cells were then washed in RPMI to remove dimethyl sulfoxide (DMSO), then recovered overnight, wherein recovery was examined by microscopy and cell counting the next day.

2.5.4 *Ex-vivo* expansion

To expand HSCs or LSCs *ex-vivo* from a c-kit enriched population, cells were seeded at between 7-9x10^5 cells/ml into a 24 well plate with 1ml of expansion media. If working from a sorted population, 500 LSK-SLAM cells were sorted into a 24 well plate with 1ml of expansion media. Media changes were made every second day (Monday, Wednesday and Friday) when working from a c-kit enriched population, or after the first 5 days, and then every other day if working from a sorted LSK-SLAM population. Media changes were conducted by gently pipetting media out of each well, followed by replacement with fresh pre-warmed media. Wells were split 1:3 when a confluence of 70% was met. When conducting assays on *ex-vivo* expanded HSCs/LSCs, cells were expanded for 7-10 days prior to use. Phenotypes in culture were measured as previously described (2.3.5).

2.5.5 Cell line culture

Cell lines were passaged every second day (Monday, Wednesday and Friday), where densities reached approximately 7x10^5 cells/ml. Cells were split to a density of 2x10^5 cells/ml.

2.5.6 Hibernation culture

To culture HSCs/LSCs in a quiescent state, bones were harvested, c-kit enriched, and sorted using LSK-SLAM markers (2.3.5) into pre-warmed media. Cells were sorted at a density of one cell per well or 1x10^3 cells/ml, in either 200µl or 1ml of hibernation culture media. Cells were cultured for between 3 and 7 days. For single cell cultures, the presence or absence of a cell was determined manually with light microscopy, while for bulk culture a combination of microscopy and flow cytometry was used. To quantify the number of hibernation HSCs used for transplant, counting beads were used, with the cell concentrations calculated as per the manual ((number of cells/number of beads) * (beads per 50µl/sample volume)).

2.5.7 Cell counting

Cells were counted using either the Casy Counter or the CellDrop, two different counters were used due to the Casy Counter being repaired for a long period of time. For the Casy Counter, 20µl of cell suspension was taken and diluted in Casy counter buffer, then run on the machine. For the CellDrop, 20µl of cell suspension was taken and mixed with 20µl of acridine orange-PI. Then 10µl of this mixture was loaded onto the imaging platform, where counts were performed at minimum in duplicate. Cell lines and murine cells were run using the appropriate programmes on each machine. The Attune was also used for counting K562s and KCL22s cultured in different BCAA media concentrations. From a final volume of 200µl in each well, a 180µl volume was acquired, with output cell numbers scaled to the full well volume.

2.5.8 Incucyte

To measure cell confluency and cell death in expanded HSCs and LSCs, the Incucyte was used. Cells were harvested and expanded as previously described, then seeded at $1x10^5$ cells/ml into various concentrations of BCAA deficient media alongside PI at a dilution of 1:1000. Measurements were taken every 4 hours until 72 hours had elapsed.

2.5.9 AA deficient culture

E*x-vivo* expanded LSCs or cell lines were counted, washed twice in PBS, then seeded in the AA deficient media at densities between $1\text{-}2x10^4$ cells/ml. *Ex-vivo* expanded cells were cultured in $BCAA^-$ deficient Plasmax medias, while cell lines were cultured in $BCAA^-$ deficient RPMI, formulation as described previously. For glucose tracing experiments, cells were cultured in the ^{13}C $BCAA^-$ deficient medias for either 24 or 48 hours.

2.5.10 Resazurin

To conduct the Resazurin assay, cell lines in different $BCAA^-$ media concentrations received 10µl of 50X Resazurin stock into 200µl of culture media at 24, 48 and 72 hours. After 2.5 hours of incubation at 20% O_2 and 5% CO_2 at 37°C, plates were removed from the incubator and left at room temperature in the dark for an

additional 30 minutes. After this period, plates were taken to the Tecan plate reader, where fluorescence values were measured. Fluorescence of wells with only media and PBS were measured as controls.

2.5.11 Processing of cells for transplant

Vials of cryopreserved DTG, STG and Bl6 cells were recovered for transplantation by the dropwise addition of warm DAMP media into cryovials. After thawing, each vial (20M cells) was washed in 10ml of DAMP 3 times (3x300 RCF for 5 mins), counted, then seeded into overnight recovery media at a 7-9x10^5 cells/ml. The next day, cells were counted, washed in PBS at 300 RCF for 10 minutes, then resuspended in a total volume of 220µl of PBS per mouse. This volume would contain a total of 1-1.65x10^6 Bl6 or STG cells, alongside either: 1-1.65x10^6 DTG cells, 1.1x10^6 *ex-vivo* expanded cells, or 220 hibernation HSCs. 200µl of suspension was used for transplant.

2.6 Metabolomics

2.6.1 Metabolite extraction for LC-MS

To prepare cells for LC-MS, they were washed with PBS at 450 RCF for 10 minutes at 4°C, resuspended in PBS, and then washed again at 12,000 RCF for 15 seconds. Ice-cold extraction solvent was added to create a final concentration of 1x10^6 cells/ml for mouse cells, 1x10^6 cells/ml for cell lines, 1:50 for BM supernatant, serum, and media. Samples were then vortexed and incubated for 5 minutes at 4°C, centrifuged at 16,100 RCF for 15 minutes, wherein supernatant was transferred into glass metabolomics vials. To extract metabolites from mouse chow, 3 pellets from either $BCAA^-$ or control diets were crushed using a mortar and pestle. From this, 20mg per sample were then weighed, and metabolites extracted using extraction solvent at a concentration of 20mg/ml. After all extraction steps, vials were stored at -80°C, and processed by Alejandro Huerta Uribe or Engy Shokry at the Beatson.

2.6.2 LC-MS setup

To conduct LC-MS, a Thermo Ultimate 3000 High pressure liquid chromatography (HPLC) system (Thermo Scientific) and a Q Exactive Orbitrap mass spectrometer

(Thermo Scientific) were used. The HPLC system was composed of a ZIC-pHILIC column (SeQuant, 150 x 2.1mm, 5µm, Merck KGaA) and a ZIC-pHILIC guard column (SeQuant, 20 x 2.1mm). An initial mobile phase of 20% 20mM ammonium carbonate, pH 9.4, and 80% acetonitrile was used, with 5µl metabolite extracts injected into the system, and separated over a 15 minute mobile phase gradient (acetonitrile content decreased to 20%), at a flow rate of 200µl/minute, and a column temperature of 45^{0}C. The mass spectrometer was then used to detect fragments, with electrospray ionization and polarity switching. All fragments were detected across a mass range of 75-1000m/z, at a resolution of 35,000 (at 200m/z), and with a mass accuracy below 5ppm.

2.6.3 Analysis of LC-MS data

Raw .raw files were obtained from the metabolomics facility, which were loaded into Skyline, alongside a molecule list with expected retention times, molecule names and structures. Peaks were manually checked, and adjusted where Skyline had chosen the wrong peak. Peak intensity files were exported, where metabolite abundance was more closely compared to solvent blank levels, suspect peaks that were not thought to be accurate with reasonable certainty were dropped. Intensity matrices were loaded into metabolite autoplotter or graphed in R to examine the raw peak intensities, and in the case of tracing data, to plot absolute and relative labelling fractions. For differential analysis and pathway analysis, data for each set of comparisons was read into Metaboanalyst, log10 transformed and mean centred. Gene level and pathway level statistics were then generated.

2.6.4 Seahorse

To measure the oxygen consumption rate (OCR) and extracellular acidification rate (ECAR) of K562s and KCL22s in full or BCAA deficient medias, the Seahorse XFe96 Analyser (Agilent) was used. The day before cells were seeded into BCAA deficient media, they were split to a concentration of $1x10^5$/ml. The next day, cells were seeded at $3x10^5$/ml into full or BCAA deficient medias. The day before running the assay, the 96 well Seahorse plate was prepared by coating with a 0.02mg/ml sodium bicarbonate (pH 7.4) Cell-Tak solution. After 30mins of incubation, the Cell-Tak solution was discarded, and plates were washed twice with water and left overnight to dry. Additionally, a XF96 sensor cartridge was

hydrated overnight with calibrant solution in a non-humidified CO_2-free incubator at 37°C. The next day, cells were counted, re-seeded at $1x10^5$/ml in the Seahorse plate in the appropriate BCAA deficient Seahorse media, then centrifuged at 200 RCF for 1 minute without brakes. Cells were then incubated in a non-humidified CO_2-free incubator at 37°C for 30mins. Media from the seeding volume was also counted again, and used for normalisation after the Seahorse was run. To prepare the drug plate, oligomycin (1.5μM), CCCP (1μM), rotenone (1μM) and antimycin A (1μM) were prepared in the appropriate BCAA deficient medias and transferred into the sensor cartridge, which was loaded into the machine for calibration. After calibration and the incubation steps, the cell plate was also loaded. Following data acquisition, raw data was loaded into Wave, with the cell counts used to normalize OCR and ECAR rates, which were subsequently exported for statistical analysis.

Chapter 3 Meta-analysis of CML LSPCs

3.1 Introduction

Over the past 20 years there have been more than 119 CML transcriptomic experiments published, covering numerous different phenotypes in both humans and mice[158]. Given the stem cell driven nature of CML, we aimed to analyse human progenitor ($CD34^+$) and stem and progenitor ($CD34^+CD38^-$, or more primitive) bulk and single cell transcriptomic datasets, with two main aims. Firstly, as proof of principle studies targeting LSC metabolism have previously been successful, we sought to define metabolic changes in progenitors and stem and progenitor populations, and to report genes which were commonly differentially expressed (DE) across datasets in the same phenotypes. Secondly, we assessed the usefulness of CML datasets in the current research landscape, by comparing gene expression changes across different phenotypes. Commonly DE genes should represent potential drug targets, as there can be increased confidence of biological relevance rather than technical or computational artifacts, and that expression of a gene is not tied to a particular cohort or ethnic group.

To achieve this, first we identified CML datasets through literature and repository searches, such as GEO, ArrayExpress and the European genome-phenome archive. Subsequently datasets which profiled cell lines were discarded, alongside those which did not contain any relevant comparisons in at least the HSPC compartment, or had fewer than 100 statistically significant genes (adjusted p value (padj) < 0.05). For microarray datasets, experiments with smaller array sizes ($<5x10^3$) were also discarded. For bulk transcriptomics where raw files were provided we began analysis from these, else the uploaded normalized matrices were used. For scRNA-seq data when available, CellRanger outputs were used to construct Seurat objects, or if .RDS objects were available these instead were utilized. Post-processing we were left with 20 datasets of interest (Table 3-1). After conducting differential expression, genes and pathways from bulk and scRNA-seq datasets within the same phenotype (referred to as a comparison set) were intersected to determine which signals could be consistently detected, and therefore may be predicted to be targetable (Figure 3-1).

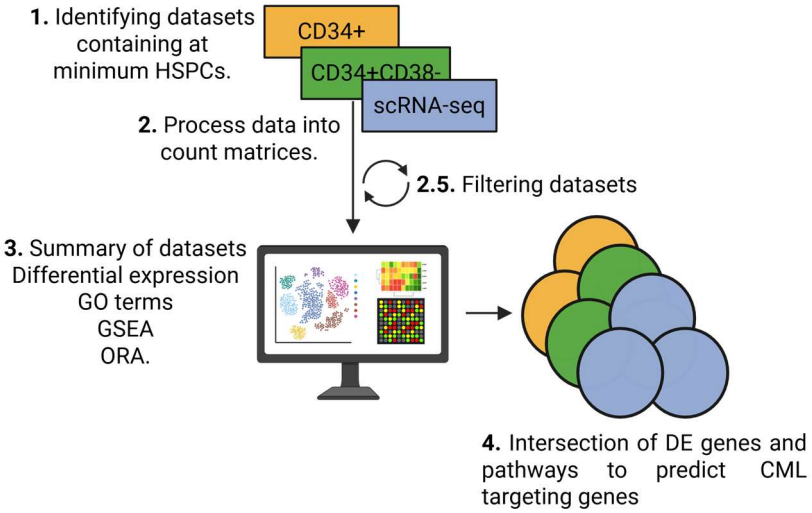


Figure 3-1 Meta analysis workflow. Suitable datasets were identified, subject to quality control, then processed into a raw count matrices, with metadata where possible. Differential testing was then conducted, followed by downstream analysis. Outputs within and between comparison sets were then intersected to discover common features.

Table 3-1 List of datasets used for meta-analysis and associated parameters.

ID	Type	Tiss	Phenotypes used	Comparison	Publication
GSE216837	RNA-seq	PB	$CD34^+$	TKI LSC vs LSC	Rattigan *et al*[94]
GSE198576	RNA-seq	PB	$CD34^+$	TKI LSC vs LSC	Castañeda *et al*[159]
GSE218183	RNA-seq	PB	$CD34^+$	TKI LSC vs LSC	Scott *et al*[160]
GSE48294	Microarray	BM	$CD34^+$	TKI LSC vs LSC	Ng *et al*[107]
E-MTAB-2594	Microarray	PB	$CD34^+$	TKI LSC vs LSC	Pellicano *et al*[161]
GSE12211	Microarray	PB	$CD34^+$	TKI LSC vs LSC	Bruennert *et al*[162]
GSE140385	RNA-seq	PB	$CD34^+$	LSC vs HSC	Sinnakannu *et al*[163]
GSE163690	RNA-seq	BM	$CD34^+$	LSC vs HSC	Youn *et al*[164]
GSE5550	Microarray	BM	$CD34^+$	LSC vs HSC	Abraham *et al*[165]
E-MTAB-2508	Microarray	PB	$CD34^+$	LSC vs HSC	Graham *et al*[166]
GSE24739	Microarray	BM	$CD34^+$	LSC vs HSC	Affer *et al*[167]
GSE4170	Microarray	NA	$CD34^+$	LSC vs HSC	Radich *et al*[168]
E-MTAB-2581	Microarray	PB	$CD34^+38^-$, $CD34^+$ CD38+	LSC vs HSC	Abraham *et al*[165]
EGAD0000100407	RNA-seq	BM	$CD34^+38^-$	LSC vs HSC	Landberg *et al*[98]
GSE97562	Microarray	BM	$Lin^-CD34^+38^-$	LSC vs HSC	Vázquez *et al*[169]
GSE47927	Microarray	BM	$CD34^+38^-CD90^+$	LSC vs HSC	Morales *et al*[170]
GSE76312	scRNA-seq	BM	$Lin^-CD34^+38^-$	LSC, HSC, TKI response	Giustacchini *et al*[114]
GSE218184	scRNA-seq	PB	$CD34^+$ CD38-	HSC vs LSC	Scott *et al*[160]
GSE236233	scRNA-seq	BM	Lin^-CD34^+/$Lin^-CD34^+CD38^-$	LSC vs LSPC	Warfvinge *et al*[89]
EGAS0000100550	scRNA-seq	BM	$CD34^+$	HSC vs LSC,	Krishnan *et al*[171]

Table 3-2 List of comparisons used from scRNA-seq datasets.

ID	Phenotypes used	Comparisons
GSE76312	$Lin^-CD34^+CD38^-$	$BCR::ABL1^+$ vs $BCR::ABL1^-$ Good vs poor responder 3 months TKI
GSE218184	$CD34^+CD38^-$ Predicted primitive	Predicted primitive $BCR::ABL1^+$ vs $BCR::ABL1^-$ LSPC vs HSPC
GSE236233	$Lin^-CD34^+CD38^-$ Predicted primitive	Predicted primitive $BCR::ABL1^+$ vs $BCR::ABL1^-$ LSPC vs $BCR::ABL1^-$ Optimal, warning and failure response groups
EGAS00001005509	$CD34^+CD38^-$	Optimal, suboptimal, failure response groups

3.2 Identifying LSCs from scRNA-seq data

Single cell technologies offer the advantage of identifying rare cell types from heterogeneous populations, however its analysis has more complexity compared to bulk methods, therefore here we describe processing steps used to predict cell types. Using CML scRNA-seq datasets within atleast the HSPC compartment, we aimed to identify LSCs with the most predicted primitive phenotype that could be used for differential expression for a dataset, taking into account whether that same comparison set existed in other datasets, making it more informative (Table 3-2). Of the 4 suitable human datasets, GSE76312 (GSE312) could not be annotated further due to limited cell numbers, while GSE236233 (GSE233) already had extensive cell type annotations. EGAS00001005509 (EGA) had coarse annotations, leaving room for improvement, while GSE218184 (GSE184) had no annotation metadata.

To predict the most stem phenotype possible from GSE184, we used gene expression profiles from GSE233 and GSE312 (Figure 3-2). First, we used the BCR::ABL1$^+$ gene signature generated by Giustacchini *et al*. to differentiate LSCs from BCR::ABL1$^-$ HSCs (bystanders), finding clusters 9 and 11 to score poorly (Figure 3-2A-C)[114]. Based on the presence of known marker genes from GSE233 and GSE312 we excluded cluster 11 as a differentiated cluster and labelled cluster 9 as BCR::ABL1$^-$ cells. Then we conducted differential expression on GSE233 between the most predicted primitive BCR::ABL1$^+$ phenotype and other Lin$^-$ CD34$^+$CD38$^-$ LSCs, and took the top 100 positively expressed features to module score GSE184. From examining the primitive LSC signature on a cluster-wise basis, we saw that cluster 0 had the highest score and was in close proximity to cluster 9, while cluster 11 had the lowest score (Figure 3-2D). We then looked at the expression of CD26, CD90 and CD93 on a patient and cluster-wise basis to discern which clusters were enriched for these markers (Figure 6-3 and Figure 6-4A). From this, we observed that cluster 0 was enriched for CD26, CD90 and CD93, followed by cluster 1, with one individual not contributing any CD90$^+$ cells. While CD93 expression was greater in cluster 0, its expression was not necessarily restricted to CD26$^+$ clusters, potentially indicating a distinct CD26$^+$CD93$^+$ phenotype. We then selected cells from cluster 0 with a primitive module score of > 0.97, representing scores corresponding to the top 75% of primitive LSCs from GSE233 (Figure 3-2B/G). We also integrated GSE233 and GSE184 to see where the predicted

primitive cells were located compared to cells from GSE233 (Figure 3-2E-F). There was some overlap between predicted primitive cells, and close proximity of bystander HSCs in both datasets, and as GSE184 was derived from the PB and GSE233 was derived from BM, we expected some transcriptional differences. Notably, GSE184 bystanders did not express CD90, while the normal HSC clusters (not shown) and GSE233 bystanders did. This potentially indicates that PB bystanders are depleted of functional HSCs. Alternately, it may be possible that bystanders transiently lose CD90 expression while travelling through the PB, as it has been shown that $Lin^{-}CD34^{+}CD38^{-}FLT3^{-}$ cells can transiently lose/gain CD90 expression, although this was cell cycle dependant[172]. Aside from cluster 0, cluster 5 was partially $CD26^{+}CD90^{+}$ positive. We chose not to include this cluster in the analysis, as it and its adjacent clusters were enriched with erythroid differentiation markers (GATA1, KLF1, CD36), indicating they may be erythroid primed LSCs (Figure 6-4A). Additionally, we could not identify this population/structure in GSE233 (Figure 3-2 black circle and Figure 6-4B). Note that for UMAPS displaying cells from GSE233, GSE184 and EGA, BCR::$ABL1^{+}$ labels refer to the most predicted primitive population identified for that dataset, while LSC/LSPC refers to the remaining predicted BCR::$ABL1^{+}$ populations.

As EGA was derived from $CD34^{+}$ cells it was not possible to identify subclusters within the $CD34^{+}CD38^{-}$ compartment (Figure 3-3). However, we demarcated a more stem population compared to the original publication ($CD34^{+}$ to $CD34^{+}CD38^{-}$) with higher BCR::$ABL1^{+}$ scores compared to other LSPC clusters. To do so, we first visualised the dataset by UMAP, then used BCR::ABL1 scoring to identify BCR::$ABL1^{+}$ clusters from the CML cells across the different responder groups. We found clusters 0, 5, 7 and 8 to have greatest scores for their respective patient groups (Figure 3-3A-C). Then we used label transfer from GSE233 (which contained a mix of both $CD34^{+}$ and $CD34^{+}CD38^{-}$ cells), alongside canonical stem cell markers to predict clusters which were most $CD34^{+}CD38^{-}$-like. These cells also resided in clusters 0, 5, 7 and 8, matching the BCR::$ABL1^{+}$ cluster predictions (Figure 3-3D/F). Lastly, we integrated GSE233 and EGA to see how closely associated predicted $CD34^{+}CD38^{-}$ clusters were between datasets (Figure 3-3E-F). We saw that the normal HSCs from EGA and the bystander HSCs from GSE233 occupied overlapping clusters, while predicted BCR::$ABL1^{+}$ $CD34^{+}CD38^{-}$ cells from EGA showed close proximity to the primitive LSCs from GSE233.

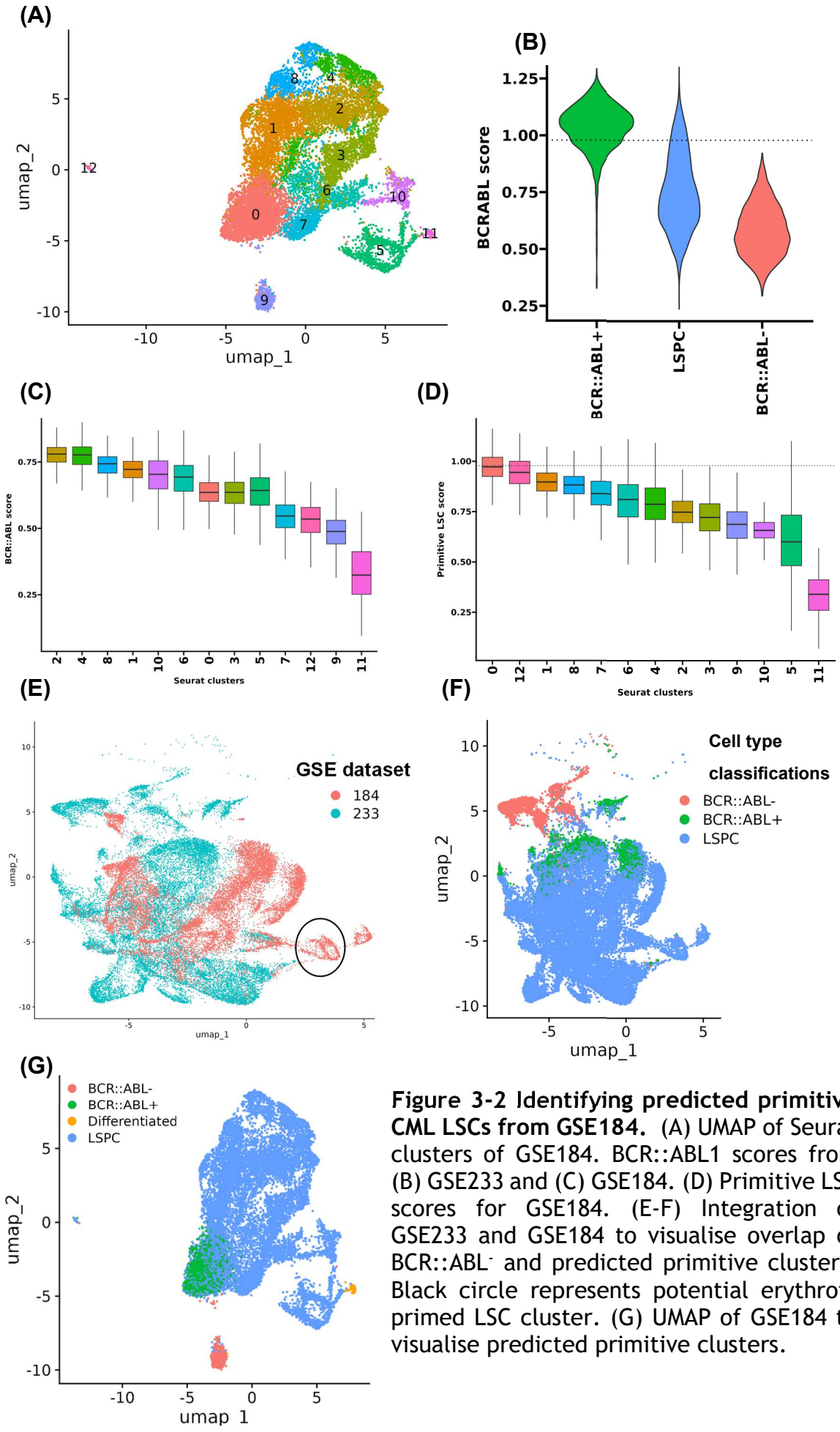


Figure 3-2 Identifying predicted primitive CML LSCs from GSE184. (A) UMAP of Seurat clusters of GSE184. BCR::ABL1 scores from (B) GSE233 and (C) GSE184. (D) Primitive LSC scores for GSE184. (E-F) Integration of GSE233 and GSE184 to visualise overlap of BCR::ABL$^-$ and predicted primitive clusters. Black circle represents potential erythroid primed LSC cluster. (G) UMAP of GSE184 to visualise predicted primitive clusters.

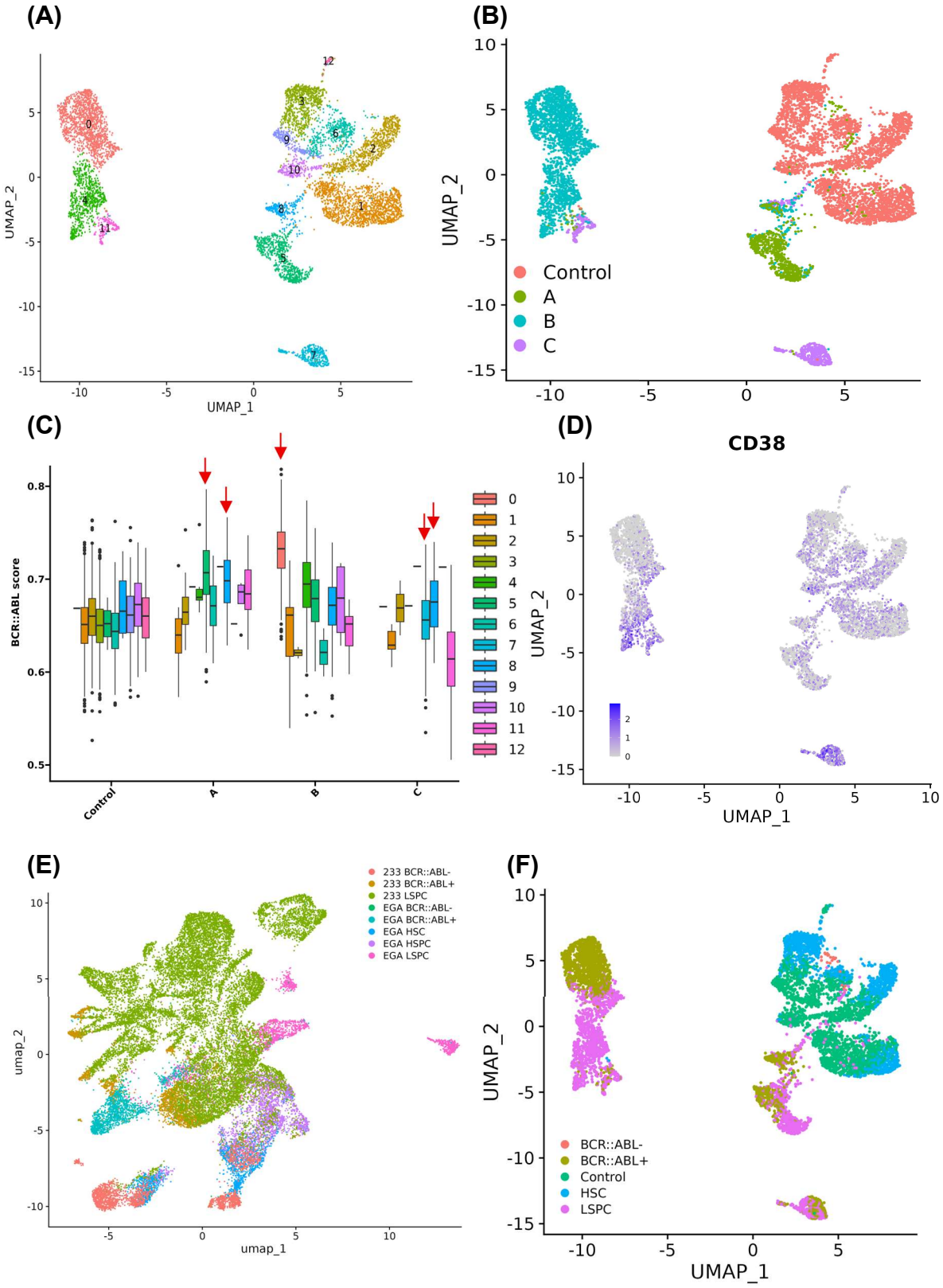


Figure 3-3 Identifying predicted LSPC populations from EGA. (A) Seurat clusters from EGA. (B) Responder groups from EGA. (C) BCR::ABL1 scores for EGA Seurat clusters. (D) Example of progenitors in predicted BCR::ABL^{+} clusters. (E) Integration of GSE233 and EGA to visualise cluster overlap. (F) UMAP of GSE184 to visualise predicted LSPC (CD34^{+}CD38^{-}) clusters.

3.3 Overlap of commonly differentially expressed genes provides predicted CML LSC targets and reveals metabolic changes in LSCs across phenotypes.

3.3.1 Bulk transcriptomics

To characterise the effect of CML transformation and TKI treatment at the $CD34^+$/$CD34^+CD38^-$ level, and to assess common signals between these phenotypes, we utilised 8 $CD34^+$, 3 $CD34^+CD38^-$ and 8 TKI treatment datasets (Table 3-1). To do so, we conducted gene set enrichment analysis (GSEA) (Figure 3-4), plotted top scoring genes and examined their commonality (Figure 3-5), then further detailed metabolic changes in $CD34^+CD38^-$ cells (Figure 3-6).

Firstly, from examining Figure 3-4A, we saw no pathway from the $CD34^+$ LSPC vs HSPC comparison set to be common across all datasets in the same enrichment direction, with heme metabolism and fatty acid metabolism gene sets showing the greatest concordance in the upregulated direction, followed by MYC and E2F targets alongside mTOR signalling. Most downregulated was IL-2, tumour necrosis factor-α (TNF-α) and interferon signalling alongside allograft rejection. Datasets derived from $CD34^+CD38^-$ LSC vs HSCs (Figure 3-4B) show a widespread upregulation of cell cycle pathways, alongside concordant changes in OXPHOS, glycolysis, fatty acid metabolism and xenobiotic metabolism (transporters and metabolism genes). There were no pathways for this comparison set that were downregulated across all datasets, however 2/3 of these datasets showed a downregulation of TNF-α, transforming growth factor beta (TGF-β) and inflammatory responses. Examining common pathways across the TKI treatment comparison set (Figure 3-4C), we observed that BCR::ABL1 inhibition resulted in a downregulation of all metabolic pathways (aside from heme metabolism which showed a mixed response) and cell cycle related genes, alongside an upregulation of genes involved in cytokine signalling and xenobiotic metabolism, in line with the known effects of TKI on CML cells.

Next, we looked at the top scoring features across these comparison sets to identify genes which could represent useful targets. A top scoring gene was required to be expressed in the greater value between 2/3 or 3 datasets, with a

padj < 0.05, and an absolute log 2 fold change (Log2FC) > 1 in at least one dataset, or all datasets for the CD34^{+}CD38^{-} comparison (Figure 3-5).

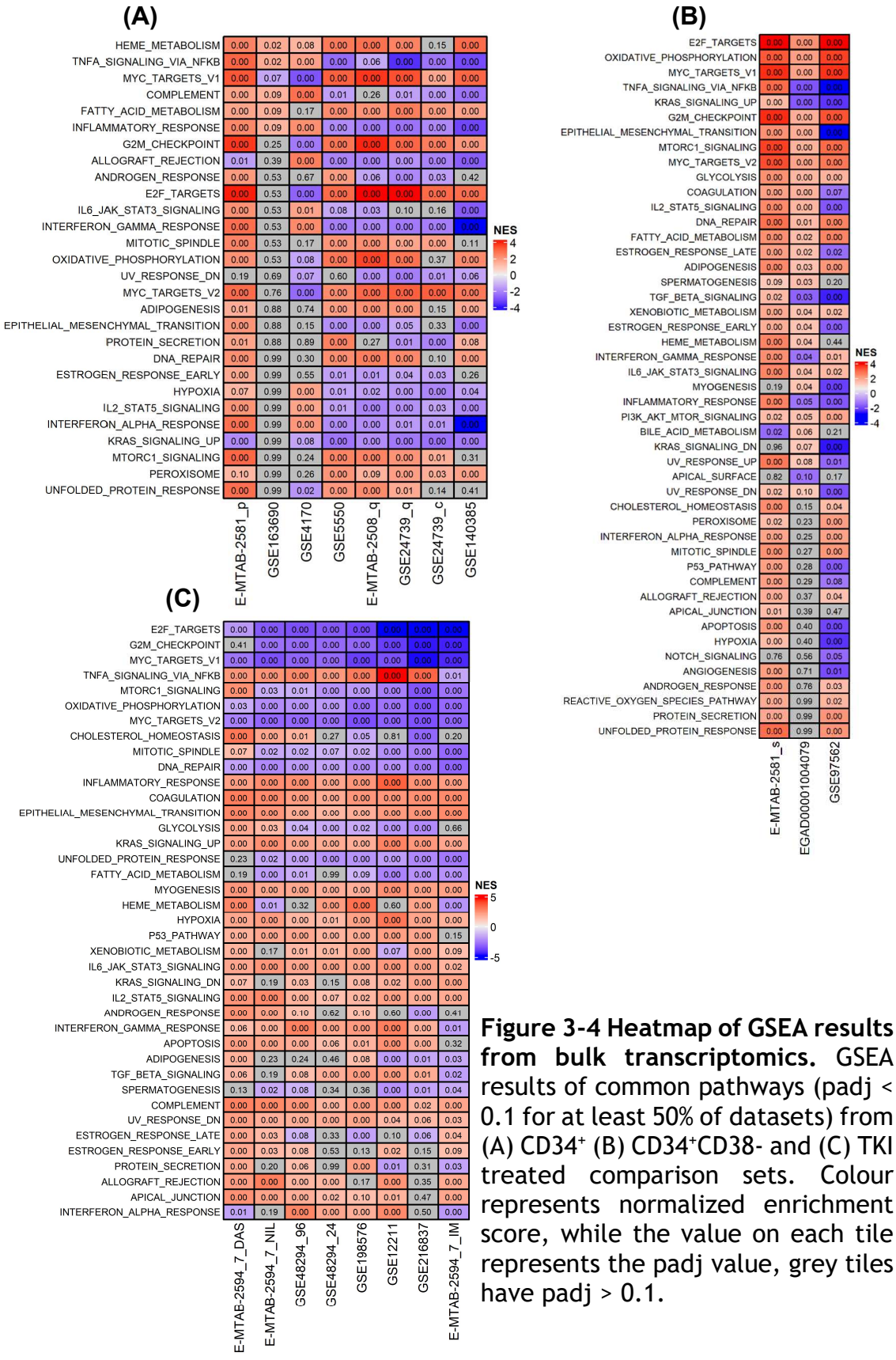


Figure 3-4 Heatmap of GSEA results from bulk transcriptomics. GSEA results of common pathways (padj < 0.1 for at least 50% of datasets) from (A) CD34^{+} (B) CD34^{+}CD38- and (C) TKI treated comparison sets. Colour represents normalized enrichment score, while the value on each tile represents the padj value, grey tiles have padj > 0.1.

We chose stringent thresholds to avoid false positives, and also employed p-value combination using the Fisher method for genes of interest. However, combining p-values lead to substantially less conservative results, therefore we did not threshold based upon this. First we examined the $CD34^{+}CD38^{-}$ comparison set, as if common genes could not be detected between this comparison set and the others, we could discount their continued analysis (Figure 3-5B). We found 21 top scoring genes for the $CD34^{+}CD38^{-}$ comparison set. Of these, 9 had previously been subject to extensive interest in CML (Growth Arrest-Specific Protein 2 (GAS2), PIEZO2, CD26/DPP4, calcium voltage-gated channel subunit alpha1 D (CACNA1D), IL1-RAP, SLC4A4, IGF1R, CDH2, PLAG1 and ECE1), while for the remaining 12 we present additional data to suggest importance, compared to previous understanding (relaxin family peptide receptor 1 (RXFP1), CCL5, Kynureninase (KYNU), FAM83A, S100A16, KCNT2, ARHGAP, SETBP1, CHRDL1, FRY, zinc finger protein 385D (ZNF385D)). We classified genes as previously implicated if they were computationally identified and explicitly discussed with more evidence than we provide here, and/or were subject to any *in-vitro*/*in-vivo* investigation. The topmost upregulated/downregulated genes of interest from this comparison set without extensive characterisation were RXFP1 and ZNF385D. In the context of CML, RXFP1 was previously identified in the publication associated with GSE97562, and on bioRxiv as of June 2025, where machine learning was conducted on GSE97562/GSE279135[173,174]. Additionally, we found RXFP1 to be upregulated in GSE138883 (another $CD34^{+}CD38^{-}$ dataset), which we discarded from this comparison set due to low numbers of statistically significant genes. RXFP1 is the cognate receptor for Relaxin, and is typically expressed by osteoclasts and osteoblasts, functioning to induce MSC differentiation into osteoblasts, with no indication of its effects upon HSCs/LSCs[175]. Functionally, RXFP1 is a GPCR, and following its activation a cyclic adenosine monophosphate (cAMP) signalling pathway is initiated, and can cause a cAMP feedback loop through PI3K recruitment to amplify its signalling[176]. ZNF385D has been identified as being downregulated in LSCs in other publications, but only noted in passing as a DE feature, and has received no *in-vitro*/*in-vivo* investigation with respect to CML. Its function has received computational speculation in the context of normal haematopoiesis, with ZNF385D being expressed in a range of haematopoietic cells, and being predicted to be involved in early erythroid development[177]. If this is the case, lack of expression could contribute to the observed lineage biases in CML.

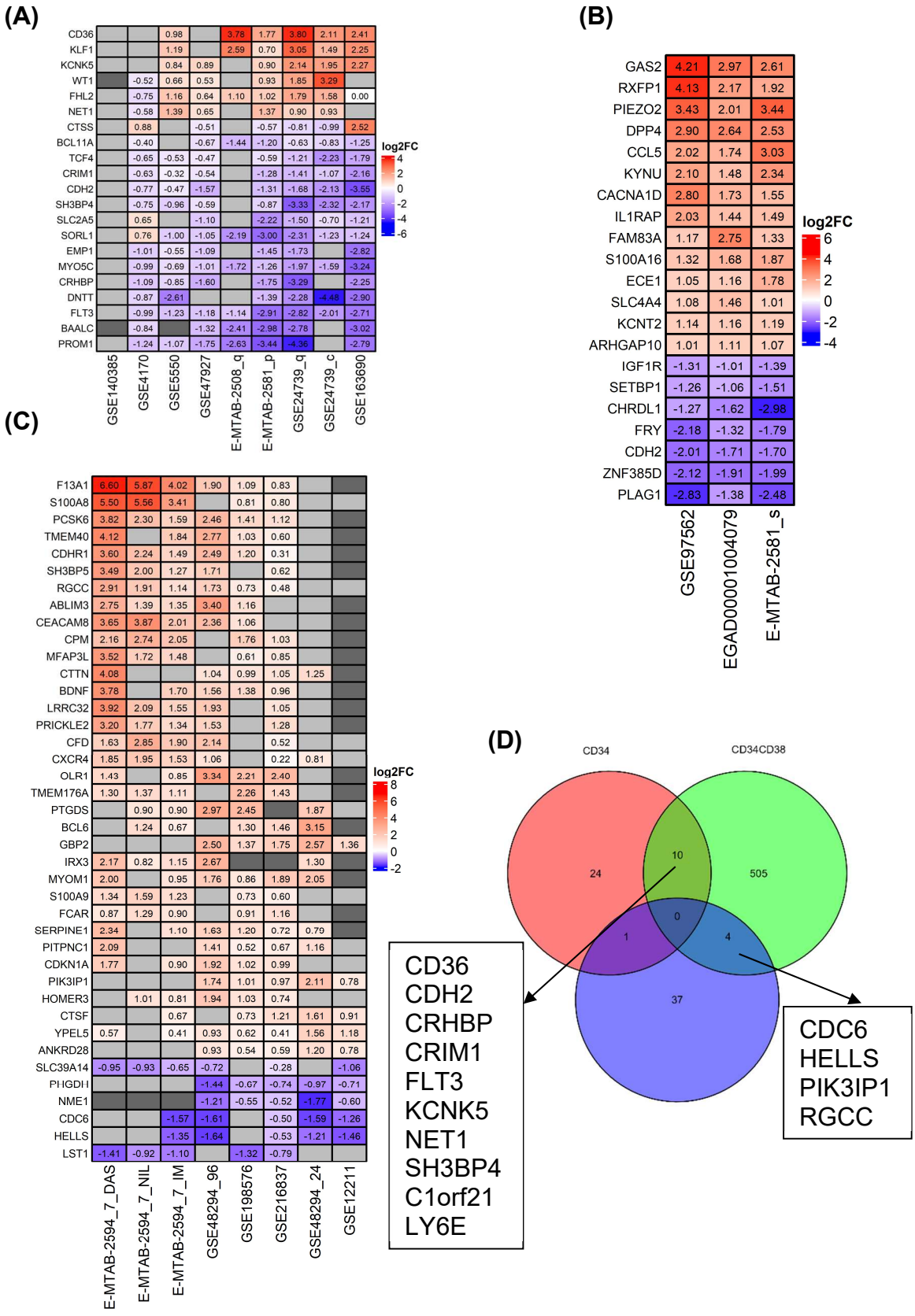
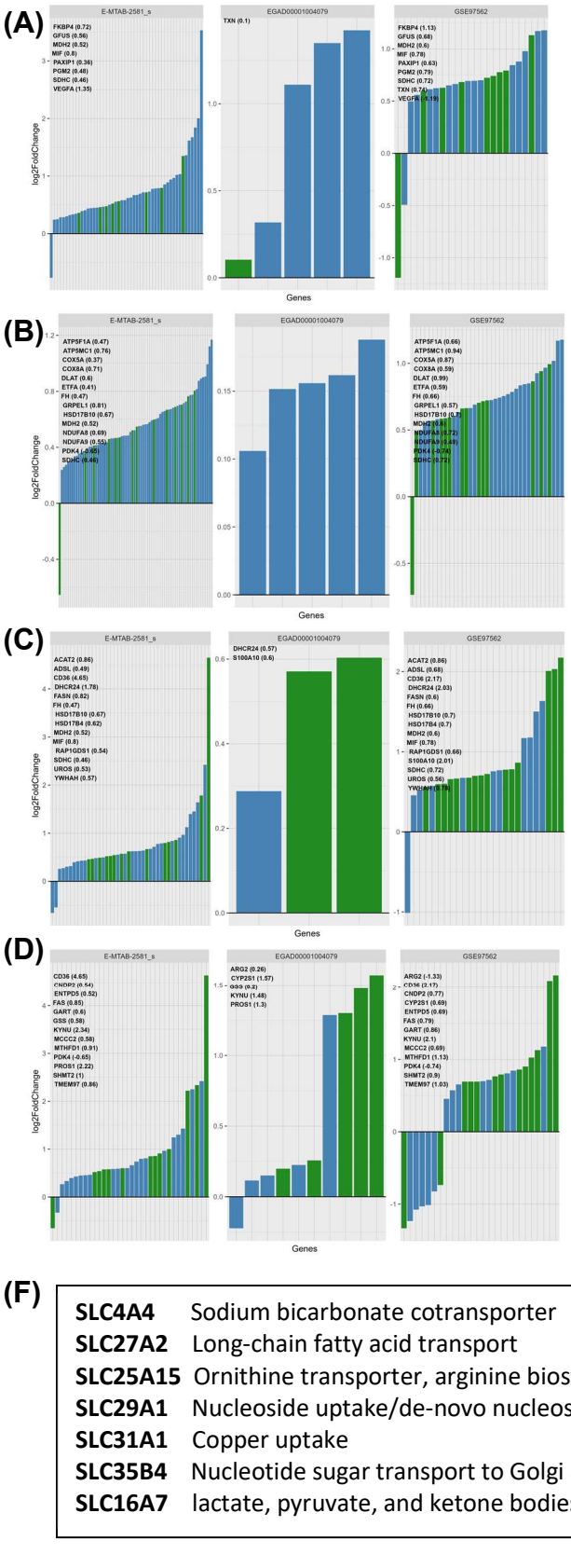


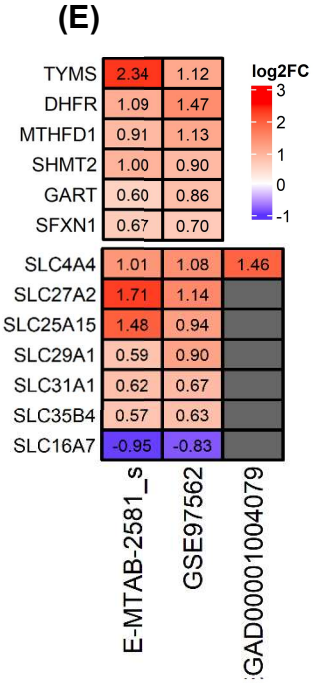
Figure 3-5 Heatmaps of top scoring genes from bulk transcriptomics. Commonly differentially expressed genes from (A) $CD34^{+}$ (B) $CD34^{+}CD38^{-}$ and (C) TKI treated CML cells. Colour and tile values represents Log2FCs, light grey represents no significance (padj >0.05) while dark grey represent a gene that was not detected. (D) Overlap of common significant genes present in 2/3 datasets in each comparison set.

To determine the usefulness of the $CD34^+$ and TKI treated datasets, we then examined whether there was any overlap between these comparison sets and the $CD34^+CD38^-$ comparison set. We found an overlap of 1 gene (CDH2) between $CD34^+$ and $CD34^+CD38^-$ comparison sets and 0 between the others, wherein we reduced stringency to see if any commonality at all could be observed (padj < 0.05 and absolute log2FC > 0.5 in 2/3 datasets) (Figure 3-5D). Of the 10 genes common to the less stringent intersection between $CD34^+CD38^-$ and $CD34^+$ comparison sets, 3 had already been subject to further analysis (CD36, FLT3 and CDH2), while 7 had not (chromosome 1 open reading frame 21 (C1orf21), lymphocyte antigen 6 family member E (LY6E), corticotropin releasing hormone binding protein (CRHBP), cysteine-rich motor neuron 1 (CRIM1), potassium two pore domain channel subfamily K member 5 (KCNK5), neuroepithelial cell transforming gene 1 (NET1), and SH3 Domain Binding Protein 4 (SH3BP4)). We saw a downregulation of CRHBP, which has been identified computationally as a downregulated gene in CML here (and elsewhere when looking at individual datasets), and was recently identified as a novel AML suppressor[178]. In normal HSPCs, CRHBP functions by regulating corticotropin releasing hormone bioavailability, however, upon supplementation its anti-AML effects were found to occur independently from this mechanism, which remains undefined[178]. CRIM1 may be involved in drug resistance in other myeloid leukaemias, KCNK5 is a potassium channel which has received passing computational interest, C1orf21 is involved in proliferation , LY6E is involved in T-cell regulation, and NET1 has been shown to be an ABL1 effector[179–182]. SH3BP4 is a known haematopoietic regulator, whose expression we found to be reduced[183]. $CD38^-$ cells make up approximately 10% of the $CD34^+$ population, and given almost no genes were found to be commonly DE between these comparison sets, gene signatures observed in $CD34^+$ cells in general do not appear to be widely applicable to $CD34^+CD38^-$ cells. The $CD34^+$ compartment is heterogeneous, with the expected cell frequencies and transcriptional programmes between normal and CML progenitors being different, therefore genes detected were most likely to reflect the most abundant cell types, limiting their usefulness to a general CML progenitor signal. Therefore, we did not continue to analyse $CD34^+$ datasets.

As we observed an upregulation of metabolic pathways in the $CD34^+CD38^-$ comparison set from the GSEA, with two of the top scoring genes directly related to energy metabolism (KYNU and SLC4A4), we further investigated changes in

individual genes underlying this enrichment (Figure 3-6). To do so, we plotted genes belonging to the hallmark OXPHOS, glycolysis, fatty acid and xenobiotic metabolism pathways (padj<0.05 common to 2/3 datasets). Upregulated and physiologically meaningful in the glycolytic pathway (Figure 3-6A) were FKBP Prolyl Isomerase 4 (FKBP4) and malate dehydrogenase 2 (MDH2). FKBP4 is a co-chaperone which has been shown in other cancers to induce aberrant glycolysis via upregulation of hexokinase-2, while MDH2 converts malate to oxaloacetate in the TCA cycle[184,185]. Hexokinase and MDH2 are critical to glycolysis and the TCA respectively. Additionally in the TCA pathway, but categorized as a xenobiotic metabolism gene was fumarate hydratase (FH), a critically important enzyme which converts fumarate to malate[186]. In the OXPHOS pathway (Figure 3-6B), aside from general upregulation of ETC genes such as ATP synthase, NADH ubiquinones and SDHC (part of complex II), we also saw an upregulation of GrpE Like 1 Mitochondrial (GRPEL1) and Pyruvate Dehydrogenase Kinase 4 (PDK4). GRPEL1 functions to mediate and stabilize mitochondrial protein import/folding, while the downregulated PDK4 enzyme is responsible for regulating pyruvate dehydrogenase activity, which converts pyruvate to acetyl-CoA[187]. Another strongly upregulated gene associated with mitochondrial metabolism that was not in the OXPHOS gene set was cyclophilin D (PPIF). PPIF is a peptidyl-prolyl cis-trans isomerase which can be directly activated by P53, and has numerous functions. For example, directly regulating ATP synthase, and formation of the mitochondrial permeability transition pore. While its functions are far from fully characterised, PPIF is gaining interest as a marker for poor prognosis in other cancers[188]. While the observed upregulated fatty acid metabolism genes showed redundancy with OXPHOS and glycolysis genes, we report changes in 7-beta-Hydroxysteroid dehydrogenase type 10/4 (HSD17B10/B4), acetyl-CoA acetyltransferase 2 (ACAT2) and 24-dehydrocholesterol reductase (DHCR24) (Figure 3-6C). HSD17B10 is partially responsible for isoleucine and short chain fatty acid metabolism and mitochondrial tRNA maturation, while HSD17B4 is responsible for very-long chain fatty acid oxidation[189–191]. Neither gene has been investigated in the context of CML. ACAT2 catalyses the conversion of acetyl-CoA to acetoacetyl-CoA, which can be subsequently used for ketogenesis or cholesterol biosynthesis, suggesting ACAT2 upregulation could provide additional cholesterol for either anabolic or catabolic processes[192]. Also DE in the cholesterol biosynthesis pathway was DHCR24 and Transmembrane Protein 97 (TMEM97).





(F)

SLC4A4	Sodium bicarbonate cotransporter
SLC27A2	Long-chain fatty acid transport
SLC25A15	Ornithine transporter, arginine biosynthesis
SLC29A1	Nucleoside uptake/de-novo nucleoside synth
SLC31A1	Copper uptake
SLC35B4	Nucleotide sugar transport to Golgi apparatus
SLC16A7	lactate, pyruvate, and ketone bodies transport

Figure 3-6 Characterisation of metabolic changes from the bulk CD34^{+}CD38^{-} comparison set. DE metabolic genes (padj < 0.05) in 2/3 comparisons (green bars) associated with (A) glycolysis (B) OXPHOS (C) fatty acid and (D) xenobiotic metabolism gene sets. Each bar represents a gene, while green bars represent genes in atleast 2 datasets. Values represent log2FC values. (E) 1C metabolism genes (top) and SLC genes (bottom). (F) SLC gene descriptions.

TMEM97 is an oxidoreductase that catalyses reactions in multiple cholesterol metabolism pathways, with its activation resulting in an increase in cholesterol and neutral lipid uptake[193,194]. Examining changes in the xenobiotic metabolism gene set (Figure 3-6D) showed an upregulation of 1C genes (Figure 3-6E, top) and some AA metabolism genes. Changes in 1C metabolism have been previously reported by our laboratory, with its perturbation being highly effective *in-vitro*/*in-vivo*, however, this serves as additional evidence to support its importance. We also found changes in other AA metabolism genes such as arginase-2 (ARG2) which converts arginine to ornithine (contrasting but low fold-changes), and a large upregulation KYNU. KYNU catalyses reactions in the tryptophan and kynurenine metabolism pathway, which contributes to NAD^+ and alanine production. It is also thought to contribute to an inflammatory phenotype, and is possibly a useful prognostic marker in CML[195,196]. Lastly, as other transporters such as CD36 were upregulated in 2/3 of the datasets, we examined the expression of SLC family transporters common to atleast 2/3 of the datasets in this comparison set. We show an upregulation of nucleoside, ornithine/arginine and fatty acid uptake transporters, and a downregulation of SLC16A7, which is a high affinity pyruvate transporter (Figure 3-6E, bottom)[197].

3.3.2 scRNA-seq of $CD34^+CD38^-$ LSCs vs bystanders

While bulk RNA-seq of $CD34^+CD38^-$ LSCs and HSCs has been highly informative in the past for hypothesis generation, this comparison has issues. For example, the $CD34^+CD38^-$ compartment itself is heterogeneous, and when taking this population for bulk sequencing a variable number of bystander HSCs are also being encapsulated, increasing the signal to noise ratio (Figure 6-5D). Bystander HSCs and normal HSCs have different transcriptional profiles, meaning that comparing normal HSCs to LSCs doesn't capture the true biological differences within the leukaemic niche, and so targets of interest may be incorrectly identified. Therefore, to better characterise the differences between $CD34^+CD38^-$ LSCs and bystander HSCs, we took 3 scRNA-seq datasets (GSE233, GSE184 and GSE312) containing this comparison, conducted differential expression and GSEA analysis, intersected top scoring genes, and characterised metabolic changes (Figure 3-7 and Figure 3-8).

From the GSEA results (Figure 1.7C) we show that in all datasets there was an upregulation of E2F and MYC targets, OXPHOS, other cell cycle pathways, and a downregulation of TNF-α signalling genes. Following this, we intersected the top scoring genes (padj < 0.05 and absolute log2FC >1) for this comparison set (Figure 3-7A), finding 23 genes of interest. From these, 6 genes (Figure 3-7D) were overlapping with the bulk $CD34^{+}CD38^{-}$ comparison set (GAS2, CACNA1D, RXFP1, IL1-RAP, FRY and ZNF385D), 7 had already received some level of investigation (RAB31, PTPRD, CD33, CDCA7, galectin-1 (LGALS1), PCNA and STAP1), while the remaining 9 had little to no relevant literature (LYZ, C1orf54, ORC6, AC002454.1, MAD2L1, DNAJC12, SORL1 and PCDH9) (Figure 6-5B). Also displayed is the mean expression of a selection of top scoring genes for each pseudobulked patient to provide a better visualisation of variability (Figure 3-7B).

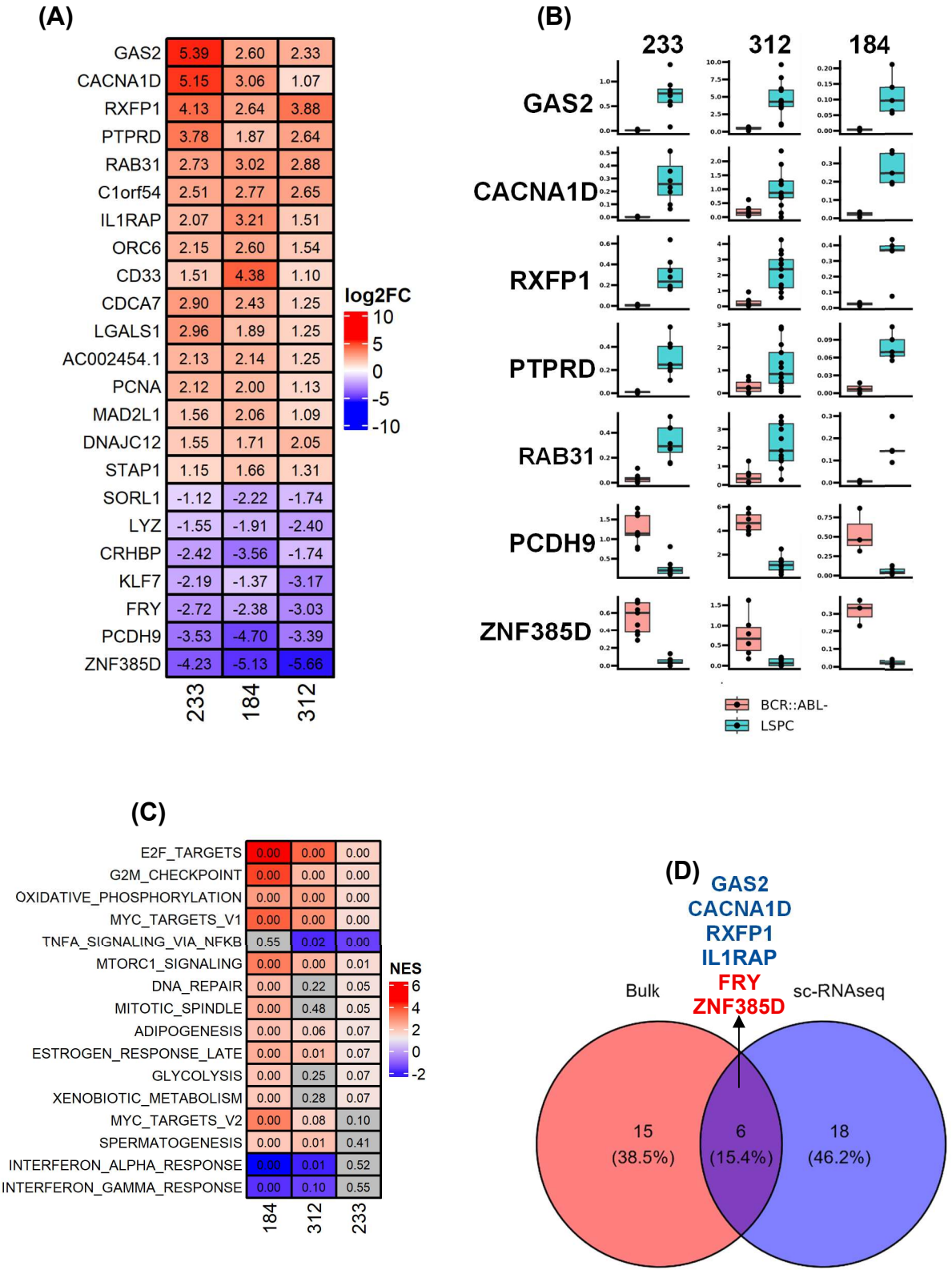
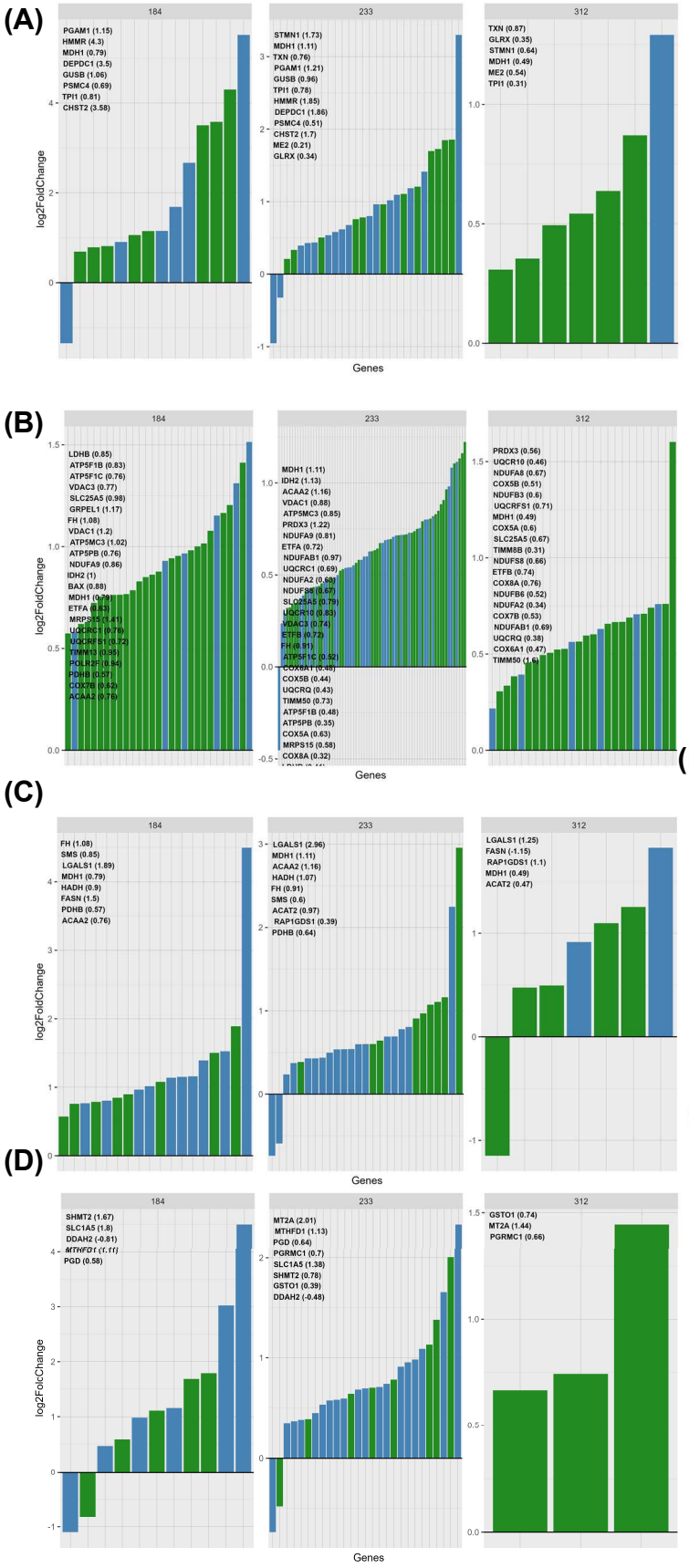


Figure 3-7 Top scoring genes and pathways from the CD34⁺CD38⁻ BCR::ABL⁺ LSCs vs bystander HSCs comparison set. (A) Top scoring genes. (B) Mean expression across patients of a selection of top scoring genes. (C) GSEA for this comparison set. (D) Overlap of top scoring genes from CD34⁺CD38⁻ BCR::ABL⁺ LSCs vs bystanders and bulk CD34⁺CD38⁻ comparison sets.

Given that we observed positive enrichment scores for metabolic pathways, we next aimed to characterise genes underlying this enrichment (Figure 3-8). Examining changes in glycolytic genes (Figure 3-8A) showed an upregulation of MDH1 rather than MDH2, alongside upregulations of thioredoxin (TXN), phosphoglycerate mutase-1 (PGAM1) and DEP domain containing 1 protein (DEPDC1). TXN functions to regulate redox state, is involved in cysteine nitrosylation, and has been implicated in AML and ALL[198]. PGAM1 is a core glycolytic enzyme, while DEPDC1 indirectly regulates glycolysis through AKT/mTOR/HIF-1α[199,200]. Out of OXPHOS genes (Figure 3-8B), we again observed an upregulation of ATP synthase components, NADH ubiquinones, cytochrome genes, PPIF (log2FC of 1.9 and 3.2 from GSE233 and GSE184 respectively), mitochondrial translation genes (MRPLs) (Figure 3-8E), translocase of the inner membrane genes, isocitrate dehydrogenase-1 (IDH1), HSD17B10, peroxiredoxin-3 (PRDX3), voltage dependent anion channel (VDAC) components, and BCL2 associated-X (BAX). This again indicates increased mitochondrial activity, for example IDH1 is a key TCA enzyme and PRDX3 regulates H_2O_2 levels in the mitochondria[201]. BAX is most well known for being a Bcl-2 family protein, acting to promote apoptosis, however, it also is involved in mitochondrial fusion[202]. VDAC genes comprise a voltage gated ion channel located in the outer mitochondrial membrane that plays a key role in mitochondrial mediated apoptosis, metabolite and calcium import/export, and facilitates the exit of oxidized mt-DNA, which has shown to cause an inflammatory phenotype in macrophages[203,204]. It has been speculated that VDAC may work in conjunction with BAX to regulate mitochondrial morphology and fusion through DRP1 (a key protein involved in mitochondrial dynamics), however the mechanisms underlying this interaction are yet to be defined[188]. Additionally, VDAC and PPIF may interact when working to open the mitochondrial permeability transition pore, further indicating potential importance of mitochondrial dynamics. In the fatty acid metabolism pathway (Figure 3-8C) we observed changes in LGALS1, and again detected the upregulation of ACAT2 and FH. LGALS1 has been identified as a gene of interest computationally, but has received only passing interest *in-vitro* and while categorised as a fatty acid metabolism gene, its main function appears to related to modulation of T-cell activity, which is highly relevant to the CML research landscape[205–207]. Lastly, examining xenobiotic metabolism genes, we detected similar changes in 1C metabolism, further supporting evidence of importance,

alongside MRPL genes and SLC genes (Figure 3-8D-F). Notably, we did not detect any similar SLC family genes compared to the bulk CD34^{+}CD38^{-} comparison set, suggesting these changes may be less reliable.



(E)

	184	233	312
TYMS	3.96	2.66	
DHFR	2.47	1.92	
SHMT2	1.67	0.78	
MTHFD1	1.11	1.13	
MRPL3	1.20	0.00	0.94
MRPS15	1.41	0.58	
MRPL12	1.38	0.89	
MRPL13	0.93	0.90	0.56
MRPL27	1.05	0.81	
MRPL14	0.85	0.63	
MRPS14	0.64	0.81	
SLC29A1	2.28	1.88	
SLC1A5	1.80	1.38	
SLC16A1	1.90	0.74	
SLC25A5	0.98	0.79	0.67
SLC38A5	1.09	0.90	
SLC12A2		0.57	0.92
SLC25A4	-1.32	-1.04	

(F)

SLC29A1	Nucleoside uptake
SLC1A5	Neutral AA transporter
SLC16A1	Carboxylate Transporter
SLC25A5	ADP:ATP antiporter
SLC38A5	Neutral AA transporter
SLC12A2	Cation transporter
SLC40A1	Iron transporter

Figure 3-8 Differential expression of metabolic genes between CD34^{+}CD38^{-} BCR::ABL^{+} LSCs and bystanders. DE metabolic genes (padj < 0.05) in 2/3 *comparisons* (green bars) associated with (A) glycolysis (B) OXPHOS (C) fatty acid and (D) xenobiotic metabolism gene sets. (E) 1C metabolism genes (top), MRPL genes (middle) and SLC genes (bottom). Colours and values represent log2FC values. (F) SLC gene descriptions.

3.3.3 scRNA-seq of predicted primitive LSCs vs bystander HSCs

The $CD34^+CD38^-$ compartment is highly enriched for functional HSCs, but still contains a multitude of cell types. From this population, the addition of CD90 allows ST/LT-HSCs to be demarcated from MPPs. The exact phenotype of the LSC within the $CD34^+CD38^-$ compartment responsible for disease propagation and relapse is undefined, however $CD26^+CD90^+$ LSCs are a candidate population. For this reason, we aimed to further investigate the features of this population ($CD26^+CD90^+$) using the corresponding cluster from GSE184 (Figure 3-20), alongside the already annotated GSE233. Additionally, we also used the bulk transcriptomic dataset GSE47927, which among other phenotypes, compared $Lin^-CD34^+CD38^-CD90^+$ cells from CML patients to normal cells. While this phenotype does not directly correspond to the predicted primitive phenotype, and does capture bystanders, the $CD26^+$ clusters identified in GSE184 and GSE233 are relatively enriched for CD90. Therefore, we thought GSE47927 would be most suited for this comparison set. In order to investigate these predicted primitive cells, we identified top scoring genes, conducted GSEA, and investigated metabolic changes (Figure 3-9). Additionally, we plotted all top scoring genes and metabolic genes of interest from $CD34^+CD38^-$ or more primitive comparison sets against the same genes from the TKI treated comparison set, to see if TKI treatment altered their expression (Figure 3-10).

From the GSEA results (Figure 3-9C), we observed numerous statistically significant pathways in GSE184 and GSE47927, but none from GSE233, despite having the greatest number of DE genes from the comparison set (Figure 6-2C and Figure 6-5C). Looking more closely at the fold-change distribution of GSE233 showed that the majority of genes had log2FCs of close to 0 (11692 genes with absolute log2FC < 0.01), reducing the capacity to detect DE pathways. From GSE233 and GSE47927, we saw an upregulation of E2F and cell cycle related genes, MYC targets, glycolysis, OXPHOS, fatty acid metabolism and mTOR signalling. We also observed a downregulation of TNF-α and TGF-β pathways.

Intersecting the top scoring genes from GSE233 and GSE184 (padj < 0.05 and absolute log2FC >1) revealed 68 candidate genes, and from intersecting the entire comparison set, we observed 13 genes, which represented a subset of these 68 genes (Figure 3-9A/B). Of the 68 genes, 16 overlapped with the top scoring genes

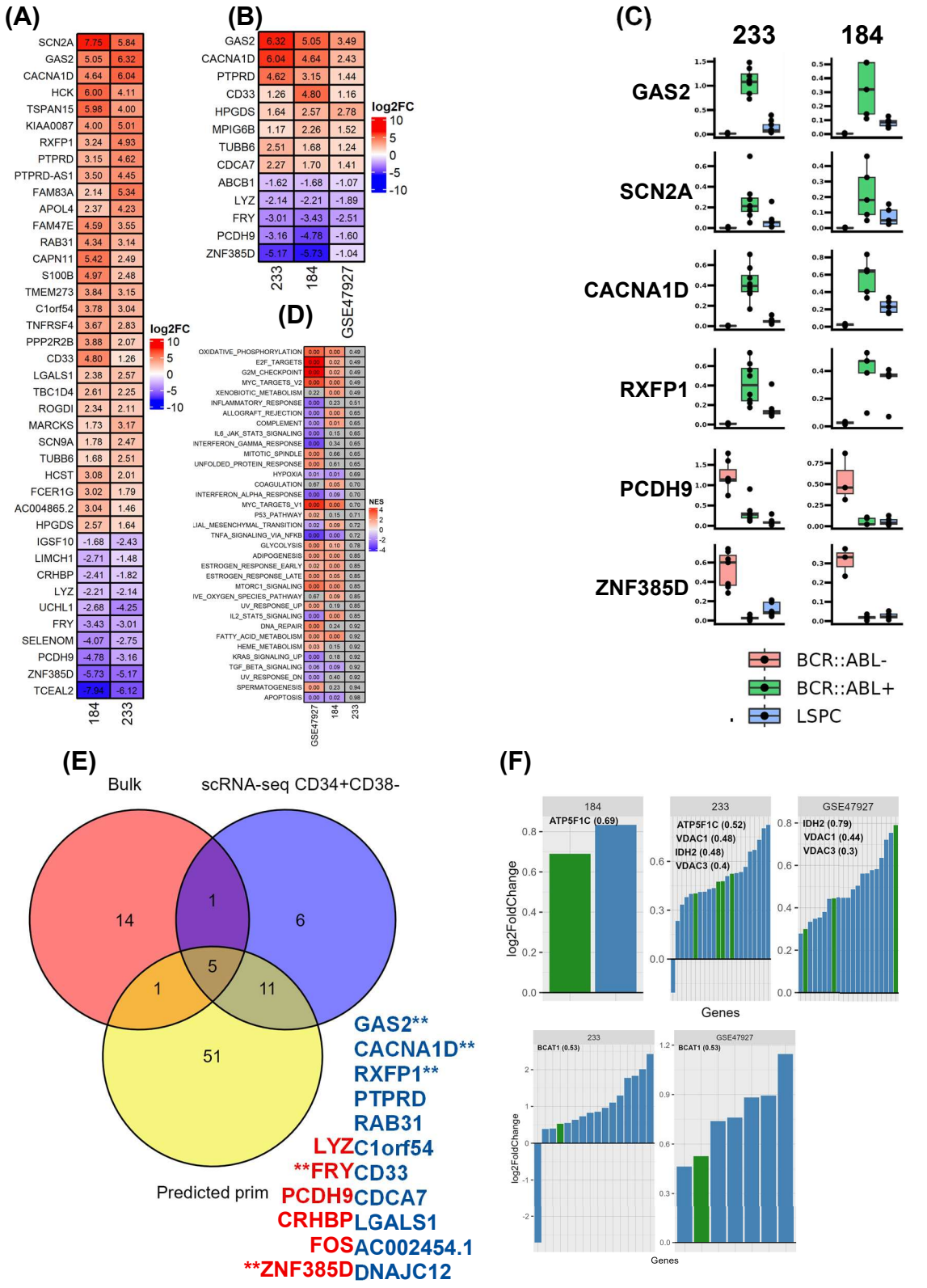


Figure 3-9 Top scoring genes and pathways from the predicted primitive BCR::ABL⁺ LSCs vs bystander HSC comparison set (A) 40 of the top scoring genes. (B) All top scoring genes when including GSE47927. (C) Mean expression of a selection of top scoring genes from patient pseudobulks. (D) GSEA results for the predicted primitive comparison set. (E) Overlap of top scoring genes across comparison sets, with genes common to all comparison sets denoted by **, and coloured based on fold change direction. (F) Changes in OXPHOS (top) and xenobiotic metabolism (bottom) pathways.

from the previous comparison set, while 5 of these 16 genes additionally

overlapped with the bulk $CD34^+CD38^-$ comparison set (RXFP1, GAS2, CACNA1D, FRY and ZNF385D (Figure 3-9A/D). Of these 5 genes, the remaining upregulated genes that have not been discussed here are GAS2 and CACNA1D. GAS2 has received *in-silico* recognition and some in-*vitro* investigation, and has been detected in the PB at diagnosis and during TKI therapy, however the effects of its inhibition have not been measured in $CD34^+CD38^-$ patient samples, nor in PDX models[89,208,209]. CACNA1D was previously identified through a drug screen and was well characterised by the late Dr Khalaf *in-vitro*/*in-vivo*, however no steps toward a translational outcome have been taken, despite promising findings *in-vitro* and in PDX models[210].

When examining metabolic changes for this comparison set, despite positive GSEA results, we detected proportionally fewer changes compared to most other comparison sets (Figure 6-5). This was somewhat expected as the $CD34^+CD38^-CD90^+CD26^+$ LSC has been shown to be highly quiescent[89]. We observed an upregulation (padj < 0.05 in atleast 2/3 datasets) of VDAC components, PPIF, IDH2, ATP5F1C, LGALS1 and BCAT1. Changes in BCAT1 expression were not consistently detected at any other phenotype, indicating that BCAA metabolism may be highly specific to primitive LSCs. Lastly, we aimed to visualize genes which were differentially expressed across comparison sets (padj < 0.05 and absolute log2FC > 1 in atleast 2/3 datasets per comparisons set). To do so, we took the $CD34^+CD38^-$ bulk and/or $CD34^+CD38^-$ scRNA-seq and predicted primitive comparison sets, and plotted top scoring genes, alongside metabolic genes of interest against the same genes from the TKI treated comparison set (Figure 3-10). From this, we see the majority of top scoring genes remain unchanged by TKI treatment, however, TKI treatment appeared to substantially reduce GAS2 expression, which is somewhat contrasting with corresponding data from human samples acquired by Warfvinge *et al*[89]. Based upon this, we would predict genes displayed in Figure 3-10 to represent worthwhile targets, particularly the 5 genes which were significant across all LSPC datasets and comparison sets, and the other genes which were significant in all datasets from the scRNA-seq and predicted primitive comparison sets. While changes in metabolic genes were much more modest at the predicted primitive level, we would predict that investigating the interplay between VDACs and PPIF (and BAX) to be worthwhile, as their signalling

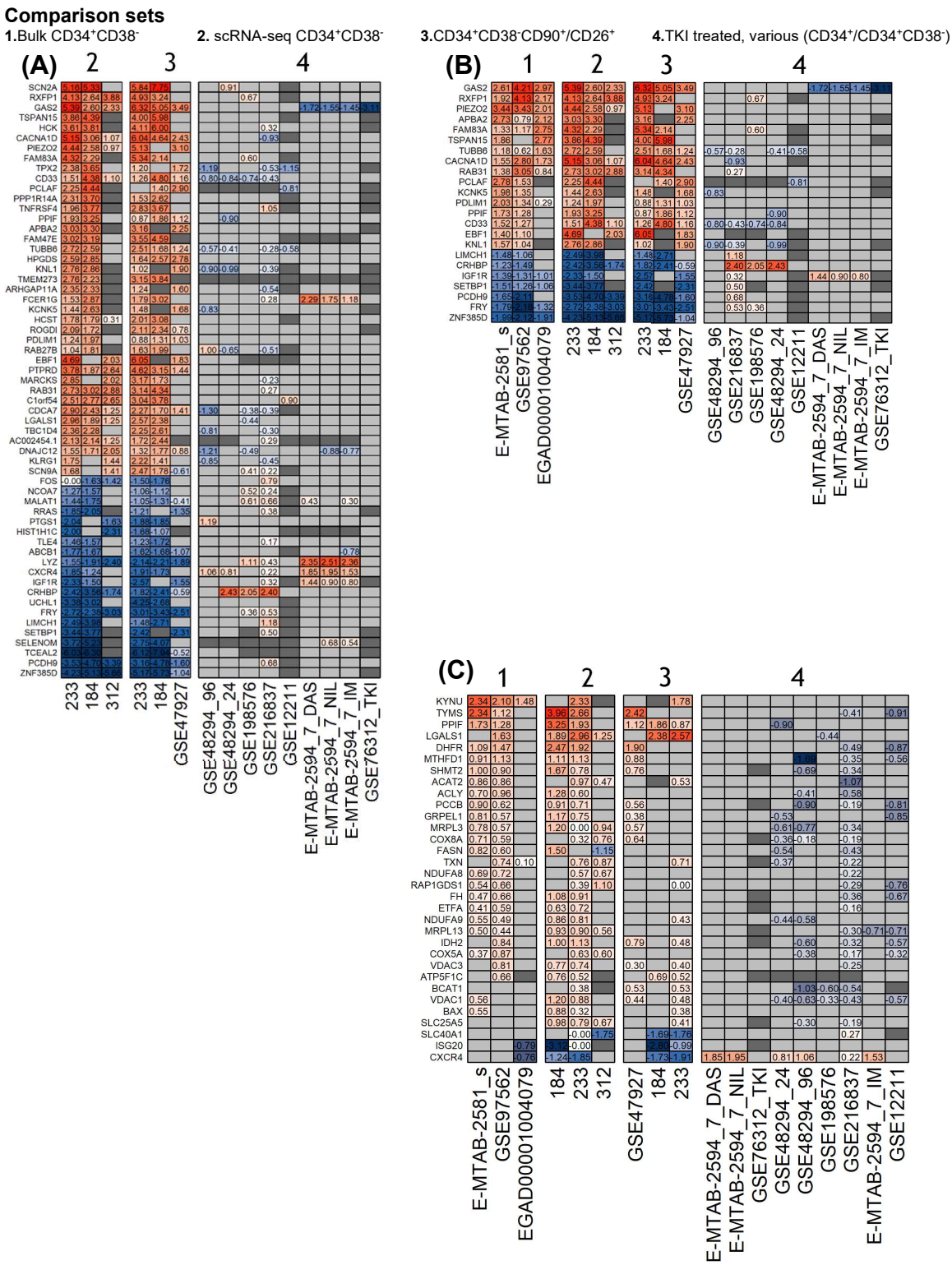


Figure 3-10 Expression of top scoring genes, close to top scoring genes, and metabolic genes of interest for stem comparison sets. Genes with padj <0.05 and absoloute log2FC > 1 expressed in (A) at least 2/3 datasets from the scRNA-seq CD34^{+}CD38^{-} comparison set and the predicted primitve comparison set, alongside these genes in TKI treated datasets. (B) Genes present in atleast 2/3 datasets for each LSPC comparison set, alongside these genes in TKI treated datasets. (C) Metabolic genes of interest across LSPC comparison sets. Comparison set key at top of the figure, with tile colour and values representing log2FC values.

is intimately linked with Ca^{2+} signalling/mitochondrial metabolism, both of which have already proven to be effective targets. Inhibition of BCAA metabolism, IDH's metabolism and LGALS1 signalling may also be worthwhile to investigate. For less primitive cells, we detected changes in 1C metabolism across bulk and scRNA-seq comparison sets, alongside other previously discussed genes such as ACAT2, DHCR24, HSD17B10, SLC25A5, ACLY, MRPL genes, and possibly KYNU.

HSC/LSC genetic and transcriptional programmes are somewhat flexible as previously discussed, across this continuum it would be reasonable to assume that stem LSCs from both the BM and the PB should share gene expression representative of a common state of stemness, regardless of the tissue of origin. From these comparison sets, we predicted genes that may represent part of this common state of leukaemic stemness, and make suggestions regarding which metabolic genes could represent worthwhile targets.

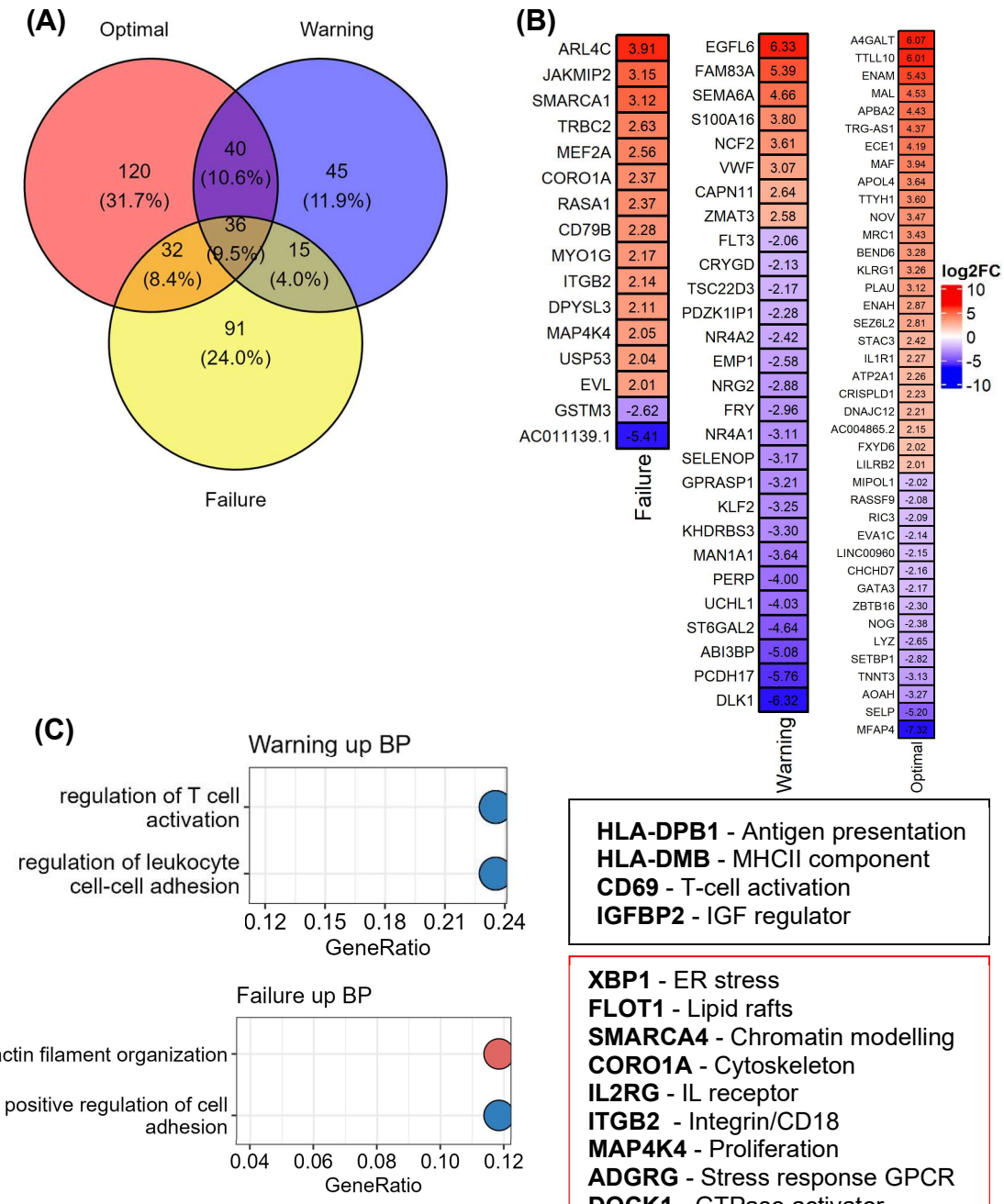
3.3.4 Transcriptomics poorly defines the differences between responder groups at predicted $CD34^{+}CD38^{-}$ and primitive level

Recently, there have been numerous publications using gene signatures to predict treatment responses to CML. Therefore, we aimed to determine whether we could identify relevant gene signatures at the $CD34^{+}CD38^{-}$ level using EGA, GSE233 and GSE312, and at the predicted primitive level using GSE233. The response groups associated with GSE233 and EGA were classified similarly to the 2020 ELN criteria: A (Optimal), B (did not achieve CCyR until 18 months of IM treatment but did respond optimally to second/third generation TKI), Warning, or C (Failure). Patients from GSE312 were categorised as good or poor responders (Table **1-2**).

First, we made comparisons between predicted BCR::$ABL1^{+}$ $CD34^{+}CD38^{-}$ cells from Warning/ group B vs group A, group C vs group A, and group C vs Warning/group B for GSE233 and EGA, and between poor and good responders from GSE312. Very few changes in gene expression were detected between Warning/group B and group A individuals, or between good and poor responders for any of the comparisons, with group C being most different to both groups A and Warning/B, based on number of DE genes (Figure 6-7). Intersecting DE genes between datasets, we saw no overlap between EGA and GSE233 responder group

comparisons, and 2 genes in total overlapping between GSE233/GSE312 and EGA/GSE312, being S100A10 and HNRNPL.

As we could not detect consistent changes at the $CD34^{+}CD38^{-}$ level, we made the same ABC comparisons using predicted primitive cells from GSE233, wherein group B refers to individuals classified in the Warning risk group in this instance. Again, we couldn't detect any major changes between groups A and B, and <15 DE genes (Figure 6-8) between groups B and C, and A and C (mean expression of top genes shown by Figure 6-9). Instead we took a different approach, considering that different responder groups have different BM environmental and immunological landscapes, we decided to compare the different responder groups to their respective bystander HSCs, then examine genes that were unique to each of these comparisons (Figure 3-11). From this, we identified 120, 45 and 91 genes unique to groups A, B and C respectively. In an attempt to characterise the unique genes for each comparison, we displayed top scoring genes for each comparison (padj < 0.05 and absolute log2FC > 2), and conducted over representation analysis (ORA) on statistically significant genes, with only groups B and C yielding significant results (Figure 3-11B/C). From the ORA results from group B LSCs, there was an enrichment of immunological activators, while group C LSCs showed upregulation of adhesion and cytoskeletal components. Perhaps when additional scRNA-seq datasets within the same phenotype are published, and with greater numbers of replicates, more consistent signals could be identified.



HLA-DPB1 - Antigen presentation
HLA-DMB - MHCII component
CD69 - T-cell activation
IGFBP2 - IGF regulator

XBP1 - ER stress
FLOT1 - Lipid rafts
SMARCA4 - Chromatin modelling
CORO1A - Cytoskeleton
IL2RG - IL receptor
ITGB2 - Integrin/CD18
MAP4K4 - Proliferation
ADGRG - Stress response GPCR
DOCK1 - GTPase activator

TMSB10 - PI3K/Akt signalling
STMN1 - Microtubules
EVL - Actin remodelling
ABI2 - Motility, adhesion
DPYSL3 - Actin organisation
RASA1 - RAS protein
MYO1G - Adhesion/immunity
IQGAP1 - Scaffold

Figure 3-11 Differences between treatment response groups from primitive GSE233 cells. (A) Overlap between gene expression when comparing responder groups to their corresponding bystanders. (B) Top scoring genes unique to these comparisons. (C) ORA for the response groups for warning/B (black) and failure/C groups (red), alongside annotations for genes.

3.4 Discussion

3.4.1 Discovery of genes common within the CD34$^+$CD38$^-$ compartment, and TKI treatment.

We identified and processed 20 transcriptomic datasets within the CD34$^+$, CD34$^+$CD38$^-$, or more stem compartments, aiming to compare CML LSCs to HSCs or bystanders, and to TKI treated LSCs. In doing so, we show that CML LSCs transcriptionally upregulate metabolic genes compared to their HSPC or bystander counterparts, even at a predicted primitive phenotype. Additionally, through intersecting DE genes across comparison sets, we provide evidence indicating the importance of numerous genes that could represent therapeutic targets, many of which were previously unreported (section 3.4.3). While this research did not validate the expression levels of the top scoring genes, at minimum it provides a standardised analysis and report, consolidating findings from the past 20 years of CML research.

There were numerous technical limitations that made this analysis challenging. Firstly, for GSE312 most patients had very few BCR::ABL1$^+$ cells, meaning that psuedobulking risked capturing an incomplete transcriptome for that patient. For example, does a pseudobulk not express a certain gene, or was it not captured in that quantity of cells? The alternative of down sampling to reduce patient bias was similarly unappealing given this would reduce the capacity of the analysis to detect DE genes, while retaining only individuals with high numbers of cells would drastically reduce the number of biological replicates. Additionally, there were only two biological replicates from GSE233 failure/group C, meaning that any analysis should be considered exploratory. For EGAD00001004079 (CD34$^+$CD38$^-$ LSC vs HSC), differential expression was calculated using limma on log-transformed RPKM values, rather than DESeq2 with raw counts, as this dataset was not publicly available. This type of normalisation is not ideal for differential expression, given the altered mean-variance relationship inferred by the RPKM unit. This dataset showed the least similarity compared to the other datasets in its comparison set (Figure 6-5A), and this is possibly why. If this dataset could be reprocessed into raw counts, additional gene discovery could be possible. The comparison set between primitive LSCs and bystanders from GSE184 and GSE233 was hampered by only having two phenotypically similar datasets, which came from different

tissues. Additional datasets would allow for a more thorough comparison between tissues to be conducted, as with only one dataset from each tissue, technical variation cannot be discerned from biological variation. For example, cells from GSE184 were cultured overnight after thawing with low levels of cytokines (SCF, G-CSF, GM-CSF, IL-6, LIF and MIP-α), whereas cells from GSE233 were sequenced almost immediately after thawing. A more comprehensive comparison between cell types could be conducted by doing cluster-wise cell type annotation, then comparing similarly annotated clusters across multiple datasets, preferably where cells are exposed to similar cytokine sets. Additionally, as more datasets with sufficient numbers of $CD34^{+}CD38^{-}$ cells are published, it may become possible to recontextualize the findings presented here.

Considering TKI treated datasets, and datasets that were generated from *in-vitro* culture, output cell types were typically not measured directly, nor was the fraction of remaining BCR::ABL1^{+} cells. This makes interpretation challenging as there isn't a straightforward way to discern whether a change in a gene of interest is occurring due to a difference in the ratio of remaining cell types, and if BCR::ABL1^{-} bystanders are surviving TKI treatment preferentially. For example, in Figure 3-10C, are TKI treated LSCs upregulating CXCR4, or are LSCs dying, and so the higher expression of CXCR4 on bystanders/HSCs is the signal being detected. Similar issues are still present in recently published TKI datasets, such as GSE279135. This dataset was comprised of BM $Lin^{-}CD34^{+}$ cells from 79 CP patients at diagnosis, and after 12 months of TKI treatment[211]. As confirmed by fluorescent in situ hybridization (FISH), no Ph^{+} nuclei were detected at 12 months. Therefore, nothing can be learned about the LSC response to TKI, the focus of the publication, as no BCR::ABL1^{+} population has been captured with sufficient granularity. There were likely BCR::ABL1^{+} LSCs within these samples at a very low abundance, however the phenotype sequenced was not stem enough to be informative about that either. Sequencing of $CD34^{+}$ cells is clearly informative for predicting patient outcomes, as shown by Warfvinge *et al* and Krishnan *et al*, however when the aim is to mechanistically investigate TKI resistance in LSCs, this approach is not informative. There is the possibility of using bulk deconvolution on such datasets to estimate LSC proportions, as there are an abundance of LSC and HSC gene signatures. However, whether this can be used to do more than make predictions regarding outcomes is yet to be shown.

When suggesting predicted targets based on transcriptomic data, the limitations of the technology should be considered, especially for metabolic genes and pathways. For example, most reactions in glycolysis and the TCA can occur in either direction, and are controlled at the post translational level. Neither metabolic flux nor post-translational modification is captured by this method. Similarly, expression of transcripts is not necessarily correlated to protein expression. Post transcriptional modification can have numerous effects on transcript stability, and post translational modification may alter functionality and localization. Ideally a polyomic approach, such as simultaneous metabolome and transcriptome profiling would be conducted. Alternately, a multiway FACS sort on different populations within the $Lin^{-}CD34^{+}CD38^{-}CD90^{+}$ compartment, followed by bulk transcriptomics and metabolomics from these cells could be conducted.

3.4.2 Features of treatment resistant LSCs

To see if we could identify potential targets and metabolic differences between responder groups A, B and C, we compared cells from these patient groups at the $CD34^{+}CD38^{-}$ level from GSE233, EGA and GSE312, and at the predicted primitive level from GSE233. We were unable to detect consistent changes between these datasets, and were unable to detect many changes between the different responder groups at a predicted primitive level. Comparing the different responder groups to their corresponding bystanders did yield additional genes, however without additional datasets to compare these results to, it was not possible to evaluate their consistency.

As discussed previously, it is clearly possible to make relatively accurate predictions regarding the response of an individual to TKI based upon transcriptomic signatures. Similar to Krishnan *et al.*, we conducted single-cell regulatory network inference and clustering analysis to try and identify regulons separating treatment groups at the $CD34^{+}CD38^{-}$ level (not shown), however, we were unable to find consistent regulons, possibly due to low numbers of biological replicates. The usefulness of transcriptomics to discern targetable features specific to treatment response groups is somewhat dependant on which dogma regarding CML disease origin one finds most convincing: Whether the LSC which initially becomes Ph^{+} and $BCR::ABL1^{+}$ is most important, as it would define the

molecular features of all progeny, or whether the immune/inflammatory state of the individual prior to developing the disease is most important. Regardless, if this question could be resolved, it would make interpretation of -omics datasets with treatment response groups more straightforward. Our inability to detect common genes across patient groups between EGA and GSE233 was somewhat addressed by Warfvinge *et al*. in response to reviewer critique, where they show a metric of cluster stability between the same cell types, where they mostly appear relatively similar[89]. However, this analysis was based on canonical marker genes that should be expressed in certain cell types and not stratified by response group, meaning that the consistency of potentially more subtle changes between these groups may not be detected. An analysis that we did not conduct that may have been worthwhile was the comparison between bystanders from the different treatment groups. As HSC phenotypes and gene expression programmes are far more defined than that of LSCs, this could remove a layer of variability and provide information regarding the state of the BM environment, based on changes in cytokine signalling genes and pathways. Lastly, the 2020 ELN guidelines that these datasets have used for responder groups lack classification for certain patient subsets. This lack of stratification possibly introduces noise into analysis based on potentially biologically distinct groups having the same response classification.

3.4.3 Novelty and comparison to previous -omic findings

Given that the meta-analysis was comprised of previously published datasets, a benchmark of the potential relevance of the top scoring genes was the re-identification of known drug targets (discounting $CD34^+$ datasets given their lack of biological relevance). While a direct comparison between the generated and original gene lists was not always possible due to some publications not providing lists of DE genes, a comparison between genes of interest displayed in the corresponding publications could be made at minimum. By examining the published $CD34^+CD38^-$ bulk data and comparing that to the analysis conducted here, generally there was a high similarity on a dataset vs dataset basis. There were however some discrepancies, particularly from EGAD00001004079. This however was to be expected given a proprietary software (Qlucore Omics Explorer) was used, and that differential testing was conducted using 1418 surface markers rather than all genes, corresponding to less stringent p-adjustment. For GSE97562 the top 20 DE genes were displayed in the supplementary data, which

were highly similar to this re-analysis, aside from some differences in rank order. Similarly, fold changes provided in the supplementary data corresponding to E-MTAB-2581 were near identical to the values calculated by this analysis. When considering the scRNA-seq data, DE genes derived from GSE233 and GSE312 were highly similar to what was displayed or shown in the corresponding supplementary data, while at the $CD34^{+}CD38^{-}$ level from GSE184 we also saw an enrichment of MYC and E2F signatures. This similarity was unsurprising, given that perhaps one of the largest factors influencing differential expression analysis is cell type annotation, which was not required in these instances. The comparisons made using EGA, and GSE184 at the $CD26^{+}CD90^{+}$ level had not been made in the original publications, meaning benchmarking was not possible.

When looking at DE genes across all comparison sets (padj < 0.05 and absolute log2FC >1) 5 genes were identified. As discussed previously, GAS2 and CACNA1D are widely known in the field, while RXFP1, ZNF385D and FRY have been of little to no interest. While 5 genes appears to be a low number, considering variations in sample origin/processing (e.g. cryopreserved, fresh, culture or no culture, cytokine milieu, blood or BM and time to sort), technologies used (e.g. different sequencing platforms and different probe sets) and contaminants in cell types (e.g. bystanders within bulk LSC datasets, inaccuracies in cell sorting and variations in LSC phenotypes) it's not overly surprising that more genes were not identified. However, when lowering the stringency to be DE in 2/3 datasets per comparison set (comparing bystanders to LSCs), ~60 genes were identified (Figure 3-10A). While a lower log2FC threshold was also considered for this analysis, the gene list became too large to be manually inspected, yet not large enough for ORA to be effective. While Figure 3-10A displays many genes which are known to be DE in LSCs (e.g. GAS2, CACNA1D, CXCR4, CD33, PIEZO2), it doesn't contain target genes from corresponding publications which have been validated to be DE through additional methods (e.g. CD36, EZH components). Aside from the previously mentioned technical limitations of comparing transcriptomic experiments, this issue could potentially arise from the use of arbitrary fold change thresholds rather than thresholding on a per-dataset basis derived from the distribution of fold-changes. Another weakness of this analysis was a larger focus at the gene level rather than at the pathway level. That said, pathway analysis is still inferior to unsupervised methods when aiming to identify novel

genes from previously published datasets, as knowledge of a genes function for a specific tissue is requisite, and the likelihood of identifying the same genes as from the publication is greater.

Chapter 4 *Ex-vivo* culture of CML LSCs and the effects of BCAA restriction *in-vitro* and *in-vivo*

4.1 Introduction

As previously outlined, targeting altered metabolism of CML LSCs has been a successful strategy in proof of principle studies. For example, the enzymatic degradation of arginine, inhibition of pyruvate anaplerosis, and inhibition of *de-novo* purine synthesis[94,95,212]. However, most hypothesis generating experiments are conducted in cell lines, commonly K562s and KCL22s, which are derived from BP patient samples. These cells do not have an LSPC phenotype, and are transcriptionally distinct from LSCs[114,213,214]. This means that when testing new ways to target CML LSCs, agents which target stemness may not show the same effectiveness given the phenotypic and transcriptional dissimilarity. Therefore, we aimed to utilize *ex-vivo* culture of LSCs to improve the phenotypic stemness at which LSCs could be investigated *in-vitro*, and to potentially reduce mouse usage.

To assess the usefulness of *ex-vivo* LSC culture, we cultured HSCs and LSCs under expansion and hibernation conditions, and conducted initial transplantation experiments. To decide how to target CML LSCs, we mimicked the AA dropout screen conducted by Rattigan *et al.*, showing that BCAA restriction had a strong anti-leukaemic effect in expanded cells alongside K562s and KCL22s[212]. We then metabolically profiled K562s and KCL22s using LC-MS and Seahorse to mechanistically understand how BCAA restriction was causing a reduction in cell number, however the exact reason remains unclear. Then, to determine whether BCAA restriction could be used to treat CML *in-vivo,* we tested the effectiveness of dietary intervention in mouse models, finding that when combined with CML, BCAA restriction was poorly tolerated, causing rapid weight loss. Additionally, we held a focus group to determine whether leukaemia patients would be interested in diet-based therapy, and in what form this might take.

4.2 *Ex-vivo* culture of CML LSCs

4.2.1 CML LSCs can be expanded *ex-vivo* from sorted or enriched populations.

Previously, Patel *et al. 2022* showed that LSCs could be expanded from a sorted population of $Flt3^{-}$LSK-SLAM cells, however FACS can often act as a bottleneck depending on whether a sorter and a flow facility member is available.[129]. For this reason we aimed to see if LSC culture could be initiated from a c-kit enriched population over the course of 14 days, and whether this model could be useful for determining the metabolic weaknesses of CML LSCs (Figure 4-1). In Figure 4-1A we show that after 14 days of *ex-vivo* expansion from a c-kit enriched population, there was a mean fold change increase of 2.7, 6.5, 6.1 and 14.1 across live, Lin-, LSK and $CD150^{+}$LSK fractions respectively, with the frequency of stem phenotypes becoming more prevalent within the cultures. Figure 4-1B shows that sorted LSCs can be cultured in hypoxia, increasing stemness at the expense of reduced cell numbers, indicating that the effects of hypoxia could be studied using this model. While the frequency of LSKs is increased in hypoxic cultures, the proportion of Lin- cells is significantly decreased, indicating that hypoxia combined with these culture conditions supports the growth of lineage committed leukaemic cells, an effect which was also shown in expanded HSCs[122].

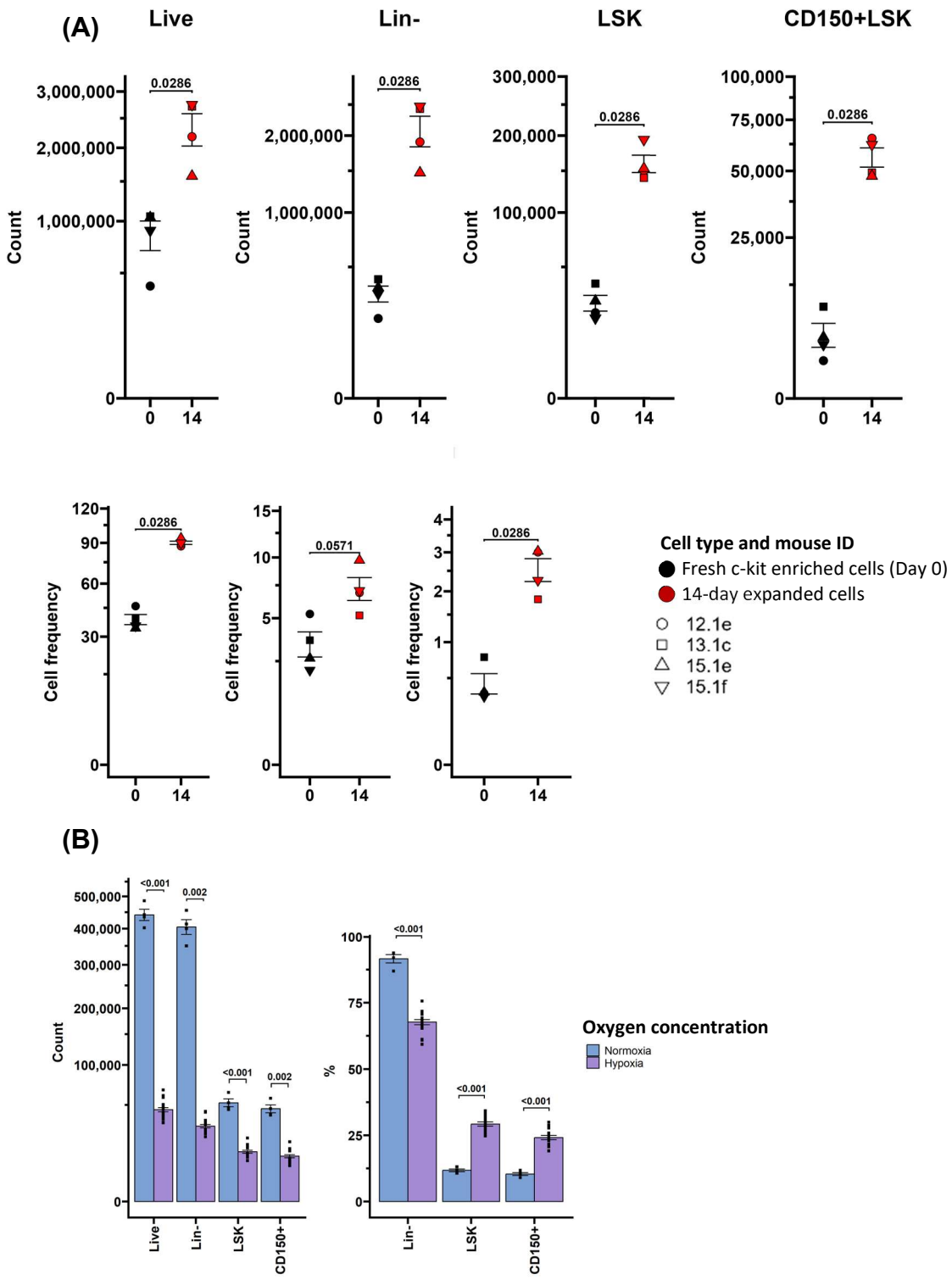


Figure 4-1 *Ex-vivo* expansion of CML LSCs from bulk and sorted LSCs. (A) Number of (top) or frequency of (bottom) expanded LSCs at 14 days compared to starting number of cells from a c-kit enriched population. 4 biological replicates (BRs) and 1 technical replicate (TR) of pooled cells at 14 days, statistics calculated by Mann-Whitney U (MW) test. (B) Number of (left) or frequency of (right) expanded LSCs cultured in normoxia or hypoxia starting from 500 sorted LSCs. 1 BR and 3-7 TR, statistics calculated by MW test.

Given that expanded LSCs show poor engraftment and leukaemogenic potential, we intersected bulk transcriptomic datasets GSE242036 and GSE189368, to discern pathways that may be involved in loss of stemness. GSE242036 compares DTG induced LSKs to uninduced counterparts, while GSE189368 compares *ex-vivo* expanded CD150^{+}LSKs LSCs to expanded HSC counterparts, rationale being genes unique to each comparison could describe why expanded LSCs are losing stemness (Figure 4-2A). While a similar comparison was made by Patel *et al.*, the fresh HSC/LSC phenotype used was not specified (but could be presumed to be LSK-SLAM), and their comparison of these cell types yielded >10,000 differentially expressed genes, which is an unconvincingly high number, as this represents ~1/3 to 1/4 of most reference genomes prior to any filtering steps.

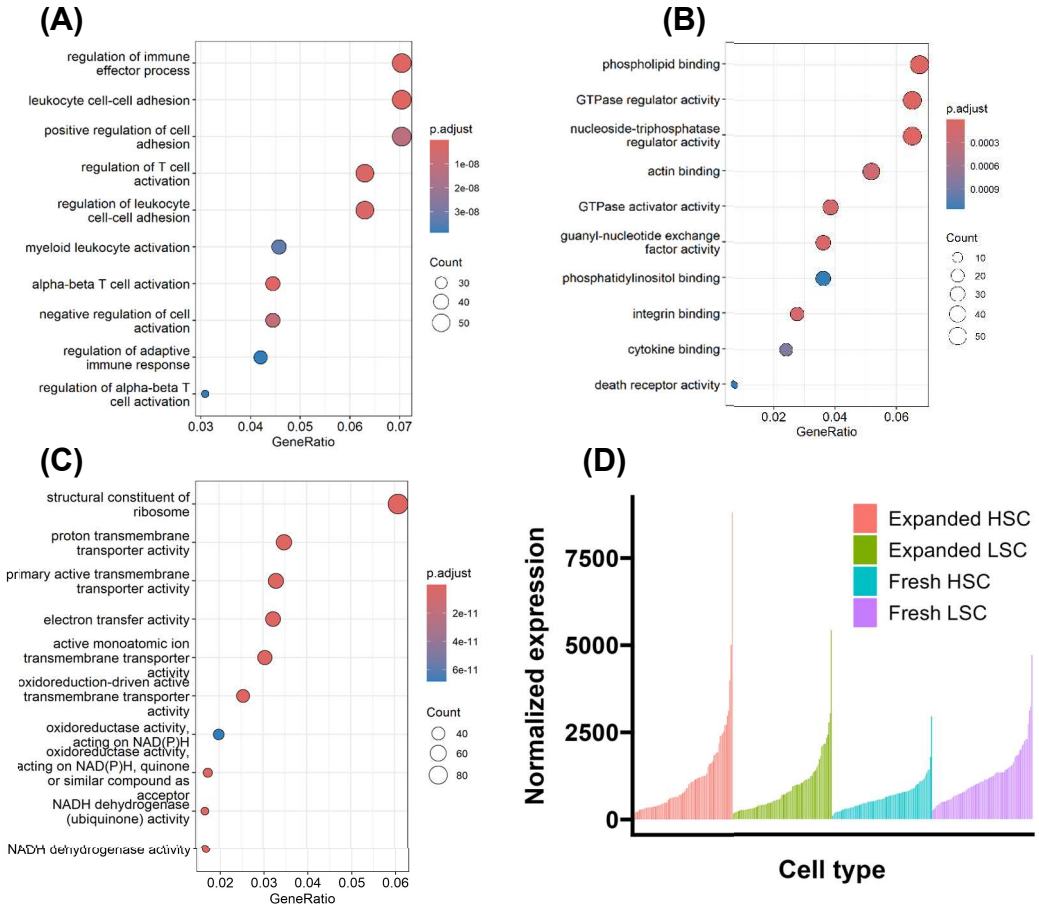


Figure 4-2 Comparison of expanded LSC vs expanded HSCs, and fresh LSCs vs fresh HSCs. (A-B) ORA of genes unique to the expanded LSC vs HSC comparison, using molecular function and cell component ontologies. (C) ORA of genes unique to the fresh HSC vs LSC comparison, using the molecular function ontology. (D) Normalized expression levels of OXPHOS genes from GSE189368 and GSE242036, when processed together with experiment used as a covariate.

Conducting ORA on the output gene lists showed that the most upregulated pathways only expressed in the expanded LSCs were related to cell adhesion,

GTPase activity, and cytoskeletal components (Figure 4-2A/B). Analysing genes upregulated only in fresh LSKs was less informative as mostly metabolic signals were detected (Figure 4-2C/D). This was because while expanded HSCs and LSCs had greater expression of mitochondrial genes than fresh HSCs and LSCs, there was little difference between these groups, while fresh LSCs had greater expression of mitochondrial genes compared to fresh HSCs. This suggests that when studying the effects of mitochondrial inhibitors on expanded LSCs, using hypoxic culture or modifying the cytokine milieu to reduce cycling could be more physiologically relevant. Given changes in adhesion genes were prevalent from intersecting the two datasets, we tested if changing the type of plate could improve LSC culture, using CellBind plates which enhance *ex-vivo* HSC culture. This was initially successful, however when we conducted two-week culture we observed substantial changes in morphology (data not shown). It is possible that using an agarose coating or plates with different TC coating could alter gene expression toward a more stem phenotype, however this remains untested[128].

Previously it has been shown that SCF expands normal LT-HSCs, but causes LT-LSCs to produce MPPs and LSKs. Additionally, as disease causing LSCs are MPL^{hi} (the cognate receptor for TPO), we decided to test whether LSCs in culture were dependent on TPO under normoxic and hypoxic conditions (Figure 4-3A/B). Decreasing TPO concentrations did not have a strong effect in either normoxia or hypoxia, however under hypoxic conditions decreased TPO concentrations resulted in a downward trend in cell number, while under normoxia a slight upward trend in cell counts but a reduction in stem phenotype was observed. This indicates that culturing LSCs in hypoxia may better preserve stem features in expanded LSCs, as under normoxic conditions cells showed poor sensitivity to TPO.

To work around the poor engraftment of expanded LSCs, we aimed to expand DTG/STG cells maintained on doxycycline (DOX) to prevent BCR::ABL1 expression, followed by non-conditioned transplant, to generate a non-irradiative model of CML (Figure 4-4). To this end, BM was pooled from two $tomato^{+}$ DTG mice (male and female), and from 2 $tomato^{+}$ STG mice (male and female), wherein cells were c-kit enriched and cultured for 28 days. DTG cells were also supplemented with DOX (0.5µg/ml), while $tomato^{+}$ mice were used as this was done in collaboration with Dr Zerbst, with the additional aim of conducting *in-vivo* imaging. DOX was

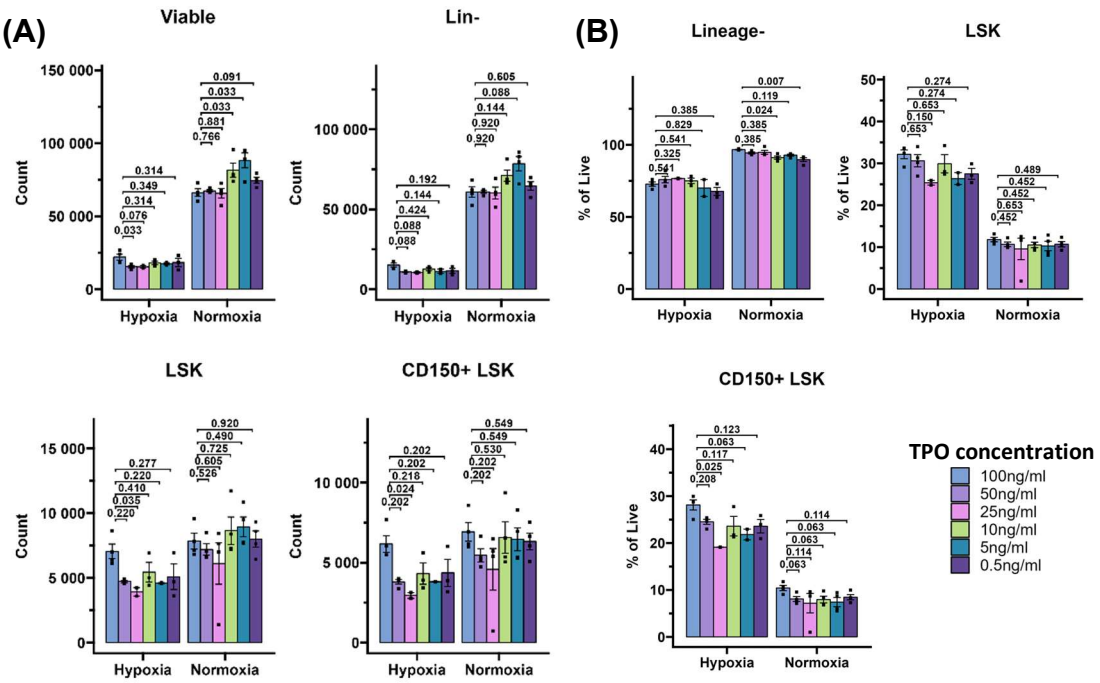


Figure 4-3 Effect of different TPO concentrations on expanded LSC growth in normoxia or hypoxia. Cell counts (A) and frequencies (B) of live, Lin-, LSK and CD150^{+} LSK fractions. 4TRs and 1BR, statistics calculated by Kruskal-Wallis (KW) test followed by Dunns test with Benjamini-Hochberg (BH) adjustment.

removed from a well of DTG cells after 14 days. Stem phenotypes were measured prior to the transplantation of 1x10^{6} STG or DOX^{+}DTG cells into 12 Bl6 mice (Figure 4-4A). In a separate experiment, 1x10^{6} STG cells cultured in hypoxia for 21 days were transplanted into Bl6 mice, however CD48 and EPCR expression was not measured for this experiment. This was because expanded LSCs only retain CD150 expression, therefore this was the resolution at which we typically monitored corresponding expanded HSCs (Figure 4-4E). From Figure 4-4A we observed that STGs and DOX^{+}DTG cells had a similar frequency of SLAM cells, however DOX^{+}DTG cells had an increased frequency of SLAM-EPCR^{+} cells. This may have been due to an off-target effect of DOX, which has been shown to cause cell lines to become more glycolytic[215]. Increased glycolysis and decreased OXPHOS in expanded HSCs has been shown to cause an increase in EPCR^{+} cell frequency, possibly explaining this effect[122]. DOX^{-}DTG wells had reduced LSK, SLAM and SLAM-EPCR^{+} frequencies, suggesting that DOX supplementation was successfully controlling BCR::ABL1 expression, and its withdrawal was inducing BCR::ABL1, resulting in reduced phenotypic stemness. When measuring engraftment (Figure 4-4B), we found that by week 3 PB chimerism was lost in all females transplanted with either DTG or STG cells, while male mice retained low levels of engraftment. When analysing

BM chimerism (Figure 3.4C/D), only mouse 14.4a from the STG group exceeded the 1% PB chimerism threshold, which could be considered successful engraftment. Examination of the LSK gate from the two mice showing the greatest engraftment (14.4a and 105.2b), showed that 14.4b had an abnormally high proportion of transplanted LSKs, whereas 105.2b showed LSK frequencies similar to that of host CD45.1 LSKs. When Bl6 mice were transplanted with $1x10^6$ hypoxia cultured STG cells (Figure 4-4E), 1 of 3 mice (all males) showed PB chimerism >1% at 6 weeks, wherein the experiment was ended. As some level of engraftment did occur across both experiments, this suggests that there were definitely functional stem cells in culture. It is possible that engraftment may have been inconsistent due to stochastic factors such as whether a mouse received a functional HSC, and whether it would contribute to PB chimerism in the short term in a primary recipient.

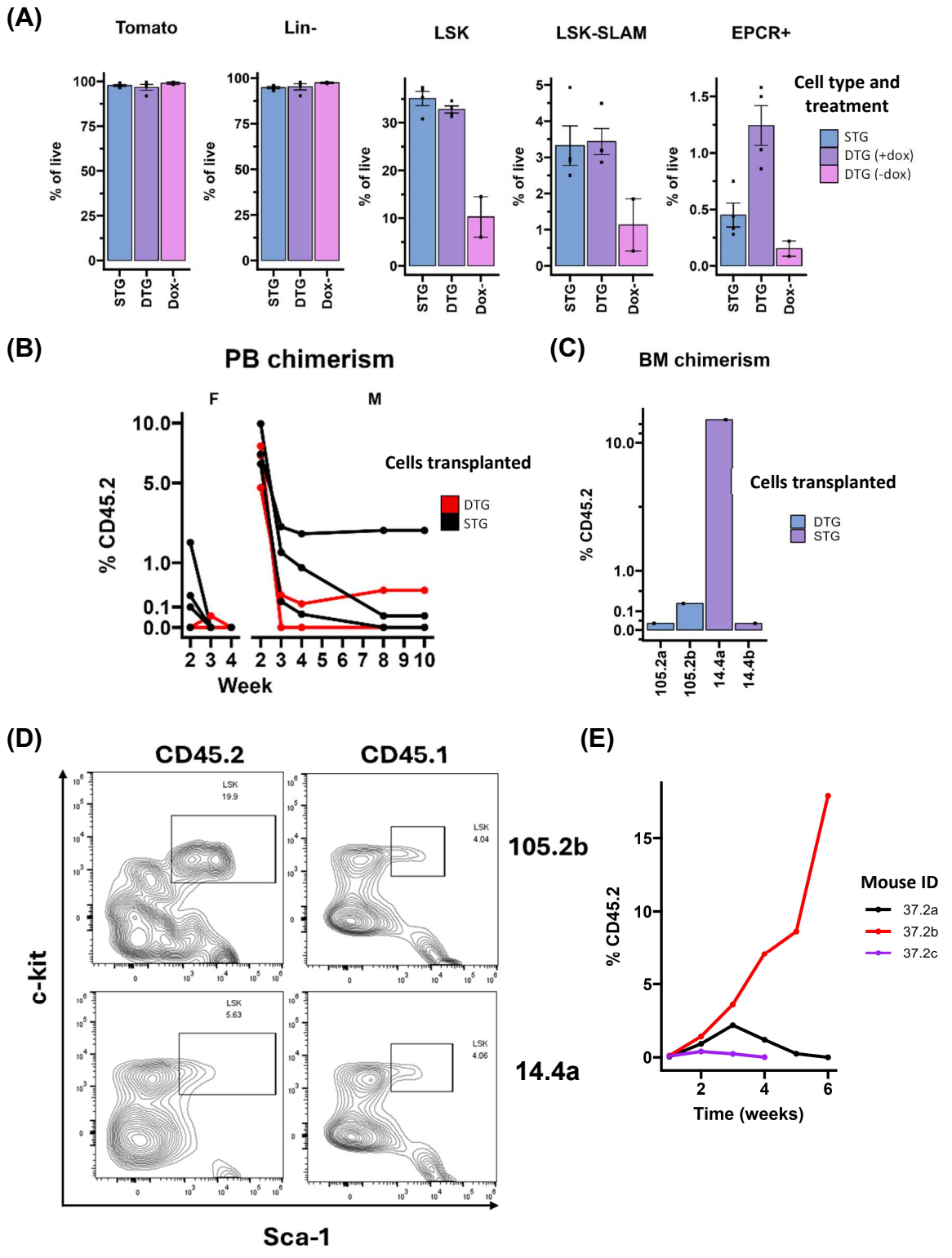


Figure 4-4 Non-conditioned transplant of *ex-vivo* expanded cells. (A) Flow cytometry of stem phenotypes from 28 day expanded STG or DTG cells +/- DOX. (B) PB chimerism of females (left) and males (right) when transplanted with 1x10^6 *ex-vivo* expanded cells. (C) BM chimerism of male mice, only mice which showed chimerism were plotted. (D) LSK gate of CD45.1 and CD45.2 cells from 105.2b and 14.4a, for host (CD45.1) and transplanted (CD45.2) cells. (E) PB chimerism from mice transplanted with 1x10^6 *ex-vivo* expanded cells cultured under hypoxic conditions for 28 days.

4.2.2 LSCs can be grown in hibernation culture and retain sensitivity to a combination of TKI and HCQ

As previously outlined, hibernation culture allows single HSCs to retain their molecular identities and engraftment potential, with the only cytokine present being IL-11. We therefore aimed to see if LSCs (LSK-SLAM) could be cultured under these conditions, and whether this could be suitable for investigating the metabolic properties of LSCs (Figure 4-5). First, single HSCs were sorted and cultured for 7 days and quantified by microscopy at days 3 and 7, with SCF (100ng/ml) treated wells to control for sorting accuracy (Figure 4-5A). The purpose of SCF treatment being twofold, first to get a baseline of how many wells were successfully sorted into, and to provide a depth of field suitable for spotting the single HSCs. When sorting HSCs, we also found similar results to Oedekoven *et al*, where HSC viability decreased over the course of 7 days, and that SCF supplementation expanded around 80% of single HSCs[127]. Next, cells from a Bl6 and an induced DTG mouse were sorted into SCF supplemented wells (and not wells without SCF due to a sorter error), showing that LSCs could tolerate this combination of cytokines and media (Figure 4-5B). To determine the applicability of screening to this method, we sorted HSCs or LSCs from a Bl6 and an induced DTG mouse into control media, media supplemented with HCQ (10μM), Nilotinib (2μM) or a combination of both, and quantified the number of surviving cells at day 3 (Figure 4-5C). We found that LSCs were more resilient than HSCs in untreated media, and that Nilotinib and HCQ singularly had a weak effect, but in combination were highly effective.

Given that counting single HSCs was extremely time consuming, and that visual conformation is neither scalable nor in the long term reliable, we bulk sorted HSCs or LSCs into hibernation culture so that phenotypic measurements could be made (Figure 4-6). Additionally, LSCs were cultured without IL-11 to see if phenotypic LSCs required IL-11 to survive (Figure 3.6A), as LSC cytokine usage differs compared to HSCs. HSCs or LSCs were cultured for 3 days, pooled, then 220 phenotypic HSCs (alongside the other cells in culture) were transplanted into 3 irradiated STG mice (one of which died due to irradiation poisoning) alongside $1.5x10^6$ CD45.2 STG cells (Figure 4-6B/C). From this transplant, we saw that PB chimerism of hibernation HSCs peaked at weeks 2 and 3, which then dropped to undetectable levels by week 6. After 12 weeks these mice were culled, showing

<1% BM chimerism. Oedekoven *et al*, show that single cell transplants from 7-day cultured HSCs had >1% PB chimerism at 16-22 weeks in 33/69 mice, suggesting at least one element of this experiment was unsuccessful. To harvest cells for transplant, cells were centrifuged then supernatant discarded, rather than sorting directly into the transplant volume, therefore it is possible that loss of cells prior to transplant may have caused low engraftment. Still, this seems unlikely given the transient PB chimerism of 4-8% that was detected at week 3.

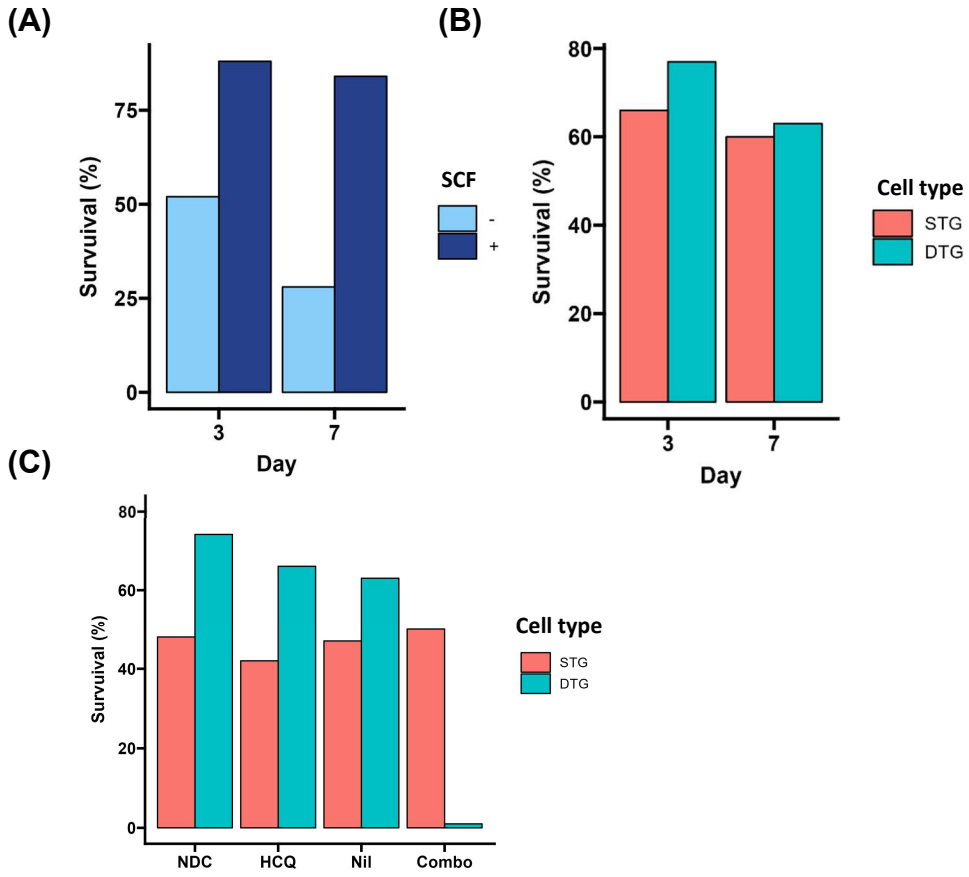


Figure 4-5 Hibernation culture of single HSCs and LSCs. (A) Survival rate of HSCs +/- SCF at 3 and 7 days of culture (n=108 cells for +/- SCF). (B) Survival rate of HSCs or LSCs from SCF^+ wells. (n=36 cells for DTG/STG). (C) Survival rate of HSCs or LSCs treated with vehicle, HCQ (n=66), Nilotinib (n=66), or both (n=66), for 3 days. Statistics were not calculated due to single biological replicates being used per experiment.

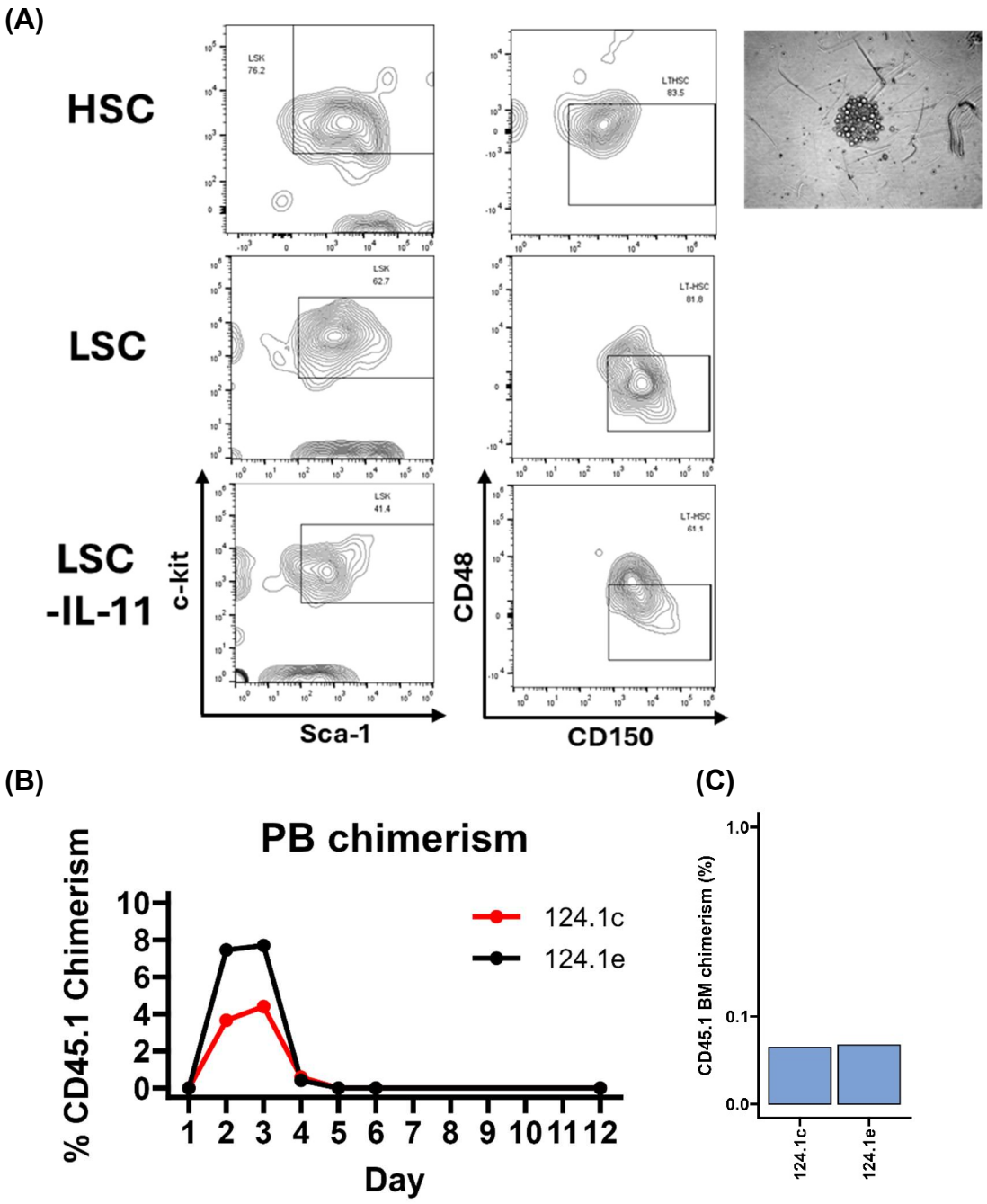


Figure 4-6 Flow cytometry and transplantation of hibernation cultured HSCs and LSCs. (A) LSK and SLAM gate of HSCs, LSCs and LSCs without IL-11, alongside an image of hibernation HSCs. (B) PB chimerism in mice transplanted with 220 CD45.1 LT-HSCs after 72 hours of hibernation culture, alongside $1.5x10^6$ CD45.2 whole bone marrow cells. (C) BM chimerism at 12 weeks.

4.3 Exploring the metabolic dependencies of CML LSCs

4.3.1 Amino acid dropout screen identifies BCAAs as a metabolic weakness in expanded CML LSCs

Given the difficulty of hibernation culture and lower cell numbers, we decided to use the Wilkinson *et al*. expansion method to investigate the metabolic properties of CML LSCs. To do so, we took an approach previously applied by our laboratory, where Plasmax, a physiologically similar media to blood, was formulated missing each of the AAs, followed by 72 hours of cell culture[212]. LSCs from induced DTG mice were expanded then seeded at $2x10^4$ cells/well into Plasmax missing each of the AAs, with relative abundances measured by flow cytometry (Figure 4-7). We found that dropout of essential AAs reduced cell counts (Figure 4-7A/B), however only modest statistical significance was detected. Across phenotypes the strongest effects were observed due to the dropout of essential AAs methionine, isoleucine and tryptophan. This was similar to the findings of Rattigan *et al.*, who observed the strongest effects from methionine, isoleucine and histidine when profiling K562s. Dropouts of non-essential AAs which showed similar cell numbers compared to full Plasmax (aspartic acid, asparagine, glutamic acid, alanine, citrulline and glycine) also showed little to no effect size in K562s. As valine, leucine and isoleucine showed strong effect sizes in cell lines and expanded cells, we decided to further investigate the BCAAs. We also selected the BCAAs based on the identification of BCAT1 as a candidate gene in our meta-analysis, relevance to BC CML/BCR::ABL1^{T315I}, AML, and the fact that BCAA restriction had been reported to be tolerated in both humans and mice[49,147,150,157].

To see if similar effects to the AA dropouts could be achieved pharmacologically, we use the LAT1 inhibitor JPH203 (20μM) on expanded HSCs and LSCs (Figure 4-8). In both expanded HSCs and LSCs JPH203 treatment increased the number of cells compared to controls. In HSCs, treatment caused a depletion of LSKs into the non-LSK Lin- fraction, while this did not occur in expanded LSCs. This suggests that in HSCs JPH203 was causing LSKs to divide and differentiate, however no anti-leukaemic effect in expanded LSCs was detected. This, alongside unpublished data generated by Dr Kevin Rattigan, which showed that JPH203 could lower intracellular AA concentrations in cell lines but not reduce viability, suggested that JPH203 was not a suitable replacement for a BCAA dropout.

To investigate the effects of restricting BCAAs, we formulated BCAA- (and valine-) Plasmax (Figure 6-10a), expanded HSCs and LSCs, then seeded them at a density of $2x10^4$ cells/well in BCAA deficient media for 69 hours (Figure 4-9A/B). Additionally, apoptosis was conducted on expanded LSCs at 48 hours, while whether IM remained effective in both Hams F12 nutrient mix, media typically used for this expansion culture, and Plasmax, was also evaluated (Figure 4-9C/D).

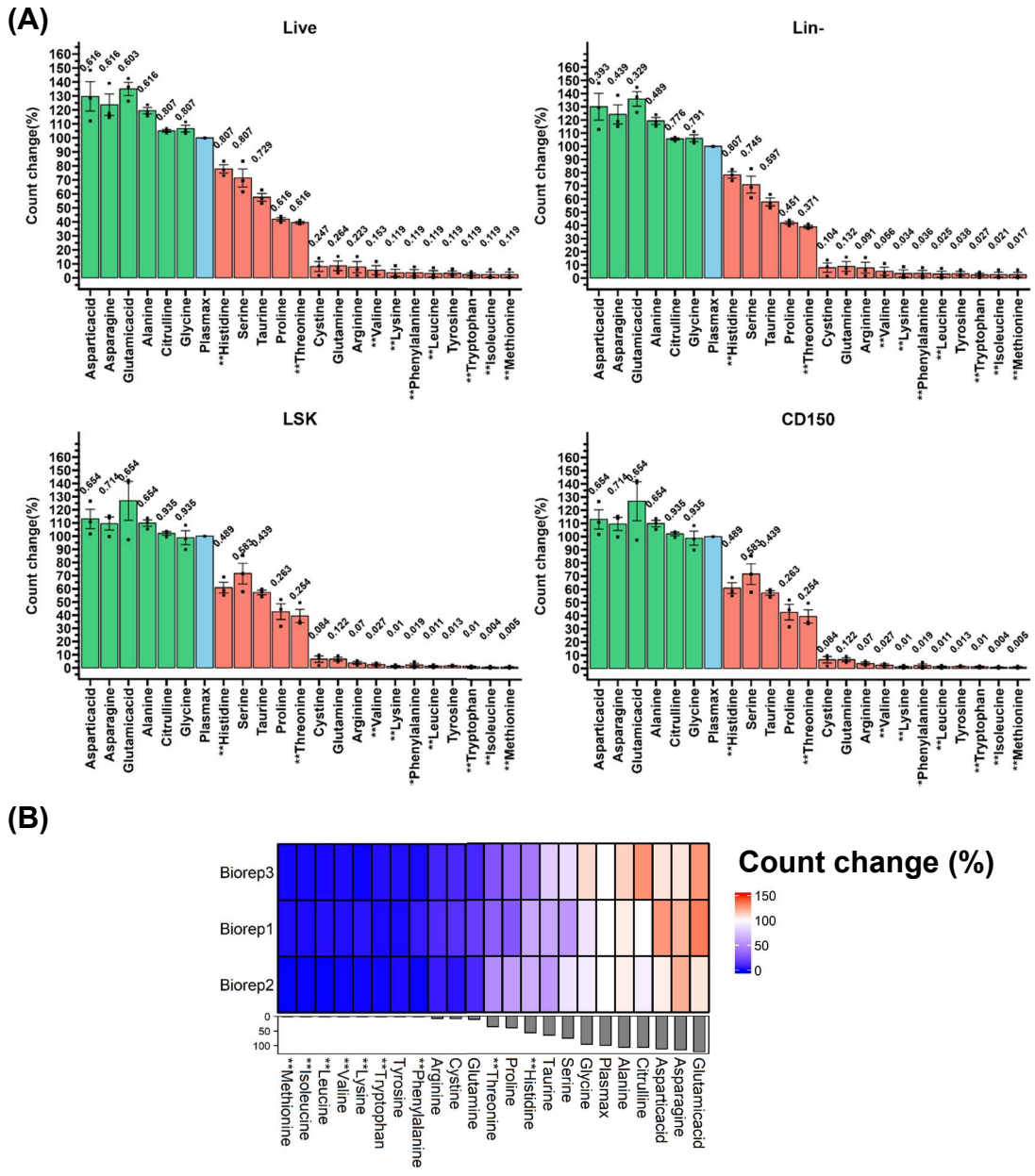


Figure 4-7 Changes in cell numbers when expanded LSCs were grown in Plasmax missing each of the AA. (A) Relative reduction in in cell count for cells compared to their corresponding biological control, from all cells, Lin- cells, LSKs, and CD150$^+$LSKs. 3BRs and 3TRs, statistics calculated using a KW test followed by a Dunn's test with BH adjustment (B) Heatmap of the changes in cell counts in the CD150$^+$LSK compartment. Essential AAs marked with *.

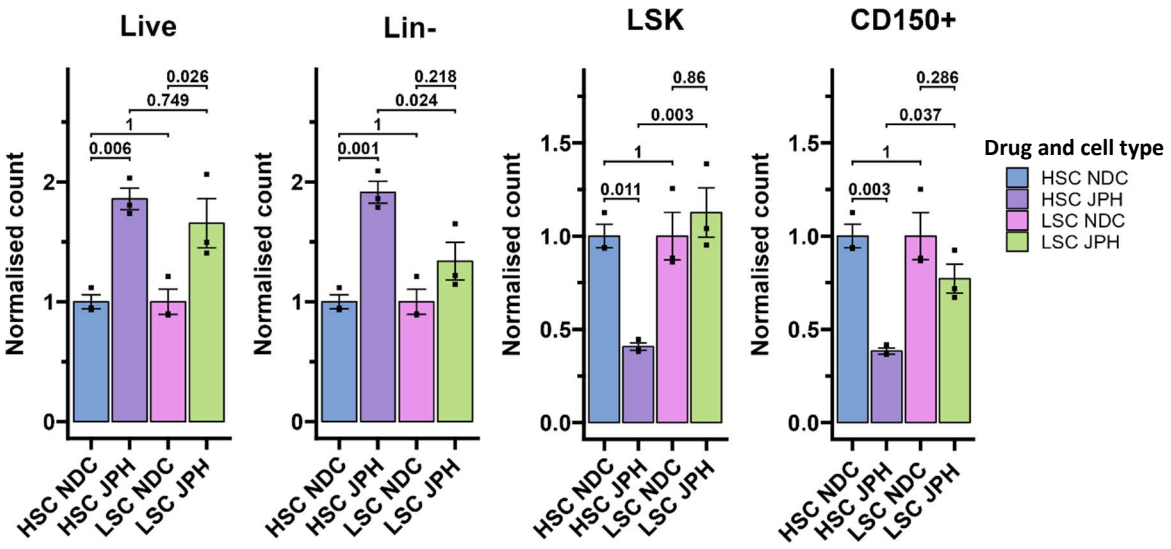


Figure 4-8 Effect of JPH203 treatment in expanded HSCs and LSCs. Change in relative cell count in expanded HSCs or LSCs treated with JPH203. 1BR, and 3TRs, statistics calculated by 2-way ANOVA followed by Šidák correction.

By measuring cell growth, in line with previous experiments we saw that BCAA restriction reduced the confluency in both expanded HSCs and LSCs, without a window to selectively target LSCs. We found that culture in 50% BCAA media had little effect on confluency or cell death at any time point for both expanded HSCs and LSCs, however 25% BCAA media or lower had strong effects at 48 and 69 hours compared to full media. Apoptosis in BCAA$^-$ media showed a similar effect size to that of the individual BCAA dropouts, with a viability reduction of ~80%. IM treatment at all concentrations showed a decrease in relative cell numbers compared to the NDC control at all phenotypes, with a somewhat dose dependant effect.

We theorized that autophagy could be a method by which LSCs in BCAA/valine restricted media could be surviving, therefore the lysosomal content and relative cell counts of expanded LSCs +/-Bafilomycin (Baf) and +/-BCAAs were measured after 24 hours of culture (Figure 4-10). We found that BCAA restriction in the absence of Baf increased lysosomal content, and that in non-LSKs Baf treatment reverted this in all media types (Figure 4-10A). Singularly, Baf treatment reduced the number of live and Lin- cells in all medias but not in LSK and CD150$^+$LSK populations, with BCAA$^-$ or valine- media synergizing with Baf treatment (Figure 4-10B).

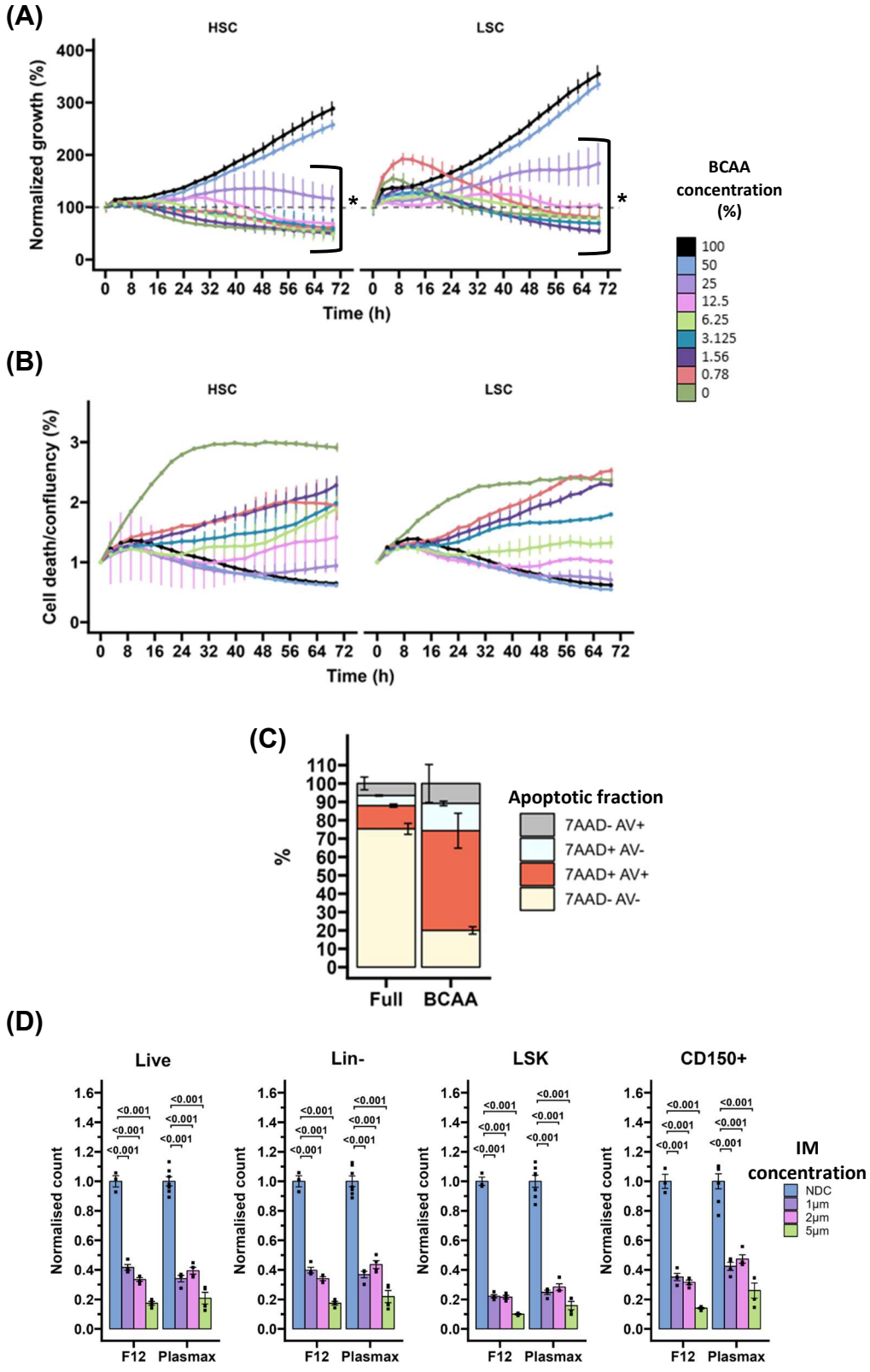


Figure 4-9 Growth and cell death of expanded HSCs and LSCs cultured in full media or BCAA deficient medias. (A) Confluence of HSCs or LSCs in BCAA deficient media, measurements made every 4 hours for 69 hours. 2BRs with 2TRs for HSCs and LSCs, statistics calculated at 48 and 69 hours by KW test followed by Dunns test with BH adjustment. (B) Cell death/confluency for these cells. (C) Apoptosis of expanded cells in full or $BCAA^-$ media, 2BRs, 1TR. (D) Effect of TKI treatment on expanded LSCs in either Hams F12 nutrient mix or Plasmax. 1BR and 3TRs, statistics calculated by ANOVA followed by Dunnet test.

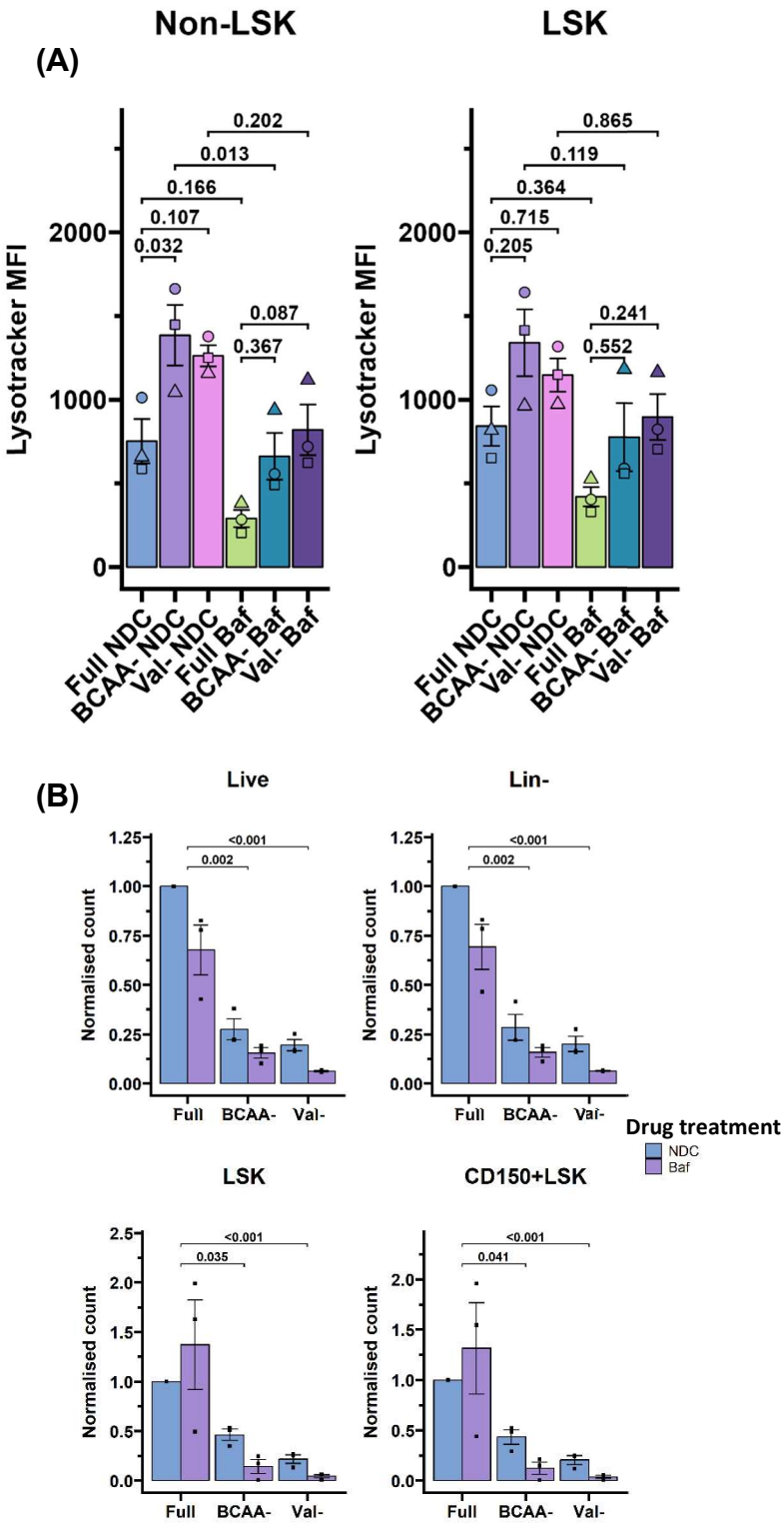


Figure 4-10 Effect of autophagy inhibition on expanded LSCs grown in BCAA or valine deficient media. (A) MFI of lysotracker in full, BCAA$^-$ and valine$^-$ media with or without bafilomycin for 24 hours in LSKs or non-LSKs. 3BRs and 3TRs, statistics calculated using 2-way ANOVA with Sidak adjustment. (B) Changes in cell counts compared to corresponding biological controls associated with lysotracker MFI in the live, Lin-, LSK and CD150$^+$LSK compartments. Statistics calculated by Aligned Rank Transform ANOVA, comparing the interaction between Baf treatment and BCAA$^-$ or Val$^-$ media to that of full media.

4.3.2 BCAA restriction reduced cell number and metabolic outputs in K562s and KCL22s

Next, we aimed to see if BCAA restriction had the same effect in expanded cells compared to K562s and KCL22s. To facilitate this, we formulated custom RPMI (Figure 6-10B). K562s and KCL22s were seeded at $1x10^4$/well and cultured in full, 50%, 25% and 0% BCAA RPMI, where growth was assessed over 72 hours using cell counting and resazurin (Figure 4-11A/B). Additionally, the correlation between relative cell counts and relative fluorescence was conducted, alongside measurement of apoptosis and cell cycle (Figure 4-12). Similar to expanded LSCs, a 50% reduction in BCAA concentration had little effect on cell number in K562s but had a stronger effect in KCL22s, with further reductions in BCAA concentration further reducing growth. When examining the correlations (Figure 4-12A/B), there was a linear relationship between reduction in cell count and reduction in fluorescence at 48 and 72 hours, suggesting a dose dependant reduction in metabolic output. However, K562s cultured in 50% media at 48 hours clearly deviate from the line of best fit, suggesting that additional datapoints between 50% and full media would have improved the modelling of this effect. Conducting apoptosis on these cell lines (Figure 4-12B) shows that the apoptotic effect of BCAA reduction only occurs strongly at a complete restriction of BCAAs in K562s and KCL22s. Further investigation using cell cycle (Figure 4-12C) analysis showed that BCAA restriction in KCL22s caused an increased proportion of cells in G1 and a reduction in S phase, while in K562s this trend was only partially true.

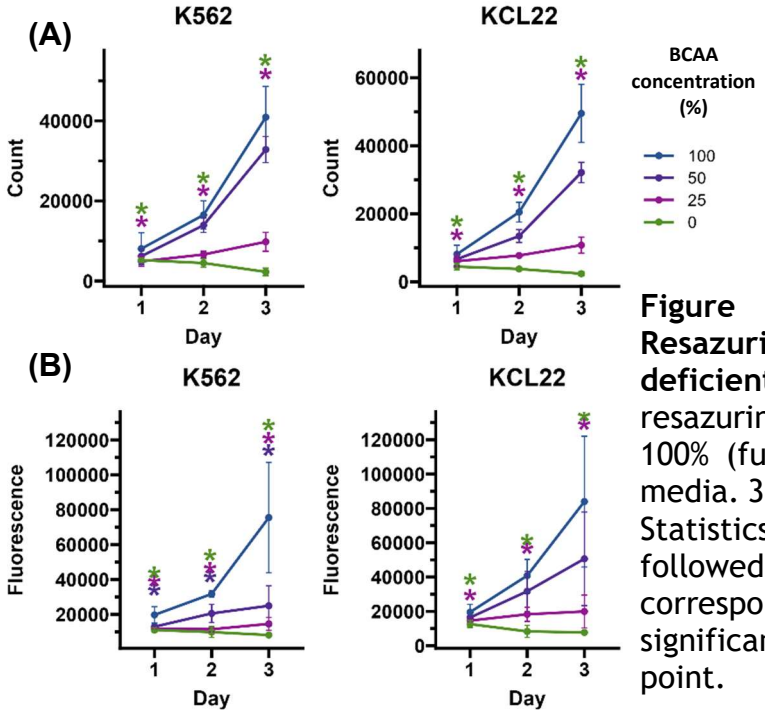


Figure 4-11 Cell counts and Resazurin assay of cell lines in BCAA deficient media. Cell counts (A) and resazurin (B) of K562 and KCL22s at 100% (full), 50%, 25% and 0% BCAA media. 3BRs and 3TRs per time point. Statistics calculated by ANOVA followed by Dunnet test. Stars of the corresponding colour denote significance ($p < 0.05$) for that time point.

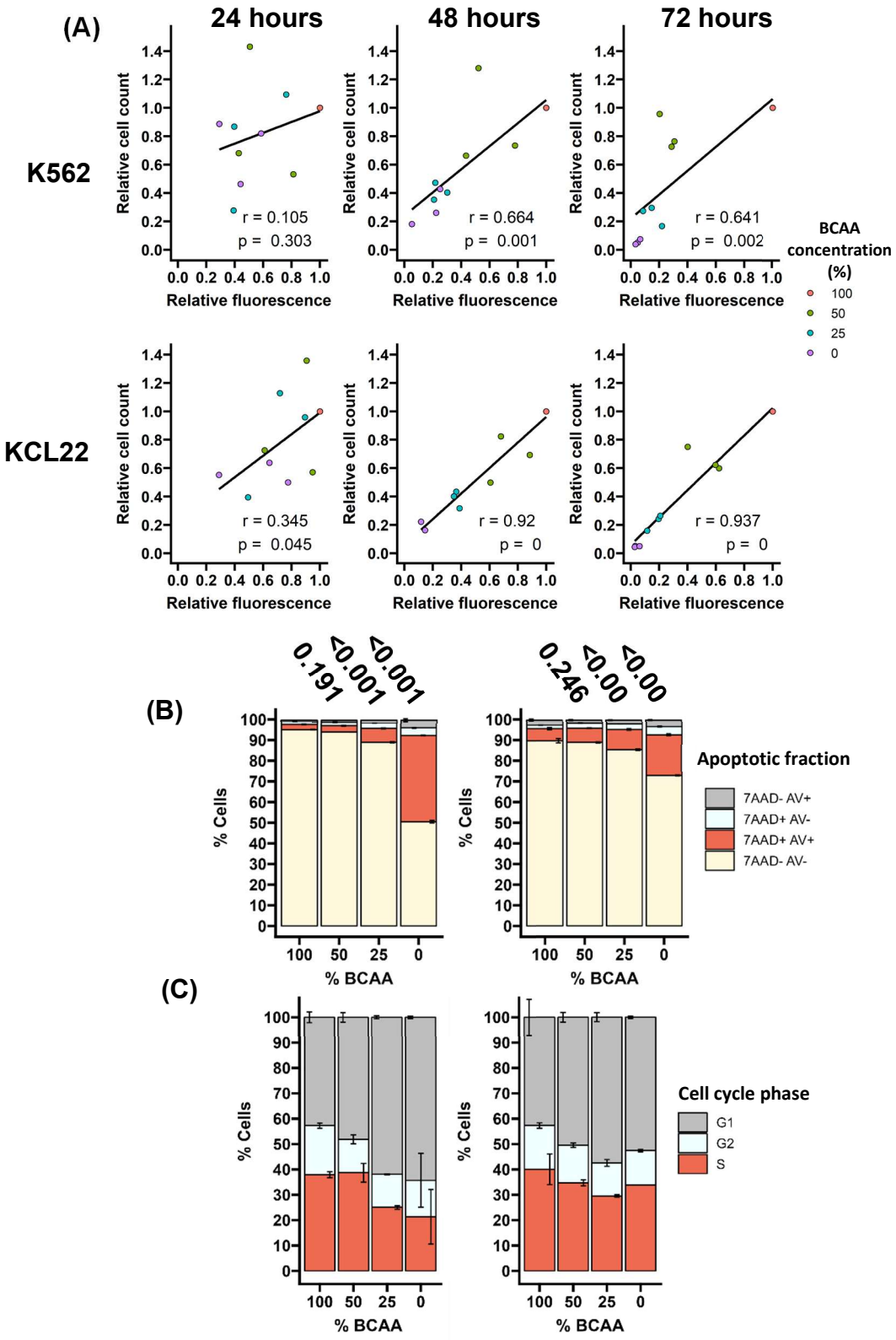


Figure 4-12 BCAA culture kinetics. (A) Linear model of the relationship between cell counts and Resazurin fluorescence. (B) Apoptosis of KCL22s (left) and K562s (right) cultured in BCAA deficient medias for 48 hours. 2BRs and 3TRs, aside from 0% BCAA media (1BR 3TR). Statistics calculated by KW test from one set of TR and BRs, followed by Dunns test with BH adjustment. (C) Cell cycle for of KCL22s (left) and K562s (right) cultured in BCAA deficient medias over 48 hours, 1BR and 3TRs.

4.3.3 Metabolic profiling of K562 and KCL22s cultured in BCAAs

Given that in Figure 4-12 it can be seen that the decrease in metabolic capacity caused by BCAA restriction is somewhat linear, we aimed to characterise these changes by conducting the Seahorse mito stress test on K562s and KCL22s grown in different concentrations of BCAAs at 24 and 48 hours (Figure 4-13 and Figure 4-14). In short, the Seahorse assay allows for the measurement of the OCR and ECAR as a function of a cells oxidative and glycolytic capacity, and when combined with drugs which perturb metabolism, additional parameters can be derived. We used Oligomycin, an inhibitor of ATP synthase to determine the basal rate of OXPHOS, CCCP, a mitochondrial uncoupler to determine the maximal contribution of OXPHOS, and rotenone alongside antimycin-A to completely prevent OXPHOS through the inhibition of complex I/III, allowing for the non-mitochondrial OCR to be determined.

From this, we found that KCL22s at 24 hours at 50% or 25% BCAA content had reduced basal respiration rates. The same trends were present for maximal, ATP-linked and spare respiration, however not all these changes reached statistical significance. At 48 hours the same trends were present, but with reduced statistical significance. This suggests that KCL22s can mitigate the effects of BCAA restriction at the expense of reduced cycling (Figure 4-13B and 4-14B). K562s at 24 hours showed a different trend, with 50% BCAA cultured cells showing similar basal and ATP linked respiration, but reduced maximal and spare respiratory capacity compared to full media (Figure 4-13A and 4-14C). At 48 hours this pattern changed slightly, with the basal respiration rate of 50% BCAA cultured cells being greater than that of full media cultured cells. However, maximal and spare respiratory capacity from 50% BCAA cultured cells recovered to be less than cells cultured in full media, but greater than when cultured in 25% BCAA media.

Combining the OCR data with the ECAR data (Figure 4-14A) suggests that BCAA restriction in KCL22s at 24 hours results in reduced oxidative and glycolytic respiration, however at 48 hours OXPHOS is most restricted. For K562s at 24 hours cells cultured in 50% BCAA media have a modestly increased OCR with reduced ECAR, while when cultured in 25% BCAA media a similar OCR but reduced ECAR was detected compared to full media. At 48 hours a similar trend is present, with K562s cultured in 50% BCAA media showing increased OCAR but decreased ECAR,

and 25% BCAA cultured cells showing reduced OCAR and ECAR. Together, this data suggests that K562s and KCL22s respond differently to BCAA depletion, with a reduction in OCR being more prevalent in KCL22s, while in K562s an increased ECAR compensates for the OCR decrease at 50% BCAA.

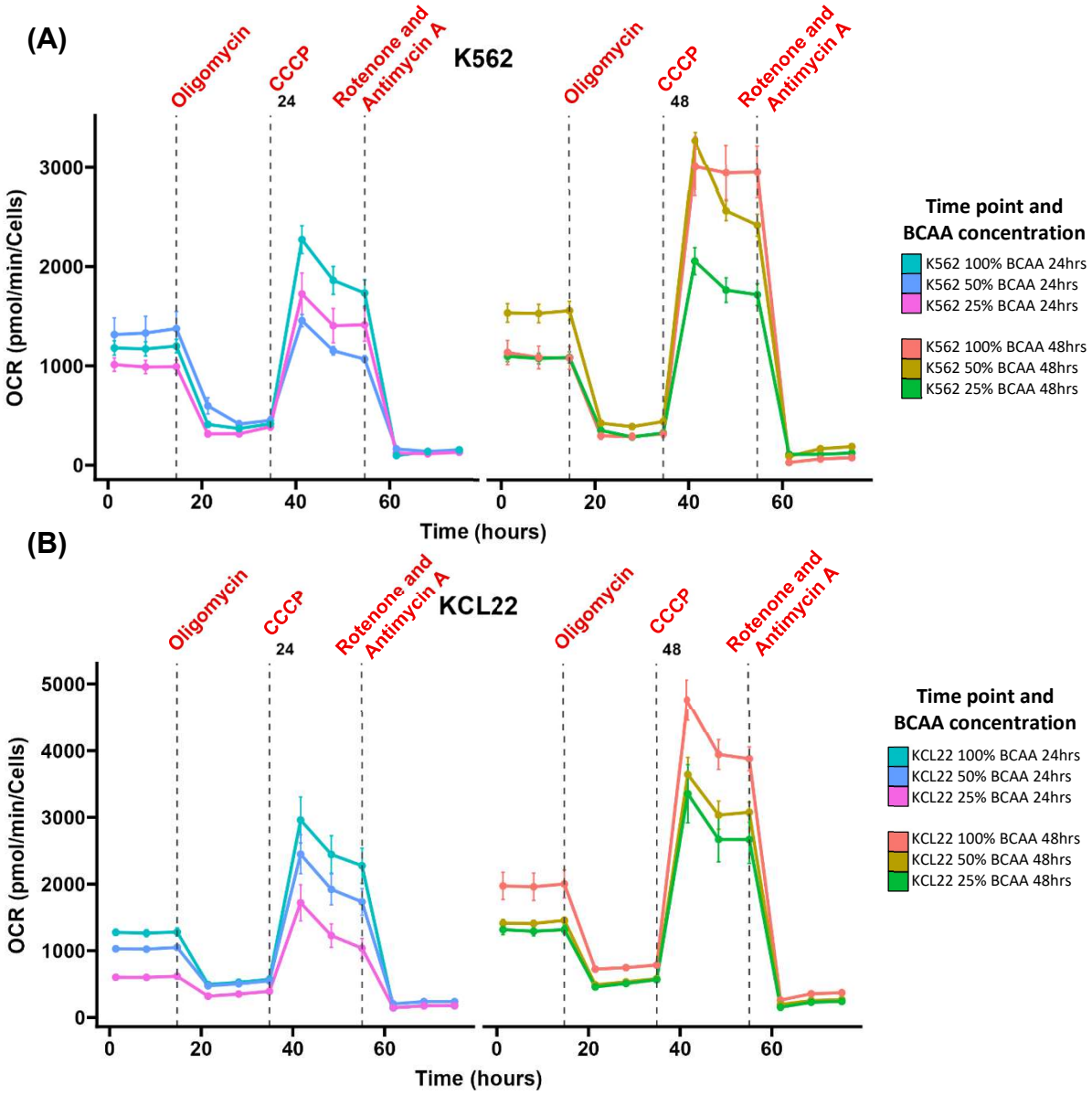


Figure 4-13 OCR trace of K562s and KCL22s cultured in BCAA deficient medias at 24 and 48 hours. OCR rates of (A) K562s or (B) KCL22s at 24 or 48 hours in full, 50% or 25% BCAA RPMI. 6 TRs, 1BR per condition. Red labels and lines represent the times at which the corresponding drugs were injected.

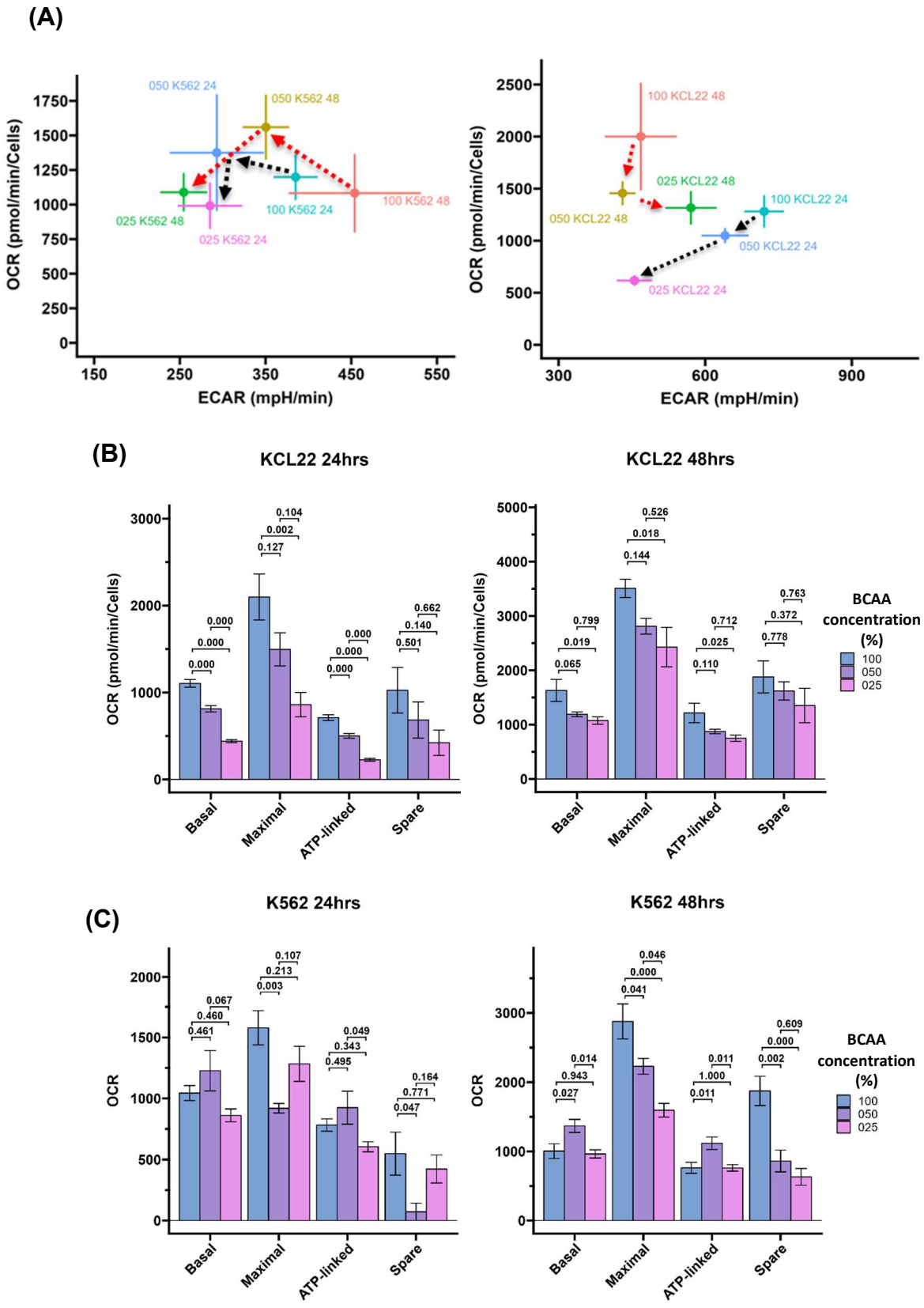


Figure 4-14 Seahorse assay of K562s and KCL22s cultured in BCAA deficient medias for 24 or 48 hours. (A) Energy map of basal respiration of K562s (left) or KCL22s (right) at 24 or 48 hours in BCAA deficient medias. Comparison of Seahorse derived metabolic parameters of (B) KCL22s or (C) K562s at 24 or 48 hours. Statistics calculated by ANOVA followed by Tukey test.

After quantifying changes in OXPHOS and glycolysis in K562s and KCL22s, we then wanted to better discern the metabolic pathways that may be underlying these changes. Therefore, we conducted a glucose tracing experiment using K562s and KCL22s, where cells were seeded at $3x10^5$ cells/ml, grown in 100%, 50% and 25% BCAA concentrations for 24 or 48 hours.

We saw that the BCAAs did not receive any labelling from glucose, which was unsurprising as BCAAs receive only an amine group during synthesis from BCKAs, and that BCAA levels were very similar between 24 and 48 hours for both cell lines in full media, despite high variation between other metabolites (Figure 4-15A/B and Figure 6-12). Examining PCA plots for this data (Figure 4-15C) shows that there was more variation between 24 and 48 hour cultured cells compared to the differences between media types within a day, aside from outlying samples. Cells cultured in BCAA deficient media had lower BCAA abundances, however a statistically significant effect (padj < 0.1) was observed only in 50% and 25% media in KCL22s at 48 hours, and in 25% BCAA media at 48 hours in K562s. This is in line with the growth curves and correlations, as at 24 hours the effect size and r^2 values are comparatively weaker. However, this contrasts the Seahorse data, as at least in KCL22s, the effect size was greater at 24 hours. This suggests that 48 hours is the minimum time required to detect a reduction in BCAAs in these cell lines, and that 25% BCAA media should be used if the aim is to reduce cell numbers, while 50% media should be used if the aim is to reduce metabolic capacity.

Conducting pathway analysis on these cells (Figure 6-14) showed only a few modified pathways (note that BCAA biosynthesis is detected given the significant reduction in BCAA concentration), with the majority having poor padj values. However, in K562s at 48 hours there was a reduction in TCA metabolites. For example, malate, citrate and αKG, and a reduction in glycolytic metabolites such as pyruvate and phosphoenolpyruvate, with glucose and glucose-6-phosphate identified non-significantly at a slightly lower abundance compared to control (Figure 4-16). Labelling fractions within glycolysis and the TCA were relatively unchanged, with the only substantial change in labelling coming from alanine (Figure 6-12). We also detected a reduction in AICAR, which is involved in *de-novo* purine synthesis, and can act as an AMPK activator. AMPK activation would be

expected to occur during times of stress to inhibit mTOR, preventing anabolic processes. However, the status of AMPK/mTOR in BCAA deficient conditions remains undefined.

Aside from the BCAAs, increases in creatine and carnitine were also detected. BCAAs have been shown to inhibit creatine kinase in rat brain cells cultured *in-vitro*, which could explain increased creatine levels where a reduction would be expected if cells are struggling to meet energy demands, if the same holds true in BM cells[216]. Increased carnitine levels suggest fatty acid transport to the mitochondria could be enhanced to meet energy demands, supported by unchanged abundances of acetyl-carnitine, indicating that increased carnitine levels are not being used to export acetyl groups out of the mitochondria. While outside of statistical significance, SAH levels were increased while SAM levels were modestly reduced, the reaction to convert SAH to SAM (part of the methionine cycle) provides the intermediates required for creatine synthesis. Aside from a reduction in BCAAs and an increase in creatine and carnitine, far fewer changes were detected in KCL22s in 25% media at 48 hours, which is somewhat representative of the Seahorse data.

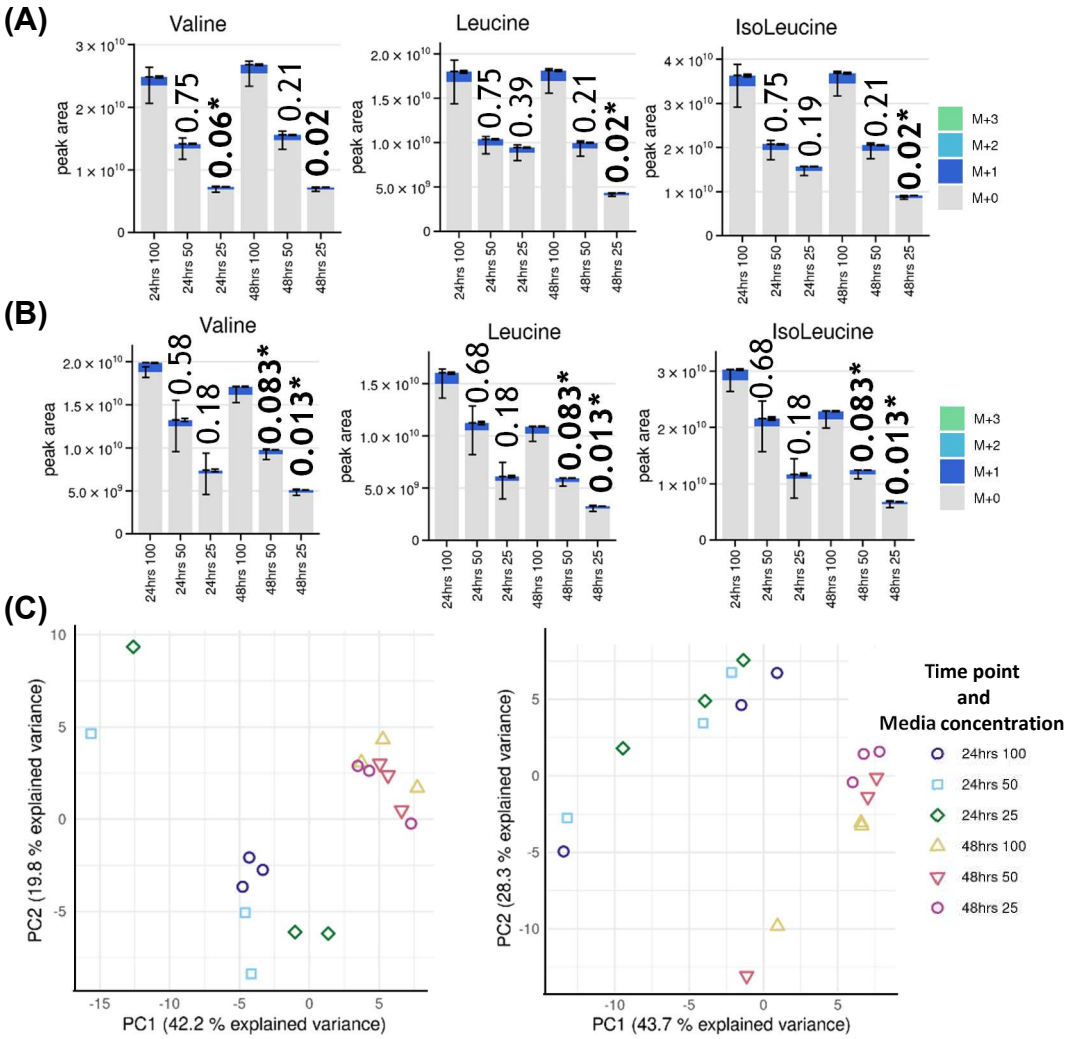


Figure 4-15 Glucose tracing of the BCAAs. Peak areas of intracellular BCAAs from (A) K562s or (B) KCL22s cultured in full, 50% or 25% BCAA media for 24 or 48 hours. 3BR, statistics calculated by Metaboanalyst t-test function with BH adjustment. (C) PCA plots of the metabolomic data from KCL22s (left) or K562s (right). Heatmaps of top features for full vs 50% and 25% comparisons shown in Figure 6-13.

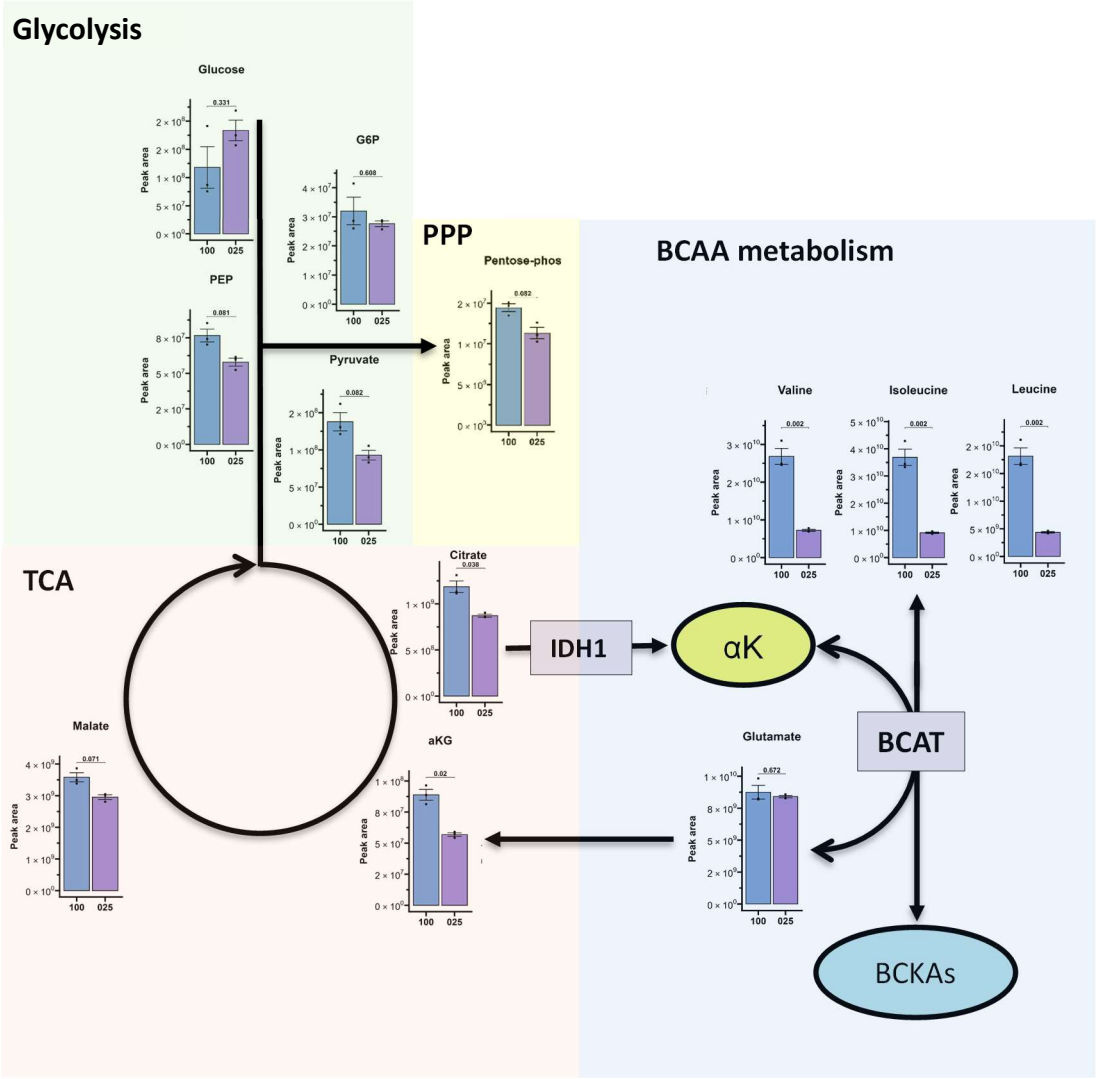


Figure 4-16 Pathway analysis of metabolomics data for K562s at 48 hours in 25% BCAA media. Differentially expressed metabolites involved in central carbon metabolism in K562s at 48 hours in 25% BCAA media.

4.4 HSC mobilization

4.4.1 Mobilized HSCs could be used for *ex-vivo* expansion

Haematopoietic stem cell transplant is a method which can be used to treat myeloid malignancies. However, finding a matching donor is challenging and the irradiation required for the procedure has multiple harmful effects. We therefore speculated whether it would be possible to mobilize HSCs to the blood of mice, sort for LSKs and expand them *ex-vivo*, then use BCAA depletion alongside re-transplant of the expanded blood cells to treat CML (Figure 4-17). To determine which HSC mobilization agents may be most efficient to use, we utilized AMD3100 and G-CSF, which are used clinically for this purpose. Additionally, we used P-selectin, a more experimental drug which has been shown to be able to mobilize murine HSCs[217].

We found that LSKs were nearly undetectable in a 100µl volume of untreated blood, with AMD3100 on its own, or with P-selectin only marginally increasing this number. AMD3100 and G-CSF together for 2 or 5 days had a far stronger effect, with 5 days of treatment mobilizing the most phenotypic LSKs. Bl6 mice were used for mobilisation, aside from G-CSF and AMD3100 for 5 days, which was instead an STG mouse, to facilitate transplant. We chose a volume of 100µl to analyse as this could be obtained from a single bleed, potentially allowing an autologous transplant experiment to be simulated in mouse models. From the 5-day AMD3100 G-CSF treated mouse, 100µl of mobilized bloods alongside $1x10^6$ CD45.1 competitor cells were transplanted into a Bl6 mouse, while LSKs from the remaining blood volume (400µl) were sorted for expansion culture (Figure 4-17C/D). As literature suggested that 5% oxygen (physoxia) was best for transplantation, we used a hypoxia chamber at this oxygen concentration, however no cells survived culture, and in a follow-up experiment using BM c-kit$^+$ cells, expansion also failed. The transplanted mobilized blood cells showed 25% PB and BM chimerism over the course of 16 weeks, displaying that the culture initiating blood volume contained functional stem cells. For numerous reasons, such as poor tolerance to a BCAA restricted diet in conjunction with CML (Figure 4-21), and time constraints to repeat HSC mobilization and expansion, the end goal of mobilizing, expanding and transplanting HSCs was not realised. However, as 100µl of blood from a 5-day treated AMD3100 and G-CSF mouse engrafted well,

we speculate that after expansion this starting volume and cell number is likely sufficient to singularly reconstitute conditioned BM.

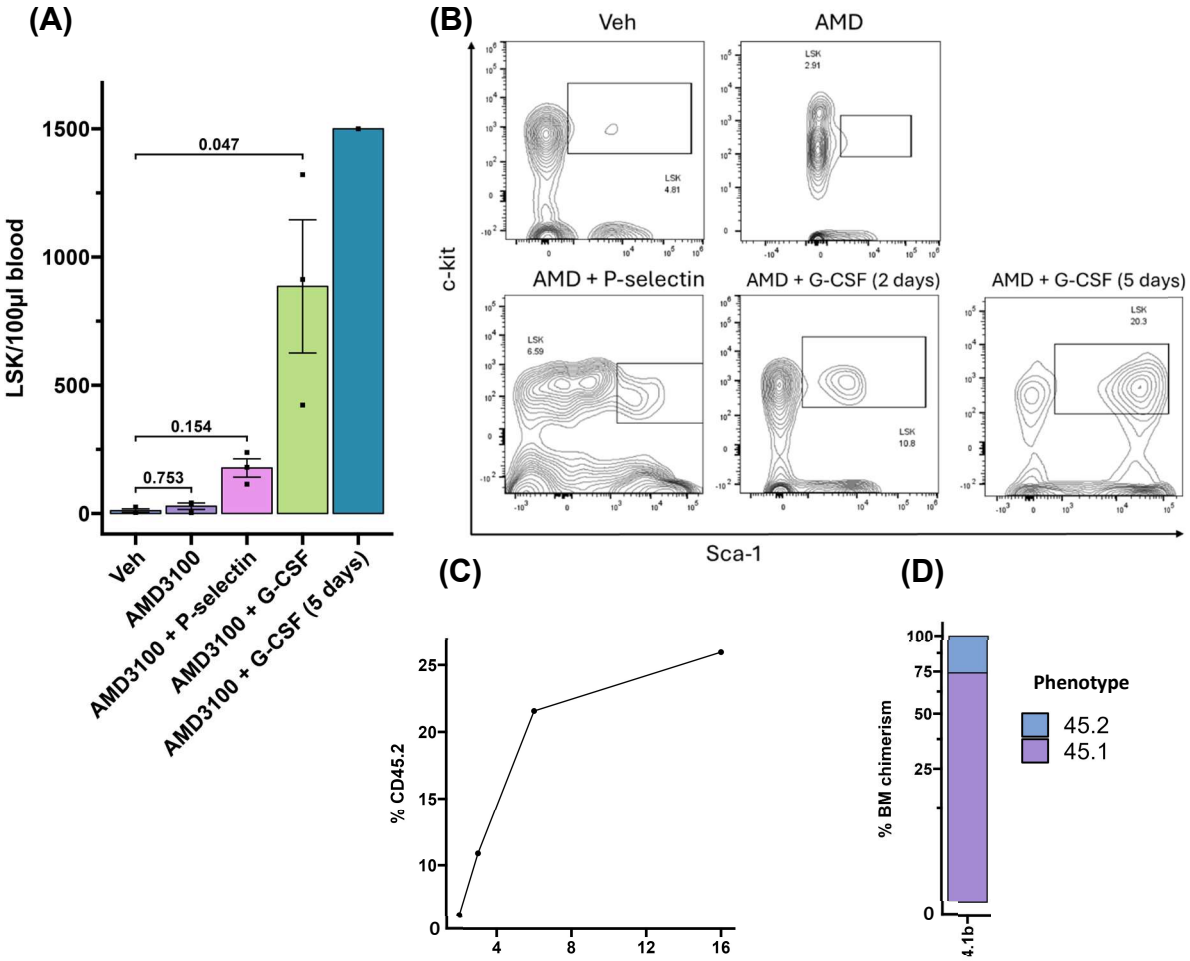


Figure 4-17 LSK prevalence in mobilised bloods. (A) Number of LSKs in 100µl of blood from mice treated with vehicle, and combinations of AMD3100, P-selectin and G-CSF. 3BRs, statistics calculated by KW test followed by Dunns test with BH adjustment (AMD3011 + G-CSF 5 days was excluded) (B) Flow cytometry plots of the LSK gates for these conditions. (C) PB and (D) BM chimerism over 16 weeks where 100µl of mobilised bloods from an STG mouse was transplanted into a Bl6 mouse, alongside $1x10^6$ CD45.1 competitors.

4.5 Dietary intervention to treat CML *in-vivo*

4.5.1 BCAA restriction causes weight loss and reduced haematopoietic potential in healthy NRGW and STG mice

As BCAA restriction proved effective *in-vitro*, we thought to test the effects of restricting BCAAs through diet *in-vivo*. Although these effects have previously been characterised in healthy mice, some parameters in these publications such as weight change have not been shown. This, the requirement for pilot studies using new treatments, alongside the capacity for variability seen *in-vivo* led to investigation into the effects of BCAA restriction in healthy STG and NRGW mice. To do so, 16 mice (10 STG and 6 NRGW) were fed either a control or $BCAA^-$ diet (diets were isocaloric and isonitrogenous), where weight and food intake was measured daily (Figure 4-18). Additionally, blood was taken the day before culling, which was used for IDEXX and LC-MS, while post-cull BM and BM supernatant was harvested for flow cytometry and LC-MS. Additionally, LC-MS was performed on control and BCAA- pellets to confirm that effects seen were due to BCAA restriction alone (Figure 4-19C/D and Figure 6-10C).

From examining weight changes, we observed that the BCAA diet caused steady weight loss over the course of 9 days in most animals, from both NRGW and STG groups (Figure 4-18B). This indicated that a long-term diet would not be feasible under UK project licences. On the control diet, most STG mice gained weight, however all NRGW and some STG mice suffered minor weight loss. An additional comparison to a set of mice on the diet provided by the mouse facility could have helped determine the reason for this. To prove that the weight loss in $BCAA^-$ diet mice was due to BCAA restriction rather than lack of appetite or mice not consuming diet due to preference, we measured food intake for control and BCAA diets over the course of the experiment (Figure 4-18C). From this, we found that there was no significant difference between rates of diet consumption, with day 4 appearing to be the day with the greatest difference. Sometimes mice would crumble large portions of diet, which was visible on the cage floor, datapoints where this was detected were omitted. We also detected a decrease in spleen weight between control and BCAA diets for STG and NRGW mice, with the exception of one mouse. When measuring differences in blood cells via IDEXX, we saw a reduction of monocytes and reticulocytes (Figure 4-18D/E). We saw

numerous effects of BCAA restriction in the BM (Figure 4-19A/B); a reduction in cells within the first gate representing haematopoietic cells, reduced $CD45^+$ cells, reduced Lin- cells in NRGW mice, and reduced LT-HSC numbers. Despite having lost substantially less weight than other mice in the STG BCAA group, mouse 45.1a still showed a similar reduction in its spleen weight proportional to body weight (spleen weight/body weight), alongside a similar reduction in haematopoietic cells. Examining LC-MS data of control vs BCAA pellets confirmed that BCAAs were absent (Figure 6-10C). However, some metabolites were upregulated, which was surprising. Analysis of LC-MS data from these mice showed no statistically significant changes from whole BM cells for either mouse strain, while in the plasma of STG and NRGW mice there was a significant decrease in valine (padj<=0.1), and the other BCAAs for NRGW mice (Figure 4-20A/B). In the BM environment of STG mice, widespread metabolic changes were detected, including 11 AAs (Figure 4-20C), while no statistically significant changes were detected in the NRGW BM environment. As dietary restriction had the capacity to reduce plasma BCAA concentrations, and BCAA concentrations in other tissues in under 10 days, we decided this was a sufficient effect size to test in conjunction with CML.

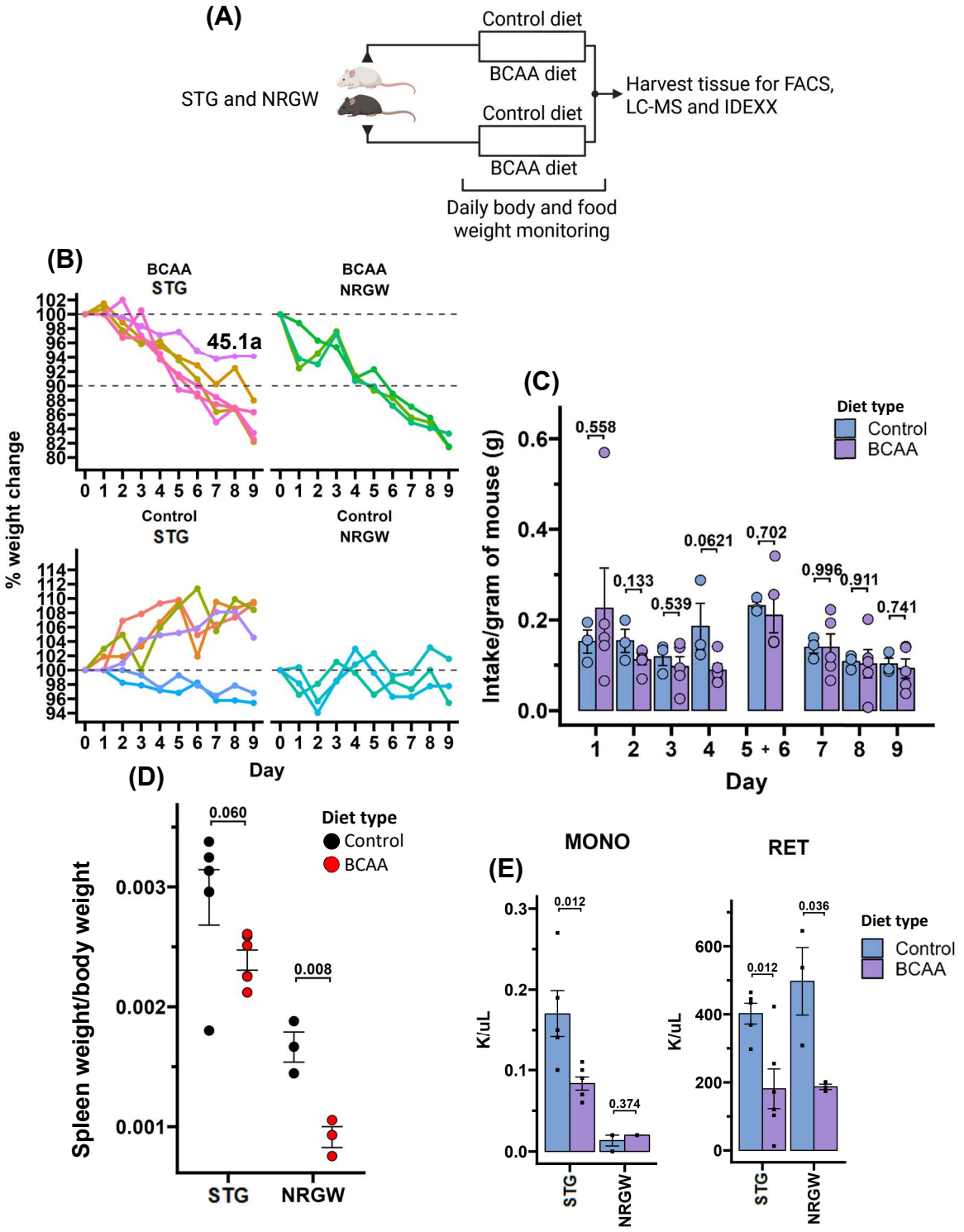


Figure 4-18 Weight, food intake and blood and spleen parameters of control or BCAA diet mice. (A) Schematic of experiment, STG and NRGW mice were put on either control or BCAA diet, where food intake and body weight were measured daily. Tissue was harvested for FACS, LC-MS or IDEXX. (B) Percentage weight change of control or BCAA- diet STG or NRGW mice. (C) Food intake normalised to mouse weight. Each datapoint represents a cage (D) Spleen weight as a proportion of body weight for STG and NRGW control or BCAA- diet mice. (E) Significant changes in the blood measured by IDEXX for STG and NRGW control or BCAA- diet mice. Statistics for (C-E) calculated using multiple t-tests.

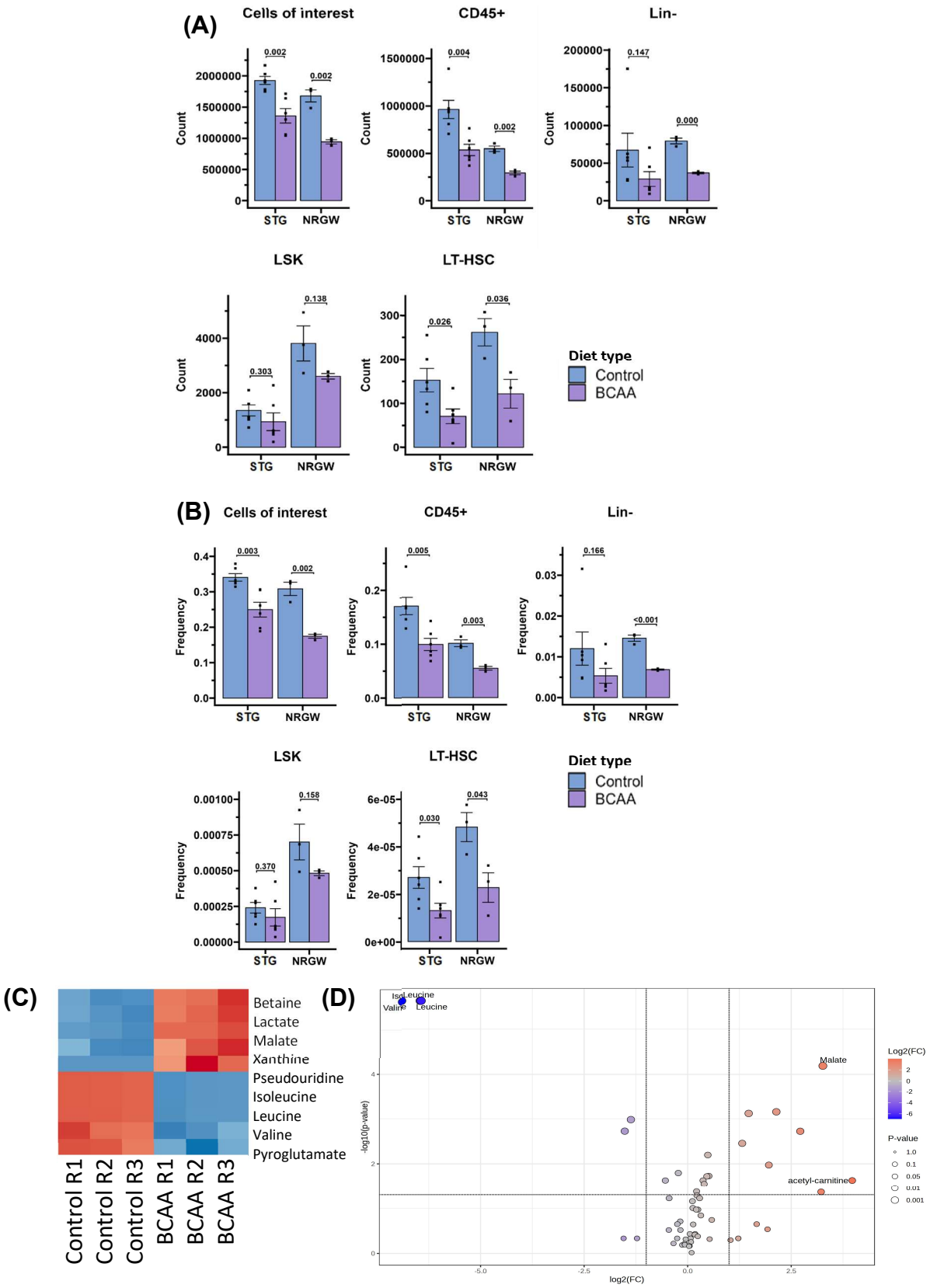


Figure 4-19 Flow cytometric parameters of control and BCAA mice, and LC-MS of control and BCAA diets. Number of (A) or frequency of (B) cells in different phenotypes/5 million cells. Cells within the first gate where haematopoietic cells can be identified via scatter, CD45$^+$ cells, Lin- cells, LSKs and LT-HSCs. 5 BRs for STG Control and BCAA groups, 3 BRs for NRGW control and BCAA groups. Statistics calculated using multiple t-tests. (C) Heatmap of the top DE metabolites between control and BCAA diets. (D) Volcano plot of DE metabolites between control and BCAA diets. Statistics calculated using ANOVA from Metaboanalyst.

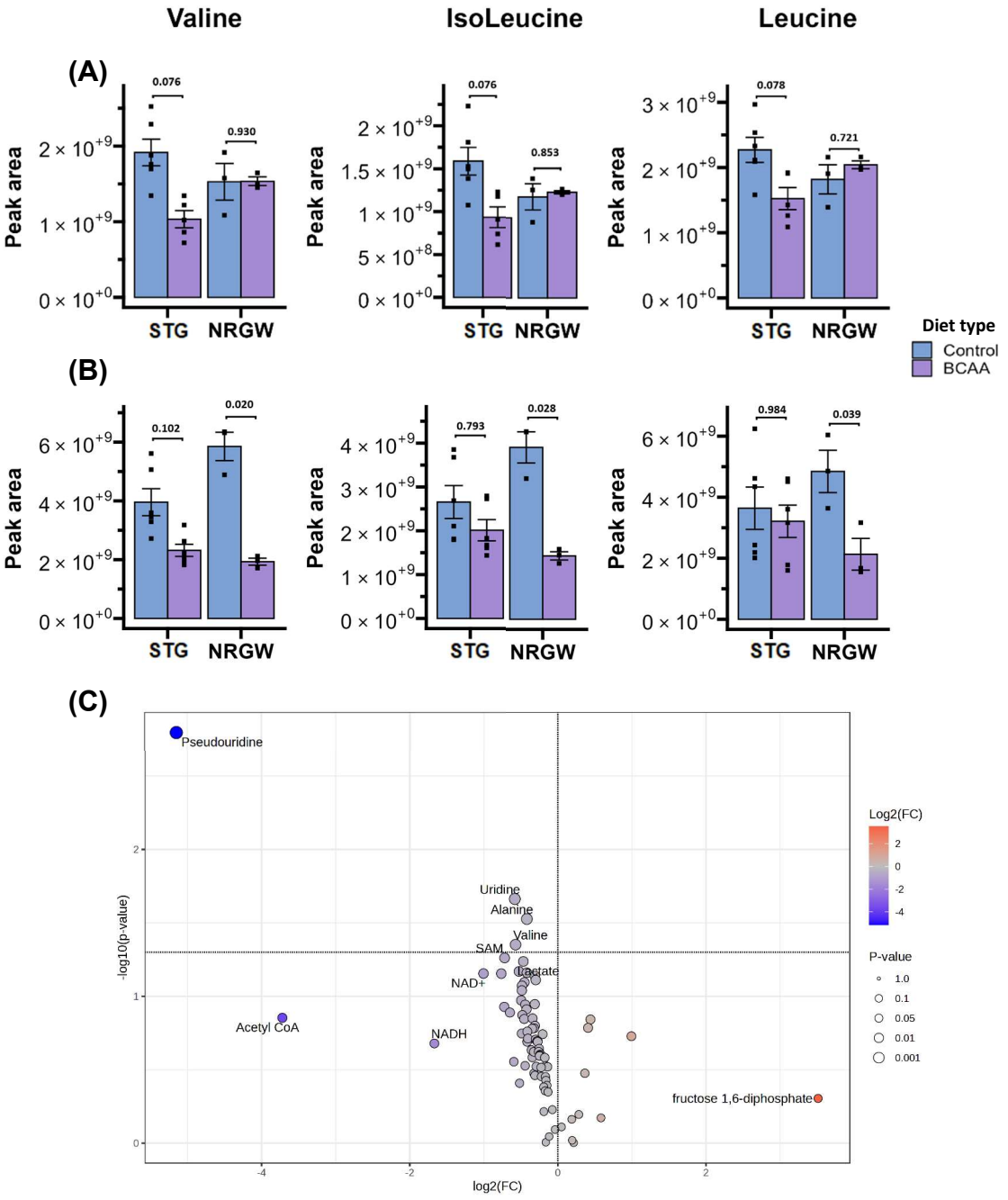


Figure 4-20 LC-MS data from STG and NRGW mice on control or BCAA$^-$ diets for 9 days. Peak area of the BCAAs from the (A) BM supernatant or (B) plasma, from mice on control or BCAA$^-$ diet. Statistics calculated by Metaboanalyst t-test function with BH adjustment. (C) Volcano plot of BM supernatant from B6 mice.

4.6 BCAA restriction alongside CML is poorly tolerated, with or without TKI

To determine whether BCAA restriction could be tolerated in leukaemic mice, and if it had an anti-leukaemic effect, 10 Bl6 mice (3 of which had to be culled prior to experiment start due to irradiation side effects) were transplanted with $1.5x10^6$ CD45.1 Bl6 and $1.5x10^6$ CD45.2 DTG whole BM cells. After recovery, leukaemia was induced, and one week later, mice were switched to either control or BCAA$^-$ diets. Mouse weights were measured, while blood, bone and spleen were harvested for IDEXX (data not shown) and flow cytometry (Figure 4-21 and Figure 4-22).

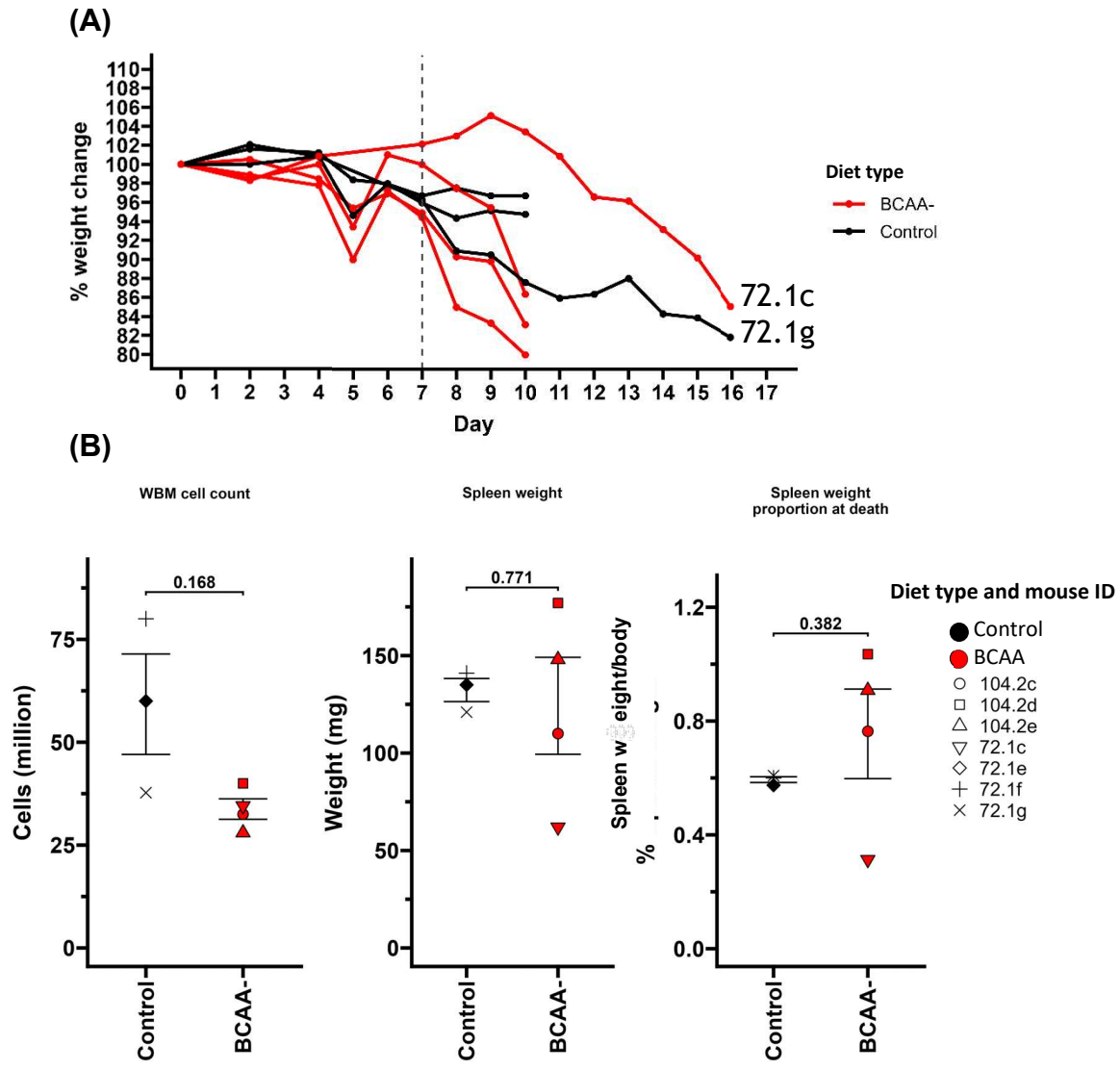


Figure 4-21 Change in physical parameters of induced DTG mice on control or BCAA$^-$ diet. (A) Mouse weight over time, day 0 represents time of induction, day 7 represents a switch to either control or BCAA diet. (B) Cell counts measured using cell counter, spleen weight and spleen weight as a % of body weight at death. 3 control diet mice and 4 BCAA$^-$ mice, statistics calculated using multiple t-tests.

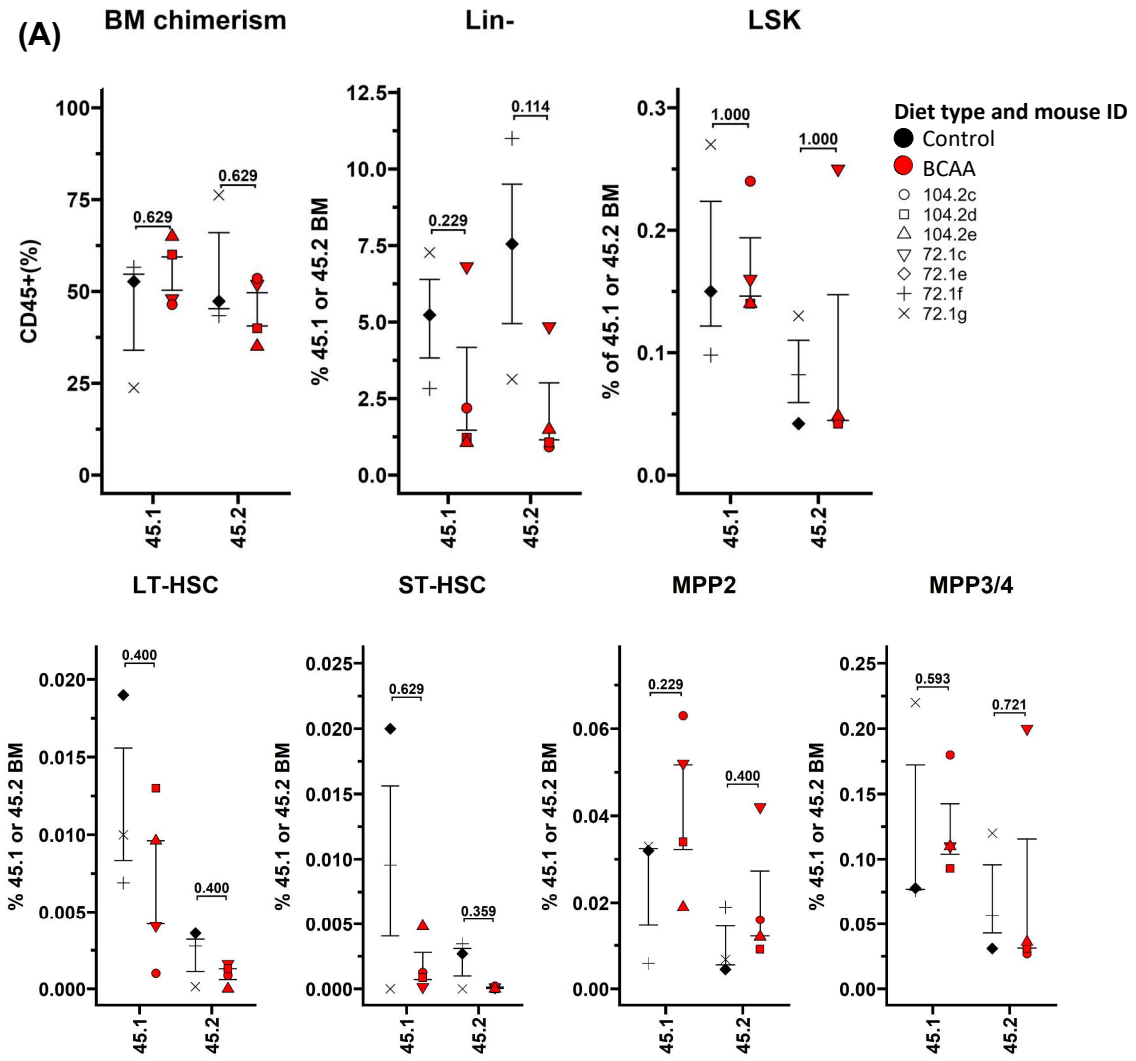


Figure 4-22 Cell frequencies of different phenotypes from the BM of CML mice on control or BCAA$^-$ diet. (A) BM chimerism, and frequency of stem and progenitor phenotypes as a proportion of CD45.1 or 45.2 cells. Statistics calculated using MW tests.

Examining weight change of control vs BCAA$^-$ diet mice (Figure 4-21) shows that in combination with leukaemia induction, BCAA restriction is not well tolerated, with one mouse (72.1c) showing a greater tolerance (9 days). Spleen weight after BCAA restriction was highly variable, with mouse 72.1c showing a substantially lower spleen weight than other BCAA$^-$ and control mice, but also retained a low cell count which CML alongside a BCAA$^-$ diet seems to cause. From the BM flow cytometry data (Figure 4-22) we saw no significant effects for either CD45.1 or CD45.2 cells, but examining 72.1c alongside 72.1g individually shows a reduction in CD45.2 BM chimerism. This, alongside the reduced spleen size suggests that mice that can tolerate CML alongside BCAA restriction have a reduced leukaemic

burden. However, to confirm this a much larger cohort would be required, and the outcome, even if positive, would not be practically useful.

We theorized that the damage BCR::ABL1 causes to non-leukaemic bystander cells in the BM may be responsible for the more rapid weight loss observed. To test if this could be mitigated with TKI treatment, and to see if BCAA and TKI treatment could produce a stronger effect size, Bl6 mice were transplanted with CD45.1 healthy and CD45.2 DTG cells, induced with leukaemia, treated with TKI then put on either a control or $BCAA^-$ diet (Figure 4-23 and Figure 4-24). We found that after approximately 1 week of TKI treatment, weight loss began to lessen, suggesting that BCR::ABL1 had been sufficiently inhibited, potentially allowing for dietary intervention without rapid weight loss. However, the $BCAA^-$ diet still caused extremely rapid weight loss, and so mice were culled after the third day (Figure 4-23B). Examining cell counts showed no significant difference between treatment groups, however there was a significant decrease in spleen weight as a proportion of body weight (Figure 4-23C). From the flow cytometry data (Figure 4-24), no significant differences between haematopoietic cell fractions were detected. It is possible that a longer duration of IM may have increased tolerance to BCAA restriction. However, previous data suggests that IM treatment does not completely prevent weight loss (either due to low levels of CML, or oral gavage or TKI side effects reducing appetite), which limits the window at which the experiment could be conducted. Alternately, a 2^{nd} or 3^{rd} generation TKI may have yielded better results, given their greater efficacy.

As BCAA tolerance was not improved by TKI, we aimed to determine if a short-term BCAA diet followed by return to control, then back to BCAA diet again could combat weight loss but also kill LSCs. To test this, mice were transplanted with healthy or leukaemic cells, induced, then put on control or $BCAA^-$ diet, then returned to control diet if/when weight change was reversed (Figure 6-16). This approach did not appear to work, however due to multiple mice suffering from irradiation complications in this cohort, this experiment would require to be repeated to confirm this.

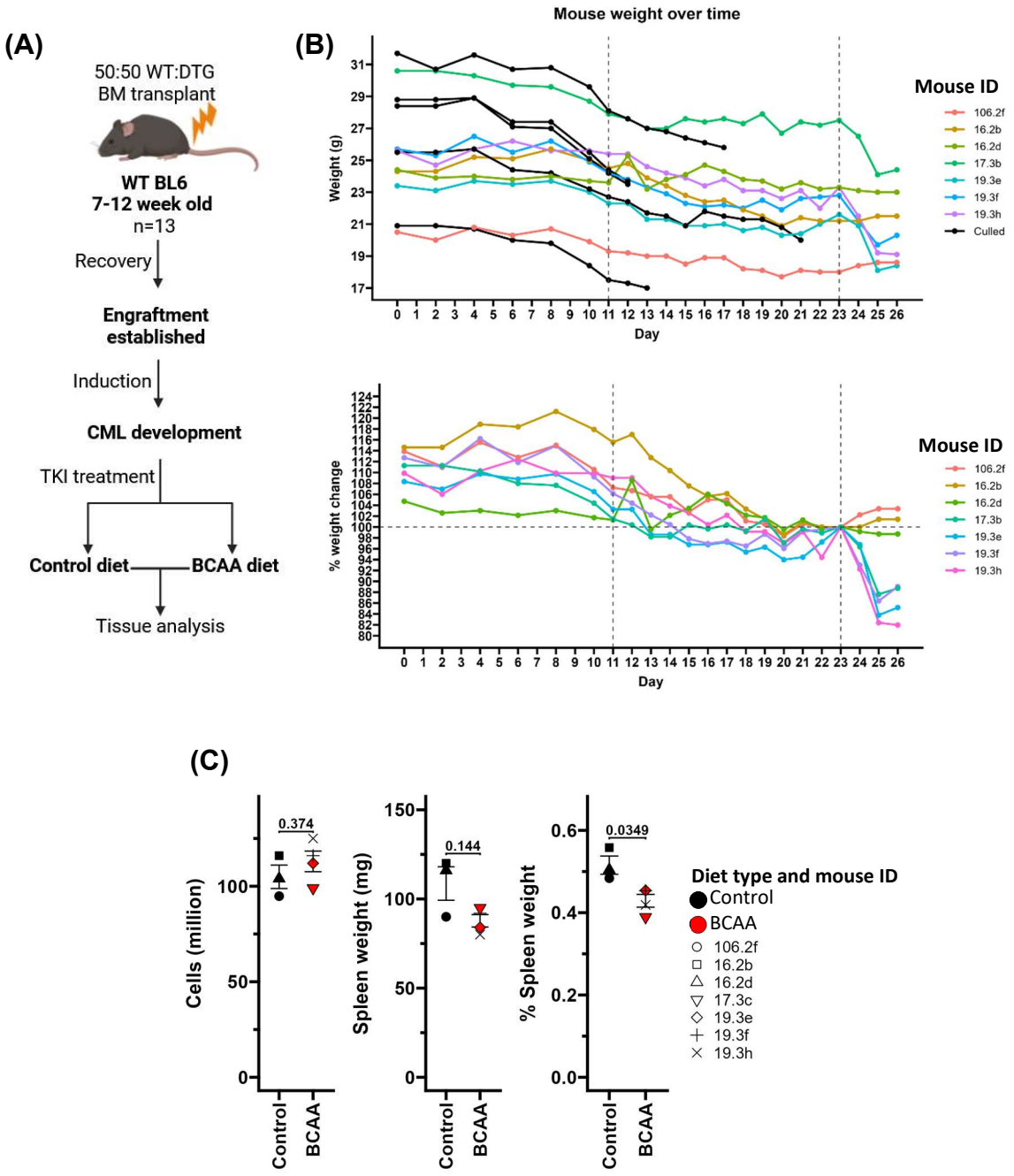


Figure 4-23 Interaction between TKI, CML and BCAA restriction. (A) Schematic of experimental plan. Mice were transplanted with BL6 and DTG cells, allowed to recover, and then induced for 1 week. TKI treatment was then administered, followed by either a control or BCAA- diet. (B) Absoloute (top) and relative (bottom) weight change. Black lines represent mice that were culled prior to dietary intervention. (C) BM cell counts, spleen weight and spleen weight as a proportion of body weight. Statistics calculated using MW tests.

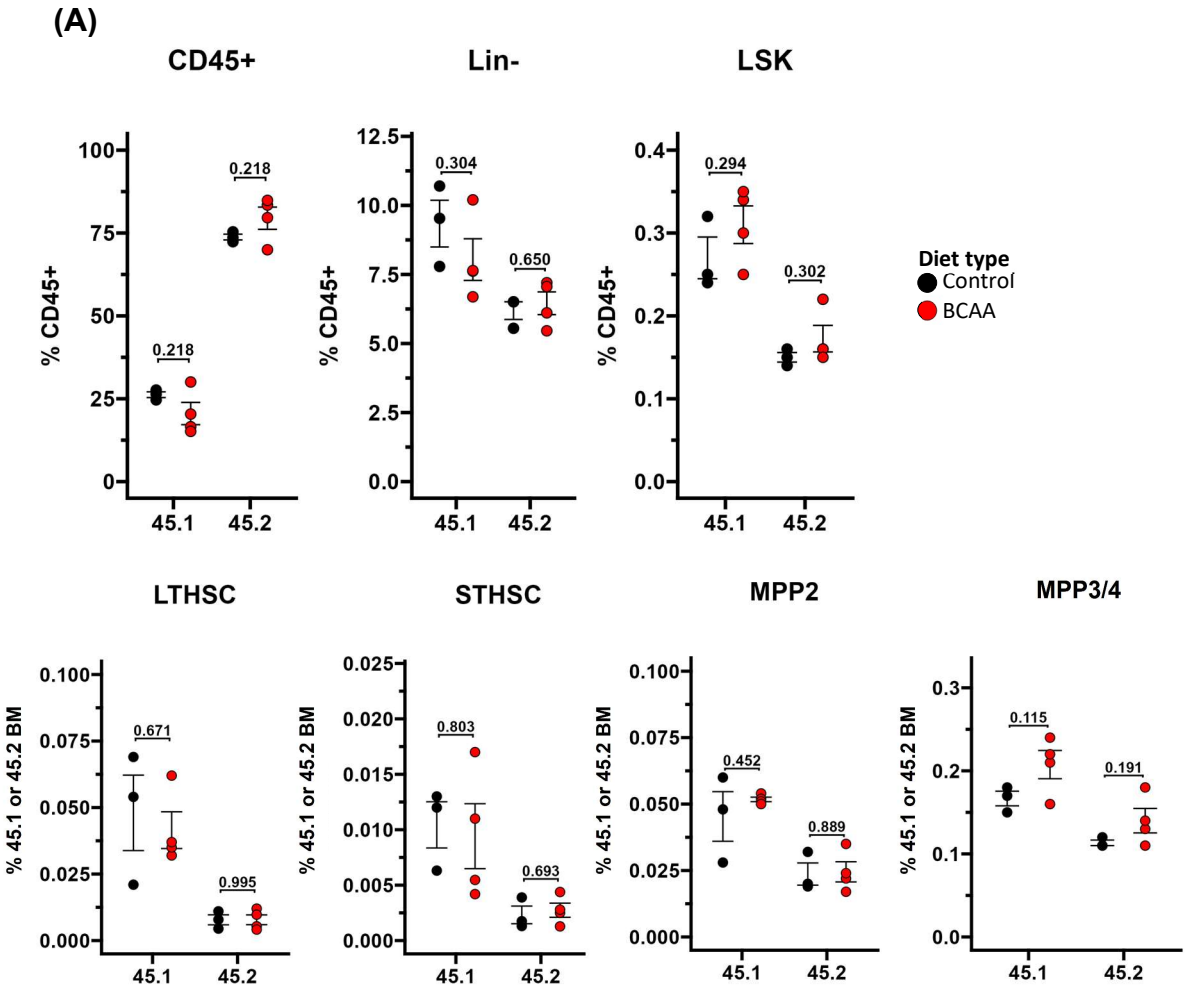


Figure 4-24 Flow cytometry data from CML mice treated with TKI on control or BCAA$^-$ diet. BM chimerism, and frequency of stem and progenitor phenotypes as a proportion of CD45.1 or 45.2 cells. Statistics calculated using MW tests.

4.6.1 Focus group with leukaemia patients to determine amenability to dietary interventions

As Doctor Stephen Wilkie and I were investigating the feasibility of dietary intervention to treat myeloid malignancies, we aimed to talk to individuals with myeloid malignancies (CLL, non-Hodgkin's lymphoma and CML) to determine if this was something that would interest patients. The aim of this was not to provide any scientific or mechanistic information regarding research conducted by the Helgason Lab, or any other lab to these individuals, but to understand their perspectives, wants, and feelings toward diet/nutrition and cancer. Two sessions were run, the first of which was in person in the Wolfson Wohl (6 attendees), while the second was online (8 attendees). We asked two questions, and constructively guided the discussion, questions being: How did your diagnosis and treatment affect your diet, and what would you want in a diet-based treatment? Due to ethics, we cannot provide granular information regarding individual responses to these questions, however the general consensus participants came to will be outlined.

In response to the first question, the consensus was that there was no dietary or lifestyle advice provided at the point of diagnosis, leaving them unsure about day-to-day lifestyle choices. Participants were aware of the dangers of taking dietary advice from social media influencers or individuals aiming to profiteer without any scientific proof. Unanimously, participants stated that they had altered their diets post-diagnosis, and had at least attempted to increase physical activity. This activity tended to be basic, such as walking and body weight exercises rather than weight training and long runs. Unanimously participants stated they felt making these changes was empowering and made them feel as if they could regain agency in a situation they felt was out of their control. All participant's stated additional stress during the period of providing blood samples prior to receiving the results, but said this did not specifically affect diet.

When addressing the second question, participants unanimously stated they would prefer dietary intervention to exist as an additional supplement rather than removal of certain elements from their food, or a complete replacement. Most participants said despite their diagnosis, they were often responsible for providing meals for their families, and would prefer any dietary intervention to not disrupt

this aspect of their life where possible. Some participants noted that food and mealtimes were an important part of connecting with or spending time with friends and family, and were concerned that they would miss out on bonding experiences, or possibly be singled out as someone who had cancer. Additionally, no individuals wanted to stay on an altered diet indefinitely. Conversely, participants stated they felt dietary intervention could provide the opportunity for friends and family to feel more involved in their treatment, which is not possible with other forms of treatment. No participant wanted meal replacement to be in the form of a shake, however they did state that if it was scientifically proven to be effective they would still accept dietary intervention in any form, if it offered the possibility of reduced drug or chemotherapeutic dose. Participants also wondered how dietary intervention would be handled when travelling, aside from stockpiling food.

In conclusion, all participants felt positive toward the idea of dietary intervention, and that if it was an option to take singularly or as a combination therapy that they would. The most prevalent concerns were the possibility of being isolated from friends and family, and having to cook for family/children alongside themselves.

4.7 Discussion

4.7.1 Expansion and hibernation culture

To investigate the utility of *ex-vivo* expansion and hibernation culture, we showed that LSCs could be expanded under normoxia or hypoxia, and from a sorted or enriched population. We also showed that both methods are capable of being used for screening LSCs, with varying degrees of success. Lastly, we aimed to create a non-irradiative model of transplantable CML, and tested whether the hibernation culture HSCs would engraft, both of which were mostly unsuccessful.

The main problem with LSC *ex-vivo* expansion using SCF and TPO is that SCF does not expand phenotypic LT-LSCs, meaning that the leukaemogenic potential of these cells cannot easily be measured. Additionally, from a preliminary experiment (Figure 4-3), reducing TPO concentration had very little effect on the cell count and phenotype. To gain greater insight into this, adding a culture condition with no TPO, measuring cell cycle and MPL expression, alongside a greater culture period could be next steps. Similar to Sakurai *et al*. who examined downstream signalling pathways activated by SCF and TPO in $CD34^{+}CD38^{-}$ cells, whether the same level of activation exists in expanded LSCs compared to expanded HSCs could be measured[131]. Subsequently, the efficacy of different agonists against these targets could be tested to more optimally expand LSCs. As previously mentioned, the long-term engrafting and leukaemogenic LSC is very rare in the DTG model, so it is possible that when sorting or enriching for LSCs to expand, these cells are not always present in each well, which is problematic when considering transplantation. When considering starting from a more pure population it has been separately shown that LSK-SLAM HSCs that are Sca-1^{Hi} and c-kit^{Lo} have greater self-renewal and functional potential, and that the leukaemogenic LSC population resides within the c-kit^{Lo} population[130,218,219]. Whether Sca-1^{Hi} LSCs are more leukaemogenic could be tested, with the possibility that Sca-1^{Hi}/c-kit^{Lo} LSCs could demarcate a population with greater leukaemogenic potential.

Given that we could not expand engraftable disease causing LSCs, we cultured DTG cells with DOX to suppress BCR::ABL1 expression. To properly validate this model a number of additional experiments need to be conducted, for example

detection of BCR::ABL1 using western blot or q-PCR alongside transplantation into irradiated hosts. Additionally, the effects of DOX requires more careful study, as it appears to influence at least phenotypic stemness, likely through altering metabolism, as has been shown in cell lines (Figure 4-4)[215]. For the non-conditioned engraftment, in both the normoxic and hypoxic culture functional stem cells were present. However, engraftment was far from consistent. This could be improved by sex-matching donors, isotype matching transplant cells, and the use of hypoxic culture. Bhattacharya *et al*. show that when 4000 CD45.1 HSCs (LSK-Thy1.1^{low}Flk2^{-}CD34^{-}) were transplanted into either CD45.2 or CD45.1XCD45.2 mice without irradiation, PB chimerism was only detected in the CD45.1XCD45.2 mice, and that this rejection of CD45.2 cells was due to host CD4^{+} T-cells[220]. Detecting donor cells with the same CD45 isotype as the acceptor could be achieved using tomato^{+} DTG or STG mice, making this possible improvement straightforward. If this method was proven to work it would have numerous applications, for example, phenotypically defining the leukaemia initiating cell. Uninduced LT-HSCs from DTG mice could be index sorted, expanded, DNA barcoded then transplanted. From there, the dominant clone that expands *in-vivo* could be mapped back to a specific phenotype, increasing the resolution at which LSCs could be identified. Alternately, genetic manipulation or CRISPR screens could be used for a targeted/untargeted approach to define which genes are essential for CML development. For example, knocking down a gene of interest followed by transplantation, or doing a large scale screen similar to Haney *et al*., where cells were expanded, barcoded, subject to CRISPR knockdown followed by transplant and analysis of which clones had the greatest/least repopulating capacity[221].

While we successfully cultured HSCs and LSCs under normoxia and hypoxia, we could not get cells to expand under physoxia, despite published literature[122]. This was likely due to the use of a hypoxia chamber rather than a hypoxia incubator, which Igarashi *et al*. used. Where the hypoxia chamber retains a steady oxygen concentration, incubators do not, due to the need to remove cells for media changes and due to other users opening the incubator. It could be that bursts of oxygen provided by media changes provide some level of signalling to survive and or divide. This could be measured using CellTrace and viability dyes to determine how many divisions occur directly after a media change where media has been

equilibrated or not, if cells cycle more rapidly after supplementation with equilibrated media, and if this has any strong effect on stemness and viability.

An additional aim was to see if HSCs from mobilized blood could be expanded *ex-vivo* and then transplanted, as a model of autologous stem cell transplant. This could then have been used in conjunction with BCAA restriction to both kill leukaemic cells and provide space for BM reconstitution with a highly pure HSC population (Figure 4-17). To better validate this model, these experiments would require to be repeated with more mice for both HSC mobilization and transplantation, alongside a transplant and sort including vehicle control blood. Additionally, sorted HSCs should be cultured in normoxia and hypoxia, so that the optimal conditions for the greatest level of engraftment from a set number of starting cells could be determined.

Exactly why hibernation HSCs showed no long-term engraftment was unclear. However, if this experiment was to be repeated, EPCR expression should be measured given its selectivity for functional stemness, alongside use of a cell sorter prior to transplant to ensure an exact cell count into the final volume used for transplantation. If this resolved the engraftment issue, hibernation culture of LSCs could have numerous applications. For example, studying the metabolic dependencies of CML LSCs in the actual leukaemogenic cell population, where experiments such as an AA dropout screen or knockdown of nutrient transporters, or metabolic genes, alongside transplantation could be utilized. Conducting large screens using this method would likely not be suitable, given the low starting number of cells which only reduces over time. Hibernation conditions could also be a good platform for studying exactly which cytokines and nutrients are required for LSC maintenance and expansion. Additionally, the effects of hypoxia could be added to this model, however this would likely result in even smaller cell numbers.

4.7.2 BCAA restriction

We showed that BCAA restriction had strong effects *in-vitro* on expanded HSCs and LSCs, and in K562s and KCL22s. In cell lines the metabolic effect was exerted through a reduction in both OXPHOS at 24 hours, and a reduction in glycolysis in K562s at 24 and 48 hours. BCAA restriction *in-vivo* was tolerated in healthy mice, but not in combination with leukaemia with or without TKI.

Despite the metabolic profiling of BCAA$^-$ cultured cells, it is still mechanistically unclear why BCAA restriction reduced cell numbers to the extent observed. We observed a decrease in citrate and αKG levels but no change in glutamine, with the contribution of BCAAs to the TCA not being quantifiable through glucose tracing. Ideally, tracing with each of the labelled BCAAs and glutamine would have been conducted, alongside use of BCKA/BCAA standards. We also attempted to rescue BCAA cultured cell lines with αKG supplementation, however this was unsuccessful (data not shown). Additionally, it could be informative to conduct a similar experiment measuring the contribution of autophagy to BCAA intermediates. It has been shown that labelling of BCKAs in the blood when using labelled BCAAs reached ~50% of that of the BCAAs within 60 seconds without changes in TCA metabolites, suggesting that recycling of proteins to produce BCAAs could possibly be better measured through a combination of BCKA and BCAA abundances. As discussed in section 1.5.2, it would be worthwhile to perform western blot or q-PCR of PRC2 components and molecules essential for nutrient sensing and metabolism, such as mTOR, and to examine rates of protein synthesis using single cell energetic metabolism by profiling translation inhibition (SCENITH). We profiled expanded cells and K562s using SCENITH (not shown), however our results showed high technical variation, with SCENITH results from Shaowei *et al* also looking to show high variability[96].

While the body condition and behaviour of mice on BCAA diet did not deviate from control diet healthy mice, the main difficulty with BCAA restriction *in-vivo* was rapid weight loss. From Figure 4-9 and Figure 4-12, it could be possible that a diet with partially reduced BCAA content could circumvent weight loss, while possibly leaving LSCs sensitized to other forms of metabolic or BCR::ABL1 inhibition. Due to cost constraints, purchasing multiple different diets with varying BCAA concentrations was not feasible. However, it might be possible to mix ratios of BCAA and control pellets to a desired concentration, however diet preference within single cages would need to be measured, and introduces an impractical degree of randomness. A more practical solution could be to purchase a liquid diet, and supplement that with the desired concentration of BCAAs, possibly allowing for a more economical solution. While liquid diets are less prevalent than solid foods, Sun *et al*. utilized a liquid diet to provide stable isotopes for tracing in a non-invasive way, which would be a desirable addition to any *in-vivo*

experiment aiming to investigate metabolic perturbation[222]. From the focus group that we ran, no individuals were enthusiastic about a shake-style meal replacement. Solid foods would therefore be our recommendation for dietary interventions in humans.

A shortcoming of the *in-vivo* experiments is that only phenotypic and quantitative readouts (e.g., weight, relative metabolite abundance, proportion of stem cells) were measured, not functional readouts. Therefore, conducting secondary transplants using healthy/CML and Control/BCAA diet treated mice could provide additional information. Aside from LC-MS, to assess the functional effects of BCAA restriction it could be possible to take blood samples and conduct flow cytometry using an mTOR/p-mTOR antibody, or use SCENITH to measure rates of protein translation. Lastly, while we were unable to find a practical use for complete BCAA restriction in CP murine models, given that Hattori *et al*. show that BCAT1 is required for BC transformation, dietary intervention could be used prophylactically in individuals with poor TKI response predicted to progress to the BP, providing success in murine models.

Chapter 5 Conclusions and future perspectives

5.1 Introduction

This thesis used existing transcriptomic datasets to identify common gene expression changes and metabolic weaknesses in CML LSCs. Additionally, we used HSC expansion culture to profile the metabolic weaknesses of LSCs. The following chapter explores the potential future directions of the key concepts presented by this thesis, and how this might improve our understanding of CML LSCs.

5.1.1 Phenotypes, xenografts and -omics

The most phenotypically relevant population that we aimed to profile was the Lin^- $CD34^+CD38^-CD90^+(CD26^+)$ compartment. Most CML -omics have profiled $CD34^+$ cells, where stem cell gene signatures are virtually undetectable. However, we have reached the point where continuing to sequence at the $CD34^+CD38^-$ level may not be providing as good a resolution as might be required. If we believe the leukaemogenic LSC to share a similar profile to the LT-HSC, from 1000 cells in the $CD34^+CD38^-$ fraction there may be ~2 LT-LSCs (1/617), from the $CD34^+CD38^-$ $CD90^+CD49f^-$ fraction there may be ~6 LT-LSCs (1/157), while at the $CD34^+CD38^-$ $CD90^+CD49f^+$ fraction there may be ~120 LT-LSCs (1/8.37), generously assuming similar stem cell frequencies in CML compared to normal HSCs[223–225]. If the functional LT-LSC contains features that demarcate it from other multipotent LSCs that require specific targeting, then sequencing at any other level that was not the $CD34^+CD38^-CD90^+CD49f^+$ (or $EPCR^+$) phenotype would not be informative, if aiming to answer this question. This is additionally complicated compared to profiling HSCs for a number of reasons: We might expect greater heterogeneity (more biological replicates required), as LT-LSCs from different response groups could have differing molecular profiles, and leukaemic markers require to be profiled, such as CD26, CD33, CD35, CD93 and IL1-RAP. Additionally, there are functional stem cells within the $CD90^-/CD49f^-$ populations, albeit at a much lower frequency than their positive counterparts, and possibly leukaemogenic potential within $CD26^-$ cells. It might be possible to conduct CITE-seq or other multiomics to dissect gene expression profiles within the $CD34^+CD38^-CD90^+CD49f^+$ compartment, however, until the functional and leukaemogenic potential of cells within/out with the CD26/CD90/CD49f/EPCR fractions have been measured, the

usefulness of single cell sequencing is limited to marker discovery. Based upon this uncertainty, it would be reasonable to assume that the next step should not be to sequence predicted primitive LSC populations, but instead to conduct rigorous transplantation studies on FACS purified populations. HSC characterisation has substantially outpaced that of CML LSCs, which at least provides a basis of where to begin. For example, the $Lin^-CD34^+CD38^-CD90^+CD45RA^-$ phenotype has been profiled in HSCs for over a decade, with recent improvements in HSC identification allowing for functional stem cells to be identified at a purity of ~1/3 ($CD34^+EPCR^+CD38^-CD45RA^-$), with subfractions of the $CD90^+/CD49f^+$ compartment also being profiled[226,224]. In LSCs, the most primitive cells transplanted have been $CD34^+CD38^-$[227]. Therefore, the most sensible way to improve our understanding of leukaemogenic LSCs would be to increase our functional understanding of CML LSCs near the top of the haematopoietic hierarchy. An example set of experiments could be to transplant cells with putative LSC markers (previously discussed) within the $CD90^+CD49F^{-/+}/EPCR^{-/+}$ fractions, in conjunction with BCR::ABL1 detection (e.g. FISH or digital PCR). After this, the most leukaemogenic/transplantable cells which yield $BCR::ABL1^+$ colonies once re-sorted from recipient mice could be considered as candidate LSCs. This could be followed up by any number of experiments, such as sequencing, TKI treatment and secondary transplantation. The technical difficulty and the amount of work this would require cannot be overstated, given the cell type rarity, difficulty of conducting large scale *in-vivo* experiments, range of techniques required, and the possibility for less stringent FACS gating to provide false positive results. However, without rigorous characterisation of stem LSC phenotypes, there is no way to definitively prove which LSCs we should be specifically targeting. The use of scRNA-seq to initially identify a population of interest most similar to HSCs, followed by transplantation of those cell types would clearly be less laborious. However, this works under the assumption that the leukaemogenic LSC actually has the same phenotype as the LT-HSC, which may not be true. Alternately, the haematopoietic hierarchy of LSCs close to the LT-HSC phenotype might not be organized in the same way as it is for normal HSCs. For example, cell types outside of a hypothetically predicted primitive LSC population identified through scRNA-seq may have the capacity to become functional stem cells under certain circumstances, similar to abnormal HSC behaviour when steady state haematopoiesis is disrupted. Though it would be

extremely challenging, to improve target identification from disease causing LSCs, more functional studies are required. It may also be worth considering whether the recipient mice used for PDX experiments still need additional optimization. For example, do CML LSCs poorly engraft and not produce disease because of an extremely low functional stem cell burden, or because the mouse BM microenvironment does not have the supporting cells/cytokine conditions required by LSCs. While this has been partially addressed through mouse models which produce human cytokines, something more extensive may be required, e.g. use of an ossicle niche[228]. Another consideration for why CML LSC engraftment is so poor, could be related to how cells are harvested/stored. For example, as discussed previously, when normal human and murine HSCs were harvested under hypoxic conditions, they retained a substantially larger proportion of functional stem cells[133]. We also know that LSCs are particularly sensitive to HIF degradation, so it stands to reason that the method by which LSCs are harvested is critically important[96,108,229]. While a hypoxic harvest is not practical in a clinical setting, Mantel *et al*. used a PPIF inhibitor while harvesting HSCs (as a way to reduce oxidative stress), resulting in a 4x increase in the number of LT-HSCs harvested[133]. Other drugs that reduce mitochondrial metabolism have also been shown to have similar effects, for example IACS treatment increased the stemness of *ex-vivo* expanded HSCs[122]. While this is speculation, it could have far reaching implications, and have the potential to vastly improve transplantation studies.

While characterisation of the CML metabolome is linked to the phenotypic stemness with which LSCs can be identified, in humans this remains comparatively less defined. The most stem human LSC population that has been metabolically profiled to date was by Rattigan *et al*., who utilized $CD34^+$/$CD34^+CD38^-$ cells[94]. Obtaining sufficient cell numbers is the main bottleneck for this process, however, a feasible next step could be to profile the difference between bystanders and BCR::$ABL1^+$ LSCs.

5.1.2 *Ex-vivo* expansion and CML mouse models

Previously, we showed the expansion of murine LSCs for a limited duration, and that LSCs could also be cultured under hibernation conditions. Additionally, we used *ex-vivo* expanded LSCs to screen for AAs which when absent from culture media, were capable of reducing LSC cell numbers.

Murine *ex-vivo* HSC culture can be conducted with relative ease compared to other stem cell types, with a high yield of functional HSCs in the short term. Additionally, the phenotype of functional *ex-vivo* expanded HSCs has been well defined in recent years ($EPCR^+$LSK-SLAM), with purity reaching as high as 66% from *ex-vivo* culture (33 of 50 single HSC clones PB engraftment > 1% at 12 weeks)[124]. This allows for greater stability when conducting transplantation studies, and when gene editing specific HSC clones. Contrastingly, murine LSC expansion is much less defined, with Patel *et al*. showing expanded LSCs to have poorer engraftment and leukaemogenic potential[129]. A current issue with the DTG mouse model is the inability to separate long term engrafting and disease-causing LT-LSCs from those which are not disease causing. Current understanding indicates such a cell should be LSK^{Lo}-SLAM, and most likely $EPCR^+$, although the latter remains undefined. If the aforementioned culture with DOX^+ expanded DTG cells was fully characterised, as discussed previously, this could be used to identify the phenotype of the leukaemia initiating cell. However, from discussion with the Bhatia lab, who is working on further optimization, they indicate that murine LSCs most likely require supporting niche cells (e.g. MSCs) for consistent and stable expansion. Whether it is worth devoting resources to the characterisation of murine LSCs at the most primitive level, rather than from humans is somewhat debatable. For example, while only two mouse datasets were shown in section 4.2.1, we processed an additional 4 bulk RNA-seq and 2 scRNA-seq mouse datasets in atleast the LSK phenotype. From these mouse datasets, we could not detect changes in most of the top scoring genes identified from the meta-analysis. Additionally, if a specific pre-leukaemic HSC clone is required for CML to occur/influences treatment outcomes, there is no guarantee that the equivalent clone with similar molecular features exists in the mouse models. This means that murine LSCs would be derived from murine HSCs with fundamentally different biology to their human counterparts, meaning that studies investigating primitive murine cells may not translate to human cells.

When considering human HSC expansion, while the capacity to expand $CD34^+CD38^-$ cells has improved, the cost of doing so compared to murine HSCs is greater. This seems somewhat contradictory as it would be expected that the cost of cytokines would be greater than that of chemically synthesized agonists, representing a substantial barrier to attempting these experiments using LSCs. If human LSCs

could be expanded *ex-vivo,* previously discussed experiments that would have been highly challenging or unfeasible due to restricted cell numbers would become easily possible.

5.1.3 The future of BCAA restriction to treat CML

From the meta-analysis, we showed that $CD34^{+}CD38^{-}CD90^{+}(CD26^{+})$ LSCs specifically had an upregulation of BCAT1, the first enzyme in the BCAA metabolism pathway. Additionally, and as previously shown by Wilkinson *et al.*, we showed BCAA restriction to cause the depletion of haematopoietic cells[49]. However, we also show that in both NRGW and STG mice, weight loss thresholds were approached after 9 days when on $BCAA^{-}$ diet, and when combined with CML, 3 days, even in conjunction with TKI. As discussed previously, literature shows that a BCAA restricted diet was tolerable in humans with no reported side effects[157]. Additionally, it was also tolerated with no side effects when used to treat AML xenografts, aside from minor weight loss (weight data not shown by publication)[148]. Given that CML PDX models do not cause aggressive disease, it is possible that we could have conducted a PDX experiment, and this may have been tolerated. However, using a precious resource for an experiment that was shown to be unsuccessful using murine cells did not seem like a reasonable allocation of resources. The future of BCAA restriction/BCAT inhibition for CML will likely be $BC/BCR::ABL1^{T315I}$ focused, given that BCAT activity appears to be required for BC transformation[147]. If research into BCAAs would continue, we would recommend not to use a full dropout, as in both cell lines and expanded cells a 50% reduction over 48 hours was sufficient to cause a metabolic depression. This could possibly be followed up with subsequent TKI treatment, PDX, or secondary transplant to better characterise the potential anti-leukaemic effects of BCAA restriction.

Chapter 6 Appendix

6.1 Appendix index

Figure 6-1 and Figure 6-2 show volcano plots for $CD34^+$, $CD34^+CD38^-$ bulk and scRNA-seq datasets, and predicted primitive phenotypes, providing an overview of the number of DE genes and the distribution of log2FC and padj values.

Figure 6-3 displays the distribution of stem markers and CD26 across Seurat clusters and patients for GSE184. Figure 6-4 shows the distribution of these markers, and other markers overlaid onto UMAPs for GSE233 and GSE184.

Figure 6-5 shows the number of common genes between datasets part of the same comparison sets, the percentage of BCR::$ABL1^+$ positive cells on a dataset and patient-wise basis, alongside the number of differentially expressed metabolic genes as a proportion of the number of up and downregulated genes for a dataset, for all comparison sets.

Figure 6-6 shows the expression of top scoring genes overlaid onto UMAPs for GSE233 and GSE184.

Figure 6-7, Figure 6-8 and Figure 6-9 display genes informative to differentiating different responder groups from GSE312, GSE233 and EGA, at the $CD34^+CD38^-$ and predicted primitive level.

Figure 6-10 shows LC-MS of relative AA concentrations from the different medias formulated, and BCAA/control mouse diets.

Figure 6-11 shows biological replicates for the Seahorse experiments.

Figure 6-12, Figure 6-13, and Figure 6-14 show aspects of the LC-MS experiment investigating metabolic changes in K562 and KCL22 culture in BCAA deficient media. Displaying relative labelling fractions of the BCAAs, top metabolites for comparisons, and metabolite set enrichment analysis.

Figure 6-15 shows metabolite set enrichment analysis results for tissues taken from BCAA or control diet STG or NRGW mice.

Figure 6-16 shows mouse weights for DTG or STG transplanted mice on control or $BCAA^-$ diets, where we aimed to see if sequential diets could be a feasible option.

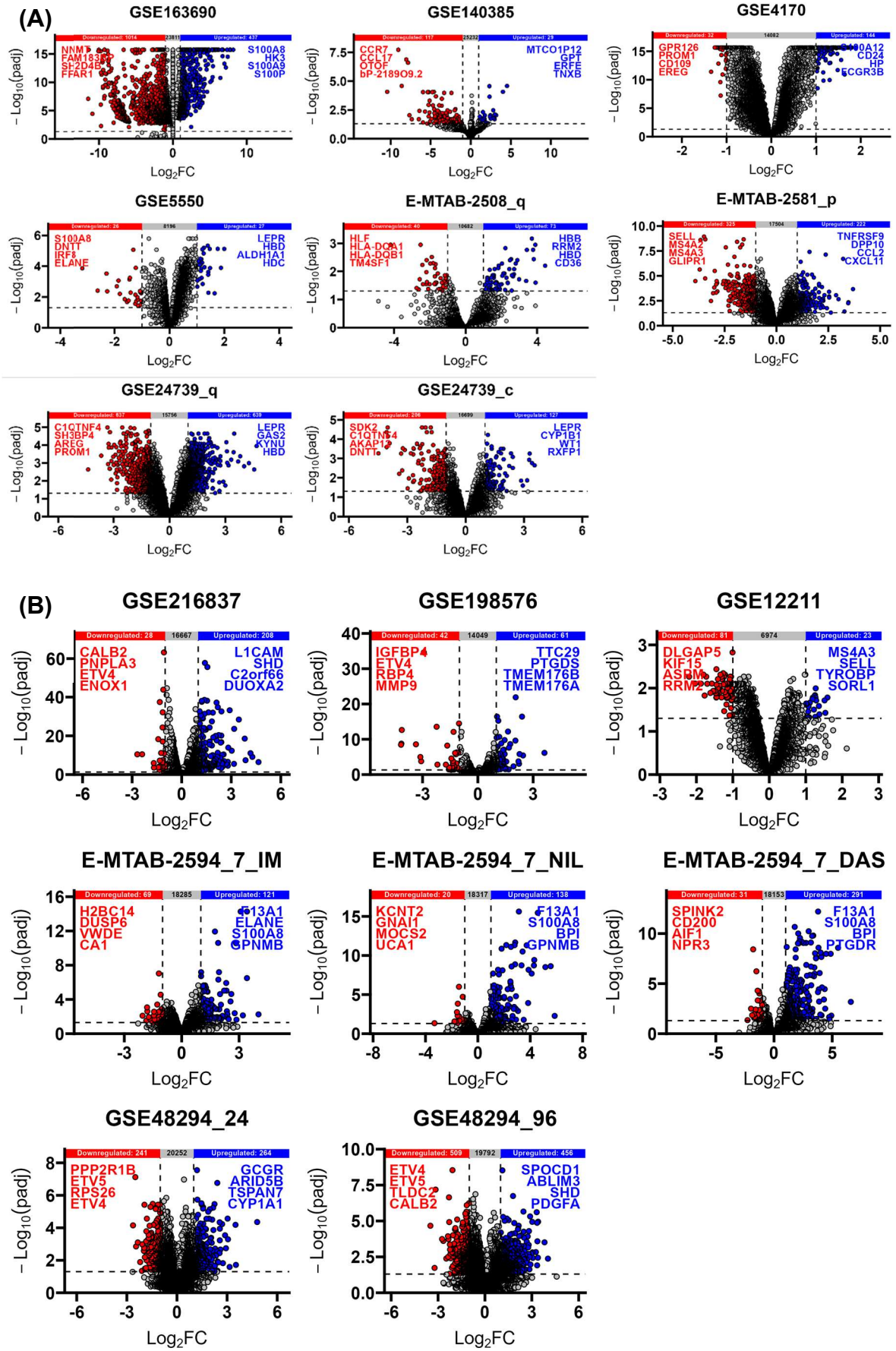


Figure 6-1 Gene expression profiles for CD34⁺ and TKI treated comparison sets. (A) Volcano plots displaying gene expression profiles and numbers of up and downregulated genes (padj < 0.05 and log2FC >1) for (A) CD34⁺ datasets and (B) TKI treated datasets.

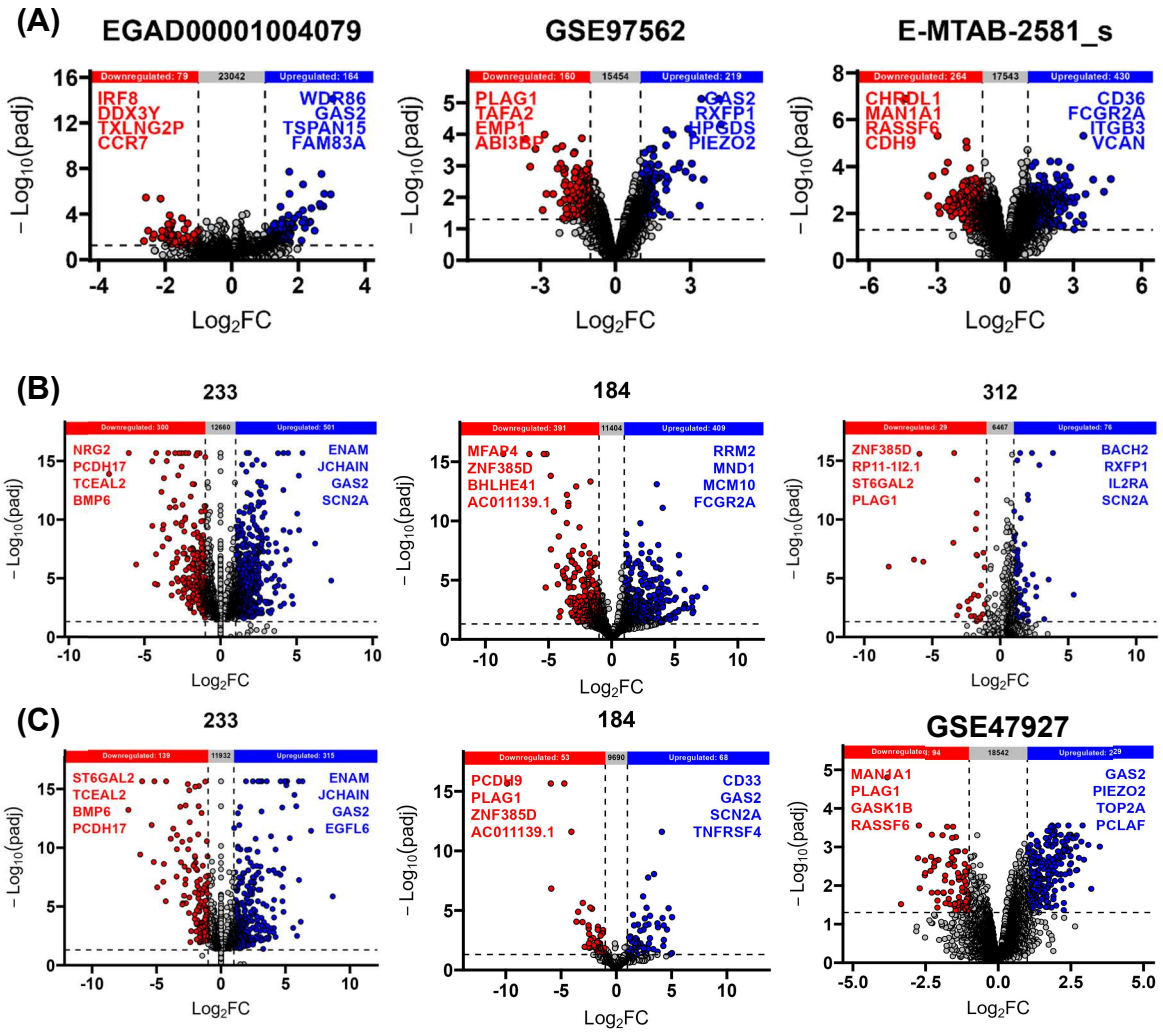


Figure 6-2 Gene expression profiles for CD34⁺CD38⁻ and predicted primitive comparison sets. (A) Volcano plots displaying gene expression profiles and numbers of up and downregulated genes (padj < 0.05 and log2FC >1) for (A) the CD34⁺CD38⁻ bulk RNA-seq comparison set (B) scRNA-seq CD34⁺CD38⁻ comparison set and (C) the predicted primitive comparison set.

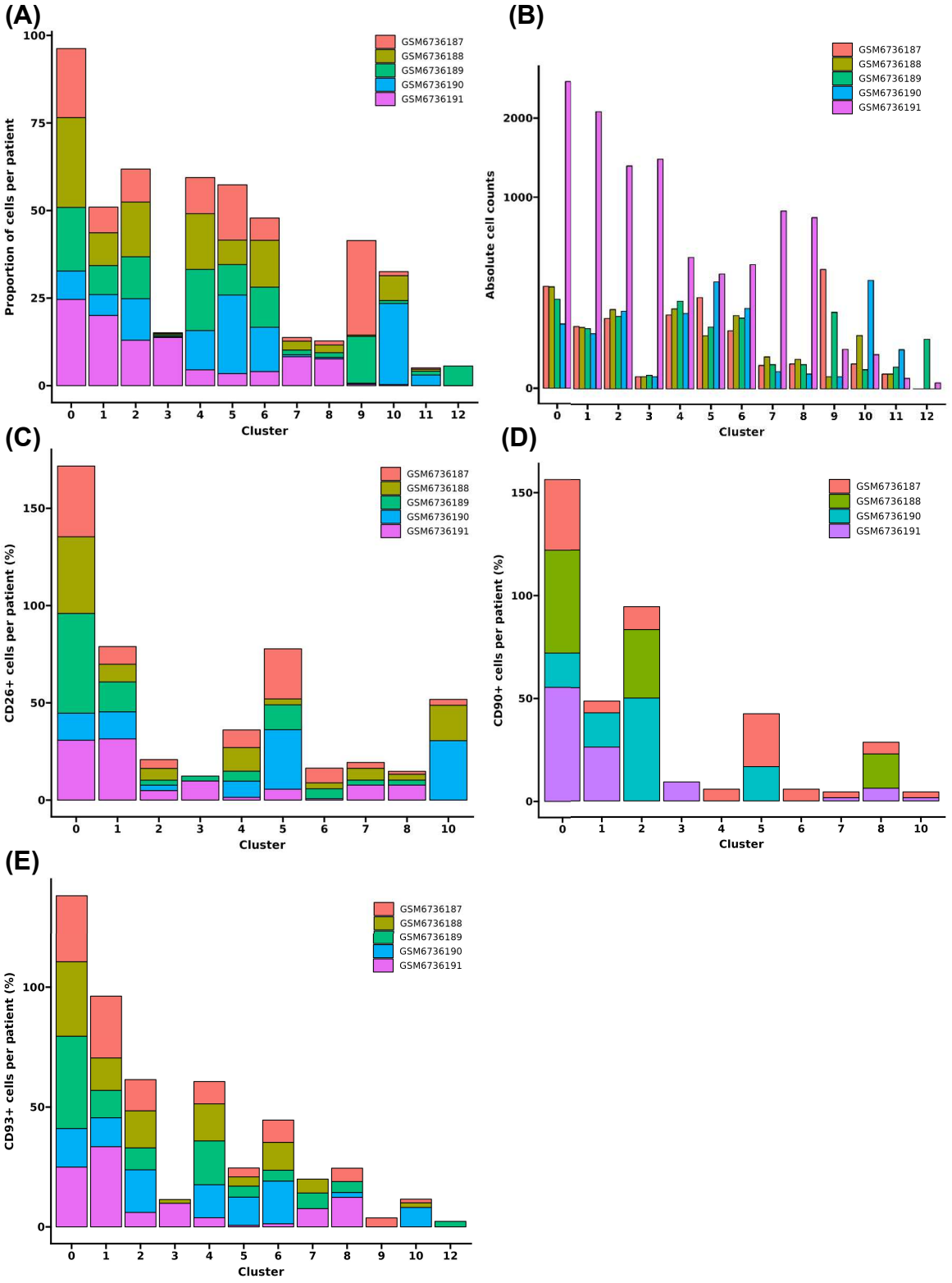


Figure 6-3 Patient contribution to clusters and patient contribution to marker positivity in clusters. (A) Contribution of each patient to each cluster as a percentage of an individual's cell number. (B) Number of cells in each cluster per patient. Percentage of (C) $CD26^{+}$ (D) $CD90^{+}$ or (E) $CD93^{+}$ cells on a per patient and per cluster basis.

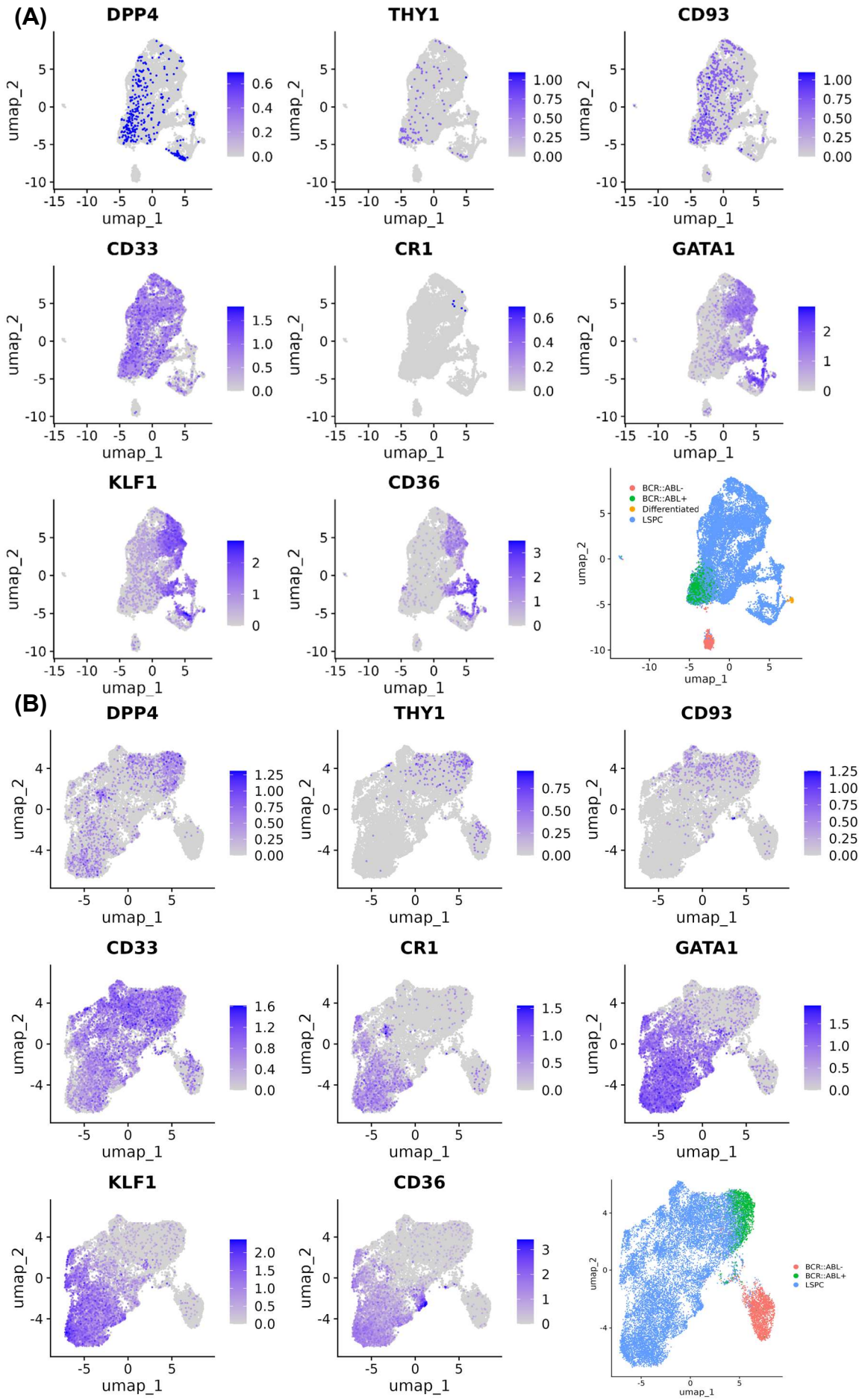


Figure 6-4 Stem genes and erythroid genes of interest for GSE233 and GSE184. Stem markers and genes of interest for (A) GSE184 and (B) GSE233.

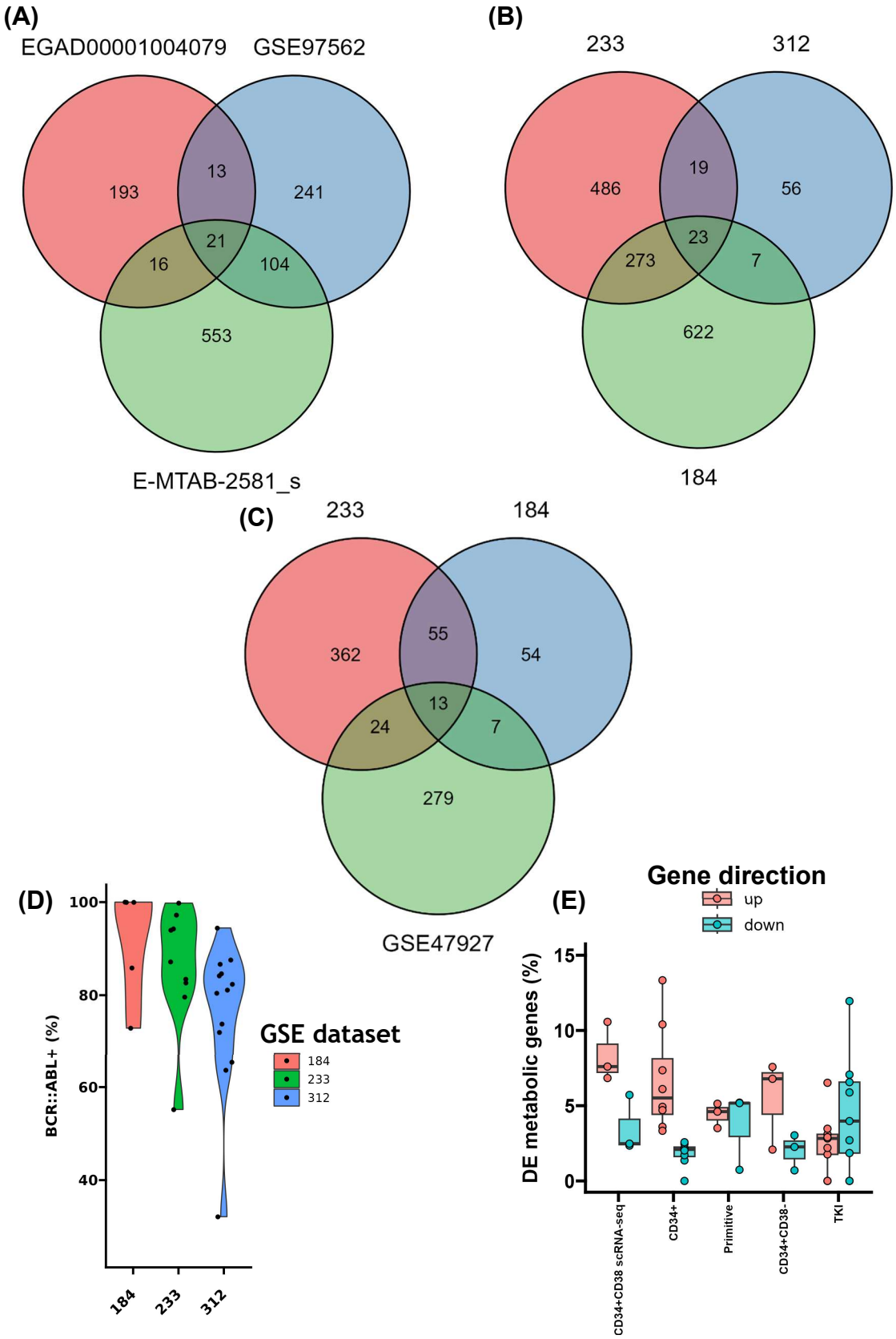


Figure 6-5 Overlap of DE genes within comparison sets, and various metrics. Overlap of differentially expressed genes (padj < 0.05 and log2FC > 1) between (A) $CD34^{+}CD38^{-}$ bulk (B) $CD34^{+}CD38^{-}$ scRNA-seq and (C) Predicted primitive comparison sets. (D) Percent $BCR::ABL1^{+}$ cells on a patient wise basis for GSE184, GSE233 and GSE312. (E) Percentage of metabolic genes changing as a percentage of upregulated and downregulated genes for each dataset in a comparison set.

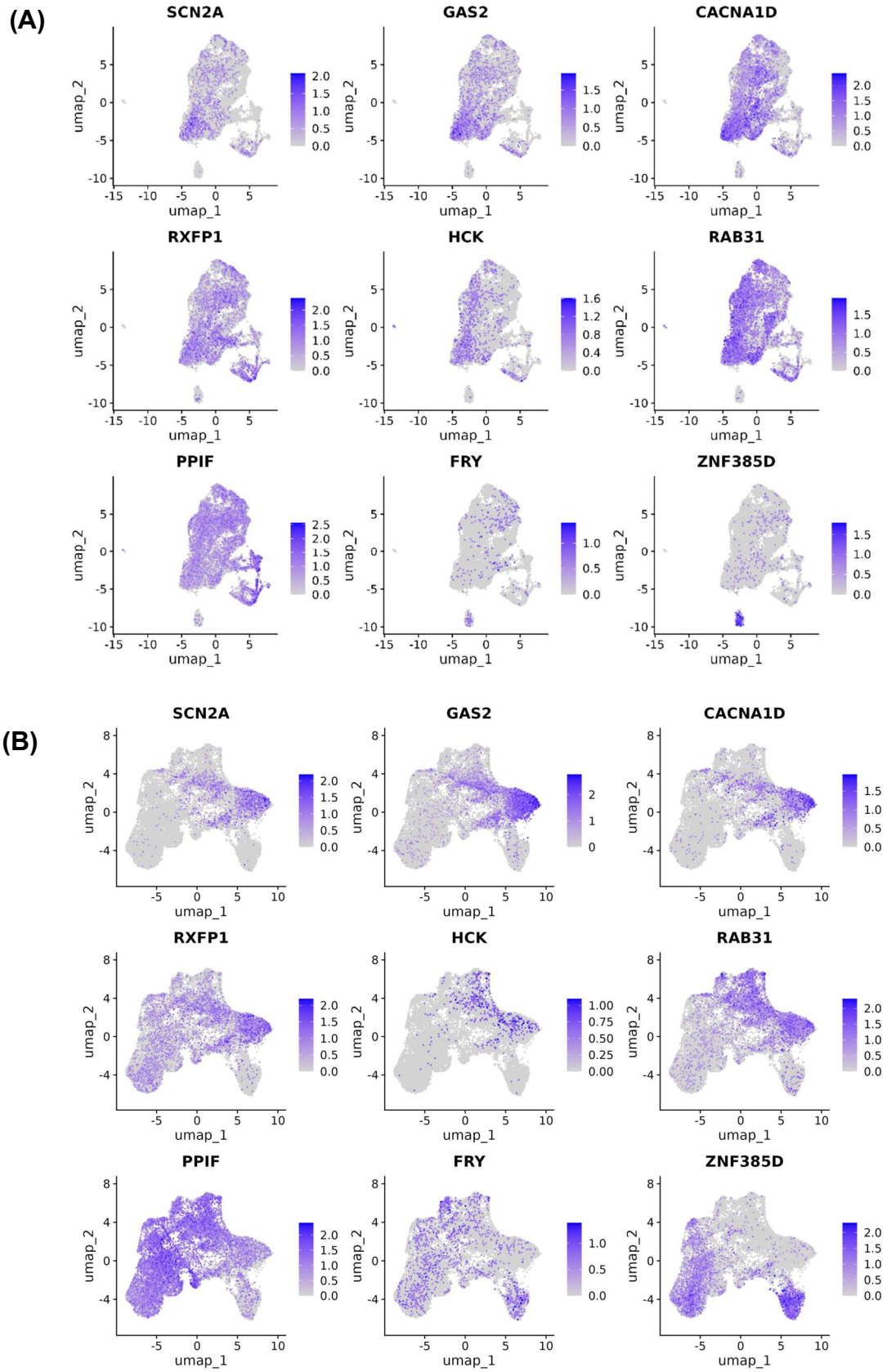


Figure 6-6 UMAPs of expression levels of top scoring genes for predicted primitive LSCs vs bystanders. UMAPs of expression levels for top scoring genes for predicted primitive LSCs vs bystanders from (A) GSE184 and (B) GSE233.

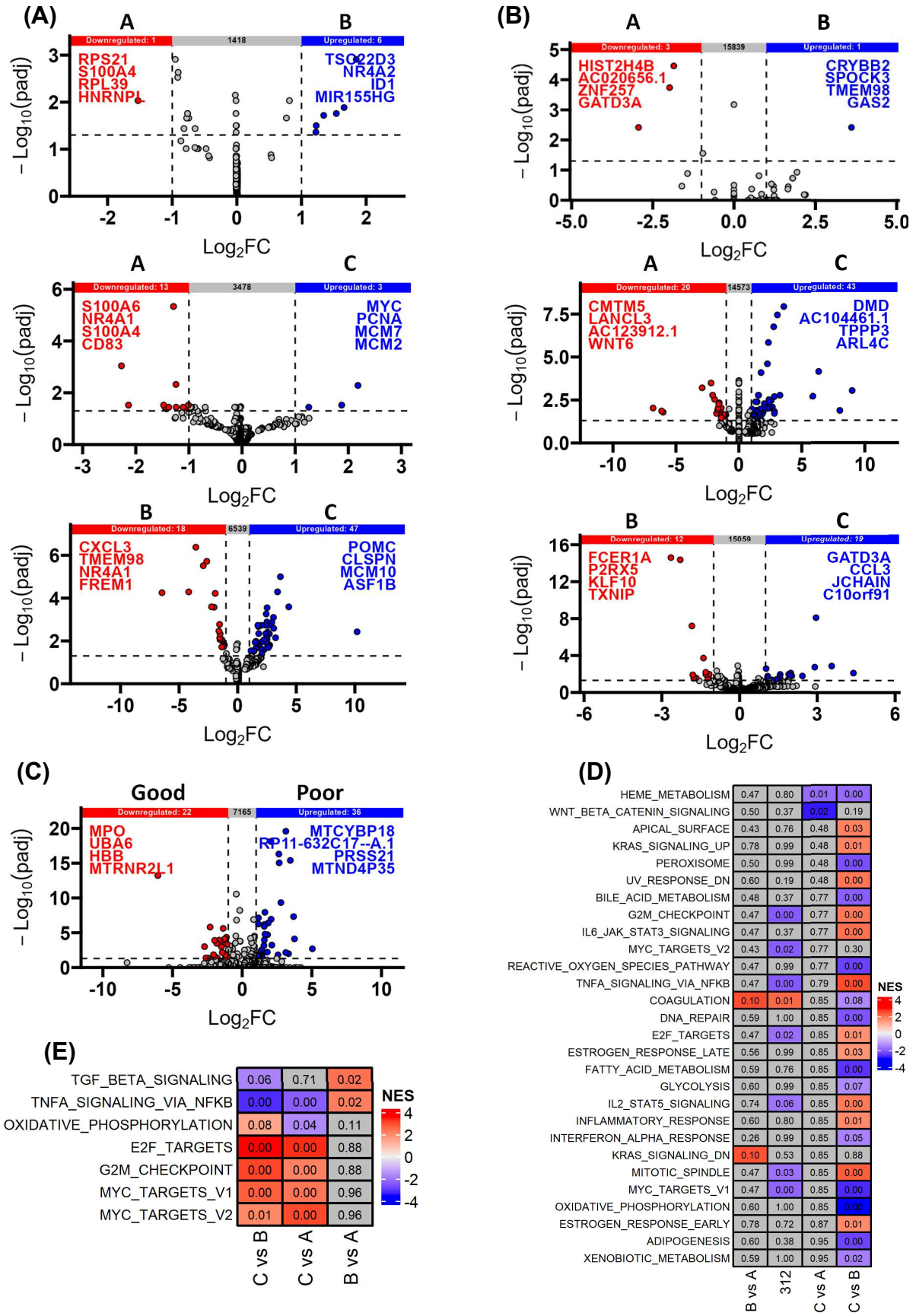


Figure 6-7 Differences in gene expression between groups A, B and C CD34^{+}CD38^{-} LSCs. (A) Volcano plots showing differential expression between groups A,B and C from (A) EGA and (B) GSE233. (C) Difference in gene expression between poor and good responders from GSE312. GSEA results for corresponding comparisons from (D) EGA and GSE312 and (E) GSE233.

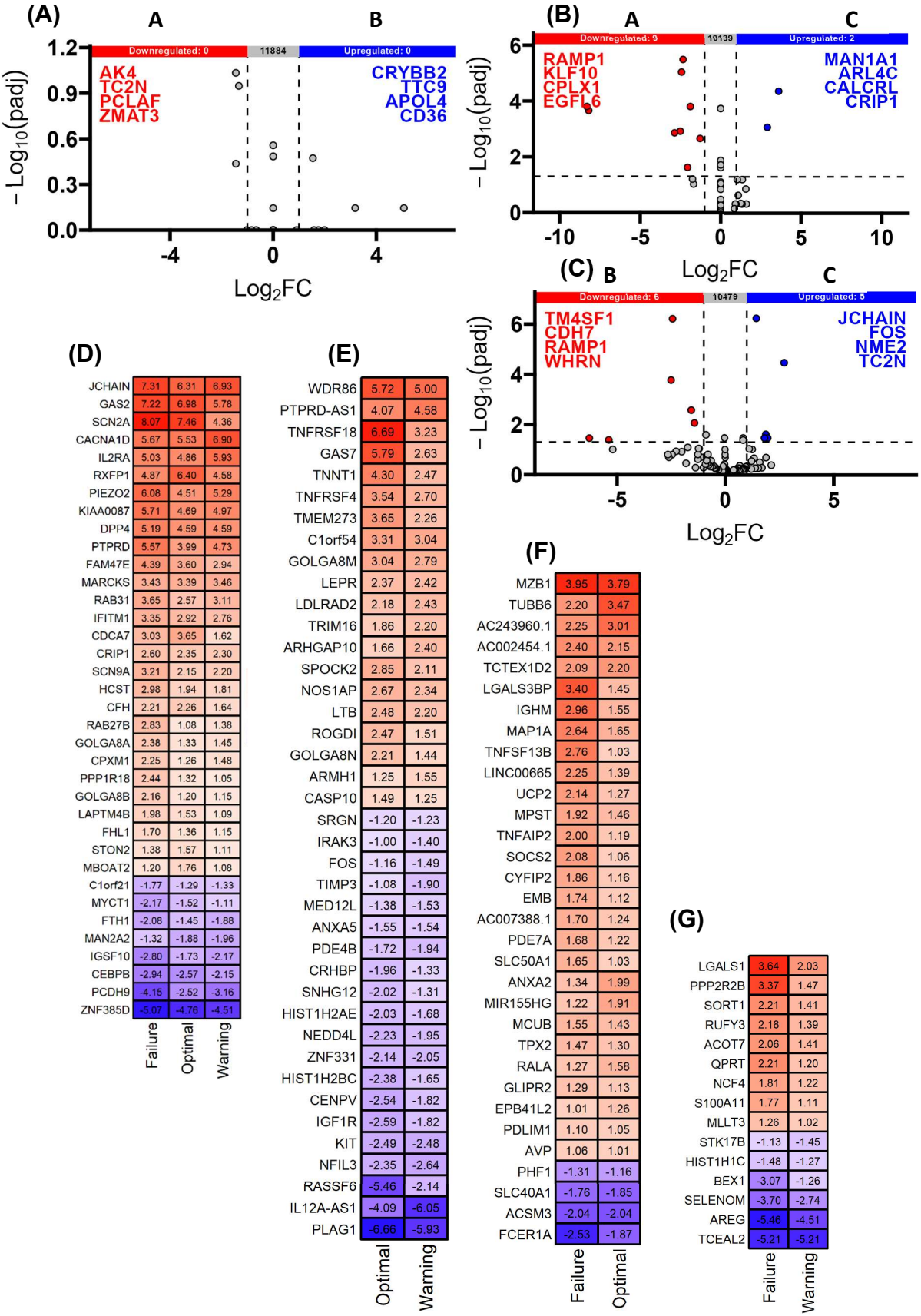


Figure 6-8 DE genes between and common to predicted primitive responder groups when compared to their specific bystanders. Volcano plot showing differences between (A) group B vs group A (B) group C vs group A and (C) group C vs group B. Genes common to (D) all comparisons (E) Optimal and warning (F) Failure and optimal and (g) Failure and warning comparisons, when compared to corresponding bystanders. Colour and values represent log2FC values.

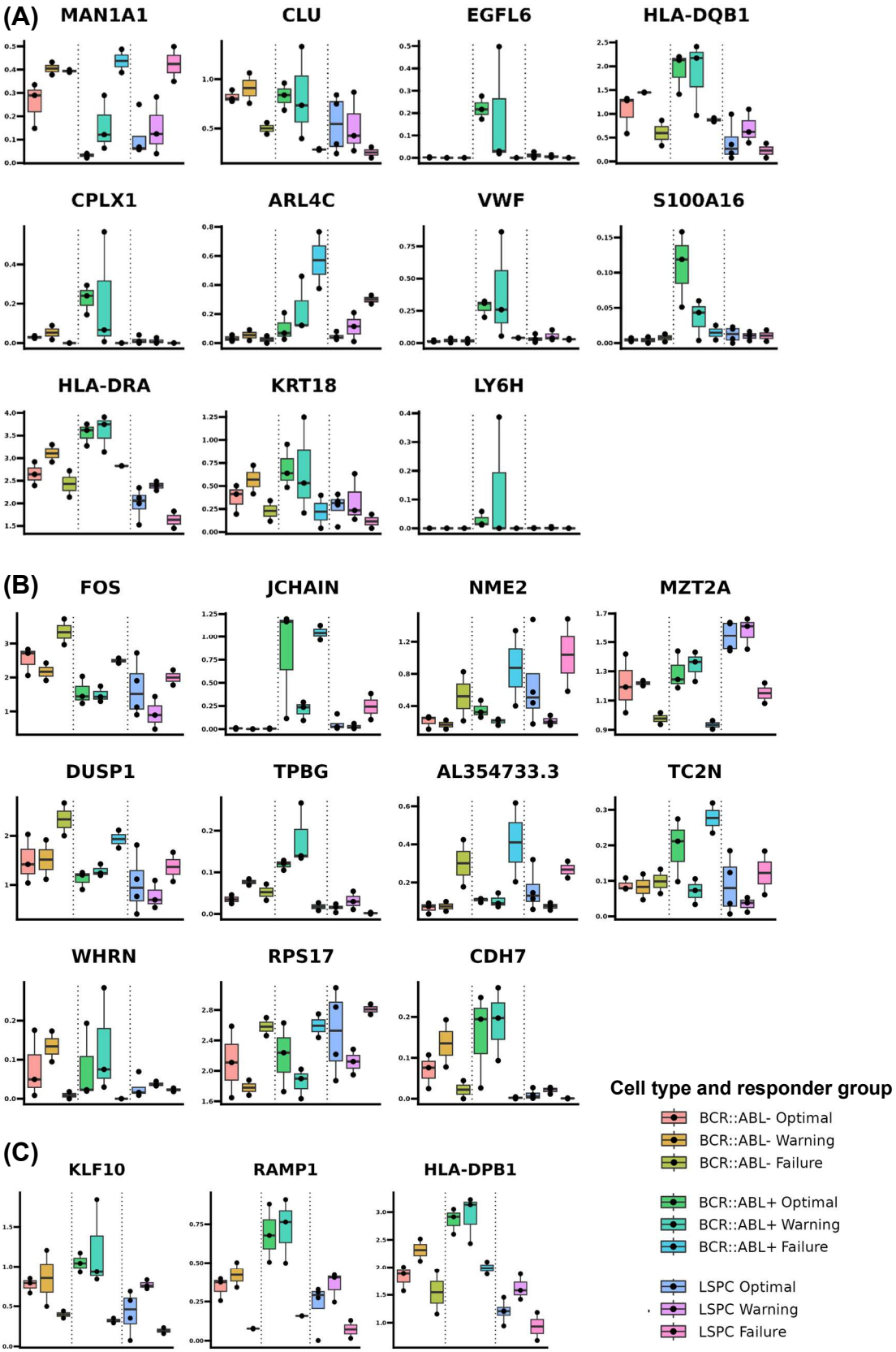


Figure 6-9 Difference between predicted primitive failure BCR::ABL⁺ LSCs, and optimal and warning and BCR::ABL⁺ LSCs. Genes capable of discriminating primitive failure BCR::ABL⁺ LSCs from (A) Optimal (B) warning or (C) both groups A and B. Corresponding BCR::ABL⁻ cells and LSPCs for these groups shown in the left hand and right hand of each plot respectively. Y axis represents normalised expression.

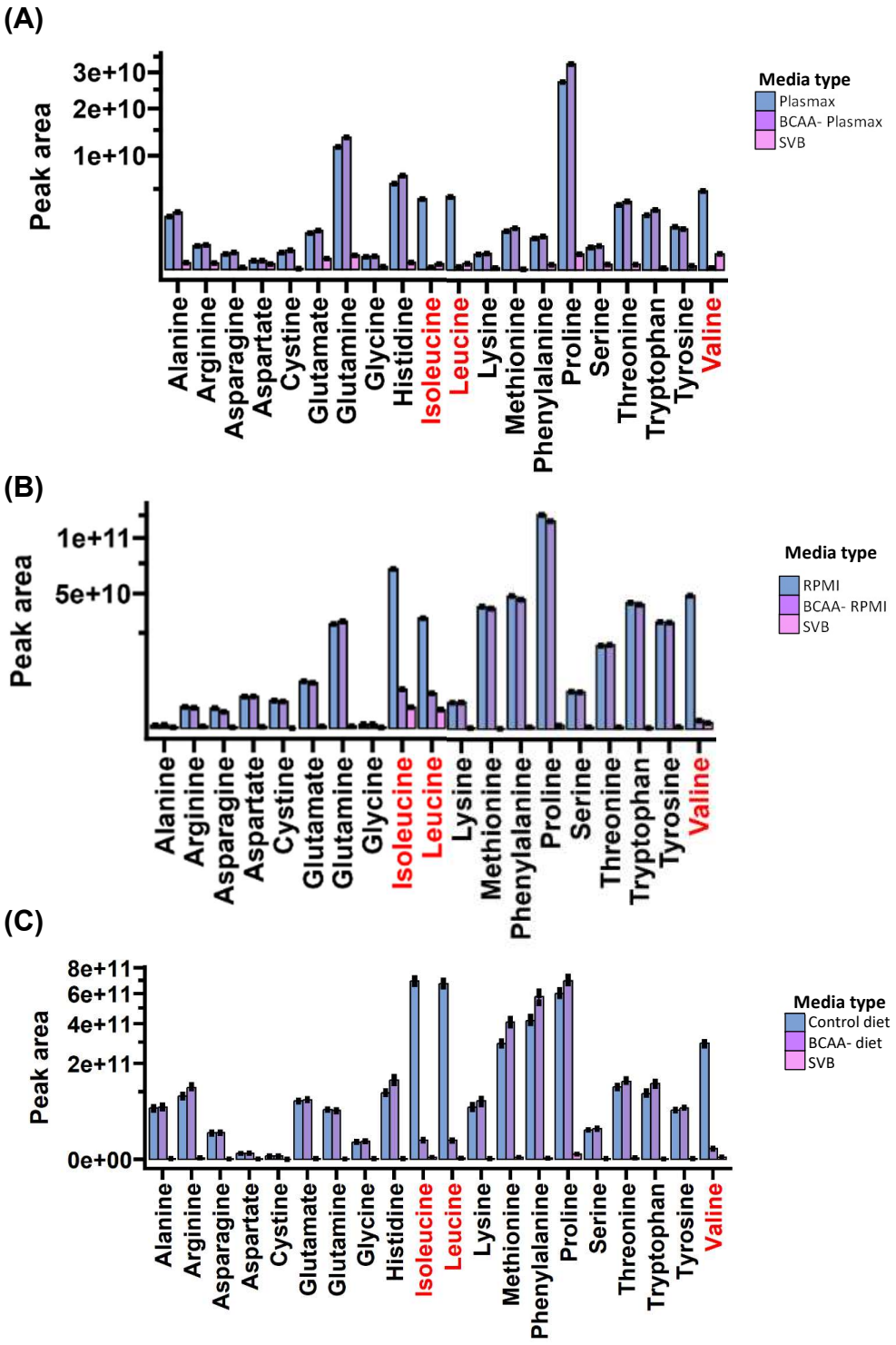


Figure 6-10 Relative AA concentrations measured by LC-MS of medias and diet pellets. Peak areas of (A) Full Plasmax and BCAA$^-$ Plasmax (n=1), (B) RPMI and BCAA$^-$ RPMI (n=1) and (C) Control and BCAA$^-$ diet pellets (n=3), alongside respective SVBs.

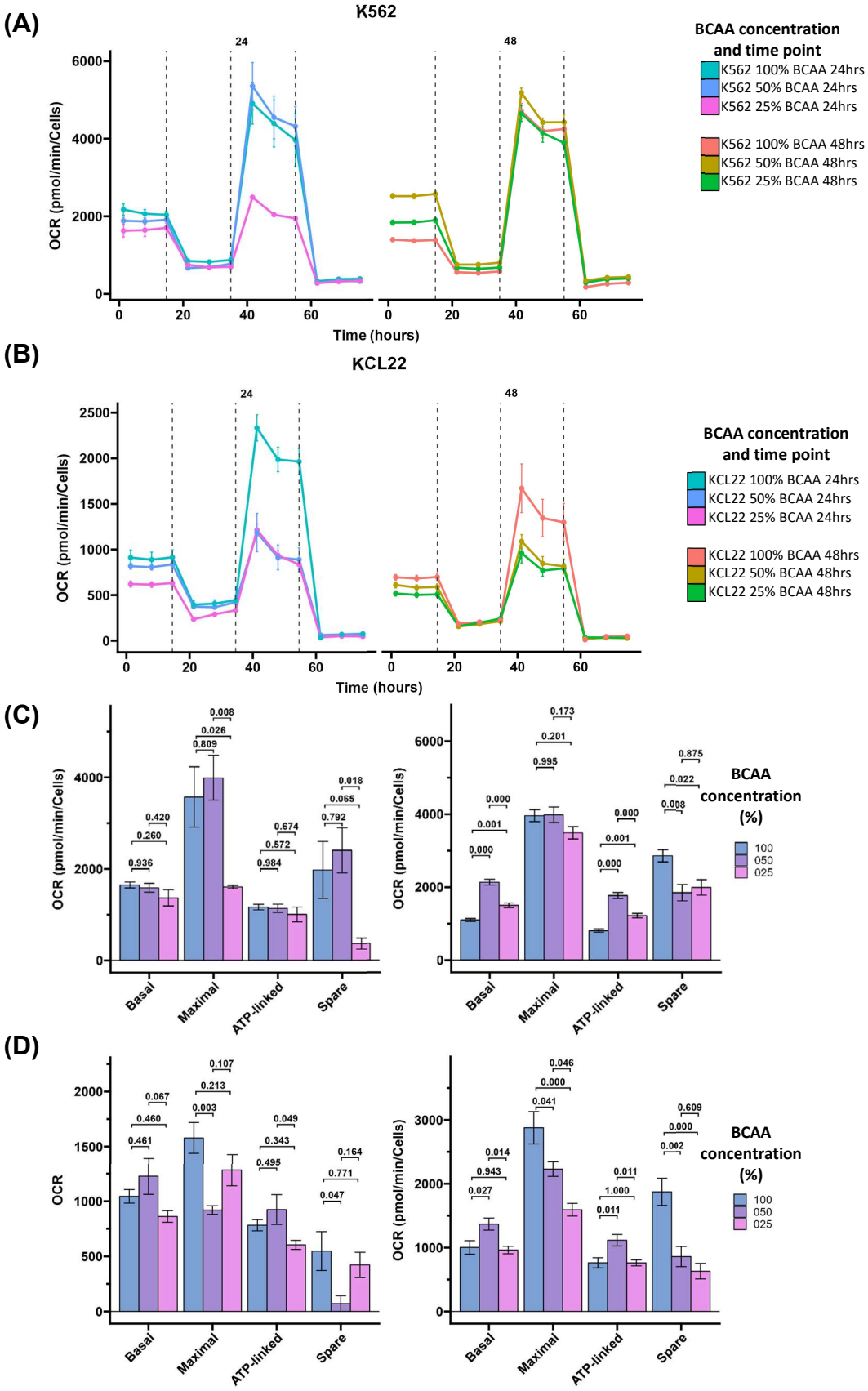


Figure 6-11 Biological replicates of seahorse assay of K562 and KCL22s cultured in BCAA deficient medias at 24 and 48 hours OCR rates for K562s (A) or KCL22s (B) at 24 or 48 hours. 6 TRs, 1BR per condition. Comparison of seahorse derived metabolic parameters of (C) K562s or (D) KCL22s, at 24 or 48 hours. Statistics calculated by ANOVA followed by Tukey test.

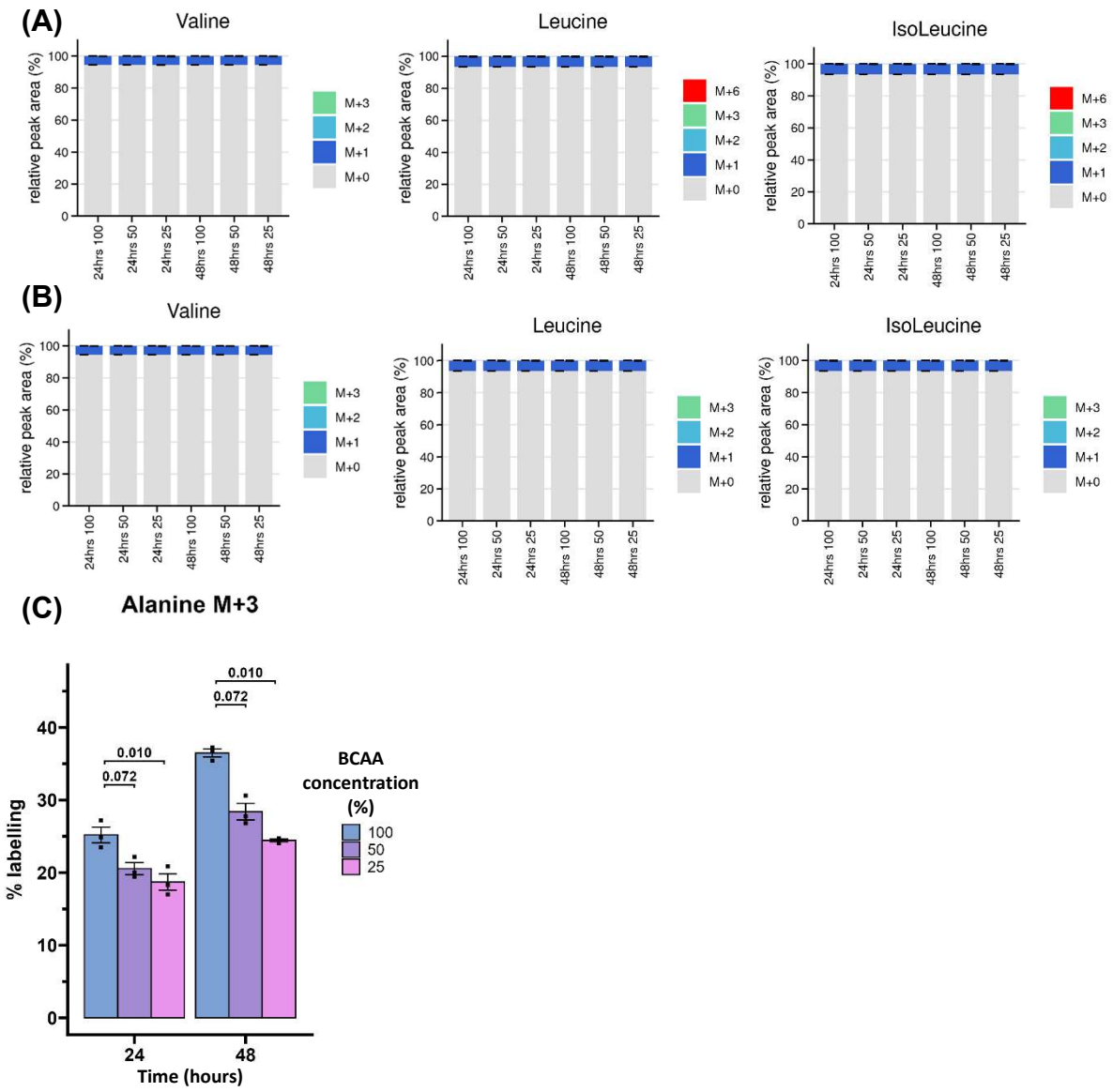


Figure 6-12 Glucose tracing of BCAAs from K562s and KCL22s cultured for 24 or 48 hours in full, 50% or 25% BCAA media. Labelling fractions for BCCAs from (A) K562s or (B) KCL22s. (C) Changes in M+3 labelling of alanine from K562s cultured in BCAA deficient medias at 24 and 48 hours.

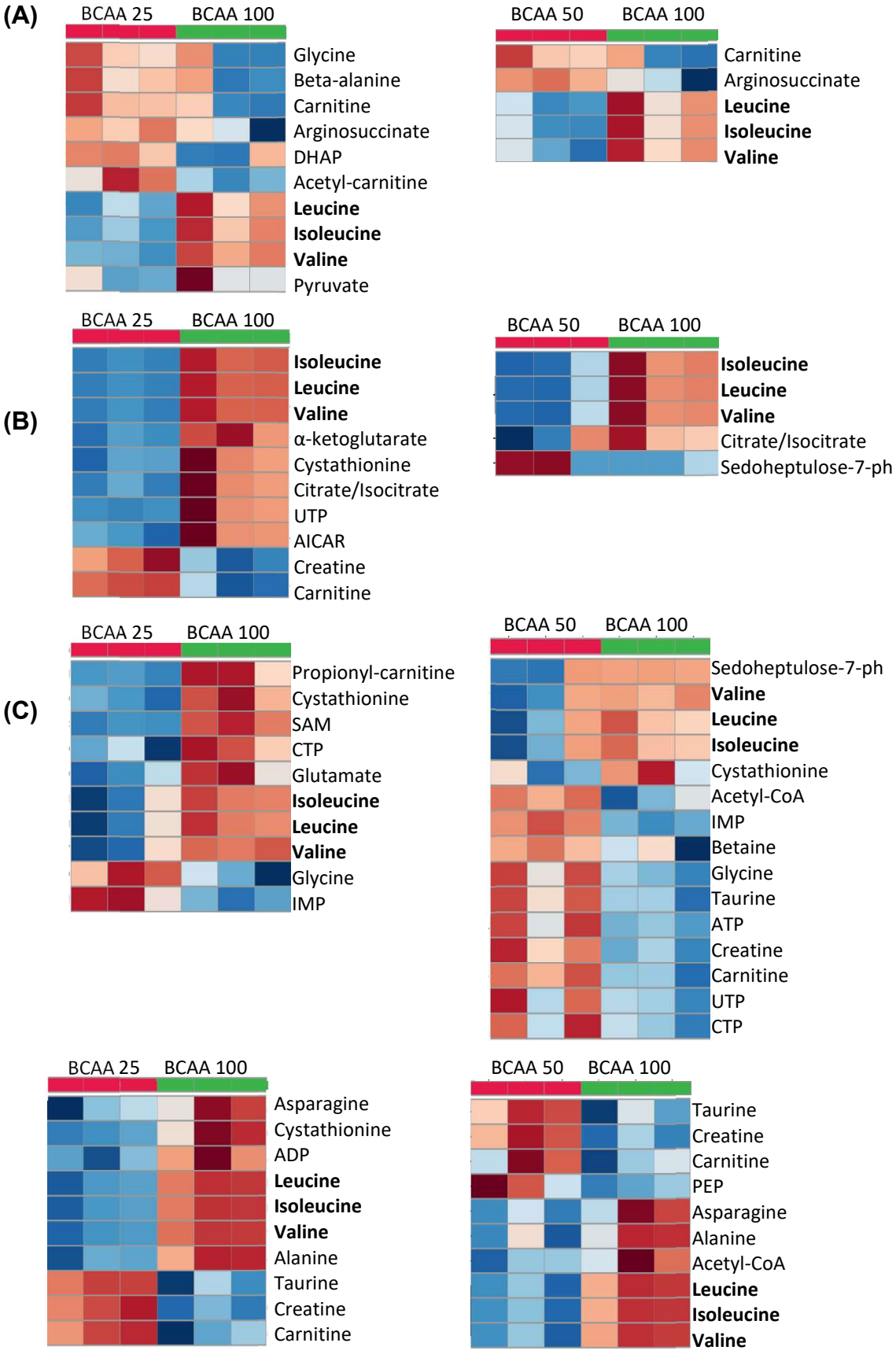


Figure 6-13 Heatmaps of top differentially expressed metabolites in K562s and KCL22s at 24 or 48 hours cultured in BCAA deficient medias. Top metabolites when comparing (A) full media to 50% or 25% media at 24 hours and (B) full media to 50% or 25% media at 48 hours in K562s. Top metabolites when comparing (C) full media to 50% or 25% media at 24 hours and (D) full media to 50% or 25% media at 48 hours in KCL22s.

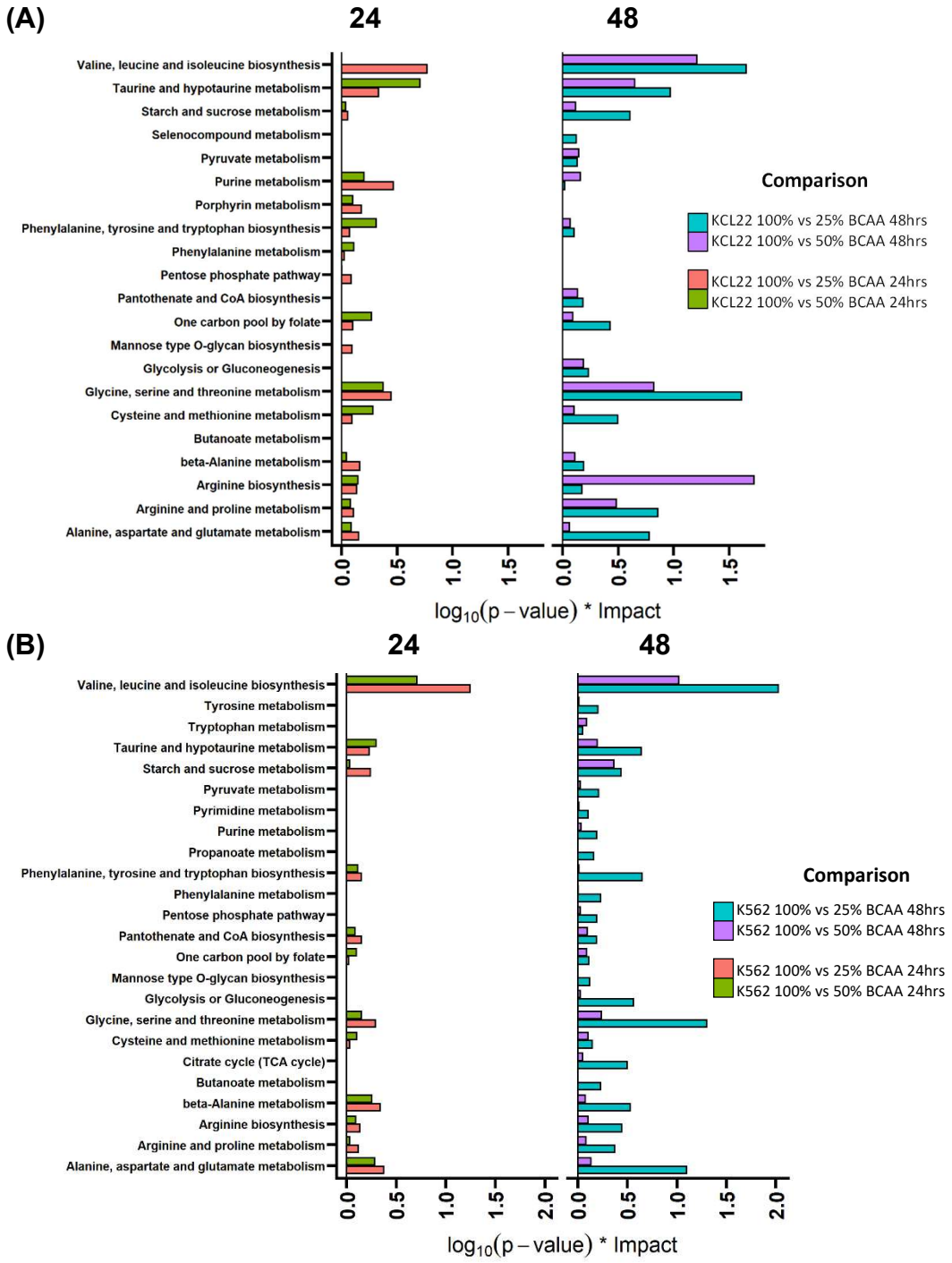


Figure 6-14 Pathway analysis from K562s or KCL22s cultured in BCAA deficient medias for 24 or 48 hours. DE pathways between full, 50% and 25% BCAA cultured cells at 24 or 48 hours. (A) KCL22s and (B) K562s. Pathways displayed have a padj < 0.25 in atleast one comparison set for a cell line. Each bar represents the metabolite enrichment score for the corresponding pathway and mouse/tissue type.

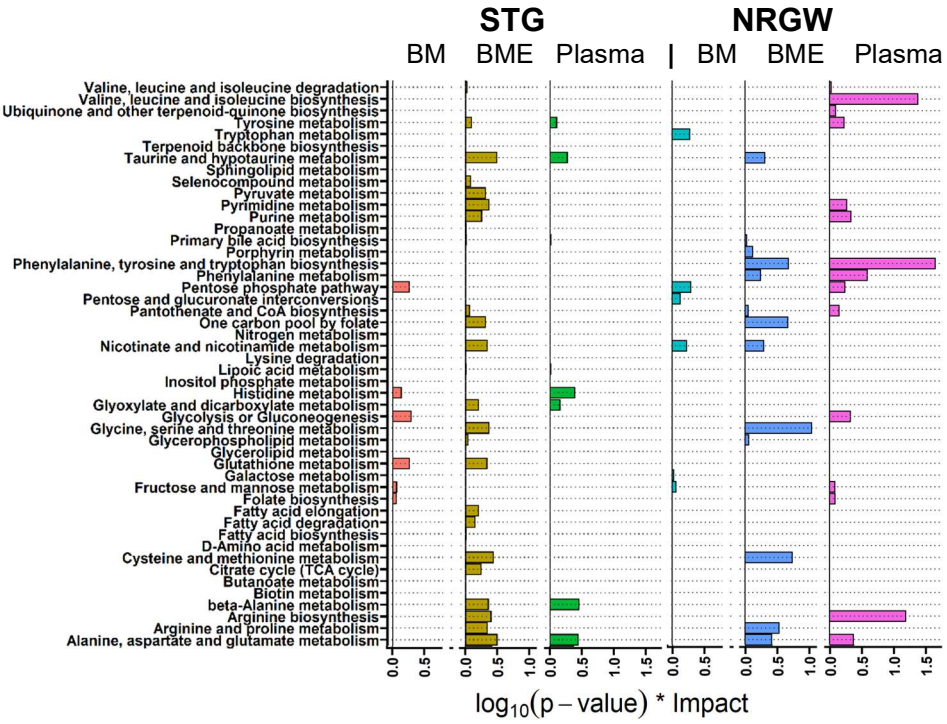


Figure 6-15 Pathway analysis comparing tissue from BCAA$^-$ diet mice to control diet mice. Pathway analysis of BM, BM environment (BME) and plasma from STG and NRGW mice. Pathways displayed have a padj < 0.25 in atleast one comparison.

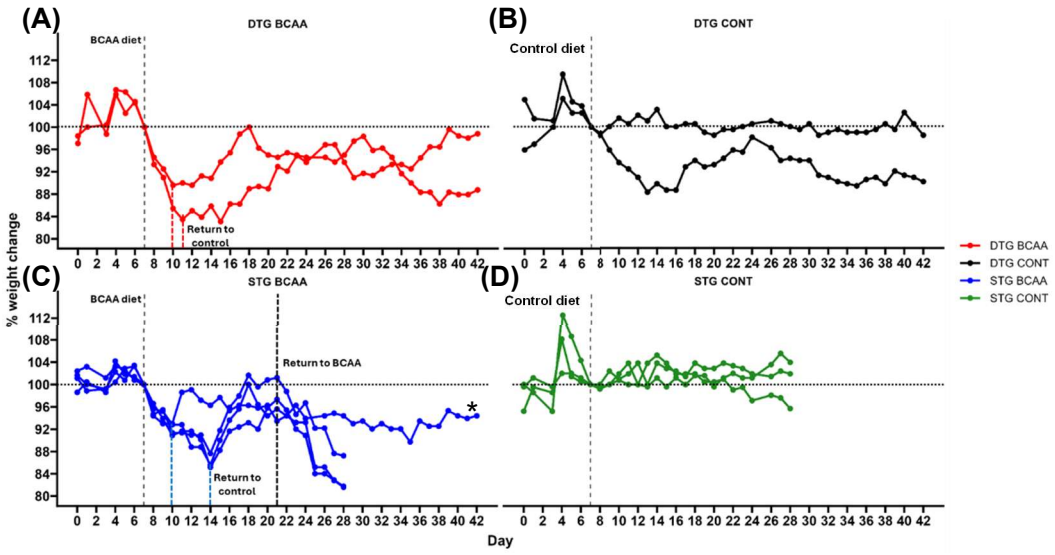


Figure 6-16 Relative change in mouse weights during two periods of BCAA restriction in healthy and leukaemic mice. Weights of (A) DTG transplanted mice on BCAA$^-$ diet or (B) control diet or (C) STG transplanted mice on BCAA$^-$ diet (D) or control diet. Labelled lines represent when mice were switched to either a control or BCAA$^-$ diet. Mice were maintained on a BCAA$^-$ diet until weight loss thresholds were approached, wherein they were returned to control diet. If or when a cage of BCAA$^-$ mice returned to control diet reached within 5% of their starting cage weight, they were returned to the BCAA$^-$ diet until weight loss thresholds were met. One STG BCAA$^-$ mouse was kept to mimic DTG BCAA$^-$ mice for the duration of the experiment, denoted by *.

References

1. Ng, A. P. & Alexander, W. S. Haematopoietic stem cells: past, present and future. *Cell Death Discov.* **3**, 17002 (2017).
2. Till, J. E. & McCulloch, E. A. A Direct Measurement of the Radiation Sensitivity of Normal Mouse Bone Marrow Cells. *Radiation Research* **14**, 213–222 (1961).
3. Bodine, D. M., Seidel, N. E. & Orlic, D. Bone Marrow Collected 14 Days After In Vivo Administration of Granulocyte Colony-Stimulating Factor and Stem Cell Factor to Mice Has 10-Fold More Repopulating Ability Than Untreated Bone Marrow. *Blood* **88**, 89–97 (1996).
4. Bernitz, J. M., Daniel, M. G., Fstkchyan, Y. S. & Moore, K. Granulocyte colony-stimulating factor mobilizes dormant hematopoietic stem cells without proliferation in mice. *Blood* **129**, 1901–1912 (2017).
5. Liang, R. *et al*. Restraining Lysosomal Activity Preserves Hematopoietic Stem Cell Quiescence and Potency. *Cell Stem Cell* **26**, 359-376.e7 (2020).
6. Wilkinson, A. C., Igarashi, K. J. & Nakauchi, H. Haematopoietic stem cell self-renewal in vivo and ex vivo. *Nat Rev Genet* **21**, 541–554 (2020).
7. Carrelha, J. *et al*. Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature* **554**, 106–111 (2018).
8. Yamamoto, R. *et al*. Clonal Analysis Unveils Self-Renewing Lineage-Restricted Progenitors Generated Directly from Hematopoietic Stem Cells. *Cell* **154**, 1112–1126 (2013).
9. Rodriguez-Fraticelli, A. E. *et al*. Clonal analysis of lineage fate in native hematopoiesis. *Nature* **553**, 212–216 (2018).
10. Sanjuan-Pla, A. *et al*. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* **502**, 232–236 (2013).
11. Loughran, S., Haas, S., Wilkinson, A. C., Klein, A. & Brand, M. Lineage commitment of hematopoietic stem cells and progenitors: insights from recent single cell and lineage tracing technologies. *Exp Hematol* **88**, 1–6 (2020).
12. Loughran, S., Haas, S., Wilkinson, A. C., Klein, A. & Brand, M. Lineage commitment of hematopoietic stem cells and progenitors: insights from recent single cell and lineage tracing technologies. *Exp Hematol* **88**, 1–6 (2020).
13. Meaker, G. A. & Wilkinson, A. C. Ex vivo hematopoietic stem cell expansion technologies: recent progress, applications, and open questions. *Experimental Hematology* **130**, 104136 (2024).
14. Rix, B., Maduro, A. H., Bridge, K. S. & Grey, W. Markers for human haematopoietic stem cells: The disconnect between an identification marker and its function. *Front. Physiol.* **13**, (2022).
15. Majeti, R., Park, C. Y. & Weissman, I. L. Identification of a Hierarchy of Multipotent Hematopoietic Progenitors in Human Cord Blood. *Cell Stem Cell* **1**, 635–645 (2007).
16. Takeda, H., Yamamoto, M., Morita, N. & Tanizawa, T. Relationship between Thy-1 expression and cell-cycle distribution in human bone marrow hematopoietic progenitors. *American Journal of Hematology* **79**, 187–193 (2005).
17. Hua, P. *et al*. Single-cell assessment of transcriptome alterations induced by Scriptaid in early differentiated human haematopoietic progenitors during ex vivo expansion. *Sci Rep* **9**, 5300 (2019).

18. Purton, L. E. Adult murine hematopoietic stem cells and progenitors: an update on their identities, functions, and assays. *Experimental Hematology* **116**, 1–14 (2022).
19. Morita, Y., Ema, H. & Nakauchi, H. Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. *J Exp Med* **207**, 1173–1182 (2010).
20. Grinenko, T. *et al*. Clonal expansion capacity defines two consecutive developmental stages of long-term hematopoietic stem cells. *J Exp Med* **211**, 209–215 (2014).
21. Wilson, N. K. *et al*. Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. *Cell Stem Cell* **16**, 712–724 (2015).
22. Rodriguez-Fraticelli, A. E. *et al*. Single-cell lineage tracing unveils a role for Tcf15 in haematopoiesis. *Nature* **583**, 585–589 (2020).
23. Martin, G. H. & Park, C. Y. EPCR: a novel marker of cultured cord blood HSCs. *Blood* **129**, 3279–3280 (2017).
24. Subramaniam, A., Talkhoncheh, M. S., Magnusson, M. & Larsson, J. Endothelial protein C receptor (EPCR) expression marks human fetal liver hematopoietic stem cells. *Haematologica* **104**, e47–e50 (2019).
25. Yin, A. H. *et al*. AC133, a Novel Marker for Human Hematopoietic Stem and Progenitor Cells. *Blood* **90**, 5002–5012 (1997).
26. Christopher, A. C. *et al*. Preferential Expansion of Human CD34+CD133+CD90+ Hematopoietic Stem Cells Enhances Gene-Modified Cell Frequency for Gene Therapy. *Hum Gene Ther* **33**, 188–201 (2022).
27. Notta, F. *et al*. Isolation of Single Human Hematopoietic Stem Cells Capable of Long-Term Multilineage Engraftment. *Science* **333**, 218–221 (2011).
28. Kricun, M. E. Red-yellow marrow conversion: Its effect on the location of some solitary bone lesions. *Skeletal Radiol* **14**, 10–19 (1985).
29. Yang, M., Büsche, G., Ganser, A. & Li, Z. Morphology and quantitative composition of hematopoietic cells in murine bone marrow and spleen of healthy subjects. *Ann Hematol* **92**, 587–594 (2013).
30. Cowan, P. T., Launico, M. V. & Kahai, P. Anatomy, Bones. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2025).
31. Busch, C., Nyamondo, K. & Wheadon, H. Complexities of modeling the bone marrow microenvironment to facilitate hematopoietic research. *Experimental Hematology* **135**, 104233 (2024).
32. Acar, M. *et al*. Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. *Nature* **526**, 126–130 (2015).
33. Fonseca, L. N. *et al*. Cell surface markers for mesenchymal stem cells related to the skeletal system: A scoping review. *Heliyon* **9**, e13464 (2023).
34. Li, T. & Wu, Y. Paracrine Molecules of Mesenchymal Stem Cells for Hematopoietic Stem Cell Niche. *Bone Marrow Research* **2011**, 353878 (2011).
35. Safi, R. *et al*. Direct Interaction Between CD34+ Hematopoietic Stem Cells and Mesenchymal Stem Cells Reciprocally Preserves Stemness. *Cancers (Basel)* **16**, 3972 (2024).
36. Saleh, M., Shamsasanjan, karim, Movassaghpourakbari, A., Akbarzadehlaleh, P. & Molaeipour, Z. The Impact of Mesenchymal Stem Cells on Differentiation of Hematopoietic Stem Cells. *Adv Pharm Bull* **5**, 299–304 (2015).

37. Aquino, J. B., Sierra, R. & Montaldo, L. A. Diverse cellular origins of adult blood vascular endothelial cells. *Developmental Biology* **477**, 117–132 (2021).
38. Pinho, S. & Frenette, P. S. Haematopoietic stem cell activity and interactions with the niche. *Nat Rev Mol Cell Biol* **20**, 303–320 (2019).
39. Xu, C. *et al*. Stem cell factor is selectively secreted by arterial endothelial cells in bone marrow. *Nat Commun* **9**, 2449 (2018).
40. Itkin, T. *et al*. Distinct bone marrow blood vessels differentially regulate hematopoiesis. *Nature* **532**, 323–328 (2016).
41. Khatib-Massalha, E. & Méndez-Ferrer, S. Megakaryocyte Diversity in Ontogeny, Functions and Cell-Cell Interactions. *Front Oncol* **12**, 840044 (2022).
42. Megakaryocytes regulate hematopoietic stem cell quiescence through CXCL4 secretion | Nature Medicine. https://doi.org/doi.org/10.1038/nm.3707 doi:doi.org/10.1038/nm.3707.
43. Zhao, M. *et al*. Megakaryocytes maintain homeostatic quiescence and promote post-injury regeneration of hematopoietic stem cells. *Nat Med* **20**, 1321–1326 (2014).
44. Huang, Z. *et al*. Cellular crosstalk in the bone marrow niche. *J Transl Med* **22**, 1096 (2024).
45. Kocabas, F. *et al*. Hypoxic metabolism in human hematopoietic stem cells. *Cell & Bioscience* **5**, 39 (2015).
46. Merchant, S. *et al*. Different effects of fatty acid oxidation on hematopoietic stem cells based on age and diet. *Cell Stem Cell* **32**, 263-275.e5 (2025).
47. Li, C. *et al*. Amino acid catabolism regulates hematopoietic stem cell proteostasis via a GCN2-eIF2α axis. *Cell Stem Cell* **29**, 1119-1134.e7 (2022).
48. Taya, Y. *et al*. Depleting dietary valine permits nonmyeloablative mouse hematopoietic stem cell transplantation. *Science* **354**, 1152–1155 (2016).
49. Wilkinson, A. C., Morita, M., Nakauchia, H. & Yamazaki, S. Branched-chain amino acid depletion conditions bone marrow for hematopoietic stem cell transplantation avoiding amino acid imbalance-associated toxicity. *Exp Hematol* **63**, 12-16.e1 (2018).
50. Parmar, K., Mauch, P., Vergilio, J.-A., Sackstein, R. & Down, J. D. Distribution of hematopoietic stem cells in the bone marrow according to regional hypoxia. *Proceedings of the National Academy of Sciences* **104**, 5431–5436 (2007).
51. Nombela-Arrieta, C. *et al*. Quantitative Imaging of Hematopoietic Stem and Progenitor Cell localization and hypoxic status in the Bone Marrow microenvironment. *Nat Cell Biol* **15**, 533–543 (2013).
52. Ziello, J. E., Jovin, I. S. & Huang, Y. Hypoxia-Inducible Factor (HIF)-1 Regulatory Pathway and its Potential for Therapeutic Intervention in Malignancy and Ischemia. *Yale J Biol Med* **80**, 51–60 (2007).
53. Zhang, C. C. & Sadek, H. A. Hypoxia and Metabolic Properties of Hematopoietic Stem Cells. *Antioxid Redox Signal* **20**, 1891–1901 (2014).
54. Takubo, K. *et al*. Regulation of the HIF-1α Level Is Essential for Hematopoietic Stem Cells. *Cell Stem Cell* **7**, 391–402 (2010).
55. Rehn, M. *et al*. Hypoxic induction of vascular endothelial growth factor regulates murine hematopoietic stem cell function in the low-oxygenic niche. *Blood* **118**, 1534–1543 (2011).
56. Dengler, V. L., Galbraith, M. & Espinosa, J. M. Transcriptional Regulation by Hypoxia Inducible Factors. *Crit Rev Biochem Mol Biol* **49**, 1–15 (2014).

57. Vukovic, M. *et al*. Adult hematopoietic stem cells lacking Hif-1α self-renew normally. *Blood* **127**, 2841–2846 (2016).
58. Guitart, A. V. *et al*. Hif-2α is not essential for cell-autonomous hematopoietic stem cell maintenance. *Blood* **122**, 1741–1745 (2013).
59. Ludin, A. *et al*. Reactive Oxygen Species Regulate Hematopoietic Stem Cell Self-Renewal, Migration and Development, As Well As Their Bone Marrow Microenvironment. *Antioxid Redox Signal* **21**, 1605–1619 (2014).
60. Bigarella, C. L. *et al*. FOXO3 Transcription Factor Is Essential for Protecting Hematopoietic Stem and Progenitor Cells from Oxidative DNA Damage. *J Biol Chem* **292**, 3005–3015 (2017).
61. Ito, K. *et al*. Regulation of oxidative stress by ATM is required for self-renewal of haematopoietic stem cells. *Nature* **431**, 997–1002 (2004).
62. Vannini, N. *et al*. The NAD-Booster Nicotinamide Riboside Potently Stimulates Hematopoiesis through Increased Mitochondrial Clearance. *Cell Stem Cell* **24**, 405-418.e7 (2019).
63. Kalaitzidis, D. *et al*. mTOR Complex 1 Plays Critical Roles in Hematopoiesis and *Pten*-Loss-Evoked Leukemogenesis. *Cell Stem Cell* **11**, 429–439 (2012).
64. Magee, J. A. *et al*. Temporal Changes in PTEN and mTORC2 Regulation of Hematopoietic Stem Cell Self-Renewal and Leukemia Suppression. *Cell Stem Cell* **11**, 415–428 (2012).
65. Montazersaheb, S., Ehsani, A., Fathi, E., Farahzadi, R. & Vietor, I. An Overview of Autophagy in Hematopoietic Stem Cell Transplantation. *Front. Bioeng. Biotechnol.* **10**, (2022).
66. Jung, C. H. *et al*. ULK-Atg13-FIP200 Complexes Mediate mTOR Signaling to the Autophagy Machinery. *MBoC* **20**, 1992–2003 (2009).
67. Kim, J., Kundu, M., Viollet, B. & Guan, K.-L. AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. *Nat Cell Biol* **13**, 132–141 (2011).
68. Mortensen, M. *et al*. The autophagy protein Atg7 is essential for hematopoietic stem cell maintenance. *J Exp Med* **208**, 455–467 (2011).
69. Jabbour, E. & Kantarjian, H. Chronic myeloid leukemia: 2025 update on diagnosis, therapy, and monitoring. *American Journal of Hematology* **99**, 2191–2212 (2024).
70. Risk Factors for CML | American Cancer Society. https://www.cancer.org/cancer/types/chronic-myeloid-leukemia/causes-risks-prevention/risk-factors.html.
71. Chronic myeloid leukaemia (CML) statistics. *Cancer Research UK* https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-cml (2015).
72. Holyoake, T. L. & Vetrie, D. The chronic myeloid leukemia stem cell: stemming the tide of persistence. *Blood* **129**, 1595–1606 (2017).
73. Ross, D. M. & Hughes, T. P. Treatment-free remission in patients with chronic myeloid leukaemia. *Nat Rev Clin Oncol* **17**, 493–503 (2020).
74. Nicolini, F. E. *et al*. The BCR-ABLT315I mutation compromises survival in chronic phase chronic myelogenous leukemia patients resistant to tyrosine kinase inhibitors, in a matched pair analysis. *Haematologica* **98**, 1510–1516 (2013).
75. Huang, X., Cortes, J. & Kantarjian, H. Estimations of the increasing prevalence and plateau prevalence of chronic myeloid leukemia in the era of tyrosine kinase inhibitor therapy. *Cancer* **118**, 3123–3127 (2012).

76. Jabbour, E. & Kantarjian, H. Chronic Myeloid Leukemia: A Review. *JAMA* https://doi.org/10.1001/jama.2025.0220 (2025) doi:10.1001/jama.2025.0220.
77. Casey, M. *et al*. Global Trial Representation and Availability of Tyrosine Kinase Inhibitors for Treatment of Chronic Myeloid Leukemia. *Cancers (Basel)* **16**, 2838 (2024).
78. O'Hare, T., Deininger, M. W. N., Eide, C. A., Clackson, T. & Druker, B. J. Targeting the BCR-ABL Signaling Pathway in Therapy-Resistant Philadelphia Chromosome-Positive Leukemia. *Clin Cancer Res* **17**, 212–221 (2011).
79. Shaul, Y. c-Abl: activation and nuclear targets. *Cell Death Differ* **7**, 10–16 (2000).
80. Dixon, A. S., Constance, J. E., Tanaka, T., Rabbitts, T. H. & Lim, C. S. Changing the Subcellular Location of the Oncoprotein Bcr-Abl Using Rationally Designed Capture Motifs. *Pharm Res* **29**, 1098–1109 (2012).
81. Biernaux, C., Loos, M., Sels, A., Huez, G. & Stryckmans, P. Detection of major bcr-abl gene expression at a very low level in blood cells of some healthy individuals. *Blood* **86**, 3118–3122 (1995).
82. Jaiswal, S. *et al*. Expression of BCR/ABL and BCL-2 in myeloid progenitors leads to myeloid leukemias. *Proceedings of the National Academy of Sciences* **100**, 10002–10007 (2003).
83. Mojtahedi, H., Yazdanpanah, N. & Rezaei, N. Chronic myeloid leukemia stem cells: targeting therapeutic implications. *Stem Cell Research & Therapy* **12**, 603 (2021).
84. Vetrie, D., Helgason, G. V. & Copland, M. The leukaemia stem cell: similarities, differences and clinical prospects in CML and AML. *Nat Rev Cancer* **20**, 158–173 (2020).
85. Koschmieder, S. *et al*. Reversible Expansion of Hematopoietic Stem Cells (HSC) and CML-Like Disease in Transgenic Mice Expressing BCR-ABL under the Control of the SCL 3′ Enhancer. *Blood* **104**, 715 (2004).
86. Kinstrie, R. *et al*. CD93 is expressed on chronic myeloid leukemia stem cells and identifies a quiescent population which persists after tyrosine kinase inhibitor therapy. *Leukemia* **34**, 1613–1625 (2020).
87. Herrmann, H. *et al*. Dipeptidylpeptidase IV (CD26) defines leukemic stem cells (LSC) in chronic myeloid leukemia. *Blood* **123**, 3951–3962 (2014).
88. Ebian, H. F. *et al*. Peripheral blood CD26 positive leukemic stem cells as a possible diagnostic and prognostic marker in chronic myeloid leukemia. *Leukemia Research Reports* **17**, 100321 (2022).
89. Warfvinge, R. *et al*. Single-cell multiomics analysis of chronic myeloid leukemia links cellular heterogeneity to therapy response. *eLife* **12**, RP92074.
90. Camacho, M. F. *et al*. Evaluation of leukemic stem cell (CD26+) in chronic myeloid leukemia patients with different molecular responses and in treatment-free remission. *Clin Exp Med* **25**, 93 (2025).
91. Kinstrie, R. *et al*. CD93 is expressed on chronic myeloid leukemia stem cells and identifies a quiescent population which persists after tyrosine kinase inhibitor therapy. *Leukemia* **34**, 1613–1625 (2020).
92. Baquero, P., Dawson, A. & Helgason, G. V. Autophagy and mitochondrial metabolism: insights into their role and therapeutic potential in chronic myeloid leukaemia. *The FEBS Journal* **286**, 1271–1283 (2019).
93. Kuntz, E. M. *et al*. Targeting mitochondrial oxidative phosphorylation eradicates therapy-resistant chronic myeloid leukemic stem cells. *Nat Med* **23**, 1234–1240 (2017).

94. Rattigan, K. M. *et al*. Pyruvate anaplerosis is a targetable vulnerability in persistent leukaemic stem cells. *Nat Commun* **14**, 4634 (2023).
95. Zarou, M. M. *et al*. Inhibition of mitochondrial folate metabolism drives differentiation through mTORC1 mediated purine sensing. *Nat Commun* **15**, 1931 (2024).
96. Qiu, S. *et al*. Metabolic adaptation to tyrosine kinase inhibition in leukemia stem cells. *Blood* **142**, 574–588 (2023).
97. Mancini, C. *et al*. BCR::ABL1 expression in chronic myeloid leukemia cells in low oxygen is regulated by glutamine via CD36-mediated fatty acid uptake. *Cancer Cell International* **25**, 176 (2025).
98. Landberg, N. *et al*. CD36 defines primitive chronic myeloid leukemia cells less responsive to imatinib but vulnerable to antibody-based therapeutic targeting. *Haematologica* **103**, 447–455 (2018).
99. Sheng, Z., Ma, L., Sun, J. E., Zhu, L. J. & Green, M. R. BCR-ABL suppresses autophagy through ATF5-mediated regulation of mTOR transcription. *Blood* **118**, 2840–2848 (2011).
100. Rothe, K. *et al*. The core autophagy protein ATG4B is a potential biomarker and therapeutic target in CML stem/progenitor cells. *Blood* **123**, 3622–3634 (2014).
101. Baquero, P. *et al*. Targeting quiescent leukemic stem cells using second generation autophagy inhibitors. *Leukemia* **33**, 981–994 (2019).
102. Altman, B. J. *et al*. Autophagy is Essential to Suppress Cell Stress and to Allow BCR-Abl-Mediated Leukemogenesis. *Oncogene* **30**, 1855–1867 (2011).
103. Bellodi, C. *et al*. Targeting autophagy potentiates tyrosine kinase inhibitor–induced cell death in Philadelphia chromosome–positive cells, including primary CML stem cells. *J Clin Invest* **119**, 1109–1123 (2009).
104. Karvela, M. *et al*. ATG7 regulates energy metabolism, differentiation and survival of Philadelphia-chromosome-positive cells. *Autophagy* **12**, 936–948 (2016).
105. Horne, G. A. *et al*. A randomised phase II trial of hydroxychloroquine and imatinib versus imatinib alone for patients with chronic myeloid leukaemia in major cytogenetic response with residual disease. *Leukemia* **34**, 1775–1786 (2020).
106. Ianniciello, A. *et al*. ULK1 inhibition promotes oxidative stress–induced differentiation and sensitizes leukemic stem cells to targeted therapy. *Sci Transl Med* **13**, eabd5016 (2021).
107. Ng, K. P. *et al*. Physiologic hypoxia promotes maintenance of CML stem cells despite effective BCR-ABL1 inhibition. *Blood* **123**, 3316–3326 (2014).
108. Wang, J. *et al*. HIF-2α inhibition disrupts leukemia stem cell metabolism and impairs vascular microenvironment to enhance chronic myeloid leukemia treatment. *Cancer Letters* **597**, 217060 (2024).
109. Rouault-Pierre, K. *et al*. HIF-2α Protects Human Hematopoietic Stem/Progenitors and Acute Myeloid Leukemic Cells from Apoptosis Induced by Endoplasmic Reticulum Stress. *Cell Stem Cell* **13**, 549–563 (2013).
110. Ohmine, K. *et al*. Characterization of stage progression in chronic myeloid leukemia by DNA microarray with purified hematopoietic stem cells. *Oncogene* **20**, 8249–8257 (2001).
111. Zheng, C. *et al*. Gene expression profiling of CD34+ cells identifies a molecular signature of chronic myeloid leukemia blast crisis. *Leukemia* **20**, 1028–1034 (2006).

112. Diaz-Blanco, E. *et al*. Molecular signature of CD34(+) hematopoietic stem and progenitor cells of patients with CML in chronic phase. *Leukemia* **21**, 494–504 (2007).
113. Scott, M. T. *et al*. Epigenetic reprogramming sensitizes CML stem cells to combined EZH2 and tyrosine kinase inhibition. *Cancer Discov* **6**, 1248–1257 (2016).
114. Giustacchini, A. *et al*. Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nat Med* **23**, 692–702 (2017).
115. Warfvinge, R. *et al*. Single-cell multiomics analysis of chronic myeloid leukemia links cellular heterogeneity to therapy response. *Elife* **12**, RP92074 (2024).
116. Scott, M. T. *et al*. Activating p53 abolishes self-renewal of quiescent leukaemic stem cells in residual CML disease. *Nat Commun* **15**, 651 (2024).
117. Krishnan, V. *et al*. A single-cell atlas identifies pretreatment features of primary imatinib resistance in chronic myeloid leukemia. *Blood* **141**, 2738–2755 (2023).
118. Hochhaus, A. *et al*. European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia. *Leukemia* **34**, 966–984 (2020).
119. Kantarjian, H. *et al*. Dasatinib versus Imatinib in Newly Diagnosed Chronic-Phase Chronic Myeloid Leukemia. *New England Journal of Medicine* **362**, 2260–2270 (2010).
120. Saglio, G. *et al*. Nilotinib versus Imatinib for Newly Diagnosed Chronic Myeloid Leukemia. *New England Journal of Medicine* **362**, 2251–2259 (2010).
121. Wilkinson, A. C. *et al*. Long-term ex vivo haematopoietic-stem-cell expansion allows nonconditioned transplantation. *Nature* **571**, 117–121 (2019).
122. Igarashi, K. J. *et al*. Physioxia improves the selectivity of hematopoietic stem cell expansion cultures. *Blood Adv* **7**, 3366–3377 (2023).
123. Zhang, Q., Olofzon, R., Konturek-Ciesla, A., Yuan, O. & Bryder, D. Ex vivo expansion potential of murine hematopoietic stem cells is a rare property only partially predicted by phenotype. *eLife* **12**, RP91826.
124. Che, J. L. C. *et al*. Identification and characterization of in vitro expanded hematopoietic stem cells. *EMBO reports* **23**, e55502 (2022).
125. Becker, H. J. *et al*. Controlling genetic heterogeneity in gene-edited hematopoietic stem cells by single-cell expansion. *Cell Stem Cell* **30**, 987-1000.e8 (2023).
126. Kobayashi, H. *et al*. Environmental Optimization Enables Maintenance of Quiescent Hematopoietic Stem Cells Ex Vivo. *Cell Rep* **28**, 145-158.e9 (2019).
127. Oedekoven, C. A. *et al*. Hematopoietic stem cells retain functional potential and molecular identity in hibernation cultures. *Stem Cell Reports* **16**, 1614–1628 (2021).
128. Doma, E. *et al*. A robust approach for the generation of functional hematopoietic progenitor cell lines to model leukemic transformation. *Blood Adv* **5**, 39–53 (2021).
129. Patel, S. B. *et al*. Ex Vivo Expansion of Phenotypic and Transcriptomic Chronic Myeloid Leukemia Stem Cells. *Exp Hematol* **115**, 1–13 (2022).
130. Shah, M. *et al*. Low c-Kit expression identifies primitive, therapy-resistant CML stem cells. *JCI Insight* **8**, e157421.
131. Sakurai, M. *et al*. Chemically defined cytokine-free expansion of human haematopoietic stem cells. *Nature* **615**, 127–133 (2023).

132. Johnson, C. S. *et al*. Adaptation to ex vivo culture reduces human hematopoietic stem cell activity independently of the cell cycle. *Blood* **144**, 729–741 (2024).
133. Mantel, C. R. *et al*. Enhancing hematopoietic stem cell transplantation efficacy by mitigating oxygen shock. *Cell* **161**, 1553–1565 (2015).
134. Oostendorp, R. A. J. *et al*. Stromal cell lines from mouse aorta-gonads-mesonephros subregions are potent supporters of hematopoietic stem cell activity. *Blood* **99**, 1183–1189 (2002).
135. Jing, D. *et al*. Hematopoietic stem cells in co-culture with mesenchymal stromal cells - modeling the niche compartments in vitro. *Haematologica* **95**, 542–550 (2010).
136. Sood, S. *et al*. 3195 – LONG-TERM EX VIVO HEMATOPOIETIC STEM CELL (HSC) EXPANSION USING NOVEL BONE-LINING REINVIGORATING MESENCHYMAL STROMAL CELLS (RMSCS). *Experimental Hematology* **111**, S142 (2022).
137. Rubio-Lara, J. A. *et al*. Expanding hematopoietic stem cell ex vivo: recent advances and technical considerations. *Experimental Hematology* **125–126**, 6–15 (2023).
138. Bai, T. *et al*. Expansion of primitive human hematopoietic stem cells by culture in a zwitterionic hydrogel. *Nat Med* **25**, 1566–1575 (2019).
139. Hao, J. *et al*. Membrane-bound SCF and VCAM-1 synergistically regulate the morphology of hematopoietic stem cells. *Journal of Cell Biology* **220**, e202010118 (2021).
140. Cuchiara, M. L., Horter, K. L., Banda, O. A. & West, J. L. Covalent immobilization of stem cell factor and stromal derived factor 1α for in vitro culture of hematopoietic progenitor cells. *Acta Biomaterialia* **9**, 9258–9269 (2013).
141. Gvaramia, D. *et al*. Combined influence of biophysical and biochemical cues on maintenance and proliferation of hematopoietic stem cells. *Biomaterials* **138**, 108–117 (2017).
142. Donnelly, H. *et al*. Bioengineered niches that recreate physiological extracellular matrix organisation to support long-term haematopoietic stem cells. *Nat Commun* **15**, 5791 (2024).
143. Santos Souza, H. F., Marsiccobetre, S., Souza, R. O. O., Luévano-Martínez, L. A. & Silber, A. M. Branched chain amino acids catabolism as a source of new drug targets in pathogenic protists. *Experimental Parasitology* **249**, 108499 (2023).
144. Peng, H., Wang, Y. & Luo, W. MULTIFACETED ROLE OF BRANCHED-CHAIN AMINO ACID METABOLISM IN CANCER. *Oncogene* **39**, 6747–6756 (2020).
145. Mann, G., Mora, S., Madu, G. & Adegoke, O. A. J. Branched-chain Amino Acids: Catabolism in Skeletal Muscle and Implications for Muscle and Whole-body Metabolism. *Front. Physiol.* **12**, (2021).
146. Jiang, Y. *et al*. BCAT1 contributes to the development of TKI-resistant CML. *Cell Oncol.* **48**, 411–424 (2025).
147. Hattori, A. *et al*. Cancer progression by reprogrammed BCAA metabolism in myeloid leukemia. *Nature* **545**, 500–504 (2017).
148. Kikushige, Y. *et al*. Human acute leukemia uses branched-chain amino acid catabolism to maintain stemness through regulating PRC2 function. *Blood Adv* **7**, 3592–3603 (2022).
149. Brunmeir, R. *et al*. EZH2 modulates mRNA splicing and exerts part of its oncogenic function through repression of splicing factors in CML. *Leukemia* **39**, 650–662 (2025).

150. Raffel, S. *et al*. BCAT1 restricts αKG levels in AML stem cells leading to IDHmut-like DNA hypermethylation. *Nature* **551**, 384–388 (2017).
151. Budhathoki, S. *et al*. Association of plasma concentrations of branched-chain amino acids with risk of colorectal adenoma in a large Japanese population. *Ann Oncol* **28**, 818–823 (2017).
152. Katagiri, R. *et al*. Increased Levels of Branched-Chain Amino Acid Associated With Increased Risk of Pancreatic Cancer in a Prospective Case-Control Study of a Large Cohort. *Gastroenterology* **155**, 1474-1482.e1 (2018).
153. Chen, S., He, G., Zhang, M., Tang, N. & Zeng, Y. Causal relationship between branched-chain amino acids and leukemia risk: insights from a two-sample Mendelian randomization study. *Hematology* **29**, 2433904 (2024).
154. Wilkinson, A. C., Morita, M., Nakauchi, H. & Yamazaki, S. Branched-chain amino acid depletion conditions bone marrow for hematopoietic stem cell transplantation avoiding amino acid imbalance-associated toxicity. *Experimental Hematology* **63**, 12-16.e1 (2018).
155. Xie, H. *et al*. Polycomb repressive complex 2 regulates hematopoietic stem cell maintenance and differentiation in a developmental stage-specific manner. *Cell Stem Cell* **14**, 68–80 (2014).
156. Ramzan, I. *et al*. The Association between Circulating Branched Chain Amino Acids and the Temporal Risk of Developing Type 2 Diabetes Mellitus: A Systematic Review & Meta-Analysis. *Nutrients* **14**, 4411 (2022).
157. Karusheva, Y. *et al*. Short-term dietary reduction of branched-chain amino acids reduces meal-induced insulin secretion and modifies microbiome composition in type 2 diabetes: a randomized controlled crossover trial. *Am J Clin Nutr* **110**, 1098–1107 (2019).
158. MetaSRA. https://metasra.biostat.wisc.edu/?species=human&assay=RNA-seq.
159. Gómez-Castañeda, E. *et al*. Tyrosine Kinase Inhibitor Independent Gene Expression Signature in CML Offers New Targets for LSPC Eradication Therapy. *Cancers (Basel)* **14**, 5253 (2022).
160. Scott, M. T. *et al*. Activating p53 abolishes self-renewal of quiescent leukaemic stem cells in residual CML disease. *Nat Commun* **15**, 651 (2024).
161. Pellicano, F. *et al*. hsa-mir183/EGR1–mediated regulation of E2F1 is required for CML stem/progenitor cell survival. *Blood* **131**, 1532–1544 (2018).
162. Bruennert, D. *et al*. Early in vivo changes of the transcriptome in Philadelphia chromosome-positive CD34+ cells from patients with chronic myelogenous leukaemia following imatinib therapy. *Leukemia* **23**, 983–985 (2009).
163. Sinnakannu, J. R. *et al*. SRSF1 mediates cytokine-induced impaired imatinib sensitivity in chronic myeloid leukemia. *Leukemia* **34**, 1787–1798 (2020).
164. Youn, M. *et al*. Comparison of the Transcriptomic Signatures in Pediatric and Adult CML. *Cancers (Basel)* **13**, 6263 (2021).
165. Abraham, S. A. *et al*. Dual targeting of p53 and c-MYC selectively eliminates leukaemic stem cells. *Nature* **534**, 341–346 (2016).
166. Graham, S. M., Vass, J. K., Holyoake, T. L. & Graham, G. J. Transcriptional analysis of quiescent and proliferating CD34+ human hemopoietic cells from normal and chronic myeloid leukemia sources. *Stem Cells* **25**, 3111–3120 (2007).

167. Affer, M. *et al*. Gene Expression Differences between Enriched Normal and Chronic Myelogenous Leukemia Quiescent Stem/Progenitor Cells and Correlations with Biological Abnormalities. *J Oncol* **2011**, 798592 (2011).
168. Radich, J. P. *et al*. Gene expression changes associated with progression and response in chronic myeloid leukemia. *Proc Natl Acad Sci U S A* **103**, 2794–2799 (2006).
169. Avilés-Vázquez, S. *et al*. Global gene expression profiles of hematopoietic stem and progenitor cells from patients with chronic myeloid leukemia: the effect of in vitro culture with or without imatinib. *Cancer Med* **6**, 2942–2956 (2017).
170. Cramer-Morales, K. *et al*. Personalized synthetic lethality induced by targeting RAD52 in leukemias identified by gene mutation and expression profile. *Blood* **122**, 1293–1304 (2013).
171. Krishnan, V. *et al*. A single-cell atlas identifies pretreatment features of primary imatinib resistance in chronic myeloid leukemia. *Blood* **141**, 2738–2755 (2023).
172. Takeda, H., Yamamoto, M., Morita, N. & Tanizawa, T. Relationship between Thy-1 expression and cell-cycle distribution in human bone marrow hematopoietic progenitors. *American Journal of Hematology* **79**, 187–193 (2005).
173. Avilés-Vázquez, S. *et al*. Global gene expression profiles of hematopoietic stem and progenitor cells from patients with chronic myeloid leukemia: the effect of in vitro culture with or without imatinib. *Cancer Med* **6**, 2942–2956 (2017).
174. Khan, M. S. AI-Driven Identification and Validation of RXFP1 as a CML Biomarker Using Gene Expression and Integrated GUI Tools. 2025.06.06.658261 Preprint at https://doi.org/10.1101/2025.06.06.658261 (2025).
175. Escobar, L. M., Bendahan, Z., Garcia, C. & Castellanos, J. E. Relaxin treatment stimulates the differentiation of mesenchymal stem cells into osteoblasts. *J Dent Sci* **18**, 1786–1793 (2023).
176. Illiano, S. *et al*. Characterization of a new potent and long-lasting single chain peptide agonist of RXFP1 in cells and in vivo translational models. *Sci Rep* **12**, 20435 (2022).
177. Xing, C. *et al*. FT-Kernel: An innovative kernel for decoding cellular secrets related time. 2025.03.01.640966 Preprint at https://doi.org/10.1101/2025.03.01.640966 (2025).
178. Barbieri, D. A., Telonis, A. G. & Figueroa, M. K. E. Corticotropin Releasing Hormone Binding Protein (CRHBP) Regulates Hematopoietic Function and Acts As a Novel AML Tumor Suppressor through a Crh-Independent Mechanism. *Blood* **140**, 5840–5841 (2022).
179. Prenkert, M., Uggla, B., Tidefelt, U. & Strid, H. CRIM1 is expressed at higher levels in drug-resistant than in drug-sensitive myeloid leukemia HL60 cells. *Anticancer Res* **30**, 4157–4161 (2010).
180. Sprenger, A., Carr, H. S., Ulu, A. & Frost, J. A. Src stimulates Abl-dependent phosphorylation of the guanine exchange factor Net1A to promote its cytosolic localization and cell motility. *J Biol Chem* **299**, 104887 (2023).
181. Alvarez-Baron, C. P., Jonsson, P., Thomas, C., Dryer, S. E. & Williams, C. The two-pore domain potassium channel KCNK5: induction by estrogen receptor alpha and role in proliferation of breast cancer cells. *Mol Endocrinol* **25**, 1326–1336 (2011).

182. Hailin, L. *et al*. Ly6E on tumor cells impairs anti-tumor T-cell responses: a novel mechanism of tumor-induced immune exclusion. *Cancer Immunol Immunother* **74**, 4 (2024).
183. Kim, Y.-M. *et al*. SH3BP4 is a negative regulator of amino acid-Rag GTPase-mTORC1 signaling. *Mol Cell* **46**, 833–846 (2012).
184. Zeng, Z., Xu, S., Wang, R. & Han, X. FKBP4 promotes glycolysis and hepatocellular carcinoma progression via p53/HK2 axis. *Sci Rep* **14**, 26893 (2024).
185. Priestley, J. R. C. *et al*. Malate dehydrogenase 2 deficiency is an emerging cause of pediatric epileptic encephalopathy with a recognizable biochemical signature. *Mol Genet Metab Rep* **33**, 100931 (2022).
186. Picaud, S. *et al*. Structural basis of fumarate hydratase deficiency. *J Inherit Metab Dis* **34**, 671–676 (2011).
187. Srivastava, S. *et al*. Regulation of mitochondrial protein import by the nucleotide exchange factors GrpEL1 and GrpEL2 in human cells. *J Biol Chem* **292**, 18075–18090 (2017).
188. Elrod, J. W. & Molkentin, J. D. Physiologic Functions of Cyclophilin D and the Mitochondrial Permeability Transition Pore. *Circ J* **77**, 1111–1122 (2013).
189. Yang, S.-Y. *et al*. Mental retardation linked to mutations in the HSD17B10 gene interfering with neurosteroid and isoleucine metabolism. *Proc Natl Acad Sci U S A* **106**, 14820–14824 (2009).
190. Liu, L. *et al*. Deacetylation of HSD17B10 by SIRT3 regulates cell growth and cell resistance under oxidative and starvation stresses. *Cell Death Dis* **11**, 563 (2020).
191. Bae, E. Y. *et al*. First Case of Peroxisomal D-bifunctional Protein Deficiency with Novel HSD17B4 Mutations and Progressive Neuropathy in Korea. *J Korean Med Sci* **35**, e357 (2020).
192. Ma, Z. *et al*. Hepatic Acat2 overexpression promotes systemic cholesterol metabolism and adipose lipid metabolism in mice. *Diabetologia* **66**, 390–405 (2023).
193. Fu, X. & Wang, Z. DHCR24 in Tumor Diagnosis and Treatment: A Comprehensive Review. *Technol Cancer Res Treat* **23**, 15330338241259780 (2024).
194. Parente, M., Tonini, C., Caputo, S., Fiocchetti, M. & Pallottini, V. Mechanisms of Sigma-2/TMEM97 Involvement in Cholesterol Metabolism. *J Cell Biochem* **125**, e30645 (2024).
195. Mohania, D., Shokeen, Y., Jauhri, M. & Aggarwal, S. 338P Association of kynurenine, tryptophan and key enzymes involved in kynurenine pathway of tryptophan catabolism with Imatinib duration in patients with chronic myeloid leukemia. *Annals of Oncology* **27**, ix106 (2016).
196. KYNU kynureninase [Homo sapiens (human)] - Gene - NCBI. https://www.ncbi.nlm.nih.gov/gene?Cmd=DetailsSearch&Db=gene&Term=8942.
197. PubChem. SLC16A7 - solute carrier family 16 member 7 (human). https://pubchem.ncbi.nlm.nih.gov/gene/SLC16A7/human.
198. Kamal, A. M., Hamdy, N. M., Hegab, H. M. & El-Mesallamy, H. O. Expression of thioredoxin-1 (TXN) and its relation with oxidative DNA damage and treatment outcome in adult AML and ALL: A comparative study. *Hematology* **21**, 567–575 (2016).
199. Han, J. M. & Jung, H. J. Cyclophilin A/CD147 Interaction: A Promising Target for Anticancer Therapy. *International Journal of Molecular Sciences* **23**, 9341 (2022).

200. Di, S.-C. *et al*. DEPDC1 as a metabolic target regulates glycolysis in renal cell carcinoma through AKT/mTOR/HIF1α pathway. *Cell Death Dis* **15**, 533 (2024).
201. Gomes, F., Turano, H., Haddad, L. A. & Netto, Luis. E. S. Human mitochondrial peroxiredoxin Prdx3 is dually localized in the intermembrane space and matrix subcompartments. *Redox Biology* **78**, 103436 (2024).
202. Hoppins, S. *et al*. The soluble form of Bax regulates mitochondrial fusion via MFN2 homotypic complexes. *Mol Cell* **41**, 150–160 (2011).
203. Shoshan-Barmatz, V. *et al*. VDAC, a multi-functional mitochondrial protein regulating cell life and death. *Mol Aspects Med* **31**, 227–285 (2010).
204. Xian, H. *et al*. Oxidized DNA fragments exit mitochondria via mPTP- and VDAC-dependent channels to activate NLRP3 inflammasome and interferon signaling. *Immunity* **55**, 1370-1385.e8 (2022).
205. Xu, J., Wu, M., Zhu, S., Lei, J. & Gao, J. Detecting the stable point of therapeutic effect of chronic myeloid leukemia based on dynamic network biomarkers. *BMC Bioinformatics* **20**, 202 (2019).
206. Fei, F. *et al*. Galectin-1 and Galectin-3 in B-Cell Precursor Acute Lymphoblastic Leukemia. *Int J Mol Sci* **23**, 14359 (2022).
207. Cedeno-Laurent, F. & Dimitroff, C. J. Galectin-1 research in T cell immunity: Past, present and future. *Clin Immunol* **142**, 107–116 (2012).
208. Zhou, H. *et al*. Growth Arrest Specific 2 Is Up-Regulated in Chronic Myeloid Leukemia Cells and Required for Their Growth. *PLoS One* **9**, e86195 (2014).
209. Ramirez-Guzman, L. A., Huang, W., Cole, J. J. & Jørgensen, H. G. GAS2 Upregulation Is a Targetable Vulnerability in Chronic Myeloid Leukemia. *International Journal of Translational Medicine* **4**, 354–368 (2024).
210. Khalaf, A. *et al*. Nutrient-sensitizing drug repurposing screen identifies lomerizine as a mitochondrial metabolism inhibitor of chronic myeloid leukemia. *Science Translational Medicine* **16**, eadi5336 (2024).
211. Trojani, A. *et al*. Bone Marrow CD34+/lin− Cells of Patients with Chronic-Phase Chronic Myeloid Leukemia (CP-CML) After 12 Months of Nilotinib Treatment Exhibit a Different Gene Expression Signature Compared to the Diagnosis and the Corresponding Cells from Healthy Subjects. *Cancers (Basel)* **17**, 1022 (2025).
212. Rattigan, K. M. *et al*. Arginine dependency is a therapeutically exploitable vulnerability in chronic myeloid leukaemic stem cells. *EMBO reports* **24**, e56279 (2023).
213. Lozzio, C. B. & Lozzio, B. B. Human Chronic Myelogenous Leukemia Cell-Line With Positive Philadelphia Chromosome. *Blood* **45**, 321–334 (1975).
214. Kubonishi, I. & Miyoshi, I. Establishment of a Ph1 chromosome-positive cell line from chronic myelogenous leukemia in blast crisis. *The International Journal Of Cell Cloning* **1**, 105–117 (1983).
215. Ahler, E. *et al*. Doxycycline Alters Metabolism and Proliferation of Human Cell Lines. *PLoS One* **8**, e64561 (2013).
216. Pilla, C. *et al*. Creatine kinase activity from rat brain is inhibited by branched-chain amino acids in vitro. *Neurochem Res* **28**, 675–679 (2003).
217. Wang, T.-F. *et al*. Platelet-derived circulating soluble P-selectin is sufficient to induce hematopoietic stem cell mobilization. *Stem Cell Res Ther* **14**, 300 (2023).
218. Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. **16**, 712–724 (2015).
219. Shin, J. Y., Hu, W., Naramura, M. & Park, C. Y. High c-Kit expression identifies hematopoietic stem cells with impaired self-renewal and

megakaryocytic bias. *The Journal of Experimental Medicine* **211**, 217 (2014).
220. Bhattacharya, D., Rossi, D. J., Bryder, D. & Weissman, I. L. Purified hematopoietic stem cell engraftment of rare niches corrects severe lymphoid deficiencies without host conditioning. *J Exp Med* **203**, 73–85 (2006).
221. Haney, M. S. *et al.* In vivo CRISPR screening identifies SAGA complex members as key regulators of hematopoiesis. *bioRxiv* 2022.07.22.501030 (2025) doi:10.1101/2022.07.22.501030.
222. Sun, R. C. *et al.* Noninvasive liquid diet delivery of stable isotopes into mouse models for deep metabolic network tracing. *Nat Commun* **8**, 1646 (2017).
223. Bhatia, M., Wang, J. C. Y., Kapp, U., Bonnet, D. & Dick, J. E. Purification of primitive human hematopoietic cells capable of repopulating immune-deficient mice. *Proceedings of the National Academy of Sciences* **94**, 5320–5325 (1997).
224. Anjos-Afonso, F. *et al.* Single cell analyses identify a highly regenerative and homogenous human CD34+ hematopoietic stem cell population. *Nat Commun* **13**, 2048 (2022).
225. Wong, W. M. *et al.* Expression of Integrin α2 Receptor in Human Cord Blood CD34+CD38−CD90+ Stem Cells Engrafting Long-Term in NOD/SCID-IL2Rγcnull Mice. *Stem Cells* **31**, 360–371 (2013).
226. Notta, F. *et al.* Isolation of Single Human Hematopoietic Stem Cells Capable of Long-Term Multilineage Engraftment. *Science* **333**, 218–221 (2011).
227. Dolinska, M. *et al.* Characterization of the bone marrow niche in patients with chronic myeloid leukemia identifies CXCL14 as a new therapeutic option. *Blood* **142**, 73–89 (2023).
228. Reinisch, A. *et al.* A Humanized Ossicle-niche Xenotransplantation Model Enables Improved Human Leukemic Engraftment. *Nat Med* **22**, 812–821 (2016).
229. Zhang, H., Li, H., Xi, H. S. & Li, S. HIF1α is required for survival maintenance of chronic myeloid leukemia stem cells. *Blood* **119**, 2595–2607 (2012).